

CHAPMAN & HALL/CRC APPLIED MATHEMATICS  
AND NONLINEAR SCIENCE SERIES

# Introduction to Quantum Control and Dynamics



Domenico D'Alessandro



Chapman & Hall/CRC  
Taylor & Francis Group

CHAPMAN & HALL/CRC APPLIED MATHEMATICS  
AND NONLINEAR SCIENCE SERIES

---

# Introduction to Quantum Control and Dynamics

---

Domenico D'Alessandro

Iowa State University  
Ames, U.S.A.



Chapman & Hall/CRC

Taylor & Francis Group

Boca Raton London New York

---

Chapman & Hall/CRC is an imprint of the  
Taylor & Francis Group, an informa business

Chapman & Hall/CRC  
Taylor & Francis Group  
6000 Broken Sound Parkway NW, Suite 300  
Boca Raton, FL 33487-2742

© 2008 by Taylor & Francis Group, LLC

Chapman & Hall/CRC is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works

Printed in the United States of America on acid-free paper

10 9 8 7 6 5 4 3 2 1

International Standard Book Number-13: 978-1-58488-884-0 (Hardcover)

This book contains information obtained from authentic and highly regarded sources. Reprinted material is quoted with permission, and sources are indicated. A wide variety of references are listed. Reasonable efforts have been made to publish reliable data and information, but the author and the publisher cannot assume responsibility for the validity of all materials or for the consequences of their use.

No part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access [www.copyright.com](http://www.copyright.com/) (<http://www.copyright.com/>) or contact the Copyright Clearance Center, Inc. (CCC) 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

**Trademark Notice:** Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

---

**Library of Congress Cataloging-in-Publication Data**

---

D'Alessandro, Domenico.

Introduction to quantum control and dynamics / Domenico D'Alessandro.

p. cm. -- (Chapman & Hall/CRC applied mathematics and nonlinear  
science series)

Includes bibliographical references and index.

ISBN 978-1-58488-884-0 (hardcover : acid-free paper)

1. Quantum theory--Mathematics. 2. Control theory. 3. Lie groups. 4. Algebras,  
Linear. I. Title. II. Series.

QC174.125.D2 2008

530.1201'51--dc22

2007017047

---

Visit the Taylor & Francis Web site at  
<http://www.taylorandfrancis.com>

and the CRC Press Web site at  
<http://www.crcpress.com>

---

# Contents

<b>1</b>	<b>Quantum Mechanics</b>	<b>1</b>
1.1	States and Operators . . . . .	1
1.1.1	State of a quantum system . . . . .	1
1.1.2	Linear operators . . . . .	5
1.1.3	State of composite systems and tensor product spaces . . . . .	8
1.1.4	State of an ensemble; Density operator . . . . .	11
1.1.5	Vector and matrix representation of states and operators . . . . .	13
1.2	Observables and Measurement . . . . .	15
1.2.1	Observables . . . . .	15
1.2.2	The measurement postulate . . . . .	20
1.2.3	Measurements on ensembles . . . . .	24
1.3	Dynamics of Quantum Systems . . . . .	25
1.3.1	Schrödinger picture . . . . .	26
1.3.2	Heisenberg and interaction picture . . . . .	29
1.4	Notes . . . . .	30
1.4.1	Interpretation of quantum dynamics as information processing . . . . .	30
1.4.2	Direct sum versus tensor product for composite systems . . . . .	31
1.5	Exercises . . . . .	32
<b>2</b>	<b>Modeling of Quantum Control Systems; Examples</b>	<b>35</b>
2.1	Quantum Theory of Interaction of Particles and Field . . . . .	35
2.1.1	Classical electrodynamics . . . . .	36
2.1.2	Canonical quantization . . . . .	46
2.1.3	An example of canonical quantization: The quantum harmonic oscillator . . . . .	50
2.1.4	Quantum mechanical Hamiltonian . . . . .	53
2.2	Approximations and Modeling; Molecular Systems . . . . .	56
2.2.1	Approximations for molecular and atomic systems . . . . .	56
2.2.2	Controlled Schrödinger wave equation . . . . .	59
2.3	Spin Dynamics and Control . . . . .	62
2.3.1	Introduction of the spin degree of freedom in the dynamics of matter and fields . . . . .	63
2.3.2	Spin networks as control systems . . . . .	66
2.4	Mathematical Structure of Quantum Control Systems . . . . .	69
2.5	Notes and References . . . . .	72

2.6 Exercises . . . . .	73
<b>3 Controllability</b>	<b>75</b>
3.1 Lie Algebras and Lie Groups . . . . .	76
3.1.1 Basic definitions for Lie algebras . . . . .	76
3.1.2 Lie groups . . . . .	79
3.2 Controllability Test: The Dynamical Lie Algebra . . . . .	81
3.2.1 Procedure to generate a basis of the dynamical Lie algebra . . . . .	82
3.2.2 Uniform finite generation of compact Lie groups and universal quantum gates . . . . .	83
3.3 Notions of Controllability for the State . . . . .	84
3.4 Pure State Controllability . . . . .	85
3.4.1 Lie transformation groups . . . . .	86
3.4.2 Coset spaces and homogeneous spaces . . . . .	87
3.4.3 The special unitary group and its action on the unit sphere . . . . .	88
3.4.4 The symplectic group and its action on the unit sphere . . . . .	89
3.4.5 Test for pure state controllability . . . . .	93
3.5 Equivalent State Controllability . . . . .	94
3.6 Equality of Orbits . . . . .	95
3.6.1 Density matrix controllability . . . . .	97
3.7 Notes and References . . . . .	98
3.7.1 Alternate tests of controllability . . . . .	98
3.7.2 Pure state controllability and existence of constants of motion . . . . .	100
3.7.3 Bibliographical notes . . . . .	102
3.7.4 Some open problems . . . . .	102
3.8 Exercises . . . . .	103
<b>4 Observability and State Determination</b>	<b>107</b>
4.1 Quantum State Tomography . . . . .	107
4.1.1 Example: Quantum tomography of a spin- $\frac{1}{2}$ particle . . . . .	107
4.1.2 General quantum tomography . . . . .	109
4.1.3 Example: Quantum tomography of a spin- $\frac{1}{2}$ particle (ctd.) . . . . .	111
4.2 Observability . . . . .	113
4.2.1 Equivalence classes of indistinguishable states; Partition of the state space . . . . .	114
4.3 Observability and Methods for State Reconstruction . . . . .	118
4.3.1 Observability conditions and tomographic methods . . . . .	118
4.3.2 System theoretic methods for quantum state reconstruction . . . . .	119
4.4 Notes and References . . . . .	121
4.5 Exercises . . . . .	121

<b>5 Lie Group Decompositions and Control</b>	<b>123</b>
5.1 Decompositions of $SU(2)$ and Control of Two Level Systems	125
5.1.1 The Lie groups $SU(2)$ and $SO(3)$ . . . . .	125
5.1.2 Euler decomposition of $SU(2)$ and $SO(3)$ . . . . .	126
5.1.3 Determination of the angles in the Euler decomposition of $SU(2)$ . . . . .	127
5.1.4 Application to the control of two level quantum systems	128
5.2 Decomposition in Planar Rotations	130
5.3 Cartan Decompositions	131
5.3.1 Cartan decomposition of semisimple Lie algebras . . . . .	132
5.3.2 The decomposition theorem for Lie groups . . . . .	132
5.3.3 Refinement of the decomposition; Cartan subalgebras	133
5.3.4 Cartan decompositions of $su(n)$ . . . . .	135
5.3.5 Cartan involutions of $su(n)$ and quantum symmetries	137
5.3.6 Computation of the factors in the Cartan decompositions of $SU(n)$ . . . . .	139
5.4 Levi Decomposition	145
5.4.1 Ideals and normal subgroups . . . . .	145
5.4.2 Solvable Lie algebras . . . . .	146
5.4.3 Levi decomposition . . . . .	147
5.5 Examples of Application of Decompositions to Control	147
5.5.1 Control of two coupled spin- $\frac{1}{2}$ particles with Ising interaction . . . . .	148
5.5.2 Control of two coupled spin- $\frac{1}{2}$ particles with Heisenberg interaction . . . . .	150
5.6 Notes and References	153
5.7 Exercises	154
<b>6 Optimal Control of Quantum Systems</b>	<b>157</b>
6.1 Formulation of the Optimal Control Problem	158
6.1.1 Optimal control problems of Mayer, Lagrange and Bolza	158
6.1.2 Optimal control problems for quantum systems . . . . .	160
6.2 The Necessary Conditions of Optimality	162
6.2.1 General necessary conditions of optimality . . . . .	162
6.2.2 The necessary optimality conditions for quantum control problems . . . . .	166
6.3 Example: Optimal Control of a Two Level Quantum System	166
6.4 Time Optimal Control of Quantum Systems	169
6.4.1 The time optimal control problem; Bounded control .	171
6.4.2 Minimum time control with unbounded control; Riemannian symmetric spaces . . . . .	175
6.5 Numerical Methods for Optimal Control of Quantum Systems	182
6.5.1 Methods using discretization . . . . .	183
6.5.2 Iterative methods . . . . .	183

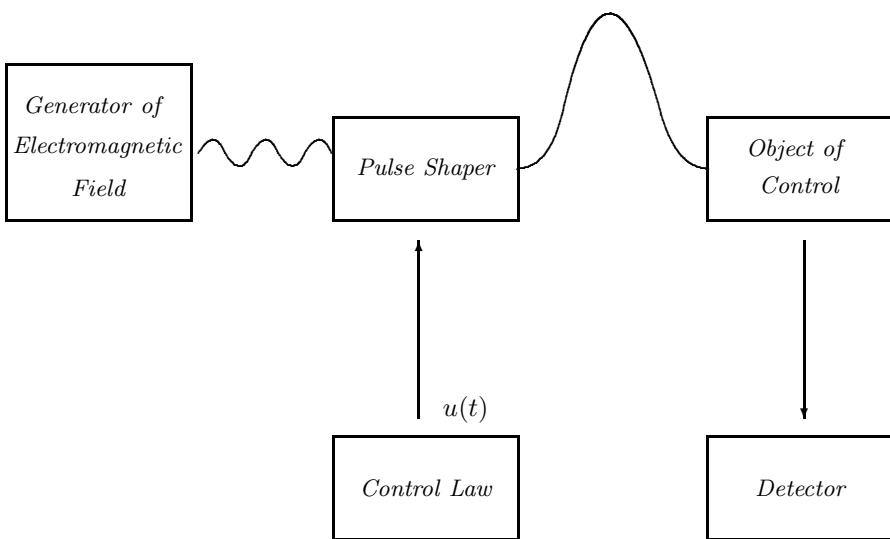
6.5.3	Numerical methods for two points boundary value problems . . . . .	186
6.6	Notes and References . . . . .	187
6.7	Exercises . . . . .	188
<b>7</b>	<b>More Tools for Quantum Control</b>	<b>191</b>
7.1	Selective Population Transfer via Frequency Tuning . . . . .	191
7.2	Time Dependent Perturbation Theory . . . . .	196
7.3	Adiabatic Control . . . . .	198
7.4	STIRAP . . . . .	201
7.5	Lyapunov Control of Quantum Systems . . . . .	205
7.5.1	Quantum control problems in terms of a Lyapunov function . . . . .	205
7.5.2	Determination of the control function . . . . .	208
7.5.3	Study of the asymptotic behavior of the state $\rho$ . . . . .	208
7.6	Notes and References . . . . .	214
7.7	Exercises . . . . .	215
<b>8</b>	<b>Analysis of Quantum Evolutions; Entanglement, Entanglement Measures and Dynamics</b>	<b>217</b>
8.1	Entanglement of Quantum Systems . . . . .	218
8.1.1	Basic definitions and notions . . . . .	218
8.1.2	Tests of entanglement . . . . .	223
8.1.3	Measures of entanglement and concurrence . . . . .	231
8.2	Dynamics of Entanglement . . . . .	238
8.2.1	The two qubits example . . . . .	240
8.2.2	The odd-even decomposition and concurrence dynamics	243
8.2.3	Recursive decomposition of dynamics in entangling and local parts . . . . .	248
8.3	Local Equivalence of States . . . . .	251
8.3.1	General considerations on dimensions . . . . .	252
8.3.2	Invariants and polynomial invariants . . . . .	255
8.3.3	Some solved cases . . . . .	257
8.4	Notes and References . . . . .	257
8.5	Exercises . . . . .	259
<b>9</b>	<b>Applications of Quantum Control and Dynamics</b>	<b>261</b>
9.1	Nuclear Magnetic Resonance Experiments . . . . .	261
9.1.1	Basics of NMR . . . . .	261
9.1.2	2-Dimensional NMR . . . . .	266
9.1.3	Control problems in NMR . . . . .	268
9.2	Molecular Systems Control . . . . .	269
9.2.1	Pulse shaping . . . . .	269
9.2.2	Objectives and techniques of molecular control . . . . .	270

<b>9.3 Atomic Systems Control; Implementations of Quantum Information Processing with Ion Traps . . . . .</b>	<b>272</b>
9.3.1 Physical set-up of the trapped ions quantum information processor . . . . .	273
9.3.2 Classical Hamiltonian . . . . .	274
9.3.3 Quantum mechanical Hamiltonian . . . . .	275
9.3.4 Practical implementation of different interaction Hamiltonians . . . . .	277
9.3.5 The control problem: Switching between Hamiltonians	282
9.4 Notes and References . . . . .	282
9.5 Exercises . . . . .	283
<b>A Positive and Completely Positive Maps, Quantum Operations and Generalized Measurement Theory . . . . .</b>	<b>287</b>
A.1 Positive and Completely Positive Maps . . . . .	287
A.2 Quantum Operations and Operator Sum Representation . . . . .	288
A.3 Generalized Measurement Theory . . . . .	289
<b>B Lagrangian and Hamiltonian Formalism in Classical Electrodynamics . . . . .</b>	<b>291</b>
B.1 Lagrangian Mechanics . . . . .	291
B.2 Extension of Lagrangian Mechanics to Systems with Infinite Degrees of Freedom . . . . .	296
B.3 Lagrangian and Hamiltonian Mechanics for a System of Interacting Particles and Field . . . . .	299
<b>C Cartan Semisimplicity Criterion and Calculation of the Levi Decomposition . . . . .</b>	<b>305</b>
C.1 The Adjoint Representation . . . . .	305
C.2 Cartan Semisimplicity Criterion . . . . .	306
C.3 Quotient Lie Algebras . . . . .	306
C.4 Calculation of the Levi Subalgebra in the Levi Decomposition	307
C.5 Algorithm for the Levi Decomposition . . . . .	307
<b>D Proof of the Controllability Test of Theorem 3.2.1 . . . . .</b>	<b>309</b>
<b>E The Baker-Campbell-Hausdorff Formula and Some Exponential Formulas . . . . .</b>	<b>315</b>
<b>F Proof of Theorem 6.2.1 . . . . .</b>	<b>317</b>
<b>References . . . . .</b>	<b>321</b>

---

# Preface

In many experimental setups, an electromagnetic field interacts with a system, such as an atom, a nucleus or an electron, whose dynamics follows the laws of quantum mechanics. While it is appropriate to treat the latter system as a quantum mechanical one, the electromagnetic field can often be treated as a classical field, giving predictions that agree with macroscopic observation. This is the so-called *semiclassical approximation*. With technological advances of the last decades, it is nowadays possible to shape the interacting electromagnetic field almost at will. This leads to a point of view which considers these experiments as *control experiments* where the electromagnetic field plays the role of the control, and the atom, nucleus or electron, or any other quantum mechanical system, is the object of the control.



**FIGURE 0.1:** Scheme of a quantum control experiment

Along with the above mentioned technological opportunities, in the last two or three decades many more applications of quantum mechanical systems have emerged. These applications require the manipulation of quantum systems in a precise way. They include *control of chemical reactions*, *quantum computa-*

*tion and quantum metrology*, which is the science that studies how to measure physical quantities using quantum systems. The introduction of control theory in these sciences is important not only in the concrete implementation of experiments but also because it gives new insight into several theoretical questions at the heart of quantum physics. These include questions arising in the newly developed *quantum information theory*.

The state of a quantum system  $\vec{\psi}$  is represented by a vector in a Hilbert space  $\mathcal{H}$  whose dynamics follows a linear differential equation called the *Schrödinger equation*,

$$i \frac{d}{dt} \vec{\psi} = H(u(t)) \vec{\psi}, \quad \vec{\psi}(0) = \vec{\psi}_0. \quad (1)$$

Here  $H(u)$  is a linear Hermitian operator on the Hilbert space  $\mathcal{H}$  which we have made a function of a control  $u = u(t)$ . This models the external action on the system and, in most cases we will treat, has the physical meaning of an interacting electromagnetic field. The state vector  $\vec{\psi}$  is constrained to have norm equal to one and therefore it can be seen as a point on the complex sphere of radius one in  $\mathbb{C}^n$ , if the Hilbert space  $\mathcal{H}$  has dimension  $n$ . The solution of (1) has the form

$$\vec{\psi}(t) = X(t) \vec{\psi}_0, \quad (2)$$

where  $X(t)$  is the solution of the *Schrödinger operator equation*

$$i \dot{X} = H(u(t)) X, \quad X(0) = \mathbf{1}. \quad (3)$$

Here  $\mathbf{1}$  denotes the identity operator. In the finite dimensional case, where the dimension of the Hilbert space is  $n$ ,  $X$  is an element of the Lie group of unitary matrices  $U(n)$ .

The first question one faces, when dealing with an experimental apparatus, concerns the type of experiments that can be performed. If the only available degree of freedom is the control, the corresponding mathematical question concerns what type of state transfer can be obtained, with (2) and (3), by opportunely varying the control,  $u = u(t)$ . This is a question of *controllability* for system (3). This question can be answered by applying results on the controllability of right invariant control systems on Lie groups. In particular, the *Lie Algebra Rank Condition* says that (under some weak assumptions on the set of the possible control functions) the set of values reachable from the identity for system (3) is the connected Lie group corresponding to the Lie algebra generated by the matrices  $iH(u)$ , as  $u$  varies in the set of possible values for the controls. This fundamental result is nothing but an application of general results of geometric control theory, which were developed by several authors in the seventies. Not only, it allows us to give an answer to a very natural question in quantum control experiments but also it opens up the possibility of applying results and ideas of Lie group and Lie algebra theory to the analysis and control of quantum systems.

This book is an introduction to the analysis and control of quantum dynamics which emphasizes the application of Lie algebra and Lie group theory. It was developed from lecture notes written during a course given at Iowa State University in the spring semester of 2004. This course was intended for advanced undergraduate and beginning graduate students in mathematics, physics and engineering. These lecture notes were expanded during a sabbatical visit at the Institute of Quantum Electronics at ETH Zurich in the fall 2004 and in the following two years. The book is written for an audience with a general mathematics and physics knowledge and introduces many concepts from the beginning. In particular, it presents introductory notions of quantum mechanics. The emphasis is on the mathematical concepts and the physics and applications are presented mostly at the level of fundamental physics rather than of precise implementation of given control experiments. However, a chapter at the end of the book presents some physical implementations of quantum control and dynamics from a more applied point of view. The subject is interdisciplinary in nature as it is at the intersection of several fields: quantum mechanics, control theory, Lie group theory, linear algebra, etc.

The book is organized as follows. [Chapter 1](#) introduces the basics of quantum mechanics, with an emphasis on dynamics and a quantum information theory point of view. This chapter could be skipped by a reader already familiar with quantum mechanics. In [Chapter 2](#), from fundamental physics, a class of models for quantum control systems is derived. This class describes many experimental situations and is the main class of models considered in the book. In [Chapter 3](#), we study the controllability of quantum systems and, in the process, we introduce many concepts of Lie algebra and Lie group theory as well as Lie transformation groups. [Chapter 4](#) deals with observability of quantum systems and the related problem of quantum state determination and measurement. [Chapter 5](#) introduces Lie group decompositions as a tool to analyze dynamics and design control algorithms. In [Chapters 6 and 7](#) we describe more methods for control. In particular, in Chapter 6 we deal with optimal control theory applied to quantum systems. Several optimal control design strategies are described there giving both analytic and numerical algorithms for optimal control. Chapter 7 presents several more methods for control, including Lyapunov control, adiabatic and STIRAP methods. All the control methods described in this book are *open loop*, i.e., the control law is designed before the experiment and then used during the experiment without modifications. [Chapter 8](#) presents some topics in quantum information theory that are relevant to the study of the dynamics of quantum systems, in particular concerning *entanglement* and *entanglement dynamics*. [Chapter 9](#) discusses physical set-ups of implementations and applications of quantum control and dynamics. Chapters 1, 2, 3, 5, and 8 should be read in the given order. The remaining chapters could be, for the most part, read independently after the first three chapters.

In writing this book I have benefited from the interaction with many people. In particular, I would like to remember the late Mohammed Dahleh, who was my Ph.D. advisor and first introduced me to the subject of control of quantum systems. His view of control theory as an interdisciplinary field and his curiosity for other areas of science and technology greatly stimulated my own interest in these areas. I was fortunate to work for many years with Francesca Albertini. She has been a very competent and effective collaborator in much of the research that led to this book. I learned a lot from the interaction with Raffaele Romano who has been an excellent research collaborator and gave me many useful suggestions on this book. I would like to thank Justin Peters, the chair of my department, who has always done his best to create the optimal conditions for me to work. Faculty and students at my department have been very kind and patient in helping me with several issues during the writing of this book. I am indebted to Wolfgang Kliemann for several useful discussions and his continuous encouragement in this work. I would like to thank Atac Imamoglu for his kind hospitality at ETH Zurich in the fall of 2004 and Geza Giedke for teaching me many interesting things during that visit. I had useful discussions on the subject of this book with many people. A partial list includes Claudio Altafini, Adriano Batista, Ferdinando Borsa, Ugo Boscain, Laura Cattaneo, Mehmet Dagli, Viatcheslav Dobrovitski, Richard Ng, Michele Pavon, Murti Salapaka, Johnathan Smith, Pablo Tamborenea, and Lorenza Viola. My thanks also go to department of Physics at the University of Lecce (Italy) for its hospitality. In particular, I would like to thank Giulio Landolfi who invited me there, in the fall of 2006, to give a series of lectures, which helped to improve some parts of this book. I would like to thank James Fiedler for a careful reading of an early draft. Mehmet Dagli also read some parts of this manuscript and gave me useful suggestions. Finally, I would like to thank CRC Press for its assistance and for agreeing to publish this book and the National Science Foundation for financial support during this work.<sup>1</sup>

Credit goes to Ahmet Ozer for Figures 1.1, 2.6, 6.1, 6.2, 6.6, 7.1, 7.2, 7.3, 7.4, 8.1, 8.4, 9.2, 9.3, 9.4, 9.5, 9.6, 9.7, 9.8, 9.9, 9.10, and 9.11, while all the other figures were drawn by the author except Figures 1.3, 4.2, 6.4, and 6.5, which are by Mariano Di Maria at [www.elabografica.it](http://www.elabografica.it).

*Domenico D'Alessandro  
Ames, Iowa, U.S.A.*

---

<sup>1</sup>This material is based upon work supported by the National Science Foundation under Grant No. ECS0237925. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation (NSF).

# Chapter 1

---

## Quantum Mechanics

We present some of the main notions of quantum mechanics and introduce some of the notation that will be used in the book. Some more concepts of interest in general quantum mechanics will also be introduced in other parts of the book as needed. There are several introductions to quantum mechanics at various levels. Some of the most popular are [49], [80], [84], [88], [145], [185].

---

### 1.1 States and Operators

#### 1.1.1 State of a quantum system

##### 1.1.1.1 States and Hilbert spaces

The **state** of a quantum mechanical system is represented by a vector in a separable Hilbert space. The word ‘represented’ means here that knowledge of the state vector gives complete information on the properties of the system. A **Hilbert space**  $\mathcal{H}$  (see, e.g., [179] pg. 76 ff.) is a complex vector space with an *inner product* that is an operation  $(\cdot, \cdot) : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{C}$  which satisfies the following properties

1.

$$(\vec{y}, \vec{x}) = (\vec{x}, \vec{y})^*, \quad (1.1)$$

where  $a^*$  denotes the complex conjugated of  $a$ ,

2.

$$(\vec{x} + \vec{y}, \vec{z}) = (\vec{x}, \vec{z}) + (\vec{y}, \vec{z}),$$

3.

$$(\alpha \vec{x}, \vec{y}) = \alpha^* (\vec{x}, \vec{y}), \quad (1.2)$$

for any complex scalar  $\alpha$ , and,

4.

$$(\vec{x}, \vec{x}) \geq 0,$$

where equality holds if and only if  $\vec{x} = 0$ .

The *norm* of a vector  $\vec{x}$ ,  $\|\vec{x}\|$ , is defined as  $\|\vec{x}\| := \sqrt{(\vec{x}, \vec{x})}$ . The *metric* of  $\mathcal{H}$  is defined using this norm, the distance between the vectors  $\vec{x}$  and  $\vec{y}$  being given by  $\|\vec{x} - \vec{y}\|$ . A Hilbert space is required to be *complete* with respect to this metric<sup>1</sup>. A Hilbert space is called *separable* if there exists a dense countable set  $\{\vec{x}_j\}$ ,  $j$  in a set  $J$  which is an orthonormal basis, i.e., it is such that

$$(\vec{x}_j, \vec{x}_k) = \delta_{jk}, \quad (1.3)$$

and every vector  $\vec{x} \in \mathcal{H}$  can be written in a unique way as

$$\vec{x} = \sum_{j \in J} \vec{x}_j (\vec{x}_j, \vec{x}). \quad (1.4)$$

Here we used the *Kronecker symbol*  $\delta_{jk}$

$$\begin{aligned} \delta_{jk} &= 0, && \text{if } j \neq k, \\ \delta_{jk} &= 1, && \text{if } j = k. \end{aligned}$$

**Remark 1.1.1** A *form* on a complex vector space  $\mathcal{H}$  is a map  $: \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{C}$ . The inner product as defined above is a *sesquilinear* form on  $\mathcal{H}$  in that it is antilinear in the first argument, i.e.,  $(\alpha\vec{x} + \beta\vec{y}, \vec{z}) = \alpha^*(\vec{x}, \vec{z}) + \beta^*(\vec{y}, \vec{z})$  and linear in the second one, i.e.,  $(\vec{x}, \alpha\vec{y} + \beta\vec{z}) = \alpha(\vec{x}, \vec{y}) + \beta(\vec{x}, \vec{z})$ . Some authors give the definition by imposing linearity in the first argument, replacing  $\alpha^*$  with  $\alpha$  in (1.2).

### 1.1.1.2 Dirac Notation

In quantum mechanics, the state vectors are denoted using Dirac's notation, that is, the state  $\psi$  is denoted as  $|\psi\rangle$  and called **ket**. The inner product between two kets  $|\phi\rangle$  and  $|\psi\rangle$  is denoted by  $\langle\phi|\psi\rangle$ . Notice that from (1.1)

$$\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*.$$

### 1.1.1.3 Physical states and normalization

Vectors  $|\psi\rangle$  and  $\alpha|\psi\rangle$ , for any  $\alpha \in \mathbb{C}$ ,  $\alpha \neq 0$ , represent the same **physical state**. For this reason, it is often more appropriate to speak about *rays* or *directions* in the Hilbert space  $\mathcal{H}$  as representing the state of a quantum system. It will be often assumed that the vector representing the state of the system has norm equal to one, i.e., for every  $|\psi\rangle$ ,  $\langle\psi|\psi\rangle = \|\psi\|^2 = 1$ . Physical states can be viewed as points on a complex sphere with radius one in a Hilbert space, with points differing by a phase factor treated as the same state.

---

<sup>1</sup>Recall that a Cauchy sequence  $\{\vec{x}_n\}$  is such that, for every  $\epsilon > 0$ , there is an  $N$  such that  $\|\vec{x}_n - \vec{x}_m\| < \epsilon$  if  $n > N$  and  $m > N$ . A metric space  $\mathcal{H}$  is complete if every Cauchy sequence in  $\mathcal{H}$  converges to an element of  $\mathcal{H}$ .

#### 1.1.1.4 Dimension of a quantum system

In modeling a quantum system with an appropriate Hilbert space one chooses a set of states that 1) have a definite physical meaning, e.g., states corresponding to a given energy or states corresponding to a given position and 2) form an orthonormal basis of the Hilbert space which means that conditions (1.3), (1.4) are verified, for a countable set or their generalization for an uncountable set. In particular, in the latter case, the sum is replaced by an integral (cf. 1.1.2.4). According to the meaning of this preferred basis, we say that we work in a given *representation*. For example, if the chosen vectors represent state with a given value of the energy, we say that we work in the energy representation. The **dimension** refers to the cardinality of the preferred set of states. It could be finite, countably infinite or noncountably infinite. In the energy representation the dimension is called the **number of levels**. This terminology is justified by the measurement postulate which will be discussed in [section 1.2](#), in that the preferred states correspond to eigenvectors of linear operators on the Hilbert space.

The simplest, but a very important, example of a quantum system is a two dimensional (or two level) system which has a pair of distinguished orthonormal states, say  $|0\rangle$  and  $|1\rangle$ . Every physical state  $|\psi\rangle$  of such a system can be written as a superposition of these two states as

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad |\alpha|^2 + |\beta|^2 = 1.$$

A simple example of an uncountably infinite dimensional quantum system is a particle which moves along a line whose only degree of freedom is the position  $x$ . The preferred set of states can be denoted by  $\{|x\rangle\}$ . The state  $|x\rangle$  represents a state which, when position is measured, is found with certainty in position  $x$ .

#### 1.1.1.5 Example: Spin- $\frac{1}{2}$ particles

Several quantum systems possess a property called **spin**. Spin is a purely quantum mechanical feature which has no direct classical counterpart. The spin  $j$  of a quantum mechanical system may be a positive integer or a half integer. It is a fixed quantity associated to a system, just like its mass. For a quantum system with spin, one can measure a **spin angular momentum** because the system will be subjected to a force in a magnetic field just like a classical system of a charged particle with orbital angular momentum (cf. Lorentz equation (2.10)). If the spin angular momentum of a system with spin  $j$  is measured, the result (in appropriate units) is in the set  $\{-j, -(j-1), \dots, (j-1), j\}$ . The Hilbert space associated to a system with spin  $j$  and all the other degrees of freedom neglected has dimension  $2j+1$ . The simplest example is  $j = \frac{1}{2}$ . Particles that have spin  $\frac{1}{2}$  include, for example, electrons. In this case there are two distinguished states which, upon measurement of

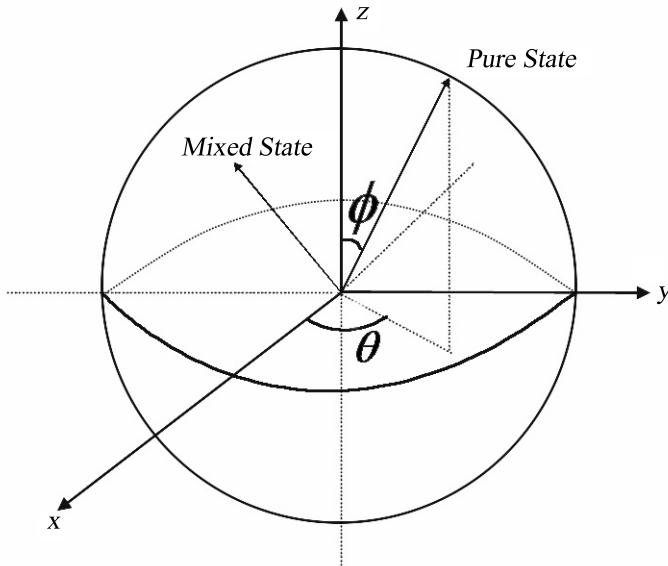
the spin angular momentum, will give  $+\frac{1}{2}$  and  $-\frac{1}{2}$ , with certainty. These two states span the underlying Hilbert space of the system.

One of the early experimental demonstrations of the existence of spin is the Stern-Gerlach experiment which is discussed in detail in several textbooks (see, e.g., [185]). The Stern-Gerlach experiment was also one of the first experimental demonstrations of physical phenomena that cannot be explained by classical mechanics alone. It will be discussed later in this chapter when we discuss measurement (see subsection 1.2.2.2).

The physical state of a spin  $\frac{1}{2}$  particle (as well as the state of any two level quantum system) is often represented by an arrow from the origin to a point on the unit sphere called the *Bloch sphere*. Let  $|0\rangle$  ( $|1\rangle$ ) be the state which produces with certainty the result  $\frac{1}{2}$  ( $-\frac{1}{2}$ ) when the spin angular momentum is measured. Then the state

$$|\psi\rangle := \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\phi} \sin\left(\frac{\theta}{2}\right)|1\rangle$$

is represented by an arrow to a point with spherical coordinates  $\theta$  and  $\phi$ . See the pure state in Figure 1.1.



**FIGURE 1.1:** Pure and mixed states on the Bloch sphere. The mixed state vector has length strictly less than one; that is its endpoint is inside the Bloch sphere. The endpoint of the pure state vector is on the surface of the Bloch sphere.

### 1.1.2 Linear operators

Given two Hilbert spaces  $\mathcal{H}_1$  and  $\mathcal{H}_2$  a linear operator  $X$ ,  $X : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ , is a linear map from  $\mathcal{H}_1$  to  $\mathcal{H}_2$ . It is called *bounded* if  $\|X\| := \sup\{\|X\vec{x}\| : \|\vec{x}\| = 1\}$  is finite. If  $\mathcal{H}_1$  and  $\mathcal{H}_2$  are both finite dimensional every linear operator  $X$  is bounded.<sup>2</sup>

#### 1.1.2.1 Bras

To any ket  $|\psi\rangle$  is associated a linear operator  $\langle\psi| : \mathcal{H} \rightarrow \mathbb{C}$  which, when applied to the vector  $|\phi\rangle$ , gives the number  $\langle\psi|\phi\rangle$ . This linear operator is called the **bra** associated with the ket  $|\psi\rangle$ . It is an element of the dual space of  $\mathcal{H}$ . If  $\mathcal{H}$  has a countable basis, then we can find a countable orthonormal basis  $\{|e_j\rangle\}$  and write  $|\psi\rangle = \sum_j \alpha_j |e_j\rangle$ . The bra corresponding to  $|\psi\rangle$  is then given by (cf. Exercise 1.2)

$$\langle\psi| = \sum_j \alpha_j^* \langle e_j|. \quad (1.5)$$

#### 1.1.2.2 Outer product

Given a ket  $|\psi\rangle$  and a bra  $\langle\phi|$ , we define a linear operator  $|\psi\rangle\langle\phi| : \mathcal{H} \rightarrow \mathcal{H}$  which is called the **outer product** of  $\langle\phi|$  with  $|\psi\rangle$ . It maps the vector  $|\chi\rangle$  to the vector  $|\psi\rangle\langle\phi|\chi\rangle$ , the vector  $|\psi\rangle$  multiplied by the scalar  $\langle\phi|\chi\rangle$ . Notice this vector is in general not normalized.

#### 1.1.2.3 Sums, linear combinations and compositions of operators

Given two (or more) linear operators  $X$  and  $Y$  we can construct a new **sum** operator  $X + Y$  which acts on a ket  $|\psi\rangle$  as

$$(X + Y)|\psi\rangle = X|\psi\rangle + Y|\psi\rangle. \quad (1.6)$$

If  $\alpha$  is a scalar,  $\alpha \in \mathbb{C}$ , and  $X$  a linear operator, the operator  $\alpha X$  is defined as

$$(\alpha X)|\psi\rangle = \alpha(X|\psi\rangle). \quad (1.7)$$

Using the sum of two operators and product of an operator by a scalar, we can construct more linear operators using operators of the type  $|\phi\rangle\langle\psi|$ . For example the operator  $|\phi_1\rangle\langle\psi_1| + \frac{1}{2}|\phi_2\rangle\langle\psi_2|$  is defined as

$$(|\phi_1\rangle\langle\psi_1| + \frac{1}{2}|\phi_2\rangle\langle\psi_2|)|\psi\rangle = |\phi_1\rangle\langle\psi_1|\psi\rangle + \frac{1}{2}|\phi_2\rangle\langle\psi_2|\psi\rangle.$$

<sup>2</sup>Cf. Example 1.3. in [51].

Another way to construct new operators from old ones is by *composition* or *product*. If  $X$  and  $Y$  are two linear operators, the linear operator  $XY$  is defined as

$$XY|\psi\rangle = X(Y|\psi\rangle). \quad (1.8)$$

#### 1.1.2.4 State expansion and completeness relations

Given a countable orthonormal basis in the Hilbert space  $\mathcal{H}$ ,  $|e_j\rangle$ ,  $j = 1, 2, \dots$ , we have

$$\sum_j |e_j\rangle\langle e_j| = \mathbf{1}, \quad (1.9)$$

where  $\mathbf{1}$  is the identity operator. In fact, by writing  $|\psi\rangle$  according to the **state expansion**,

$$|\psi\rangle = \sum_k \alpha_k |e_k\rangle, \quad (1.10)$$

and applying  $\sum_j |e_j\rangle\langle e_j|$  to both sides, by using orthonormality of the  $|e_j\rangle$ 's, one obtains

$$\left( \sum_j |e_j\rangle\langle e_j| \right) |\psi\rangle = \sum_{j,k} \alpha_k |e_j\rangle\langle e_j| |e_k\rangle = \sum_{j,k} \alpha_k |e_j\rangle \delta_{jk} = |\psi\rangle,$$

where we used  $\langle e_j | e_k \rangle = \delta_{jk}$ . By multiplying (1.10) on the left by  $\langle e_j |$  and using again  $\langle e_j | e_k \rangle = \delta_{jk}$ , we have  $\alpha_j = \langle e_j | \psi \rangle$ . Relation (1.9) is called the **completeness relation**.

In the case of uncountably infinite dimensional Hilbert space, an orthonormal basis is replaced by a basis  $|x\rangle$ , as  $x$  varies in an appropriate measurable set  $\Omega$  with measure  $dx$ , such that

$$\langle x | x_1 \rangle = \delta(x - x_1), \quad (1.11)$$

where  $\delta(x - x_1)$  is the Dirac delta function centered at  $x_1$ .<sup>3</sup> The state  $|\psi\rangle$  is expanded as

$$|\psi\rangle = \int_{\Omega} \psi(x) |x\rangle dx, \quad (1.12)$$

for some function  $\psi(x)$  on  $\Omega$ . The function  $\psi(x)$  is called the **wave function**. By multiplying both sides by  $\langle x_1 |$  and using (1.11) we obtain that  $\psi(x_1) =$

---

<sup>3</sup>A rigorous definition of the Dirac's delta function centered at  $x_1$ ,  $\delta(x - x_1)$ , requires the introduction of the theory of distributions (see, e.g., [104]). Here, we shall only need the basic property, for every continuous function  $f$ ,

$$\int_a^b f(x) \delta(x - x_1) dx = \begin{cases} f(x_1) & \text{if } x \in [a, b], \\ 0 & \text{if } x \notin [a, b]. \end{cases}$$

Whenever we shall use this property, we shall make the tacit assumption that the relevant function  $f$  is continuous.

$\langle x_1|\psi\rangle$ . Also, the completeness relation in this case becomes

$$\int_{\Omega} |x\rangle\langle x| dx = \mathbf{1}.$$

### 1.1.2.5 Hermitian, unitary operators and projections

Given a linear operator  $X$  and a ket  $|\psi\rangle$ , we can construct linear operators  $\mathcal{H} \rightarrow \mathbb{C}$  in two different ways. We can consider the ket  $X|\psi\rangle$  and take the corresponding bra. Denote this bra by  $\langle X\psi|$ . Recall that we have defined above in this section a bra as a linear operator  $\mathcal{H} \rightarrow \mathbb{C}$ . Alternatively, we can consider the linear operator  $\langle\psi|X$ , which multiplies an element of  $\mathcal{H}$  by  $X$  first and then takes the inner product with  $|\psi\rangle$ , i.e.,  $\langle\psi|X : |\phi\rangle \rightarrow \langle\psi|(X|\phi\rangle)$ .  $\langle X\psi|$  and  $\langle\psi|X$  are not always the same. However if they are the same for every ket  $|\psi\rangle$  in  $\mathcal{H}$ , then the operator  $X$  is said to be **Hermitian** or **self adjoint**. Given a bounded linear operator  $X$ , there exists a unique<sup>4</sup> linear operator, denoted by  $X^\dagger$ , acting on  $\mathcal{H}$ , such that for every  $|\psi\rangle$  and  $|\phi\rangle$ ,

$$\langle X\psi|\phi\rangle = \langle\psi|X^\dagger|\phi\rangle.$$

$X^\dagger$  is called the **adjoint** of  $X$ . An Hermitian operator  $X$  is such that  $X = X^\dagger$ .

**Example 1.1.2** Consider the three dimensional Hilbert space  $\mathcal{H}$  spanned by an orthonormal basis  $\{|0\rangle, |1\rangle, |2\rangle\}$ . The linear operator  $X$ ,  $X : \mathcal{H} \rightarrow \mathcal{H}$ , defined by

$$X|0\rangle := |0\rangle,$$

$$X|1\rangle := |0\rangle + |1\rangle,$$

$$X|2\rangle := |0\rangle,$$

is not self-adjoint. To see this consider the ket  $|\psi\rangle := |1\rangle$ . We have, from (1.5),  $\langle X\psi| = \langle 0| + \langle 1|$ . Given that state  $|0\rangle$ , we calculate  $\langle X\psi|0\rangle = 1$ , while  $\langle\psi|X|0\rangle = 0$ .

A special case of a Hermitian operator is a *projection*, which is a linear operator  $P$  having the property

$$P = P^2 = P^\dagger.$$

A **unitary operator**  $X$  is defined by the relation  $X^\dagger(t)X(t) = \mathbf{1}$ . This implies that, for every state  $|\psi\rangle$ ,

$$\langle X\psi|X|\psi\rangle := \langle\psi|X^\dagger X|\psi\rangle = \langle\psi|\psi\rangle = 1.$$

---

<sup>4</sup>See, e.g., [51] Theorem 2.2.

### 1.1.3 State of composite systems and tensor product spaces

Consider two systems  $\Sigma_1$  and  $\Sigma_2$  with states represented by vectors in the Hilbert spaces  $\mathcal{H}_1$  and  $\mathcal{H}_2$ , respectively. The state of the composite system ( $\Sigma_1$  together with  $\Sigma_2$ ) is represented by a vector in a Hilbert space which is the **tensor product** of  $\mathcal{H}_1$  and  $\mathcal{H}_2$ , denoted by  $\mathcal{H}_1 \otimes \mathcal{H}_2$ . The tensor product of  $\mathcal{H}_1$  and  $\mathcal{H}_2$  is constructed as follows: Consider an orthonormal basis  $|e_j\rangle$  of  $\mathcal{H}_1$  and an orthonormal basis  $|f_k\rangle$  of  $\mathcal{H}_2$ . Then a basis of  $\mathcal{H}_1 \otimes \mathcal{H}_2$  is constructed by putting together, in all the possible ways, one element of the basis of  $\mathcal{H}_1$  and one element of the basis of  $\mathcal{H}_2$ . In the special case where  $\mathcal{H}_1$  and  $\mathcal{H}_2$  are finite dimensional, with dimensions  $n$  and  $m$  respectively, a basis of  $\mathcal{H}_1 \otimes \mathcal{H}_2$  is given by all the possible pairs  $|e_j\rangle|f_k\rangle$ ,  $j = 1, \dots, n$ ;  $k = 1, \dots, m$ . A generic element of the basis is denoted by  $|e_j f_k\rangle$  or  $|e_j\rangle \otimes |f_k\rangle$ . When appearing between two vectors the symbol  $\otimes$  is called the *tensor product* between vectors. If  $|e_j\rangle$  and  $|f_k\rangle$  belong to bases of the Hilbert spaces  $\mathcal{H}_1$  and  $\mathcal{H}_2$ , respectively, then  $|e_j\rangle \otimes |f_k\rangle$  is the corresponding element of the basis of  $\mathcal{H}_1 \otimes \mathcal{H}_2$ . By requiring the tensor product operation between vectors to be linear with respect to both arguments the definition extends to arbitrary vectors in  $\mathcal{H}_1$  and  $\mathcal{H}_2$ . In particular, if system  $\Sigma_1$  is in the state  $|\psi_1\rangle$  and system  $\Sigma_2$  is in the state  $|\psi_2\rangle$ , we say that the overall system is in the state  $|\psi_1\rangle \otimes |\psi_2\rangle$ . Expand  $|\psi_1\rangle$  and  $|\psi_2\rangle$  in the corresponding bases as  $|\psi_1\rangle := \sum_j \alpha_j |e_j\rangle$  and  $|\psi_2\rangle = \sum_k \beta_k |f_k\rangle$ , respectively. Using linearity, we have that

$$|\psi_1\rangle \otimes |\psi_2\rangle := \sum_{j,k} \alpha_j \beta_k |e_j\rangle \otimes |f_k\rangle,$$

which shows that  $|\psi_1\rangle \otimes |\psi_2\rangle$  can be expressed in the basis of  $\mathcal{H}_1 \otimes \mathcal{H}_2$ , i.e., is an element of  $\mathcal{H}_1 \otimes \mathcal{H}_2$ .

#### 1.1.3.1 Entanglement

If the state of a composite system ( $\Sigma_1$  and  $\Sigma_2$ ) can be written as the tensor product of two states  $|\psi_1\rangle$  and  $|\psi_2\rangle$ , i.e., as  $|\psi_1\rangle \otimes |\psi_2\rangle$ , then we may say that system  $\Sigma_1$  is in the state  $|\psi_1\rangle$  and system  $\Sigma_2$  is in the state  $|\psi_2\rangle$ . However there are vectors in the space  $\mathcal{H}_1 \otimes \mathcal{H}_2$  that cannot be written as tensor products of two vectors (see the [Example 8.1.1](#) in Chapter 8). If the composite system is in one of these states then it is not possible to separate the states of systems  $\Sigma_1$  and  $\Sigma_2$ . In these cases, the two systems are said to be **entangled**. Entanglement is one of the main features of quantum mechanical systems which is used in quantum computation. There are several theoretical questions related to it including criteria for the characterization and measure of the entanglement of two systems. We shall treat entanglement and entanglement dynamics in Chapter 8. Independent surveys can be found in [92], and [155].

### 1.1.3.2 Inner product in tensor spaces

The inner product  $(\cdot, \cdot)$  in  $\mathcal{H}_1 \otimes \mathcal{H}_2$  is constructed from the inner products in the Hilbert spaces  $\mathcal{H}_1$  and  $\mathcal{H}_2$ . One first defines the inner product for tensor product states  $|\psi_1\rangle \otimes |\psi_2\rangle$ ,  $|\phi_1\rangle \otimes |\phi_2\rangle$ , as

$$(|\psi_1\rangle \otimes |\psi_2\rangle, |\phi_1\rangle \otimes |\phi_2\rangle) := \langle \psi_1 | \phi_1 \rangle \langle \psi_2 | \phi_2 \rangle, \quad (1.13)$$

and then extends by (sesqui)linearity to linear combinations of tensor products (cf. Remark 1.1.1). In particular, if

$$|\psi_1\rangle := \sum_{j,k} \alpha_{j,k} |e_j\rangle \otimes |f_k\rangle,$$

and

$$|\psi_2\rangle := \sum_{j,k} \beta_{j,k} |e_j\rangle \otimes |f_k\rangle,$$

we calculate

$$\begin{aligned} (|\psi_1\rangle, |\psi_2\rangle) &= \sum_{j,k} \alpha_{j,k}^* \left( |e_j\rangle \otimes |f_k\rangle, \sum_{l,m} \beta_{l,m} |e_l\rangle \otimes |f_m\rangle \right) \\ &= \sum_{j,k,l,m} \alpha_{j,k}^* \beta_{l,m} (|e_j\rangle \otimes |f_k\rangle, |e_l\rangle \otimes |f_m\rangle), \end{aligned}$$

where in the first equality we used antilinearity with respect to the first argument and in the second equality we used linearity with respect to the second argument. Since

$$(|e_j\rangle \otimes |f_k\rangle, |e_l\rangle \otimes |f_m\rangle) := \langle e_j | e_l \rangle \langle f_k | f_m \rangle = \delta_{jl} \delta_{km},$$

we obtain

$$(|\psi_1\rangle, |\psi_2\rangle) = \sum_{j,k,l,m} \alpha_{j,k}^* \beta_{l,m} \delta_{jl} \delta_{km} = \sum_{j,k} \alpha_{j,k}^* \beta_{j,k}.$$

### 1.1.3.3 Linear operators in tensor spaces

Consider two vector spaces  $\mathcal{V}_1$  and  $\mathcal{V}_2$  along with a linear operator  $A : \mathcal{H}_1 \rightarrow \mathcal{V}_1$  and a linear operator  $B : \mathcal{H}_2 \rightarrow \mathcal{V}_2$ . The **linear operator**  $A \otimes B$  maps elements of  $\mathcal{H}_1 \otimes \mathcal{H}_2$  to elements of  $\mathcal{V}_1 \otimes \mathcal{V}_2$  as follows. For any product state,  $|\psi_1\rangle \otimes |\psi_2\rangle$ , we have

$$(A \otimes B)|\psi_1\rangle \otimes |\psi_2\rangle := (A|\psi_1\rangle) \otimes (B|\psi_2\rangle) \quad (1.14)$$

and the definition extends to general states by linearity.

This definition includes some important special cases. If  $A$  is a linear operator  $A : \mathcal{H}_1 \rightarrow \mathbb{C}$  (for example a bra) and  $B$  a linear operator  $B : \mathcal{H}_2 \rightarrow \mathbb{C}$ , then  $A \otimes B$  is defined as

$$(A \otimes B)|\psi_1\rangle \otimes |\psi_2\rangle := (A|\psi_1\rangle)(B|\psi_2\rangle), \quad (1.15)$$

and then extended by linearity to states that are not tensor products. In the right hand side of (1.15), the tensor product coincides with the product of two numbers since both spaces have dimension 1. A special case of this type of operator is the bra corresponding to  $|\psi_1\rangle \otimes |\psi_2\rangle := |\psi_1, \psi_2\rangle$  since from (1.13) it follows that  $\langle\psi_1, \psi_2| = \langle\psi_1| \otimes \langle\psi_2|$ , where the symbol  $\otimes$  has the meaning of tensor product of two operators.

Another special case of (1.14) is the outer product of  $|\psi_1\rangle \otimes |\psi_2\rangle$  and  $|\phi_1\rangle \otimes |\phi_2\rangle$ , for states  $|\psi_1\rangle, |\phi_1\rangle$  in  $\mathcal{H}_1$  and  $|\psi_2\rangle, |\phi_2\rangle$  in  $\mathcal{H}_2$ . This can be written as

$$|\psi_1, \psi_2\rangle \langle\phi_1, \phi_2| := |\psi_1\rangle \langle\phi_1| \otimes |\psi_2\rangle \langle\phi_2|. \quad (1.16)$$

The operator on the right hand side transforms  $|v\rangle \otimes |w\rangle$  into  $|\psi_1\rangle \langle\phi_1|v\rangle \otimes |\psi_2\rangle \langle\phi_2|w\rangle$  and it is linear.

From tensor product operators one can construct more operators by using the linear combination and composition rules in [subsection 1.1.2.3](#) (cf. (1.6), (1.7), (1.8)). In particular, notice that (assuming  $\mathcal{V}_1 = \mathcal{H}_1$  and  $\mathcal{V}_2 = \mathcal{H}_2$ )

$$(A_1 \otimes B_1)(A_2 \otimes B_2) = A_1 A_2 \otimes B_1 B_2.$$

**Example 1.1.3** We consider a special case to illustrate some of the calculations introduced in this subsection and concerning composite systems. Assume system  $\Sigma_1$  and  $\Sigma_2$  are two spin- $\frac{1}{2}$  particles with Hilbert spaces  $\mathcal{H}_1$  and  $\mathcal{H}_2$ , both two dimensional and spanned by the orthonormal basis  $\{|0\rangle, |1\rangle\}$ . Consider the vector  $|\psi_1\rangle$  in  $\mathcal{H}_1$ , given by  $|\psi_1\rangle := \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ , and the vector  $|\psi_2\rangle$  in  $\mathcal{H}_2$ , given by  $|\psi_2\rangle := i|0\rangle$ . We have

$$|\psi_1, \psi_2\rangle := |\psi_1\rangle \otimes |\psi_2\rangle = \frac{i}{\sqrt{2}} (|0\rangle \otimes |0\rangle + |1\rangle \otimes |0\rangle).$$

The state

$$|\phi\rangle := \frac{1}{\sqrt{2}} (|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle)$$

is an entangled state as it cannot be expressed as the tensor product of two states (cf. [Chapter 8](#)). We calculate

$$(|\psi_1\rangle \otimes |\psi_2\rangle, |\phi\rangle) = \frac{-i}{2},$$

and

$$\begin{aligned} & |\psi_1, \psi_2\rangle \langle\phi| \\ &= \frac{i}{2} (|0\rangle \langle 0| \otimes |0\rangle \langle 0| + |1\rangle \langle 0| \otimes |0\rangle \langle 0| + |0\rangle \langle 1| \otimes |0\rangle \langle 1| + |1\rangle \langle 1| \otimes |0\rangle \langle 1|). \end{aligned}$$

We have illustrated the definitions concerning composite systems for the case of systems composed of two subsystems, that is, for *bipartite systems*. The definitions extend naturally to the case of *multipartite systems* that are systems composed by more than two subsystems.

### 1.1.4 State of an ensemble; Density operator

It is often the case that the object of study consists of an **ensemble** of a large number of identical systems in different states. In these cases, we need to have information on the fraction of systems in a given state. This information is recorded using an operator rather than a vector. This operator, called the **density operator**, is defined in the following.

If there is a fraction  $0 < w_j \leq 1$  of systems in the ensemble with state  $|\psi_j\rangle$ ,  $j$  in a set  $\mathcal{I}$ ,  $\sum_{j \in \mathcal{I}} w_j = 1$ , then the density operator describing the state of the ensemble is given by

$$\rho := \sum_{j \in \mathcal{I}} w_j |\psi_j\rangle\langle\psi_j|. \quad (1.17)$$

It is a linear operator  $\mathcal{H} \rightarrow \mathcal{H}$  as it is a linear combination of outer products. The density operator completely describes the state of the ensemble. Special cases are **pure ensembles** or **pure states** which are such that  $w_j = 1$ , for some  $j$ . They are described by density operators of the form,

$$\rho := |\psi_j\rangle\langle\psi_j|.$$

On the contrary, **mixed ensembles** (or **mixed states**) are described by density operators (1.17) with more than one  $w_j$  different from zero. The term *density matrix* is often used instead of density operator. The density operator has the following fundamental properties that can be proved directly using the definition.<sup>5</sup>

1.  $\rho$  is Hermitian and positive semidefinite.<sup>6</sup>

2.

$$\text{Tr}(\rho) = 1. \quad (1.18)$$

3.

$$\rho^2 = \rho$$

if and only if  $\rho$  represents a pure ensemble.

---

<sup>5</sup>Recall that the trace of an operator  $X$  on a Hilbert space with countable orthonormal basis  $|e_k\rangle$  is given by  $\text{Tr}(X) = \sum_k \langle e_k | X | e_k \rangle$  and is independent of the choice of the orthonormal basis. In the uncountable case, the sum is replaced by an appropriate integral  $\text{Tr}(X) = \int_{\Omega} \langle x | X | x \rangle dx$ , where  $|x\rangle$ ,  $x \in \Omega$ , is an orthonormal basis (cf. (1.11)).

<sup>6</sup>A Hermitian operator  $X : \mathcal{H} \rightarrow \mathcal{H}$  is called positive semidefinite if  $\langle \psi | X | \psi \rangle \geq 0$ , for every  $|\psi\rangle \in \mathcal{H}$ .

4.

$$0 < \text{Tr}(\rho^2) < 1$$

if  $\rho$  represents a mixed ensemble.

The set of density matrices is a convex subset of the space of Hermitian operators whose ‘extreme’ points represent pure states. By ‘extreme points’ we mean here operators that cannot be written as convex combinations of two operators in the set. This is a consequence of properties 3 and 4 above.<sup>7</sup>

#### 1.1.4.1 Example: Ensembles of two level systems

The density matrix representing an ensemble of two level systems can be written as

$$\rho = \frac{1}{2}(\mathbf{1} + x\sigma_x + y\sigma_y + z\sigma_z), \quad (1.19)$$

where the matrices  $\sigma_{x,y,z}$  are the *Pauli matrices*

$$\sigma_x := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y := \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad \sigma_z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (1.20)$$

The state is represented by a vector (called *Bloch vector*) as in Figure 1.1 whose endpoint has coordinates  $x$ ,  $y$ , and  $z$ . In the case of a pure state the representative point on the Bloch sphere is the same as the one obtained with polar coordinates in Figure 1.1. If the state is a mixed state,  $\text{Tr}(\rho^2) < 1$ , i.e.,  $x^2 + y^2 + z^2 < 1$ , the endpoint is *inside* the Bloch sphere.

#### 1.1.4.2 Ensembles of composite systems

Consider two subsystems with associated orthogonal bases  $\{|e_j\rangle\}$  and  $\{|f_k\rangle\}$ , respectively. A pure state of the overall system  $|\psi\rangle$  is described by

$$|\psi\rangle = \sum_{j,k} \alpha_{j,k} |e_j\rangle \otimes |f_k\rangle.$$

---

<sup>7</sup>Write a density operator  $\rho$  as

$$\rho := \lambda\rho_1 + (1 - \lambda)\rho_2, \quad 0 \leq \lambda \leq 1,$$

and calculate  $\rho^2$ . We have

$$\text{Tr}(\rho^2) = \lambda^2 \text{Tr}(\rho_1^2) + (1 - \lambda)^2 \text{Tr}(\rho_2^2) + \lambda(1 - \lambda) \text{Tr}(\rho_1\rho_2).$$

If  $\rho$  is a pure state then  $\text{Tr}(\rho^2) = 1$ , but to have this we need  $\text{Tr}(\rho_1^2) = \text{Tr}(\rho_2^2) = 1$  and  $\rho_1$  proportional to  $\rho_2$  which implies by (1.18)  $\rho_1 = \rho_2$ . We also need  $\lambda = 0$  or  $\lambda = 1$ . We have used here the property of the Frobenius inner product of Hilbert-Schmidt operators,  $\text{Tr}(\rho_1\rho_2) \leq \text{Tr}(\rho_1^2)\text{Tr}(\rho_2^2)$  with equality if and only if  $\rho_1$  and  $\rho_2$  are linearly dependent. See, e.g., [158] pg. 118.

If we have an ensemble of systems, with states

$$|\psi_l\rangle := \sum_{j,k} \alpha_{j,k,l} |e_j\rangle \otimes |f_k\rangle,$$

each with probability  $w_l$ , using linearity and the definition of outer product for tensor product kets (1.16), we derive the density operator  $\rho$  for the ensemble,

$$\rho = \sum_l \sum_{j,k} \sum_{r,s} w_l \alpha_{j,k,l} \alpha_{r,s,l}^* |e_j\rangle \langle e_r| \otimes |f_k\rangle \langle f_s|.$$

#### 1.1.4.3 Density matrix describing a single system

Density matrices are also used to describe an unknown state of a single system. If a quantum system is in the state

$$\rho = \sum_k w_k |\psi_k\rangle \langle \psi_k|,$$

with  $\sum_k w_k = 1$  then there is probability  $w_k$  that the system is in the pure state  $|\psi_k\rangle$ .

The use of a density matrix to describe both pure and mixed states allows a unified and elegant treatment.

#### 1.1.5 Vector and matrix representation of states and operators

In the following we shall often work with finite dimensional systems whose Hilbert space is spanned by an orthonormal basis  $|e_j\rangle$ ,  $j = 1, \dots, n$ . Associating with  $|e_j\rangle$  the column vector having all the entries equal to zero and a 1 in the  $j$ -th position, a ket  $|\psi\rangle := \sum_{j=1}^n \alpha_j |e_j\rangle$  is represented by a column vector denoted by  $\vec{\psi}$ ,

$$\vec{\psi} := \begin{pmatrix} \alpha_1 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_n \end{pmatrix}.$$

The corresponding bra  $\langle\psi|$  is represented by  $\vec{\psi}^\dagger$ , i.e., the conjugate transposed of  $\vec{\psi}$ , and we have

$$\langle\psi|\phi\rangle = \vec{\psi}^\dagger \vec{\phi}.$$

It is known from linear algebra that once a basis is fixed linear operators are in one to one correspondence with  $n \times n$  matrices. While  $\vec{\psi}$  denotes the column vector associated with  $|\psi\rangle$ , we shall use the same notation for linear operators and the matrices that represent them. Therefore, if  $|\phi\rangle = X|\psi\rangle$ , with an operator  $X$ , then  $\vec{\phi} = X\vec{\psi}$ , with the matrix  $X$  in the corresponding vector-matrix notation.

### 1.1.5.1 Composite systems and Kronecker product of matrices

Consider two systems with Hilbert spaces  $\mathcal{H}_1$  and  $\mathcal{H}_2$  spanned respectively by orthonormal bases  $\{|e_j\rangle\}$ ,  $j = 1, \dots, n$  and  $\{|f_k\rangle\}$ ,  $k = 1, \dots, m$ . Consider now the corresponding basis in  $\mathcal{H}_1 \otimes \mathcal{H}_2$ ,  $\{|e_j, f_k\rangle\}$ . In Exercise 1.4, the reader is asked to verify that if  $|\psi\rangle \in \mathcal{H}_1$  has column vector representation  $\vec{\psi}$  and  $|\phi\rangle \in \mathcal{H}_2$  has column vector representation  $\vec{\phi}$ , then  $|\psi\rangle \otimes |\phi\rangle$  has column vector representation  $\vec{\psi} \otimes \vec{\phi}$  where  $\otimes$  here denotes the Kronecker product of the two vectors. Here we have ordered the elements of the basis  $\{|e_j, f_k\rangle\}$  so that  $|e_{j_1}, f_{k_1}\rangle$  comes before  $|e_{j_2}, f_{k_2}\rangle$  if  $j_1 < j_2$  or, when  $j_1 = j_2$ , if  $k_1 < k_2$ . The *Kronecker product* ( $\otimes$ ) of two general matrices  $A$  and  $B$  (with arbitrary dimensions) is defined as

$$A \otimes B := \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ a_{r1}B & a_{r2}B & \cdots & a_{rn}B \end{pmatrix}.$$

A comprehensive discussion of the properties of the Kronecker product can be found in [100] pg. 242 ff. Some of the properties we shall use in the following are

1.

$$(A \otimes B)(C \otimes D) = AC \otimes BD, \quad (1.21)$$

2.

$$(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger,$$

3.

$$\text{Tr}(A \otimes B) = \text{Tr}(A) \text{Tr}(B),$$

4. If  $\mathbf{1}$  denotes the identity matrix of appropriate dimensions

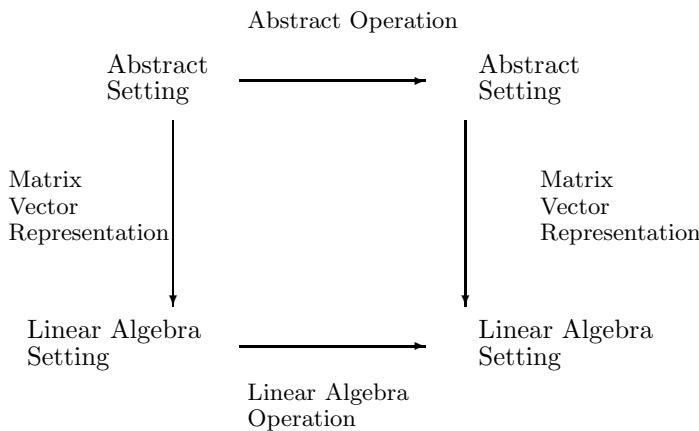
$$e^{A \otimes \mathbf{1} + \mathbf{1} \otimes B} = e^A \otimes e^B. \quad (1.22)$$

With the above ordering convention on the basis of the tensor product space, the matrix representative of the tensor product  $A \otimes B$  of two operators is the Kronecker product of the two corresponding matrices.

### 1.1.5.2 Correspondence between abstract and linear algebra operations

Generalizing what said above, all the operations that we have defined in the abstract setting of vectors and operators have a counterpart in linear algebra.

This is summarized in [Table 1.1](#). Notice in particular from the table that Hermitian operators correspond to Hermitian matrices. The *abstract operation* with vectors and operators, and the corresponding *linear algebra operation*, with column vectors and matrices, are such that the following commutative diagram is satisfied.



## 1.2 Observables and Measurement

Measurement theory in quantum mechanics is the subject of much study. A summary and references can be found in [33]. There are in general several types of measurements that can be performed in a laboratory. In this section, we summarize the concepts concerning the standard Von Neumann-Lüders measurement. Extensions to more general types of measurements can be obtained using the formalism of operations and effects and this is treated in [Appendix A](#).

### 1.2.1 Observables

In quantum mechanics, observed quantities are called **observables** and are associated with Hermitian operators on the Hilbert space  $\mathcal{H}$ . Rules are postulated to associate a physical quantity with a Hermitian operator acting on states. The form of this operator will depend on the specific representation used for the system just like the matrix representing a linear operator on a finite dimensional vector space depends on the basis chosen for the vector

**TABLE 1.1:** Abstract operations with vectors and operators and the corresponding linear algebra counterparts

Abstract Operation	Linear Algebra Operation
$\langle \psi   \phi \rangle$ , inner product in $\mathcal{H}$	$\vec{\psi}^\dagger \vec{\phi}$ , inner product in $\mathbb{C}^n$
$X \psi\rangle$ , linear operator $X$ acting on vector $ \psi\rangle$	$X\vec{\psi}$ , matrix $X$ multiplying vector $\vec{\psi}$
$ \psi\rangle\langle\phi $ , outer product	$\vec{\psi}\vec{\phi}^\dagger$ , column-row product
$XY$ , composition of operators $X, Y$	$XY$ , product of matrices $X, Y$
$\alpha X + \beta Y$ , linear combination of operators $X, Y$	$\alpha X + \beta Y$ linear combination of matrices $X, Y$
$X^\dagger$ , adjoint operator of $X$	$X^\dagger$ transposed, conjugate of matrix $X$
$ \psi\rangle \otimes  \phi\rangle$ , tensor product of vectors $ \psi\rangle,  \phi\rangle$	$\vec{\psi} \otimes \vec{\phi}$ , Kronecker product of vectors $\vec{\psi}, \vec{\phi}$
$A \otimes B$ , tensor product of operators $A, B$	$A \otimes B$ , Kronecker product of matrices $A, B$

space. This process is called *canonical quantization* and will be discussed in the next chapter in the modeling of a system of charged particles and fields. In several situations, specifically in several infinite dimensional cases, the operator defined in this fashion does not map *all* elements of  $\mathcal{H}$  to elements of  $\mathcal{H}$ . We illustrate this with an example.

**Example 1.2.1** Consider a particle with a one dimensional motion along the  $x$  axis. Assume the motion is not constrained to a bounded region so that the position  $x$  can assume all the values in  $(-\infty, +\infty)$ . In the position representation  $\{|x\rangle\}$ , the state  $|\psi\rangle$  is written as (cf. (1.12))

$$|\psi\rangle = \int_{-\infty}^{+\infty} \psi(x) |x\rangle dx. \quad (1.23)$$

Since  $\||\psi\rangle\|$  is finite, the wave function  $\psi(x)$  must be in  $L^2(\mathbf{R})$ .<sup>8</sup> The position observable is defined as the operator  $X$  which multiplies the wave function  $\psi$  by  $x$ . However, it is easy to construct examples where  $\psi$  is in  $L^2(\mathbf{R})$  while  $x\psi(x)$  is not.<sup>9</sup> Therefore this operator does not map any element with finite norm to an element with finite norm.

To circumvent this problem one defines a *domain* of a linear operator  $A$ ,  $\mathcal{D}(A)$  as the set of states in  $\mathcal{H}$  where the linear operator  $A$  is bounded, i.e.,  $\|A|\psi\rangle\| < \infty$ . A Hermitian operator is such that  $\mathcal{D}(A) = \mathcal{D}(A^\dagger)$ , and  $A$  and  $A^\dagger$  coincide on the common domain. It is assumed that  $\mathcal{D}(A)$  is dense in  $\mathcal{H}$ .

### 1.2.1.1 Spectrum and spectral theorem

Consider  $\mathbf{R}$  as a measurable space.<sup>10</sup> A *projection valued measure*,  $dP$ , on  $\mathbf{R}$  is defined as a map which associates with every measurable set  $\Omega \subseteq \mathbf{R}$ , a projection operator  $P(\Omega)$ , and satisfies the following two axioms

1.  $P(\mathbf{R}) = \mathbf{1}$ ,
2. If  $\{\Omega_n\}$  is a countable family of disjoint measurable sets  $P(\cup_n \Omega_n) = \sum_n P(\Omega_n)$ .

---

<sup>8</sup>Given a measure space  $\Omega$ , with associated measure  $dx$ , the vector space  $L^2(\Omega)$  is the space of square integrable functions  $\psi$ , on  $\Omega$ , i.e., functions satisfying

$$\left( \int_{\Omega} \psi^*(x)\psi(x) dx \right)^{\frac{1}{2}} < +\infty.$$

$L^2(\Omega)$  is a Hilbert space when equipped with the inner product

$$(\psi, \phi) := \int_{\Omega} \psi^*(x)\phi(x) dx.$$

An introduction to general  $L^p$  spaces, for general  $p$ , and general measure spaces, can be found in [179].

<sup>9</sup>Consider for example  $\psi(x) := \frac{1}{x}$ , for  $x \geq 1$  and  $\psi(x) \equiv 0$ , for  $x < 1$ .

<sup>10</sup>In most of the discussion here we follow [33].

Associated with a projection valued measure is a *spectral family* that is a family of commuting projections  $P_\lambda$ , with  $\lambda \in \mathbf{R}$ , given by the Lebesgue integral

$$P_\lambda := \int_{-\infty}^\lambda dP.$$

The spectral family has several properties. In particular  $\lambda_1 \leq \lambda_2$  implies  $\text{range}(P_{\lambda_1}) \subseteq \text{range}(P_{\lambda_2})$ . Moreover  $P_\lambda$  is continuous from the right, i.e.,  $\lim_{\epsilon \rightarrow 0} P_{\lambda+\epsilon} = P_\lambda$ , and  $\lim_{\lambda \rightarrow -\infty} P_\lambda = \mathbf{1}$ ,  $\lim_{\lambda \rightarrow +\infty} P_\lambda = \mathbf{0}$ , with  $\mathbf{0}$  the zero operator which maps every vector in  $\mathcal{H}$  to the zero vector. A point in  $\lambda \in \mathbf{R}$  is said to be *stationary* for the spectral family  $\{P_\lambda\}$  if, for every  $\epsilon > 0$ ,

$$P_{\lambda+\epsilon} - P_{\lambda-\epsilon} = \mathbf{0}.$$

The **spectral theorem** (see, e.g., [3]) gives a resolution of every self-adjoint operator in terms of a projection valued measure and a spectral family. It is stated as follows.

**Theorem 1.2.2** *To every self-adjoint operator  $A$  there corresponds a unique projection valued measure  $dP_A$  and a spectral family  $\{P_{A\lambda}\}$ , such that*

$$A = \int_{\mathbf{R}} \lambda dP_{A\lambda}, \quad (1.24)$$

on every  $|\psi\rangle$  in  $\mathcal{D}(A)$ , the domain of  $A$ .

The **spectrum** of a self-adjoint operator  $A$  is defined as the set of non-stationary points of the corresponding spectral family. Points in the spectrum are called *eigenvalues*.

### 1.2.1.2 Special cases of the spectral decomposition

A self-adjoint operator  $A$  is said to have *discrete spectrum* if all the points in the spectrum of  $A$  are isolated points. An isolated point  $\lambda_0$  in the spectrum is, by definition, such that there exists an  $\epsilon > 0$  with  $P_{A\lambda} - P_{A\lambda_0} = \mathbf{0}$  for all  $\lambda_0 \leq \lambda < \lambda_0 + \epsilon$ , and  $P_{A\lambda_0} - P_{A\lambda} \neq \mathbf{0}$ , for all  $\lambda_0 - \epsilon < \lambda < \lambda_0$  (cf. [Figure 1.2](#)). In this case, the spectrum is a countable set,  $\{\lambda_j\}$ .

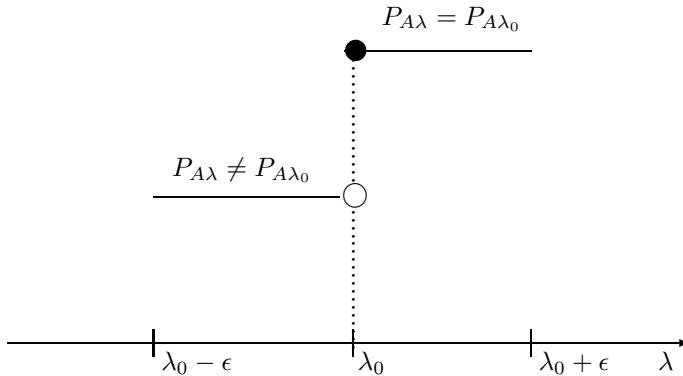
Define<sup>11</sup>(cf. Exercise 1.9)

$$P_j := P_{A\lambda_j} - P_{A\lambda_{j-1}}. \quad (1.25)$$

The spectral theorem takes the form

$$A = \sum_j \lambda_j P_j. \quad (1.26)$$

<sup>11</sup>Choose  $P_{A\lambda_{j-1}} = \mathbf{0}$  if  $\lambda_j$  is the minimum value in the spectrum.



**FIGURE 1.2:** The spectral family has a **jump** at an isolated point of the spectrum.

A vector  $|e_{jk}\rangle$  in the range of  $P_j$  is called an *eigenvector* corresponding to the eigenvalue  $\lambda_j$ . It has the property that  $A|e_{jk}\rangle = \lambda_j|e_{jk}\rangle$ . Choosing an orthonormal basis of  $\text{range}(P_j)$ ,  $\{|e_{jk}\rangle\}$ , we can write  $P_j$  as

$$P_j = \sum_k |e_{jk}\rangle \langle e_{jk}|.$$

If  $\text{rank } P_j = 1 (>1)$  then  $\lambda_j$  is said to be a *nondegenerate (degenerate) eigenvalue*.

A continuous spectrum is such that it contains no isolated points. Generalizing the definitions given above for the discrete case, an eigenvector  $|e\rangle$  corresponding to the eigenvalue  $\lambda_0$  is such that, for every  $\lambda > \lambda_0$ ,  $|e\rangle$  is in the range of

$$\int_{\lambda_0}^{\lambda} dP_A,$$

and we have, using (1.24),

$$A|e\rangle = \lambda_0|e\rangle.$$

As an example of an operator with continuous spectrum consider the position operator  $X$  of Example 1.2.1. The spectral family in this case (cf. [33] section 2.1) is such that, with  $|\psi\rangle$  in (1.23),

$$P_{X\lambda}|\psi\rangle = \int_{-\infty}^{\lambda} \psi(x)|x\rangle dx. \quad (1.27)$$

The corresponding spectral measure can be written as

$$dP_X = |x\rangle \langle x|dx, \quad (1.28)$$

which if we use (1.11) gives (1.27). Accordingly the position operator  $X$  is written as

$$X = \int_{-\infty}^{+\infty} x|x\rangle\langle x|dx. \quad (1.29)$$

There are also observables with mixed discrete and continuous spectrum.

### 1.2.2 The measurement postulate

#### 1.2.2.1 Results of measurement and modification of the state

The Von Neumann-Lüders **measurement postulate** states that, when an observable  $A$  is measured, the result is an eigenvalue of  $A$ . Let  $J$  be a measurable subset of the spectrum of  $A$ , and let  $P_J$  be the projection

$$P_J := \int_J dP_A,$$

where  $dP_A$  is the projection measure associated to  $A$ . If the state at the moment of the measurement is  $|\psi\rangle$ , the probability of obtaining a value  $\lambda$  in  $J$ , as the result of the measurement, is given by

$$\Pr(\lambda \in J) = \langle\psi|P_J|\psi\rangle = \|P_J|\psi\rangle\|^2. \quad (1.30)$$

If the measurement gives a value in  $J$  the state is modified according to the rule

$$|\psi\rangle \rightarrow \frac{P_J|\psi\rangle}{\langle\psi|P_J|\psi\rangle}.$$

The result of a measurement is certain if and only if the system is already in an eigenvector of the observable being measured, in which case we obtain the corresponding eigenvalue and the state of the system is left unchanged.

Denote by  $\Pr(\lambda_j)$  the probability that the result of the measurement is  $\lambda_j$ . The *expectation value* for the measurement of  $A$ , when the system is in the state  $|\psi\rangle$ , is given by, in the discrete spectrum case,

$$\langle A \rangle_\psi := \sum_j \lambda_j \Pr(\lambda_j) = \sum_j \lambda_j \langle\psi|P_j|\psi\rangle = \langle\psi| \sum_j \lambda_j P_j |\psi\rangle = \langle\psi|A|\psi\rangle, \quad (1.31)$$

where we used the spectral decomposition (1.26). The same formula holds for the continuous spectrum case (cf. Exercise 1.10).

It is common practice to denote by  $|\lambda\rangle$  the eigenvector corresponding to the eigenvalue  $\lambda$ . With this notation, for a discrete spectrum  $\{\lambda_j\}$  the formula of the spectral theorem can be written (in terms of one dimensional projectors) as

$$A = \sum_j \lambda_j |\lambda_j\rangle\langle\lambda_j|,$$

where the  $\lambda_j$  are possibly repeated. For a continuous spectrum  $\Omega$  the operator  $A$  can be written as

$$A = \int_{\Omega} x|x\rangle\langle x|dx,$$

where the projection valued measure  $dP_A$  is written in terms of the eigenvectors  $|x\rangle$ , i.e.,  $dP_A := |x\rangle\langle x|dx$ . This is the generalization of what was said in (1.28), (1.29), for the position operator.

Consider an observable with discrete nondegenerate spectrum  $\{\lambda_j\}$ , namely the rank of each projection  $P_j$  associated with  $\lambda_j$  is one.<sup>12</sup> In terms of the eigenvalue  $|\lambda_j\rangle$ , a state  $|\psi\rangle$  is expanded as

$$|\psi\rangle = \sum_j \alpha_j |\lambda_j\rangle,$$

and the projection associated to  $\lambda_j$  is  $P_j = |\lambda_j\rangle\langle\lambda_j|$ . According to (1.30) the probability of a result  $\lambda_j$  when  $A$  is measured and the state  $|\psi\rangle$  is

$$\langle\psi|P_j|\psi\rangle = |\alpha_j|^2.$$

Generalizing this to a continuous time spectrum, if  $|\psi\rangle$  is expanded as in (1.12) in terms of the eigenvectors,  $|x\rangle$ , of  $A$  and writing the projection corresponding to a result in the set  $J$  as

$$P_J = \int_J |x\rangle\langle x|dx,$$

using (1.11), we obtain that the probability of having a result in the set  $J$  is given by (cf. (1.30))

$$\Pr(\lambda \in J) = \langle\psi|P_J|\psi\rangle = \int_J \psi^*(x)\psi(x)dx. \quad (1.32)$$

### 1.2.2.2 Example of measurement: The Stern-Gerlach experiment

In the Stern-Gerlach experiment, an oven contains an ensemble of particles with spin  $\frac{1}{2}$  (silver atoms in the original version of the experiment). These particles exit the oven, are aligned through a narrow slit, and pass through an inhomogeneous magnetic field.<sup>13</sup> In the magnetic field they experience a

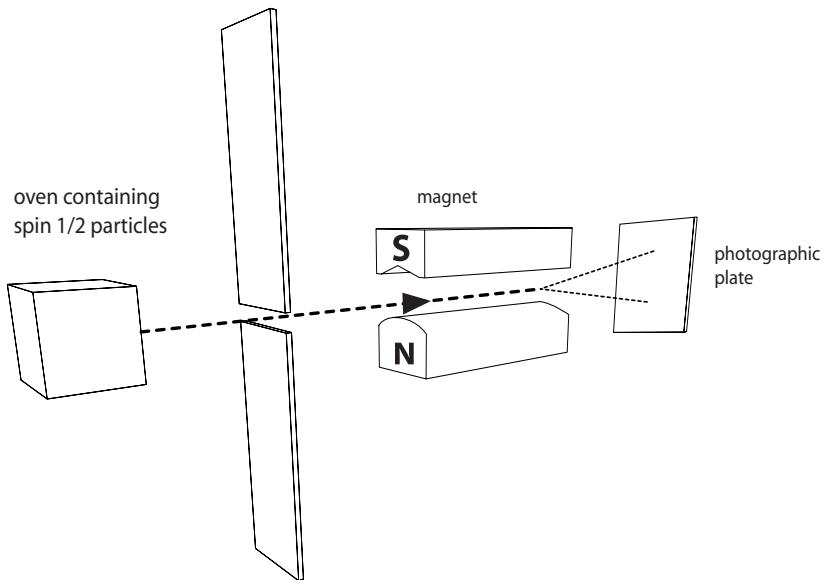
<sup>12</sup>The generalization of what we are going to say to the general possibly degenerate case is straightforward.

<sup>13</sup>If we think of this experiment with classical electrodynamics we have spinning charges in a magnetic field. The spinning charges act as magnets which tend to align themselves with the field. In fact, the torque experienced is proportional to their angular momentum. Moreover the particles experience a magnetic force depending on their orientation when considered as magnets due to the fact that the external field is not homogeneous. The South pole of the magnet-particle will be repelled from the South pole of the external magnet a little less (or more) than the North pole does.

force which deflects their paths. In particular, the  $z$ -component of the force  $F_z$  is approximately proportional to the (spin) angular momentum in the  $z$  direction  $S_z$ , according to the relation

$$F_z \approx kS_z \frac{\partial B_z}{\partial z},$$

for some constant  $k$  and where  $B_z$  is the  $z$  component of the magnetic field. In the Stern-Gerlach experiment the particles are then collected on a plate. Therefore, the position of a particle collected on the plate indicates the force acting on the particle and therefore its angular momentum. As the angular momenta of the particles in the oven are completely random, in principle, classically one would expect them to be distributed continuously on the wall. However the Stern-Gerlach experiment has shown that only two locations on the wall are possible for the particles. These correspond to the values of the spin angular momentum  $+\frac{1}{2}$  and  $-\frac{1}{2}$ . The Stern-Gerlach experiment was one of the first experimental demonstrations of the quantum mechanical measurement postulate. A scheme of the Stern-Gerlach experiment is shown in Figure 1.3. Further discussion can be found, e.g., in [185].



**FIGURE 1.3:** Scheme of a Stern-Gerlach experiment

In the rest of this section we shall refer to the case of measurement of observables with discrete spectrum. The generalization to continuous spectrum

can be obtained without much difficulty using the formalism of the spectral theorem.

### 1.2.2.3 Simultaneous measurement of observables; Compatible and incompatible observable and uncertainty relations

Consider observables with discrete spectrum. A set of  $r$  observables,

$$\{S_1, \dots, S_r\},$$

which are pairwise commuting, i.e.,  $[S_j, S_k] := S_j S_k - S_k S_j = 0$ ;  $j, k = 1, \dots, r$ , are said to be **compatible**. It follows from results in linear algebra and operator theory (see, e.g., [99], [51]) that a set of observables are compatible if and only if they have the same eigenvectors. These eigenvectors form a countable basis of  $\mathcal{H}$ ,  $\{|e_k\rangle\}$ . In this case, each observable  $S_j$  can be expanded, with possibly repeated eigenvalues, as

$$S_j = \sum_k \lambda_{kj} |e_k\rangle\langle e_k|. \quad (1.33)$$

The projections  $|e_k\rangle\langle e_k|$  are common to all the observables, while the eigenvalues  $\lambda_{kj}$  may be different for different  $j$ .

Consider two observables  $A$  and  $B$  and assume, for simplicity, that they are nondegenerate (i.e., every eigenspace has dimension 1). The measurement of  $A$  (on a pure state) gives a result  $\lambda$  and the state collapses into the corresponding eigenstate  $|\lambda\rangle$ . This is an eigenstate of  $B$  as well, so a measurement of  $B$  does not change the state. After the measurement of  $B$  a measurement of  $A$  gives again the value  $\lambda$ , with certainty. In other words, the measurement of  $B$  does not change the result for the measurement of  $A$ . For this reason commuting observables are called *compatible*.

The measurement postulate for the measurement of compatible observables is a generalization of the one for a single observable discussed above. It says that if the state immediately before the measurement is  $|\psi\rangle$  and the results of the measurements are the eigenvalues  $\lambda_1, \dots, \lambda_l$  of the observables  $S_1, S_2, \dots, S_l$ , with projectors (as in (1.25))  $P_1, P_2, \dots, P_l$ , respectively, then the state after the measurement will be

$$|\psi_1\rangle := \frac{P_l P_{l-1} \dots P_1 |\psi\rangle}{\|P_l P_{l-1} \dots P_1 |\psi\rangle\|}. \quad (1.34)$$

The probability of obtaining the  $l$ -ple of results,  $\lambda_1, \dots, \lambda_l$ , is given by

$$\Pr(\lambda_1, \dots, \lambda_l) = \langle \psi | P_l P_{l-1} \dots P_1 | \psi \rangle.$$

Notice that if  $P_k P_j = 0$ , where  $P_k$  and  $P_j$  correspond to two eigenvalues  $\lambda_k$  and  $\lambda_j$  for observables  $S_k$  and  $S_j$ , then the two values  $\lambda_k$  and  $\lambda_j$  cannot appear simultaneously as result of the measurement. The probability of obtaining them at the same time is zero.

**Remark 1.2.3** The maximum number of linearly independent compatible observables for a system of dimension  $n$  is  $n$ , i.e., the maximum number of linearly independent pairwise commuting Hermitian matrices.

**Incompatible** observables are such that

$$AB - BA := [A, B] \neq 0.$$

The **Heisenberg uncertainty relation** quantitatively describes the uncertainty in the measurement of two incompatible observables. Consider an observable  $A$  and its expectation value in some state  $|\psi\rangle$ ,  $\langle A \rangle_\psi$ . Defining  $\Delta A := A - \langle A \rangle_\psi \mathbf{1}$ , the expectation value of  $\Delta A^2$ ,  $\langle \Delta A^2 \rangle_\psi$  measures how much  $A$  is expected to deviate from its mean value at the state  $|\psi\rangle$  and therefore the uncertainty in the measurement of  $A$  at the state  $|\psi\rangle$ . Defining the same quantity for the observable  $B$ ,  $\langle \Delta B^2 \rangle_\psi$ , the uncertainty relation is<sup>14</sup>

$$\langle \Delta A^2 \rangle_\psi \langle \Delta B^2 \rangle_\psi \geq \frac{1}{4} | \langle [A, B] \rangle_\psi |^2.$$

Therefore, a decrease in the uncertainty in the measurement of  $A$  implies an increase in the uncertainty in the measurement of  $B$  and vice versa.

### 1.2.3 Measurements on ensembles

If we perform a measurement of an observable  $A$  on an ensemble described by the density operator

$$\rho := \sum_k w_k |\psi_k\rangle\langle\psi_k|, \quad (1.35)$$

the systems in the state  $|\psi_k\rangle$  giving the result  $\lambda_j$  will be in the state  $\frac{P_j |\psi_k\rangle}{\sqrt{\langle\psi_k|P_j|\psi_k\rangle}}$  immediately after the measurement. The fraction of these systems is proportional both to  $w_k$  and to  $\Pr(\lambda_j)$ , that is, it is proportional to  $w_k \Pr(\lambda_j) = w_k \langle\psi_k|P_j|\psi_k\rangle$ . Therefore the density matrix representing the sub-ensemble of systems which have given result  $\lambda_j$  is

$$\rho_j := \sum_k f w_k \langle\psi_k|P_j|\psi_k\rangle \frac{P_j |\psi_k\rangle\langle\psi_k|P_j}{||P_j|\psi_k\rangle||^2},$$

for some factor  $f$ . Since  $||P_j|\psi_k\rangle||^2 = \langle\psi_k|P_j|\psi_k\rangle$ , this can be written as

$$\rho_j := f P_j \left( \sum_k w_k |\psi_k\rangle\langle\psi_k| \right) P_j = f P_j \rho P_j,$$

---

<sup>14</sup>Note that the operator  $[A, B]$  is in general not Hermitian. However we can still define the ‘expectation value’ as  $\langle [A, B] \rangle_\psi := |\langle\psi|[A, B]|\psi\rangle|$ .

and  $f = [\text{Tr}(P_j \rho P_j)]^{-1}$  so that  $\text{Tr}(\rho_j) = 1$ . From (1.30) and the definition of the density matrix (1.35), it follows (cf. Exercise 1.6) that  $\text{Tr}(P_j \rho P_j)$  is the probability of finding the result  $\lambda_j$  for a measurement on an ensemble  $\rho$ . We have

$$\rho_j = [\text{Tr}(P_j \rho P_j)]^{-1} P_j \rho P_j. \quad (1.36)$$

A measurement where systems that give the same result are collected in sub-ensemble is called **selective**. In a **nonselective** measurement, we may assume to observe the expectation value of the observable  $A$  given by

$$\langle A \rangle_\rho := \sum_k w_k \langle \psi_k | A | \psi_k \rangle = \sum_k w_k \text{Tr}(|\psi_k\rangle\langle\psi_k|A) = \text{Tr}(\rho A). \quad (1.37)$$

After the measurement the state  $\rho$  is transformed into

$$\rho' = \sum_j P_j \rho P_j, \quad (1.38)$$

as the sub-ensembles  $\rho_j$  corresponding to the different outcomes are combined together with coefficients  $\text{Tr}(P_j \rho P_j)$  to form the new ensemble.

Selective and nonselective measurements can also be defined on single systems whose state is represented by a density matrix (cf. 1.1.4.3). In this situation, a selective measurement which gives a result  $\lambda_j$  will project the state onto the state  $\rho_j$  as in (1.36). If the measurement is performed but the result is not read then the state after the measurement will be of the form  $\rho'$  in (1.38).

Because of the property

$$\text{Tr}(\rho A) = \text{Tr}(A \rho),$$

it is often stated that observables and states are ‘dual’ concepts. Both are represented by Hermitian operators and the expectation value of the observable  $A$  if the system is in a state  $\rho$  is equal to the expectation of the observable  $\rho$  in a state  $A$ .

### 1.3 Dynamics of Quantum Systems

The energy of a quantum system is a very important observable that plays a fundamental role in the description of the dynamics of the system. The associated operator is called the **Hamiltonian operator** and it is usually denoted by  $H$ .  $H$  might be time dependent and in fact this is typically the case for quantum control systems as studied here. A time dependent Hamiltonian models an energy exchange with an external system which plays the role of a controller.

### 1.3.1 Schrödinger picture

#### 1.3.1.1 Schrödinger equation

Let  $|\psi(t)\rangle$  be the state of the system at time  $t$ .  $|\psi(t)\rangle$  evolves according to the **Schrödinger equation**

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H(t) |\psi(t)\rangle, \quad (1.39)$$

where  $i$  is the imaginary unit and  $\hbar$  is the Planck constant. If  $|\psi(0)\rangle$  is the initial state we have

$$|\psi(t)\rangle = X(t) |\psi(0)\rangle, \quad (1.40)$$

with the linear operator  $X(t)$  satisfying the differential equation

$$i\hbar \frac{d}{dt} X(t) = H(t) X(t), \quad (1.41)$$

with initial condition equal to the identity operator. The operator  $X(t)$  is called the **evolution operator** or **propagator** and equation (1.41) will be referred to as **Schrödinger operator equation**. It follows from the fact that  $H$  is Hermitian that, at every time  $t$ ,  $X(t)$  is a unitary operator. In particular this implies that the evolution does not modify the norm of the state vector, which we have set equal to one. In the finite dimensional case, in the vector matrix representation,  $X$  can be represented as a unitary matrix of dimension  $n \times n$  ( $n$  being the dimension of the underlying Hilbert space  $\mathcal{H}$ ). In this case  $X$  varies in the Lie group of unitary matrices  $U(n)$ .

#### 1.3.1.2 Schrödinger wave equation

Consider the Schrödinger equation (1.39) for the state  $|\psi\rangle$  in an infinite dimensional Hilbert space. The state  $|\psi\rangle$  has the expansion (1.12). In particular applying  $\langle x|$  to both sides of (1.39), and using  $\psi(x, t) := \langle x|\psi(t)\rangle$ , we obtain

$$i\hbar \frac{\partial \psi(x, t)}{\partial t} = \langle x|H(t)|\psi(t)\rangle. \quad (1.42)$$

Since  $|\psi(t)\rangle$  has the linear expansion (1.12) and  $H$  is a linear operator acting on  $|\psi(t)\rangle$   $H$  is uniquely determined by a linear operator on the coefficient  $\psi(x, t)$  which we still denote by  $H$  and call Hamiltonian operator. Therefore the right hand side of (1.42) can be written, using (1.12), as

$$\langle x|H(t)|\psi(t)\rangle = \int_{\Omega} (H(t)\psi(x_1, t)) \langle x|x_1\rangle dx_1.$$

Using (1.11) and the fundamental property of the Dirac delta function

$$\int_{\Omega} f(x)\delta(x - x_1) dx = f(x_1),$$

for every continuous function  $f = f(x)$ , if  $x_1 \in \Omega$  we obtain the **Scrödinger wave equation**

$$i\hbar \frac{\partial \psi(x, t)}{\partial t} = H(t)\psi(x, t). \quad (1.43)$$

Notice that  $H$  here represents here a linear operator on a space of functions.

In practice, the Hamiltonian operator  $H = H(t)$  in (2.2) is derived from an expression of the energy of the system in the classical setting and a process called *canonical quantization* which associates quantum mechanical operators to classical quantities.

**Example 1.3.1** (Particle in a well of potential) Consider a particle constrained to moving along the direction of the  $x$ -axis. Assume the (classical) total energy given by

$$H_{cl} = \frac{1}{2}p^2 + V(x), \quad (1.44)$$

where  $p$  is the kinematical momentum, in the  $x$  direction, i.e.,  $p = m\dot{x}$ , with  $m$  the mass of the particle, so that the first term represents the kinetic energy.  $V$  is the potential energy. Let us assume that  $V(x)$  has a square shape, i.e., it is zero in an interval  $(0, L)$  and equal to  $V_0$  at the locations 0 and  $L$ . Assuming  $V_0$  very large, we obtain the potential well in [Figure 1.4](#). This scheme may represent a particle with negative charge such as an electron confined between two negatively charged plates. The system is also referred to as a particle in a box. Classically, in the absence of friction and viscous forces, the particle moves back and forth between the two walls of a box. To describe the system in quantum mechanics, we choose a Hilbert space spanned by the position eigenvectors  $\{|x\rangle\}$  so that the state  $|\psi\rangle$  can be written as  $|\psi(t)\rangle := \int_{\mathbf{R}} \psi(x, t)|x\rangle dx$ , with  $\psi(x, t)$ . As we shall see in [Chapter 2](#), the rules of canonical quantization, applied to this situation, associate with  $p$  the operator  $-i\hbar \frac{d}{dx}$ , i.e., differentiation of the wave function by  $x$  and multiplication by  $-i\hbar$ , and with  $V(x)$  the operator multiplying the wave function by  $V(x)$ . Therefore, the quantum mechanical Hamiltonian is given by

$$H = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x), \quad (1.45)$$

so that Schrödinger wave equation (2.2) reads as

$$i\hbar \frac{\partial \psi(x, t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x, t)}{\partial x^2} + V(x)\psi(x, t).$$

As we assume that energy is conserved, we assume that when we measure energy we obtain a given value  $E$  with certainty, namely the system is in an eigenstate of the Hamiltonian operator. Therefore we look for solutions  $\psi$  of the Schrödinger wave equation satisfying  $H\psi = E\psi$ , for some constant  $E$ . Neglecting the dependence on  $t$ , we need to solve

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) + V(x)\psi(x) = E\psi(x), \quad (1.46)$$

which is called the *time independent Schrödinger wave equation*. If  $0 < x < L$ , we set  $V(x) = 0$  and we impose boundary conditions at  $x = 0$  and  $x = L$ . The general solution of (1.46) is of the form

$$\psi(x) = A \sin(kx) + B \cos(kx),$$

with  $k^2 = \frac{2mE}{\hbar^2}$ , with free coefficients  $A$  and  $B$ . Since  $|\psi(x)|^2$  is the probability density of finding the particle in position  $x$ , and the particle is very unlikely to be found very close to the walls, the boundary conditions are chosen as  $\psi(0) = \psi(L) = 0$ , which gives  $B = 0$ . Moreover since  $A$  cannot be zero (it would mean that the particle could not be found anywhere in the well), we must have  $kL = \sqrt{\frac{2mE}{\hbar^2}}L = n\pi$ , for  $n = 1, 2, \dots$ . This gives that the possible values for the energy  $E$  are in a discrete set and are given by

$$E_n = \frac{n^2\pi^2\hbar^2}{2mL^2}.$$

The solution  $\psi$ , corresponding to energy  $E_n$ , is given by

$$\psi_n(x) = A \sin\left(\frac{n\pi}{L}x\right),$$

where  $A$  is any number with absolute value  $\sqrt{\frac{2}{L}}$  obtained by the requirement that  $\int_0^L \psi^*(x)\psi(x)dx = 1$ . Plots of  $\psi_n$  for the first three energy levels are shown in [Figure 1.4](#). Notice in particular, for  $n \geq 2$ , the existence of ‘nodes’, i.e., points in the interval  $(0, L)$  where  $\psi_n$  is equal to zero, i.e., the particle has probability zero to be found.

The system of a particle in a quantum well of potential, and various generalizations of it, serve as models for several physical systems and in particular for quantum dots (see, e.g., [68]). It is possible to introduce a control in the model by shaping the potential energy with time (see, e.g., [200]) in order to obtain a desired behavior for the wave function.

### 1.3.1.3 Liouville’s equation and evolution of ensembles

Assume the initial state of the system is represented by a density matrix

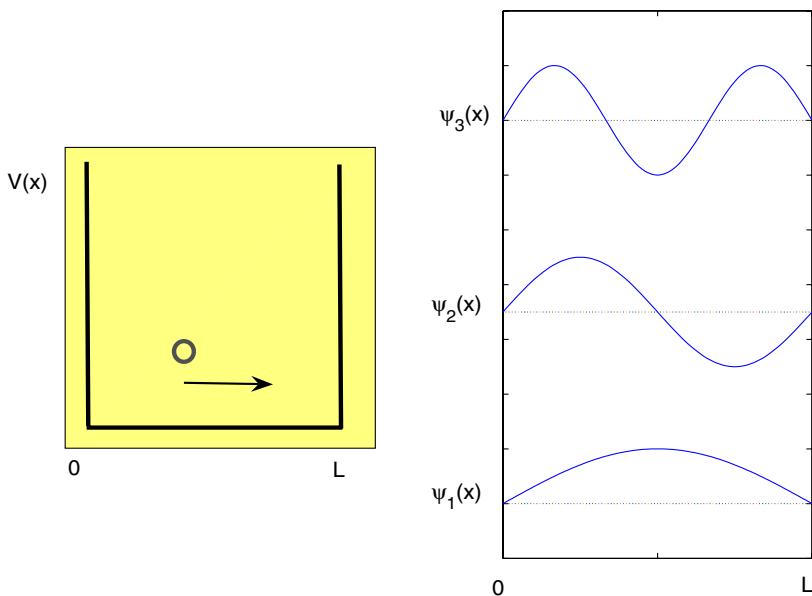
$$\rho(0) := \sum_k w_k |\psi_k(0)\rangle\langle\psi_k(0)|.$$

Then, according to (1.40),  $\rho$  varies with time as

$$\rho(t) = \sum_k w_k X(t) |\psi_k(0)\rangle\langle\psi_k(0)| X^\dagger(t) = X(t)\rho(0)X^\dagger(t), \quad (1.47)$$

where  $X(t)$  is the solution of (1.41). By differentiating (1.47) taking into account (1.41) we obtain the differential equation for the density matrix known as **Liouville’s equation**

$$i\hbar \frac{d}{dt} \rho = [H(t), \rho] := H(t)\rho - \rho H(t). \quad (1.48)$$



**FIGURE 1.4:** One dimensional motion of a single particle in a quantum well and resulting wave functions for the first three values of the energy.

The commutator  $[A, B] := AB - BA$  of two operators (or matrices) will play a crucial role in what follows.

### 1.3.2 Heisenberg and interaction picture

In classical mechanics, the description of a system is done through observables (e.g., momentum, position, energy) rather than states. Therefore, it seems reasonable that a better parallel with classical mechanics could be obtained if we study the dynamics of observables rather than that of states. This approach is called the **Heisenberg picture**. Given an observable  $A$ , we define an operator function of time,  $A^H = A^H(t)$ , as

$$A^H(t) := X^\dagger(t)AX(t), \quad (1.49)$$

where  $X$  is the propagator in (1.41). From (1.49) and (1.41), it follows that  $A^H(t)$  satisfies

$$i\hbar \frac{dA^H}{dt} = [A^H, H], \quad (1.50)$$

where  $H$  is the Hamiltonian of the system in (1.41), possibly function of time. This is known as the **Heisenberg equation**. Notice that the expectation

value of the observable  $A$  at time  $t$  can be calculated both as  $\text{Tr}(A^H(t)\rho(0))$  and as  $\text{Tr}(A\rho(t))$  since from (1.47) we have

$$\text{Tr}(A\rho(t)) = \text{Tr}(AX(t)\rho(0)X^\dagger(t)) = \text{Tr}(X^\dagger(t)AX(t)\rho(0)) = \text{Tr}(A^H(t)\rho(0)),$$

using the property  $\text{Tr}(AB) = \text{Tr}(BA)$ . More generally, in the Heisenberg picture, one considers the state of the system constant and the observables varying with time and studies the dynamics of observables using Heisenberg equation (1.50).

In many cases the Hamiltonian  $H$  has the form

$$H := H_0 + H_I(t),$$

with  $H_0$  a constant Hamiltonian (sometimes called *internal Hamiltonian*) and  $H_I$  a (possibly) time varying Hamiltonian (sometimes called *interaction Hamiltonian*). Given the Schrödinger equation

$$i\hbar \frac{d}{dt}|\psi(t)\rangle = (H_0 + H_I(t))|\psi(t)\rangle,$$

it is possible to eliminate the term  $H_0$  by making a change of coordinates

$$|\psi\rangle \rightarrow e^{\frac{iH_0 t}{\hbar}}|\psi\rangle. \quad (1.51)$$

Defining  $|\tilde{\psi}\rangle := e^{\frac{iH_0 t}{\hbar}}|\psi\rangle$ , we obtain from (1.39)

$$i\hbar \frac{d}{dt}|\tilde{\psi}\rangle = H'_I(t)|\tilde{\psi}\rangle, \quad (1.52)$$

with

$$H'_I(t) := e^{\frac{iH_0 t}{\hbar}} H_I(t) e^{-\frac{iH_0 t}{\hbar}}. \quad (1.53)$$

This description of the dynamics is called **interaction picture**. In many cases, one considers a basis of eigenvectors of the Hamiltonian  $H_0$  which is therefore diagonal in this basis. Each eigenvector corresponds to a state having energy given by the corresponding eigenvalue of  $H_0$ . In the given basis, the matrix representation of  $H_0$  is diagonal. Therefore the coordinate transformation (1.51) does not affect the relative magnitude of the components of the state vector which are the same for  $|\psi\rangle$  and  $|\tilde{\psi}\rangle$ .

## 1.4 Notes

### 1.4.1 Interpretation of quantum dynamics as information processing

In quantum information theory, the state  $|\psi\rangle$  of a (finite dimensional) quantum system encodes information. In particular, in typical implementations,

the information is encoded in a number of two level systems called *qubits*. If a particular system performs a given unitary evolution  $X_T$  during an interval of time  $[0, T]$ , the information in  $|\psi(0)\rangle$  is manipulated to obtain

$$|\psi(T)\rangle = X_T |\psi(0)\rangle.$$

$X_T$  is also called a *quantum logic gate* as it may be seen as performing a *logic operation* on  $|\psi(0)\rangle$ .

As an example, consider a two level system with basis given by  $|0\rangle$  and  $|1\rangle$ , and with constant Hamiltonian  $H$  defined by  $H|0\rangle = -i\hbar|1\rangle$ ,  $H|1\rangle = i\hbar|0\rangle$ . If we write  $|\psi(t)\rangle = \alpha_1(t)|0\rangle + \alpha_2(t)|1\rangle$  the differential equation corresponding to the Schrödinger equation for  $\vec{\psi}(t) = \begin{pmatrix} \alpha_1(t) \\ \alpha_2(t) \end{pmatrix}$  is

$$\frac{d}{dt} \vec{\psi} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \vec{\psi}.$$

This equation can be solved explicitly. The solution is given by

$$\vec{\psi}(t) = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix} \vec{\psi}(0).$$

For  $T = \frac{\pi}{2}$  we obtain

$$\vec{\psi}(T) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \vec{\psi}(0).$$

This dynamics performs the logic operation *NOT* in that it *inverts* the state which is transferred from  $|1\rangle$  to  $|0\rangle$  and from  $|0\rangle$  to  $-|1\rangle$  (notice the minus sign plays the role of an overall phase factor. The state  $-|1\rangle$  is physically the same as the state  $|1\rangle$ ). In general, and this is one of the main features of *quantum computation* as compared to *classical computation*, the state  $|\psi\rangle$  may not be in one of the two basis states but in a superposition of the two before and/or after the computation (evolution). This fact has been shown useful in several algorithms, known as *quantum algorithms* which are shown to be more powerful than classical algorithms (see, e.g., [155] for an introduction to quantum computation). A computation more general than the one described above is performed by a *cascade* of  $r$  quantum logic gates  $X_1, \dots, X_r$ . After  $r$  computations with an overall time  $T$ ,  $|\psi(T)\rangle = X_r X_{r-1} \cdots X_1 |\psi(0)\rangle$ . The computations  $X_1, \dots, X_r$  can be performed by the same system by using different Hamiltonians  $H$  in different intervals. The Hamiltonians may be changed by changing the experimental setup (cf. [section 9.3](#)) and/or by interaction with an external field. The choice of an appropriate sequence of Hamiltonians is a control problem.

#### 1.4.2 Direct sum versus tensor product for composite systems

We have said that the Hilbert space describing the total state of two systems with Hilbert spaces  $\mathcal{H}_1$  and  $\mathcal{H}_2$  is the tensor product space  $\mathcal{H}_1 \otimes \mathcal{H}_2$ . This is

often accepted as a postulate of quantum mechanics. One might wonder why the direct sum of vector spaces,  $\mathcal{H}_1 \oplus \mathcal{H}_2$ , is not used instead.

Consider for simplicity two finite dimensional systems of dimension  $n_1$  and  $n_2$ , and energies represented by Hamiltonians  $H_1$  and  $H_2$ . Consider the generic case where these Hamiltonians are not degenerate so that there are  $n_1$  possible eigenvalues (values for energy) for system 1 and  $n_2$  possible values of the energy for system 2. It is intuitive that the possible values for the energies for the total system are the sums of the two values of the energies. The number of these, in the generic case, will be  $n_1 n_2$ , the dimension of the tensor product, rather than  $n_1 + n_2$ , the dimension of the direct sum. The total energy will have to be an operator on an  $n_1 n_2$ -dimensional space.<sup>15</sup>

---

## 1.5 Exercises

The goal of the following exercises is to acquaint the student with the notations of quantum mechanics and with translating them into the language of linear algebra.

### Exercise 1.1

Prove formula (1.5).

In the Exercises 1.2 through 1.5, we assume finite dimensional quantum systems.

### Exercise 1.2

Consider a Hilbert space  $\mathcal{H}$ , with basis  $\{|e_j\rangle\}$ ,  $j = 1, \dots, n$ . A vector  $|\psi\rangle = \sum_{j=1}^n \alpha_j |e_j\rangle$  is represented by the column vector  $\vec{\psi} := \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$ . If  $|\phi\rangle = \sum_{j=1}^n \beta_j |e_j\rangle$  and  $|\psi\rangle = \sum_{j=1}^n \alpha_j |e_j\rangle$ , find the matrix representation of the outer product  $|\psi\rangle\langle\phi|$ . Verify that it is equal to  $\vec{\psi}\vec{\phi}^\dagger$ .

**Exercise 1.3** An ensemble of two level systems has  $\frac{1}{3}$  of the systems in the state

$$|\psi_1\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{i}{\sqrt{2}}|1\rangle,$$

and  $\frac{2}{3}$  in the state

$$|\psi_2\rangle = |1\rangle.$$

---

<sup>15</sup>The interpretation reported in this subsection was suggested to the author by Raffaele Romano.

Write an expression for the density matrix  $\rho$  representing the total system in terms of the states  $|0\rangle$  and  $|1\rangle$  and then use this along with the result of Exercise 1.2 to write a matrix for the operator  $\rho$  in the basis  $|0\rangle$  and  $|1\rangle$ .

**Exercise 1.4 (Tensor Product and Kronecker Product)** Consider two vectors  $|\psi\rangle$  and  $|\phi\rangle$  in Hilbert spaces  $\mathcal{H}$  and  $\mathcal{H}_2$ , respectively, with given bases  $\{|e_j\rangle\}$  and  $\{f_k\}$ , respectively. Consider the column vector representations of  $|\psi\rangle$  and  $|\phi\rangle$ , i.e.,  $\vec{\psi}$  and  $\vec{\phi}$ . Consider the basis of the space  $\mathcal{H}_1 \otimes \mathcal{H}_2$ , defined as  $\{|e_j\rangle \otimes |f_k\rangle\}$  with the ordering convention defined as follows:  $|e_{j_1}, f_{k_1}\rangle$  comes before  $|e_{j_2}, f_{k_2}\rangle$  if  $j_1 < j_2$  or, when  $j_1 = j_2$ , if  $k_1 < k_2$ . Verify that the matrix vector representation of the tensor product  $|\psi\rangle \otimes |\phi\rangle$ , in the basis  $\{|e_j\rangle \otimes |f_k\rangle\}$ , is the Kronecker product of the two matrix vector representations. In other terms, verify that the following commutative diagram, which is a special case of the one in subsection 1.1.5, holds:

$$\begin{array}{ccc}
 & \text{Tensor Product} & \\
 \{|\psi\rangle, |\phi\rangle\} & \xrightarrow{\hspace{2cm}} & |\psi\rangle \otimes |\phi\rangle \\
 \downarrow \text{Matrix Vector Representation} & & \downarrow \text{Matrix Vector Representation} \\
 \{\vec{\psi}, \vec{\phi}\} & \xrightarrow{\hspace{2cm}} & \vec{\psi} \otimes \vec{\phi} \\
 & \text{Kronecker Product} &
 \end{array}$$

**Exercise 1.5** The matrix representation of the energy operator  $H$  of a two level system is

$$H = \begin{pmatrix} \frac{3}{2} & \frac{i}{2} \\ -\frac{i}{2} & \frac{3}{2} \end{pmatrix}, \quad (1.54)$$

in a given orthonormal basis  $|0\rangle, |1\rangle$ .

- Find the two possible values  $E_1$  and  $E_2$  of the energy for this system.
- Using the spectral decomposition of the matrix (1.54) write the energy operator in Dirac notation, namely as the sum of appropriate outer products of states. What are the projectors  $P_j$  in this case?

- Assume the state of the system is  $|0\rangle$ . In what states can we find the system after a measurement of the energy, and with what probability?
- Write a differential equation which gives the evolution of the components of the column vector representing an arbitrary state  $|\psi\rangle$  (in the basis  $|0\rangle, |1\rangle$ ). Assume  $\hbar = 1$ .

**Exercise 1.6** Use the measurement postulate (1.30) and the definition of density matrix (1.35), to show that  $\text{Tr}(P_j \rho P_j)$  is the probability of finding the result  $\lambda_j$  for a measurement on an ensemble  $\rho$ .

**Exercise 1.7** The spectral theorem discussed in subsection 1.2.2 is a generalization of the spectral decomposition for matrices. Consider the matrix

$$A := \frac{1}{3} \begin{pmatrix} 4 & 2 & 1 \\ 2 & 4 & -1 \\ 1 & -1 & 1 \end{pmatrix},$$

and find its spectral decomposition (1.26) by identifying eigenvalues  $\lambda_j$  and corresponding projections  $P_j$ .

**Exercise 1.8** Derive equation (1.52).

**Exercise 1.9** Prove that the operators defined in (1.25) are such that the product of two of them is equal to  $\mathbf{0}$ .

**Exercise 1.10** Prove formula (1.31), in the continuous spectrum case, i.e., that the expectation value of  $A$  at the state  $|\psi\rangle$ ,  $\langle A \rangle_\psi$ , is equal to  $\langle \psi | A | \psi \rangle$ . Use the spectral theorem and the fact that the element of the spectral family corresponding to an interval  $[\lambda_1, \lambda_2]$  of the spectrum is

$$\int_{\lambda_1}^{\lambda_2} dP_A.$$

# Chapter 2

---

## *Modeling of Quantum Control Systems; Examples*

Quantum systems whose dynamics depends on one or more control functions will be named *quantum control systems*. In a typical situation the objects of control are charged particles while the control functions are appropriately shaped electromagnetic fields. Therefore, to derive the models, in this chapter we describe elements of the theory of field-particle interaction. We start with the classical theory of electrodynamics and show how to obtain the quantum mechanical description through a process called *canonical quantization*. We present physical examples of (finite dimensional) quantum control systems of interest in applications and we highlight the fact that such systems have a common mathematical structure, as bilinear systems whose evolution is determined by an element of the Lie group of unitary matrices. In a parallel to classical control theory, we identify a *state*, an *input*, and an *output* in our description of quantum control systems.

---

### 2.1 Quantum Theory of Interaction of Particles and Field

Consider a system of  $N$  charged particles in an electromagnetic field. The Hilbert space describing this system may be chosen to be spanned by the eigenvectors of the *position observable*. This is called the *position representation*. More precisely, if there are  $N$  particles involved, with coordinates  $\{x_j, y_j, z_j\}$ ,  $j = 1, 2, \dots, N$ , the Hilbert space is spanned by the eigenvectors  $|\vec{r}\rangle := |x_1, y_1, z_1, \dots, x_N, y_N, z_N\rangle$ . These eigenvectors correspond to the eigenvalues  $x_1, y_1, z_1, \dots, x_N, y_N, z_N$  of the observables associated to the coordinates of particles 1, 2, ...,  $N$ . The state of the system  $|\psi(t)\rangle$  is expanded as (see (1.12) in Chapter 1)

$$|\psi(t)\rangle = \int_{\mathbf{R}^{3N}} \psi(\vec{r}, t) |\vec{r}\rangle d\vec{r}. \quad (2.1)$$

The function  $\psi(\vec{r}, t) = \langle \vec{r} | \psi \rangle$ ,  $\in L^2(\mathbf{R}^{3N})$ ,  $\forall t$ , is the wave function and the function  $|\psi(\vec{r}, t)|^2$  is the (time dependent) probability density for the particles to be in a particular region of space. More precisely, if there are  $N$  particles involved in the process, with coordinates  $\{x_j, y_j, z_j\}$ ,  $j = 1, 2, \dots, N$ , let  $d\vec{r}$  be the infinitesimal volume element in  $3N$ -dimensional space. Then the probability of being in the  $3N$ -dimensional region  $I$  is

$$\Pr(I, t) := \int_I |\psi(\vec{r}, t)|^2 d\vec{r}.$$

This is a special case of (1.32).

The wave function  $\psi(\vec{r}, t)$  satisfies the *Schrödinger wave equation*

$$i\hbar \frac{\partial \psi(\vec{r}, t)}{\partial t} = H(t)\psi(\vec{r}, t). \quad (2.2)$$

The Hamiltonian operator is obtained from the classical expression of the energy of the system<sup>1</sup> and the process of canonical quantization which associates quantum mechanical operators to physical quantities. The quantum mechanical modeling of a system of charged particles interacting with an electromagnetic field involves the study of Lagrangian and Hamiltonian mechanics (cf. [Appendix B](#)) for systems with an uncountable number of degrees of freedom. This study is the subject of books on (nonrelativistic) quantum electrodynamics (see, e.g., [50] [95]). In the remainder of this section, we describe the main points.

**Notation:** In the remainder of this section and in the following two sections we shall often deal with classical quantities and their corresponding quantum mechanical operators. We shall use the convention of denoting by  $\hat{A}$  the quantum mechanical operator corresponding to the quantity  $A$ . In particular, in these sections, the classical energy is denoted by  $H$  while the quantum mechanical Hamiltonian is denoted by  $\hat{H}$ .

### 2.1.1 Classical electrodynamics

#### 2.1.1.1 Total energy and Maxwell equations

We start with the classical expression for the total energy  $H$  of a system of moving particles and the associated electromagnetic field. This is the sum of

---

<sup>1</sup>More precisely, it is obtained from the associated Hamiltonian of classical Hamiltonian mechanics. There are cases where this Hamiltonian does not coincide with the energy (cf. [50]) but these cases will not be treated here.

the kinetic energy of the particles and the energy of the field, i.e.,<sup>2</sup>

$$H = \sum_{j=1}^N \frac{1}{2} m_j \left( \frac{d\vec{r}_j}{dt} \right)^2 + \frac{\epsilon_0}{2} \int_{\mathbf{R}^3} \vec{E}^2(\vec{r}) + c^2 \vec{B}^2(\vec{r}) d\vec{r}, \quad (2.3)$$

where  $\vec{E}$  and  $\vec{B}$  are the electric and magnetic fields respectively,  $c$  is the speed of light in vacuum,  $\epsilon_0$  is the free space *electric permittivity*. The quantity

$$\mu_0 := \frac{1}{\epsilon_0 c^2}$$

is called the *magnetic permeability*.<sup>3</sup> The potential energy of the particles is not explicitly present in (2.3) because it is contained in the field energy.

Fields and particles interact in that the motion of the particles is determined by the field and at the same time the particles establish the field. Moreover, the interaction among particles happens through the field.

More specifically, once the motion of the particles is given the electromag-

---

<sup>2</sup>A justification of this formula and the following formula (2.27) in terms of Lagrangian and Hamiltonian mechanics is presented in [Appendix B](#). In particular see (B.31).

<sup>3</sup>In the following we shall often use the symbol  $\nabla$  of vector calculus. The symbol  $\nabla$  is defined as

$$\nabla := \frac{\partial}{\partial x} \vec{i} + \frac{\partial}{\partial y} \vec{j} + \frac{\partial}{\partial z} \vec{k}.$$

If  $\phi$  is a function  $\nabla\phi$  is the vector field

$$\nabla\phi = \frac{\partial\phi}{\partial x} \vec{i} + \frac{\partial\phi}{\partial y} \vec{j} + \frac{\partial\phi}{\partial z} \vec{k}.$$

If  $\vec{A}$  is a vector field

$$\vec{A} = M\vec{i} + N\vec{j} + P\vec{k}$$

then  $\nabla \cdot \vec{A}$  is a function equal to the formal scalar product of  $\nabla$  and  $\vec{A}$ . It is defined by

$$\nabla \cdot \vec{A} = \frac{\partial M}{\partial x} + \frac{\partial N}{\partial y} + \frac{\partial P}{\partial z}.$$

Analogously,  $\nabla \times \vec{A}$  is a vector field equal to the formal cross product of  $\nabla$  and  $\vec{A}$ , which is

$$\nabla \times \vec{A} = \left( \frac{\partial P}{\partial y} - \frac{\partial N}{\partial z} \right) \vec{i} + \left( \frac{\partial M}{\partial z} - \frac{\partial P}{\partial x} \right) \vec{j} + \left( \frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} \right) \vec{k}.$$

The Laplacian operator  $\nabla^2$  acting on a vector field  $\vec{A} = M\vec{i} + N\vec{j} + P\vec{k}$  gives  $\nabla^2 \vec{A} = \nabla^2 M\vec{i} + \nabla^2 N\vec{j} + \nabla^2 P\vec{k}$  with, for a smooth function  $f$ ,  $\nabla^2 f := \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$ .

netic field evolves according to Maxwell's equations:<sup>4</sup>

$$\nabla \cdot \vec{E} = \frac{\rho}{\epsilon_0}, \quad (2.4)$$

$$\frac{1}{\mu_0}(\nabla \times \vec{B}) - \epsilon_0 \frac{\partial \vec{E}}{\partial t} = \vec{J}, \quad (2.5)$$

$$\nabla \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0, \quad (2.6)$$

$$\nabla \cdot \vec{B} = 0. \quad (2.7)$$

Here we have used the symbol  $\rho = \rho(\vec{r}, t)$  (function of position and time) for the *charge density*<sup>5</sup> (charge per unit volume), and  $\vec{J} = \vec{J}(\vec{r}, t)$  for the *current density* (current across unit area). The units used in (2.4)-(2.7) are *SI* units. In particular  $\vec{J}$  is measured in  $Coul/(m^2 \times s)$  while the charge density  $\rho$  is measured in  $Coul/m^3$ .

For a discrete distribution of  $N$  charges  $q_j$  at locations  $\vec{r}_j$ ,  $j = 1, \dots, N$ , the charge density is given by

$$\rho(\vec{r}) = \sum_{j=1}^N q_j \delta(\vec{r} - \vec{r}_j), \quad (2.8)$$

while  $\vec{J}$  is given by

$$\vec{J}(\vec{r}) = \sum_{j=1}^N q_j \dot{\vec{r}}_j \delta(\vec{r} - \vec{r}_j). \quad (2.9)$$

---

<sup>4</sup>In the Maxwell equations, the first equation (2.4) is called *Gauss's law*. It expresses the fact that the flux integral across any closed surface is equal to the net charge enclosed inside the surface (multiplied by the factor  $\frac{1}{\epsilon_0}$ ). In formulas we have

$$\int_S \vec{E} \cdot \vec{n} dS = \frac{1}{\epsilon_0} \int_V \rho dV,$$

which is obtained by performing a space integral of both sides of (2.4) and then applying the divergence theorem to the left hand side. The third equation (2.6) expresses the so called Faraday's electromagnetic induction namely: a time varying magnetic field creates an electric field. It is also called *Faraday's law*. The second equation (2.5), called *Ampere's law*, expresses the reciprocal fact, that a time varying electric field, along with a moving charge  $\vec{J}$ , generates a magnetic field. The last equation, (2.7), expresses the fundamental property that, unlike the electric field, which has electric charges as point sources, the magnetic field has no point source. It can also be written applying the divergence theorem as

$$\int_S \vec{B} \cdot \vec{n} dS = 0,$$

where  $S$  is any closed surface.

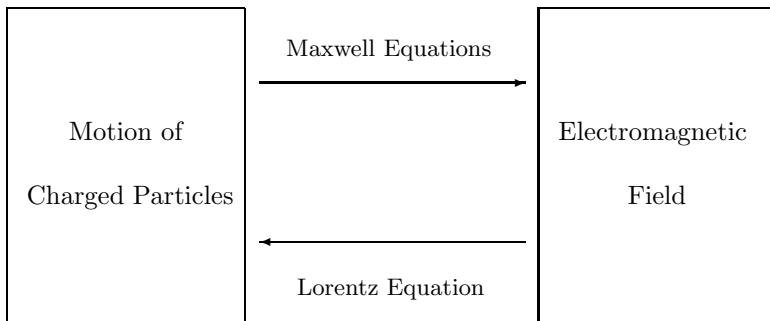
<sup>5</sup>To follow standard notations, we denote the charge density function by  $\rho$  as there is no possibility of confusion with the density matrix introduced in the previous chapter.

Given an electromagnetic field, the dynamics of a particle of charge  $q$  is determined by *Lorentz equation*

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B}), \quad (2.10)$$

where  $\vec{F}$  denotes the force acting on the particles and the fields are evaluated at the point where the particle is located.

Summarizing the system of particles and field evolves according to the following diagram. The motion of the particles determines the field according to Maxwell's equations. The field has a back-action on the charges according to Lorentz equation. The variables associated with the charged particles are the positions  $\vec{r}_j$  and the velocities  $\dot{\vec{r}}_j$  of the (discrete) set of particles  $j = 1, \dots, N$ . The variables associated with the field are the electric field  $\vec{E} := \vec{E}(\vec{r})$ , and the magnetic field  $\vec{B} := \vec{B}(\vec{r})$ , parametrized by the continuous space variable  $\vec{r}$ .



### 2.1.1.2 Vector and scalar potential

It is common practice, and it is instrumental for the transition from the classical to the quantum theory, to express the electric and magnetic field in terms of vector and scalar potentials. Doing this, the two *homogeneous* Maxwell equations, (2.6) and (2.7), are automatically satisfied. More specifically, equation (2.7) and the vector identity

$$\nabla \cdot (\nabla \times \vec{A}) \equiv 0, \quad \forall \vec{A} := \vec{A}(\vec{r}), \quad (2.11)$$

suggest that (2.7) is automatically satisfied if we express  $\vec{B}$  as the curl of another vector field, i.e., we set

$$\vec{B} := \nabla \times \vec{A}. \quad (2.12)$$

The vector field  $\vec{A}$  is called **vector potential**. Replacing (2.12) into (2.6), we can write (2.6) as

$$\nabla \times (\vec{E} + \frac{\partial \vec{A}}{\partial t}) = 0. \quad (2.13)$$

The vector identity

$$\nabla \times \nabla \phi = 0, \quad (2.14)$$

valid for every (smooth) scalar function  $\phi = \phi(\vec{r})$ , suggests to write  $\vec{E}$  as

$$\vec{E} = -\frac{\partial \vec{A}}{\partial t} - \nabla \phi, \quad (2.15)$$

for some scalar function  $\phi$ . This way, (2.13) and therefore (2.6) is automatically satisfied. The function  $\phi$  is called the **scalar potential**. The scalar and vector potentials have to satisfy equations that are obtained by placing (2.12) and (2.15) into the (inhomogeneous) first two Maxwell equations (2.4), (2.5). The equations obtained are

$$\nabla^2 \phi + \frac{\partial}{\partial t} (\nabla \cdot \vec{A}) = -\frac{\rho}{\epsilon_0}, \quad (2.16)$$

$$\nabla(\nabla \cdot \vec{A}) - \nabla^2 \vec{A} + \frac{1}{c^2} \frac{\partial}{\partial t} \nabla \phi + \frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} = \mu_0 \vec{J}, \quad (2.17)$$

where, in the last one, we used the vector identity

$$\nabla \times (\nabla \times \vec{A}) = \nabla(\nabla \cdot \vec{A}) - \nabla^2 \vec{A}.$$

The variables  $\vec{A}$  and  $\phi$  replace  $\vec{E}$  and  $\vec{B}$  as field variables and equations (2.16) and (2.17) replace the Maxwell's equations in determining how the motion of charged particles affects the field.

### 2.1.1.3 Gauge and gauge transformation

The observable quantities in the Maxwell equations are the electric and magnetic fields  $\vec{E}$  and  $\vec{B}$ . The vector and scalar potentials  $\vec{A}$  and  $\phi$  themselves are not observable but have been defined to express the Maxwell equations in a simpler form. There exists a certain freedom in the choice of the **gauge**, i.e., the pair  $\{\vec{A}, \phi\}$ . In particular, given a gauge,  $\vec{A}$  and  $\phi$ , one can define a new one,  $\vec{A}'$  and  $\phi'$ , equivalent to  $\vec{A}$  and  $\phi$  in that it gives the same electric and magnetic fields. This is achieved through the **gauge transformation**

$$\vec{A}' := \vec{A} + \nabla \eta, \quad (2.18)$$

$$\phi' := \phi - \frac{\partial \eta}{\partial t}, \quad (2.19)$$

where  $\eta$  is a function of  $\vec{r}$  and  $t$ . Different gauge transformations can be chosen which simplify equations (2.16) and (2.17) in different ways and are more or less appropriate depending on the physical situation.

The **Coulomb gauge** is the most used gauge in problems where the classical electromagnetic theory is translated into quantum mechanics. It is defined

by choosing  $\vec{A}$  so that<sup>6</sup>

$$\nabla \cdot \vec{A} = 0. \quad (2.20)$$

If we use (2.20) in (2.16), (2.17), we obtain the two equations

$$\nabla^2 \phi = -\frac{\rho}{\epsilon_0}, \quad (2.21)$$

$$-\nabla^2 \vec{A} + \frac{1}{c^2} \frac{\partial}{\partial t} \nabla \phi + \frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} = \mu_0 \vec{J}, \quad (2.22)$$

the first of which is *Poisson's equation*.

#### 2.1.1.4 Separation of the electrostatic and radiative energy term

Consider now the term describing the electromagnetic field energy in (2.3),

$$\frac{\epsilon_0}{2} \int_{\mathbf{R}^3} \vec{E}^2(\vec{r}) + c^2 \vec{B}^2(\vec{r}) d\vec{r}. \quad (2.23)$$

A portion of the energy in (2.23) is due to the Coulomb interaction between particles. This is called *electrostatic energy*. The remaining energy is called *radiative energy*.

To separate the electrostatic and radiative energy in (2.23), we recall (see, e.g., [14] pp. 92-93) that every sufficiently smooth vector field  $\vec{V}$ , going to zero sufficiently fast at infinity, can be written as<sup>7</sup> (cf. Exercise 2.4)

$$\vec{V} = \vec{V}_L + \vec{V}_T. \quad (2.24)$$

The *irrotational* (also called *longitudinal*) part,  $\vec{V}_L$ , is by definition a vector field such that

$$\nabla \times \vec{V}_L = 0,$$

while the *divergence free* part (also called *transversal* part),  $\vec{V}_T$ , is a vector field satisfying

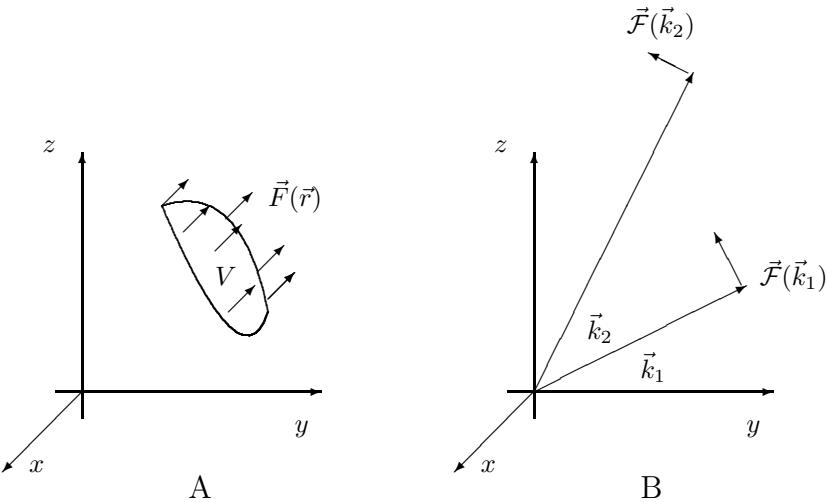
$$\nabla \cdot \vec{V}_T = 0.$$

Condition (2.20) for the Coulomb gauge means that the vector potential is chosen purely transversal (cf. [Figures 2.1](#) and [2.2](#)).

<sup>6</sup>Given an arbitrary  $\vec{A}$ , the vector potential  $\vec{A}'$  in the Coulomb gauge is obtained by choosing  $\eta$  solution of Poisson's equation

$$-\nabla \cdot \vec{A} = \nabla^2 \eta.$$

<sup>7</sup>This result is known as *Helmholtz decomposition theorem*.



**FIGURE 2.1:** Assume  $\vec{F}$  represents the velocity of a fluid flow with constant density. For any region  $V$  in space the balance of material in and out of  $V$  is given by the surface integral  $\int_{\partial V} \vec{F} \cdot \vec{n} dS$ . According to Gauss's divergence theorem this integral is equal to the volume integral  $\int_V \nabla \cdot \vec{F} d\vec{r}$ . For a divergence free vector field  $\vec{F}$  this integral is always zero for every region  $V$ . So the balance of material in and out of  $V$  is zero for any region  $V$  (cf. Part A). The terminology ‘transversal’ derives from consideration of the spatial Fourier transform (cf. Exercise 2.4) of the vector field  $\vec{F}$ ,  $\vec{\mathcal{F}}(\vec{k})$ . This is perpendicular to  $\vec{k}$  for every vector  $\vec{k}$ , as in Part B.

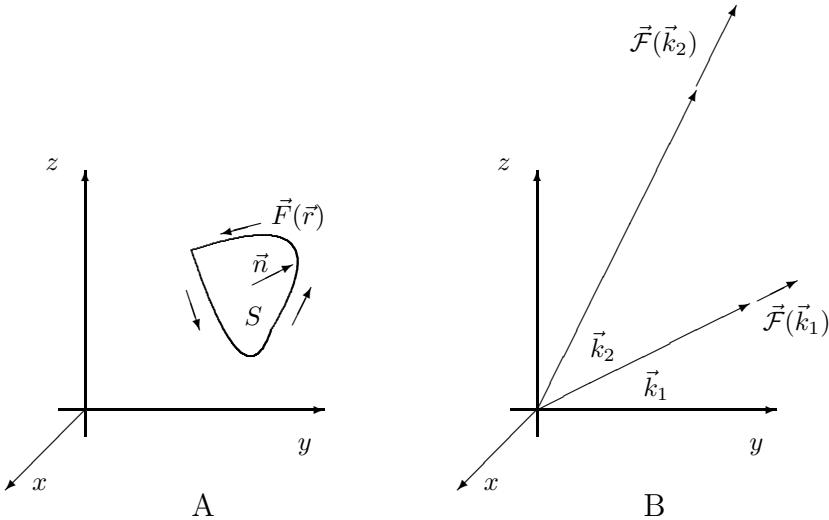
Under appropriate boundary conditions (the vector fields go to zero at infinity faster than  $\frac{1}{|\vec{r}|^2}$ ), we have<sup>8</sup>

$$\int_{\mathbf{R}^3} \vec{V}_L(\vec{r}) \cdot \vec{V}_T(\vec{r}) d\vec{r} = 0. \quad (2.25)$$

<sup>8</sup>This fact is an immediate consequence of *Parseval-Plancherel theorem*

$$\int_{\mathbf{R}^3} \vec{F}^*(\vec{r}) \vec{G}(\vec{r}) d\vec{r} = \int_{\mathbf{R}^3} \vec{\mathcal{F}}^*(\vec{k}) \cdot \vec{\mathcal{G}}(\vec{k}) d\vec{k},$$

where  $\vec{\mathcal{F}}$  and  $\vec{\mathcal{G}}$  are the spatial Fourier transforms of the fields  $\vec{F}$  and  $\vec{G}$ , respectively (cf. Exercise 2.4). It follows observing that the spatial Fourier transform of a longitudinal (transversal) vector field is parallel (orthogonal) to  $\vec{k}$  (at every point  $\vec{k}$ ) so that the Fourier transforms of longitudinal and transverse fields are orthogonal (see also, e.g., [50] section 1.B).



**FIGURE 2.2:** According to Stokes' theorem the *circulation* of a vector field  $\vec{F}$  along the border of a surface  $S$  in  $\mathbf{R}^3$ , i.e., the line integral  $\int_{\partial S} \vec{F} \cdot d\vec{l}$ , is equal to the surface integral across  $S$  of the curl of  $\vec{F}$ , that is  $\int_S \nabla \times \vec{F} \cdot \vec{n} dS$ . For an irrotational vector field  $\vec{F}$ ,  $\nabla \times \vec{F}(\vec{r}) = 0$  for every  $\vec{r}$ . Therefore the circulation is zero for every surface. The terminology ‘transversal’ derives from the consideration of the spatial Fourier transform (cf. Exercise 2.4) of the vector field  $\vec{F}$ ,  $\vec{F}(\vec{k})$ . This is parallel to  $\vec{k}$  for every vector  $\vec{k}$ , as in Part B.

Applying this to the energy of the field in (2.3) we can write

$$\frac{\epsilon_0}{2} \int_{\mathbf{R}^3} \vec{E}^2(\vec{r}) + c^2 \vec{B}^2(\vec{r}) d\vec{r} = \frac{\epsilon_0}{2} \int_{\mathbf{R}^3} \vec{E}_L^2(\vec{r}) d\vec{r} + \frac{\epsilon_0}{2} \int_{\mathbf{R}^3} \vec{E}_T^2(\vec{r}) + c^2 \vec{B}_T^2(\vec{r}) d\vec{r}.$$

The first term is the electrostatic potential energy of the system due to the Coulomb interaction among the particles. In fact, for a discrete distribution of charges (2.8) we can prove that (see the following Remark 2.1.1)

$$\frac{\epsilon_0}{2} \int_{\mathbf{R}^3} \vec{E}_L^2(\vec{r}) d\vec{r} = \frac{1}{8\pi\epsilon_0} \sum_{k \neq j} \frac{q_k q_j}{|\vec{r}_k - \vec{r}_j|} := V_{coul}. \quad (2.26)$$

We write the total energy  $H$  of the system as

$$H = \sum_{j=1}^N \frac{1}{2} m_j \left( \frac{d\vec{r}_j}{dt} \right)^2 + V_{coul} + \frac{\epsilon_0}{2} \int_{\mathbf{R}^3} \vec{E}_T^2(\vec{r}) + c^2 \vec{B}_T^2(\vec{r}) d\vec{r}, \quad (2.27)$$

separating the kinetic energy of the charged particles (first term), the electrostatic potential energy due to the Coulomb interaction among charged

particles (second term, defined in (2.26)) and the radiative energy of the field (third term).

**Remark 2.1.1 (Proof of formula (2.26))** Formula (2.26) is valid modulo a constant term (theoretically infinity) which is independent of the mutual positions of the particles. This term is called *self-energies* and is typically not included in the expression of the energy. To derive (2.26), use (2.15), and the fact that in the Coulomb gauge  $\vec{A}_L = 0$ , to write

$$\frac{\epsilon_0}{2} \int_{\mathbf{R}^3} \vec{E}_L^2(\vec{r}) d\vec{r} = \frac{\epsilon_0}{2} \int_{\mathbf{R}^3} (\nabla\phi)^2 d\vec{r}, \quad (2.28)$$

since  $\nabla\phi$  is purely longitudinal from (2.14). Recall now the integration by parts formula given by Green's first identity. For every volume  $V$  and smooth function  $\phi$ ,

$$\int_V (\nabla\phi)^2 d\vec{r} = \int_{\partial V} \phi \vec{n} \cdot \nabla\phi dS - \int_V \phi \nabla^2\phi d\vec{r},$$

where the first integral on the right hand side is a surface integral over  $\partial V$ , the boundary of  $V$ . Using this in (2.28) we can write

$$\frac{\epsilon_0}{2} \int_{\mathbf{R}^3} (\nabla\phi)^2 d\vec{r} = \frac{\epsilon_0}{2} \left( \lim_{V \rightarrow \infty} \int_{\partial V} \phi \vec{n} \cdot \nabla\phi dS - \int_V \phi \nabla^2\phi d\vec{r} \right). \quad (2.29)$$

We assume that the potential  $\phi$  is chosen to be zero at infinity so that the limit of the first integral of the right hand side is zero. Using this and equation (2.21), we obtain

$$\frac{\epsilon_0}{2} \int_{\mathbf{R}^3} \vec{E}_L^2(\vec{r}) d\vec{r} = \frac{1}{2} \int_{\mathbf{R}^3} \phi \rho d\vec{r}. \quad (2.30)$$

Poisson equation (2.21), assuming a discrete distribution of charges (2.8), has solution vanishing at infinity

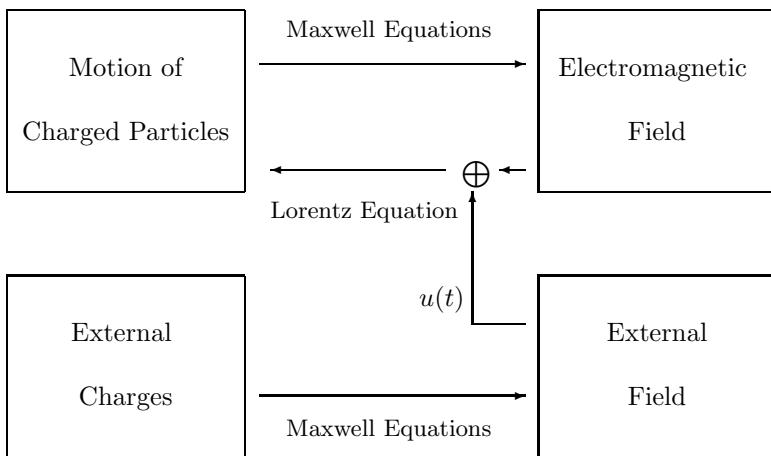
$$\phi(\vec{r}) = \frac{1}{4\pi\epsilon_0} \sum_{j=1}^N \frac{q_j}{|\vec{r} - \vec{r}_j|}. \quad (2.31)$$

By plugging expressions (2.8) and (2.31) into (2.30) some terms become infinity. These terms are called *self-energies*. They are not considered as they do not depend on the mutual positions of the various charges. Neglecting these terms, we obtain (2.26).

### 2.1.1.5 System in the presence of an externally applied electromagnetic field

The most interesting situation for our purposes is when the system of particles and field is interacting with an *external system of particles and field*. In the external system, the motion of the charges, i.e., the functions  $\rho(\vec{r}, t)$

and  $\vec{J}(\vec{r}, t)$ , are fixed and independent of the field. This means in practice that the back reaction of the field on the particles is very weak and/or it is compensated by other forces in the experiment. These functions determine the external field according to Maxwell's equations. Because of the linearity of the Maxwell's equations, the total field will be the sum of the proper field of the particles and the external field. Also, from the linearity of (2.16) and (2.17) the vector and scalar potentials are the sum of the ones proper of the system and the ones of the external system. The situation is summarized in the following diagram which extends the one given in (2.1.1.1). The external field plays the role of the control  $u(t)$ .



In the presence of an external field the energy of the system of particles and field (excluding the external system) is not constant with time and it has to be modified (cf. [50] Complement  $C_{II}$  and pg. 174) by adding a potential energy term which, in the case of a discrete system of particles, takes the form

$$V_e := \sum_{j=1}^N q_j \phi_e(\vec{r}_j, t), \quad (2.32)$$

where  $\phi_e(\vec{r}_j, t)$  is the scalar potential associated with the external system. In conclusion the energy associated with a system subject to an external field is given by (cf. (2.27))

$$H = \sum_{j=1}^N \frac{1}{2} m_j \left( \frac{d\vec{r}_j}{dt} \right)^2 + V_{coul} + \sum_{j=1}^N q_j \phi_e(\vec{r}_j, t) + \frac{\epsilon_0}{2} \int_{\mathbf{R}^3} \vec{E}_T^2(\vec{r}) + c^2 \vec{B}^2(\vec{r}) d\vec{r}, \quad (2.33)$$

with  $\phi_e$  the external potential.

**Example 2.1.2** Let us reconsider the system of a particle in a potential well of Example 1.3.1 and its possible realization as a system of a charged particle in an external field. In particular, assume the particle to be an electron confined between two layers of negative charges at  $x = 0$  and  $x = L$ , for example in a semiconductor. The charges of the two layers represent the *external charges* which set up the external field. In particular, in a one dimensional problem, the  $x$  component of the electric external field  $E_x := E_x(x)$  has large magnitude and it is negative close to  $x = 0$  and it has large magnitude and it is positive close to  $x = L$ , while it is about zero for  $0 < x < L$ . Consider equation (2.22), since there are no external currents, i.e.,  $\vec{J} = 0$ , we can take the external vector potential  $\vec{A}_e$  equal to zero.<sup>9</sup> The external scalar potential  $\phi_e$  satisfies (cf. (2.15))

$$-E_x(x) = \frac{d\phi_e}{dx}, \quad (2.34)$$

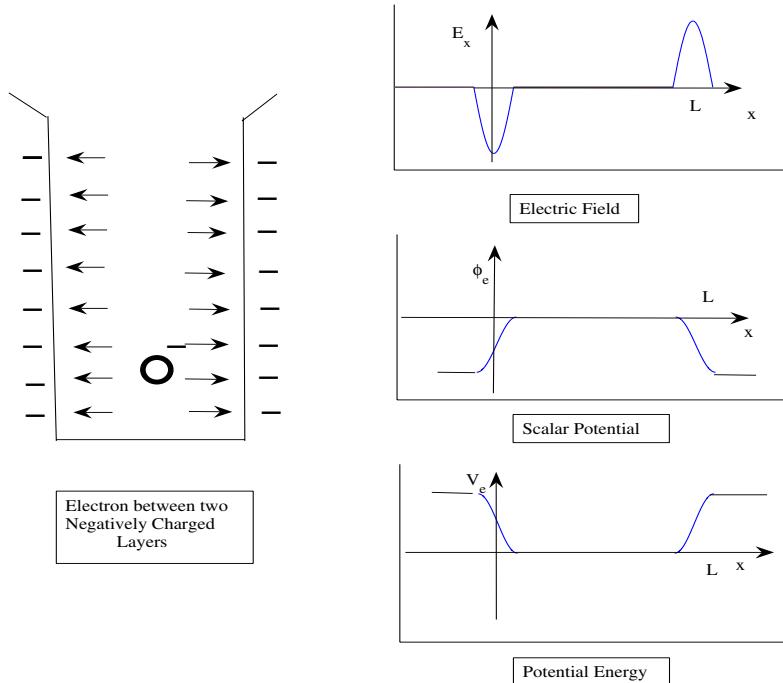
and integrating this we obtain a shape of  $\psi_e$  which is constant (and negative) equal to some  $-\phi_{e0}$  until  $x = 0$  increases until a value zero and then decreases at  $x = L$  again to  $-\phi_{e0}$ . The potential energy for a particle at position  $x$  is obtained according to (2.32) as a function of the position of the particle  $x$  as  $V_e = e\phi_e(x)$  where  $e$  is the charge of the electron. It has the shape of a quantum well as considered in example 1.3.1 in the case  $V_0 := -e\phi_{e0}$  very large. In fact the general expression of the energy (2.33) gives (1.44) in this case with  $V = V_e$ . The discussion is summarized in [Figure 2.3](#). Since there is only one charge in the well  $V_{coul} = 0$ . If there is more than one charge the term  $V_{coul}$  is added to the energy. This scheme can be modified by adding a further external *control* electric field possibly time varying which modifies the shape of the well, for example by placing the system in a capacitor. In that case the model extends naturally there will be an extra (control) potential energy obtained integrating (2.34) with the control field playing the  $E_x$ . This is a common scheme to control electron motion in semiconductors quantum dots (see, e.g., [200]). Some analysis on the effect of an external field on the motion of a single electron in a quantum well is given in [154].

### 2.1.2 Canonical quantization

**Canonical quantization** is the process which associates quantities in classical mechanics with operators in quantum mechanics. This process allows us to translate the classical description of a system to the quantum mechanical one and in particular to obtain the quantum mechanical Hamiltonian from the classical Hamiltonian. For a general system one considers the canonical variables of Lagrangian mechanics and the associated canonical momenta (cf. [Appendix B](#) and [50]). For a canonical variable  $x$  and associated canonical

---

<sup>9</sup>We can assume Coulomb gauge here so that  $\vec{A}_e$  is purely transversal while  $\frac{\partial}{\partial t}\nabla\phi_e$  is purely longitudinal from (2.14).



**FIGURE 2.3:** Electrostatic realization of the motion of a single charged particle in a potential well. An electron is confined between two negatively charged layers. A possible shape of the electric field  $E_x$ , the scalar potential  $\phi_e$  and the potential energy  $V_e$  is indicated.

momentum  $p$ , one associates operators  $\hat{x}$  and  $\hat{p}$  and the underlying Hilbert space can be chosen as spanned by the eigenvectors of  $\hat{x}$  (or  $\hat{p}$ ). The way  $\hat{x}$  and  $\hat{p}$  are defined has to be such that<sup>10</sup>

$$[\hat{x}, \hat{p}] = i\hbar\mathbf{1}. \quad (2.35)$$

The Hamiltonian operator, which determines the dynamics of the system according to the Schrodinger equation, is obtained by transforming the expression of the Hamiltonian of classical mechanics (which is written only in terms of canonical variables and associated momenta) by replacing all the canonical variables and momenta by the corresponding operator. In this process *sums* are replaced by *sums of operators* and *products* are replaced by *compositions*

<sup>10</sup>For a discussion and motivations for this requirement and relation with classical mechanics see, e.g., Chapter 15 of [144].

of operators (cf. 1.1.2.3). Care is needed due to the fact that even though the classical variable  $x$  commutes with its associated momentum  $p$ , the corresponding operators  $\hat{x}$  and  $\hat{p}$  do not commute. A symmetrized form for  $xp$  can be used, that is  $\frac{1}{2}(\hat{x}\hat{p} + \hat{p}\hat{x})$ . Generalizing what is said for sums and products, every analytic function of the canonical variables and momentum is replaced by the corresponding formula in terms of operators.

In the case of the system of charged particles and fields one should first identify the canonical variables and momenta and express the Hamiltonian (2.33) in terms of these. The system consists of two parts, the particles and the field and there will be canonical variables and momenta associated with both particles and field. The corresponding operators will act on the Hilbert space corresponding to the particles motion,  $\mathcal{H}_P$  or on the Hilbert space associated with the field  $\mathcal{H}_S$ , respectively. The total Hilbert space is  $\mathcal{H}_P \otimes \mathcal{H}_F$  (cf. subsection 1.1.3), while  $\mathcal{H}_P$  itself is the tensor product of Hilbert spaces each associated to one particle. The variables corresponding to the field in the Hamiltonian, i.e.,  $\vec{E}_T$ ,  $\vec{B}$ ,  $\vec{A}$  are replaced by the corresponding operators in terms of the operators associated to canonical variables and momentum of the field. Only the external field is treated as a classical variable and the external vector potential  $\vec{A}_e$  and scalar potential  $\phi_e$  are treated as given functions.

We shall not treat the quantization of the electromagnetic field in this book as this is the object of study of *quantum optics*. We refer to [50] for a complete treatment. However, since the charged particles are the object of the control we shall focus on the particles part of the Hilbert space  $\mathcal{H}_P$ . We start by identifying the canonical variables and momenta for the particles.

### 2.1.2.1 Canonical variables and momentum for the charged particles

The canonical variables for a system of  $N$  charged particles in an electromagnetic field are given by  $\vec{r}_j$ ,  $j = 1, \dots, N$ . The **canonical momentum** associated with  $\vec{r}_j$ , which we denote by  $\vec{p}_j$ , for particle  $j$  is defined as the sum of the *kinematical momentum*  $m_j \dot{\vec{r}}_j$  and a term proportional to the vector potential at the location of the particle, given by  $q_j \vec{A}_{tot}(\vec{r}_j)$ .<sup>11</sup> One has

$$\vec{p}_j := m_j \dot{\vec{r}}_j + q_j \vec{A}_{tot}(\vec{r}_j). \quad (2.36)$$

Here  $\vec{A}_{tot}$  is the total vector potential sum of the vector potential  $\vec{A}$  of the system and the vector potential of the applied external field  $\vec{A}_e$ , i.e.,

$$\vec{A}_{tot}(\vec{r}, t) := \vec{A}(\vec{r}, t) + \vec{A}_e(\vec{r}, t).$$

---

<sup>11</sup>In a description in terms of Lagrangian mechanics definition (2.36) follows from taking the derivative of the Lagrangian associated with the system with respect to  $\dot{\vec{r}}_j$ . This derivation is presented in Appendix B (cf. formula (B.25)) along with an overview of Lagrangian and Hamiltonian mechanics as applied to the system of interacting charged particles and fields. A complete treatment can be found in [86].

With this definition, the energy (2.33) reads

$$H = \sum_{j=1}^N \frac{1}{2m_j} (\vec{p}_j - q_j \vec{A}_{tot}(\vec{r}_j))^2 + V_{coul} + \sum_{j=1}^N q_j \phi_e(\vec{r}_j, t) \\ + \frac{\epsilon_0}{2} \int_{\mathbf{R}^3} \vec{E}_T^2(\vec{r}) + c^2 \vec{B}^2(\vec{r}) d\vec{r}. \quad (2.37)$$

### 2.1.2.2 Rules of canonical quantization for variables concerning the charged particles

We work in the position representation and therefore we describe the canonical quantization specifying how the quantum mechanical operators act on the wave function. In particular the full system of particles and field lives in the Hilbert space  $\mathcal{H}_P \otimes \mathcal{H}_F$  and assume that field space has countable orthonormal basis  $\{|k\rangle\}$ ,  $k = 1, 2, \dots$ . The state of the full system is written in the position (and  $\{|k\rangle\}$ ) representation as

$$|\psi_{full}\rangle := \sum_k \int_{\mathbf{R}^{3N}} \psi_{full}(\vec{r}, k) |\vec{r}\rangle \otimes |k\rangle d\vec{r}.$$

The state associated to the particles is obtained by calculating the partial trace with respect to the field degrees of freedom (cf. 8.1.1.3), which gives

$$|\psi\rangle = \int_{\mathbf{R}^{3N}} \psi(\vec{r}) |\vec{r}\rangle d\vec{r},$$

where the *wave function*  $\psi(\vec{r})$  is defined by

$$\psi(\vec{r}) := \sum_k \psi_{full}(\vec{r}, k).$$

The *rules of canonical quantization* associate with the canonical momentum  $\vec{p}_j$  of the  $j$ -th particle in (2.36) the operator  $\hat{p}_j$ , which is a triple of operators  $\hat{p}_j := \{\hat{p}_{jx}, \hat{p}_{jy}, \hat{p}_{jz}\}$  each corresponding to one component of  $\vec{p}_j$ .  $\hat{p}_j$  acts on the state of the particles  $|\psi(t)\rangle$  in (2.1) according to

$$\hat{p}_{jl} \psi(\vec{r}, t) := -i\hbar \frac{\partial}{\partial l_j} \psi(\vec{r}, t), \quad l_j = x_j, y_j, z_j, \quad j = 1, \dots, N,$$

or, in compact vector notation,

$$\hat{p}_j \psi(\vec{r}, t) = -i\hbar \nabla_j \psi(\vec{r}, t), \quad j = 1, \dots, N,$$

where  $\nabla_j$  denotes the gradient with respect to coordinates of the  $j$ -th particle.

The rules of canonical quantization associate with the position

$$\vec{r}_j := \{x_j, y_j, z_j\}$$

of the  $j$ -th particle the operator  $\hat{r}_j := \{\hat{x}_j, \hat{y}_j, \hat{z}_j\}$ .  $\hat{r}_j$  acts on the state of the particles  $|\psi(t)\rangle$  as a multiplication by the corresponding component, i.e.,

$$\hat{m}_j \psi(\vec{r}, t) = m_j \psi(\vec{r}, t), \quad m_j = x_j, y_j, z_j, \quad j = 1, \dots, N,$$

or, in compact vector notation,

$$\hat{r}_j \psi(\vec{r}, t) = \vec{r}_j \psi(\vec{r}), \quad j = 1, \dots, N.$$

From these definitions it follows that the following commutation relation (cf. (2.35)) holds (cf. Exercise 2.2)

$$[\hat{p}_{jl}, \hat{m}_k] = -i\hbar \delta_{lm} \delta_{jk}, \quad l, m = x, y, z, \quad j, k = 1, \dots, N. \quad (2.38)$$

**Example 2.1.3** Going back to Example 2.1.2 and applying these quantization rules with the Hamiltonian (2.33) specialized to this case, we obtain the Hamiltonian of the form (1.45) of example 1.3.1. This Hamiltonian can be modified by adding an extra (time varying) control electric field which results in an extra term for the scalar potential as already discussed in Example 2.1.2.

### 2.1.3 An example of canonical quantization: The quantum harmonic oscillator

The quantum harmonic oscillator is one of the most important and most studied quantum systems as it is the model for several phenomena in molecular dynamics, solid state physics, quantum optics etc. It is obtained from a procedure of canonical quantization starting from a classical harmonic oscillator. For simplicity, we consider only the one dimensional harmonic oscillator whose only degree of freedom is the  $x$  coordinate. Extensions to higher dimensional case are natural. For this system the conjugate canonical momentum  $p$ , calculated according to Lagrangian mechanics as in [Appendix B](#) (B.10), is found to be coinciding with the kinematical momentum, i.e.,<sup>12</sup>

$$p = m\dot{x}. \quad (2.39)$$

The total energy is given by the Hamiltonian

$$H_{ho} = \frac{p^2}{2m} + \frac{1}{2}m\omega^2x^2, \quad (2.40)$$

---

<sup>12</sup>The associated Lagrangian is given by

$$L_{ho} = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}m\omega^2x^2,$$

where (cf. (B.4))  $T := \frac{1}{2}m\dot{x}^2$  is the kinetic energy while  $V := \frac{1}{2}m\omega^2x^2$  represents the potential energy of the system and  $\omega$  is a constant having units of a frequency. Using this Lagrangian, the Euler-Lagrange equations (B.2) give the equation of motion

$$\ddot{x} - \omega^2x = 0.$$

The Hamiltonian (2.40) is found from the definition (B.12) using (2.39).

which is the sum of the kinetic energy (first term) and the potential energy of the system (second term).

To obtain the quantum mechanical description of the harmonic oscillator, one defines the position,  $\hat{x}$ , and momentum,  $\hat{p}$ , operators, associated with  $x$  and  $p$ , respectively, according to the rules of canonical quantization. In particular<sup>13</sup>  $\hat{x}$  corresponds to multiplication of the wave function  $\psi(x)$  by  $x$  while  $\hat{p}$  transforms  $\psi(x)$  according to  $\psi(x) \rightarrow -i\hbar \frac{d}{dx} \psi(x)$ . They satisfy the commutation relation

$$[\hat{x}, \hat{p}] = i\hbar,$$

which is a special case of (2.38). From (2.40), the Hamiltonian operator is given by

$$\hat{H}_{ho} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2.$$

We can express the state in the *energy representation*, i.e., with an expansion in terms of the energy eigenstates.<sup>14</sup> The eigenvalues of the energy operators can be obtained by solving the time independent Schrödinger wave equation (cf. (1.46)) which in this case is

$$\hat{H}\psi(x) = -\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + \frac{1}{2}m\omega^2x^2\psi(x) = E\psi(x).$$

This equation can be solved explicitly and gives the values of the energy eigenvalues and the corresponding wave functions. In particular the requirement that  $\psi(x)$  be a probability density leads to only a discrete set of possible values for the energy which are equally spaced by a quantity  $\hbar\omega$ . We shall illustrate this result using Dirac's method which works in terms of state kets rather than wave functions.

Define the **destruction operator**,  $\hat{a}$ , as

$$\hat{a} := (2m\hbar\omega)^{-\frac{1}{2}}(m\omega\hat{x} + i\hat{p}), \quad (2.41)$$

and the **creation operator**,  $\hat{a}^\dagger$ , as

$$\hat{a}^\dagger := (2m\hbar\omega)^{-\frac{1}{2}}(m\omega\hat{x} - i\hat{p}). \quad (2.42)$$

They satisfy the commutation relation

$$[\hat{a}, \hat{a}^\dagger] = \mathbf{1}, \quad (2.43)$$

where  $\mathbf{1}$  is the identity operator (cf. Exercise 2.3).

---

<sup>13</sup>Recall we are working in the position representation, i.e., we write the state as  $|\psi\rangle = \int_{-\infty}^{\infty} \psi(x)|x\rangle dx$ .

<sup>14</sup>For the relation between the two representations as well as for a more extensive study of the quantum harmonic oscillator see, e.g., [185] Sect. 2.3.

The Hamiltonian  $\hat{H}_{ho}$  can be rewritten in terms of these operators as

$$\hat{H}_{ho} = \frac{1}{2}\hbar\omega(\hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a}) = \hbar\omega(\hat{a}^\dagger\hat{a} + \frac{1}{2}). \quad (2.44)$$

Consider now the eigenstates of  $\hat{H}_{ho}$  and denote by  $|n\rangle$  a generic eigenstate and  $E_n$  the corresponding value of the energy. This notation is chosen because we shall show below that we can put energy values and eigenvectors in one-to-one correspondence with nonnegative integer values. The energy eigenvalue equation (which corresponds to the Schrödinger time independent Schrödinger equation) reads

$$\hbar\omega(\hat{a}^\dagger\hat{a} + \frac{1}{2})|n\rangle = E_n|n\rangle.$$

Applying  $\hat{a}^\dagger$  to both terms and replacing  $\hat{a}^\dagger\hat{a}$  with  $\hat{a}\hat{a}^\dagger - \mathbf{1}$  according to (2.43), we obtain

$$\hbar\omega(\hat{a}^\dagger\hat{a}\hat{a}^\dagger - \hat{a}^\dagger + \frac{1}{2}\hat{a}^\dagger)|n\rangle = E_n|n\rangle.$$

Taking the term  $\hbar\omega\hat{a}^\dagger|n\rangle$  to the right hand side and recalling the expression of the Hamiltonian  $\hat{H}_{ho}$ , we obtain

$$\hat{H}_{ho}\hat{a}^\dagger|n\rangle = (E_n + \hbar\omega)\hat{a}^\dagger|n\rangle.$$

This shows that  $|n+1\rangle := \hat{a}^\dagger|n\rangle$  is also an eigenvector of  $\hat{H}_{ho}$  with eigenvalue  $E_n + \hbar\omega$ . Analogously one can see that  $|n-1\rangle := \hat{a}|n\rangle$  is an eigenvector of  $\hat{H}_{ho}$  with eigenvalue  $E_n - \hbar\omega$ . The elementary energy quantity  $\hbar\omega$ , which represents the gap between two successive energy levels, is called a *phonon* in the case where the harmonic oscillator represents a mechanical system.<sup>15</sup>

Let us now study more in depth the nature of the set of eigenvalues and eigenvectors. Let us assume we start from a certain eigenvalue  $E$  and eigenvector  $|e\rangle$ . Applying  $\hat{a}$   $n$  times we obtain the eigenvector  $(\hat{a})^n|e\rangle$ , with eigenvalue  $\bar{E} - n\hbar\omega$ . This cannot be negative and denote by  $\bar{n}$  the largest value  $n$  so that  $\bar{E} - n\hbar\omega$  is nonnegative. Call  $|0\rangle := (\hat{a})^{\bar{n}}|e\rangle$ . Since  $\hat{H}_{ho}\hat{a}|0\rangle = (E - (\bar{n} + 1)\hbar\omega)\hat{a}|0\rangle$  and  $\bar{E} - (\bar{n} + 1)\hbar\omega < 0$ ,  $\hat{a}|0\rangle$  cannot be an eigenvalue of  $\hat{H}_{ho}$ , therefore we must have

$$\hat{a}|0\rangle = 0,$$

where 0 represents the zero vector of the associated Hilbert space. Call  $E_0$  the eigenvalue corresponding to  $|0\rangle$ . Using the expression of the energy operator  $\hat{H}_{ho}$  (2.44) and  $\hat{a}|0\rangle = 0$  we obtain

$$\hbar\omega(\hat{a}\hat{a} + \frac{1}{2})|0\rangle = \frac{\hbar\omega}{2}|0\rangle = E_0|0\rangle,$$

---

<sup>15</sup>Harmonic oscillators also model an electromagnetic field as a quantum system (see, e.g., [50], [95], [138]). In that context  $\hbar\omega$  is called a *photon*.

which shows that  $E_0 = \frac{\hbar\omega}{2}$ . We have then shown that  $E_n := E_0 + n\hbar\omega$ ,  $n = 0, 1, 2, \dots$  are eigenvalues of  $\hat{H}_{ho}$  and called  $|n\rangle$  the corresponding eigenvectors. We want to show now that these are the *only* possible eigenvalues. If there was any value different from these there would be associated a sequence of energy values. Call  $|0'\rangle$  the one corresponding to the lowest possible value of the energy  $E'_0$ . As above we have  $\hat{a}|0'\rangle$  and therefore  $E'_0 = \frac{1}{2}\hbar\omega = E_0$ . Therefore the sequence of eigenvalues coincides with  $E_n := E_0 + n\hbar\omega$ ,  $n = 0, 1, 2, \dots$  and the minimum value is  $E_0$ . It is an important result that the minimum value of the energy is not zero. The state  $|0\rangle$  is called the *vacuum state* and it corresponds to no phonon. The creation (destruction) operator<sup>16</sup> transforms an eigenstate of the Hamiltonian into an eigenstate corresponding to energy higher (lower) by a quantity  $\hbar\omega$ . It ‘creates’ (‘destructs’) a *quantum* of energy  $\hbar\omega$ , that is, a phonon.

#### 2.1.4 Quantum mechanical Hamiltonian

Let us consider again the full classical Hamiltonian for a system for interacting particles and fields in the presence of an external driving field (2.33) with the definition of the canonical momentum (2.36). We have

$$\begin{aligned} H = & \sum_{j=1}^N \frac{1}{2m_j} (\vec{p}_j - q_j A_{tot}(\vec{r}_j, t))^2 + V_{coul} + \sum_{j=1}^N q_j \phi_e(\vec{r}_j, t) \\ & + \frac{\epsilon_0}{2} \int_{\mathbf{R}^3} \vec{E}_T^2(\vec{r}) + c^2 \vec{B}^2(\vec{r}) d\vec{r}. \end{aligned} \quad (2.45)$$

In anticipation to applying canonical quantization we identify the variables concerning the particles,  $\vec{p}_j$  and  $\vec{r}_j$ , and the ones concerning the field  $\vec{E}$ ,  $\vec{A}$ ,  $\vec{B}$ . We write the energy Hamiltonian (2.45) as

$$H := H_R + H_P + H_I,$$

with

$$\begin{aligned} H_R &:= \frac{\epsilon_0}{2} \int_{\mathbf{R}^3} \vec{E}_T^2(\vec{r}) + c^2 \vec{B}^2(\vec{r}) d\vec{r}, \\ H_P &:= \sum_{j=1}^N \frac{1}{2m_j} (\vec{p}_j)^2 + V_{coul}, \\ H_I &:= - \sum_{j=1}^N \frac{q_j}{m_j} \vec{p}_j \cdot \vec{A}_{tot}(\vec{r}_j, t) + \sum_{j=1}^N \frac{q_j^2}{2m_j} \vec{A}_{tot}^2(\vec{r}_j, t) + \sum_{j=1}^N q_j \phi_e(\vec{r}_j, t). \end{aligned} \quad (2.46)$$

$H_R$  is the energy of the field only. It is transformed into an Hamiltonian operator which acts only on the part of the Hilbert space  $\mathcal{H}_P \otimes \mathcal{H}_F$  concerning

---

<sup>16</sup>These operators are also called the *raising* and *lowering* operators, respectively.

the field, i.e., on  $\mathcal{H}_F$ , i.e., it has the form  $\hat{H}_R := \mathbf{1} \otimes \hat{H}_{RF}$  for an operator  $\hat{H}_{RF}$  on  $\mathcal{H}_F$ . The term  $H_P$  transforms into a quantum mechanical operator  $\hat{H}_P$  according to the rules of canonical quantization.  $\hat{H}_P$  is given by

$$\hat{H}_P = \sum_{j=1}^N -\frac{1}{2m_j} \hbar^2 \nabla_j^2 + \hat{V}_{coul},$$

where from (2.26) the operator  $\hat{V}_{coul}$  multiplies the wave function by

$$V_{coul} := \frac{1}{8\pi\epsilon_0} \sum_{k \neq j} \frac{q_k q_j}{|\vec{r}_k - \vec{r}_j|}.$$

All the terms appearing in  $H_I$  in (2.46) contain both quantities concerning the particles and the field, both the external field and the field created by the particles. This corresponds to the interaction part of the quantum mechanical Hamiltonian which contains operators acting simultaneously on the particles Hilbert space  $\mathcal{H}_P$  and the field Hilbert space  $\mathcal{H}_F$ . These are possibly time varying operators as a consequence of the presence of the external field. Since  $\vec{A}_{tot} = \vec{A} + \vec{A}_e$ , we write  $H_I$  as

$$H_I = H_{IED} + H_{Ie},$$

where

$$H_{IED} := \sum_{j=1}^N \frac{q_j^2}{2m_j} \left( \vec{A}^2(\vec{r}_j, t) + 2\vec{A} \cdot \vec{A}_e(\vec{r}_j, t) \right) - \sum_{j=1}^N \frac{q_j}{m_j} \vec{p}_j \cdot \vec{A}(\vec{r}_j), \quad (2.47)$$

and

$$H_{Ie} : \sum_{j=1}^N \frac{q_j^2}{2m_j} \vec{A}_e^2(\vec{r}_j, t) - \sum_{j=1}^N \frac{q_j}{m_j} \vec{p}_j \cdot \vec{A}_e(\vec{r}_j, t) + \sum_{j=1}^N q_j \phi_e(\vec{r}_j, t). \quad (2.48)$$

The term  $H_{IED}$  contains the *interaction* between particles and field possibly time varying because of the presence of the external field. This term is responsible for *electrodynamic interaction* among the particles, i.e., interaction which happens through the vector potential  $\vec{A}$  as opposed to the *electrostatic interaction* which is the coulomb interaction  $V_{coul}$ . The term  $H_{Ie}$  models the interaction with the external field only. The quantum mechanical operator associated to this term is

$$\hat{H}_{Ie} = \sum_{j=1}^N \frac{q_j^2}{2m_j} \hat{A}_e^2(\vec{r}_j, t) + \sum_{j=1}^N \frac{q_j}{m_j} i\hbar \nabla_j \hat{A}_e(\vec{r}_j, t) + \sum_{j=1}^N q_j \hat{\phi}_e(\vec{r}_j, t).$$

According to the rules of canonical quantization  $\hat{\phi}_e$  is obtained from the function  $\phi_e = \phi_e(\vec{r}_j, t)$ , by replacing every instance of  $\vec{r}_j := [x_j, y_j, z_j]$  by the

corresponding operator (multiplication of the wave function  $\psi(\vec{r})$  by  $x_j$ ,  $y_j$  or  $z_j$ ). Therefore  $\hat{\phi}_e(\vec{r}_j, t)\psi = \phi_e(\vec{r}_j, t)\psi$ , where the second expression is multiplication of two functions. The operator  $\hat{A}_e = \hat{A}_e(\vec{r}_j, t)$  is obtained in the same way. The only difference is that it is a *vector* operator and it has to be seen as a triple of operators. We have  $\hat{A}_e(\vec{r}_j, t)\psi = \vec{A}_e(\vec{r}_j, t)\psi$ , where the second term represents multiplication of a function by a scalar. When applied to  $\hat{A}_e(\vec{r}_j, t)\psi$ ,  $\nabla_j$  returns a  $\nabla_j \cdot \vec{A}_e(\vec{r}_j, t)\psi$ , which is a scalar function again. Notice that there is ambiguity in this quantization process. Since we are quantizing the product  $\vec{p}_j \cdot \vec{A}_e(\vec{r}_j, t)$  we might as well have used the operator  $i\hbar \vec{A}_e(\vec{r}_j, t) \cdot \nabla_j$  or the symmetrized form  $\frac{1}{2}i\hbar(\vec{A}_e(\vec{r}_j, t) \cdot \nabla_j + \nabla_j \cdot \vec{A}_e(\vec{r}_j, t))$ . However as we assume  $\vec{A}_e$  is chosen in the Coulomb gauge all these choices are equivalent. In fact, from the product rule, we have

$$\nabla_j \cdot (\vec{A}_e(\vec{r}_j)\psi) = (\nabla_j \cdot \vec{A}_e(r_j))\psi + \vec{A}_e(\vec{r}_j) \cdot \nabla_j \psi,$$

and using the fact that, in the Coulomb gauge,  $\nabla_j \cdot \vec{A}_e(\vec{r}_j) = 0$ , we have

$$\nabla_j \cdot (\vec{A}_e(\vec{r}_j)\psi) = \vec{A}_e(\vec{r}_j) \cdot \nabla_j \psi.$$

We illustrate the mathematical definitions with a simple example without making reference to a particular physical situation.

**Example 2.1.4** Consider a system of two particles in 3-D space. The coordinates are  $\vec{r} := \{x_1, y_1, z_1, x_2, y_2, z_2\}$ . Let the wave function be

$$\psi = \psi(\vec{r}) = \psi(x_1, y_1, z_1, x_2, y_2, z_2) = \frac{1}{\pi} e^{-(x_1^2 + x_2^2)}.$$

Consider an external scalar and vector potential  $\phi_e$  and  $\vec{A}_e$ , independent of time and given by

$$\phi_e = 2 - x^2,$$

and

$$\vec{A}_e = y\vec{i} + x\vec{j},$$

which is in the Coulomb gauge since  $\nabla \cdot \vec{A}_e = 0$ . We calculate

$$\hat{\phi}_e(\vec{r}_1)\psi = \frac{2 - x_1^2}{\pi} e^{-(x_1^2 + x_2^2)},$$

$$\hat{A}_e(\vec{r}_1)\psi = \frac{1}{\pi} e^{-(x_1^2 + x_2^2)}(y_1\vec{i} + x_1\vec{j}),$$

and

$$\nabla_1 \cdot \hat{A}_e(\vec{r}_1)\psi = \frac{-2x_1y_1}{\pi} e^{-(x_1^2 + x_2^2)}.$$

## 2.2 Approximations and Modeling; Molecular Systems

In modeling real systems of charged particles and fields the Hamiltonian (2.45) is almost never considered in full and appropriate approximations are introduced depending on the particular situation at hand. Consider the *control of molecular or atomic dynamics* where the objectives to be pursued include maximizing the amount of a desired product in a chemical reaction, expediting a given reaction, or dissociating a molecule. The system of interest consists of the electrons and nuclei of the (reacting) atoms. These are treated as charged particles in an electromagnetic field, as in the previous section. The underlying Hilbert space for the particles is spanned by the position eigenvectors of the electrons, and therefore it is uncountably infinite dimensional.

### 2.2.1 Approximations for molecular and atomic systems

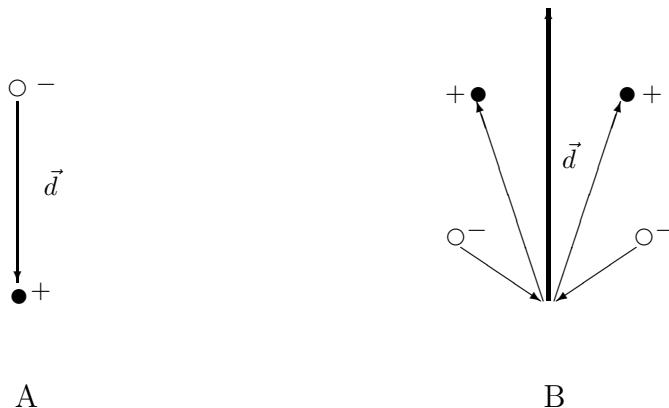
#### 2.2.1.1 Neglecting electrodynamic interactions

In the control of molecular or atomic systems it can be assumed that the particles (electrons and nuclei) are sufficiently near so that the Coulomb interaction alone is the dominant term which takes into account the interaction among particles. The electrodynamic interaction terms corresponding to (2.47) are therefore considered small and neglected. This also allows to decouple the part of the system referring to the own system field, corresponding to the part of the Hilbert space  $\mathcal{H}_F$ . Therefore we can consider only the part of the system concerning the particles and the Hamiltonian contains no field other than the external fields. The classical Hamiltonian after this approximation is written as

$$\begin{aligned} H_P + H_{Ie} &= \sum_{j=1}^N (\vec{p}_j)^2 + V_{coul} - \sum_{j=1}^N \frac{q_j}{m_j} \vec{p}_j \cdot \vec{A}_e(\vec{r}_j, t) \\ &\quad + \sum_{j=1}^N \frac{q_j^2}{2m_j} \vec{A}_e^2(\vec{r}_j, t) + \sum_{j=1}^N q_j \phi_e(\vec{r}_j, t). \end{aligned} \quad (2.49)$$

The corresponding Hamiltonian operator after this approximation is written as

$$\begin{aligned} \hat{H}_P + \hat{H}_{Ie} &= \sum_{j=1}^N -\frac{1}{2m_j} \hbar^2 \nabla_j^2 + \hat{V}_{coul} + \sum_{j=1}^N \frac{q_j}{m_j} i\hbar \nabla_j \hat{A}_e(\vec{r}_j, t) \\ &\quad + \sum_{j=1}^N \frac{q_j^2}{2m_j} \hat{A}_e^2(\vec{r}_j, t) + \sum_{j=1}^N q_j \hat{\phi}_e(\vec{r}_j, t). \end{aligned}$$



**FIGURE 2.4:** The electric dipole moment  $\vec{d}$  of a system of one positive and one negative charge (Part A) such as an hydrogen atom and the electric dipole moment of a system of two positive and two negative charges (Part B).

### 2.2.1.2 The long wavelength approximation and the electric dipole interaction

The last two terms in the Hamiltonian (2.49) model the interaction between the transversal field and the motion of the particles. In fact, they depend on both  $\vec{A}_e$ ,  $\phi_e$  and the positions of the  $j$ -th particle  $\vec{r}_j$ . It is common to rewrite them in terms of the dot product of the electric field  $\vec{E}$  and the dipole moment

$$\vec{d} = \sum_{j=1}^N q_j \vec{r}_j, \quad (2.50)$$

where, in the case of a molecule or atom, the  $q_j$ 's correspond to the charge of the electrons and protons.<sup>17</sup>

This requires a gauge transformation and an approximation. First perform a gauge transformation of the form (2.18), (2.19), with  $\eta$  given by

$$\eta = \eta(\vec{r}, t) = -\vec{r} \cdot \vec{A}_e(\vec{r}_0, t),$$

<sup>17</sup>For a single atom, the protons may be considered giving no contribution as they may be assumed concentrated in one point.

where  $\vec{r}_0$  is a fixed position that can be taken to be at the center of mass of the molecule. The new vector and scalar potentials are given by

$$\vec{A}'_e(\vec{r}, t) = \vec{A}_e(\vec{r}, t) - \vec{A}_e(\vec{r}_0, t),$$

$$\phi'_e(\vec{r}, t) = \phi_e(\vec{r}, t) + \vec{r} \cdot \frac{\partial \vec{A}_e(\vec{r}_0, t)}{\partial t}.$$

We notice that this gauge transformation keeps  $\vec{A}$  in the Coulomb gauge, as  $\nabla \cdot \vec{A}' = 0$  when  $\nabla \cdot \vec{A} = 0$ . Take  $\vec{r}_0 = \vec{0}$  and assume, without loss of generality, that  $\phi_e$  had been chosen so that  $\phi_e(\vec{0}, t) \equiv 0$  for every  $t$ .<sup>18</sup> In the expression of the classical Hamiltonian (2.49)  $\phi_e$ ,  $\vec{A}_e$  and  $\vec{p}_j$  have to be replaced by the new  $\phi'_e$ ,  $\vec{A}'_e$  and  $\vec{p}'_j$ . We now make the *long wavelength approximation* assuming that the field does not change significantly near  $\vec{0}$  and that the particles stay close to  $\vec{0}$ , i.e.,  $\vec{r}_j \approx \vec{0}$ , for all  $j = 1, \dots, N$ . Therefore we have

$$\vec{A}'_e(\vec{r}_j, t) := \vec{A}_e(\vec{r}_j, t) - \vec{A}_e(\vec{0}, t) \approx 0, \quad (2.51)$$

and we expand to first order in  $\vec{r}$

$$\phi'_e(\vec{r}, t) \approx \phi_e(\vec{0}, t) + \frac{\partial}{\partial t} \vec{A}_e(\vec{0}, t) \cdot \vec{r} + \nabla \phi_e(\vec{0}, t) \cdot \vec{r}.$$

Using (2.15) and the assumption on the value of  $\phi_e$  at the origin, we have

$$\phi'_e(\vec{r}, t) \approx -\vec{E}_e(\vec{0}, t) \cdot \vec{r},$$

where  $\vec{E}_e$  is the external electric field. Replacing this and (2.51) in the energy (2.49) (written for in the new gauge) we obtain

$$H_P + H_{Ie} = \sum_{j=1}^N (\vec{p}'_j)^2 + V_{coul} - \sum_{j=1}^N q_j \vec{E}_e(\vec{0}, t) \cdot \vec{r}_j,$$

and using the definition of the dipole moment (2.50), we have

$$H_P + H_{Ie} = \sum_{j=1}^N (\vec{p}'_j)^2 + V_{coul} - \vec{d} \cdot \vec{E}_e(\vec{0}, t). \quad (2.52)$$

We denote by  $\hat{d} := \{\hat{d}_x, \hat{d}_y, \hat{d}_z\}$  the quantum mechanical operator corresponding to  $\vec{d}$ . We have

$$\hat{d} = \sum_{j=1}^N q_j \hat{r}_j.$$

The quantum mechanical (dipole) Hamiltonian corresponding to (2.52) is

$$\hat{H}_P + \hat{H}_{Ie} = \sum_{j=1}^N -\frac{1}{2m_j} \hbar^2 \nabla_j^2 + \hat{V}_{coul} - \hat{d} \cdot \vec{E}_e(\vec{0}, t). \quad (2.53)$$

---

<sup>18</sup>This could have been obtained by a preliminary gauge transformation which leaves  $\vec{A}_e$  unchanged.

### 2.2.2 Controlled Schrödinger wave equation

It is convenient to separate the Hamiltonian of a quantum mechanical system interacting with an external field into a constant part and a time varying part depending on the external field. We denote the relevant Hamiltonian (after appropriate approximations) by  $\hat{H}$  and write it as

$$\hat{H} = \hat{H}_0 + \hat{H}_I(t). \quad (2.54)$$

The operator  $\hat{H}_0$ , which represents sums of kinetic and potential energy in absence of the external transversal field, is called the *internal Hamiltonian*. The operator  $\hat{H}_I$  represents the interaction energy between the system and the external field and it is called *interaction Hamiltonian*. For  $\hat{H} = \hat{H}_P + \hat{H}_{Ie}$  in (2.53) we have

$$\hat{H}_0 = \hat{H}_P = \sum_{j=1}^N -\frac{1}{2m_j} \hbar^2 \nabla_j^2 + \hat{V}_{coul},$$

and

$$\hat{H}_I = \hat{H}_{Ie} = -\vec{d} \cdot \vec{E}(\vec{0}, t).$$

The *controlled Schrödinger wave equation* reads

$$i\hbar \frac{\partial \psi(\vec{r}, t)}{\partial t} = (\hat{H}_0 + \hat{H}_I)\psi(\vec{r}, t). \quad (2.55)$$

In the absence of the external field, the state of the molecule will be described by a wave function  $\psi_s(\vec{r}, t)$ , solution of the Schrödinger wave equation (cf. (2.2))

$$i\hbar \frac{\partial \psi_s(\vec{r}, t)}{\partial t} = \hat{H}_0 \psi_s(\vec{r}, t). \quad (2.56)$$

A wave function  $\psi_k(\vec{r}, t)$  corresponding to a fixed value of the energy  $E_k$  has to satisfy the Schrodinger wave equation

$$\hat{H}_0 \psi_k(\vec{r}, t) = E_k \psi_k(\vec{r}, t). \quad (2.57)$$

There are various possible values for the energy of the system which might form a discrete set or a continuous set of a mixed (continuous and discrete) set. The state at the lowest value of the energy is called the *ground state* while the other states are called *excited states*.

Placing (2.57) into (2.55) and expressing  $\psi_k(\vec{r}, t)$  as  $\psi_k(\vec{r}, t) = \alpha_k(t)\phi_k(\vec{r})$ , we can solve the partial differential equation by separation of variables and obtain the solution

$$\psi_k(\vec{r}, t) = \phi_k(\vec{r}) e^{\frac{-iE_k t}{\hbar}},$$

where  $\phi_k(\vec{r})$  is solution of the *time independent Schrödinger wave equation* (cf. (2.57))

$$\hat{H}_0 \phi_k(\vec{r}) = E_k \phi_k(\vec{r}).$$

The (approximate) solution of this equation in molecular dynamics is obtained by methods of quantum chemistry (see, e.g., [199]). This way, one obtains the initial equilibrium state  $\phi_k(\vec{r})$ . Let us assume that there is a discrete set of energy values  $E_k$  along with the eigenstates  $\phi_k(\vec{r})$  associated with them. It is also customary to assume that we work in a finite range of energies. This is referred to as **multilevel approximation** or **truncation**.<sup>19</sup> The control problem in molecular dynamics can be formulated as transferring the state of a compound between two eigenstates corresponding to two different values of the energy.

Writing the general form of the wave function solution of (2.2) in terms of the eigenstates  $\phi_k$  of the Hamiltonian operator  $\hat{H}_0$ , we obtain

$$\psi(\vec{r}, t) = \sum_{k=1}^n c_k(t) \phi_k(\vec{r}), \quad (2.58)$$

with  $c_k$  complex coefficients and  $\sum_k |c_k|^2 = 1$ .  $|c_k|^2$  are the probabilities for the system to be in eigenstate  $\phi_k(\vec{r})$  and have energy  $E_k$ .<sup>20</sup>

### 2.2.2.1 The control system

By placing (2.58) into (2.55), we find that the complex, norm 1, vector  $\vec{c} = \{c_1, c_2, \dots, c_d\}$  satisfies the linear ordinary differential equation<sup>21</sup>

$$i\hbar \frac{d\vec{c}}{dt} = (\hat{H}_0 + \hat{H}_I(t))\vec{c}, \quad (2.59)$$

with  $\hat{H}_0$  and  $\hat{H}_I(t)$  Hermitian matrices.<sup>22</sup> In particular  $\hat{H}_0$  is diagonal and the values on the diagonal are the possible values for the energy. Using the dipole approximation, the matrix  $\hat{H}_I$  is the matrix representation of the operator (cf. (2.53))  $-\hat{d} \cdot \vec{\epsilon}E(\vec{0}, t)$ , where  $\vec{\epsilon}$  is the direction of the electric field.<sup>23</sup> The control problem of transferring the state between two different energy eigenstates can now be naturally formulated for system (2.59). This system

<sup>19</sup>See more on the multilevel approximation in Chapter 7, [section 7.1](#).

<sup>20</sup>This is easily seen by recalling that the probability for a state  $|\psi\rangle$  to be found with a value of an observable corresponding to an eigenvector  $|\phi_n\rangle$  is  $|\langle\phi_n|\psi\rangle|^2$ . In our case

$$|\psi\rangle = \int \psi(\vec{r}, t) |\vec{r}\rangle d\vec{r} = \sum_k^n c_k(t) \int \phi_k(\vec{r}) |\vec{r}\rangle d\vec{r} := \sum_k^n c_k(t) |\phi_k\rangle,$$

where  $|\phi_k\rangle := \int \phi_k(\vec{r}) |\vec{r}\rangle d\vec{r}$  and taking the product of both sides with  $\langle\phi_n|$  we obtain  $c_n = \langle\phi_n|\psi\rangle$  (since  $\hat{H}_0$  is Hermitian,  $|\phi_n\rangle$  are orthonormal).

<sup>21</sup>Also, cf. Exercise 2.1.

<sup>22</sup>Continuing with the notation introduced in [Chapter 1](#), we denote with the same symbol an operator and its matrix representation.

<sup>23</sup>For simplicity we assume here this direction fixed although what we say can be easily adapted for a direction  $\vec{E}$  variable with time.

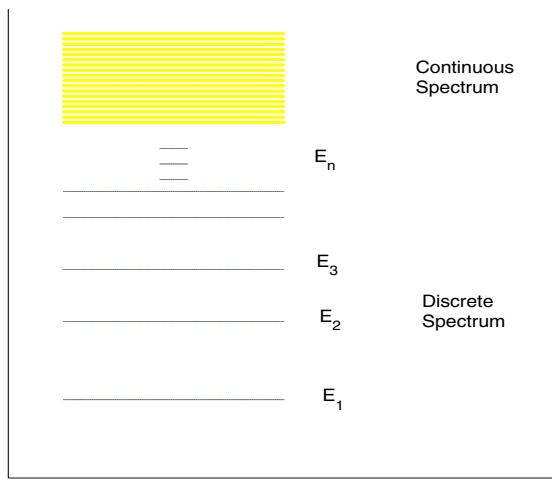
has state varying on a complex sphere of radius 1 and if we treat the electric field as the control  $u$  it has the form of a *bilinear system*

$$\frac{d\vec{c}}{dt} = (A + Bu)\vec{c},$$

with  $A$  and  $B$  Hermitian matrices.

### 2.2.2.2 Summary of the modeling procedure

We summarize the procedure for modeling a quantum control system outlined in this and the previous section. The procedure follows the following points



**FIGURE 2.5:** A typical situation in the dissociation of molecule is when the energy spectrum has the structure outlined in this figure. This is a mixed structure discrete and continuous. The ground state  $E_1$  and the excited states  $E_2, E_3, \dots, E_n, \dots$  correspond to a discrete spectrum of energy levels. There then exists a continuum of energy levels. Dissociation of the molecule happens in the passage from a state corresponding to a value of energy in the discrete spectrum to one corresponding to a value in the continuous spectrum.

1. Obtain the classical Hamiltonian (the energy) of the system under consideration identifying the time varying parameters corresponding to an external action.

2. Introduce appropriate approximations neglecting small terms (e.g., electrodynamic interactions) in the expression of the Hamiltonian, in order to obtain a simplified form.
3. Express the Hamiltonian in terms of canonical variables and canonical momenta only.
4. Perform canonical quantization to obtain the quantum mechanical Hamiltonian and the relevant Hilbert space for the quantum mechanical system.
5. Identify the internal  $\hat{H}_0$  and interaction Hamiltonian  $\hat{H}_I$  in (2.54) and calculate the energy eigenvalues and eigenstates of  $\hat{H}_0$  from the time independent Schrödinger equation (2.57).
6. Identify the relevant energy levels (multilevel approximation) and expand the wave function in terms of the corresponding eigenstates as in (2.58) (assumed finite).
7. Obtain a finite dimensional control system for the coefficient  $c_k$  in (2.58).

Notice that the modeling procedure remains valid up to Schrodinger wave equation (2.55) even though we do not assume a finite number of levels. In this case, the problem is a control problem for an infinite dimensional system. Control problems of this type arise for example in molecular dissociation as discussed in [Figure 2.5](#).

---

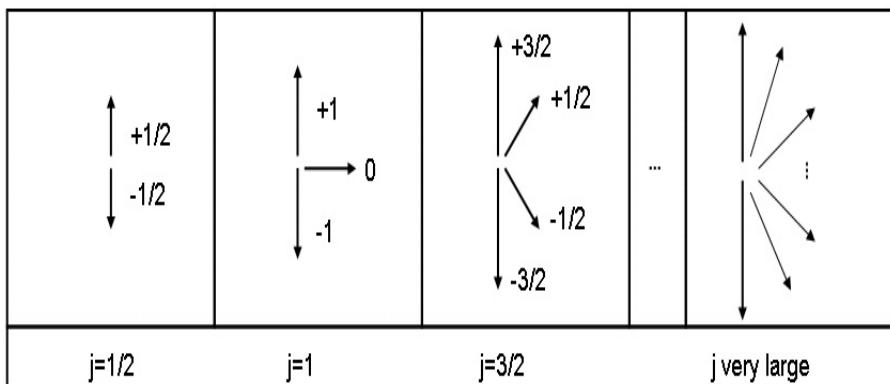
### 2.3 Spin Dynamics and Control

Several particles such as electrons, some types of nuclei<sup>24</sup> and molecules display an *intrinsic* (as opposed to *orbital*) angular momentum referred to as **spin angular momentum**. We denote the observable spin angular momentum in the  $x, y, z$ -direction by  $S_{x,y,z}$ , respectively. If  $S_{x,y,z}$  is measured, there are  $2j + 1$  possible results,  $-\hbar j, -\hbar(j-1), \dots, \hbar(j-1), \hbar j$ , where  $j$  is either a positive integer or a positive half integer whose value depends on the system under consideration (cf. [Figure 2.6](#)). It is called the **spin** of the system and it is one of its intrinsic properties.

Particles with spin (and other degrees of freedom neglected) are of great theoretical and practical interest in quantum physics. They are the prototypical example of finite dimensional quantum systems, being similar to classical

---

<sup>24</sup>In particular all isotopes which contain an odd number of nucleons and some of the ones with an even number of nucleons have nuclear spin.



**FIGURE 2.6:** Pictorial representation of particles with spin for several values of the spin  $j$ . For  $j = \frac{1}{2}$  only two values of the spin are possible when measured, as in a Stern-Gerlach experiment. For spin  $j = 1$  there are three possible values of the spin. As the intrinsic spin of the particle increases, the behavior resembles more the one of classical orbital angular momentum, which may give a continuum of possible values, when measured.

systems for large values of  $j$  (many possible values of the angular momentum) and displaying distinctive quantum mechanical behavior for small values of the spin  $j$ . Several experimental scenarios for the manipulation of particles with spin include *nuclear magnetic resonance* (see, e.g., [76]) and *electron paramagnetic resonance* (see, e.g., [87]), with a large variety of applications. Chapter 9, in section 9.1 presents a discussion of control problems in nuclear magnetic resonance. It should be mentioned that techniques for the control of spins using electromagnetic fields have been in place in these areas for several decades. Often these techniques have been suggested by physical intuition and the introduction of control theory in these areas is more recent.

### 2.3.1 Introduction of the spin degree of freedom in the dynamics of matter and fields

The existence of the spin degree of freedom, which was first demonstrated in the Stern-Gerlach experiment (see previous chapter), can be derived theoretically using relativistic quantum electrodynamics (see, e.g., [50]). In this way, one can also derive the form of the Hamiltonian modeling the interaction between electromagnetic field and particles with spin. In a heuristic way, this Hamiltonian can be obtained by using the correspondence between orbital angular momentum and spin (intrinsic) angular momentum. After having rewritten the Hamiltonian of classical electrodynamics in terms of orbital angular momentum, one ‘augments’ the orbital angular momentum with

an additional spin term (which in some cases is the dominant term). We illustrate this for the term which models the interaction of the particles with the external field, that is for  $H_{Ie}$  in (2.48).

The definition of the **orbital angular momentum**  $\vec{L}_j$  for the particle  $j$  is given in terms of the canonical momentum  $\vec{p}_j$ , as

$$\vec{L}_j := \vec{p}_j \times \vec{r}_j. \quad (2.60)$$

Let us make an assumption of *low intensity of the field*. Under this condition, in the expression of  $H_{Ie}$  in (2.48), the magnitude of the first term on the right is much smaller than the second term, and so we shall neglect it in the following discussion. Let us also make a long wavelength approximation by assuming that the wavelength of the field is much larger than the size of the system of particles which is assumed localized in a small region of space, just as it was done in 2.2.1.2. This allows us to neglect the dependence of the magnetic field on space. Let  $\vec{0}$  be the position of the center of mass of the system of particles. The expression (2.12) for the magnetic field in terms of the vector potential  $\vec{A}$  suggests writing  $\vec{A}_e$  as

$$\vec{A}_e(\vec{r}, t) = -\frac{1}{2}\vec{r} \times \vec{B}_e(\vec{0}, t),$$

where  $\vec{B}_e$  represents the external magnetic field. Notice that this leaves  $\vec{A}_e$  in the Coulomb gauge since  $\nabla \cdot \vec{A}_e = 0$ . Using this formula for the vector potential  $\vec{A}_e$  in (2.48),  $H_{Ie}$  (with low field intensity approximation) takes the form

$$H_{Ie} = \frac{1}{2} \sum_{j=1}^N \frac{q_j}{m_j} \vec{p}_j \cdot (\vec{r} \times \vec{B}_e(\vec{0}, t)) + \sum_{j=1}^N q_j \phi_e(\vec{r}_j, t).$$

Using the vector identity

$$\vec{P} \cdot (\vec{R} \times \vec{B}) = \vec{B} \cdot (\vec{P} \times \vec{R}),$$

and the definition of the orbital angular momentum (2.60), we write  $H_{Ie}$  as

$$H_{Ie} = \sum_{j=1}^N \frac{q_j}{2m_j} \vec{B}_e(\vec{0}, t) \cdot \vec{L}_j + \sum_{j=1}^N q_j \phi_e(\vec{r}_j, t).$$

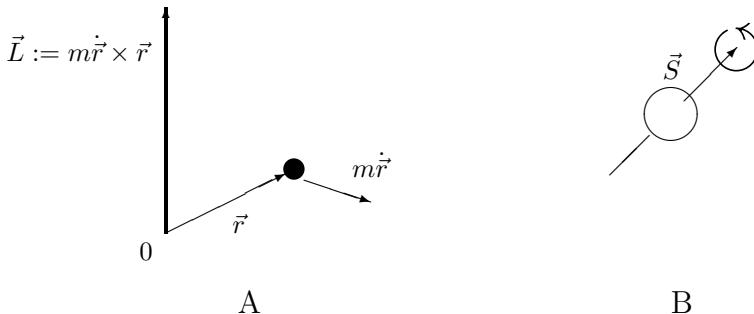
The operator

$$\hat{L}_j := -i\hbar \vec{r}_j \times \nabla_j \quad (2.61)$$

is the *orbital angular momentum* operator for the  $j$ -th particle. With this definition we rewrite the operator corresponding to  $H_{Ie}$  as

$$\hat{H}_{Ie} = \sum_{j=1}^N \frac{q_j}{2m_j} (\hat{L}_j) \cdot \vec{B}_e(\vec{r}_0, t),$$

and this is another way of writing the part of the Hamiltonian which models the interaction between particles and fields. The spin angular momentum is an additional, intrinsic angular momentum of the particle. If the particle has spin, the Hamiltonian is therefore heuristically modified by adding to  $\hat{L}_j$  a term  $g_j \hat{S}_j$ , where  $\hat{S}_j$  is the spin angular momentum operator and  $g_j$  is an extra factor called the *Lande factor*. Therefore, we replace the term  $\sum_{j=1}^N \frac{q_j}{2m_j} (\hat{L}_j) \cdot \vec{B}_e(\vec{r}_0, t)$ , with  $\sum_{j=1}^N \frac{q_j}{2m_j} (\hat{L}_j + g_j \hat{S}_j) \cdot \vec{B}_e(\vec{r}_0, t)$ .



**FIGURE 2.7:** If the conjugate momentum  $\vec{p}$  is equal to the kinematical momentum  $m\dot{\vec{r}}$  the *orbital angular momentum*  $\vec{L}$  is given by the cross product  $m\dot{\vec{r}} \times \vec{r}$  (Part A). In many cases, a quantum mechanical treatment has to consider an intrinsic angular momentum, the spin  $\vec{S}$  depicted in Part B.

The state of a general system of moving particles that have spin will be described by a vector in a Hilbert space  $\mathcal{H}_P \otimes \mathcal{H}_S$ , where  $\mathcal{H}_P$  refers to the position degrees of freedom and is infinite dimensional, while  $\mathcal{H}_S$  refers to the spin degrees of freedom and is finite dimensional.  $\mathcal{H}_S$ , in turn, is the tensor product of Hilbert spaces  $\mathcal{H}_{Sk}$ , where  $Sk$  refers to the spin of the  $k$ -th particle. The state  $|\psi\rangle$  is then written as

$$|\psi(t)\rangle = \sum_k \int \psi_k(\vec{r}, t) |\vec{r}\rangle \otimes |k\rangle d\vec{r},$$

where the states  $|k\rangle$  form an orthogonal basis in  $\mathcal{H}_S$ . The vector operator  $\hat{S}_k$  only acts on the  $\mathcal{H}_{Sk}$  part of the Hilbert space.

### 2.3.2 Spin networks as control systems

#### 2.3.2.1 Spin in an electromagnetic field

The simplest example of a control system involving particles with spin is a *single* particle, with all other degrees of freedom neglected, interacting with a classical magnetic field  $\vec{B}_e$ .  $\vec{B}_e$ , which plays the role of the control, is assumed constant in space and varying with time. In the associated Hamiltonian, the only relevant term is the one which models the interaction of the spin angular momentum with the external field. As we have seen in the previous subsection, under long wavelength approximation, this term has the form of a dot product of the spin operator with the external field. Given  $\vec{B}_e = \vec{B}_e(t) := \{B_x, B_y, B_z\}$  and  $\hat{S} := \{\hat{S}_x, \hat{S}_y, \hat{S}_z\}$ , the Hamiltonian can therefore be written as

$$\hat{H}(t) := \gamma \vec{S} \cdot \vec{B}_e(t) = \gamma(\hat{S}_x B_x(t) + \hat{S}_y B_y(t) + \hat{S}_z B_z(t)), \quad (2.62)$$

for some proportionality factor  $\gamma$  called the gyromagnetic ratio.<sup>25</sup>

In nuclear magnetic resonance (NMR) [76] and electron paramagnetic resonance (EPR)[87] experiments, a sinusoidal magnetic field at different frequencies is applied to a particle with Hamiltonian  $\hat{H}(t)$  as in (2.62). Only sinusoids at a given frequency induce a state transition. This frequency contains the information on the nature of the particle under consideration (as it depends on the parameter  $\gamma$  in (2.62)).<sup>26</sup> This method has enjoyed for several decades important applications. More recently particles with spin have been proposed as elementary components of quantum computers. In fact, particles with spin  $\frac{1}{2}$  are attractive devices to be used as quantum bits in quantum information theory.

Let us consider (2.62) and, for sake of simplicity, the case of a particle with spin  $j = \frac{1}{2}$ . We consider a basis of eigenstates of  $S_z$ . The corresponding eigenvalues (possible values of the spin when measured, setting  $\hbar = 1$ ) are  $\pm \frac{1}{2}$  and the eigenvectors are denoted by  $|\pm \frac{1}{2}\rangle$ . Let us express the state  $|\psi(t)\rangle$  as  $|\psi(t)\rangle = c_1(t)|\frac{1}{2}\rangle + c_2(t)|-\frac{1}{2}\rangle$ . From the Schrödinger equation and the Hamiltonian (2.62), we derive the differential equation for the coefficients  $\vec{c} = (c_1, c_2)^T$

$$i \frac{d}{dt} \vec{c} = \gamma \frac{1}{2} (\sigma_x B_x(t) + \sigma_y B_y(t) + \sigma_z B_z(t)) \vec{c}. \quad (2.63)$$

The matrices  $\frac{1}{2}\sigma_{x,y,z}$  are the matrix representations of the operators  $S_{x,y,z}$ , where  $\sigma_{x,y,z}$  are the Pauli matrices already introduced in (1.20).

$$\sigma_x := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2.64)$$

---

<sup>25</sup>The value of the factor  $\gamma$  depends on the Lande factor  $g$ , but often its value has to be modified to take into account the effect of neighboring particles. This modification is called *chemical shift* (cf. Chapter 9, section 9.1).

<sup>26</sup>cf. section 9.1 in Chapter 9 for more discussion on methods of nuclear magnetic resonance.

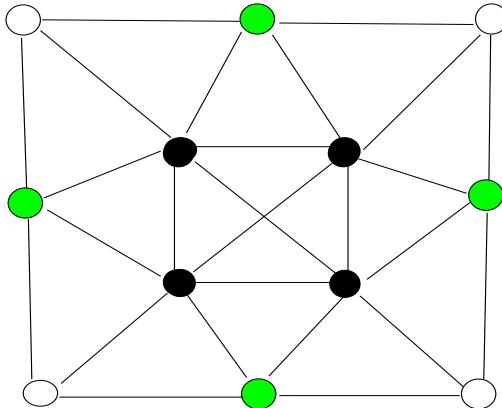
In the case of an ensemble with state described by a density matrix  $\rho$ ,  $\rho$  satisfies Liouville's equation

$$i \frac{d}{dt} \rho = \gamma \frac{1}{2} [\sigma_x B_x(t) + \sigma_y B_y(t) + \sigma_z B_z(t), \rho] := \gamma \frac{1}{2} [\vec{\sigma} \cdot \vec{B}, \rho]. \quad (2.65)$$

It is conceivable to measure the spin in one direction in space, i.e., one of  $S_{x,y,z}$  with a selective measurement or to measure its mean value  $\text{Tr}(\rho S_{x,y,z})$  with a nonselective measurement.

### 2.3.2.2 Control systems of several spins

A more general situation is the one of a *network* of particles with spin (with the position degrees of freedom neglected) interacting with each other and with an external magnetic field. This can be realized for example in a molecule when two or more nuclei (and/or electrons) interact.



**FIGURE 2.8:** The magnetic molecule  $Mn_{12}$  can be seen as a spin network. Three different types of Manganese ions, which are represented by the white, grey and black circles, all have spin and interact with each other. The lines connecting circles represent interactions among them (cf. [30] and references therein).

In this case, the Hamiltonian operator is composed of a part modeling the

interaction among the particles and a part modeling the interaction with the external field. This, under the appropriate approximations, as seen above, has the form

$$\hat{H}_I(t) = \sum_{j=1}^N \gamma_j \hat{S}_j \cdot \vec{B}(t),$$

where  $\hat{S}_j$  ( $\gamma_j$ ) is the spin angular momentum (gyromagnetic ratio) of the  $j$ -th particle and  $N$  is the number of particles. The total Hilbert space for this system is the tensor product of the Hilbert spaces for the single subsystems, each one having possibly different dimensions according to the values of the spins.<sup>27</sup> The  $x, y, z$ -component of the vector operator  $\hat{S}_j := \{\hat{S}_{xj}, \hat{S}_{yj}, \hat{S}_{zj}\}$  acting on the tensor product Hilbert space,  $j = 1, \dots, N$ , is

$$\hat{S}_{xj,yj,zj} = \mathbf{1} \otimes \mathbf{1} \otimes \cdots \otimes \mathbf{1} \otimes \hat{S}_{x,y,z} \otimes \mathbf{1} \otimes \cdots \otimes \mathbf{1}, \quad (2.66)$$

where  $\hat{S}_{x,y,z}$  appears in the  $j$ -th position.

The part of the Hamiltonian which models the interaction among particles may take different forms according to different situations. It is in principle obtained with a procedure similar to the one described in the previous subsection for the interaction with the magnetic field. Starting from the electrodynamic interaction (2.47) one expresses the interaction in terms of angular momentum interactions and then introduces the spin. Various approximations need to be introduced and the subject of modeling spin-spin interaction has a vast literature. The book [2] contains several examples of Hamiltonians modeling the interaction among particles with spin. A common form for the Hamiltonian modeling interaction among spins is *Heisenberg interaction* given by (cf. (2.66))

$$\hat{H}_0 = \sum_{k < j} J_{kj} \hat{S}_k \cdot \hat{S}_j, \quad (2.67)$$

where  $k$  and  $j$  vary over all the spin particles (modulo the constraint  $k < j$ ) and  $J_{kj}$  are the *coupling constants*. Another form is the *Ising interaction* which has the form

$$\hat{H}_0 = \sum_{k < j} J_{kj} \hat{S}_{zk} \hat{S}_{zj}, \quad (2.68)$$

where  $\hat{S}_{z(k,j)}$  is the  $z$  component of the spin angular momentum for particle  $k, j$ . The dynamics is described by the Schrödinger equation with the appropriate Hamiltonian.

To write the Schrödinger equation as a differential equation in matrix form, one considers the Hilbert space of the system of  $n$  spins which is the tensor product of the Hilbert spaces of the single spins. With the matrix form for the Hamiltonian, one can write the differential equation for the components

---

<sup>27</sup>In particular, for a spin  $j$  the associated Hilbert space has dimension  $2j + 1$ .

of the state, as well as for the density matrix. It is easily seen that in such a differential equation, the control  $\vec{B}_e(t)$  multiplies the state.

As the output of a putative experiment, one may consider the result of a selective measurement of the spin angular momentum in the  $z$  direction of the  $j$ -th spin. This is a selective measurement of the observable

$$\hat{M}_{jz} := \mathbf{1} \otimes \cdots \otimes \mathbf{1} \otimes \hat{S}_z \otimes \mathbf{1} \otimes \cdots \otimes \mathbf{1},$$

where  $\hat{S}_z$  appears in the  $j$ -th position. Another possible output in a nonselective measurement experiment is the expectation value of such an observable,  $\text{Tr}(\mathbf{1} \otimes \cdots \otimes \mathbf{1} \otimes \sigma_z \otimes \mathbf{1} \otimes \cdots \otimes \mathbf{1}\rho)$ . Another possible output in a nonselective measurement is the (expectation) value of the total spin in the  $z$  direction, which is proportional to  $\text{Tr}(\sum_j \hat{M}_{jz}\rho)$ , where the sum is taken over all the spins.

More discussion on NMR experiments and spin manipulation is given in [Chapter 9](#).

## 2.4 Mathematical Structure of Quantum Control Systems

A finite dimensional **control system** is a system of ordinary differential equations which displays one or more time varying functions of time,  $u := u(t)$ , the *controls*, to be chosen in an appropriate set of functions. Therefore a general control system has the form

$$\dot{x} = f(t, x, u). \quad (2.69)$$

Here  $x$  represents the *state* of the system, in general varying on an appropriate manifold, and  $f$  is a vector field. The state  $x$  completely describes the system under consideration. In most cases, the whole state  $x$  is not accessible to an outside observer. Only a (vector) function of  $x$ ,  $y := y(x)$ , is accessible and this is called the *output*. It models the interface between the system and the external world.

For **quantum control systems**, the relevant equation describing the dynamics is the Schrödinger equation. Let  $\vec{\psi}$  denote the vector column corresponding to the ket  $|\psi\rangle$  in a given basis (cf. [subsection 1.1.5](#)). In this book, we shall focus on finite dimensional quantum control systems, after appropriate approximations, so that  $\vec{\psi}$  is a vector in a finite dimensional vector space. For quantum control systems, the energy of the system depends on some functions of time which play the role of controls. In the examples discussed in the previous sections, these model time-dependent electromagnetic fields can be assumed *constant in space*. As a consequence, the Hamiltonian depends on

time through the controls and the Schrödinger equation in the column vector form can be written (setting  $\hbar = 1$ ) as

$$\frac{d}{dt}\vec{\psi} = -iH(u(t))\vec{\psi}. \quad (2.70)$$

Here  $H(u)$  is a matrix function of  $u$  which is Hermitian for every value of  $u$ . Equation (2.70) should be compared with (2.69) with  $\vec{\psi}$  playing the role of the state. In this case, because of the condition  $|||\psi\rangle|| = 1$ , the state  $\vec{\psi}$  varies on a complex sphere with radius equal to one in the space  $\mathbb{C}^n$ . The classes of examples treated in the previous sections show that  $H(u)$  often has the form

$$H(u) = H_0 + \sum_k H_k u_k. \quad (2.71)$$

Therefore, quantum control systems can be modeled as **bilinear systems** (i.e., the right hand side,  $f$ , of (2.69) is a linear function of the state and an affine function of the control) with state varying on a complex sphere with radius equal to one. Moreover, since the matrices  $H_0, H_k$  are Hermitian, the matrices  $-iH_0, -iH_k$  are skew-Hermitian. In some cases the control allows to *switch* among Hamiltonians in a finite set  $\{H_1, \dots, H_p\}$ . This can be obtained for example by adjusting the experimental set-up so as to induce different evolutions.<sup>28</sup> In these cases, we can write the Hamiltonian as in (2.71) with  $H_0 \equiv 0$  and the  $u_k$ 's attaining values such that only one of them is equal to 1 and the others are equal to zero. In these cases, the controls do not necessarily have a direct physical meaning (of an electromagnetic field). They are only a mathematical tool to model the switching between different configurations.

#### 2.4.0.3 Control of ensembles

If the system is an ensemble, the relevant equation describing the dynamics is Liouville's equation with the Hamiltonian depending on the control,

$$\dot{\rho} = [-iH(u(t)), \rho]. \quad (2.72)$$

In this case, the state  $\rho$  varies in the set of Hermitian, semidefinite positive matrices with trace equal to one, i.e., in the set of density matrices. This set is not a manifold but rather a manifold with boundary. However, we can assume that the state  $\rho$  is just varying in the manifold of Hermitian matrices with trace equal to one as equation (2.72) preserves the property of the matrix  $\rho$  being positive semidefinite. In fact, the solution of (2.72) can be written as

$$\rho(t) = X(t)\rho(0)X^\dagger(t), \quad (2.73)$$

---

<sup>28</sup>An example is given by the quantum information processor based on ion traps discussed in section 9.3 of Chapter 9.

where  $X(t)$  is the unitary matrix representing the propagator, which is a solution at time  $t$  of the Schrödinger operator equation. This is given by

$$\dot{X} = -iH(u)X, \quad X(0) = \mathbf{1}_{n \times n}, \quad (2.74)$$

where  $\mathbf{1}_{n \times n}$  is the  $n \times n$  identity matrix.

#### 2.4.0.4 Control of the evolution operator

System (2.74) may be taken as the object of control. The state  $X$  is now varying in the Lie group (see [Chapter 3](#)) of unitary matrices of dimension  $n$ , with  $n$  being the dimension of the system. Considering  $X$  as the object of control can be more natural and convenient in a number of situations, for the following reasons:

1. Systems of the form (2.74) have been among the most studied in *geometric control theory* [112], [197] and therefore many tools developed there can be directly applied.
2. As we have seen in [subsection 1.4.1](#), in applications to quantum computing, the state  $\vec{\psi}$  represents the *information* and the matrix  $X$  describes the *operation* to be carried out according to

$$\vec{\psi}(t) = X(t)\vec{\psi}(0).$$

Therefore it is a natural formulation, in this context, to pose the control problem in terms of the matrix  $X$ , when the desired evolution has to perform a given (logic) operation. This problem is of interest when we want to design a gate transforming the state  $\psi$  in a desired way. The problem is to choose the control  $u$  in (2.74) so that the solution  $X$  of (2.74) attains a desired value at a given time.

3. If we are able to solve a control problem for  $X$ , we are also able to control pure states  $\psi$  and ensembles  $\rho$  for the given system.

#### 2.4.0.5 Output of a quantum control system

So far, quantum control systems appear to be just a special case of classical control systems with a specific structure to their dynamics. However, in describing the interaction of a quantum control system with an observer, new elements have to be taken into account which are a consequence of the measurement postulate discussed in [subsection 1.2.2](#).

The output of a quantum control system consists of one or more observables that can be *measured simultaneously* no matter what the current state of the system is, i.e., the output consists of the results of compatible measurements (cf. [subsection 1.2.2](#)). The measurement may be selective or nonselective (cf.

[subsection 1.2.3](#)) and can be modeled differently according to the specific experimental apparatus. In general, for selective measurements, the probability of each result will depend on the state. Therefore, even though the dynamics is described by a deterministic model, one should consider probability theory in the modeling of the output. This is not the case in a nonselective measurement if one considers as the output  $y$  the expectation value of one or more compatible observables. The output  $y$  can be written, in the case of a Von Neumann-Lüders measurement, as  $y = \text{Tr}(S\rho)$ , where  $S$  is the corresponding observable and  $\rho$  the density matrix. All these measurements are discrete measurements, i.e., one-time (theoretically) instantaneous measurements. After the measurement, the state is modified according to the measurement postulate, and therefore the dynamics has to be considered starting from a different initial condition. For these schemes, the modification of the state due to measurement prevents the implementation of feedback control, and this is the main difference between classical and quantum schemes. In fact, all the control schemes to be applied to the systems described in the previous sections must be **open loop control schemes** and all the methods described in this book will be open loop schemes. This means that the control law is completely determined prior to the control experiment. It is used and *not modified* during the experiment.

---

## 2.5 Notes and References

The models of quantum control systems described in this chapter are simplified, not only because of the various approximations we have introduced (semiclassical approximation, dipole approximation, nonrelativistic treatment, etc.) but also because we have neglected the presence of the environment and assumed the *system* to be isolated. In fact, the design of controls which would render the evolution robust to the presence of an external environment is one of the main theoretical and practical problems in quantum control. An excellent treatment of dynamical models of open quantum systems is given in [33]. More generally, the set of models and potential applications presented in this chapter are only a subset of the models of quantum control systems of interest in applications. However, it is a large subset which encompasses several important applications, as we have pointed out. It is amenable to applications of techniques from geometric control theory, which will be the main object of the following chapters.

The book [39] contains several examples of quantum control systems, while the two recent books [174] [192] focus on the control of molecular dynamics. The paper [139] presents further motivations and applications to study quantum control systems and contains an extensive literature. In recent years,

several monographs have been written on the neighboring field of quantum information and computation. In this area, we refer to [155], [125], [35], [19].

As we have mentioned, the measurement postulate prevents the direct implementation of feedback for the models described here. One can circumvent this limitation in several ways. For example, one could imagine having a very large number of identical quantum systems evolving (simultaneously) in the same conditions starting from the same initial value. Measuring each system at slightly spaced instants of time simulates a continuous measurement. At each time the systems on which the measurement has already been performed are not considered anymore and the control is modified to influence the dynamics of the remaining systems, so as to simulate feedback control. More genuine quantum feedback schemes on a single system can be implemented in the case of continuous weak measurement, where the modification of the state (called ‘*quantum state reduction*’) is slow enough so that it can be compensated for by appropriate control. In these cases, the model of the dynamics has to be modified to include the continuous back-action of the measurement. The resulting model is a stochastic differential equation and the tools for control methods are based on the methods of stochastic control theory. This approach has received great interest in recent years starting from the seminal work of Belavkin until the more recent investigations and experimental realizations by Wiseman, Mabuchi et al. (see, e.g., [214], [221] and references therein). This approach is not the main topic of this book, which focuses on deterministic quantum models, open loop control design, and analysis of quantum dynamics with geometric methods.

## 2.6 Exercises

**Exercise 2.1** Repeat the discussion in section 2.2 leading to formula (2.59) in the energy representation, namely expressing the state  $|\psi\rangle$  in terms of energy eigenstates.

**Exercise 2.2** Use the definition of the position and canonical momentum operator to prove the commutation relation (2.38). Which ones are the compatible observables and which ones are incompatible? Observe that from this formula and the discussion in 1.2.2.3 one obtains Heisenberg uncertainty relation for the position  $\hat{x}$  and momentum  $\hat{p}_x$  (consider a single particle moving in the  $x$  direction, vector potential equal to zero), that is

$$\langle \Delta \hat{x}^2 \rangle_\psi \langle \Delta \hat{p}_x^2 \rangle_\psi \geq \frac{1}{4} \hbar^2.$$

**Exercise 2.3** Using the definitions of the destruction and creation operator

$\hat{a}$  and  $\hat{a}^\dagger$ , (2.41), (2.42), and the commutation relation between  $\hat{x}$  and  $\hat{p}$  prove (2.43) and (2.44).

**Exercise 2.4** A sufficient condition for the validity of *Helmholtz theorem* 2.24 is the existence of the *spatial Fourier transform* of the field  $\vec{V} = \vec{V}(\vec{r})$ , defined as

$$\vec{\mathcal{V}}(\vec{k}) = \frac{1}{(2\pi)^{\frac{3}{2}}} \int_{\mathbf{R}^3} \vec{V}(\vec{r}) e^{-i\vec{k}\cdot\vec{r}} d\vec{r},$$

with inversion formula

$$\vec{V}(\vec{r}) = \frac{1}{(2\pi)^{\frac{3}{2}}} \int_{\mathbf{R}^3} \vec{\mathcal{V}}(\vec{k}) e^{i\vec{k}\cdot\vec{r}} d\vec{k}.$$

To prove Helmholtz decomposition first prove the vector identity

$$\vec{\mathcal{V}}(\vec{k}) = (\vec{k} \cdot \vec{\mathcal{V}})\vec{k} + (\vec{k} \times \vec{\mathcal{V}}) \times \vec{k}.$$

Then take the anti-transform of both sides and show that the anti-transform of the first term on the right hand side is irrotational (longitudinal) and the one of the second term is divergence free (transversal).

# Chapter 3

---

## Controllability

Given a quantum control system describing an experimental set-up, a natural question concerns the type of feasible experiments that can be performed. This question can be formulated in terms of the kind of state transfers admissible with the given system, namely, it is a question of *controllability*.

In this chapter, we shall first consider this question for the Schrödinger operator equation in the general form<sup>1</sup>

$$\dot{X} = -iH(u)X, \quad X(0) = \mathbf{1}_{n \times n}. \quad (3.1)$$

We shall study the set of unitary matrices that can be obtained by changing the control for this system, i.e., the reachable set of states for this system. This set is always a subset of the group of unitary matrices of dimension  $n$ ,  $U(n)$ . As  $-iH(u)$  is skew Hermitian for every  $u$ ,  $X$  must be unitary. If the set of possible matrices that can be obtained for system (3.1) is the set of all the unitary matrices, the system is said to be *controllable*.

The study of controllability for the Schrödinger operator equation gives the crucial information to study the controllability of the state vector for Schrödinger equation and of the density matrix for Liouville's equation. We shall define notions of controllability for these systems and give criteria for controllability.

We first give the definition of reachable sets. Consider system (3.1), where the control  $u$  is assumed to belong to a space of functions  $\bar{\mathcal{U}}$ . Denote by  $X(t, u)$  the solution of (3.1), with control  $u$ , at time  $t$ . The **reachable set at time**  $T > 0$  for system (3.1),  $\mathcal{R}(T)$ , is the set of all the unitary matrices  $\bar{X}$  such that there exists a control  $u \in \bar{\mathcal{U}}$  with  $X(T, u) = \bar{X}$ . The reachable set  $\mathcal{R}(\leq T)$  is defined as  $\mathcal{R}(\leq T) := \cup_{0 \leq t \leq T} \mathcal{R}(t)$ . The **reachable set**  $\mathcal{R}$  is defined as<sup>2</sup>

$$\mathcal{R} = \cup_{T \geq 0} \mathcal{R}(T). \quad (3.2)$$

In this chapter, we shall describe the reachable set for system (3.1) in the case where  $\bar{\mathcal{U}}$  is the set of piecewise constant functions with values in a set

<sup>1</sup>We set  $\hbar = 1$  (This is done without loss of generality by changing units, or incorporating  $\hbar$  into the Hamiltonian  $H$ ) and do not assume, in general, any special dependence of the Hamiltonian  $H$  on the control  $u$ .

<sup>2</sup>These definitions extend naturally to other types of systems although one typically should specify the initial state.

$\mathcal{U}$ . Notice that if  $X(t, u)$  is the solution of (3.1) with initial condition equal to the identity, then the solution with initial condition equal to  $S$  is given by  $X(t, u)S$ . Therefore if we can control the system to a state  $X_1$  with control  $u_1$  and to a state  $X_2$  with control  $u_2$ , we can also control it to a state  $X_2X_1$ , by using a control which is the concatenation of  $u_1$  and  $u_2$ . In other words,  $\mathcal{R}$  is a *semigroup*. It turns out that  $\mathcal{R}$  is actually a *Lie group*, namely a group with an additional differentiable structure. We will be able to characterize this Lie group exactly. Before we do that, we need to recall the basic definitions and facts about Lie algebras and Lie groups.

---

### 3.1 Lie Algebras and Lie Groups

We summarize in this section the basic definitions concerning Lie groups and Lie algebras. In the process, we shall recall some elementary concepts of linear algebra and differential geometry but omit some other elementary definitions for which we refer, e.g., to [184].

#### 3.1.1 Basic definitions for Lie algebras

##### 3.1.1.1 Lie algebras

**Definition 3.1.1** A **Lie algebra**  $\mathcal{L}$  over a field  $\mathcal{F}$  is a vector space over  $\mathcal{F}$  with an additional binary operation  $\mathcal{L} \times \mathcal{L} \rightarrow \mathcal{L}$ . This operation associates with an ordered pair of elements  $\{x, y\}$  in  $\mathcal{L}$  an element  $[x, y]$ . It is called the *Lie bracket* or *commutator* and it is required to satisfy the following axioms:

1. Bilinearity:

$$\begin{aligned}[x + y, z] &= [x, z] + [y, z], & [x, y + z] &= [x, y] + [x, z], \\ [\alpha x, y] &= \alpha[x, y], & \forall \alpha \in \mathcal{F}. \end{aligned}\tag{3.3}$$

2.

$$[x, x] = 0, \quad \forall x \in \mathcal{L}.\tag{3.4}$$

3. Jacobi identity:

$$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0.$$

In the cases of interest to us where the field  $\mathcal{F}$  is the set of real numbers  $\mathbf{R}$  or the set of complex numbers  $\mathbb{C}$ , condition (3.4) can be replaced (cf. Exercise 3.5) by the *skew-symmetry condition*, that is:

$$[x, y] = -[y, x].$$

The structure of a Lie algebra is unambiguously determined by the commutation relations on a basis, since these determine the value of the commutator of any pair of elements in the Lie algebra.

Perhaps the simplest nontrivial example of a Lie algebra is the set of vectors in three dimensional space,  $\mathbf{R}^3$ , with the cross product playing the role of the commutator. If we choose a basis  $\{\vec{i}, \vec{j}, \vec{k}\}$ , then the commutation relations are given by

$$\vec{i} \times \vec{j} = \vec{k}, \quad \vec{j} \times \vec{k} = \vec{i}, \quad \vec{k} \times \vec{i} = \vec{j}. \quad (3.5)$$

We shall be concerned with Lie algebras of matrices which are called *linear Lie algebras*, where the commutator  $[A, B]$  is the standard matrix commutator

$$[A, B] := AB - BA.$$

The Lie algebra of all the  $n \times n$  matrices with real (complex) entries is denoted by  $gl(n, \mathbf{R})$  ( $gl(n, \mathbb{C})$ ) and it is called the *general linear* algebra over the real (complex) numbers.<sup>3</sup>

### 3.1.1.2 Subalgebras

**Definition 3.1.2** Given a Lie algebra  $\mathcal{L}$ , consider a subspace  $\mathcal{A} \subseteq \mathcal{L}$ . If  $\mathcal{A}$  with the commutator defined on  $\mathcal{L}$  is a Lie algebra, then  $\mathcal{A}$  is called a **subalgebra** of  $\mathcal{L}$ .

Subalgebras of  $gl(n, \mathbf{R})$  or  $gl(n, \mathbb{C})$  are called *linear Lie algebras*. An example of a subalgebra of  $gl(n, \mathbf{R})$  is the *special linear algebra*  $sl(n, \mathbf{R})$  (or  $sl(n)$ ) which is the Lie algebra of  $n \times n$  real matrices with trace equal to zero. Another example is the *orthogonal Lie algebra*  $so(n, \mathbf{R})$  (or  $so(n)$ ) which is the Lie algebra of skew-symmetric matrices. In the following, we shall be interested in the **Lie algebra  $u(n)$  of skew-Hermitian  $n \times n$  matrices** considered as a Lie algebra over the real field. For instance, all matrices  $-iH(u)$  in (3.1) are in  $u(n)$ . The subalgebra  $su(n)$  of  $u(n)$  will play an important role. It consists of the matrices in  $u(n)$  with zero trace. The simplest but very important example is  $su(2)$  which is spanned by the the multiples of the Pauli matrices (1.20)

$$\bar{\sigma}_x := \frac{1}{2} \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad \bar{\sigma}_y := \frac{1}{2} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \bar{\sigma}_z := \frac{1}{2} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}. \quad (3.6)$$

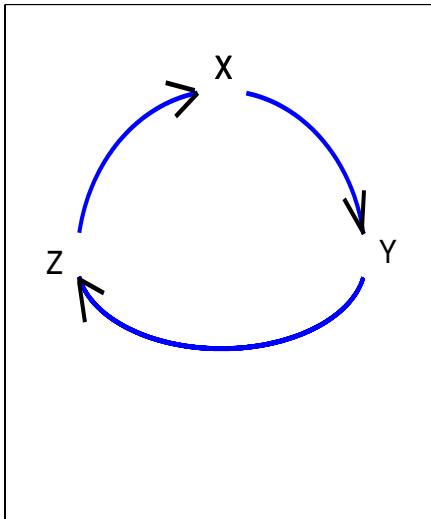
The matrices  $\bar{\sigma}_{x,y,z}$  satisfy the commutation relations

$$[\bar{\sigma}_x, \bar{\sigma}_y] = \bar{\sigma}_z, \quad [\bar{\sigma}_y, \bar{\sigma}_z] = \bar{\sigma}_x, \quad [\bar{\sigma}_z, \bar{\sigma}_x] = \bar{\sigma}_y, \quad (3.7)$$

which completely describe the Lie algebra  $su(2)$ .

---

<sup>3</sup> $gl(n, \mathbf{R})$  is a vector space over the real field while  $gl(n, \mathbb{C})$  is a vector space over the complex field.



**FIGURE 3.1:** This graph is useful to remember the commutation relations among  $\{\bar{\sigma}_x, \bar{\sigma}_y, \bar{\sigma}_z\}$ . If we want to find the commutator  $[\bar{\sigma}_A, \bar{\sigma}_B]$ , we follow a branch from  $A$  to  $B$  on the graph. The following node of the graph is the result except possibly for the sign. If our path follows the same direction as the arrows then the sign is positive otherwise it is negative.

### 3.1.1.3 Homomorphisms and Isomorphisms

**Definition 3.1.3** Given two Lie algebras  $\mathcal{L}_1, \mathcal{L}_2$ , a **homomorphism**  $\phi$ ,  $\phi : \mathcal{L}_1 \rightarrow \mathcal{L}_2$ , is a linear map which preserves the Lie bracket of  $\mathcal{L}_1$ , namely

$$\phi([x, y]_1) = [\phi(x), \phi(y)]_2, \quad (3.8)$$

where  $[\cdot, \cdot]_{1,2}$  represent the Lie bracket on  $\mathcal{L}_{1,2}$ . A bijective homomorphism is called an **isomorphism**, and the two Lie algebras are called isomorphic if such an isomorphism exists. A homomorphism of  $\mathcal{L}_1$  onto a subalgebra of  $gl(n, \mathbf{R})$  or  $gl(n, \mathbf{C})$  is called a *representation* of  $\mathcal{L}_1$ .

It is easily seen that, for a linear map to be a homomorphism, (3.8) needs to be verified only for the elements of the basis of  $\mathcal{L}_1$ .

**Example 3.1.4** As a simple example of two isomorphic Lie algebras, consider the Lie algebras  $\mathcal{L}_1$  and  $\mathcal{L}_2$  defined as follows.  $\mathcal{L}_1$  is the Lie algebra of the vectors in  $\mathbf{R}^3$  with the cross product as the commutator. It is spanned by

$\vec{i}, \vec{j}, \vec{k}$ .  $\mathcal{L}_2 := su(2)$  is spanned by  $\bar{\sigma}_{x,y,z}$  in (3.6). Because of the commutation relations (3.7), (3.5), the following correspondence gives an isomorphism:

$$\vec{i} \leftrightarrow \bar{\sigma}_x, \quad \vec{j} \leftrightarrow \bar{\sigma}_y, \quad \vec{k} \leftrightarrow \bar{\sigma}_z.$$

### 3.1.1.4 Lie algebra generated by a set of elements

Given a set of vectors of a Lie algebra  $\mathcal{L}$ ,  $\{x_1, \dots, x_n\}$ , the set of all the (possibly repeated) commutators of  $\{x_1, \dots, x_n\}$  spans a subalgebra of  $\mathcal{L}$  which is called the **Lie algebra generated by**  $\{x_1, \dots, x_n\}$ . It will be indicated here by  $\{x_1, \dots, x_n\}_{\mathcal{L}}$ . The Lie algebra  $\{x_1, \dots, x_n\}_{\mathcal{L}}$  is the smallest subalgebra of  $\mathcal{L}$  containing  $\{x_1, \dots, x_n\}$ . Notice that, in general, in order to prove that  $\{x_1, \dots, x_n\}_{\mathcal{L}}$  is equal to a given subalgebra  $\mathcal{A}$ , we only need to produce, as (possibly repeated) commutators of  $\{x_1, \dots, x_n\}$ , a basis of  $\mathcal{A}$ . For example consider  $\mathcal{L}$  the Lie algebra of complex  $2 \times 2$  matrices over the complex numbers. The Lie algebra generated by  $\{\bar{\sigma}_x, \bar{\sigma}_y, \bar{\sigma}_z\}$  in  $\mathcal{L}$  is equal to  $su(2)$ . Because of (3.7), the Lie algebra generated by  $\{\bar{\sigma}_x, \bar{\sigma}_y\}$ ,  $\{\bar{\sigma}_x, \bar{\sigma}_y\}_{\mathcal{L}}$ , is equal to  $su(2)$ , as well.

## 3.1.2 Lie groups

### 3.1.2.1 General definition of a Lie group

**Definition 3.1.5** A **Lie group** is a group which is also an analytic differentiable manifold and such that the group operations  $\{x, y\} \rightarrow xy$ , and  $x \rightarrow x^{-1}$  are analytic.

An important example of a Lie group is the *general linear group*,  $Gl(n, \mathbb{C})$ , defined as the group of nonsingular matrices of dimension  $n$  with complex entries. This is a group under matrix multiplication since  $\det(AB) = \det(A)\det(B)$  and we can give it the structure of a differentiable manifold by mapping the real and imaginary parts of each entry of an element to open sets in  $\mathbf{R}^{2n^2}$ . Because the matrix multiplication and inversion only require analytic operations, this is a Lie group.

**Definition 3.1.6** Given two Lie groups  $G_1$  and  $G_2$ , an analytic map

$$\phi : G_1 \rightarrow G_2$$

is a **homomorphism of Lie groups** if it preserves the group operations of  $G_1$ , i.e., for every  $g_1, g_2 \in G_1$ ,

$$\phi(g_1g_2) = \phi(g_1)\phi(g_2), \quad \phi(g_1^{-1}) = [\phi(g_1)]^{-1}.$$

The operations on the left and right hand sides of these conditions refer to the operations in  $G_1$  and  $G_2$ , respectively.

A homomorphism of Lie groups which is also bijective is called an **isomorphism of Lie groups**.

### 3.1.2.2 Lie subgroups

The Lie groups we shall be most interested in are the Lie group of unitary matrices of dimension  $n$  and the Lie group of unitary matrices of dimension  $n$  with determinant equal to 1. The first is called the **unitary group** and denoted by  $U(n)$ ; the second is called the **special unitary group** and denoted by  $SU(n)$ . These are Lie subgroups of  $Gl(n, \mathbb{C})$  in the following sense:

**Definition 3.1.7** A subgroup  $H$  of a Lie group  $G$  which is also an analytic submanifold of  $G$  and a Lie group with the structure inherited from  $G$  is called a **Lie subgroup** of  $G$ .

There are various definitions of submanifold in differential geometry. We adopt the definition of [26] pg. 75 ff. (see also Theorem 6.6 pg. 83, therein). For us a submanifold of  $M$  is a subset  $N$  of  $M$  with the relative topology and such that, for every point  $p \in N$ , there is a chart  $(U, \phi)$  of  $M$  such that i)  $\phi(p) = 0$ , ii)  $\phi$  maps  $U$  onto an open ball of radius  $\epsilon$  with center at the origin in  $\mathbf{R}^m$ ,  $B_\epsilon^m$ ,  $m$  being the dimension of  $M$ , iii)  $\phi(U \cap N)$  is the set of points in  $B_\epsilon^m$  with the last  $m - n$  coordinates equal to zero,  $n$  being the dimension of  $N$ .

### 3.1.2.3 Lie algebra of a Lie group

Consider a Lie group  $G$  and a neighborhood  $N$  of the identity. A *tangent vector at the identity*  $I$  is defined as a class of equivalent curves  $\gamma : \mathbf{R} \rightarrow N$ . Two curves are equivalent if

$$\gamma_1(0) = \gamma_2(0) = I$$

and, for every function  $f : N \rightarrow \mathbf{R}$ ,

$$\frac{d}{dt}\Big|_{t=0}(f \circ \gamma_1(t)) = \frac{d}{dt}\Big|_{t=0}(f \circ \gamma_2(t)). \quad (3.9)$$

There is a one-to-one correspondence between classes of equivalent curves and *derivations* on the space of  $C^\infty$  functions  $N \rightarrow \mathbf{R}$ . The derivation corresponding to the curve  $\gamma$  maps  $f$  to  $\frac{d}{dt}|_{t=0}(f \circ \gamma(t))$ . Using the vector space structure on the space of derivations we can give the space of tangent vectors the structure of a vector space, the *tangent space*. This has the same dimension as a vector space as the dimension of the Lie group  $G$  as a manifold.

The tangent space at the identity of a *Lie group of matrices* can be given the structure of a Lie algebra over the reals as follows. Consider a parametrization of the group  $G$  as a subgroup of  $Gl(n, \mathbb{C})$  or  $Gl(n, \mathbf{R})$ , namely with  $2n^2$  or  $n^2$  (possibly redundant) parameters. Then a curve in  $G$  can be written, in these coordinates, as a map  $\mathbf{R} \rightarrow Gl(n, \mathbb{C})$  or  $\mathbf{R} \rightarrow Gl(n, \mathbf{R})$ ,  $t \rightarrow \tilde{X}(t)$ . The derivative  $\frac{d}{dt}$  of  $\tilde{X}(t)$  at zero is another matrix,  $\tilde{A}$ . It uniquely determines the derivatives in (3.9). The curve  $\{e^{\tilde{A}t} | t \in \mathbf{R}\}$  is in the same equivalence

class as  $\{\tilde{X}(t) | t \in \mathbf{R}\}$ . It is called the *one parameter subgroup* of  $G$  and it can be proved to belong to  $G$  for every  $t \in \mathbf{R}$ . In fact, this defines a one-to-one correspondence between tangent vectors and one parameter subgroups. Consider now the Lie algebra generated by the matrices  $\tilde{A}$  obtained as above. To each matrix  $\tilde{A}$  there corresponds a one parameter subgroup and therefore a tangent vector. In fact one can prove if  $e^{At}$  and  $e^{Bt}$  are in  $G$  so are  $e^{(A+B)t}$  and  $e^{[A,B]t}$ . Proofs of the above statements can be found in any of the standard introduction to Lie groups and Lie algebras such as [96], [184], [217]. The important facts for our purposes are summarized in the following.

**FACT. (Correspondence between Lie groups and Lie algebras)** Given a Lie group of matrices, the matrices obtained by differentiating curves at the identity at time  $t = 0$  generate the corresponding Lie algebra. This Lie algebra is isomorphic (in the sense of vector spaces) to the tangent space of the Lie group at the identity. In particular, it has the same dimension as the tangent space which is the dimension of the Lie group seen as an analytic manifold. Conversely, given a Lie algebra of matrices  $\mathcal{L}$ , the associated one dimensional subgroups generate a Lie group

$$e^{\mathcal{L}} := \{e^{A_1} e^{A_2} \cdots e^{A_m}, \quad A_1, A_2, \dots, A_m \in \mathcal{L}\}$$

which is the unique connected Lie subgroup of  $Gl(n, \mathbf{R})$  ( $Gl(n, \mathbb{C})$ ) with Lie algebra  $\mathcal{L}$ .

More notions of Lie algebras and Lie group theory will be presented in this book as they are needed. In particular see [Chapter 5](#) and [Chapter 8](#). The map from the Lie algebra  $\mathcal{L}$  to the corresponding Lie group  $e^{\mathcal{L}}$  which associate to  $A \in \mathcal{L}$ ,  $e^A \in e^{\mathcal{L}}$  is called the *exponential map*. It plays a prominent role in all the theory of Lie groups and Lie algebras and in what follows. In general  $e^A e^B \neq e^{A+B}$  unless  $A$  and  $B$  commute. There exist however several known formulas relating matrix exponentials known as *exponential formulas*. Some of the most useful are collected in [Appendix E](#) (some more formulas can be found for example in [184] section 3 Chapter 5).

### 3.2 Controllability Test: The Dynamical Lie Algebra

The main test for controllability of quantum systems (3.1) is the following:

**Theorem 3.2.1** *The set of reachable states for system (3.1) is the connected Lie group associated with the Lie algebra  $\mathcal{L}$  generated by  $\text{span}_{u \in \mathcal{U}}\{-iH(u)\}$ . In short*

$$\mathcal{R} = e^{\mathcal{L}}.$$

The Lie algebra  $\mathcal{L}$  is called **the dynamical Lie algebra** associated with the system. This is always a subalgebra of  $u(n)$ . In the case  $\dim(\mathcal{L}) = n^2 = \dim(u(n))$ , which is equivalent to  $\mathcal{L} = u(n)$  and  $e^{\mathcal{L}} = U(n)$ , the system is said to be **controllable**. In this case  $\mathcal{R} = U(n)$ , which means that every unitary matrix can be obtained by choosing an appropriate control in (3.1). We shall say that the system is controllable even in the case where  $\dim(\mathcal{L}) = n^2 - 1 = \dim(su(n))$  which is equivalent to  $\mathcal{L} = su(n)$  and  $e^{\mathcal{L}} = SU(n)$ . Sometimes we shall use the terminology **operator controllability** or **complete controllability** to distinguish this case (controllability of system (3.1)) from the case where we have controllability of the state  $|\psi\rangle$  which will be treated later in this chapter. The condition of Theorem 3.2.1 is often referred to as the **Lie Algebra Rank Condition**. The proof of Theorem 3.2.1 is presented in [Appendix D](#).

### 3.2.1 Procedure to generate a basis of the dynamical Lie algebra

To generate a basis of the dynamical Lie algebra  $\mathcal{L}$ , one first takes a basis in  $\text{span}_{u \in \mathcal{U}} \{-iH(u)\}, \{A_1, \dots, A_s\}$ . A *repeated Lie bracket* is an element of  $\mathcal{L}$  of the form  $[X_1, [X_2, [\dots [X_{p-1}, X_p]]]]$ , with  $X_1, \dots, X_p \in \{A_1, \dots, A_s\}$ . One defines the *depth* of a repeated Lie bracket  $B$  obtained with  $\{A_1, \dots, A_s\}$  as the number of Lie brackets performed to obtain  $B$ , i.e.,  $p - 1$  in the above definition. The recursive procedure is as follows:

**Step 0:** List the vectors of depth 0,  $\{A_1, \dots, A_s\}$ .

**Step k:**

1. Calculate the Lie brackets of the elements of depth  $k - 1$  obtained at step  $k - 1$ , with the elements of depth 0. This way one obtains elements of depth  $k$ .
2. Out of the elements obtained, select the ones that form a linearly independent set along with the ones obtained up to step  $k - 1$ .
3. Stop the procedure if there is no new vector or if the dimension of the linearly independent set is  $n^2 - 1$  or  $n^2$ .

The set of matrices obtained this way forms a basis of the dynamical Lie algebra  $\mathcal{L}$ .

A justification of this procedure is the object of Exercise 3.1. Notice that the procedure above described will always terminate since the dynamical Lie algebra  $\mathcal{L}$  is a subalgebra of  $u(n)$  and therefore it has at most dimension  $n^2$ .

**Example 3.2.2** Consider the system of two spin- $\frac{1}{2}$  particles with Ising interaction (2.68), interacting with a magnetic field which has zero component only

in the  $x$ -direction. The magnetic field plays the role of the control. Assuming two equal gyromagnetic ratios for the two particles, the set of matrices one obtains at Step 0 of the above procedure is given by the span of

$$\{A_1, A_2\} = \{i\sigma_z \otimes \sigma_z, i\sigma_x \otimes \mathbf{1}_{2 \times 2} + i\mathbf{1}_{2 \times 2} \otimes \sigma_x\},$$

where the matrices  $\sigma_x, \sigma_y, \sigma_z$  are the Pauli matrices defined in (1.20). At Step 1, using (1.21), we find the matrix  $A_3 = [A_1, A_2]$ , with

$$A_3 = 2i\sigma_y \otimes \sigma_z + 2i\sigma_z \otimes \sigma_y.$$

At Step 2 we do not find any new linearly independent matrix. Therefore, the dynamical Lie algebra  $\mathcal{L}$  is the span of  $\{A_1, A_2, A_3\}$ . It is a 3-dimensional proper subalgebra of  $su(4)$  and the system is not controllable. In fact,  $\mathcal{L}$  is isomorphic to  $su(2)$  if we set the correspondence

$$\frac{1}{2}i\sigma_x \otimes \mathbf{1}_{2 \times 2} \leftrightarrow \bar{\sigma}_x, \quad \frac{1}{2}i(\sigma_y \otimes \sigma_z + \sigma_z \otimes \sigma_y) \leftrightarrow \bar{\sigma}_y, \quad \frac{1}{2}i\sigma_z \otimes \sigma_z \leftrightarrow \bar{\sigma}_z,$$

as it is verified using (1.21) and the commutation relations (3.7).

### 3.2.2 Uniform finite generation of compact Lie groups and universal quantum gates

The proof of the controllability test presented in [Appendix D](#) shows, among other things, the two facts in the following corollary. The crucial assumption is that  $e^{\mathcal{L}}$  is a Lie subgroup of  $U(n)$  and  $U(n)$  is compact.

**Corollary 3.2.3** Consider a connected Lie group  $e^{\mathcal{L}}$  corresponding to a Lie algebra  $\mathcal{L}$ . Then

- a) Every element  $X_f$  in  $e^{\mathcal{L}}$  can be written in the form

$$X_f = e^{At_1} \cdots e^{At_r}, \quad (3.10)$$

with the indeterminates  $A$  in the set  $\mathcal{S} := \{A_1, \dots, A_s\}$  of generators of  $\mathcal{L}$  and  $t_1, \dots, t_r \geq 0$ . The number  $r$  will depend on  $X_f$ .

b) Every element  $X_f$  in a neighborhood  $N$  of the identity in  $e^{\mathcal{L}}$  can be expressed as a product of the form (3.10) with uniformly bounded  $r$ .

If the Lie group  $e^{\mathcal{L}}$  is *compact*, for example  $e^{\mathcal{L}} = U(n)$  or  $e^{\mathcal{L}} = SU(n)$ , then it is possible to show that the number of factors  $r$  is uniformly bounded as  $X_f$  varies in  $e^{\mathcal{L}}$ . In this case, the Lie group is said to be *uniformly finitely generated* (see, e.g., [126]) by the set of generators  $\mathcal{S}$ . Therefore every compact Lie group is uniformly generated by a set of generators of the corresponding Lie algebra. To see this, notice that, for an open neighborhood  $N$  of the identity in  $e^{\mathcal{L}}$ , we can write

$$e^{\mathcal{L}} = \cup_{K \in e^{\mathcal{L}}} KN.$$

This is an open cover of  $e^{\mathcal{L}}$  and, by compactness of  $e^{\mathcal{L}}$ , contains a finite subcover. Therefore we can write

$$e^{\mathcal{L}} = \cup_{i=1}^l K_i N,$$

for finite  $l$ . Since, from part a) of the above corollary,  $K_i$  is itself the finite product of elements of the form  $e^{A_t}$ , the result follows.

A physical interpretation of the above uniform generation result concerns universality of quantum gates. Consider two Hamiltonians associated with the skew-Hermitian matrices  $A_1$  and  $A_2$ . Each of them represents a possible evolution of the quantum system which modifies the state. We have interpreted each evolution  $e^{A_1 t}$  or  $e^{A_2 t}$  as a quantum logic gate (see subsection 1.4.1). Assume it is possible to switch between these two evolutions. We can call the set of gates  $\{A_1, A_2\}$  *universal* if, by switching between them, it is possible to generate all the (special) unitary evolutions. According to the above results, this is possible if and only if  $\{A_1, A_2\}$  generate the whole Lie algebra  $u(n)$  or  $su(n)$ . Moreover, since  $U(n)$  and  $SU(n)$  are compact Lie groups, the number of gates needed to generate every (special) unitary evolution is uniformly bounded. It is possible to show that almost every couple of skew-Hermitian matrices generate  $u(n)$ . Therefore the above result is often quoted by saying that ‘almost every quantum logic gate is universal’ (see [137], [218] for more discussion on this).

---

### 3.3 Notions of Controllability for the State

In the previous section, we saw that a calculation of the dynamical Lie algebra gives information on the set of all the possible transformations between states achieved with a physical apparatus. In this section we investigate in what cases this set of transformations allows any desired transfer between two states. In particular, we shall investigate the following notions of controllability.

**Definition 3.3.1** The Schrödinger equation

$$\frac{d}{dt}\vec{\psi}(t) = -iH(u(t))\vec{\psi}(t) \quad (3.11)$$

is **pure state controllable (PFC)** if for every pair of initial and final states,  $\vec{\psi}_0$  and  $\vec{\psi}_1$ , there exist control functions  $u$  and a time  $T > 0$  such that the solution of (3.11) at time  $T$ , with initial condition  $\vec{\psi}_0$ , is  $\vec{\psi}(T) = \vec{\psi}_1$ . Here  $\vec{\psi}_0$  and  $\vec{\psi}_1$  are two vectors on the complex sphere of radius 1,  $S_{\mathbb{C}}^{n-1}$ .

Since, for any vector  $\vec{\psi}_1$ ,  $\vec{\psi}_1$  and  $e^{i\phi}\vec{\psi}_1$  represent the same physical state, for any  $\phi \in \mathbb{R}$ , from a physics viewpoint, the following property is equivalent to PFC.

**Definition 3.3.2** The system (3.11) is **equivalent-state-controllable (ESC)** if, for every pair of initial and final states,  $\vec{\psi}_0$  and  $\vec{\psi}_1$  in  $S_{\mathbb{C}}^{n-1}$ , there exist controls  $u$  and a phase factor  $\phi$  such that the solution  $\vec{\psi}$  of (3.11), with  $\vec{\psi}(0) = \vec{\psi}_0$ , satisfies  $\vec{\psi}(T) = e^{i\phi}\vec{\psi}_1$ , at some  $T > 0$ .

Assume the system is an ensemble and the dynamics is described by Liouville equation for the density matrix  $\rho$  (cf. (1.48))

$$\frac{d}{dt}\rho(t) = [-iH(u(t)), \rho(t)]. \quad (3.12)$$

Then if the initial condition is  $\rho_0$ ,  $\rho(t)$  is unitarily equivalent<sup>4</sup> to  $\rho_0$  for every  $t$  (cf. (1.47)). Therefore the following definition is appropriate.

**Definition 3.3.3** The system (3.12) is **density matrix controllable (DMC)** if, for each pair of unitarily equivalent density matrices  $\rho_1$  and  $\rho_2$ , there exists a control  $u$  and a time  $T > 0$ , such that the solution  $\rho(t)$  of (3.12) with initial condition equal to  $\rho_1$  at time  $T$  satisfies

$$\rho(T) = \rho_2.$$

In the rest of this chapter, we will show how to use the dynamical Lie algebra to check whether or not the system under consideration possesses any of the controllability properties above defined.

### 3.4 Pure State Controllability

The solution of (3.11) with initial condition  $\vec{\psi}_0$  can be written in matrix form as

$$\vec{\psi}(t) = X(t)\vec{\psi}_0,$$

where  $X(t)$  is the matrix solution of the Schrödinger operator equation.  $X(t)$  attains the values in the Lie group  $e^{\mathcal{L}}$ , where  $\mathcal{L}$  is the dynamical Lie algebra

---

<sup>4</sup>Two Hermitian matrices  $A, B$  are said to be unitarily equivalent if there exists a matrix  $C \in U(n)$  such that  $CAC^\dagger = B$ . This is equivalent to the two matrices having the same spectrum.

associated to the system. Therefore, all the states that can be obtained as a solution of (3.11) with initial condition equal to  $\vec{\psi}_0$  are the states of the set

$$\mathcal{O}_{\psi_0} := \{X\vec{\psi}_0 | X \in e^{\mathcal{L}}\}.$$

The system is pure state controllable if for every pair of complex vectors with radius equal to 1,  $\{\vec{\psi}_0, \vec{\psi}_1\}$ , that is for every pair of elements of the complex sphere  $S_{\mathbb{C}}^{n-1}$  of radius 1 in  $\mathbb{C}^n$ , there exists a matrix  $X$  in  $e^{\mathcal{L}}$ , such that

$$\vec{\psi}_1 = X\vec{\psi}_0.$$

In view of this observation, the question of pure state controllability reduces to an analysis of the properties of  $e^{\mathcal{L}}$  acting as a **Lie transformation group** on  $S_{\mathbb{C}}^{n-1}$ .  $e^{\mathcal{L}}$  will be called **transitive** if the above property holds. In the next two subsections, we give the main definitions concerning Lie transformation groups.

### 3.4.1 Lie transformation groups

Given a Lie group  $G$  and a manifold  $M$ , an *action* of  $G$  on  $M$  is defined as an analytic map

$$\theta : G \times M \rightarrow M,$$

satisfying the conditions

1.

$$\theta(I, x) = x$$

for every  $x \in M$ , where  $I$  is the identity element of the group;

2.

$$\theta(g_2, \theta(g_1, x)) = \theta(g_2g_1, x). \quad (3.13)$$

It is more convenient to use the notation  $gx$  instead of  $\theta(g, x)$ , so that property (3.13) reads

$$g_2(g_1x) = (g_2g_1)x.$$

If an action exists,  $G$  is called a Lie transformation group on  $M$ . It is called **transitive** if, for every pair of points in  $M$ ,  $x_1, x_2$ , there exists a  $g$  in  $G$  such that

$$x_2 = gx_1.$$

If  $G$  is transitive on  $M$ ,  $M$  is called a **homogeneous space** of  $G$ . The terminology derives from the fact that if  $P$  is a *property* on  $M$  that is invariant for any element  $g$  of  $G$ , namely a function

$$P : M \rightarrow \mathbf{R}, \quad Pg = P, \quad \forall g \in G,$$

then  $P$  is constant over  $M$ . In fact, if for every two points  $x_1$  and  $x_2$  we have  $x_2 = gx_1$  for some  $g \in G$ , then  $P(x_2) = P(gx_1) = P(x_1)$ . We shall see in subsection 3.4.3 that the special unitary group  $SU(n)$ , with action given by the standard matrix-vector multiplication, is transitive on the sphere. The norm of a vector  $v$  is a property which is invariant for any element  $X \in SU(n)$ , since  $\|Xv\| = \|v\|$ , and it is constant on the sphere.

### 3.4.2 Coset spaces and homogeneous spaces

Consider a Lie group  $G$  and a closed Lie subgroup  $H$ , obtained as the connected Lie subgroup corresponding to a Lie subalgebra of the Lie algebra associated with  $G$ . Given  $g \in G$ , all the elements of  $G$  that can be written as  $gX_h$ , with  $X_h \in H$ , form an equivalence class in  $G$  which is called a **left coset** and is denoted by  $\{gH\}$ . Analogously one defines a **right coset** as the equivalence class of all the elements in  $G$  that can be written as  $X_hg$  with  $X_h \in H$ . The right coset containing  $g$  is denoted by  $\{Hg\}$ . In the following definitions we shall refer to left cosets but only natural modifications are needed for right cosets. The **coset space**  $G/H$  is the set of left cosets. There is a natural mapping  $\pi$ ,

$$\pi : G \rightarrow G/H, \quad (3.14)$$

which maps  $g \in G$  to  $\{gH\}$  and a topology can be given to  $G/H$  from the requirement that  $\pi$  is continuous and open. In fact,  $G/H$  can be given the structure of an analytic manifold.

To see this, consider, for simplicity, the case which is of interest to us of matrix Lie groups  $G$  and  $H$ , with  $H$  a Lie subgroup of  $G$  and let  $\mathcal{H}$  and  $\mathcal{G}$ , with  $\mathcal{H} \subseteq \mathcal{G}$ , be the Lie algebras of  $H$  and  $G$ , of dimension  $n_h$  and  $n_g$ , respectively. If  $\mathcal{M}$  is a subspace of  $\mathcal{G}$  complementary<sup>5</sup> to  $\mathcal{H}$ , with basis  $A_1, \dots, A_{n_g - n_h}$ , then the map

$$\phi : \pi(e^{A_1 t_1} e^{A_2 t_2} \cdots e^{A_{n_g - n_h} t_{n_g - n_h}} g) \rightarrow (t_1, \dots, t_{n_g - n_h})$$

homeomorphically maps a neighborhood of  $\{gH\} \in G/H$  onto a neighborhood of the origin in  $\mathbf{R}^{n_g - n_h}$  and provides the required chart. This gives  $G/H$  the structure of a differentiable manifold of dimension  $n_g - n_h$ .

There is a natural way of defining an action of  $G$  on  $G/H$  which associates to  $g \in G$  and  $\{bH\} \in G/H$  the coset  $\{gbH\}$ . With this action and the above analytic structure of  $G/H$ ,  $G$  acts as a *Lie transformation group* on  $G/H$ . It is clear that  $G$  is transitive and therefore  $G/H$  is a homogeneous space of  $G$ .

Coset spaces are canonical examples of homogeneous spaces in a sense that will now be explained. Consider a connected transitive Lie transformation group  $G$  on a connected manifold  $M$  which is therefore a homogeneous space

---

<sup>5</sup>i.e.,  $\mathcal{M} \oplus \mathcal{H} = \mathcal{G}$ .

of  $G$ . Select a point  $p$  in  $M$ . The group of elements of  $G$  that leaves  $p$  fixed is called the **isotropy group** or the **stabilizer** of  $p$  and is denoted by  $G_p$ . It is a closed Lie subgroup of  $G$ . The correspondence

$$\{gG_p\} \leftrightarrow gp \quad (3.15)$$

for  $g \in G$  is a diffeomorphism and therefore  $M$  and  $G/G_p$  are diffeomorphic. In particular, they have the same dimension.

In the following we will be interested in matrix Lie subgroups of the unitary group acting on the complex sphere  $S_{\mathbb{C}}^{n-1}$  via the standard matrix vector multiplication. To show that a Lie group  $e^{\mathcal{L}}$  is transitive, it is enough to fix a  $\vec{\psi}_0 \in S_{\mathbb{C}}^{n-1}$  and show that, for every  $\vec{\psi}_1 \in S_{\mathbb{C}}^{n-1}$ , there exists an  $X \in e^{\mathcal{L}}$  such that

$$\vec{\psi}_1 = X\vec{\psi}_0.$$

### 3.4.3 The special unitary group and its action on the unit sphere

The **special unitary group** of dimension  $n$ ,  $SU(n)$ , is the Lie group of unitary  $n \times n$  matrices with determinant equal to one. It is the connected Lie group associated to the Lie algebra  $su(n)$  of skew-Hermitian  $n \times n$  matrices, with trace equal to zero. Its dimension as a manifold, which is equal to the dimension of its Lie algebra as a vector space, is  $n^2 - 1$ .

$SU(n)$  is transitive on the sphere  $S_{\mathbb{C}}^{n-1}$ . In order to see this, consider the point  $\vec{\psi}_0 := [1, 0, \dots, 0]^T$ . Every matrix  $X$  in  $SU(n)$  whose first column is equal to  $\vec{\psi}_f$  is such that

$$\vec{\psi}_f = X\vec{\psi}_0. \quad (3.16)$$

We can always construct such a matrix by choosing the remaining columns so as to form an orthonormal set with  $\vec{\psi}_f$ . This gives a unitary matrix whose determinant will in general be equal to  $e^{i\phi}$  for some  $\phi \in \mathbf{R}$ . By multiplying one of the columns 2 through  $n$  by  $e^{-i\phi}$ , we still have a unitary matrix  $X$  with the property (3.16) and with determinant equal to one.

The isotropy group of  $\vec{\psi}_0 = [1, 0, \dots, 0]^T$  is the Lie group of matrices

$$\tilde{X} = \begin{pmatrix} 1 & 0 \\ 0 & X \end{pmatrix},$$

with  $X \in SU(n-1)$ , clearly isomorphic to  $SU(n-1)$ . The diffeomorphism (3.15) between  $SU(n)/SU(n-1)$  and the sphere  $S_{\mathbb{C}}^{n-1}$  maps a representative  $g$  of  $\{gSU(n-1)\}$  to  $g\vec{\psi}_0$ . It is easily seen that this map does not depend on the representative chosen. The inverse map associates to  $\vec{\psi}_f$  the equivalence class of elements of  $SU(n)$  which map  $\vec{\psi}_0$  to  $\vec{\psi}_f$ . The dimension of  $SU(n)/SU(n-1)$

is given by  $n^2 - 1 - ((n-1)^2 - 1) = 2n - 1$  which is indeed the dimension of  $S_{\mathbb{C}}^{n-1}$ .

### 3.4.4 The symplectic group and its action on the unit sphere

**Definition 3.4.1** The **symplectic group**  $Sp(k)$  is defined as the Lie subgroup of  $SU(2k)$  of matrices  $X$  satisfying

$$XJX^T = J, \quad (3.17)$$

where the  $2k \times 2k$  matrix  $J$  is defined as

$$J = \begin{pmatrix} 0 & -\mathbf{1}_{k \times k} \\ \mathbf{1}_{k \times k} & 0 \end{pmatrix}, \quad (3.18)$$

with  $\mathbf{1}_{k \times k}$  the  $k \times k$  identity matrix.

$Sp(k)$  is the group of matrices in  $SU(2k)$  leaving invariant the bilinear form

$$(y, z)_J := y^T J z$$

on  $\mathbb{C}^{2k}$ , since, for every  $X \in Sp(k)$ , (3.17) implies  $(Xy, Xz)_J = (y, z)_J$ .

#### 3.4.4.1 Structure of the matrices in $Sp(k)$

From equation (3.17), since the conjugate of  $X$ ,  $\bar{X}$ , is the inverse of  $X^T$ , we have

$$XJ = J\bar{X}. \quad (3.19)$$

If we write  $X$  by partitioning its columns into sub-columns of dimension  $k$ ,

$$X := \begin{pmatrix} x_1 & x_2 & \dots & x_k & x_{k+1} & x_{k+2} & \dots & x_{2k} \\ y_1 & y_2 & \dots & y_k & y_{k+1} & y_{k+2} & \dots & y_{2k} \end{pmatrix},$$

with  $x_l, y_l \in \mathbb{C}^k$ ,  $l = 1, \dots, 2k$ , then equation (3.19) implies

$$x_{k+j} = -\bar{y}_j, \quad y_{k+j} = \bar{x}_j, \quad j = 1, \dots, k. \quad (3.20)$$

Therefore symplectic matrices are special unitary matrices (cf. Exercise 3.6) of the form

$$X = \begin{pmatrix} A & -\bar{B} \\ B & \bar{A} \end{pmatrix}. \quad (3.21)$$

#### 3.4.4.2 The symplectic Lie algebra

The Lie algebra of the symplectic Lie group  $Sp(k)$  is denoted by  $sp(k)$  and it is called the **symplectic Lie algebra**. Its elements have to satisfy a

condition obtained from (3.17). By letting  $X$  depend on  $t$ , with  $X(0) = \mathbf{1}_{n \times n}$ , and differentiating (3.17), we obtain

$$\dot{X}(0)J + J\dot{X}^T(0) = 0.$$

Therefore the Lie algebra of *skew-Hermitian* matrices  $A$  satisfying

$$AJ + JA^T = 0 \quad (3.22)$$

is the Lie algebra of  $Sp(k)$ . It is clear that this is a Lie algebra since if  $A$  and  $B$  satisfy (3.22) so does  $[A, B]$ .

If we partition  $A$  in  $k \times k$  blocks as

$$A := \begin{pmatrix} L_1 & L_2 \\ -L_2^\dagger & L_3 \end{pmatrix},$$

with

$$L_1 = -L_1^\dagger, \quad L_3 = -L_3^\dagger \quad (3.23)$$

then conditions (3.22), (3.23) imply

$$L_3 = \bar{L}_1, \quad L_2 = L_2^T;$$

therefore any matrix  $A$  in  $sp(k)$  has the form

$$A := \begin{pmatrix} L_1 & L_2 \\ -\bar{L}_2 & \bar{L}_1 \end{pmatrix},$$

with  $L_1$  a skew-Hermitian matrix but otherwise arbitrary, and  $L_2 = L_2^T$ . In  $L_1$  we have  $k^2$  free real parameters while, for  $L_2$ , we can choose  $\frac{k(k+1)}{2}$  complex numbers and therefore  $k(k+1)$  real parameters since  $L_2$  is symmetric. The total number of free parameters that is the *dimension* of  $sp(k)$  is

$$\dim sp(k) = k^2 + k(k+1) = k(2k+1).$$

Therefore the dimension of  $Sp(k)$  is  $k(2k+1)$  as well.

### 3.4.4.3 Transitivity of the symplectic group

An important property, for our purposes, of the symplectic group  $Sp(k)$  is that it is transitive on the complex unit sphere in the  $2k$ -dimensional complex space,  $S_{\mathbb{C}}^{2k-1}$ .

Consider the point  $\vec{\psi}_0 := [1, 0, \dots, 0]^T$  in  $S_{\mathbb{C}}^{2k-1}$  and let  $\vec{\psi}_f : [v_1^T v_2^T]^T$  be the target point in  $S_{\mathbb{C}}^{2k-1}$ , where  $v_1$  and  $v_2$  are  $k$ -dimensional complex vectors. To construct a matrix  $X \in Sp(k)$  such that  $\vec{\psi}_f = X\vec{\psi}_0$ , we choose the first column of  $X$ ,  $[x_1^T y_1^T]^T$ , equal to  $\vec{\psi}_f := [v_1^T v_2^T]^T$ . The  $(k+1)$ -th column

is chosen according to (3.20). The second column is chosen as a unit vector which is orthogonal to the first and  $(k+1)$ -th columns. The  $(k+2)$ -th column is chosen according to (3.20). Proceeding this way for the following columns, one obtains the desired matrix.

The isotropy group in  $Sp(k)$  for  $\vec{\psi}_0 = [1, 0, \dots, 0]$  is given by the special unitary matrices of the form

$$X = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & A & 0 & -\bar{B} \\ 0 & 0 & 1 & 0 \\ 0 & B & 0 & \bar{A} \end{pmatrix},$$

where  $A$  and  $B$  are  $(k-1) \times (k-1)$  blocks and

$$\begin{pmatrix} A & -\bar{B} \\ B & \bar{A} \end{pmatrix} \in Sp(k-1).$$

The coset space  $Sp(k)/Sp(k-1)$  has dimension  $k(2k+1)-(k-1)(2(k-1)+1)=4k-1$  which is in fact the real dimension of the sphere  $S^{2k-1}_{\mathbb{C}}$ .

#### 3.4.4.4 Properties of the symplectic Lie algebra

The symplectic Lie algebra  $sp(k)$  is one of the most studied Lie algebras, being one of the so called *classical Lie algebras* (along with  $su(n)$  and  $so(n)$ , the Lie algebra of skew-symmetric matrices of dimension  $n$ ). It has several properties. In the following three propositions, we list here three more properties which will be used in the following treatment. Proofs can be found in [5], [40].

**Proposition 3.4.2**  $sp(\frac{n}{2})$  is a **maximal subalgebra** of  $su(n)$ , i.e., there is no subalgebra of  $su(n)$  properly containing  $sp(\frac{n}{2})$  other than  $su(n)$  itself. Equivalently, the Lie algebra generated by  $sp(\frac{n}{2})$  and any other element  $F$  of  $su(n)$ ,  $F \notin sp(\frac{n}{2})$ , is equal to  $su(n)$ .

Another example of a maximal subalgebra of  $su(n)$  is  $so(n)$ , namely the Lie algebra spanned by the real matrices in  $su(n)$ .

The main consequence for us of the property in Proposition 3.4.2 is that there is no connected proper Lie subgroup of  $SU(n)$  transitive on the sphere that contains  $Sp(\frac{n}{2})$  properly. If this were the case, the corresponding Lie algebra would contain  $sp(\frac{n}{2})$  properly (see [Exercise 3.4](#)), which is not possible. Therefore, if we are trying to obtain transitive Lie groups as Lie groups which contain  $Sp(\frac{n}{2})$  properly we can find only  $SU(n)$ .

In [subsection 3.1.1.3](#) we defined isomorphisms between two Lie algebras. If we have a matrix Lie group  $e^{\mathcal{L}}$  corresponding to a Lie algebra  $\mathcal{L}$ , it is possible

to define an isomorphism  $\phi_U$  between two subalgebras of  $\mathcal{L}$ ,  $\mathcal{L}_1$  and  $\mathcal{L}_2$ , i.e.,  $\phi_U : \mathcal{L}_1 \rightarrow \mathcal{L}_2$ , as follows

$$\phi_U(A) := UAU^{-1}, \quad \forall A \in \mathcal{L}_1,$$

where  $U$  is an element of  $e^{\mathcal{L}}$ . If this is the case,  $\mathcal{L}_1$  and  $\mathcal{L}_2$  are said to be **conjugate in  $\mathcal{L}$**  and we write  $\mathcal{L}_2 = U\mathcal{L}_1U^{-1}$ . For the corresponding Lie groups we have

$$Ue^{\mathcal{L}_1}U^{-1} = e^{\mathcal{L}_2},$$

and the map  $X \rightarrow UXU^{-1}$  is an isomorphism of the two Lie groups  $e^{\mathcal{L}_1}$  and  $e^{\mathcal{L}_2}$ .

**Proposition 3.4.3** Every subalgebra of  $su(n)$  isomorphic to  $sp(\frac{n}{2})$  is conjugate to  $sp(\frac{n}{2})$  in  $su(n)$ .

Since we know that  $sp(\frac{n}{2})$  gives rise to a transitive Lie group, it is natural to investigate what happens for subalgebras of  $su(n)$  which are isomorphic to  $sp(\frac{n}{2})$ . The property in Proposition 3.4.3 says that we only have to investigate Lie algebras *conjugate* to  $sp(\frac{n}{2})$ . It is clear that these subalgebras,  $\mathcal{L}$ , give rise to transitive Lie groups as well. In fact if  $e^{\mathcal{L}} = USp(\frac{n}{2})U^\dagger$ , we can find  $X \in e^{\mathcal{L}}$  such that

$$\vec{\psi}_f = X\vec{\psi}_0,$$

for any  $\vec{\psi}_0$  and  $\vec{\psi}_f$ . We can choose  $X$  as  $X = U\tilde{X}U^\dagger$  where  $\tilde{X}$  is a matrix in  $Sp(\frac{n}{2})$  such that

$$U^\dagger\vec{\psi}_f = \tilde{X}U^\dagger\vec{\psi}_0.$$

Another property of the Lie algebra  $sp(\frac{n}{2})$  which will be used in the following is given in the Proposition 3.4.6 below. We need first the following definitions.

**Definition 3.4.4** An **ideal**  $\mathcal{I}$  of a Lie algebra  $\mathcal{L}$  is a subspace of  $\mathcal{L}$  such that

$$[\mathcal{I}, \mathcal{L}] \subseteq \mathcal{I}.$$

It is an immediate consequence of the definition that an ideal of a Lie algebra  $\mathcal{L}$  is a Lie subalgebra of  $\mathcal{L}$ .

**Definition 3.4.5** A Lie algebra  $\mathcal{L}$  is **simple** if it has no ideals other than the trivial ones:  $\{0\}$  and  $\mathcal{L}$  itself. A **semisimple** Lie algebra  $\mathcal{L}$  is the direct Lie algebra sum  $\oplus$  of simple Lie algebras  $\mathcal{L}_j$ ,  $j = 1, \dots, r$ , i.e.,

$$\mathcal{L} = \mathcal{L}_1 \bar{\oplus} \mathcal{L}_2 \bar{\oplus} \dots \bar{\oplus} \mathcal{L}_r,$$

where  $\mathcal{L}_j \bar{\oplus} \mathcal{L}_k$  means  $\mathcal{L}_j \oplus \mathcal{L}_k$  with  $[\mathcal{L}_j, \mathcal{L}_k] = 0$ .

Whether or not a Lie algebra is semisimple can be checked using *Cartan semisimplicity criterion* (cf. [Appendix C](#)).

**Proposition 3.4.6** The Lie algebras  $su(n)$  and  $sp(\frac{n}{2})$  are simple. The Lie algebra  $so(n)$  is simple for  $n \neq 4$ , and semisimple for  $n = 4$ .<sup>6</sup>

### 3.4.5 Test for pure state controllability

From what we have seen so far, if the dynamical Lie algebra  $\mathcal{L}$  is either  $su(n)$  or a Lie algebra conjugate to  $sp(\frac{n}{2})$ , then the quantum control system is PSC. The following theorem which was proved in [5] says that these are essentially the only cases.

**Theorem 3.4.7** *The quantum system is PSC if and only if the corresponding dynamical Lie algebra  $\mathcal{L}$  satisfies one of the following*

1.

$$\mathcal{L} = su(n).$$

2.  $\mathcal{L}$  is conjugate to  $sp(\frac{n}{2})$ .

3.

$$\mathcal{L} = u(n).$$

4.

$$\mathcal{L} = \text{span}\{i\mathbf{1}_{n \times n}\} \oplus \tilde{\mathcal{L}},$$

where  $\tilde{\mathcal{L}}$  is a Lie algebra conjugate to  $sp(\frac{n}{2})$ .

The case 1 and 3 above can be checked by checking the dimension of  $\mathcal{L}$ . If the dimension is  $n^2 - 1$  or  $n^2$ , then we have case 1 and 3, respectively. To have cases 2 and 4,  $n$  must be even and the dimension of  $\mathcal{L}$  must be equal to  $\frac{n}{2}(n+1)$  or  $\frac{n}{2}(n+1) + 1$ , respectively. However this is not enough to conclude that we are in case 2 or 4. Let us consider case 2 first. We need to check that  $\mathcal{L}$  is conjugate to  $sp(\frac{n}{2})$  and therefore that there exists a unitary matrix  $U$  such that

$$UAU^\dagger J + J(UAU^\dagger)^T = 0, \quad (3.24)$$

for every  $A \in \mathcal{L}$ . In fact it is enough to check (3.24) for a set of generators  $\{A_1, \dots, A_s\}$  of  $\mathcal{L}$ , namely

$$UA_k U^\dagger J + J(UA_k U^\dagger)^T = 0, \quad k = 1, \dots, s,$$

which, defining

$$\tilde{J} = U^\dagger J \bar{U}, \quad (3.25)$$

---

<sup>6</sup>In this case it is the direct sum of two copies of  $so(3)$ .

means

$$A_k \tilde{J} + \tilde{J} A_k^T = 0, \quad k = 1, \dots, s. \quad (3.26)$$

System (3.26) is a linear system of equations in the unknown  $\tilde{J}$ . If there exists a solution which is related to  $J$  as in (3.25), then  $\mathcal{L}$  is conjugate to  $sp(\frac{n}{2})$ . This method was discussed in [190]. Another method to check the given isomorphism (conjugacy) is to use the structure theory of Lie algebras (see, e.g., [96], [106]). Another, much more direct, method will be described in [section 3.6](#) below (cf. Theorem 3.6.2) after we develop some more theory. To check possibility 4, one writes  $\mathcal{L}$  as

$$\mathcal{L} = \{\text{span}\{i\mathbf{1}_{n \times n}\}\} \oplus (\text{span}\{i\mathbf{1}_{n \times n}\})^\perp,$$

and check the above conjugacy for  $(\text{span}\{i\mathbf{1}_{n \times n}\})^\perp$ .

---

### 3.5 Equivalent State Controllability

From a physics point of view, having equivalent state controllability (ESC) is equivalent to having pure state controllability. We expect that the mathematical conditions to have ESC would be the same as the ones for PSC. This is indeed the case as we shall see now.

An equivalent definition of ESC which involves the dynamical Lie algebra  $\mathcal{L}$  is as follows.

**Definition 3.5.1** The system is ESC if and only if, for any pair of elements on the complex sphere,  $\vec{\psi}_0$  and  $\vec{\psi}_1$ , there exists an  $X \in e^{\mathcal{L}}$  and  $\phi \in \mathbf{R}$  such that

$$\vec{\psi}_1 = e^{-i\phi} X \vec{\psi}_0.$$

It is clear that PSC implies ESC. Conversely, assume that ESC is verified. Consider the (augmented) Lie algebra

$$\bar{\mathcal{L}} := \text{span}\{i\mathbf{1}_{n \times n}\} + \mathcal{L}.$$

The Lie group  $e^{\bar{\mathcal{L}}}$  clearly contains all the elements  $e^{-i\phi} X$ , with  $\phi \in \mathbf{R}$  and  $X \in e^{\mathcal{L}}$ . Therefore  $e^{\bar{\mathcal{L}}}$  is transitive on the complex sphere. Since  $\bar{\mathcal{L}}$  contains multiples of the identity, according to Theorem 3.4.7, it has to be equal to  $u(n)$  or  $\text{span}\{i\mathbf{1}_{n \times n}\} \oplus \bar{\mathcal{L}}$ , with  $\bar{\mathcal{L}}$  conjugate to  $sp(\frac{n}{2})$ . This implies that

$$[\bar{\mathcal{L}}, \bar{\mathcal{L}}] = [\mathcal{L}, \mathcal{L}] = [h, h] \subseteq \mathcal{L}, \quad (3.27)$$

where  $h$  stands for  $su(n)$  or  $\tilde{\mathcal{L}}$  (conjugate to  $sp(\frac{n}{2})$ ). However, since both these Lie algebras are simple  $[h, h] = h$  (otherwise  $[h, h]$  would be an ideal). Therefore (3.27) implies

$$su(n) \subseteq \mathcal{L},$$

or

$$sp\left(\frac{n}{2}\right) \subseteq \mathcal{L},$$

which clearly implies PSC.

We summarize this discussion in the following proposition.

**Proposition 3.5.2** Equivalent state controllability (ESC) and pure state controllability (PSC) are equivalent properties.

---

### 3.6 Equality of Orbits

Let  $\mathcal{L}$  be the dynamical Lie algebra associated with the quantum control system at hand. If  $\rho_0$  is the initial condition of an ensemble, after the evolution  $X \in e^{\mathcal{L}}$ , the state of the ensemble is equal to  $X\rho_0X^\dagger$ . Therefore the set of states reachable from  $\rho_0$  is the **orbit**

$$\mathcal{O}_{\mathcal{L}}(\rho_0) := \{X\rho_0X^\dagger | X \in e^{\mathcal{L}}\}. \quad (3.28)$$

The question we shall deal with in this section is to find conditions so that this set is equal to the largest possible set, namely the orbit<sup>7</sup>

$$\mathcal{O}_{U(n)}(\rho_0) := \{X\rho_0X^\dagger | X \in U(n)\}. \quad (3.29)$$

Notice this depends in a critical way on the initial density matrix  $\rho_0$ . If  $\rho_0 = \frac{1}{n}\mathbf{1}_{n \times n}$ , a perfectly mixed ensemble of systems, then  $\mathcal{O}_{\mathcal{L}} = \mathcal{O}_{U(n)} = \{\rho_0\}$ , no matter what  $\mathcal{L}$  is. In the special case where  $\rho_0$  is a pure state, namely a matrix of rank equal to one, the equality  $\mathcal{O}_{\mathcal{L}} = \mathcal{O}_{U(n)}$  is equivalent to ESC, and therefore to PSC, and it is satisfied only in the cases described in Theorem 3.4.7.

In the general case, we can proceed as follows. The isotropy group of  $\rho_0$  in  $U(n)$  is a closed Lie subgroup of  $U(n)$  whose Lie algebra is the space of matrices in  $u(n)$  commuting with  $i\rho_0$ . This Lie algebra is called the *centralizer* of  $i\rho_0$  in  $u(n)$  and it is denoted here by  $\mathcal{C}_{\rho_0}$ . Therefore the isotropy group is denoted here by  $e^{\mathcal{C}_{\rho_0}}$ . The isotropy group of  $\rho_0$  in  $e^{\mathcal{L}}$  is a closed Lie subgroup of  $e^{\mathcal{L}}$  whose Lie algebra is  $\mathcal{C}_{\rho_0} \cap \mathcal{L}$ . This isotropy group is denoted here by  $e^{\mathcal{C}_{\rho_0} \cap \mathcal{L}}$ . Now, it follows from the general facts on coset spaces and homogeneous spaces discussed in subsection 3.4.2 (applied to the manifolds  $\mathcal{O}_{\mathcal{L}}$  and  $\mathcal{O}_{U(n)}$ ) that

<sup>7</sup>In Chapter 8 section 8.3, in the context of entanglement theory, we shall consider the question of testing whether given a Lie algebra  $\mathcal{L}$  and two density matrices  $\rho_1$  and  $\rho_2$ ,  $\mathcal{O}_{\mathcal{L}}(\rho_1) = \mathcal{O}_{\mathcal{L}}(\rho_2)$ . In entanglement theory, the Lie algebra of interest  $\mathcal{L}$  is the one of local Hamiltonians of a multipartite system (cf. Chapter 8).

$\mathcal{O}_{\mathcal{L}}$  and  $\mathcal{O}_{u(n)}$  are diffeomorphic to  $e^{\mathcal{L}}/e^{\mathcal{C}_{\rho_0} \cap \mathcal{L}}$  and  $U(n)/e^{\mathcal{C}_{\rho_0}}$  respectively. If they are equal

$$\dim \mathcal{O}_{u(n)} = \dim U(n)/e^{\mathcal{C}_{\rho_0}}$$

and

$$\dim \mathcal{O}_{\mathcal{L}} = \dim e^{\mathcal{L}}/e^{\mathcal{C}_{\rho_0} \cap \mathcal{L}}$$

are also equal. Therefore we must have

$$\dim u(n) - \dim \mathcal{C}_{\rho_0} = \dim \mathcal{L} - \dim(\mathcal{C}_{\rho_0} \cap \mathcal{L}). \quad (3.30)$$

This condition is also sufficient to have  $\mathcal{O}_{\mathcal{L}} = \mathcal{O}_{u(n)}$ . The proof of this fact is presented in [5]. We conclude with two theorems which give conditions for equality of orbits.

**Theorem 3.6.1** *The orbits (3.28) and (3.29) are equal if and only if condition (3.30) is verified.*

We can rewrite condition (3.30) in a different, more compact, form. Consider the linear map  $ad_{i\rho_0} : u(n) \rightarrow u(n)$ , which associates to an element  $A \in u(n)$  the element  $[i\rho_0, A]$ .  $\mathcal{C}_{\rho_0}$  is the kernel of this map and, by standard results of linear algebra, the number on the left hand side of (3.30) is the dimension of the range of  $ad_{i\rho_0}, [i\rho_0, u(n)]$ . Analogously, the number on the right hand side is the dimension of  $[i\rho_0, \mathcal{L}]$ . Since  $[i\rho_0, \mathcal{L}] \subseteq [i\rho_0, u(n)]$ , condition (3.30) can be rewritten as

$$[i\rho_0, \mathcal{L}] = [i\rho_0, u(n)],$$

or

$$\dim[i\rho_0, \mathcal{L}] = \dim[i\rho_0, u(n)]. \quad (3.31)$$

Condition (3.31) is always satisfied if  $i\rho_0$  is a multiple of the identity. If  $\rho_0$  is any matrix of rank one it gives an alternate and practical way to check pure state controllability and therefore equivalent state controllability and it will be satisfied only in the cases listed in Theorem 3.4.7. We state this test formally in the following Theorem.

**Theorem 3.6.2** *A quantum control system with dynamical Lie algebra  $\mathcal{L}$  is pure state controllable and equivalent state controllable if and only if condition (3.31) is satisfied for a rank 1 matrix  $\rho_0$ , for example  $\rho_0 = \text{diag}(1, 0, \dots, 0)$ .*

Test (3.31) also gives a practical way to check the isomorphism between  $\mathcal{L}$  and  $sp(\frac{n}{2})$ .

The dimension of  $[i\rho_0, u(n)]$  can be calculated in terms of the eigenvalues of  $\rho_0$ . It is zero if  $\rho_0$  has only one eigenvalue. In all the other cases it is equal to

$$\dim[i\rho_0, u(n)] = 2 \sum_{j < k} n_j n_k, \quad (3.32)$$

where  $n_{j,k}$  are the multiplicities of the eigenvalues of  $\rho_0$  (cf. Exercise 3.10). In the special case where  $\rho_0 = \text{diag}(1, 0, \dots, 0)$ , the number on the right hand side of (3.32) is  $2(n - 1)$ . Therefore we have this corollary to Theorem 3.6.2.

**Corollary 3.6.3** The quantum control system with dynamical Lie algebra  $\mathcal{L}$  is PSC if and only if

$$\dim[i\rho_0, \mathcal{L}] = 2(n - 1),$$

with  $\rho_0 = \text{diag}(1, 0, \dots, 0)$ .

**Example 3.6.4** (cf. [5]) Consider the dynamical Lie algebra  $\mathcal{L}$  spanned by skew Hermitian matrices of the form

$$F := \begin{pmatrix} L + Z & T + C \\ -\bar{T} + \bar{C} & -L + Z^T \end{pmatrix}.$$

All of the sub-matrices are  $2 \times 2$ , with  $L$  and  $T$  diagonal and  $Z$  having all entries on the main diagonal equal to zero, and  $C$  such that  $C^T = -C$ . The Lie algebra  $\mathcal{L}$  so defined is isomorphic to  $sp(2)$ . However instead of verifying this isomorphism we follow Corollary 3.6.3 and calculate  $[i\rho_0, \mathcal{L}]$  with  $\rho_0 := \text{diag}(1, 0, \dots, 0)$ . This is given by matrices of the form

$$G := \begin{pmatrix} 0 & a + ib & -c + id & -e + if \\ -a + ib & 0 & 0 & 0 \\ c + id & 0 & 0 & 0 \\ e + if & 0 & 0 & 0 \end{pmatrix},$$

with  $a, b, c, d, e$  and  $f$  free parameters. Therefore  $\dim[i\rho_0, \mathcal{L}] = 6$  which is equal to  $2(n - 1)$  since  $n = 4$ . Therefore this system is PSC.

Consider now the orbit associated to  $\rho_0 := \text{diag}(\frac{1}{2}, \frac{1}{2}, 0, 0)$  under the dynamical Lie algebra  $\mathcal{L}$ . To verify the conditions of Theorem 3.6.1, we use formula (3.32) and calculate

$$\dim[i\rho_0, u(n)] = 8.$$

However  $[i\rho_0, \mathcal{L}]$  is spanned by matrices of the type

$$G := \begin{pmatrix} 0 & 0 & a + ib & c + id \\ 0 & 0 & -c - id & e + if \\ -a + ib & c - id & 0 & 0 \\ -c + id & -e + if & 0 & 0 \end{pmatrix},$$

and it is therefore 6-dimensional. Therefore the two orbits are not the same.

### 3.6.1 Density matrix controllability

The definition of density matrix controllability is equivalent to the equality of the orbits  $\mathcal{O}_{\mathcal{L}}(\rho_0)$  and  $\mathcal{O}_{u(n)}(\rho_0)$ , for every  $\rho_0$ . This is a very strong

requirement and it is possible if and only if  $\mathcal{L}$  is equal to  $su(n)$  or  $u(n)$ . In fact since this has to be true in particular for  $\rho_0$  of rank one, the only other possibility would be  $n$  even and the dynamical Lie algebra  $\mathcal{L}$  conjugate to the symplectic Lie algebra. However, taking for example

$$\rho_0 = \begin{pmatrix} L & \mathbf{0}_{n \times n} \\ \mathbf{0}_{n \times n} & L \end{pmatrix},$$

with  $L = \text{diag}(\frac{1}{2}, 0, \dots, 0)$ , one sees that equality (3.31) is not verified. Therefore we have the following condition for density matrix controllability of Definition 3.3.3.

**Theorem 3.6.5** *A quantum control system with dynamical Lie algebra  $\mathcal{L}$  is density matrix controllable if and only if  $\mathcal{L} = su(n)$  or  $\mathcal{L} = u(n)$ , i.e., it is operator controllable.*

The diagram in [Figure 3.2](#) summarizes the relations among the various notions of controllability and the conditions under which they are verified.

---

## 3.7 Notes and References

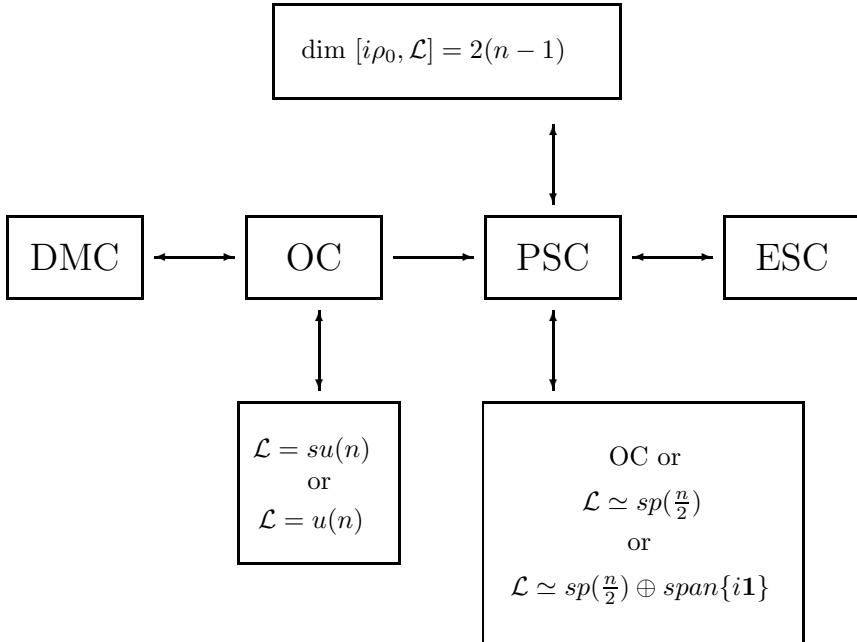
### 3.7.1 Alternate tests of controllability

There have been several attempts in the literature to give tests of controllability which avoid the direct computation of the dynamical Lie algebra. This is particularly valuable in high dimensional situations and when the physics of the problem itself can be used to check controllability. Most of the alternate methods are concerned with special classes of systems of interest in applications. Many results concern the use of conditions on a graph associated with the system to determine controllability (see, e.g., [4], [13]).

To give an example of the type of results in this area, we briefly review the main result of [207]. In this paper, a system of the type

$$\dot{\psi} = (A + Bu)\psi \tag{3.33}$$

is considered, with  $A$  diagonal and  $B$  real and therefore skew-symmetric. A graph is associated with the system. The  $n$  vertices of the graph represent the  $n$  eigenstates of  $A$ , which are the energy eigenstates when the control is equal to zero. An edge connects vertices  $j$  and  $k$  if the element  $j, k$  (and therefore the element  $k, j$ ) of the matrix  $B$  is different from zero. This means that there is *coupling* between the eigenstates corresponding to the  $j$ -th and  $k$ -th eigenvalue. If the remaining entries of  $B$  were zero (and, for simplicity,  $A$  were zero) the solution of (3.33) would be  $\psi(t) = X(t)\psi(0)$  where  $X(t)$  is a matrix



**FIGURE 3.2:** Relation among the various notions of controllability. DMC (density matrix controllability) is equivalent to OC (operator controllability) which implies PSC (pure state controllability) which is in turn equivalent to ESC (equivalent state controllability). Operator controllability is verified if and only if the dynamical Lie algebra  $\mathcal{L}$  is equal to  $su(n)$  or  $u(n)$ . PSC is verified if and only if the system is OC or  $\mathcal{L}$  is isomorphic ( $\simeq$ ) and therefore conjugate (cf. Proposition 3.4.3) to  $sp(\frac{n}{2})$  or  $sp(\frac{n}{2}) \oplus \mathbf{1}_{n \times n}$ . PSC is equivalent to the condition  $\dim[i\rho_0, \mathcal{L}] = 2(n - 1)$ , for  $\rho_0 := \text{diag}(1, 0, \dots, 0)$ .

equal to the identity everywhere except for the elements at the intersection of the  $j$ -th and  $k$ -th rows and columns which form the matrix (assuming the corresponding nonzero element of  $B$  is equal to one)

$$U_{j,k}(t) = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}.$$

Therefore the dynamics transfer magnitude between the  $j$ -th and  $k$ -th eigenstates. The above graph is called in [207] a *connectivity graph*. The edges connecting the vertices corresponding to the eigenvalues of  $A$ ,  $-i\lambda_j$  and  $-i\lambda_k$  are labeled by  $\omega_{jk} := |\lambda_j - \lambda_k|$ . Since the  $\lambda_j$ 's are the eigenvalues of the Hamiltonian without control, these are the energy differences for the various eigenstates. In Figure 3.3 we report, as an example, the connectivity graph of the pair

$$A := \begin{pmatrix} 3i & 0 & 0 & 0 \\ 0 & 2i & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & 0 & 0 & 10i \end{pmatrix}, \quad B := \begin{pmatrix} 0 & 1 & 1 & 0 \\ -1 & 0 & 1 & 1 \\ -1 & -1 & 0 & 2 \\ 0 & -1 & -2 & 0 \end{pmatrix}. \quad (3.34)$$

The main result of [207] is as follows.

**Theorem 3.7.1** *If the connectivity graph remains connected even after removing the edges which have equal labels, then system (3.33) is pure state controllable.*

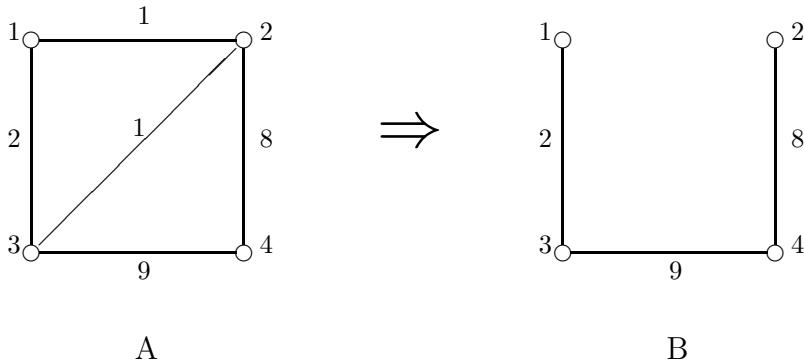
### 3.7.2 Pure state controllability and existence of constants of motion

It is intuitive that the existence of a constant of motion for the quantum control system (3.11) would prevent pure state controllability. A **constant of motion** is an observable that gives the same value with certainty when measured along a trajectory. In finite dimensional quantum mechanics, we associate Hermitian matrices with observables. To a constant of motion there corresponds a matrix  $C$  which commutes with  $H(u)$  in (3.11) for every value of  $u$ , and therefore with the whole dynamical Lie algebra  $\mathcal{L}$ , i.e.,

$$[C, \mathcal{L}] = 0. \quad (3.35)$$

In order to explain this correspondence first assume that (3.35) is true. If we measure  $C$ , we obtain the eigenvalue  $\lambda$ . According to the measurement postulate (see subsection 1.2.2) the state will collapse into an eigenstate  $\vec{\psi}_0$  of  $C$ , i.e.,

$$C\vec{\psi}_0 = \lambda\vec{\psi}_0. \quad (3.36)$$



**FIGURE 3.3:** The connectivity graph for the pair  $\{A, B\}$  in (3.34), Part A, and the graph obtained by removing equal labels, Part B. According to Theorem 3.7.1, the system is pure state controllable in this case.

After an evolution  $X \in e^{\mathcal{L}}$ , the state will be  $X\vec{\psi}_0$  which is again an eigenvector of  $C$  with eigenvalue  $\lambda$  since, from (3.36) and the fact that  $CX = XC$ , we obtain

$$CX\vec{\psi}_0 = XC\vec{\psi}_0 = \lambda X\vec{\psi}_0. \quad (3.37)$$

Therefore a measurement of  $C$  will give the value  $\lambda$  again, with certainty. This shows that if (3.35) is true,  $C$  is a constant of motion. Conversely, if  $C$  is a constant of motion for every eigenvector  $\vec{\psi}_0$  and eigenvalue  $\lambda$  we have  $C\vec{\psi}_0 = \lambda\vec{\psi}_0$  and  $CX\vec{\psi}_0 = \lambda X\vec{\psi}_0$ , for every  $X \in e^{\mathcal{L}}$ . This gives, for every eigenvector  $\vec{\psi}_0$  and for every  $X \in e^{\mathcal{L}}$ ,

$$C\vec{\psi}_0 = X^\dagger CX\vec{\psi}_0.$$

Being this true for every eigenvector  $\vec{\psi}_0$ , we have  $C = X^\dagger CX$ , for every  $X \in e^{\mathcal{L}}$  which, in turn, implies (3.35).

We now wish to investigate the relation between existence of a nontrivial constant of motion, namely one that is not a multiple of the identity, and the pure state controllability of the system. We want to give a mathematical explanation of the intuition that the existence of such a constant of motion prevents controllability. Assume there exists a constant of motion  $C$  and write  $iC$  as  $i\frac{\text{Tr}(C)}{n}\mathbf{1} + \tilde{C}$ , with  $\tilde{C} \in su(n)$ . Then, if the system is PSC because  $\mathcal{L} = u(n)$  or  $\mathcal{L} = su(n)$ , we have

$$[\tilde{C}, su(n)] = 0.$$

However, this is possible only if  $\tilde{C} = 0$  because otherwise  $\text{span}\{\tilde{C}\}$  would be a nontrivial ideal of  $su(n)$  and  $su(n)$ , being simple, has no nontrivial ideals. If the system is PSC because  $\mathcal{L}$  is (conjugate to)  $\text{span}\{i\mathbf{1}_{n \times n}\} \oplus sp(\frac{n}{2})$ , then we must have (up to a coniugacy of  $\tilde{C}$ ),

$$[\tilde{C}, sp(\frac{n}{2})] = 0,$$

but this is not possible. If  $\tilde{C} \in sp(\frac{n}{2})$  it would contradict the fact that  $sp(\frac{n}{2})$  is simple. If  $\tilde{C} \notin sp(\frac{n}{2})$  it would contradict the fact that  $sp(\frac{n}{2})$  is a maximal subalgebra in  $su(n)$  (cf. Proposition 3.4.2). Therefore we can conclude with the following proposition.

**Proposition 3.7.2** If a nontrivial constant of motion exists then the system is not PSC.

The converse of Proposition 3.7.2 is in general not true, namely it is possible that the system is not PSC and there is no nontrivial constant of motion. For example, if the dynamical Lie algebra  $\mathcal{L}$  is equal to  $so(n)$ , then the system is not PSC. However, since  $so(n)$ , for  $n \neq 4$  is also a simple maximal subalgebra in  $su(n)$ , reasoning as above for  $sp(\frac{n}{2})$ , we conclude that there is no nontrivial constant of motion.

### 3.7.3 Bibliographical notes

Much of the treatment presented here is based on [5], in particular for what concerns controllability of the state  $|\psi\rangle$  or  $\rho$ . This work was based on a previous paper [148] on the transitive action of transformation groups on spheres. A related work is [190]. The books [96], [116], [149], [217] are very good references for the basic material concerning Lie transformation groups. Various concepts on Lie algebras introduced can be found in some of the standard books on Lie algebras and Lie groups such as [94], [96] and [106]. The result on maximal subalgebras of a Lie algebra in Proposition 3.4.2 follows from the work of Dynkin [73] and was summarized in [40]. The discussion on the relation between constants of motion and controllability in the previous subsection 3.7.2 was motivated by a conjecture in [207] and was elaborated by the author in collaboration with Francesca Albertini.

### 3.7.4 Some open problems

There are several open problems concerning the study of controllability for quantum systems. In particular, the tests described in this chapter do not allow us in general to decide whether given two density matrices it is possible to go from one to the other in the case where the system is not controllable. In other terms, a general test is needed to decide whether a density matrix is in an orbit of the form (3.28). An instance of this problem in the context

of entanglement theory is described in [section 8.3](#). Another area of interest is the characterization of the reachable sets. This problem is related to the minimum time optimal control in Chapter 6 (see in particular [subsection 6.4.2](#)). The extension of the Lie algebraic approach presented here to infinite dimensional systems presents many interesting mathematical challenges (see, e.g., [24] and references therein). A problem of practical interest is the study of controllability for classes of control functions which do not include piecewise constant functions. This is motivated by the fact that in practice a jump in the control corresponds to very high frequency Fourier components. High frequency components correspond to large differences between the energy levels of the system (cf. [section 7.1](#)) and controls with high frequency may induce population transfer between levels that have been ignored in the modeling procedure.

---

## 3.8 Exercises

### Exercise 3.1

The Lie algebra  $\{x_1, \dots, x_n\}_{\mathcal{L}}$  contains not only the repeated Lie brackets of  $\{x_1, \dots, x_n\}$  but also their Lie brackets and Lie brackets of their Lie brackets. As an example it contains  $[[x_1, x_2], [x_3, x_4]]$  which is *not* a repeated Lie bracket as defined in [subsection 3.2.1](#). However these can be expressed as linear combinations of repeated Lie brackets of  $\{x_1, \dots, x_n\}$  and this justifies the algorithm presented in subsection 3.2.1. This exercise concerns the proof of this fact.

It is enough to show that given a Lie bracket  $[P, Q]$  where  $P$  is a repeated Lie bracket of the elements  $\{x_1, \dots, x_n\}$  then  $[P, Q]$  is a linear combination of elements of the type <sup>8</sup>

$$[x_{j_1}, [x_{j_2}, \dots [x_{j_l}, Q] \dots]] := ad_{x_{j_1}} ad_{x_{j_2}} \cdots ad_{x_{j_l}} Q, \quad (3.38)$$

where  $j_1, \dots, j_l$  vary in the set of indices  $\{1, \dots, n\}$ . This can be shown by induction on the depth of  $P$  and using Jacobi identity. By doing the same thing for  $[x_{j_l}, Q]$  in 3.38 one shows that  $[P, Q]$  is a linear combination of vectors of the type  $ad_{x_{j_1}} ad_{x_{j_2}} \cdots ad_{x_{j_k}} x_h$ , with  $j_1, j_2, \dots, j_k, h$  in  $\{1, \dots, n\}$ , namely of repeated Lie brackets.

Provide the details and fill the gaps in this proof.

**Exercise 3.2** As we have seen in [Chapter 2](#), many quantum control systems

---

<sup>8</sup>The operation  $ad_x$  on elements of a Lie algebra  $\mathcal{L}$  is defined as  $ad_x Q := [x, Q]$ .

have a bilinear structure

$$\dot{X} = AX + \sum_{k=1}^m B_k X u_k.$$

Assume that the set of the possible values for the controls contains a neighborhood of the origin in  $\mathbf{R}^k$ . Show that the dynamical Lie algebra is the one generated by  $A, B_1, \dots, B_m$ .

**Exercise 3.3** Write a MATLAB code to test the operator controllability of finite dimensional quantum systems based on the algorithm of subsection 3.2.1. The algorithm should accept  $s$  matrices and return an answer YES or NO according to whether the system is controllable or not. Use this algorithm to verify whether or not the following systems are operator controllable:

1. The system of two interacting spin  $\frac{1}{2}$  particles with nonzero Ising interaction and all the components of the magnetic field possibly used for control.
2. The system of two interacting spin  $\frac{1}{2}$  particles with interaction equal to zero and all the components of the magnetic field possibly used for control.
3. The system of two interacting spin  $\frac{1}{2}$  particles with nonzero Ising interaction different from zero and only the  $x$  component of the magnetic field possibly used for control.

**Exercise 3.4** Use the correspondence between Lie algebras and connected Lie groups discussed in subsection 3.1.2.3 to show that if  $\mathcal{L}_1$  and  $\mathcal{L}_2$  are two Lie algebras and  $e^{\mathcal{L}_1}$  and  $e^{\mathcal{L}_2}$  the corresponding Lie groups,

$$\mathcal{L}_1 \subseteq \mathcal{L}_2$$

if and only if

$$e^{\mathcal{L}_1} \subseteq e^{\mathcal{L}_2}.$$

Show that proper inclusion for the Lie algebras is equivalent to proper inclusion for the Lie groups.

**Exercise 3.5** Prove that the bilinearity condition (3.3) along with condition (3.4) implies, for a general Lie algebra  $\mathcal{L}$ , that  $[y, x]$  is the inverse element of  $[x, y]$  in the vector space.

**Exercise 3.6** Prove that a unitary matrix of the form (3.21) is automatically special unitary namely its determinant is equal to one.

**Exercise 3.7** Give a basis of  $su(n)$  and a basis of  $sp(\frac{n}{2})$ .

**Exercise 3.8** Find a transformation  $X$  in  $Sp(2)$  which performs the state transfer

$$\psi_1 = X\psi_0,$$

where  $\psi_0 = [\frac{1}{\sqrt{2}}, 0, 0, \frac{1}{\sqrt{2}}]^T$  and  $\psi_1 = [0, 0, \frac{1}{\sqrt{2}}, i\frac{1}{\sqrt{2}}]^T$ .

**Exercise 3.9** Display a basis of a subalgebra of  $su(4)$  which is isomorphic to  $sp(2)$  but it is not  $sp(2)$ . Use the test in [Subsection 3.6](#) to show that the corresponding group is transitive on the complex sphere.

**Exercise 3.10** Prove formula (3.32).

**Exercise 3.11** Study the controllability of the system in Example 3.2.2 in the case  $\gamma_1 \neq \gamma_2$ .

# Chapter 4

---

## *Observability and State Determination*

As is known from measurement theory (cf. section 1.2), the measurement of an observable gives a certain result with a probability depending on the state. This suggests that measurements of appropriate observables on several copies of the same system can be used to determine the state of the system with some level of confidence. In systems theory, the study of the *observability* of a system describes to what extent it is possible to determine the initial state by a combined action of dynamics and output measurement.

In this chapter, we shall first briefly summarize the main ideas of *quantum state tomography* which is the technique used to determine the quantum state. Then, we shall turn to the system theoretic treatment of the problem of state determination namely to the study of the observability of quantum systems. We shall partition the space of density matrices into equivalence classes of states that cannot be distinguished by appropriate control and measurement. This allows us to consider the equation that describes the dynamics (Liouville's equation in this case) on a lower dimensional space. We shall discuss the impact of the observability analysis on methods for the determination of the state.

---

### 4.1 Quantum State Tomography

**Quantum state tomography** is a method to determine the state of a quantum system via a series of measurements. It also allows us to estimate the expectation value of an observable  $A$  for a system in a state  $\rho$ , that is,  $\langle A \rangle_\rho$ . The method uses the results of a large number  $N$  of *selective measurements* of a quorum of observables on  $N$  identical quantum systems.

#### 4.1.1 Example: Quantum tomography of a spin- $\frac{1}{2}$ particle

Consider the case of a spin  $\frac{1}{2}$  particle. The underlying Hilbert space is two dimensional. An orthonormal basis in the space of the Hermitian operators

on this Hilbert space is spanned by the  $2 \times 2$  multiple of the identity matrix  $\tilde{\sigma}_0 = \frac{1}{\sqrt{2}}\mathbf{1}$ , along with the multiples of the Pauli matrices defined in (1.20),  $\tilde{\sigma}_{x,y,z} := \frac{1}{\sqrt{2}}\sigma_{x,y,z}$ . Therefore, we can expand any operator  $A$  as

$$A = \sum_{j=0,x,y,z} \text{Tr}(A\tilde{\sigma}_j)\tilde{\sigma}_j. \quad (4.1)$$

In particular, this expansion is valid for the density matrix  $\rho$ . Therefore, recalling that the expectation value  $\langle \tilde{\sigma}_j \rangle_\rho$  is defined as  $\langle \tilde{\sigma}_j \rangle_\rho = \text{Tr}(\rho\tilde{\sigma}_j)$ , we can write  $\rho$  as

$$\rho = \frac{1}{2}\mathbf{1} + \sum_{j=x,y,z} \langle \tilde{\sigma}_j \rangle_\rho \tilde{\sigma}_j = \frac{1}{2}\mathbf{1} + \sum_{j=x,y,z} \sum_{m=\pm\frac{1}{\sqrt{2}}} P_\rho(m,j)m\tilde{\sigma}_j. \quad (4.2)$$

In this formula, we have denoted by  $P_\rho(m,j)$  the probability of obtaining the result  $m = \pm\frac{1}{\sqrt{2}}$  when measuring  $\tilde{\sigma}_j$ , in the state  $\rho$ . From this formula, if one is able to measure the expectation value of the *quorum* of observables  $\tilde{\sigma}_{x,y,z}$ , one reconstructs the state  $\rho$ . The observables  $\tilde{\sigma}_{x,y,z}$  correspond to the values of the spin angular momentum in the  $x, y, z$  directions respectively. Their eigenvalues are  $\pm\frac{1}{\sqrt{2}}$ .

It is convenient in the sum (4.2) to highlight the factors which depend only on the quorum of observables and the factors given by the probabilities which depend on the results of the measurements. One defines an *estimator* (or *kernel function*) as

$$\mathcal{R}(m,j) := \tilde{\sigma}_j m,$$

and writes (4.2) as

$$\rho = \frac{1}{2}\mathbf{1} + \sum_{j=x,y,z} \sum_{m=\pm\frac{1}{\sqrt{2}}} \mathcal{R}(m,j)P_\rho(m,j). \quad (4.3)$$

The expectation value of an arbitrary observable  $A$  can be obtained directly from (4.2) and (4.3). We have

$$\begin{aligned} \langle A \rangle_\rho &= \frac{1}{2} \text{Tr}(A) + \sum_{j=x,y,z} \langle \tilde{\sigma}_j \rangle_\rho \text{Tr}(A\tilde{\sigma}_j) \\ &= \frac{1}{2} \text{Tr}(A) + \sum_{j=x,y,z} \sum_{m=\pm\frac{1}{\sqrt{2}}} \mathcal{K}_A(m,j)P_\rho(m,j), \end{aligned} \quad (4.4)$$

where we have used the *scalar kernel function*  $\mathcal{K}_A(m,j) := m \text{Tr}(A\tilde{\sigma}_j)$ .

In practice, the probabilities  $P_\rho(m,j)$  are obtained by counting the results on a large number of measurements.

### 4.1.2 General quantum tomography

The scheme described above can be greatly generalized to give a unifying scheme for quantum tomography procedures.<sup>1</sup> Consider a set of operators  $\{C_k\}$ ,  $k = 1, 2, \dots$ , not necessarily Hermitian, and a dual set  $\{B_k\}$  such that every Hermitian operator  $A$  can be expanded as (cf. (4.1))

$$A = \sum_j \text{Tr}(AB_j^\dagger)C_j. \quad (4.5)$$

**Proposition 4.1.1** The sets  $\{B_k\}$  and  $\{C_k\}$  give the expansion (4.5) for every observable  $A$  if the following two conditions are satisfied:

1. The set  $\{C_k\}$  is *complete*, i.e., every Hermitian operator  $A$  can be expanded as (cf. (4.1))

$$A := \sum_k a_k C_k.$$

- 2.

$$\text{Tr}(C_j B_k^\dagger) = \delta_{jk}.$$

**Proposition 4.1.2** The sets  $\{B_k\}$  and  $\{C_k\}$  give the expansion (4.5) for every observable  $A$  if the following condition is satisfied: For every orthonormal basis in the Hilbert space of the system  $\{|j\rangle\}$ ,

$$\sum_\lambda \langle j | B_\lambda^\dagger | m \rangle \langle n | C_\lambda | k \rangle = \delta_{m,n} \delta_{j,k}. \quad (4.6)$$

Either one of the two propositions can be used to check whether the operators  $\{B_k\}$  and  $\{C_k\}$  satisfy the desired property. The proof of 4.1.1 is obvious while the proof of 4.1.2 is slightly more complicated<sup>2</sup> (also cf. Exercise 4.1). Given the sets  $\{B_j\}$  and  $\{C_j\}$ , a **quorum** of observables is defined as a set of observables  $\{Q_j\}$ ,  $j = 1, 2, \dots$ , such that there exist functions  $f_j$ , with  $f_j(Q_j) = B_j$ . In this case, from the expansion (4.5) applied to  $\rho$ , we have

$$\rho = \sum_j \text{Tr}(\rho f_j(Q_j)^\dagger)C_j. \quad (4.7)$$

---

<sup>1</sup>We consider the case of Hilbert space with countable basis. There are natural generalizations to the case of uncountable basis.

<sup>2</sup>Write  $A$  as

$$A = \sum_{j,m} \langle m | A | j \rangle | m \rangle \langle j | = \sum_{j,m,k,n} \langle m | A | j \rangle | n \rangle \langle k | \delta_{m,n} \delta_{j,k}.$$

Using (4.6), we write

$$A = \sum_\lambda \sum_{j,m,k,n} \langle m | A | j \rangle \langle j | B_\lambda^\dagger | m \rangle \langle n | C_\lambda | k \rangle | n \rangle \langle k |,$$

which, using the completeness relation (cf. (1.9))  $\sum_j |j\rangle \langle j| = \mathbf{1}$ ,  $\sum_m \langle m | AB_\lambda^\dagger | m \rangle = \text{Tr}(AB_\lambda^\dagger)$ , and  $C_\lambda = \sum_{n,k} \langle n | C_\lambda | k \rangle | n \rangle \langle k |$ , gives (4.5).

By writing  $Q_j$  in terms of its eigenvalues and eigenvectors as

$$Q_j = \sum_{m_j} m_j |m_j\rangle\langle m_j|,$$

and  $f(Q_j) = \sum_{m_j} f(m_j)|m_j\rangle\langle m_j|$  and recalling that  $\langle m_j|\rho|m_j\rangle$  is the probability of having the result  $m_j$  when measuring  $Q_j$ , i.e.,  $P_\rho(m_j, j)$ , we can rewrite (4.7) as

$$\rho = \sum_j \sum_{m_j} \mathcal{R}(m_j, j) P_\rho(m_j, j), \quad (4.8)$$

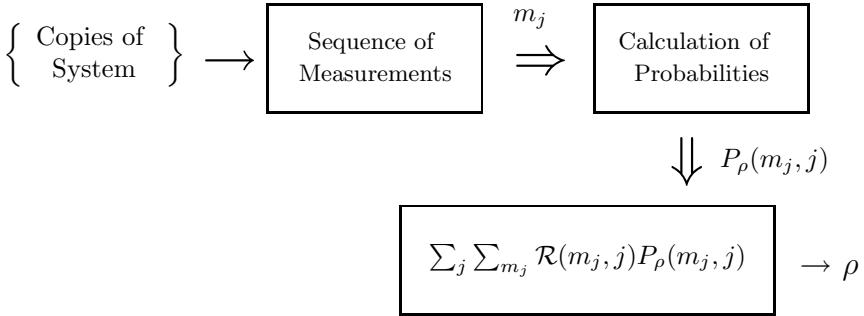
where the estimator  $\mathcal{R}(m_j, j)$  is defined by  $\mathcal{R}(m_j, j) := C_j f_j(m_j)^*$ , and it only depends on the quorum observables but not on the results of the measurements. Formula (4.8) can also be written in terms of the expectation values of the functions  $f_j$ , denoted here by  $\langle f_j \rangle_\rho := \text{Tr}(\rho f_j(Q_j))$ . We have

$$\rho = \sum_j \langle f_j \rangle_\rho^* C_j. \quad (4.9)$$

The expectation value of an observable  $A$  can be found from (4.8), (4.9), as

$$\langle A \rangle_\rho = \sum_j \sum_{m_j} \mathcal{K}_A(m_j, j) P_\rho(m_j, j) = \sum_j \langle f_j \rangle_\rho^* \text{Tr}(AC_j), \quad (4.10)$$

where the scalar kernel function is defined as  $\mathcal{K}_A(m_j, j) := \text{Tr}(A\mathcal{R}(m_j, j))$ . Once experiments are carried out to estimate the probabilities  $P_\rho(m_j, j)$ , formulas (4.8), (4.10) can be used to calculate the state  $\rho$  or the expectation value of any observable  $A$ .



**FIGURE 4.1:** Scheme of the various stages of a quantum tomographic procedure.

In the scheme described in the previous subsection,  $\tilde{\sigma}_{0,x,y,z}$  play the role of the operators  $C_j$  as well as the one of the corresponding dual operators  $B_j$  (i.e., in this case, every operator coincides with its dual). Moreover, they coincide with the quorum operators  $Q_j$ , i.e., the maps  $f_j$  above defined are all equal to the identity.

The above scheme naturally extends to the case of an uncountable number of operators  $C_\lambda$  and  $B_\lambda$ . In this case, we have a set of operators  $C_\lambda$  and  $B_\lambda$ ,  $\lambda$  in a possibly uncountable but measurable set  $\Lambda$ , such that every observable  $A$  can be expressed as<sup>3</sup> (cf. (4.5))

$$A = \int_{\Lambda} \text{Tr}(AB_\lambda^\dagger)C_\lambda d\lambda. \quad (4.11)$$

It is easy to show that this is the case if (cf. Proposition 4.1.2), for every orthonormal basis in the Hilbert space of the system  $\{|j\rangle\}$ ,

$$\int_{\Lambda} \langle j | B_\lambda^\dagger | m \rangle \langle n | C_\lambda | k \rangle d\lambda = \delta_{m,n} \delta_{j,k}. \quad (4.12)$$

Given a quorum of observables  $Q_\lambda$ , the formula for the reconstruction of  $\rho$  corresponding to (4.8) reads as (cf. (4.8))

$$\rho = \int_{\Lambda} \sum_{m_\lambda} \mathcal{R}(m_\lambda, \lambda) P_\rho(m_\lambda, \lambda) d\lambda, \quad (4.13)$$

where the estimator  $\mathcal{R}(m_\lambda, \lambda)$  is defined by  $\mathcal{R}(m_\lambda, \lambda) := C_\lambda(f_\lambda(m_\lambda))^*$ .

Quantum tomography schemes specialize the above approach to various situations.

#### 4.1.3 Example: Quantum tomography of a spin- $\frac{1}{2}$ particle (ctd.)

We conclude this section by reconsidering the problem of state determination for a spin  $\frac{1}{2}$  treated in subsection 4.1.1, but with a different quorum of observables parametrized by a continuous set  $\Lambda$ . In particular, let  $\Lambda$  be the set of angles  $\theta, \phi, \psi$ ,

$$\Lambda := \{\theta, \phi, \psi | 0 \leq \theta \leq 2\pi, 0 \leq \phi \leq \pi, -\pi \leq \psi \leq \pi\}.$$

Consider the quorum of observables given by multiples of the spin operators in all the possible directions. If the direction  $\vec{n}$  is given by

$$\vec{n} = (\cos(\theta) \sin(\phi), \sin(\theta) \sin(\phi), \cos(\phi))^T$$

---

<sup>3</sup>For an example see the following subsection, 4.1.3.

and the spin operator  $\vec{\sigma}$  is represented by the Pauli matrices (1.20)  $\vec{\sigma} := \{\sigma_x, \sigma_y, \sigma_z\}$ , then the spin operator in the direction  $\vec{n}$  can be written as  $\vec{\sigma} \cdot \vec{n}$ . The quorum of observables  $\{Q_\lambda\}$  with  $\lambda \in \Lambda$  is given by

$$Q_\lambda = \vec{\sigma} \cdot \vec{n}(\theta, \phi)\psi.$$

The function  $f_\lambda$  which maps  $Q_\lambda$  to the operator  $C_\lambda$  is taken as the exponential  $f_\lambda(Q_\lambda) := e^{-iQ_\lambda} = e^{-i\vec{\sigma} \cdot \vec{n}\psi} := C_\lambda$ . Such a unitary operator physically can be seen as a rotation around the direction  $\vec{n}$  by an angle  $2\psi$  (cf., e.g., [185] section 3.2). Its matrix expression can be easily calculated and it is given by (with  $\lambda = (\theta, \phi, \psi)$ )

$$C_\lambda = \begin{pmatrix} \cos(\psi) - i \cos(\phi) \sin(\psi) & -i \sin(\phi) e^{-i\theta} \sin(\psi) \\ -i \sin(\phi) \sin(\psi) e^{i\theta} & \cos(\psi) + i \cos(\phi) \sin(\psi) \end{pmatrix}. \quad (4.14)$$

We choose, for every  $\lambda$ ,  $B_\lambda = C_\lambda$ . With this choice, and taking as the volume element in the set  $\Lambda$ ,  $d\lambda = \frac{1}{2\pi^2} \sin^2(\psi) \sin(\phi)$ , it is a routine calculation (using as a basis the set of eigenvectors of  $\sigma_z$ ) to verify (4.12)<sup>4</sup>, i.e.,

$$\frac{1}{2\pi^2} \int_0^\pi \int_0^{2\pi} \int_{-\pi}^\pi \langle j | B_\lambda^\dagger | m \rangle \langle n | C_\lambda | k \rangle \sin^2(\psi) \sin(\phi) d\psi d\theta d\phi = \delta_{m,n} \delta_{j,k}. \quad (4.15)$$

Therefore, we can apply the continuous version of Proposition 4.1.2 and write every operator  $A$  as (cf. (4.11))

$$A = \frac{1}{2\pi^2} \int_0^\pi \int_0^{2\pi} \int_{-\pi}^\pi \text{Tr}(A e^{i\vec{\sigma} \cdot \vec{n}\psi}) e^{-i\vec{\sigma} \cdot \vec{n}\psi} \sin^2(\psi) \sin(\phi) d\psi d\theta d\phi.$$

With the given quorum of observables  $Q_\lambda$ , the formula for the reconstruction of  $\rho$  reads as (cf. 4.13)

$$\rho = \frac{1}{2\pi^2} \int_{\Lambda} \int_{-1}^1 \mathcal{R}(m_{\theta,\phi,\psi}, \theta, \phi, \psi) P_\rho(m_{\theta,\phi,\psi}) dm_{\theta,\phi,\psi} \sin^2(\psi) \sin(\phi) d\psi d\theta d\phi,$$

with the kernel  $\mathcal{R}(m_{\theta,\phi,\psi}, \theta, \phi, \psi)$  given by

$$\mathcal{R}(m_{\theta,\phi,\psi}, \theta, \phi, \psi) := e^{-im_{\theta,\phi,\psi}} e^{i\vec{\sigma} \cdot \vec{n}\psi}.$$

---

<sup>4</sup>Formula (4.15) is a consequence of more general properties and in particular of the fact that by varying  $\theta, \phi$  and  $\psi$ , (4.14) gives an irreducible representation of  $SU(2)$  and that the element of volume  $d\lambda := \frac{1}{2\pi^2} \sin^2(\psi) \sin(\phi)$  is invariant under group translations. We refer to [150], [152] for a treatment of representation theory with particular emphasis on the unitary and rotation groups. In particular the formula on pg. 216 of [150] corresponds to (4.12), while Chapters 5 and 6 of [152] give the definition of the volume element and a method to calculate it. In fact, group theory can be used to generalize the treatment here to more general systems which include, for example, spin different from  $\frac{1}{2}$  and the electromagnetic radiation field, when considered as a quantum system (rather than classically as in Chapter 2 with the semiclassical approximation). This approach is known as *Group Tomography*. We refer to [67] and the references therein.

## 4.2 Observability

As seen in the previous section, the formulas used in quantum tomography express the quantum state in terms of the probabilities and the expectation values of a quorum of observables. The evaluation of the probabilities has to be carried out by performing a large number of measurements of the quorum observables on exact copies of the same system. As the quorum observables consist typically of a large (sometimes uncountably infinite, see subsection 4.1.3) number of observables, this involves many types of measurement experiments. Often, as in the spin- $\frac{1}{2}$  example of subsections 4.1.1 and 4.1.3, the quorum of observables  $\{Q_\lambda\}$  can be obtained by a set of unitary similarity transformations,  $U_\lambda$ , on a single observable  $Q$ , i.e.,

$$Q_\lambda = U_\lambda^\dagger Q U_\lambda, \quad \lambda \in \Lambda.$$

In these cases, we have, for the expectation values  $\langle Q_\lambda \rangle_\rho$ ,

$$\langle Q_\lambda \rangle_\rho = \text{Tr}(\rho U_\lambda^\dagger Q U_\lambda) = \text{Tr}(U_\lambda \rho U_\lambda^\dagger Q), \quad (4.16)$$

and for the probabilities  $P_\rho(m, \lambda)$  in (4.8), (4.13), denoting by  $P_m$  the projection onto the eigenspace of  $Q$  corresponding to the eigenvalue  $m$ ,

$$P_\rho(m, \lambda) = \text{Tr}(\rho U_\lambda^\dagger P_m U_\lambda) = \text{Tr}(U_\lambda \rho U_\lambda^\dagger P_m) \quad (4.17)$$

(cf. subsection 1.2.3). Formulas (4.16) and (4.17) suggest that, instead of changing the measurement apparatus, for any type of measurement, we could as well rely on the dynamics of the system  $\rho \rightarrow U\rho U^\dagger$  and then measure always the same observable. This raises the question of characterizing, given a certain dynamics of the control system, the class of initial states that will give the same probabilities (expectation value) for any possible evolution.

To pose this problem in system theoretic terms as a study of observability, we consider the quantum control system (2.72) with output given by the probabilities  $y(t) := \Pr(m) = \text{Tr}(\rho(t)P_m)$ , or with output given the expectation value  $y(t) := \langle Q \rangle_\rho = \text{Tr}(\rho(t)Q)$ . In the last two formulas,  $\rho(t)$  is the solution of the differential equation (2.72). Notice that, in both cases, the output  $y$  is a linear function of the system. Therefore, we shall study the observability properties of system (2.72)

$$\dot{\rho} = [-iH(u(t)), \rho], \quad \rho(0) = \rho_0, \quad (4.18)$$

with output

$$y(t) = \text{Tr}(S\rho), \quad (4.19)$$

for some Hermitian operator  $S$  (which could be  $P_m$  or  $Q$ ). Denote by  $\rho(t, u, \tilde{\rho})$  the solution of (4.18) with initial condition  $\rho_0 = \tilde{\rho}$  and control  $u$  at time  $t$ . We have the following definition

**Definition 4.2.1** Two states  $\tilde{\rho}_1$  and  $\tilde{\rho}_2$  are called **indistinguishable** if, for any control  $u$ , and any  $t$  (cf. (4.19)),

$$\mathrm{Tr}(S\rho(t, u, \tilde{\rho}_1)) = \mathrm{Tr}(S\rho(t, u, \tilde{\rho}_2)).$$

**Definition 4.2.2** System (4.18), (4.19) is called **observable**, if two states  $\tilde{\rho}_1$  and  $\tilde{\rho}_2$  are indistinguishable if and only if  $\tilde{\rho}_1 = \tilde{\rho}_2$ .

#### 4.2.1 Equivalence classes of indistinguishable states; Partition of the state space

As seen in [Chapter 3](#), given the dynamical Lie algebra  $\mathcal{L}$  associated with the quantum system (4.18), there exists a control giving the unitary evolution  $X$  if and only if  $X$  belongs to the corresponding Lie group  $e^{\mathcal{L}}$ . Given the form (2.73) of the solution of (4.18), we easily obtain the following.

**Proposition 4.2.3** Two states  $\tilde{\rho}_1$  and  $\tilde{\rho}_2$  are indistinguishable if and only if

$$\mathrm{Tr}(X^\dagger S X \tilde{\rho}_1) = \mathrm{Tr}(X^\dagger S X \tilde{\rho}_2), \quad \forall X \in e^{\mathcal{L}}.$$

In order to state the condition of indistinguishability in linear algebra terms, it is convenient to define the *observability spaces* as follows. Let

$$\tilde{S} := S - \frac{\mathrm{Tr}(S)}{n} \mathbf{1}_{n \times n}, \quad (4.20)$$

where  $n$  is the dimension of the system. The observability space  $\mathcal{V}$  is defined as

$$\mathcal{V} := \bigoplus_{j=0}^{\infty} \mathrm{ad}_{\mathcal{L}}^j \mathrm{span}\{i\tilde{S}\}, \quad (4.21)$$

where  $\mathrm{ad}_{\mathcal{L}}^j \mathcal{A}$  is defined as spanned by all the repeated Lie brackets

$$[R_j, [\cdots [R_2, [R_1, iA]] \cdots]],$$

with  $R_1, \dots, R_j \in \mathcal{L}$  and  $iA \in \mathcal{A}$ .

With these definitions, we have:

**Theorem 4.2.4** System (4.18) with output  $y$  in (4.19) is observable if and only if

$$\mathcal{V} = su(n). \quad (4.22)$$

More generally, write  $\rho = \rho_1 + \rho_2$  where  $\rho_1$  is the component of  $\rho$  in<sup>5</sup>  $i\mathcal{V}$  and  $\rho_2$  is the component along  $i\mathcal{V}^\perp$  where  $\mathcal{V}^\perp$  is the orthogonal complement of  $\mathcal{V}$

---

<sup>5</sup> $i\mathcal{V}$  denotes the vector space of Hermitian matrices obtained by multiplying by  $i$  the skew-Hermitian matrices in  $\mathcal{V}$ .

in  $u(n)$ . Then, the following decomposition of the dynamics holds true

$$\begin{aligned}\dot{\rho}_1 &= -i[H(u), \rho_1], \\ \dot{\rho}_2 &= -i[H(u), \rho_2],\end{aligned}\tag{4.23}$$

and we have

$$y(t) := \text{Tr}(S\rho(t)) = \text{Tr}(S) + \text{Tr}(S\rho_1(t)).\tag{4.24}$$

Initial states  $\tilde{\rho}_1$  and  $\tilde{\rho}_2$  are indistinguishable if and only if  $(\tilde{\rho}_1 - \tilde{\rho}_2) \in i\mathcal{V}^\perp$ .

*Proof.* Write (4.18) with  $\rho = \rho_1 + \rho_2$ ,  $\rho_1 \in i\mathcal{V}$ ,  $\rho_2 \in i\mathcal{V}^\perp$ ,

$$\dot{\rho}_1 + \dot{\rho}_2 = [-iH(u(t)), \rho_1] + [-iH(u(t)), \rho_2],\tag{4.25}$$

and notice that, as  $i\mathcal{V}$  is invariant under the commutator operation with  $-iH(u(t))$ , the first commutator on the right hand side of (4.25) is always in  $i\mathcal{V}$ . The invariance of  $i\mathcal{V}$  also implies the invariance of  $i\mathcal{V}^\perp$ , i.e.,<sup>6</sup>

$$[-iH(u), i\mathcal{V}^\perp] \subseteq i\mathcal{V}^\perp.$$

Separating the components in  $i\mathcal{V}$  and  $i\mathcal{V}^\perp$  in (4.25), we obtain (4.23). Moreover, using the fact that  $\rho_1$  is traceless while  $\text{Tr}(\rho_2) = 1$ , one obtains (4.24). By decomposing  $\rho(t, u, \tilde{\rho}_1)$  and  $\rho(t, u, \tilde{\rho}_2)$  according to their components in  $i\mathcal{V}$  and  $i\mathcal{V}^\perp$  and applying (4.23) and (4.24) one sees that the outputs coincide for every control if the components of  $\tilde{\rho}_1$  and  $\tilde{\rho}_2$  in  $i\mathcal{V}$  coincide. This proves that if  $\tilde{\rho}_1 - \tilde{\rho}_2 \in \mathcal{V}^\perp$ ,  $\tilde{\rho}_1$  and  $\tilde{\rho}_2$  are indistinguishable. Conversely, assume that  $\tilde{\rho}_1$  and  $\tilde{\rho}_2$  are indistinguishable. Then, from Proposition 4.2.3, for every  $X \in e^{\mathcal{L}}$  we have

$$\text{Tr}(X^\dagger \tilde{S} X (\tilde{\rho}_1 - \tilde{\rho}_2)) = 0.\tag{4.26}$$

In particular, if  $R_1, \dots, R_m$  are arbitrary elements of  $\mathcal{L}$  and  $t_1, \dots, t_m$  arbitrary real numbers, we have

$$e^{R_1 t_1} e^{R_2 t_2} \cdots e^{R_m t_m} \tilde{S} e^{R_1^\dagger t_1} e^{R_2^\dagger t_2} \cdots e^{R_m^\dagger t_m} (\tilde{\rho}_1 - \tilde{\rho}_2) \equiv 0.$$

Taking (possibly higher order) derivatives with respect to  $t_1, t_2, \dots, t_m$  at  $t_1 = t_2 = \dots = t_m = 0$  and taking into account that  $R_1, \dots, R_m$  are arbitrary, we see the  $\text{Tr}(A(\tilde{\rho}_1 - \tilde{\rho}_2)) = 0$  for  $A$  varying in a spanning set of  $\mathcal{V}$ . This shows that  $\tilde{\rho}_1 - \tilde{\rho}_2 \in i\mathcal{V}^\perp$ . In the special case where  $\mathcal{V} = su(n)$ , we have  $(\tilde{\rho}_1 - \tilde{\rho}_2) \in \text{span}\{\mathbf{1}\}$  and since  $\text{Tr}(\tilde{\rho}_1) = \text{Tr}(\tilde{\rho}_2) = 1$ , we have  $\tilde{\rho}_1 = \tilde{\rho}_2$  and the system is observable.  $\square$

The observability space  $\mathcal{V}$  determines a partition of the space of density matrices into equivalence classes of indistinguishable states. Two states  $\tilde{\rho}_1$

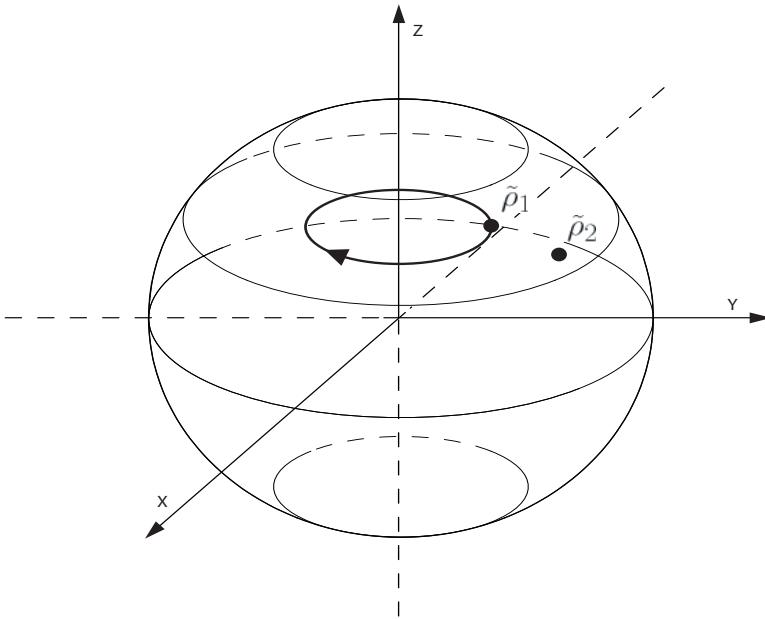
---

<sup>6</sup>If  $A \in i\mathcal{V}^\perp$ , for every  $B \in i\mathcal{V}$ , we have  $\text{Tr}([-iH, A]B) = \text{Tr}([B, -iH]A) = 0$  since  $[B, -iH] \in i\mathcal{V}$ .

and  $\tilde{\rho}_2$  belong to the same equivalence class if and only if they have the same component along  $\mathcal{V}$ . Moreover such an equivalence class is an invariant set for the dynamics as expressed by (4.23). Consider as an example the two level case where  $\rho$  in (1.19), (1.20) is represented by a point  $(x, y, z)$  in the Bloch sphere (cf. Figure 1.1). Assume  $\mathcal{V} = \text{span}\{i\sigma_z\}$ . Then the situation is the one depicted in Figure 4.2. Two points  $\tilde{\rho}_1$  and  $\tilde{\rho}_2$  with the same value of  $z$  in (1.19) are indistinguishable. Moreover, assume  $H(u) = \sigma_z$ , for every control  $u$ . Then the dynamics is given by

$$\begin{aligned}\tilde{\rho}_1 &\rightarrow e^{-i\sigma_z t} \tilde{\rho}_1 e^{i\sigma_z t} \\ &= \frac{1}{2} \mathbf{1} + z\sigma_z + (\cos(2t)x + \sin(2t)y)\sigma_x + (-\sin(2t)x + \cos(2t)y)\sigma_y,\end{aligned}$$

that is, it is a circle in a surface  $z = \text{const}$  and therefore it is such that the component along  $\mathcal{V}$  remains the same. This situation extends to higher dimensions according to Theorem 4.2.4.



**FIGURE 4.2:** Partition of the Bloch sphere into equivalence classes of indistinguishable states.  $\tilde{\rho}_1$  and  $\tilde{\rho}_2$  are indistinguishable and every indistinguishable state evolves into indistinguishable states, as for example for the trajectory starting at  $\tilde{\rho}_1$ .

The following remarks sketch a number of generalizations of the Lie algebraic characterization of observability of Theorem 4.2.4.

**Remark 4.2.5** Theorem 4.2.4 extends easily to the case of several outputs. This is typically the case when we want to find the set of initial states which will give the same probabilities for any dynamics. One modifies the definition of observability spaces (4.21) by replacing  $\text{span}\{i\tilde{S}\}$  with

$$\text{span}\{i\tilde{P}_1, i\tilde{P}_2, \dots, i\tilde{P}_m\},$$

where  $\tilde{P}_j = P_j - \text{Tr}(P_j)\mathbf{1}$ ,  $j = 1, \dots, m$ , where the  $P_j$ 's are the output operators (which might represent projections onto the eigenspaces of a given observable). The rest of the argument goes through with only formal modifications.

**Remark 4.2.6** Another generalization consists of considering generalized measurement theory as described in [Appendix A](#). In this case, all the treatment goes through by replacing the projections  $P_m$  with the effects  $F_m$  in [subsection A.3](#).

**Remark 4.2.7** One more generalization can be obtained by considering a sequence of measurements. As the state is modified at each measurement, according to measurement theory (cf. [section 1.2](#)), indistinguishability at the first measurement is not equivalent to indistinguishability at a following measurement. The treatment however can be extended to this case assuming we perform a nonselective measurement of the expectation value so that the modification of the state is assumed not to depend on the result (cf. formula (1.38) in [Chapter 1](#)). Let  $\mathcal{F}$  be the linear transformation of the state  $\rho \rightarrow \mathcal{F}(\rho)$  due to the measurement and  $\mathcal{F}^\dagger$  denote the dual map defined by  $\text{Tr}(A\mathcal{F}(\rho)) = \text{Tr}(\mathcal{F}^\dagger(A)\rho)$ , for every  $A$  and  $\rho$ . One defines indistinguishability in  $k$  steps for two states that cannot be distinguished at the  $k$ -th measurement. To give conditions for indistinguishability in  $k$  steps, one defines the *generalized observability spaces* recursively as

$$\mathcal{V}_0 := \text{span}\{i\tilde{S}\}, \quad \mathcal{V}_1 := \bigoplus_{j=0}^{\infty} ad_{\mathcal{L}}^j \mathcal{V}_0,$$

$$\mathcal{V}_k := \bigoplus_{j=0}^{\infty} ad_{\mathcal{L}}^j i\mathcal{F}^\dagger(i\mathcal{V}_{k-1}).$$

The criterion for indistinguishability (and observability) in  $k$  steps reads as Theorem 4.2.4 by replacing indistinguishability with indistinguishability in  $k$  steps and  $\mathcal{V}$  with  $\mathcal{V}_k$ . The proof is a generalization of the proof of 4.2.4 (see [57]).

Excluding the trivial case where  $S$  is a scalar matrix, if  $\mathcal{L} = su(n)$  or  $\mathcal{L} = u(n)$ , i.e., the system is operator controllable, then the observability space  $\mathcal{V}$  has to be (or to contain) the whole  $su(n)$ . This is a consequence of the fact that  $su(n)$  is a simple Lie algebra (cf. Proposition 3.4.6). As  $\mathcal{V}$  is an ideal of  $u(n)$  it can only be spanned by a multiple of the identity (a case we have excluded) or contain  $su(n)$ . We state this in the following proposition.

**Proposition 4.2.8** Operator controllability implies observability.

---

## 4.3 Observability and Methods for State Reconstruction

### 4.3.1 Observability conditions and tomographic methods

Consider now a quantum tomographic method where only a single observable is measured and the remaining information is obtained using the dynamics of the system. The practical scheme follows the same steps as in [Figure 4.1](#) where in the sequence of measurements every measurement consists of an evolution followed by the measurement of the given observable. Such a scheme is characterized in the discrete parameter set case by a formula of the type (4.8) or the continuous case by a formula (4.13) where the quorum  $\{Q_\lambda\}$  is obtained from a single observable  $Q$  by unitary similarity transformations, i.e.,  $Q_\lambda = X_\lambda^\dagger Q X_\lambda$ . Here  $X_\lambda \in e^{\mathcal{L}}$ , for  $\lambda$  in an appropriate discrete or continuous set. Therefore if  $P_m$  denotes a projection associated with the observable  $Q$  and the result  $m$ , formulas (4.8) and (4.13) can be rewritten respectively as

$$\rho = \sum_{\lambda} \sum_m f_{\lambda}(m)^* C_{\lambda} \operatorname{Tr}(X_{\lambda}^\dagger P_m X_{\lambda} \rho) := \sum_{\lambda} \mathcal{M}_{X_{\lambda}, C_{\lambda}, f_{\lambda}}(\rho), \quad (4.27)$$

or

$$\rho = \int_{\Lambda} \sum_m f_{\lambda}(m)^* C_{\lambda} \operatorname{Tr}(X_{\lambda}^\dagger P_m X_{\lambda} \rho) d\lambda := \int_{\Lambda} \mathcal{M}_{X_{\lambda}, C_{\lambda}, f_{\lambda}}(\rho) d\lambda.$$

Here we have introduced the notation  $\mathcal{M}_{X_{\lambda}, C_{\lambda}, f_{\lambda}}$  which is a function defined over a set  $\Lambda$  which characterizes the tomographic scheme. Given the dynamics and the observable which has to be measured  $Q$ , the particular scheme we adopt depends on the choice of the operators  $C_{\lambda}$ , the functions  $f_{\lambda}$ , and the unitary evolutions  $X_{\lambda}$  in  $e^{\mathcal{L}}$ . It follows from the above discussion that, unless the system is fully observable, with respect to the outputs  $P_m$  (cf. [Remark 4.2.5](#)) associated with the observable  $Q$ , determination of the state with the quantum tomographic method  $\mathcal{M}_{X_{\lambda}, C_{\lambda}, f_{\lambda}}(\rho)$  is possible only up to terms in the orthogonal of the observability space, namely  $i\mathcal{V}^\perp$ . We formalize this fact in the following proposition which we state for the case where  $\Lambda$  is a discrete set.

**Proposition 4.3.1** Let the observability space  $\mathcal{V}$  be calculated as in [Remark 4.2.5](#) starting from the projections  $P_m$  associated with the observable  $Q$ . Then  $\rho_1 - \rho_2 \in i\mathcal{V}^\perp$  if and only if, for *any* tomographic method  $\mathcal{M}_{X_{\lambda}, C_{\lambda}, f_{\lambda}}$ ,  $\mathcal{M}_{X_{\lambda}, C_{\lambda}, f_{\lambda}}(\rho_1) = \mathcal{M}_{X_{\lambda}, C_{\lambda}, f_{\lambda}}(\rho_2)$ .

*Proof.* If  $\rho_1 - \rho_2 \in i\mathcal{V}^\perp$ ,  $\rho_1$  and  $\rho_2$  are indistinguishable because of Theorem 4.2.4, and using (4.26) with the various projections  $P_m$  instead of  $\tilde{S}$ , we have

$$\mathcal{M}_{X_\lambda, C_\lambda, f_\lambda}(\rho_1) - \mathcal{M}_{X_\lambda, C_\lambda, f_\lambda}(\rho_2) = \sum_m f_\lambda(m)^* C_\lambda \text{Tr}(X_\lambda^\dagger P_m X_\lambda (\rho_1 - \rho_2)) = 0.$$

Conversely, assume  $\mathcal{M}_{X_\lambda, C_\lambda, f_\lambda}(\rho_1) = \mathcal{M}_{X_\lambda, C_\lambda, f_\lambda}(\rho_2)$  for every  $\lambda \in \Lambda$  and any tomographic method  $\mathcal{M}_{X_\lambda, C_\lambda, f_\lambda}$ . Choose  $X_1, \dots, X_r$  and projections  $P_1, \dots, P_r$  so that  $X_1 P_1 X_1^\dagger, \dots, X_r P_r X_r^\dagger$  span a basis of  $i\mathcal{V}$  (see Exercise 4.2). Choose  $f_\lambda(m_\lambda)^* C_\lambda$ ,  $\lambda = 1, \dots, r$  (where  $m_\lambda$  is the eigenvalue corresponding to the projection  $P_\lambda$ ) Hermitian with norm one, such that  $f_\lambda(m_\lambda)^* C_\lambda \in i\mathcal{V}$  is orthogonal to  $X_l P_l X_l^\dagger$ , for  $l \neq \lambda$ , which implies that  $f_\lambda(m_\lambda)^* C_\lambda$ ,  $\lambda = 1, \dots, r$ , are an orthonormal basis of  $i\mathcal{V}$ . With these choices,  $\mathcal{M}_{X_\lambda, C_\lambda, f_\lambda}(\rho_1) = \mathcal{M}_{X_\lambda, C_\lambda, f_\lambda}(\rho_2)$  using (4.27) implies  $\text{Tr}(X_\lambda P_\lambda X_\lambda^\dagger (\rho_1 - \rho_2)) = 0, \forall \lambda = 1, \dots, r$ , i.e.,  $\rho_1 - \rho_2 \in i\mathcal{V}^\perp$ .  $\square$

**Remark 4.3.2** Notice that the quantum tomographic scheme described in the second part of the proof of the theorem allows the determination of the component of  $\rho$  in  $i\mathcal{V}$ .

#### 4.3.2 System theoretic methods for quantum state reconstruction

The study of system theoretic properties, in particular observability, is not only a way to analyze the effectiveness of quantum tomographic methods, but it also gives rise to alternative procedures for the determination of the state. These can be obtained by applying the methods of system theory. The concept of observability plays a prominent role in this case, too. An integral, system theoretic formula for quantum state reconstruction based on the measurement of the expectation value of a given observable (4.19) can be obtained as follows.

Given the output (4.19) and the associated observability space  $i\mathcal{V}$  (4.20), (4.21), we can choose a control  $u = u(t)$ ,  $t \in [0, T]$  so that the corresponding solution  $X_u$  of Schrödinger operator equation

$$\dot{X} = -iH(u)X, \quad X(0) = \mathbf{1}, \quad (4.28)$$

satisfies

$$\text{span}_{t \in [0, T]} \{X_u^\dagger(t) \tilde{S} X_u(t)\} = i\mathcal{V}. \quad (4.29)$$

One way to do this is to select  $r := \dim(\mathcal{V})$  matrices  $X_1, \dots, X_r$  so that  $X_1^\dagger S X_1, \dots, X_r^\dagger S X_r$  are linearly independent and then concatenate the controls steering the matrix  $X$  in (4.28) to  $X_1, X_2 X_1^\dagger, X_3 X_2^\dagger, \dots, X_r X_r^\dagger$ .

Now, consider a trajectory  $X_u$  of (4.28) having property (4.29). From (4.19), (4.20), we obtain

$$y(t) = \text{Tr}(X_u^\dagger \tilde{S} X_u (\rho_0 - \frac{1}{n} \mathbf{1})) + \text{Tr}(S),$$

which gives

$$\begin{aligned} \int_0^T X_u^\dagger(t) \tilde{S} X_u(t) (y(t) - \text{Tr}(S)) dt = \\ \int_0^T X_u^\dagger(t) \tilde{S} X_u(t) \text{Tr} \left( X_u^\dagger(t) \tilde{S} X_u(t) (\rho_0 - \frac{1}{n} \mathbf{1}) \right) dt. \end{aligned} \quad (4.30)$$

Define the linear operator  $\mathcal{W}$  to map  $n \times n$  Hermitian matrices with zero trace into  $n \times n$  Hermitian matrices with zero trace as follows

$$\mathcal{W}_u(\hat{\rho}_0) := \int_0^T X_u^\dagger(t) \tilde{S} X_u(t) \text{Tr}(X_u^\dagger(t) \tilde{S} X_u(t) \hat{\rho}_0) dt. \quad (4.31)$$

The range of the operator  $\mathcal{W}_u$  is  $i\mathcal{V}$ . This follows from (4.29) and the fact that the kernel of  $\mathcal{W}_u$  is the component of  $i\mathcal{V}^\perp$  in  $isu(n)$ , i.e., the vector space of the traceless matrices in  $i\mathcal{V}^\perp$ . It is clear, in fact, that if  $\hat{\rho}_0 \in i\mathcal{V}^\perp$ ,  $\mathcal{W}_u(\hat{\rho}_0) = 0$ . Conversely, assume  $\mathcal{W}(\hat{\rho}_0) = 0$ . Then, from (4.31), we have

$$0 = \text{Tr}(\hat{\rho}_0 \mathcal{W}_u(\hat{\rho}_0)) = \int_0^T (\text{Tr}(X_u^\dagger(t) \tilde{S} X_u(t) \hat{\rho}_0))^2 dt,$$

which implies

$$\text{Tr}(X_u^\dagger(t) \tilde{S} X_u(t) \hat{\rho}_0) = 0, \quad a.e.$$

This, using (4.29), implies  $\hat{\rho}_0 \in i\mathcal{V}$ .

Formula (4.30) gives a linear system of equations whose solution allows one to calculate the initial state  $\rho_0$  modulo elements in  $i\mathcal{V}^\perp$ . In particular, if the system is observable, then  $\mathcal{W}_u$  is invertible and we have

$$\rho_0 = \frac{1}{n} \mathbf{1} + \mathcal{W}_u^{-1} \left( \int_0^T X_u^\dagger(t) \tilde{S} X_u(t) (y(t) - \text{Tr}(S)) dt \right). \quad (4.32)$$

Formula (4.32) represents a system theoretic alternative to methods for quantum state tomography.

We summarize the discussion in the following theorem.

**Theorem 4.3.3** *Consider system (4.18) with output (4.19). If the system is observable (in one step), then there exists a control such that formula (4.32) gives the initial state.*

**Example 4.3.4** If the system is observable, in order to obtain the initial state from formula (4.32), according to Theorem 4.3.3, the control has to be chosen so that the corresponding evolution  $X_u$  satisfies (4.29). Consider as an example a two level quantum system such that  $\tilde{S} := \sigma_x$  in (1.20) and the control is chosen so that  $H(u) := \sigma_z$  in  $[0, 2\pi]$  and  $H(u) := \sigma_y$  in  $[2\pi, 4\pi]$ . We calculate

$$\begin{aligned}\mathcal{W}_u(\hat{\rho}_0) &= \int_0^{2\pi} e^{i\sigma_z t} \sigma_x e^{-i\sigma_z t} \text{Tr}(e^{i\sigma_z t} \sigma_x e^{-i\sigma_z t} \hat{\rho}_0) dt \\ &\quad + \int_{2\pi}^{4\pi} e^{i\sigma_y t} \sigma_x e^{-i\sigma_y t} \text{Tr}(e^{i\sigma_y t} \sigma_x e^{-i\sigma_y t} \hat{\rho}_0) dt.\end{aligned}$$

We obtain

$$\mathcal{W}_u(\hat{\rho}_0) = \pi (2 \text{Tr}(\sigma_x \hat{\rho}_0) \sigma_x + \text{Tr}(\sigma_y \hat{\rho}_0) \sigma_y + \text{Tr}(\sigma_z \hat{\rho}_0) \sigma_z),$$

and

$$\mathcal{W}_u^{-1}(a\sigma_x + b\sigma_y + c\sigma_z) = \frac{1}{\pi} \left( \frac{a}{4} \sigma_x + \frac{b}{2} \sigma_y + \frac{c}{2} \sigma_z \right).$$


---

## 4.4 Notes and References

Our discussion of quantum tomography follows the general unifying scheme presented in [66], to which we also refer for a survey on the subject and relevant references. In particular, this paper presents some statistical aspects of quantum tomography as well as the connection with conventional tomographic methods in medical imaging. We have focused here on the fundamental aspects and the formulas for the reconstruction of the state emphasizing the role of the dynamics in the process of state reconstruction.

Paper [67] contains a more detailed description of spin tomography based on group theory, as well as a proposal for its practical implementation.

The treatment of observability for quantum systems is taken from [57] and [63], to which we refer for the generalizations mentioned in Remarks 4.2.5, 4.2.6, 4.2.7. This paper also presents a system theoretic treatment of observers for quantum systems related to formula (4.32). An alternative constructive algorithm which uses multiple measurements is given in [59].

---

## 4.5 Exercises

**Exercise 4.1** Prove that, if we relax the assumption that  $A$  is an Hermitian operator, the conditions of Propositions 4.1.1, 4.1.2 are also necessary for (4.5).

**Exercise 4.2** Let  $P_1, \dots, P_l$  be the orthogonal projections associated to an observable  $S$  as in the spectral decomposition (cf. 1.2.1.1). Let  $\mathcal{L}$  be the dynamical Lie algebra and  $e^{\mathcal{L}}$  the corresponding connected Lie group. Let

the observability space  $\mathcal{V}$  be calculated starting from  $P_1, \dots, P_l$ , as in Remark 4.2.5. Then prove that

$$i\mathcal{V} = \text{span}_{X_1, X_2, \dots, X_l \in e^{\mathcal{L}}} \{X_1^\dagger P_1 X_1, \dots, X_l^\dagger P_l X_l\}.$$

**Exercise 4.3** Consider a spin  $\frac{1}{2}$  particle under Hamiltonian  $H(u(t)) = \sigma_x u_x(t) + \sigma_y u_y(t)$  with measurement of the expectation value of the spin in the  $z$  direction,  $\langle \sigma_z \rangle_\rho$ . Write  $\rho$  in the Bloch representation (1.19) and consider the control  $u_x(t) \equiv 1, u_y(t) \equiv 0, t \in [0, 1]$ ,  $u_x(t) \equiv 0, u_y(t) \equiv 1, t \in (1, 2]$ . Give a formula for the determination of  $x, y$  and  $z$  in (1.19) based on (4.32).

**Exercise 4.4** Referring to Proposition 4.2.8, give an example of a system which is not operator controllable but is observable.

# Chapter 5

---

## Lie Group Decompositions and Control

Consider a Lie algebra  $\mathcal{L}$  and the corresponding connected Lie group  $e^{\mathcal{L}}$ . A *decomposition* of  $e^{\mathcal{L}}$  is a representation of all the elements  $X$  of  $e^{\mathcal{L}}$  as products

$$X = X_r X_{r-1} \cdots X_1, \quad (5.1)$$

with  $X_j \in G_j$ ,  $j = 1, \dots, r$ , and  $G_j$  are proper subsets of  $e^{\mathcal{L}}$ . These subsets typically have some further algebraic or geometric structure; for example they are Lie subgroups of  $e^{\mathcal{L}}$ .

An example of a decomposition of a Lie group can be found in the proof of the controllability test of [Chapter 3](#) which is given in [Appendix D](#). The discussion in [subsection 3.2.2](#) shows that, if the Lie group  $e^{\mathcal{L}}$  is compact, as for example  $SU(n)$ , and  $\{A_1, \dots, A_s\}$  is a set of generators of the Lie algebra  $\mathcal{L}$ , then there exists a number  $r$  such that every  $X$  in  $e^{\mathcal{L}}$  can be written as

$$X = e^{\tilde{A}_r t_r} e^{\tilde{A}_{r-1} t_{r-1}} \cdots e^{\tilde{A}_1 t_1}, \quad (5.2)$$

with  $t_j \geq 0$  and  $\tilde{A}_j$  given matrices in the set  $\{A_1, \dots, A_s\}$ ,  $j = 1, \dots, r$ . Formula (5.2) is a Lie group decomposition of the type (5.1) where the role of the subsets  $G_j$  is played by the semigroups  $\{e^{\tilde{A}_j t} | t \geq 0\}$ .

A typical control problem for the Schrödinger operator equation

$$\dot{X} = -iH(u)X, \quad X(0) = \mathbf{1}, \quad (5.3)$$

is to find the control function  $u$  which drives  $X$  from the identity  $\mathbf{1}$  to a desired target  $X_f \in e^{\mathcal{L}}$ . Now, if we have a Lie group decomposition of  $e^{\mathcal{L}}$ , we can write  $X_f$  as a product of elements  $X_{jf}$  in subsets of  $e^{\mathcal{L}}$ , i.e.,

$$X_f = X_{rf} X_{r-1f} \cdots X_{1f}. \quad (5.4)$$

Assume now that we are able for every  $X_{jf}$  to find a control  $u_j$  which drives the identity to  $X_{jf}$ . Then, by the right invariance property of system (5.4),<sup>1</sup>

---

<sup>1</sup>This means that if  $X(t)$  is the solution with initial condition equal to the identity, the solution with initial condition equal to  $S$  is  $X(t)S$ .

the concatenation of the controls  $u_1, u_2, \dots, u_r$ , in this order, drives the solution of (5.3) from the identity to the desired value  $X_f$ . So the problem of steering control is reduced to a sequence of subproblems. Lie group decompositions are of course useful if the subproblems are easier to solve than the original problem. This is usually the case if the subsets  $G_j$ ,  $j = 1, \dots, r$  are smaller subgroups or semigroups of  $e^{\mathcal{L}}$ . If they are one dimensional semigroups  $G_j := \{e^{\tilde{A}_j t} | t \geq 0\}$ ,  $j = 1, \dots, r$ , and there are values  $\tilde{u}_j$  such that  $-iH(\tilde{u}_j) = \tilde{A}_j$ , the  $j$ -th control  $u_j$  will be equal to  $\tilde{u}_j$  for an appropriate amount of time and the resulting overall control will be piecewise constant.

There are several Lie group decompositions known in the literature and developed independently of control theory. They are an important part of Lie group theory and linear algebra. For quantum systems, Lie group decompositions are not only a tool to design controls but also a way to analyze dynamics, as will be discussed in [Chapter 8](#). The subject is very vast. For much of this chapter, we shall focus on decompositions deriving from Cartan classification of the symmetric spaces of the classical Lie groups also known as *Cartan decompositions*. We shall emphasize the computational aspects which are relevant for control purposes. Further investigations on how to generate new decompositions starting from Cartan decompositions and how to use these results for quantum dynamical analysis will be presented in Chapter 8.

This chapter is organized as follows. We shall start by discussing in detail the case of two level systems and the associated Lie group  $SU(2)$ . In this case, the decomposition used is the classical Euler's resolution of a rotation. In [section 5.2](#), we give a first example of a general decomposition of  $SU(n)$ , namely the decomposition into planar rotation, which is quite simple and can be obtained by straightforward computations. This decomposition gives a direct way to parametrize the special unitary group. In [section 5.3](#) we present the basics of Cartan decompositions of semisimple Lie groups. [Section 5.4](#) discusses Levi decomposition, which shows how every Lie algebra can be decomposed as the direct sum of a solvable Lie algebra and a semisimple one. We shall discuss the consequence of this for Lie group decompositions. Finally, [section 5.5](#) gives some more examples of applications of Lie group decompositions to control which use the theory described in the previous sections. In the treatment we shall introduce some more notions of Lie groups and Lie algebra theory which complement the ones in [Chapter 3](#) and will be further developed in Chapter 8.

## 5.1 Decompositions of $SU(2)$ and Control of Two Level Systems

### 5.1.1 The Lie groups $SU(2)$ and $SO(3)$

#### 5.1.1.1 $SU(2)$

The Lie algebra associated with  $SU(2)$  (cf. subsection 3.4.3) is the Lie algebra  $su(2)$  of  $2 \times 2$ , skew-Hermitian matrices which is spanned by the matrices (cf. the Pauli matrices in (1.20))

$$\bar{\sigma}_x := \frac{1}{2} \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad \bar{\sigma}_y := \frac{1}{2} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \bar{\sigma}_z := \frac{1}{2} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad (5.5)$$

satisfying the commutation relations

$$[\bar{\sigma}_x, \bar{\sigma}_y] = \bar{\sigma}_z, \quad [\bar{\sigma}_y, \bar{\sigma}_z] = \bar{\sigma}_x, \quad [\bar{\sigma}_z, \bar{\sigma}_x] = \bar{\sigma}_y. \quad (5.6)$$

$su(2)$  is also an inner product space where the inner product between two elements  $A$  and  $B$  can be taken given by

$$\langle A, B \rangle = \text{Tr}(AB^\dagger). \quad (5.7)$$

With this inner product  $\bar{\sigma}_{x,y,z}$  are orthogonal to each other.

#### 5.1.1.2 $SO(3)$

$SO(3)$  is the Lie group of  $3 \times 3$  orthogonal matrices with determinant equal to one. The associated Lie algebra is  $so(3)$ , namely the Lie algebra of  $3 \times 3$  skew-symmetric matrices. It is spanned by the matrices

$$S_x := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \quad S_y := \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \quad S_z := \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (5.8)$$

which satisfy the commutation relations

$$[S_x, S_y] = S_z, \quad [S_y, S_z] = S_x, \quad [S_z, S_x] = S_y. \quad (5.9)$$

Every matrix  $X$  in  $SO(3)$  acting on a vector applied at the origin  $\vec{v} \in \mathbf{R}^3$  represents a rotation of  $\vec{v}$  about an axis. Any  $X \in SO(3)$  different from the identity has exactly one eigenvalue equal to one. The corresponding eigenvector  $\vec{v}$  such that  $X\vec{v} = \vec{v}$  is the axis of rotation, namely it gives the direction of the vectors which are not modified by  $X$ . The distance of every other point from the axis of rotation is unchanged by the operation  $X$ . The axis of rotation can also be determined by writing  $X$  as  $X = e^A$ , with  $A \in$

$so(3)$ , which is always possible,<sup>2</sup> and then finding a vector  $\vec{v} \in \mathbf{R}^3$  such that  $A\vec{v} = 0$ . Consider  $S_z$  in (5.8);  $e^{S_z t}$  is given by

$$e^{S_z t} = \begin{pmatrix} \cos(t) & \sin(t) & 0 \\ -\sin(t) & \cos(t) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

It rotates a vector clockwise about the  $z$  axis by an angle of  $t$  radians. Analogously,  $e^{S_x t}$  and  $e^{S_y t}$  rotate a vector about the  $x$  (clockwise) and  $y$  (counter-clockwise) axes.

### 5.1.1.3 Relation between $SU(2)$ and $SO(3)$

It follows from the commutation relations (5.6), (5.9) that the map  $\phi : su(2) \rightarrow so(3)$  defined as

$$\phi(\bar{\sigma}_x) = S_x, \quad \phi(\bar{\sigma}_y) = S_y, \quad \phi(\bar{\sigma}_z) = S_z \quad (5.10)$$

is an isomorphism (cf. Definition 3.1.3).<sup>3</sup> This isomorphism induces a homomorphism (cf. Definition 3.1.6)  $\Phi : SU(2) \rightarrow SO(3)$  which is defined as follows. For every matrix in  $X \in SU(2)$  there exists a matrix  $A \in su(2)$  such that  $X = e^A$ . If  $X = e^A$ , then

$$\Phi(X) = \Phi(e^A) := e^{\phi(A)}. \quad (5.11)$$

This map does not depend on the choice of  $A$  and it is a homomorphism. Moreover it is an onto and two-to-one map in that, for any  $S \in SU(2)$ ,  $\Phi(-S) = \Phi(S)$  and  $S$  and  $-S$  are the only matrices giving the value  $\Phi(S)$ . For the proofs of these facts see, e.g., [53], [172], [204].

### 5.1.2 Euler decomposition of $SU(2)$ and $SO(3)$

It is a well known fact of classical mechanics, and it is discussed in detail for example in [15], [185], that a rotation about an arbitrary axis can be decomposed into three rotations, the first one about the  $z$  axis, the second one about the  $y$  axis and the third one about the  $z$  axis, in that order. This means that every matrix  $X$  in  $SO(3)$  can be written as

$$X = e^{S_z t_3} e^{S_y t_2} e^{S_z t_1}, \quad (5.12)$$

where the parameters  $t_{1,2,3}$  can be chosen nonnegative and represent the angles in radians of the corresponding rotations. The decomposition (5.12) for  $SO(3)$

<sup>2</sup>The exponential map is surjective for compact connected Lie groups such as  $SU(n)$  and  $SO(n)$  (cf. [184]).

<sup>3</sup>Therefore both  $su(2)$  and  $so(3)$  are isomorphic to  $\mathbf{R}^3$  with the cross product operation, cf. Example 3.1.4.

is called **Euler's decomposition** and the parameters  $t_1$ ,  $t_2$  and  $t_3$  are called *Euler angles*. If  $\tilde{X} \in SU(2)$  is such that  $\Phi(\tilde{X}) = X$ , with  $\Phi$  in (5.11) and  $X$  given in (5.12), then  $\tilde{X}$  is given by

$$\tilde{X} = e^{\bar{\sigma}_z t_3} e^{\bar{\sigma}_y t_2} e^{\bar{\sigma}_z t_1}, \quad (5.13)$$

or

$$\tilde{X} = -e^{\bar{\sigma}_z t_3} e^{\bar{\sigma}_y t_2} e^{\bar{\sigma}_z t_1} = e^{\bar{\sigma}_z \pi} e^{\bar{\sigma}_z t_3} e^{\bar{\sigma}_y t_2} e^{\bar{\sigma}_z t_1} = e^{\bar{\sigma}_z(t_3 + \pi)} e^{\bar{\sigma}_y t_2} e^{\bar{\sigma}_z t_1}. \quad (5.14)$$

In order to see this, apply  $\Phi$  to both sides of

$$L := e^{\bar{\sigma}_z t_3} e^{\bar{\sigma}_y t_2} e^{\bar{\sigma}_z t_1}.$$

Using the definition of  $\Phi$  in (5.11) and (5.10) we obtain

$$\Phi(L) = X.$$

Therefore, from the properties of the homomorphism  $\Phi$ , we have  $\tilde{X} = \pm L$ , which gives formulas (5.13) and (5.14). The formula (5.13) is the *Euler decomposition of  $SU(2)$* . The parameters  $t_{1,2,3}$  can be chosen nonnegative and are called *Euler's angles*, as well.

**Remark 5.1.1** Notice that the Pauli matrices (5.5) can be transformed one into the other by a similarity transformation according to  $\bar{\sigma}_x \rightarrow \bar{\sigma}_y \rightarrow \bar{\sigma}_z \rightarrow \bar{\sigma}_x$ . Therefore, Euler decomposition could have been stated by using rotation about the  $z$  and  $x$  axes or  $x$  and  $y$  axes. In fact, every two rotations about two orthogonal axes could have been used (see also [subsection 5.1.4](#) for detailed calculations).

### 5.1.3 Determination of the angles in the Euler decomposition of $SU(2)$

Given a matrix  $X$  in  $SU(2)$ , we show in this subsection how to determine the Euler's angles. Notice that, from the above discussion, this also gives Euler's angles for the image of  $X$  in  $SO(3)$  under  $\Phi$ , which can also be determined independently (see, e.g., [172] Chpt. 8). An equivalent method will follow from the general methods for determination of Cartan decompositions given in [section 5.3](#) below. In fact we shall see that Euler's decomposition is a special case of Cartan decomposition.

An arbitrary matrix  $X \in SU(2)$  is written as

$$X : \begin{pmatrix} M e^{i\phi} & N e^{i\psi} \\ G e^{i\eta} & T e^{i\gamma} \end{pmatrix}.$$

Since  $M^2 + N^2 = 1$ ,  $M^2 + G^2 = 1$ , and  $N^2 + T^2 = 1$ , we can rewrite  $X$  as

$$X = \begin{pmatrix} M e^{i\phi} & \sqrt{1 - M^2} e^{i\psi} \\ -\sqrt{1 - M^2} e^{i\eta} & M e^{i\gamma} \end{pmatrix}. \quad (5.15)$$

The signs are chosen arbitrarily since we can always change them by modifying the phases in the exponentials in (5.15). In particular, we can set  $M = \cos(\theta)$ , with  $\theta \in [0, \pi]$  so that  $\sqrt{1 - M^2} = \sin(\theta)$ . Therefore we have

$$X = \begin{pmatrix} \cos(\theta)e^{i\phi} & \sin(\theta)e^{i\psi} \\ -\sin(\theta)e^{i\eta} & \cos(\theta)e^{i\gamma} \end{pmatrix}.$$

The condition  $\det(X) = 1$  gives

$$\det(X) = \cos^2(\theta)e^{i(\phi+\gamma)} + \sin^2(\theta)e^{i(\psi+\eta)} = 1,$$

which implies  $\phi + \gamma = 0$  and  $\psi + \eta = 0$  if  $\theta \notin \{0, \frac{\pi}{2}, \pi\}$ . We can always choose  $\gamma = -\phi$  and  $\eta = -\psi$  because in the case  $\theta = 0$  and  $\theta = \pi$ , the values of  $\psi$  and  $\eta$  do not matter (the phase factors are multiplied by zero) and analogously when  $\theta = \pi$ , the values of  $\gamma$  and  $\phi$  do not matter. In conclusion a general matrix  $X \in SU(2)$  can be written as

$$X = \begin{pmatrix} \cos(\theta)e^{i\phi} & \sin(\theta)e^{i\psi} \\ -\sin(\theta)e^{-i\psi} & \cos(\theta)e^{-i\phi} \end{pmatrix}, \quad (5.16)$$

with  $\theta \in [0, \pi]$  and  $\phi, \psi \in [0, 2\pi)$ . Now notice that

$$e^{\bar{\sigma}_y t} = \begin{pmatrix} \cos(\frac{t}{2}) & \sin(\frac{t}{2}) \\ -\sin(\frac{t}{2}) & \cos(\frac{t}{2}) \end{pmatrix}, \quad e^{\bar{\sigma}_z t} = \begin{pmatrix} e^{i(\frac{t}{2})} & 0 \\ 0 & e^{-i(\frac{t}{2})} \end{pmatrix}. \quad (5.17)$$

Using (5.17) and (5.16), and choosing

$$\frac{t_2}{2} = \theta, \quad \frac{t_1}{2} = \frac{\phi - \psi}{2}, \quad \frac{t_3}{2} = \frac{\phi + \psi}{2},$$

a straightforward computation shows that

$$e^{\bar{\sigma}_z t_3} e^{\bar{\sigma}_y t_2} e^{\bar{\sigma}_z t_1} = X, \quad (5.18)$$

with  $X$  in (5.16). Notice that it is always possible to choose  $t_{1,2,3} \geq 0$ , since we can always add to them multiples of  $4\pi$  without modifying the product in (5.18).

#### 5.1.4 Application to the control of two level quantum systems

Consider a two level quantum control system whose Schrödinger operator equation reads as

$$\dot{X} = (u_z \bar{\sigma}_z + u_y \bar{\sigma}_y)X, \quad (5.19)$$

with  $u_z$  and  $u_y$  two independent controls. This could represent, in appropriate units, the evolution of a spin  $\frac{1}{2}$  particle where  $u_z$  and  $u_y$  are the  $z$  and  $y$  components of the electromagnetic field, and the  $x$ -component is set to be

identically equal to zero (cf. subsection 2.3.2). Assume the problem of control is to drive the state from the identity to a target  $X_f \in SU(2)$ , and assume we have calculated Euler's decomposition of  $X_f$ ,

$$X_f = e^{\bar{\sigma}_z t_3} e^{\bar{\sigma}_y t_2} e^{\bar{\sigma}_z t_1},$$

with  $t_1, t_2, t_3 \geq 0$ . Then a control law  $u_z \equiv 1, u_y \equiv 0$ , for time  $t_1$ , followed by  $u_z \equiv 0, u_y \equiv 1$  for time  $t_2$ , followed by  $u_z \equiv 1, u_y \equiv 0$ , for time  $t_3$  will drive the identity to  $X_f$ , in time  $t_1 + t_2 + t_3$ .

The above two level system (5.19) has a very special structure which itself suggests the use of Euler's decomposition. However, the idea can be applied to much more general two level systems

$$\dot{X} = -iH(u)X. \quad (5.20)$$

The only requirement is that there exist two values of the control say  $u_1$  and  $u_2$  such that  $Z_1 := -iH(u_1)$ , and  $Z_2 := -iH(u_2)$  are *orthogonal*. If this is the case, let  $T_1$  be a unitary matrix which diagonalizes  $Z_1$  (which always exists since  $Z_1$  is skew-Hermitian). We have

$$T_1 Z_1 T_1^\dagger = \lambda \bar{\sigma}_z,$$

for some  $\lambda \in \mathbb{R}$ , and

$$T_1 Z_2 T_1^\dagger = a \bar{\sigma}_y + b \bar{\sigma}_x,$$

for some real coefficients  $a$  and  $b$ .  $T_1 Z_2 T_1^\dagger$  has zero component along  $T_1 Z_1 T_1^\dagger$  and therefore zero coefficient for  $\bar{\sigma}_z$ . A further unitary transformation

$$T_2 := \begin{pmatrix} 1 & 0 \\ 0 & \frac{a+ib}{\sqrt{a^2+b^2}} \end{pmatrix}$$

is such that

$$T_2 T_1 Z_1 T_1^\dagger T_2^\dagger = \lambda \bar{\sigma}_z, \quad T_2 T_1 Z_2 T_1^\dagger T_2^\dagger = \sqrt{a^2 + b^2} \bar{\sigma}_y.$$

Therefore, in appropriate coordinates,  $Z_2$  and  $Z_1$  are proportional to  $\bar{\sigma}_z$  and  $\bar{\sigma}_y$ , respectively, and therefore an Euler factorization of any matrix exists in terms of  $Z_1$  and  $Z_2$ , which is the same as the one in terms of  $\bar{\sigma}_z$  and  $\bar{\sigma}_y$ , except that the Euler's angles  $t_{1,2,3}$  have to be divided by  $\lambda$  or  $\sqrt{a^2 + b^2}$ . The algorithm for control can be applied in this case, too. In practice, given the control system (5.20), one chooses two controls  $u_1$  and  $u_2$  to make  $-iH(u_1)$  and  $-iH(u_2)$  orthogonal, if possible. Then one determines the matrix  $T := T_2 T_1$ , where  $T_1$  and  $T_2$  are chosen as above, and the Euler decomposition of  $TX_f T^\dagger$ , where  $X_f$  is the target. From this one finds the control algorithm.

In the most common case, two level control systems have only one control. The corresponding Schrödinger operator equation is written as

$$\dot{X} = (A + Bu)X.$$

In this case the choice  $u_1 = +k$ ,  $u_2 = -k$ , with

$$k := \sqrt{\frac{\text{Tr}(A^2)}{\text{Tr}(B^2)}},$$

is a possible choice to give two orthogonal matrices. Notice the quantity  $k$  is a measure of the ‘control authority’ of the given system.

---

## 5.2 Decomposition in Planar Rotations

The decomposition of the unitary group into planar rotations was described in [151]. It provides a simple and explicit way to parametrize the group  $SU(n)$ .

A **planar rotation** in  $SU(n)$ ,  $U_{j,k}(\theta, \beta)$ ,  $j < k$ , is an  $n \times n$  matrix which is equal to the identity except for the elements at the intersection between the  $j$ -th and  $k$ -th rows and columns which are occupied by

$$\tilde{U}(\theta, \beta) := \begin{pmatrix} \cos(\theta) & -\sin(\theta)e^{-i\beta} \\ \sin(\theta)e^{i\beta} & \cos(\theta) \end{pmatrix}.$$

An important property of planar rotations is

$$[U_{j,k}(\theta, \beta)]^{-1} = U_{j,k}^\dagger(\theta, \beta) = U_{j,k}(-\theta, \beta). \quad (5.21)$$

We shall denote by  $D(\alpha_1, \alpha_2, \dots, \alpha_{n-1})$  the diagonal matrix in  $SU(n)$

$$\text{diag}(e^{i\alpha_1}, e^{i\alpha_2}, \dots, e^{i\alpha_{n-1}}, e^{-i(\sum_{l=1}^{n-1} \alpha_l)}).$$

The following theorem gives a decomposition of any special unitary matrix into planar rotations and a matrix of the form  $D(\alpha_1, \alpha_2, \dots, \alpha_{n-1})$ . The proof is constructive in that it also shows how to determine the parameters involved in the decomposition.

**Theorem 5.2.1** *For any matrix  $X$  in  $SU(n)$  there exist  $n - 1$  parameters,  $\alpha_1, \dots, \alpha_{n-1}$ ,  $\frac{n(n-1)}{2}$  parameters,  $\theta_1, \dots, \theta_{\frac{n(n-1)}{2}}$ , and  $\frac{n(n-1)}{2}$  parameters,  $\beta_1, \dots, \beta_{\frac{n(n-1)}{2}}$ , such that*

$$X = D(\alpha_1, \dots, \alpha_{n-1})U_{1,2}(\theta_1, \beta_1)U_{1,3}(\theta_2, \beta_2)U_{2,3}(\theta_3, \beta_3) \cdots \times \quad (5.22)$$

$$\begin{aligned} & U_{1,n}(\theta_{\frac{(n-1)(n-2)}{2}+1}, \beta_{\frac{(n-1)(n-2)}{2}+1})U_{2,n}(\theta_{\frac{(n-1)(n-2)}{2}+2}, \beta_{\frac{(n-1)(n-2)}{2}+2}) \cdots \times \\ & \qquad \qquad \qquad U_{n-1,n}(\theta_{\frac{n(n-1)}{2}}, \beta_{\frac{n(n-1)}{2}}). \end{aligned}$$

Notice the planar rotation factors on the right hand side of (5.22) are ordered into subsequences according to the second index  $k$ , starting from  $k = 2$ . The subsequence corresponding to  $k$  contains factors  $U_{1,k}, U_{2,k}, \dots, U_{k-1,k}$ , in that order.

*Proof.* Consider a matrix  $X \in SU(n)$  and a general matrix  $U_{n-1,n}(\bar{\theta}, \bar{\beta})$ . Multiplication of  $X$  on the right by  $U_{n-1,n}(\bar{\theta}, \bar{\beta})$  only affects the  $(n-1)$ -th and  $n$ -th columns of  $X$ . In particular we select  $\bar{\theta}$  and  $\bar{\beta}$  so that element  $(n, n-1)$  of  $XU_{n-1,n}(\bar{\theta}, \bar{\beta})$  is zero. Now consider multiplication of  $XU_{n-1,n}$  on the right by a planar rotation  $U_{n-2,n}(\tilde{\theta}, \tilde{\beta})$ . This only affects the  $(n-2)$ -th and  $n$ -th columns of  $XU_{n-1,n}$ . In particular, it does not affect the zero in the position  $n, n-1$  which was previously introduced. We choose  $\tilde{\theta}$  and  $\tilde{\beta}$  so that the resulting matrix  $XU_{n-1,n}U_{n-2,n}(\tilde{\theta}, \tilde{\beta})$  has a zero in the position  $n, n-2$ . Continuing this way, we multiply on the right by matrices  $U_{n-3,n}, U_{n-4,n}, \dots, U_{1,n}$  and introduce zeros in the positions  $(n, n-3), (n, n-4), \dots, (n, 1)$ , in that order. The matrix  $XU_{n-1,n}U_{n-2,n} \cdots U_{1,n}$  has the  $n$ -th row all equal to zero except for the  $(n, n)$ -th element which must have magnitude equal to one, since the matrix is unitary. This also shows that the last column is zero except for the  $n, n$ -th element. Therefore we have

$$XU_{n-1,n}U_{n-2,n} \cdots U_{1,n} = \begin{pmatrix} \tilde{X}_{(n-1) \times (n-1)} & 0 \\ 0 & e^{in} \end{pmatrix},$$

for some  $\eta \in \mathbf{R}$  and  $\tilde{X}_{(n-1) \times (n-1)}$  in  $U(n-1)$ . Now, further multiplication by factors  $U_{j,k}$ , with  $k < n$ , only affects  $\tilde{X}$  and therefore, as above, we can choose  $U_{n-2,n-1}, U_{n-3,n-1}, \dots, U_{1,n-1}$ , to make the last row and column of  $\tilde{X}_{(n-1) \times (n-1)}$  equal to zero except for the diagonal element. Continuing this way, after multiplication by  $U_{k-1,k}, U_{k-2,k}, \dots, U_{1,k}$ , for all the  $k$ 's,  $k = n, n-1, n-2, \dots, 2$ , we obtain a diagonal unitary matrix with determinant equal to one. Therefore we can write

$$\begin{aligned} XU_{n-1,n}U_{n-2,n} \cdots U_{1,n}U_{n-2,n-1}U_{n-3,n-1} \cdots U_{1,n-1} \cdots U_{2,3}U_{1,3}U_{1,2} \\ = D(\alpha_1, \alpha_2, \dots, \alpha_{n-1}). \end{aligned}$$

Multiplying this expression by  $U_{1,2}^{-1}U_{1,3}^{-1}U_{2,3}^{-1} \cdots U_{n-1,n}^{-1}$  and using (5.21) we obtain (5.22) with appropriate parameters.  $\square$

### 5.3 Cartan Decompositions

In this section we shall discuss Cartan decompositions of Lie groups. We shall start giving some definitions and facts for general Lie groups associated to semisimple Lie algebras and then focus on  $SU(n)$ . Cartan decompositions will be further studied and applied in [Chapter 8](#).

### 5.3.1 Cartan decomposition of semisimple Lie algebras

Consider a semisimple (cf. Definition 3.4.5) Lie algebra  $\mathcal{L}$  with an inner product given by the Killing form<sup>4</sup> which in the case of  $su(n)$  can be taken equal to the trace inner product defined in (5.7). A decomposition of  $\mathcal{L}$  of the form

$$\mathcal{L} = \mathcal{K} \oplus \mathcal{P} \quad (5.23)$$

with  $\mathcal{P} := \mathcal{K}^\perp$  is called a **Cartan decomposition** of  $\mathcal{L}$  if the following commutation relations are verified:

$$[\mathcal{K}, \mathcal{K}] \subseteq \mathcal{K}, \quad (5.24)$$

i.e.,  $\mathcal{K}$  is a subalgebra of  $\mathcal{L}$ ,

$$[\mathcal{K}, \mathcal{P}] \subseteq \mathcal{P}, \quad (5.25)$$

and

$$[\mathcal{P}, \mathcal{P}] \subseteq \mathcal{K}. \quad (5.26)$$

Associated with a Cartan decomposition (5.23) is a **Cartan involution**, a Lie algebra isomorphism  $\theta : \mathcal{L} \rightarrow \mathcal{L}$  which is equal to the identity on  $\mathcal{K}$  and multiplies by  $-1$  the elements of  $\mathcal{P}$ , i.e.,

$$\theta(K) = K, \quad \forall K \in \mathcal{K}, \quad (5.27)$$

$$\theta(P) = -P, \quad \forall P \in \mathcal{P}. \quad (5.28)$$

More specifically, given a Cartan decomposition, relations (5.27) and (5.28) determine a Cartan involution  $\theta$  and conversely given a Cartan involution  $\theta$ , the  $+1$  and  $-1$  eigenspaces of  $\theta$  determine a Cartan decomposition (cf. Exercise 5.8). Cartan involutions for the Lie algebra  $su(n)$  play an important role in the analysis of quantum dynamics as they have a direct correspondence with quantum symmetries as we shall see below.

### 5.3.2 The decomposition theorem for Lie groups

The importance of a Cartan decomposition for control purposes lies in the following theorem whose proof can be found in standard Lie algebra and Lie group textbooks such as [96], [97].

---

<sup>4</sup>If  $ad_x$  is the matrix representation of the linear operator  $L \rightarrow [x, L]$  for a given  $x \in \mathcal{L}$ , the Killing form  $\langle x, y \rangle_K$  is defined as  $\langle x, y \rangle_K := \text{Tr}(ad_x ad_y)$ . In the case of  $su(n)$ , the Killing form is proportional to  $\text{Tr}(xy^\dagger)$  and therefore we can take this as the inner product. See, e.g., [96] for further details. See also [Appendix C](#).

**Theorem 5.3.1** Given a Cartan decomposition of a semisimple Lie algebra  $\mathcal{L}$  as above, every element  $X \in e^{\mathcal{L}}$  can be written as

$$X = PK, \quad (5.29)$$

where  $P$  is the exponential of an element of  $\mathcal{P}$  and  $K$  is an element of the Lie group corresponding to  $\mathcal{K}$ , i.e.,  $e^{\mathcal{K}}$ .

The coset space (cf. subsection 3.4.2), associated with the Cartan decomposition, is  $e^{\mathcal{L}}/e^{\mathcal{K}}$ . It is called the *globally Riemannian symmetric space*. Cartan classified all the symmetric spaces of the *classical Lie groups*, i.e.,  $SU(n)$ ,  $Sp(n)$  and  $SO(n)$  (see [42], [43], [96]). We shall review the results for  $SU(n)$  in the following subsection.

**Example 5.3.2** Consider the decomposition of  $su(n)$

$$su(n) = so(n) \oplus \mathcal{I}, \quad (5.30)$$

where  $\mathcal{I}$  is the vector space in  $su(n)$  of purely imaginary matrices and clearly  $\mathcal{I} = so(n)^\perp$ . It is obvious that the following commutation relations hold

$$[so(n), so(n)] \subseteq so(n),$$

$$[so(n), \mathcal{I}] \subseteq \mathcal{I},$$

$$[\mathcal{I}, \mathcal{I}] \subseteq so(n),$$

and therefore the decomposition (5.30) is a Cartan decomposition of  $su(n)$  of the form (5.23), where the role of  $\mathcal{K}$  is played by  $so(n)$  and the role of  $\mathcal{P}$  is played by  $\mathcal{I}$ . An application of Theorem 5.3.1 says that any matrix  $X \in SU(n)$ , i.e., any  $n \times n$  unitary matrix with determinant equal to one, can be written as

$$X = PK,$$

where  $P$  is the exponential of a skew-Hermitian purely imaginary matrix and  $K$  is a matrix in  $SO(n)$ , namely an  $n \times n$  orthogonal matrix with determinant equal to one.

### 5.3.3 Refinement of the decomposition; Cartan subalgebras

Consider now a subalgebra  $\mathcal{A}$  of  $\mathcal{L}$  which is a subspace of  $\mathcal{P}$ . Since  $[\mathcal{A}, \mathcal{A}] \subseteq \mathcal{A} \subseteq \mathcal{P}$  and  $[\mathcal{A}, \mathcal{A}] \subseteq [\mathcal{P}, \mathcal{P}] \subseteq \mathcal{K}$ , we have  $[\mathcal{A}, \mathcal{A}] = \{0\}$ , i.e.,  $\mathcal{A}$  has to be necessarily Abelian. A maximal Abelian subalgebra in  $\mathcal{P}$  (i.e., an Abelian subalgebra such that any element of  $X \in \mathcal{P}$ ,  $X \notin \mathcal{A}$  is such that  $[X, \mathcal{A}] \neq \{0\}$ ) is called a **Cartan subalgebra** of  $\mathcal{L}$ . Its dimension is called the **rank** of the symmetric space  $e^{\mathcal{L}}/e^{\mathcal{K}}$ .

**Example 5.3.3** In Example (5.3.2), for the case of  $su(n)$ , a Cartan subalgebra is spanned by the set of diagonal matrices and therefore the rank of the associated symmetric space is  $n - 1$ . There is no larger set of commuting linearly independent matrices, since every set of mutually commuting skew-Hermitian matrices are simultaneously diagonalizable (cf., e.g., [99]).

The following theorem [96] will allow us to refine the Cartan decomposition for the Lie group  $e^{\mathcal{L}}$ .

**Theorem 5.3.4** *Let  $\mathcal{A}$  be a Cartan subalgebra relative to the decomposition  $\mathcal{L} = \mathcal{K} \oplus \mathcal{P}$ . Then*

$$\bigcup_{K \in e^{\mathcal{K}}} K\mathcal{A}K^{-1} = \mathcal{P}.$$

We also have [96].

**Theorem 5.3.5** *Given two Cartan subalgebras in  $\mathcal{P}$ ,  $\mathcal{A}$  and  $\mathcal{A}'$ , there exists a  $K \in e^{\mathcal{K}}$  such that*

$$K\mathcal{A}K^{-1} = \mathcal{A}'.$$

Theorems 5.3.4 and 5.3.5 say that every element of  $\mathcal{P}$  belong to a Cartan subalgebra and that all the Cartan subalgebras are conjugate via an element  $K$  of  $e^{\mathcal{K}}$ .

**Example 5.3.6** For the Cartan decomposition of  $su(n)$  treated in Example 5.3.2, Theorem 5.3.4 says that every purely imaginary matrix  $\tilde{P}$  can be written as

$$\tilde{P} = K\tilde{A}K^T,$$

where  $\tilde{A}$  is a skew-Hermitian purely diagonal matrix and  $K$  is a special orthogonal matrix ( $\in SO(n)$ ). In other words, every skew-Hermitian purely imaginary matrix can be diagonalized using an orthogonal matrix.

**Theorem 5.3.7** *Consider a semisimple Lie algebra  $\mathcal{L}$  with a Cartan decomposition (5.23). Then, every element  $X \in e^{\mathcal{L}}$  can be written as<sup>5</sup>*

$$X = K_1 A K_2, \tag{5.31}$$

where  $K_1$  and  $K_2$  are elements of  $e^{\mathcal{K}}$  and  $A$  is an element of  $e^{\mathcal{A}}$ .

---

<sup>5</sup>We shall refer to the decomposition (5.31) as the Cartan  $KAK$  decomposition while (5.32) is called the Cartan  $PK$  decomposition.

*Proof.* If  $X \in e^{\mathcal{L}}$ , we know from Theorem 5.3.1 that

$$X = PK, \quad (5.32)$$

$P = e^{\tilde{P}}$ ,  $\tilde{P} \in \mathcal{P}$  and  $K \in e^{\mathcal{K}}$ . Using Theorem 5.3.4 we can write  $P$  as  $P = \tilde{K}e^{\tilde{A}}\tilde{K}^{-1}$ , with  $\tilde{K} \in e^{\mathcal{K}}$  and  $\tilde{A} \in \mathcal{A}$  which in (5.32) gives

$$X = \tilde{K}e^{\tilde{A}}\tilde{K}^{-1}K,$$

which is (5.31) with  $K_1 = \tilde{K}$ ,  $K_2 = \tilde{K}^{-1}K$  and  $A = e^{\tilde{A}}$ .  $\square$

**Example 5.3.8** For Example 5.3.2, Theorem 5.3.7 says that every special unitary matrix  $X \in SU(n)$  can be written as in (5.31) where  $K_1$  and  $K_2$  are orthogonal matrices and  $A$  is a diagonal matrix of the type

$$A := \text{diag}(e^{i\alpha_1}, e^{i\alpha_2}, \dots, e^{i\alpha_{n-1}}, e^{-i\sum_{k=1}^{n-1} \alpha_k}).$$

The theorem is merely an existence theorem and does not address the problem of finding the factors in (5.31). Several algorithms of numerical linear algebra exist to calculate the factors and in low dimensional cases the factors can be calculated by elementary matrix manipulations. For example, the Euler decomposition and the methods to calculate it illustrated in the previous section are a special case of a Cartan decomposition, as discussed in the following example.

**Example 5.3.9** In the special case of  $su(2)$ , a Cartan decomposition is

$$su(2) = so(2) \oplus so(2)^\perp,$$

with

$$so(2) := \text{span}\{\bar{\sigma}_y\}, \quad so(2)^\perp := \text{span}\{\bar{\sigma}_x, \bar{\sigma}_z\}$$

(cf. (5.5)). If we choose the Cartan subalgebra  $\mathcal{A} := \text{span}\{\bar{\sigma}_z\}$ , then the decomposition (5.31) reads as

$$X = e^{\bar{\sigma}_y t_3} e^{\bar{\sigma}_z t_2} e^{\bar{\sigma}_y t_1}$$

as all the Lie groups involved are one-dimensional. This is an Euler decomposition as  $\bar{\sigma}_y$  and  $\bar{\sigma}_z$  are orthogonal to each other.

### 5.3.4 Cartan decompositions of $su(n)$

A consequence of Cartan classification of the symmetric spaces of  $SU(n)$  is the classification of all the possible Cartan decompositions for the Lie algebra  $su(n)$  up to conjugation. In fact, up to conjugation, there are only three types of decompositions, denoted by **AI**, **AII**, **AIII**. The corresponding subalgebras  $\mathcal{K}$  are (conjugate to) the following:

**AI**

$$\mathcal{K} = so(n),$$

**AII**

$$\mathcal{K} = sp\left(\frac{n}{2}\right),$$

if  $n$  is even,**AI****III**

$$\mathcal{K} = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}, \quad (5.33)$$

with  $A \in u(p)$ ,  $B \in u(q)$ , with  $p + q = n$  and  $\text{Tr}(A) + \text{Tr}(B) = 0$ .

The subspace  $\mathcal{P}$  in all the cases can be found as the orthogonal complement of  $\mathcal{K}$  in  $su(n)$ . The Cartan subalgebras  $\mathcal{A}$  are as follows. In the case **AI**,  $\mathcal{A}$  is (conjugate via an element of  $SO(n)$  to) the subalgebra of diagonal matrices. Therefore this decomposition has rank  $n - 1$ . In the case **AII**,  $\mathcal{A}$  is (conjugate via an element of  $Sp(\frac{n}{2})$  to) the subalgebra of repeat diagonal matrices, i.e., the matrices of the form

$$A = \begin{pmatrix} D & 0 \\ 0 & D \end{pmatrix}, \quad (5.34)$$

with  $D \in su(\frac{n}{2})$  diagonal. Therefore the rank of this decomposition is  $\frac{n}{2} - 1$ . In the case **AI****III**,  $\mathcal{A}$  is (conjugate via an element of the Lie group of 2-block diagonal special unitary matrices (cf. (5.33)) to) the subalgebra of matrices of the form

$$A = \begin{pmatrix} 0 & B \\ -B^T & 0 \end{pmatrix}.$$

Here  $B$  is a real  $p \times q$  matrix which is zero everywhere except for the first  $p$  columns if  $p \leq q$ , or the first  $q$  rows if  $p > q$ , which are occupied by a  $p \times p$ , or respectively, a  $q \times q$ , diagonal matrix. Therefore the rank of this decomposition is  $\min\{p, q\}$ . In the following, for this decomposition we shall assume, without loss of generality,  $p \leq q$ .

The Cartan involution in the case **AI**, which we denote by  $\theta_I$ , has as the  $+1$  ( $-1$ ) eigenspace, up to a unitary similarity transformation  $T$ ,  $so(n)$  ( $so(n)^\perp$ ). Therefore, in the standard case, where  $T$  is taken to be the identity,  $\theta_I$  is given by complex conjugation. In the general case, it has the form

$$\theta_I(A) := TT^T \bar{A}(TT^T)^\dagger, \quad (5.35)$$

for a given unitary matrix  $T$ , and  $\bar{A}$  denotes the complex conjugate of the matrix  $A$ . Analogously, the Cartan involution in the case **AII**,  $\theta_{II}$ , is given by

$$\theta_{II}(A) := TJT^T \bar{A}(TJT^T)^\dagger, \quad (5.36)$$

for some unitary matrix  $T$ , where  $J$  is defined in equation (3.18). In the special case where  $T$  is the identity, it is easily seen that the matrices spanning the  $+1$

eigenspace of  $\theta_{II}$  are the skew-Hermitian matrices satisfying equation (3.22), i.e., the matrices in the symplectic Lie algebra  $sp(\frac{n}{2})$ . In the case **III**, the involution  $\theta_{III}$  has the form

$$\theta_{III}(A) = TI_{p,q}T^\dagger A(TI_{p,q}T^\dagger)^\dagger, \quad (5.37)$$

for some unitary matrix  $T$ , where  $I_{p,q}$  is the matrix

$$I_{p,q} = \begin{pmatrix} \mathbf{1}_{p \times p} & 0 \\ 0 & -\mathbf{1}_{q \times q} \end{pmatrix},$$

with  $\mathbf{1}_{p \times p}$  and  $\mathbf{1}_{q \times q}$  the  $p \times p$  and  $q \times q$  identity matrices, respectively, and  $p+q = n$ . In the special case where  $T$  is the identity matrix the matrices in the +1 eigenspace of  $\theta_{III}$  are the matrices in the subalgebra of the skew-Hermitian matrices defined in (5.33).

### 5.3.5 Cartan involutions of $su(n)$ and quantum symmetries

#### 5.3.5.1 Discrete quantum symmetries

Given a quantum system with underlying Hilbert space  $\mathcal{H}$ , a **quantum mechanical discrete symmetry** can be defined as a map  $\Theta$ ,  $\Theta : \mathcal{H} \rightarrow \mathcal{H}$ , of the form

$$\Theta := e^{i\phi}U,$$

where  $\phi$  is a constant, physically irrelevant, real parameter, and  $U$  is either a *unitary* operator or an *anti-unitary* one. An anti-unitary operator  $U$ ,  $|\alpha\rangle \rightarrow |\tilde{\alpha}\rangle := U|\alpha\rangle$  is defined as satisfying

$$\langle \tilde{\beta} | \tilde{\alpha} \rangle = \langle \beta | \alpha \rangle^*,$$

$$U(c_1|\alpha\rangle + c_2|\beta\rangle) = c_1^*U|\alpha\rangle + c_2^*U|\beta\rangle.$$

Once a basis of the Hilbert space  $\mathcal{H}$  is chosen, an anti-unitary operator  $U$  can always be written as

$$U|\alpha\rangle = X|\bar{\alpha}\rangle, \quad (5.38)$$

where  $|\bar{\alpha}\rangle$  is the operation which conjugates all the components of the vector  $|\alpha\rangle$  and  $X$  is unitary.

A symmetry  $\Theta$ , whether unitary or anti-unitary, induces a linear transformation on the space of Hermitian operators  $A$  as

$$A \rightarrow \Theta A \Theta^{-1} := \tilde{\theta}(A). \quad (5.39)$$

It is in fact easily verified that  $\tilde{\theta}(A)$  is a linear and Hermitian operator. Moreover the eigenvalues of  $\tilde{\theta}(A)$  are the same as those of  $A$  and a set of orthonormal eigenvectors is given by  $\{\Theta|\alpha_j\rangle\}$ , where  $\{|\alpha_j\rangle\}$  is an orthonormal basis of eigenvectors of  $A$ .

Description of the symmetry  $\Theta$  is usually done by specifying how  $\tilde{\theta}$  acts on Hermitian operators rather than how  $\Theta$  acts on states. This is because Hermitian operators represent physical observables and therefore the action of  $\tilde{\theta}$  on observables is suggested by physical considerations. For example, the *parity* or *space inversion symmetry* is defined such that

$$\tilde{\theta}(\hat{x}) = -\hat{x},$$

where  $\hat{x}$  denotes the position operator. On the other hand, specification of  $\tilde{\theta}$  on an irreducible set of observables uniquely determines  $\Theta$  up to a phase factor [82]. An irreducible set of observables  $\{A_j\}$  is defined such that, if an observable  $B$  commutes with all of the  $\{A_j\}$ 's, then  $B$  is a multiple of the identity.

### 5.3.5.2 The Jordan algebra $iu(n)$

Consider now the vector space of Hermitian matrices of dimension  $n$ , which represent observables for an  $n$  dimensional quantum system. This vector space (over the reals) will be denoted by  $iu(n)$  as it may be obtained by multiplying by the imaginary unit  $i$  all elements of  $u(n)$ . With the anticommutation operation

$$\{A, B\} := AB + BA,$$

it has the structure of a *Jordan algebra* (cf. Exercise 5.7). A homomorphism  $\phi : iu(n) \rightarrow iu(n)$  over this Jordan algebra is a linear map which preserves the multiplication in the algebra, i.e.,

$$\phi(\{A, B\}) = \{\phi(A), \phi(B)\}.$$

In particular, quantum symmetries  $\tilde{\theta}$  from Definition (5.39) are Jordan algebra homomorphisms.

### 5.3.5.3 Correspondence between quantum symmetries and Cartan involutions; Cartan decompositions of the Jordan algebra $iu(n)$

A special type of quantum symmetries  $\tilde{\theta}$  are the ones which are equal to the identity when applied twice, i.e.,  $\tilde{\theta}^2 = \mathbf{1}$ . We shall call these symmetries *Cartan involutory symmetries* or simply *Cartan symmetries* as they are in one-to-one correspondence with the Cartan involutions on  $u(n)$  (see formula (5.40) below). Moreover, they give decompositions of the Jordan algebra  $iu(n)$  corresponding to Cartan decompositions of  $u(n)$ . In particular, let  $i\mathcal{P}$  and  $i\mathcal{K}$ , respectively, be the  $+1$  and  $-1$  eigenspaces of  $\tilde{\theta}$  in  $iu(n)$ , where  $\mathcal{P}$  and  $\mathcal{K}$  are subspaces of  $u(n)$ . From the properties of  $\tilde{\theta}$ , it is easily seen that the following anticommutation relations are verified (cf. (5.24)-(5.26))

$$\{i\mathcal{P}, i\mathcal{P}\} \subseteq i\mathcal{P}, \quad \{i\mathcal{K}, i\mathcal{P}\} \subseteq i\mathcal{K}, \quad \{i\mathcal{K}, i\mathcal{K}\} \subseteq i\mathcal{P}.$$

If we consider the homomorphism  $\theta$  on  $u(n)$  defined by

$$\theta(A) := i\tilde{\theta}(iA), \quad (5.40)$$

it is easy to verify that this is a Cartan involution and that  $\mathcal{K}$  and  $\mathcal{P}$  are the  $+1$  and  $-1$  eigenspaces of  $\theta$  and therefore they satisfy the commutation relations (5.24)-(5.26). Conversely, given a Cartan involution  $\theta$  and a corresponding Cartan decomposition of  $u(n)$ , formula (5.40) defines a Cartan quantum symmetry and a corresponding (Cartan) decomposition of the Jordan algebra  $iu(n)$ .

Consider now an involution  $\theta$  of  $u(n)$  and  $\mathcal{K}$  and  $\mathcal{P}$  the corresponding  $+1$  and  $-1$  eigenspaces. It is easily seen that the element  $i\mathbf{1}$  must have component equal to zero in one of the two eigenspaces, i.e., it must be either an element of  $\mathcal{K}$  or an element of  $\mathcal{P}$ .<sup>6</sup> If we call  $\tilde{\mathcal{K}}$  and  $\tilde{\mathcal{P}}$  the orthogonal complements of  $\text{span}\{i\mathbf{1}\}$  in  $\mathcal{K}$  and  $\mathcal{P}$ , respectively,  $\tilde{\mathcal{K}}$  and  $\tilde{\mathcal{P}}$  are the  $+1$  and  $-1$  eigenspaces of the restriction of  $\theta$  to  $su(n)$ , which is a Cartan involution on  $su(n)$ . In conclusion, up to  $\theta(i\mathbf{1})$  which may be  $+i\mathbf{1}$  or  $-i\mathbf{1}$ , the Cartan involution of  $u(n)$  must be of the type **AI**, **AII** or **AIII**.

In conclusion, to every Cartan decomposition (involution) of the Lie algebra  $u(n)$  there corresponds a Cartan decomposition (symmetry) of the Jordan algebra  $iu(n)$ .

Tables 5.1 and 5.2 summarize the Cartan involutions of  $u(n)$  and the corresponding Cartan symmetries of the various types. In order to be consistent with the involutions defined in (5.35), (5.36) and (5.37) we have included  $\text{span}\{i\mathbf{1}\}$  in the  $-1$  eigenspace of the involution in the cases **AI** and **AII** and in the  $+1$  eigenspace for the case **AIII**. In the tables 5.1,5.2, the matrix  $T$  represents a general unitary matrix, and the orthogonal complement is taken in  $u(n)$  or  $iu(n)$

### 5.3.6 Computation of the factors in the Cartan decompositions of $SU(n)$

There exist several linear algebra algorithms to calculate the factors in the Cartan  $KAK$  decomposition (5.31). We shall summarize here some of the main ideas, omitting the details concerning numerical implementation. References can be found in [36], [37]. These algorithms treat the decompositions of the various types, **AI**, **AII** and **AIII**, in the standard form. This gives for example  $su(n) = so(n) \oplus so(n)^\perp$  for the type **AI**. Therefore, given a decomposition of  $su(n)$  the first task is to find a similarity transformation  $T$  (the matrix  $T$  in the tables 5.1,5.2) to put the two subspaces  $\mathcal{K}$  and  $\mathcal{P}$  in (5.23) in the standard form. Then, once the factors in the decomposition are

---

<sup>6</sup>In order to see this, write  $i\mathbf{1}$  as the sum of two elements, one in  $\mathcal{K}$  and one in  $\mathcal{P}$ . From the commutation relations it follows that both these components must commute with all of  $u(n)$ ; therefore they must both be multiples of the identity.

**TABLE 5.1:** Cartan Involutions and Symmetries for the Various Types of Symmetric Spaces:  
 Up to a conjugacy  $T$ , which is a general unitary matrix, there are three types of Cartan decompositions of the Lie algebra  $u(n)$  and of the Jordan algebra  $iu(n)$ . In the table  $\theta$  is the Cartan involution,  $\tilde{\theta}$  the corresponding Cartan symmetry acting on  $iu(n)$  (i.e., acting on observables) and  $\Theta$  is the corresponding quantum symmetry on states.

Type	$\theta$	$\tilde{\theta}$	$\Theta$
<b>AI</b>	$\theta_I(A) := TT^T \bar{A}(TT^T)^\dagger$	$\tilde{\theta}_I(B) = TT^T \bar{B}(TT^T)^\dagger$	$\Theta_I( \psi\rangle) = TT^T \bar{\psi}\rangle$
<b>AII</b>	$\theta_{II}(A) := TJT^T \bar{A}(TJT^T)^\dagger$	$\tilde{\theta}_{II}(B) = TJT^T \bar{B}(TJT^T)^\dagger$	$\Theta_{II}( \psi\rangle) = TJT^T \bar{\psi}\rangle$
<b>AIII</b>	$\theta_{III}(A) = TI_{p,q}T^\dagger A(TI_{p,q}T^\dagger)^\dagger$	$\tilde{\theta}_{III}(B) = -TI_{p,q}T^\dagger B(TI_{p,q}T^\dagger)^\dagger$	$\Theta_{III}( \psi\rangle) = TI_{p,q}T^\dagger \psi\rangle$

**TABLE 5.2:** Eigenspaces of the Cartan involution  $\theta$  and Cartan symmetry  $\tilde{\theta}$  for various types of Cartan decompositions.

Type	$\mathcal{K}$ , +1 eigenspace of $\theta$	$i\mathcal{P}$ , +1 eigenspace of $\tilde{\theta}$
<b>AI</b>	$Tso(n)T^\dagger$	$Tso(n)^\perp T^\dagger$
<b>AII</b>	$Tsp(\frac{n}{2})T^\dagger$	$Tisp(\frac{n}{2})^\perp T^\dagger$
<b>AIII</b>	$\left\{ K \in u(n) : K = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}, A \in u(p), B \in u(q), p+q=n \right\}$	$\left\{ K \in iu(n) : K = \begin{pmatrix} 0_{p \times p} & F \\ F^\dagger & 0_{q \times q} \end{pmatrix}, p+q=n \right\}$

found, the same similarity transformation  $T$  is used to calculate the factors in the original representation. The value of  $T$  is more easily obtained when the involution  $\theta$  (or equivalently, the quantum symmetry  $(\tilde{\theta}, \Theta)$ ) associated with the decomposition is available. This is illustrated in the following example for two coupled two-level systems. This example is a special case of a decomposition (the *Odd-Even Decomposition* (OED)) which will be further discussed in [Chapter 8](#).

**Example 5.3.10** Consider  $su(4)$ , i.e., the dynamical Lie algebra associated with a controllable system of two coupled two level systems and consider a Cartan decomposition of the form

$$su(4) = \mathcal{K} \oplus \mathcal{P},$$

with

$$\mathcal{K} := \text{span}\{\mathbf{1} \otimes \sigma, \sigma \otimes \mathbf{1}\}, \quad \sigma = \bar{\sigma}_{x,y,z},$$

with  $\bar{\sigma}_{x,y,z}$  defined in (5.5) and  $\mathcal{P} = \mathcal{K}^\perp$ . A simple dimension count shows that this decomposition must be of type **AI**. In particular, this means that  $\mathcal{K}$  is conjugate to  $so(4)$ . However, the matrices of  $\mathcal{K}$  are not purely real, i.e.,  $\mathcal{K} \neq so(n)$ . To find the unitary matrix  $T$  such that  $\mathcal{K} = Tso(4)T^\dagger$ , we first find the involution corresponding to the decomposition. Consider the Cartan symmetry  $\tilde{\theta}_{II}$ , in [Table 5.1](#) with  $T$  equal to the identity, on  $iu(2)$ . It is easy to verify that the  $+1$  eigenspace is  $\text{span}\{\mathbf{1}\}$  and the  $-1$  eigenspace is  $\text{span}\{i\bar{\sigma}_x, i\bar{\sigma}_y, i\bar{\sigma}_z\}$ . Consider now  $\tilde{\theta} := \tilde{\theta}_{II} \otimes \tilde{\theta}_{II}$  on  $iu(4)$ . This is a Cartan quantum symmetry on  $iu(4)$  which acts as (see Remark 5.3.11 below)

$$\tilde{\theta}(A) := \tilde{\theta}_{II} \otimes \tilde{\theta}_{II}(A) = J \otimes J \bar{A} J^{-1} \otimes J^{-1}. \quad (5.41)$$

Therefore the matrix  $T$  which gives  $Tso(4)T^\dagger = \mathcal{K}$  has to be unitary and, from Table 5.1, such that  $TT^T = J \otimes J$ . This matrix can be chosen up to an orthogonal factor. One possible choice is

$$T := \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & i & 1 & 0 \\ i & 0 & 0 & -1 \\ i & 0 & 0 & 1 \\ 0 & -i & 1 & 0 \end{pmatrix}. \quad (5.42)$$

The columns of the matrix (5.42) are often referred to as the *magic basis*.

**Remark 5.3.11** The operator  $\tilde{\theta}$  is the tensor product of two anti-linear operators  $\tilde{\theta}_{II}$  each acting on  $iu(2)$ . The vector space  $iu(4)$  is the tensor space of two copies of  $iu(2)$  defined according to the general definition in [subsection 1.1.3](#). A linear operator which is the tensor product of two operators is defined according to the rules in 1.1.3.3. In particular (cf. (1.14)), if  $\theta_1$  and  $\theta_2$  are linear on  $iu(2)$ , then for any  $A$  in  $iu(4)$ ,  $A := \sum_k \alpha_k B_k \otimes C_k$ ,  $B_k, C_k \in iu(2)$  we have

$$\theta_1 \otimes \theta_2(A) := \sum_k \alpha_k \theta_1(B_k) \otimes \theta_2(C_k).$$

If  $\theta_1$  and  $\theta_2$  are anti-linear then the definition is

$$\theta_1 \otimes \theta_2(A) := \sum_k \alpha_k^* \theta_1(B_k) \otimes \theta_2(C_k).$$

Applying this with  $\theta_1 = \theta_2 = \tilde{\theta}_{II}$ , with  $\tilde{\theta}_{II}(B) := J\bar{B}J^{-1}$ , one obtains formula (5.41).

In the rest of this section, we shall assume that we are dealing with a decomposition in the standard form. The main idea of the algorithms for the computation of the factors in Cartan decomposition is to reduce the computation of the factors in the decomposition to a structured eigenvalue problem for which there are several algorithms available [38], [72]. We show here how to obtain such a reduction and refer to the literature on structured eigenvalue problems for this particular aspect of the algorithms.

Consider first the case of the decomposition of the type **AI** which says that every unitary matrix  $U$  can be written as

$$U = O_1 D O_2, \quad (5.43)$$

where  $O_1$  and  $O_2$  are orthogonal matrices (with determinant equal to one) and  $D$  is diagonal. From (5.43) we obtain

$$U^T = O_2^T D O_1^T,$$

which combined with (5.43) gives

$$UU^T = O_1 D^2 O_1^T,$$

or equivalently

$$UU^T O_1 = O_1 D^2.$$

Therefore the problem amounts to finding the eigenvalues and a set of orthonormal real eigenvectors of the symmetric matrix  $UU^T$ . Once the matrices  $O_1$  and  $D^2$  have been found such that  $UU^T O_1 = O_1 D^2$ , then one can use (5.43) to find the factor  $O_2$ .

**Example 5.3.12** Consider the matrix  $U \in SU(3)$ ,

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{-i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{i}{\sqrt{2}} \\ 0 & i & 0 \end{pmatrix}.$$

We calculate

$$UU^T = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

This matrix has eigenvector  $\vec{v}_1 = [\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0]^T$  corresponding to eigenvalue 1, and eigenvectors  $\vec{v}_2 = [-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0]^T$  and  $\vec{v}_3 = [0, 0, 1]^T$  both corresponding to the eigenvalue -1. Moreover  $\vec{v}_{1,2,3}$  form an orthonormal basis. Therefore we have in (5.43)  $O_1 = [\vec{v}_1, \vec{v}_2, \vec{v}_3]$ . The first diagonal entry of  $D$  is a square root of 1 while the other two diagonal entries are square roots of -1. Choosing 1,  $i$  and  $-i$ , we find that (5.43) is verified with

$$O_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

In the case **AII**, the decomposition reads as

$$U = K_1 A K_2, \quad (5.44)$$

where  $K_1$  and  $K_2$  are in  $Sp(\frac{n}{2})$  and  $A$  is a repeat two-block (see (5.34)) diagonal unitary matrix. The matrices  $K_1$  and  $K_2$  satisfy equation (3.17), i.e., we have

$$K_{1,2} J K_{1,2}^T = J.$$

From (5.44), we have

$$U^T J U = K_2^T A K_1^T J K_1 A K_2. \quad (5.45)$$

Using the fact that equation (3.17) also holds true for  $K_{1,2}^\dagger$ , we get

$$K_{1,2}^T J K_{1,2} = J, \quad (5.46)$$

which placed into (5.45) gives

$$U^T J U = K_2^T J J^T A J A K_2.$$

As  $A$  is a repeat block diagonal matrix, we have  $J^T A J = A$ , which gives

$$J^T U^T J U = J^T K_2^T J A^2 K_2.$$

From (5.46), we obtain  $J^T K_2^T J = K_2^\dagger$ , which finally gives

$$J^T U^T J U = K_2^\dagger A^2 K_2.$$

Therefore the problem consists of diagonalizing, via a symplectic matrix  $K_2$ , the matrix  $J^T U^T J U$ . It is again a structured eigenvalue problem.

Finally, for the decomposition of type **AIII**, we can write the  $KAK$  decomposition (5.31) in the block form

$$\begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix} = \begin{pmatrix} K_{11} & 0 \\ 0 & K_{12} \end{pmatrix} \begin{pmatrix} C_1 & S \\ -S^T & C_2 \end{pmatrix} \begin{pmatrix} K_{21} & 0 \\ 0 & K_{22} \end{pmatrix},$$

where  $K_{jk}$ ,  $j = 1, 2$  are unitary matrices and  $C_1$  and  $C_2$  are diagonal with  $C_2 = \begin{pmatrix} C_1 & 0 \\ 0 & 1 \end{pmatrix}$  and only the first  $p$  rows of  $S$  are nonzero with  $C_1^2 - SS^T = \mathbf{1}$ . From this equation, the first block gives

$$U_{11}U_{11}^\dagger = K_{11}C_1^2K_{11}^\dagger.$$

This is again an eigenvalue problem and  $C_1^2$  and  $K_{11}$  are chosen to satisfy this equation. This determines  $C_1$  up to signs and therefore  $C_2$  as well as  $S$ , up to signs. The equation of block (1, 1) allows to find then  $K_{21}$ , the one of block (2, 1) allows to find  $K_{2,2}$  and so on. The last equation should be used to resolve ambiguity due to degeneracy of eigenvalues and or choices of the signs in the previous steps.

---

## 5.4 Levi Decomposition

### 5.4.1 Ideals and normal subgroups

**Definition 5.4.1**  $H$  is a **normal subgroup** of a group  $e^{\mathcal{L}}$  if and only if

$$khk^{-1} \in H,$$

for every  $k \in e^{\mathcal{L}}$  and every  $h \in H$ .

If  $\mathcal{I}$  is an ideal (see [Definition 3.4.4](#)) of the Lie algebra  $\mathcal{L}$  associated with the Lie group  $e^{\mathcal{L}}$ , it is in particular a subalgebra of  $\mathcal{L}$ . Therefore, the associated Lie group  $e^{\mathcal{I}}$  is a Lie subgroup of  $e^{\mathcal{L}}$ . There exists a fundamental correspondence between ideals of a Lie algebra and normal subgroups of the associated Lie group. This is expressed by the following theorem whose proof can be found for example in ([184], pg. 164).

**Theorem 5.4.2**  $\mathcal{I}$  is an ideal of  $\mathcal{L}$  if and only if  $e^{\mathcal{I}}$  is a normal subgroup of  $e^{\mathcal{L}}$ .

The previous Theorem has an immediate consequence for decompositions of the Lie group  $e^{\mathcal{L}}$ . If

$$\mathcal{L} = \mathcal{I} \oplus \mathcal{R},$$

where  $\mathcal{I}$  is an ideal and  $\mathcal{R}$  is an arbitrary complement to  $\mathcal{I}$  in  $\mathcal{L}$ , then every element  $X_f \in e^{\mathcal{L}}$  can be written as

$$X_f = BR, \tag{5.47}$$

where  $B$  is an element of  $e^{\mathcal{I}}$  and  $R$  is the product of exponentials of elements of  $\mathcal{R}$ . In particular, if  $\mathcal{R}$  is a subalgebra of  $\mathcal{L}$ , then  $R \in e^{\mathcal{R}}$ . In order to see

why (5.47) is true, recall that, by definition, every element of  $e^{\mathcal{L}}$  is a finite product of exponentials of elements in  $\mathcal{I}$  or  $\mathcal{R}$ . Now consider a (sub)product  $e^A e^C$  in the product of  $X_f$  with  $A \in \mathcal{R}$  and  $C \in \mathcal{I}$ . Then, we can write

$$e^A e^C = e^A e^C e^{-A} e^A.$$

But since  $e^{\mathcal{I}}$  is a normal subgroup, we have  $e^A e^C e^{-A} \in e^{\mathcal{I}}$ . Continuing this way, we can always assume that all the elements on the left in the expression of  $X_f$  are in  $e^{\mathcal{I}}$ . Clearly, a similar argument proves that we can write every  $X_f$  as  $X_f = RB$ .

### 5.4.2 Solvable Lie algebras

Given a Lie algebra  $\mathcal{L}$ , define inductively the following sequence of subalgebras of  $\mathcal{L}$

$$\mathcal{L}^{(0)} := \mathcal{L},$$

$$\mathcal{L}^{(k+1)} := [\mathcal{L}^{(k)}, \mathcal{L}^{(k)}].$$

The series  $\mathcal{L}^{(k)}$  is called the *derived series* of  $\mathcal{L}$ .

**Definition 5.4.3**  $\mathcal{L}$  is called **solvable** if there exists a  $k$  such that  $\mathcal{L}^{(k)} = \{0\}$ .

The prototypical examples (cf., e.g., [106]) of solvable Lie algebras are Lie algebras of upper (or lower) triangular matrices. For example, the Lie algebra of  $2 \times 2$  upper triangular matrices  $\mathcal{T}$  is spanned by the matrices

$$A := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad B := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad C := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

We have

$$\mathcal{T}^{(1)} = \text{span}\{[A, B], [A, C], [B, C]\} = \text{span}\{B\},$$

$$\mathcal{T}^{(2)} = \{0\}.$$

Notice that for a Lie algebra of upper triangular matrices, it is very easy to parametrize the corresponding Lie group as this involves the calculation of exponentials of triangular matrices.

A special case of solvable Lie algebra is an Abelian Lie algebra which is such that  $\mathcal{L}^{(1)} = \{0\}$ . A semisimple Lie algebra  $\mathcal{L}$  (cf. Definition 3.4.5) has the property that  $\mathcal{L}^{(k)}$  is equal to  $\mathcal{L}$  for every  $k \geq 0$ . Levi's theorem, in the following subsection, states that semisimple and solvable Lie algebras are the fundamental two cases in the structure of general Lie algebras.

### 5.4.3 Levi decomposition

We have the following fundamental result in the structure theory of Lie algebras (see, e.g., [213]).

**Theorem 5.4.4 (Levi decomposition)** *Every Lie algebra  $\mathcal{L}$  is the direct sum (of vector spaces) of a semisimple Lie algebra  $\mathcal{S}$  and a solvable Lie algebra  $\mathcal{R}$ , i.e.,*

$$\mathcal{L} = \mathcal{S} \oplus \mathcal{R}, \quad (5.48)$$

with  $\mathcal{R}$  an ideal in  $\mathcal{L}$ .

**Remark 5.4.5** The fact that  $\mathcal{R}$  is an ideal in  $\mathcal{L}$  along with (5.48) is sometimes expressed by saying that  $\mathcal{L}$  is the *semidirect sum* of  $\mathcal{S}$  and  $\mathcal{R}$ .

Levi decomposition is not unique. However the solvable part  $\mathcal{R}$  is unique and it is called the *solvable radical* of  $\mathcal{L}$ . The semisimple part of  $\mathcal{L}$  is called the *Levi subalgebra*. Moreover, for any pair of Levi's decompositions

$$\mathcal{L} = \mathcal{S}_1 \oplus \mathcal{R}, \quad \mathcal{L} = \mathcal{S}_2 \oplus \mathcal{R},$$

there is an isomorphism  $\mathcal{L} \rightarrow \mathcal{L}$  which maps  $\mathcal{S}_1$  to  $\mathcal{S}_2$ .

Levi's decomposition is a useful tool to obtain decompositions for a Lie group  $e^{\mathcal{L}}$ . Let  $\mathcal{L}$  satisfy (5.48). Since  $\mathcal{R}$  is an ideal, we can always write, as seen in subsection 5.4.1, for every  $X \in e^{\mathcal{L}}$

$$X = RS,$$

with  $R \in e^{\mathcal{R}}$  and  $S \in e^{\mathcal{S}}$ . Moreover the element  $S$  can be further factorized by elements of the subgroups corresponding to the simple ideals of  $\mathcal{S}$ .

The decomposition (5.48) can be calculated explicitly starting from a basis of  $\mathcal{L}$ . Algorithms can be found in [213] which also presents many other algorithms for Lie algebras, including algorithms to calculate the simple subalgebras of a semisimple Lie algebra. We collect some more notions of Lie algebra theory related to the computation of Levi decomposition in Appendix C.

## 5.5 Examples of Application of Decompositions to Control

In this section we present two examples of quantum control systems which can be controlled using the decompositions described in the previous sections.

Both these models concern a pair of two coupled spin- $\frac{1}{2}$  particles in a driving electromagnetic field. In the first case, we assume a weak interaction of the Ising type (2.68) and we apply Cartan decomposition. In the second case we consider a Heisenberg interaction (2.67), not necessarily weak, and we apply a combination of the previously described techniques.

### 5.5.1 Control of two coupled spin- $\frac{1}{2}$ particles with Ising interaction

A possible (simplified) model for two coupled spin- $\frac{1}{2}$ 's with Ising interaction (2.68) (cf. subsection 2.3.2) is given by the Schrödinger operator equation

$$\dot{X} = -i \left( J_{12} \sigma_z \otimes \sigma_z + \gamma_1 \left( \sum_{l=x,y,z} \sigma_l \otimes \mathbf{1} u_l(t) \right) + \gamma_2 \left( \sum_{l=x,y,z} \mathbf{1} \otimes \sigma_l u_l(t) \right) \right) X. \quad (5.49)$$

The Pauli matrices  $\sigma_{x,y,z}$  were defined in (1.20),  $J_{12}$  is the coupling constant which is assumed small (weak coupling) and  $\gamma_1$  and  $\gamma_2$  are the gyromagnetic ratios of spin 1 and 2, respectively. Equation 5.49 can be further simplified. If we scale the time by a factor  $J_{12}$ , we can eliminate  $J_{12}$  from the equation. Also, by scaling the controls by a factor  $\gamma_1$ , we can eliminate  $\gamma_1$  from the equations and replace  $\gamma_2$  with the ratio  $r := \frac{\gamma_2}{\gamma_1}$ . Another simplification is obtained in the case, which is common in experiments, where the  $z$  component of the magnetic field  $u_z$  is constant (cf. section 9.1). In this case one defines  $\tilde{X} = e^{i(\sigma_z \otimes \mathbf{1} + r \mathbf{1} \otimes \sigma_z) u_z t} X$ , and after redefining the controls  $u_x$  and  $u_y$  the equation for  $\tilde{X}$  is given by

$$\dot{\tilde{X}} = -i \left( \sigma_z \otimes \sigma_z + \left( \sum_{l=x,y} \sigma_l \otimes \mathbf{1} u_l(t) \right) + r \left( \sum_{l=x,y} \mathbf{1} \otimes \sigma_l u_l(t) \right) \right) \tilde{X}. \quad (5.50)$$

We shall make reference to equation (5.50), assuming  $|r| \neq 1$  by keeping in mind that, if we solve the control problem to drive  $\tilde{X}$  from the identity to a given value  $\tilde{X}_f$  in time  $T$ , the actual value for  $X(T)$  is  $X(T) = e^{-i(\sigma_z \otimes \mathbf{1} + r \mathbf{1} \otimes \sigma_z) u_z T} \tilde{X}_f$ .

The Lie algebra generated by the matrices which multiply the controls in (5.50), if  $r \neq \pm 1$ , is given by

$$\mathcal{K} := \text{span}\{i\mathbf{1} \otimes \sigma_{x,y,z}, i\sigma_{x,y,z} \otimes \mathbf{1}\}.$$

This would be the dynamical Lie algebra of the system if there were no inter-

action. The corresponding Lie group is the group<sup>7</sup>

$$e^{\mathcal{K}} = \{X_1 \otimes X_2 | X_1, X_2 \in SU(2)\}.$$

This means that if we did not have interaction we would be able to induce arbitrary transformations on the two spins separately, but no dynamics would ‘entangle’ the two spins. Since we have  $r \neq \pm 1$ , the two spins would react differently to the common magnetic field and this is sufficient to induce arbitrary transformations on them separately. Notice that if  $r = 1$ , then the two spins would interact exactly the same way with the external field. The corresponding Lie algebra of the matrices multiplying the controls is three-dimensional, and it is given by

$$\tilde{\mathcal{K}} := \text{span}\{i\mathbf{1} \otimes \sigma_{x,y,z} + i\sigma_{x,y,z} \otimes \mathbf{1}\},$$

and the corresponding Lie group  $e^{\tilde{\mathcal{K}}} = \{X \otimes X | X \in SU(2)\}$ . In this case, the two spins are perfectly equivalent, and the transformation induced on one spin is the same as the one induced on the other.

Going back to the case  $r \neq 1$ , it is easy to verify that

$$\mathcal{P} := \mathcal{K}^\perp = \text{span}\{i\sigma_l \otimes \sigma_j | l, j = x, y, z\},$$

where the orthogonal complement is taken in  $su(4)$ . The interaction matrix is in  $\mathcal{P}$ , and system (5.50) is controllable since the dynamical Lie algebra turns out to be equal to  $su(4)$ . Moreover, we have that the conditions (5.24)-(5.26) are verified and therefore the decomposition

$$su(4) = \mathcal{K} \oplus \mathcal{P},$$

is a Cartan decomposition of  $su(4)$ . It is in fact the decomposition of type **AI** which we have considered in Example 5.3.10. The subalgebra in  $\mathcal{P}$ ,

$$\mathcal{A} := \text{span}\{i\sigma_x \otimes \sigma_x, i\sigma_y \otimes \sigma_y, i\sigma_z \otimes \sigma_z\},$$

is a Cartan subalgebra (it is Abelian and it has dimension 3). Therefore every element  $X$  of  $SU(4)$  can be written as

$$X = X_1 \otimes X_2 e^{-i\sigma_x \otimes \sigma_x t_1} e^{-i\sigma_y \otimes \sigma_y t_2} e^{-i\sigma_z \otimes \sigma_z t_3} Y_1 \otimes Y_2, \quad (5.51)$$

for some real parameters  $t_1, t_2, t_3$ , and  $X_1, X_2, Y_1, Y_2 \in SU(2)$ . The decomposition (5.51) displays the two main ingredients in the evolution of the two spins. The factors  $X_1 \otimes X_2$  and  $Y_1 \otimes Y_2$  describe the evolutions of the spins by themselves while the three factors  $e^{-i\sigma_x \otimes \sigma_x t_1}, e^{-i\sigma_y \otimes \sigma_y t_2}, e^{-i\sigma_z \otimes \sigma_z t_3}$  are

---

<sup>7</sup>This can be easily seen by applying the formula for exponentials of Kronecker products (1.22).

responsible for the interaction between the two spins. Every unitary evolution can be decomposed in a way that makes transparent the contributions of the ‘local’ transformation on the single spins and the ‘entangling’ transformation on the pair of two spins. This will be further explored and generalized in [Chapter 8](#).

The decomposition (5.51) can be further refined by recalling that the Pauli matrices are unitarily equivalent to each other. Therefore there exists a matrix  $U_x \in SU(2)$  such that  $U_x \sigma_z U_x^\dagger = \sigma_x$  and a matrix  $U_y \in SU(2)$  such that  $U_y \sigma_z U_y^\dagger = \sigma_y$ . Therefore we have

$$U_x \otimes U_x e^{-i\sigma_z \otimes \sigma_z t_1} U_x^\dagger \otimes U_x^\dagger = e^{-i\sigma_x \otimes \sigma_x t_1}, \quad U_y \otimes U_y e^{-i\sigma_z \otimes \sigma_z t_2} U_y^\dagger \otimes U_y^\dagger = e^{-i\sigma_y \otimes \sigma_y t_2}.$$

Plugging this into (5.51), we obtain a decomposition of  $X$  of the type

$$X = X_1 \otimes X_2 e^{-i\sigma_z \otimes \sigma_z t_1} X_3 \otimes X_4 e^{-i\sigma_z \otimes \sigma_z t_2} X_5 \otimes X_6 e^{-i\sigma_z \otimes \sigma_z t_3} X_7 \otimes X_8, \quad (5.52)$$

with  $X_j \in SU(2)$ ,  $j = 1, \dots, 8$ . Now, assume that a decomposition (5.52) is known for the target state in our control problem. Methods such as the ones described in [subsection 5.3.6](#) can be used for its computation. Then the factors of the type  $X_j \otimes X_{j+1}$ ,  $j = 1, 3, 5, 7$ , can be obtained in (5.50) by very fast, high amplitude controls which would make the effect of the interaction term negligible in the whole evolution. The design of these controls is not specified here but notice that the problem without the interaction term is much easier as the dynamical Lie algebra  $\mathcal{K}$  is only 6-dimensional. In order to obtain factors of the type  $e^{-i\sigma_z \otimes \sigma_z t}$ , we only have to set the controls equal to zero for time  $t$  (notice that  $t_j$ ,  $j = 1, 2, 3$  in (5.52) can always be taken positive as the orbit  $\{e^{-i\sigma_z \otimes \sigma_z t} | t \in \mathbf{R}\}$  is periodic). Alternating fast high amplitude controls with controls identically equal to zero, we drive the state from the identity to the desired target. This control has important properties as for time optimality as we shall see in the next chapter.

### 5.5.2 Control of two coupled spin- $\frac{1}{2}$ particles with Heisenberg interaction

We consider now the model of two interacting spin- $\frac{1}{2}$  particles with Heisenberg interaction (2.67). This model will give us the opportunity to apply much of the machinery introduced in the previous sections, including Levi’s decomposition and Cartan decomposition of  $so(3)$ . In many cases Heisenberg interaction cannot be considered weak compared to the control strength and it has to be constructively used in the control algorithm. The Schrödinger operator equation can be written as (5.49) where the Ising term  $J_{12} \sigma_z \otimes \sigma_z$  is replaced by the isotropic Heisenberg term,  $J_{12}(\sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y + \sigma_z \otimes \sigma_z)$ . We perform a change of coordinates  $X \rightarrow TXT^\dagger$ , where  $T$  is the matrix (5.42). This matrix  $T$  diagonalizes the isotropic Heisenberg term. We re-scale the time by a factor  $J_{12}$  and the controls by a factor  $\gamma_1$  so that only the ratio

$r := \frac{\gamma_2}{\gamma_1}$  appears explicitly in the equation. The Schrödinger operator equation can then be written as

$$\dot{X} = (A + B_x u_x + B_y u_y + B_z u_z) X, \quad (5.53)$$

where

$$A = \text{diag}(3i, -i, -i, -i), \quad (5.54)$$

and assuming for simplicity  $r = 2$ ,

$$B_x = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -3 & 0 \\ 0 & 3 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix},$$

$$B_y = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \\ -1 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \end{pmatrix},$$

$$B_z = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3 \\ 0 & 0 & 3 & 0 \end{pmatrix}.$$

We, first of all, notice that this system is operator controllable, as the dynamical Lie algebra is equal to  $su(4)$ . The general strategy for control is to set alternatively two of the components of the control equal to zero and use the other one to control on a subgroup of the group  $SU(4)$ . In this way, the problem of control is decomposed into subproblems on lower dimensional manifolds. We start by setting  $u_x \equiv u_y \equiv 0$  (the other cases are analogous), i.e., we study the equation

$$\dot{X} = AX + B_z Xu_z. \quad (5.55)$$

The dynamical Lie algebra associated with this equation, which we denote by  $\mathcal{L}_z$ , can be calculated with the algorithm in subsection 3.2.1 of Chapter 3. It is five dimensional and it is spanned by the matrices

$$S_{x,y,z} := \begin{pmatrix} i\sigma_{x,y,z} & 0 \\ 0 & 0 \end{pmatrix}, \quad R_1 := \begin{pmatrix} i\mathbf{1}_{2 \times 2} & 0 \\ 0 & -i\mathbf{1}_{2 \times 2} \end{pmatrix}, \quad R_2 := \begin{pmatrix} 0 & 0 \\ 0 & i\sigma_y \end{pmatrix}. \quad (5.56)$$

From this basis, we can deduce the Levi decomposition of the Lie algebra  $\mathcal{L}_z$ ,

$$\mathcal{L}_z = \mathcal{S} \oplus \mathcal{R}.$$

The semisimple part,  $\mathcal{S}$ , is simple in this case. It is, in fact, isomorphic to  $su(2)$  and it is spanned by  $S_{x,y,z}$ . The solvable part  $\mathcal{R}$  is spanned by  $R_1$  and  $R_2$  and it is Abelian. Moreover we have

$$[\mathcal{S}, \mathcal{R}] = 0. \quad (5.57)$$

The Lie group  $e^{\mathcal{S}}$  consists of products of matrices of the form  $\begin{pmatrix} X & 0 \\ 0 & \mathbf{1}_{2 \times 2} \end{pmatrix}$  with  $X \in SU(2)$ , while the Lie group  $e^{\mathcal{R}}$  consists of product of matrices of the form

$$\begin{pmatrix} e^{it_1} \mathbf{1}_{2 \times 2} & 0 \\ 0 & e^{-it_1} \mathbf{1}_{2 \times 2} \end{pmatrix},$$

and

$$\begin{pmatrix} \mathbf{1}_{2 \times 2} & 0 \\ 0 & Y \end{pmatrix},$$

with

$$Y = \begin{pmatrix} \cos(t_2) & -\sin(t_2) \\ \sin(t_2) & \cos(t_2) \end{pmatrix},$$

for  $t_1$  and  $t_2$  in  $\mathbb{R}$ . We rewrite (5.55) as

$$\dot{X} = (A_S + B_{zS}u_z)X + (A_R + B_{zR}u_z)X, \quad X(0) = \mathbf{1}_{4 \times 4}$$

with  $A = A_S + A_R$  and  $B_z = B_{zS} + B_{zR}$ , with  $\{A_S, B_{zS}\} \in \mathcal{S}$ ,  $\{A_R, B_{zR}\} \in \mathcal{R}$ . Consider now the two equations

$$\dot{U} = (A_S + B_{zS}u_z)U, \quad U(0) = \mathbf{1}_{4 \times 4}, \quad (5.58)$$

$$\dot{V} = (A_R + B_{zR}u_z)V, \quad V(0) = \mathbf{1}_{4 \times 4}.$$

Because of (5.57), for the solution  $X$  of (5.55), we have  $X = UV = VU$ . The problem of control for system (5.58) is essentially a problem for a system on  $SU(2)$  of the type discussed in subsection 5.1.4. The control  $u_z$  can be chosen to drive  $U$  from the identity to any value  $\begin{pmatrix} U_f & 0 \\ 0 & \mathbf{1}_{2 \times 2} \end{pmatrix}$  with  $U_f \in SU(2)$ . The corresponding solution of (5.55) is given by

$$X_f := \begin{pmatrix} e^{it_1} \mathbf{1}_{2 \times 2} & 0 \\ 0 & e^{-it_1} \mathbf{1}_{2 \times 2} \end{pmatrix} \begin{pmatrix} \mathbf{1}_{2 \times 2} & 0 \\ 0 & e^{i\sigma_y t_2} \end{pmatrix} \begin{pmatrix} U_f & 0 \\ 0 & \mathbf{1}_{2 \times 2} \end{pmatrix}, \quad (5.59)$$

for some  $t_1, t_2 \in \mathbb{R}$ .

Now we show how the above considerations on the Lie algebraic structure of the equation (5.55) can be used to achieve control objectives for system (5.53). Assume we want to transfer a (known) state  $\vec{\psi}_0$  to an eigenvector of  $A$ , say  $\vec{\psi}_1 := [e^{i\phi}, 0, 0, 0]^T$ ,  $\phi \in \mathbb{R}$ . It is natural to require that the final state is an eigenvector of  $A$  since we want the state to remain in the desired value after we switch the control to zero. Since we can achieve any arbitrary  $U_f \in SU(2)$  in  $X_f$  in (5.59), we can control at will the first two components of the state vector. We choose  $U_f$  so as to introduce a zero in the position 2. The remaining two factors in (5.59) will introduce a common phase factor in the first two components and will modify the remaining components 3 and 4 without changing their total length. Now a similar analysis with  $u_y \neq 0$  and

$u_x \equiv u_z \equiv 0$  shows that, with  $u_y$ , we can modify the components 1 and 3 at will, while the components 2 and 4 also change but their total length remains constant. Analogously, using  $u_x$ , we can transform at will the components 1 and 4, while the components 2 and 3 evolve. We can alternate controls  $u_x$ ,  $u_y$  and  $u_z$  and, at each step, transfer magnitude from one of the components 2, 3, 4 into the first component. If at each step the component is chosen as the one with maximum magnitude, it can be shown that the state will converge to the desired state  $\vec{\psi}_1 := [e^{i\phi}, 0, 0, 0]^T$ .

The problem for the control of the evolution operator is more complicated and more general. We give here some of the ideas and refer to [6] for a complete treatment. By solving a two level problem on  $SU(2)$ , with  $u_z$  possibly different from zero and  $u_x \equiv u_y \equiv 0$ , we can control to a value  $X_f$  in (5.59) for an arbitrary  $U_f \in SU(2)$ . By choosing final time and control appropriately, it

is possible to make the extra factor  $\begin{pmatrix} e^{it_1} \mathbf{1}_{2 \times 2} & 0 \\ 0 & e^{-it_1} \mathbf{1}_{2 \times 2} \end{pmatrix} \begin{pmatrix} \mathbf{1}_{2 \times 2} & 0 \\ 0 & e^{i\sigma_y t_2} \end{pmatrix}$  in (5.59) equal to the identity, so that we obtain matrices which are equal to the identity except for the elements at the intersection of the first and second rows and columns which are occupied by an element of  $SU(2)$ . In the same way, using  $u_y$  and  $u_x$  we can obtain matrices with rows and columns 1 and 3 and 1 and 4, respectively, equal to arbitrary matrices in  $SU(2)$ . Now, we use a decomposition into planar rotations (5.22) for the target operator  $X_f \in SU(4)$ . This decomposes the problem into subproblems for a diagonal matrix  $D \in SU(4)$  and matrices of the type  $U_{1,2}$ ,  $U_{1,3}$ ,  $U_{2,3}$ ,  $U_{1,4}$ ,  $U_{2,4}$ ,  $U_{3,4}$ . Matrices of the type  $U_{1,2}$ ,  $U_{1,3}$  and  $U_{1,4}$  can be obtained as above, while matrices of the form  $U_{2,3}$  can be obtained as  $U_{1,2} U_{1,3} U_{1,2}^\dagger$ , for appropriate elements  $U_{1,2}$  and  $U_{1,3}$ . Analogously, one can obtain matrices of the form  $U_{2,4}$  and  $U_{3,4}$ . Diagonal matrices can be obtained as exponentials of  $A$  in (5.54) and exponentials of matrices that are obtained by permuting the diagonal elements of  $A$  using similarity transformations of the type  $U_{1,2}$ ,  $U_{1,3}$  and  $U_{1,4}$ .

## 5.6 Notes and References

Techniques of Lie group decompositions for the control of quantum systems have been used in many papers. In particular, for the two level problem more sophisticated results exist which incorporate in the design arbitrary bounds on the control [169] and/or minimize number of required switches [60]. The fact that Cartan decomposition could be used for the control of two interacting spins was recognized by several authors [55], [119]. In particular the paper [119] presents a geometric treatment which allows one to conclude that the controls based on Cartan's decompositions are the ones that infimize the time to reach a particular target. This will be further discussed in the next chapter,

which is devoted to the optimal control of quantum systems. Applications of Cartan decomposition to the three-spin case and general  $n$ -spin case are presented in [120], [121], while the paper [170] contains a treatment of the two-spin problem with Ising interaction which avoids very large fast control pulses and, in fact, allows for arbitrarily bounded controls. The problem with Heisenberg interaction of subsection 5.5.2 was considered in [211] and [6]. For further applications and examples of Lie group decompositions applied to quantum control we refer to [58], [188], [189] and references therein. Our presentation of the Cartan decompositions of  $su(n)$  mainly follows [61], in particular for the relation between decompositions and quantum symmetries.

---

## 5.7 Exercises

**Exercise 5.1** Consider the problem of driving the state  $\psi_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  to the state  $\psi_f = \begin{pmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}$ , for the system

$$\dot{X} = AX + BXu,$$

with  $A = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$  and  $B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ . Assume there is no a priori bound on the control. Use the Euler decomposition technique of subsection 5.1.4 to find a control law for this problem.

**Exercise 5.2** Find a decomposition into polar rotations for the special unitary matrix

$$X_f = \begin{pmatrix} \frac{i}{\sqrt{2}} & -\frac{i}{\sqrt{2}} & 0 \\ \frac{\sqrt{3}i}{2\sqrt{2}} & \frac{\sqrt{3}i}{2\sqrt{2}} & \frac{1}{2} \\ \frac{-i}{2\sqrt{2}} & \frac{-i}{2\sqrt{2}} & -\frac{\sqrt{3}}{2} \end{pmatrix}.$$

**Exercise 5.3** Let  $\mathcal{K}$  a subalgebra of  $\mathcal{L} \subseteq su(n)$ . Recall the inner product on  $\mathcal{L}$  is the one inherited from  $su(n)$  namely  $\langle A, B \rangle = \text{Tr}(AB^\dagger)$ . Prove that

$$[\mathcal{K}, \mathcal{K}^\perp] \subseteq \mathcal{K}^\perp.$$

Compare with (5.25).

**Exercise 5.4** Show that if  $\mathcal{A}$  is a maximal Abelian subalgebra in  $\mathcal{P}$ , for a given Cartan decomposition of  $\mathcal{L}$ , then it is also a maximal Abelian subalgebra of  $\mathcal{L}$ .

**Exercise 5.5** Find the factors in the **AI**, **AII** and **AIII**, with  $p = 2$ ,  $q = 2$ , Cartan decomposition of the unitary matrix  $T$  in (5.42).

**Exercise 5.6** Consider the Lie algebra spanned by the matrices in (5.56) and deduce the Levi decomposition using the methods described in [Appendix C](#), i.e., first find the solvable radical and then the Cartan subalgebra.

**Exercise 5.7** Check that the vector space  $iu(n)$  with the product  $A \diamond B := \{A, B\} := AB + BA$  has the structure of a nonassociative algebra. This means that it is a vector space with a multiplication  $A \diamond B$  which is nonassociative, i.e., there are elements  $A, B, C$  such that

$$A \diamond (B \diamond C) \neq (A \diamond B) \diamond C,$$

and that the multiplication  $A \diamond B$  satisfies the *Jordan identity*

$$(A \diamond A) \diamond (A \diamond B) = A \diamond ((A \diamond A) \diamond B).$$

**Exercise 5.8** Prove that there exists a one to one correspondence between Cartan involutions and Cartan decompositions as stated at the end of [subsection 5.3.1](#).

# Chapter 6

---

## *Optimal Control of Quantum Systems*

Optimal control theory provides a powerful set of tools and concepts that can be applied to quantum control systems. In general terms, optimal control theory (see, e.g., [81], [132], [135], [161]) is concerned with the control of a system

$$\dot{x} = f(x, u), \quad (6.1)$$

and the simultaneous minimization of a functional of the state  $x$  and the control  $u$ . Problems of control of quantum systems can naturally be formulated in this setting.

Section 6.1 presents the general formulation of optimal control problems and section 6.2 gives the basic necessary conditions of optimality, the *Pontryagin maximum principle* (PMP). An outline of the derivation of the PMP in the case of a fixed final time and free final state is presented in Appendix F. The derivation in this case is particularly simple and instructive. Section 6.3 presents an example of application of optimal control techniques to a quantum control problem that can be solved analytically. In section 6.4, we turn our attention to the optimal control problem of minimizing the *time* for a prescribed state transfer. This problem is very much motivated for quantum systems. In fact, one of the major problems in this context is that quantum systems are very sensitive to the presence of the environment which often destroys the main features of quantum dynamics. This is called *de-coherence*. One method to prevent de-coherence is to obtain the desired state transfer in the least possible time so that the interaction with the environment becomes negligible. The treatment of the time optimal control problem uses different tools depending on whether or not there is a prescribed bound on the control magnitude. For unbounded controls, the time optimal control does not exist. However an infimum can be obtained in several cases (see subsection 6.4.2) using a geometric argument based on the theory of Riemannian symmetric spaces, some of which was discussed in the previous chapter.

Very rarely, and typically only for low dimensional cases, or under special symmetries, the solution of an optimal control problem can be obtained explicitly in the form of a given function of time. Much more often, in particular for higher dimensional control problems in molecular dynamics, numerical and

iterative algorithms are used in order to find the optimal control. We shall discuss some of these methods in [section 6.5](#).

---

## 6.1 Formulation of the Optimal Control Problem

A general optimal control problem can be formulated as follows:

### Optimal Control Problem:

Given a set  $\mathcal{X}$  of (state) functions  $x : \mathbf{R} \rightarrow \mathbf{R}^n$ , and a set  $\mathcal{U}$  of (control) functions  $u : \mathbf{R} \rightarrow \mathbf{R}^m$ , find the functions  $x \in \mathcal{X}$  and  $u \in \mathcal{U}$  which minimize a cost functional  $J : \mathcal{X} \times \mathcal{U} \rightarrow \mathbf{R}$  and satisfy the *dynamical constraint* (6.1).

Virtually every control problem can be formulated as a special case of the above general optimal control problem. Here is a simple example.

**Example 6.1.1** Assume that we want to find a piecewise continuous control  $u$  to drive the state  $x$  of (6.1) from a given value  $x_0$  to a desired state  $x_f$  in time  $T$ . Then  $\mathcal{U}$  is the set of piecewise continuous functions on  $[0, T]$ ,  $\mathcal{X}$  can be taken as the set of continuous functions on  $[0, T]$  satisfying  $x(0) = x_0$ , and the cost functional we seek to minimize is

$$J(u) := \|x(T) - x_f\|^2, \quad (6.2)$$

i.e., a measure of the distance of the final state from the desired one. We can also incorporate in the cost a term which takes into consideration the energy used during the given interval of time. For example, a cost of the form

$$J(u) = \lambda \|x(T) - x_f\|^2 + \int_0^T \|u(t)\|^2 dt, \quad (6.3)$$

with  $\lambda > 0$ , will incorporate a penalty on the final state and an energy like term for the control. We can increase or decrease  $\lambda$  according to whether or not we are willing to use more energy to obtain a better final state or not.

Problems of optimal control are rarely expressed in the general form above discussed. They are expressed in one of the three equivalent forms described in the following subsection.

### 6.1.1 Optimal control problems of Mayer, Lagrange and Bolza

In the following, it is assumed that the initial value for the state  $x$  is given. Therefore, the choice of the control  $u$  and the requirement that (6.1) is satisfied determine  $x$  uniquely.

Given the system

$$\dot{x} = f(x, u), \quad x \in \mathbf{R}^n, u \in \mathbf{R}^m, \quad (6.4)$$

the **problem of Mayer** is to determine a control function  $u$ , in an appropriate set of functions, to minimize the cost functional in the form

$$J(u) := \phi(x(T), T), \quad (6.5)$$

where  $\phi$  is a (smooth) function :  $\mathbf{R}^n \times \mathbf{R} \rightarrow \mathbf{R}$ . Examples of Mayer problems are *minimum time problems* where the problem is to drive the state to a given value in minimum time. In this case the set of admissible controls consists of the ones steering to the desired target state and the cost has the form (6.5), with  $J = T$ . Another example is a problem with the cost of the form of  $J(u)$  in (6.2). Mayer problems arise when there is a particular emphasis on the final state and/or time.

In the **problem of Lagrange**, the cost functional takes the form

$$J(u) := \int_0^T L(x(t), u(t), t) dt,$$

where  $L$  is a (smooth) function :  $\mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R} \rightarrow \mathbf{R}$ . A problem of Lagrange describes a situation where the cost accumulates with time. This is the case, for example, when one wants to minimize the energy used during the control action and/or the average distance of the trajectory from a given point.

A **problem of Bolza** is a combination of problems of Mayer and Lagrange as the cost takes the form

$$J(u) := \phi(x(T), T) + \int_0^T L(x(t), u(t), t) dt, \quad (6.6)$$

with  $\phi$  and  $L$  (smooth) functions :  $\mathbf{R}^n \times \mathbf{R} \rightarrow \mathbf{R}$ ,  $\mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R} \rightarrow \mathbf{R}$ , respectively. Bolza problems arise when there is a cumulative cost which increases during the control action but special emphasis is placed on the situation at the final time. An example is the cost of the form (6.3) in Example 6.1.1.

Mayer, Lagrange and Bolza problems are all equivalent in that each of them can be converted to any other one. It is obvious that Lagrange and Mayer problems are special cases of Bolza problems. A Bolza problem for the system (6.4) with cost (6.6) can be transformed into a Mayer problem by introducing an extra component for the state vector  $y$  which satisfies the equation

$$\dot{y} = L(x, u, t), \quad y(0) = 0.$$

Using this extra variable the cost takes the Mayer form

$$J(u) = \phi(x(T), T) + y(T).$$

A Mayer problem with cost (6.5) can be converted into a Lagrange problem by rewriting the cost as

$$\begin{aligned} J(u) &= \phi(x(T), T) = \phi(x(0), 0) + \int_0^T \frac{d}{dt} \phi(x(t), t) dt \\ &= \phi(x(0), 0) + \int_0^T \left( \phi_x(x(t), t) f(x(t), u(t)) + \frac{\partial}{\partial t} \phi(x(t), t) \right) dt. \end{aligned}$$

Since  $x(0)$  is fixed, the problem is to minimize the cost

$$\bar{J}(u) = \int_0^T L(x(t), u(t), t) dt,$$

where  $L(x(t), u(t), t)$  is given by

$$L(x(t), u(t), t) = \phi_x(x(t), t) f(x(t), u(t)) + \frac{\partial}{\partial t} \phi(x(t), t),$$

which is a problem of Lagrange.

### 6.1.2 Optimal control problems for quantum systems

For quantum control systems, the state ( $x$ ) may be the density matrix, the pure state vector, or one may consider the dynamics of the evolution operator. We shall consider here the optimal control problem for the pure state in detail. Similar considerations can be made in the other cases.

The dynamics (6.1) is described by the (controlled) Schrödinger equation

$$\frac{d}{dt} \vec{\psi} = -iH(u)\vec{\psi}, \quad (6.7)$$

and a general cost of the Bolza type can be written as

$$J(u) = \phi(\vec{\psi}(T), T) + \int_0^T L(\vec{\psi}(t), u(t), t) dt. \quad (6.8)$$

In this case, both the state  $\vec{\psi}$  and the matrix  $-iH(u)$  are complex quantities. To express the problem in terms of real quantities only, we can write

$$\vec{\psi} = \vec{\psi}_R + i\vec{\psi}_I$$

and

$$-iH(u) = R(u) + iI(u),$$

with  $\vec{\psi}_R$  and  $\vec{\psi}_I$  real vectors of dimension  $n$  and  $R(u)$  and  $I(u)$  real  $n \times n$  matrix functions of  $u$ ,  $R(u)$  skew-symmetric and  $I(u)$  symmetric for every value of  $u$ . Placing this into (6.7) and separating the imaginary and real part, we obtain the two real differential equations

$$\frac{d}{dt} \vec{\psi}_R = R(u)\vec{\psi}_R - I(u)\vec{\psi}_I,$$

$$\frac{d}{dt}\vec{\psi}_I = I(u)\vec{\psi}_R + R(u)\vec{\psi}_I.$$

By defining

$$x := [\vec{\psi}_R^T, \vec{\psi}_I^T]^T, \quad (6.9)$$

and

$$\tilde{H}(u) := \begin{pmatrix} R(u) & -I(u) \\ I(u) & R(u) \end{pmatrix}, \quad (6.10)$$

we can write the differential equation describing the dynamics involving only real quantities as

$$\dot{x} = \tilde{H}(u)x. \quad (6.11)$$

$\tilde{H}(u)$  is skew-symmetric and symplectic for every  $u$ , i.e., it belongs to  $so(2n) \cap sp(n)$  (cf. Exercise 6.4). Also, the cost (6.8) can be rewritten as

$$J(u) = \tilde{\phi}(x(T), T) + \int_0^T \tilde{L}(x(t), u(t), t) dt, \quad (6.12)$$

for appropriate functions  $\tilde{\phi}$  and  $\tilde{L}$ .

A common choice for the cost functional (6.8) in molecular control is the *laser electric field fluence*

$$J(u) = k \int_0^T u^2(t) dt, \quad k > 0,$$

which measures the energy of the electric field in the interval  $[0, T]$  (cf. (2.3), (2.27)). Another possibility is to choose a cost of the type<sup>1</sup>

$$J(u) := k \int_0^T \left( \frac{du}{dt} \right)^2 dt, \quad k > 0,$$

which filters the high frequency components of the control field. When emphasis is placed on the final state also, one can minimize a cost functional in the form (6.3), or, in the form,

$$\frac{1}{2}\langle O \rangle_\psi + \frac{k}{2} \int_0^T u^2(t) dt := \frac{1}{2} \vec{\psi}^\dagger O \vec{\psi} + \frac{k}{2} \int_0^T u^2(t) dt, \quad k > 0, \quad (6.13)$$

where  $O$  is a Hermitian, negative definite, matrix. For example, we may choose  $O = -\vec{\psi}_f \vec{\psi}_f^\dagger$  if  $\vec{\psi}_f$  is the desired state.

---

<sup>1</sup>In this case one either allows the cost to be a functional of the derivative of the control or may include an extra state variable in the equations of the dynamics.

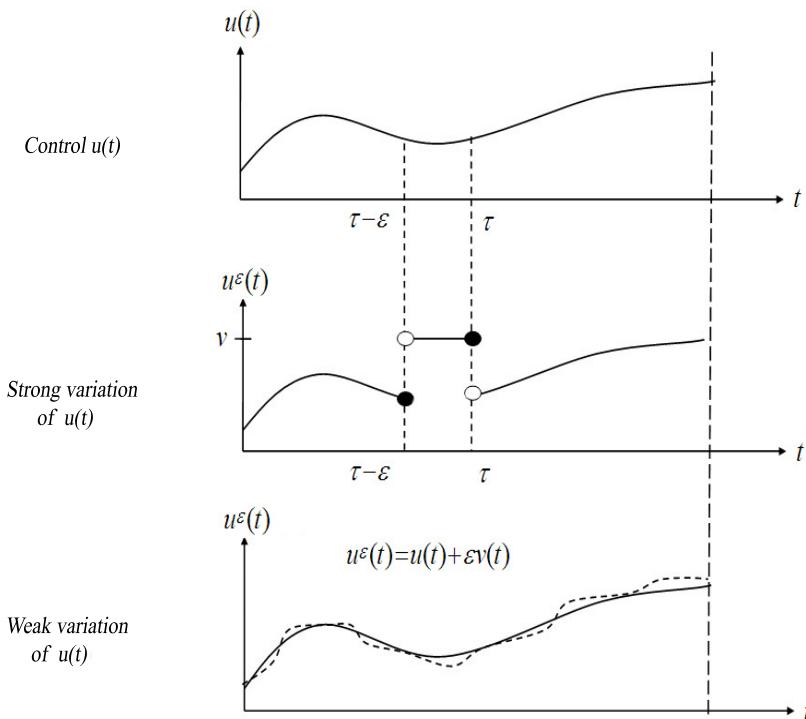
## 6.2 The Necessary Conditions of Optimality

### 6.2.1 General necessary conditions of optimality

In every derivation of **necessary conditions** of optimality, one assumes that the control  $u$  is optimal and replaces  $u$  with a slightly different control  $u^\epsilon$ , where  $\epsilon$  is a small parameter. The control  $u^\epsilon$  is still admissible, and it is called a *variation* of  $u$ . Then one imposes that, for the cost  $J = J(u)$ ,

$$J(u^\epsilon) - J(u) \geq 0, \quad (6.14)$$

which gives the desired necessary conditions on the control  $u$ . According to what type of variation one considers, different optimality conditions are obtained.



**FIGURE 6.1:** Strong and Weak Variation of a Control Function  $u = u(t)$

The basic necessary conditions of optimality in optimal control, known as the Pontryagin maximum principle (PMP), are obtained using a so-called

*strong variation.* A *strong variation*  $u^\epsilon$  of  $u$  is defined as follows: Fix  $\tau \in (0, T]$ , then  $u^\epsilon$  is a function equal to  $u$  for  $t \in [0, \tau - \epsilon]$  and for  $t \in (\tau, T]$ , and constant and equal to  $v$  for  $t \subseteq (\tau - \epsilon, \tau]$ . The value  $v$  is any value in the set of admissible values for the control,  $\mathcal{U} \subseteq \mathbb{R}^m$ . An outline of the derivation in the case where the final time  $T$  (other than initial time and state) is fixed and the final state is free is presented in [Appendix F](#). The final result is stated as follows.

**Theorem 6.2.1** (*Pontryagin maximum principle for Mayer problems. Fixed final time and free final state.*) *Assume  $u$  is the optimal control and  $x$  the corresponding trajectory solution of (6.1). Then, there exists a nonzero vector  $\lambda$  solution of the adjoint equations*

$$\dot{\lambda}^T = -\lambda^T f_x(x(t), u(t)) \quad (6.15)$$

*with terminal condition*<sup>2</sup>

$$\lambda^T(T) = -\phi_x(x(T)), \quad (6.16)$$

*such that, for every  $\tau \in (0, T]$ , we have*

$$\lambda^T(\tau) f(x(\tau), u(\tau)) \geq \lambda^T(\tau) f(x(\tau), v), \quad (6.17)$$

*for every  $v$  in the set of the admissible values for the control  $\mathcal{U}$ .*

The vector state  $\lambda$  is called the **costate**.

**Remark 6.2.2** If there is no bound on the possible value of the control  $u$ , equation (6.17) can be replaced by

$$\frac{\partial(\lambda^T f(x, u))}{\partial u} = \lambda^T f_u(x, u) = 0. \quad (6.18)$$

In this case, the necessary condition of optimality can be obtained by using a *weak variation* of the control. A weak variation is defined as

$$u(t) \rightarrow u^\epsilon := u(t) + \epsilon v(t),$$

where  $v(t)$  is a given function. If we denote by  $x^\epsilon$  the trajectory corresponding to  $u^\epsilon$  for  $\epsilon$  small, we have (for a Mayer problem)

$$\Delta J := J(x^\epsilon(T)) - J(x(T)) \approx \phi_x(x(T))(x^\epsilon(T) - x(T)). \quad (6.19)$$

Defining  $z(t) := \frac{d}{d\epsilon} x^\epsilon(t)|_{\epsilon=0}$ , we have

$$x^\epsilon(T) - x(T) \approx z(T)\epsilon.$$

---

<sup>2</sup>We drop the argument  $T$  in the function  $\phi$ , since the final time  $T$  is fixed.

Plugging this into (6.19), we obtain

$$\Delta J \approx \phi_x(x(T))z(T)\epsilon. \quad (6.20)$$

Consider the differential equation for  $x^\epsilon$ ,

$$\dot{x}^\epsilon = f(x^\epsilon, u + \epsilon v).$$

A differential equation for  $z(t)$  is obtained by differentiating this equation with respect to  $\epsilon$ , at  $\epsilon = 0$  and switching the order of differentiation between time  $t$  and  $\epsilon$ . We obtain

$$\dot{z} = f_x(x, u)z + f_u(x, u)v, \quad (6.21)$$

with initial condition  $z(0) = 0$ . Now consider the adjoint equations (same as (6.15))

$$\dot{\lambda}^T = -\lambda^T f_x(x(t), u(t)), \quad (6.22)$$

with final condition (same as (6.16))  $\lambda^T(T) = -\phi_x(x(T))$ . We can write formula (6.20) as

$$\Delta J \approx -\lambda^T(T)z(T)\epsilon. \quad (6.23)$$

Since  $z(0) = 0$ , we can write  $\lambda(T)z(T)$  as  $\lambda^T(T)z(T) = \int_0^T \frac{d}{dt} \lambda^T(t)z(t)dt$  which, using (6.21) and (6.22), gives

$$\lambda^T(T)z(T) = \int_0^T \lambda^T(t)f_u(x(t), u(t))v(t)dt.$$

This, placed in (6.23), gives

$$\Delta J \approx -\epsilon \int_0^T \lambda^T(t)f_u(x(t), u(t))v(t)dt.$$

If  $u$  is a local minimum for  $J$ , then it should not be possible by a choice of the perturbation  $v$  to obtain a smaller value. Since there is no bound on  $v$ , this requires  $\lambda^T(t)f_u(x(t), u(t)) \equiv 0$ , i.e., (6.18).

The Pontryagin maximum principle of Theorem 6.2.1 is most often used in the following way. One defines a function

$$h = h(\lambda, x, u) := \lambda^T f(x, u),$$

which is called the *optimal control Hamiltonian*.<sup>3</sup> For every  $\lambda$  and  $x$ , one maximizes this function over  $u \in \mathcal{U}$ , i.e.,  $u$  is chosen so that (cf. (6.17)), for every  $\lambda$  and  $x$ ,

$$h(\lambda, x, u) \geq h(\lambda, x, v), \quad (6.24)$$

---

<sup>3</sup>We use the notation  $h$  to avoid confusion with the quantum mechanical Hamiltonian  $H$ . The two Hamiltonians are however related as both the optimal control problem and Lagrangian mechanics are instances of variational problems (cf. Remark B.1.5).

for every  $v \in \mathcal{U}$ . This is a static optimization problem for all the values of the parameters  $\lambda$  and  $x$ . The solution will depend on these parameters and we can write it as  $u := u(x, \lambda)$ . Now, by placing this form of  $u$  into equations (6.1) and (6.15) we obtain the system of equations

$$\dot{x} = f(x, u(x, \lambda)), \quad (6.25)$$

$$\dot{\lambda}^T = -\lambda^T f_x(x, u(x, \lambda)), \quad (6.26)$$

which has to be solved with the boundary conditions  $x(0) = x_0$  and  $\lambda^T(T) = -\phi_x(x(T))$ . This is a *two points boundary value problem* as the boundary conditions are at time 0 and  $T$ . Typically, one leaves the initial condition for  $\lambda$ ,  $\lambda(0)$ , as a free parameter, solves the equations (6.25)-(6.26) and then tries to adjust the value of the parameter  $\lambda(0)$  so as to meet condition (6.16). Every control which is obtained with this procedure satisfies the necessary conditions of optimality and it is a candidate to be the optimal control. Every control which satisfies the necessary optimality conditions is called an *extremal*. By comparing the costs obtained with the various extremals one finds the minimum cost and the optimal control. This procedure can typically be carried out analytically only for low dimensional problems or problems which have special symmetries.

The above procedure assumes that an optimal control exists. The problem of existence of optimal controls is treated, for example, in ([81] Chapter III). Theorems of existence of optimal controls can be given under conditions on the equation in (6.1) and on the set of admissible control functions, which we have assumed here to include at least the set of piecewise continuous functions.<sup>4</sup>

If one has a Bolza problem with free endpoint instead of a Mayer problem, the conditions of the maximum principle can be easily adapted.

Recall from [subsection 6.1.1](#), that, by defining the extra variable  $y(t) := \int_0^t L(x(s), u(s), s) ds$ , the cost  $J(u) := \phi(x(T)) + \int_0^T L(x(s), u(s), s) ds$  can be rewritten in the Mayer form

$$J(u) = \bar{\phi}(x(T), y(T)) := \phi(x(T)) + y(T).$$

The costate now has dimension  $n + 1$ . If we call  $\lambda$  the first  $n$  components of the costate and  $\mu$  the remaining component, the adjoint equations take the form

$$\dot{\lambda}^T = -\lambda^T f_x - \mu L_x, \quad (6.27)$$

$$\dot{\mu} = 0.$$

So  $\mu$  is a constant, and since  $\mu(T) = -\bar{\phi}_y = -1$ , we have that condition (6.24) can be rewritten with the Hamiltonian

$$h(\lambda, x, u) := \lambda^T f(x, u) - L(x, u). \quad (6.28)$$

---

<sup>4</sup>This allowed us to use a strong variation as an admissible control.

Summarizing, for a Bolza problem (6.6) with fixed final time and free final state, if  $u$  is optimal, then it satisfies (6.24) with  $h$  given in (6.28).  $\lambda$  satisfies the first of (6.27) with terminal boundary condition (6.16), and  $x$  satisfies (6.1).

### 6.2.2 The necessary optimality conditions for quantum control problems

For the quantum optimal control problem (6.11), (6.12), the optimal control Hamiltonian  $h(\lambda, x, u)$  takes the form (cf. (6.28))

$$h(\lambda, x, u) = \lambda^T \tilde{H}(u)x - \tilde{L}(x, u).$$

The optimal control  $u$  has to satisfy (6.24) with this Hamiltonian. The adjoint equations are

$$\dot{\lambda}^T = -\lambda^T \tilde{H}(u) + \tilde{L}_x,$$

which, since  $\tilde{H}(u)$  is skew-symmetric, can be written as

$$\dot{\lambda} = \tilde{H}(u)\lambda + \tilde{L}_x^T, \quad (6.29)$$

and the terminal condition is given by  $\lambda(T) = -\tilde{\phi}_x^T(x(T))$  (cf. (6.16)). Notice that, if  $\tilde{L}_x \equiv 0$ , i.e., the integrand in the cost (6.12) does not depend on the state  $x$ ,  $x$  and  $\lambda$  satisfy the same differential equation and the equations for  $\lambda$  and  $x$  are coupled only through the controls.

### 6.3 Example: Optimal Control of a Two Level Quantum System

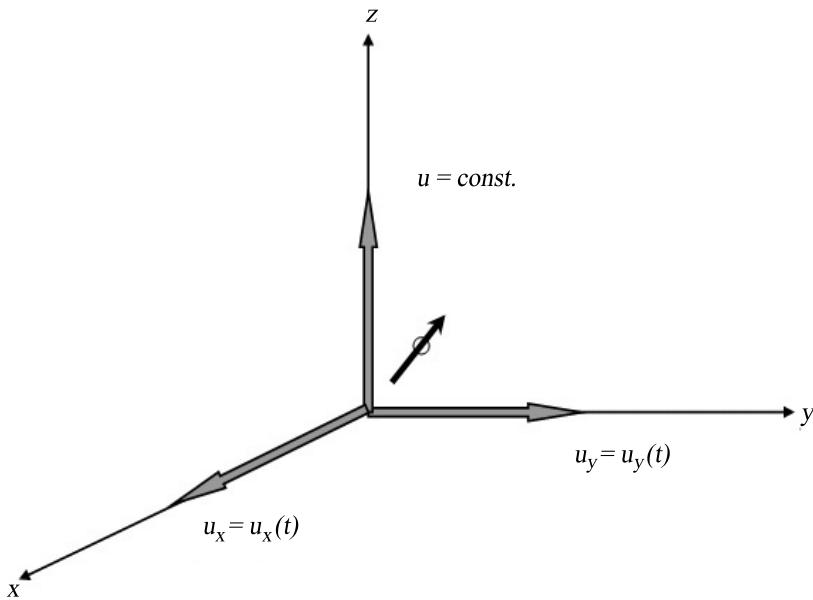
The differential equation of a spin- $\frac{1}{2}$  particle in an electromagnetic field was discussed in [subsection 2.3.2](#). It can be written, after appropriate scaling of time and magnetic field in (2.63), as

$$\dot{\psi} = (\bar{\sigma}_z u_z + \bar{\sigma}_x u_x(t) + \bar{\sigma}_y u_y(t))\psi. \quad (6.30)$$

$\bar{\sigma}_{x,y,z}$  are the (multiples of the Pauli) matrices in (5.5) satisfying the commutation relations (5.6).  $u_{x,y,z}$  represent the  $x, y, z$  components of the magnetic field which is used for control. We assume that the  $z$  component is held constant. A scheme is depicted in [Figure 6.2](#).

A possible cost to be minimized is of the form

$$J(u_x, u_y) = -Re(\vec{\psi}^\dagger(T)\vec{\psi}_f) + \eta \int_0^T u_x^2(t) + u_y^2(t) dt, \quad (6.31)$$



**FIGURE 6.2:** Scheme of a Spin  $\frac{1}{2}$  Particle in a Magnetic Field: The field is varying in the  $x$  and  $y$  direction and is constant in the  $z$  direction.

with the  $\eta > 0$  constant and  $\vec{\psi}_f$  equal to the desired final state. The cost (6.31) expresses a compromise between the goal of transferring the state to  $\vec{\psi}_f$  and the goal of keeping the energy of the field, i.e., the integral in (6.31), small. More importance is given to the second goal if the parameter  $\eta$  is large. Transforming the problem into a real one according to the procedure described in subsection 6.1.2 gives the equation corresponding to (6.30),

$$\dot{x} = (T_z u_z + T_y u_y(t) + T_x u_x(t))x, \quad (6.32)$$

where the matrices  $T_{x,y,z}$  are given by

$$T_x := \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad T_y = \frac{1}{2} \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

$$T_z = \frac{1}{2} \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}.$$

They satisfy the commutation relations

$$[T_x, T_y] = T_z, \quad [T_y, T_z] = T_x, \quad [T_z, T_x] = T_y, \quad (6.33)$$

so that the map

$$\bar{\sigma}_x \leftrightarrow T_x, \quad \bar{\sigma}_y \leftrightarrow T_y, \quad \bar{\sigma}_z \leftrightarrow T_z$$

is a Lie algebra isomorphism. The cost  $J(u_x, u_y)$  in (6.31) can be written, denoting by  $x_f$  the  $2n$  dimensional real vector corresponding to  $\vec{\psi}_f$ ,

$$J(u_x, u_y) = -x^T(T)x_f + \eta \int_0^T u_x^2(t) + u_y^2(t) dt. \quad (6.34)$$

We now apply the maximum principle to determine the form of the optimal control in this case. Notice that the integrand in (6.34) does not depend on  $x$  and therefore the adjoint equations for the costate  $\lambda$  are the same as the equations for  $x$ , i.e.,

$$\dot{\lambda} = (T_z u_z + T_y u_y(t) + T_x u_x(t))\lambda. \quad (6.35)$$

The optimal control Hamiltonian  $h(\lambda, x, u_x, u_y)$  is given by

$$h(\lambda, x, u_x, u_y) = \lambda^T(T_z u_z + T_y u_y + T_x u_x)x - \eta(u_x^2 + u_y^2).$$

Maximization with respect to  $u_x$  and  $u_y$  gives

$$u_x = \frac{1}{2\eta}\lambda^T T_x x, \quad (6.36)$$

$$u_y = \frac{1}{2\eta}\lambda^T T_y x. \quad (6.37)$$

At this point, we can place (6.36) and (6.37) into (6.32) and (6.35) and try to solve the associated two point boundary value problem. A better alternative, in this case, is to differentiate (6.36) and (6.37). Using (6.32) and (6.35) and the commutation relations (6.33), we obtain

$$\dot{u}_x = \frac{1}{2\eta}(\lambda^T T_z x u_y - u_z u_y), \quad (6.38)$$

$$\dot{u}_y = \frac{1}{2\eta}(-\lambda^T T_z x u_x + u_z u_x). \quad (6.39)$$

A similar calculation gives

$$\frac{d}{dt}\lambda^T T_z x \equiv 0,$$

so that (6.38) and (6.39) can be rewritten as

$$\dot{u}_x = k u_y,$$

$$\dot{u}_y = -ku_x,$$

where  $k$  is the constant  $k := \frac{1}{2\eta}(\lambda^T T_z x - u_z)$ . This shows that the optimal controls are trigonometric functions of the form

$$u_x(t) = M \cos(\omega t + \gamma), \quad (6.40)$$

$$u_y(t) = M \sin(\omega t + \gamma). \quad (6.41)$$

By replacing these in the equations (6.32) or (6.30) we can solve the corresponding differential equation with  $\omega$ ,  $M$  and  $\gamma$  as free parameters and then determine the parameters  $\omega$ ,  $\gamma$  and  $M$  which minimize the cost.

In conclusion, the Pontryagin maximum principle allows to transform an infinite dimensional optimization problem (the search over a set of functions) into a finite dimensional optimization problem (the search over a set of parameters) (cf. Exercise 6.2).

---

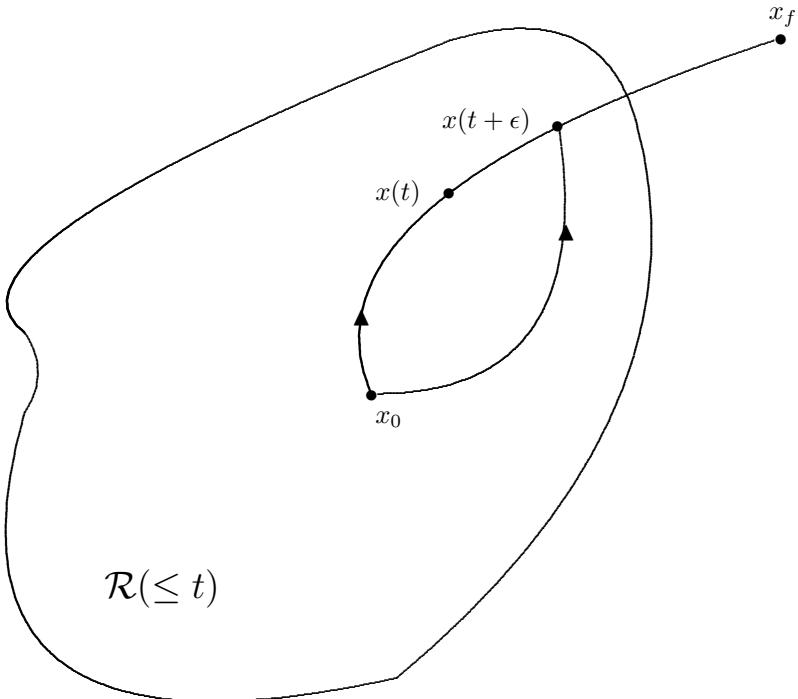
## 6.4 Time Optimal Control of Quantum Systems

Consider the problem of steering the state between two values in minimum time. The standard Filippov existence theorems (see, e.g., [81]) guarantee that the time optimal control exists, under the assumptions of controllability, i.e., existence of a control performing the desired task, Lebesgue measurable control functions with values in a compact set, and smoothness and growth conditions on the functions describing the dynamics. These smoothness and growth conditions are satisfied for bilinear quantum control models described in [Chapter 2](#).

The problem of finding the time optimal control is related to the problem of describing the reachable sets  $\mathcal{R}(t)$ ,  $\mathcal{R}(\leq t)$ ,  $\mathcal{R}$  starting from a given state, as defined in [Chapter 3](#) (see [discussion](#) before Formula (3.2)). In fact we have the following fact

**Proposition 6.4.1** Consider system (6.1) with an initial condition  $x_0$ . If  $u$  is the minimum time control to reach the target  $x_f$ ,  $T$  the minimum time and  $x = x(t)$  the corresponding trajectory, then for every  $t \in (0, T)$ ,  $x(t)$  belongs to the boundary of the reachable set  $\mathcal{R}(\leq t)$ , starting from  $x_0$ .

*Proof.* The idea of the proof is illustrated in [Figure 6.3](#). Fix  $t \in (0, T)$ . If  $x(t)$  belongs to the interior of  $\mathcal{R}(\leq t)$  then there exists an  $\epsilon > 0$  such that  $x(t + \epsilon)$  also belongs to  $\mathcal{R}(\leq t)$ . Therefore using another control and an alternative trajectory  $x_1$ , it is possible to reach  $x(t + \epsilon)$  in time  $\leq t$ . From  $x(t + \epsilon)$  we can then follow the trajectory  $x$ , for time  $T - t - \epsilon$  and reach  $x_f$  in time  $\leq T - \epsilon$  which contradicts the minimality of  $T$ .  $\square$



**FIGURE 6.3:** Illustration of the proof of Proposition 6.4.1.

For time optimal control problems, there exists a different version of the PMP (cf. Theorem 6.2.1) which singles out a set of candidate controls, the *extremals*. Nevertheless, the problem is typically very complicated because of the need to integrate a set of differential equations and then to choose the control which gives the minimum time. For this reason, the explicit expression of the time optimal controls can be obtained only in special cases and typically for low dimensional problems [29]. For some problems, there are relevant properties that simplify the search for the time optimal control (see, e.g., [28] and the references therein). An example of a class of systems will be described in the next subsection.

In the case where the controls are not required to be bounded in magnitude, the time optimal control does not exist in general for quantum control systems, and one has to look for a control which gives the infimum time. In some situations [119], such controls can be found using the Cartan decompositions of Lie groups and the Cartan classification of symmetric spaces described in the previous chapter for the case of  $SU(n)$ .

The following subsection deals with the case of bounded control by describing the necessary conditions of optimality and some examples of problems that can be solved explicitly. Subsection 6.4.2 deals with the case of unbounded control and the search for a control which makes the time of transfer arbitrarily close to an infimum.

### 6.4.1 The time optimal control problem; Bounded control

For time optimal control problems, we have the following version of the PMP which gives conditions analogous to the ones of Theorem 6.2.1, with the only difference that the condition (6.16) on the final value of the costate is now replaced by  $x(T) = x_f$ .<sup>5</sup> Recall that, for a time optimal control problem, we have fixed initial and final conditions for the state  $x$  and let the final time free and to be minimized.

**Theorem 6.4.2** *Consider the problem of driving the state  $x$  of system (6.1) from an initial condition  $x_0$  to a final condition  $x_f$  in minimum time. Consider a Lebesgue integrable control function  $u = u(t)$  such that, for almost every  $t$ ,  $u(t)$  belongs to a compact set  $\mathcal{U} \subseteq \mathbf{R}^m$ . If  $u$  is the time optimal control in the interval  $[0, T]$  and  $x$  the associated trajectory solution of (6.1) in the interval  $[0, T]$ , then there exists an  $n$ -dimensional vector of functions  $\lambda$  called the costate, not identically zero and satisfying the adjoint equation*

$$\dot{\lambda}^T = -\lambda^T f_x(x, u).$$

Define the optimal control Hamiltonian,

$$h(\lambda, x, u) := \lambda^T f(x, u).$$

Then, for every  $t$  in  $[0, T]$ ,

$$h(\lambda(t), x(t), u(t)) \geq h(\lambda(t), x(t), v),$$

for every  $v \in \mathcal{U}$ . Moreover,  $h(\lambda(t), x(t), u(t))$  is constant along an optimal trajectory.

For quantum control systems, as described in the previous section, the state  $x$  and the costate  $\lambda$  satisfy the same differential equation (cf. (6.29) where  $\tilde{L}_x \equiv 0$ ).

---

<sup>5</sup>In fact, a general version of the PMP can be given which includes as special cases Theorem 6.2.1 and Theorem 6.4.2 (see, e.g., [81]).

### 6.4.1.1 Example of a time optimal quantum control problem

For some classes of systems, properties can be proved that simplify the calculation of the time optimal control. We consider as an example the class of quantum control systems (6.7) having the Hamiltonian  $H(u)$  of the form

$$H(u_1, u_2, v_1, v_2) := \begin{pmatrix} E_1 & u_1 + iu_2 & 0 \\ u_1 - iu_2 & E_2 & v_1 + iv_2 \\ 0 & v_1 - iv_2 & E_3 \end{pmatrix}. \quad (6.42)$$

Here the controls are  $u_1, u_2, v_1, v_2$ . The real constants  $E_1, E_2, E_3$  represent the natural energy levels of the quantum system in the absence of the controls. This class of models describes the dynamics of certain three level systems coupled with two laser fields, which are the controls  $u := (u_1, u_2)^T$  and  $v := (v_1, v_2)^T$ . The laser field  $u$  ( $v$ ) activates the coupling between levels 1 and 2 (2 and 3). A straightforward calculation of the dynamical Lie algebra for this system shows that it is operator controllable. This class of models was studied in a series of papers (see [28] and references therein). It can be generalized to an  $n$  level problem and it has several properties which are useful for the calculation of the time optimal control. The following proposition holds.<sup>6</sup>

**Proposition 6.4.3** ([28]) Consider the problem of driving the state  $\vec{\psi}$  of (6.7) from an initial state  $\vec{\psi}_0$  to a final state  $\vec{\psi}_f$  in minimum time, with controls  $\|u_1 + iu_2\| \leq 1$ , and  $\|v_1 + iv_2\| \leq 1$ . Then there always exists a minimizer in resonance, namely

$$(u_1 + iu_2)(t) := u(t)e^{i[(E_2 - E_1)t + \xi_1]}, \quad (6.43)$$

$$(v_1 + iv_2)(t) := v(t)e^{i[(E_3 - E_2)t + \xi_2]}, \quad (6.44)$$

with real functions  $u(t)$  and  $v(t)$  and real phase parameters  $\xi_1$  and  $\xi_2$ .

Let us consider the problem of steering the state  $\vec{\psi}$  from a state of the form  $\vec{\psi}_0 = [e^{i\eta_1}, 0, 0]^T$ , corresponding to an energy level  $E_1$  (i.e., an eigenstate  $|E_1\rangle$ ), to a state of the form  $\vec{\psi}_f = [0, 0, e^{i\eta_2}]^T$ , corresponding to an energy level  $E_3$  (i.e., an eigenstate  $|E_3\rangle$ ). Here  $\eta_1$  and  $\eta_2$  are arbitrary real parameters. It is convenient to perform a change of coordinates

$$\vec{r} = DU(t)\vec{\psi}, \quad (6.45)$$

with  $U(t) := \text{diag}(iE_1 t, iE_2 t, iE_3 t)$  and  $D$  an appropriate diagonal, constant, unitary matrix. This transforms system (6.7), (6.42) into the system

$$\frac{d}{dt}\vec{r} = -S_z u(t)\vec{r} - S_x v(t)\vec{r}, \quad (6.46)$$

---

<sup>6</sup>The properties in Proposition 6.4.3 are valid for more general systems and costs. In particular, the existence of a minimizer in resonance is a more general property (it also holds for the problem considered in section 6.3. We refer to [28] and the references therein.

where  $S_{x,y,z}$  are defined in (5.8) and satisfy the commutation relations (5.9). In deriving (6.46), we have used Proposition 6.4.3 and in particular we have restricted ourselves to controls in resonance of the form (6.43) (6.44). In particular  $u$  and  $v$  are the amplitudes appearing in (6.43) (6.44). The change of coordinates does not change the structure of the initial state and of the final state, and we can choose them real as the phase factors  $\eta_1$  and  $\eta_2$  are arbitrary. In particular, we can assume the initial value  $\vec{r}_0 = [1, 0, 0]^T$  and the desired final value  $\vec{r}_f := [0, 0, 1]^T$ . Using the result on the existence of a minimizer in resonance, the problem is therefore reduced to a real problem of time optimal control on the 2-dimensional sphere.<sup>7</sup> The problem has low dimension and one can apply the techniques of optimal control on 2-dimensional manifolds of [29]. The final result is summarized in the following theorem which is proved in [28].

**Theorem 6.4.4** *The optimal control to steer the state  $\vec{r}$  of system (6.46) between two points is a concatenation of controls  $(u, v)$  of the type  $(u, v) = (\pm 1, \pm 1)$ ,  $(u, v) = (\pm 1, 0)$ , and  $(u, v) = (0, \pm 1)$ . In particular, the time optimal control steering from  $\vec{r}_0 = [1, 0, 0]^T$  to  $\vec{r}_f = [0, 0, 1]^T$  is given by  $u \equiv 1$ ,  $v \equiv 1$  in the interval  $[0, \frac{\pi}{\sqrt{2}}]$  (cf. Figures 6.4, 6.5).*

We shall not report the complete proof of this theorem here but we shall add some considerations to show how the PMP determines the nature of the optimal controls. In particular the arguments that follow prove the first sentence of the theorem. The optimal control Hamiltonian of Theorem 6.4.2,  $h(\lambda, \vec{r}, u, v)$ , reads in this case as

$$h(\lambda, \vec{r}, u, v) = -u\lambda^T S_z \vec{r} - v\lambda^T S_x \vec{r},$$

and it is constant along the optimal trajectory. The adjoint equations for  $\lambda$  read as

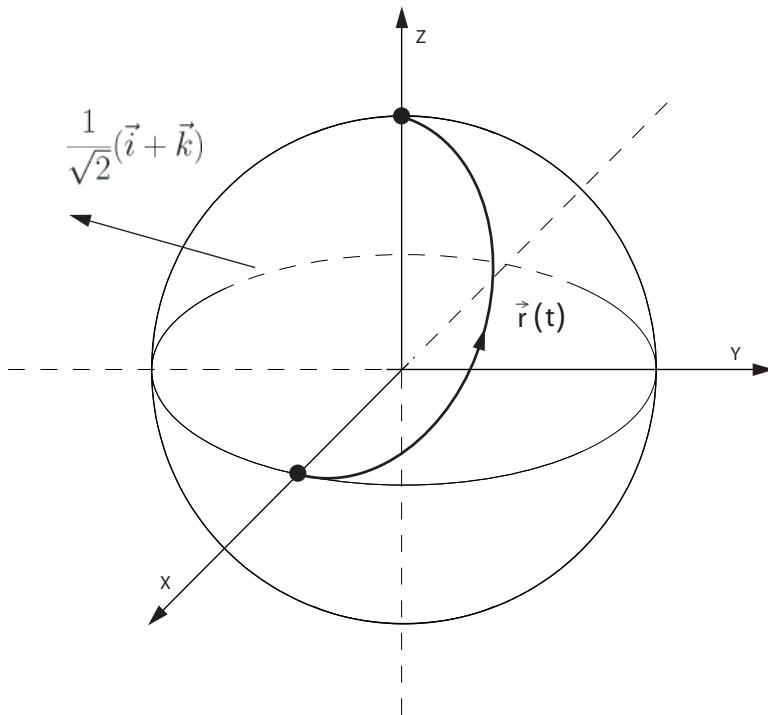
$$\dot{\lambda} = -uS_z\lambda - vS_x\lambda. \quad (6.47)$$

The procedure to find the time optimal control consists of calculating all the controls satisfying the PMP, i.e., the extremals, and then selecting, among them, the one which gives minimum time. To understand the nature of the extremals, notice that extremals which are identically zero, i.e.,  $u \equiv 0$  and  $v \equiv 0$ , in an interval of nonzero measure, must be excluded because they are clearly not time optimal. In fact the trajectory would ‘stop’ in some point for some time. From the PMP, we obtain that in intervals where  $\lambda^T S_z \vec{r} < 0$  ( $> 0$ ),  $u \equiv 1$  ( $u \equiv -1$ ), while in intervals where  $\lambda^T S_x \vec{r} < 0$  ( $> 0$ ),  $v \equiv 1$  ( $v \equiv -1$ ). We cannot have intervals where

$$\lambda^T S_z \vec{r} \equiv 0 \quad (6.48)$$

---

<sup>7</sup> As discussed in subsection 6.1.2, considering a real problem is always possible. However, in this case, we have a reduction of the dimension of the problem which is crucial in the actual solution.



**FIGURE 6.4:** Time optimal trajectory from the point  $\vec{r}_0 = [1, 0, 0]^T$  to  $\vec{r}_f = [0, 0, 1]^T$  according to Theorem 6.4.4. It is a rotation about the axis  $\frac{1}{\sqrt{2}}(\vec{i} + \vec{k})$ .

and

$$\lambda^T S_x \vec{r} \equiv 0, \quad (6.49)$$

because, in that case, by differentiating these expressions, using (6.46) and (6.47), and the fact that we have excluded extremals with  $u \equiv 0$  and  $v \equiv 0$ , we obtain

$$\lambda^T S_y \vec{r} \equiv 0. \quad (6.50)$$

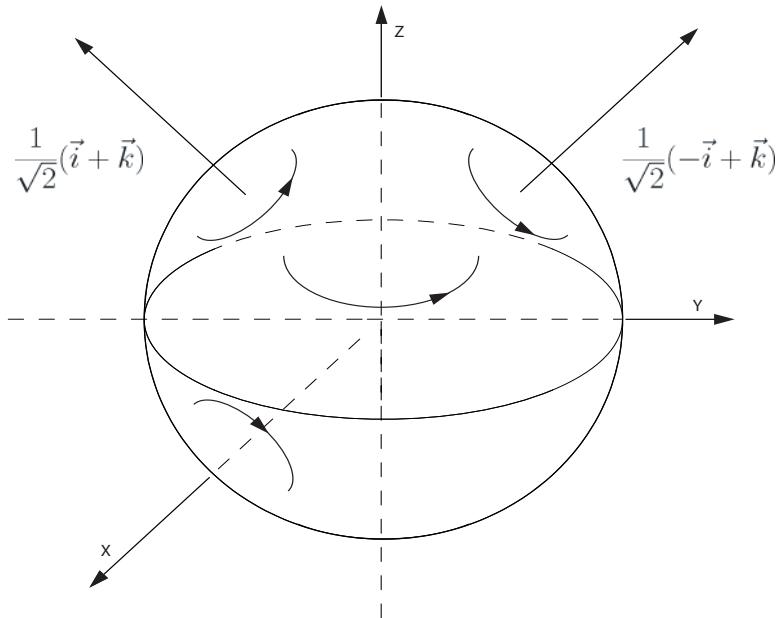
This together with (6.49) and (6.48) would give  $\lambda \equiv 0$ , which contradicts the PMP. If (6.48) is verified in some interval, differentiating (6.48) with (6.46) and (6.47) gives

$$v \lambda^T S_y \vec{r} \equiv 0,$$

and since (6.49) cannot be verified in the same interval, as just shown, we must have  $v \neq 0$  which implies (6.50). Differentiating (6.50), we obtain

$$-\lambda^T S_x \vec{r} u + \lambda^T S_z \vec{r} v \equiv 0,$$

which, with (6.48), and since  $\lambda^T S_x \vec{r} \neq 0$ , gives  $u \equiv 0$ . Analogously one proves that, when (6.49) is verified,  $v \equiv 0$ .



**FIGURE 6.5:** According to Theorem 6.4.4, every optimal trajectory is the concatenation of the four types of trajectories depicted in this picture. These are rotations about the  $x$  and  $z$  axes and about the directions of  $\frac{1}{\sqrt{2}}(\vec{i} + \vec{k})$  and  $\frac{1}{\sqrt{2}}(-\vec{i} + \vec{k})$  (all possibly in both directions).

#### 6.4.2 Minimum time control with unbounded control; Riemannian symmetric spaces

We consider now the problem of minimum time control with no a priori bound on the controls for the *Schrodinger operator equation*. The problem is to drive the *unitary evolution operator* to a desired target in minimum time. In several cases the problem can be solved explicitly using the theory of Riemannian symmetric spaces, some of which was discussed in 5.3, and concepts of Lie group theory and homogeneous spaces.

We set up the problem in general, namely, we look at a differential system on a general compact matrix Lie group  $e^{\mathcal{L}}$ , with semisimple Lie algebra  $\mathcal{L}$  of the form

$$\dot{X} = AX + \sum_{j=1}^m B_j X u_j, \quad X(0) = \mathbf{1}, \quad (6.51)$$

where  $\mathbf{1}$  is the identity matrix in the group,  $u_j$ ,  $j = 1, \dots, m$ , are the controls and the matrices  $A$  and  $B_j$ ,  $j = 1, \dots, m$  generate the Lie algebra  $\mathcal{L}$ . The problem is to steer  $X$  from the identity to a target  $X_f$  in minimum time and there is no a priori bound on the control. As the minimum time does not exist

in general for this problem, we shall look for the infimum time. In formulas, the infimum time  $t_f^* := t_f^*(X_f)$  is defined as

$$t_f^*(X_f) := \inf\{T \geq 0 | X_f \in \bar{\mathcal{R}}(\leq T)\},$$

with the reachable set  $\mathcal{R}(\leq T)$  defined in [Chapter 3](#) (cf. the discussion before formula (3.2)) and  $\bar{\mathcal{R}}$  here means the closure of the set  $\mathcal{R}$ . In words, for every  $\epsilon > 0$  it is always possible to drive the state arbitrarily close to  $X_f$  in time less than  $t_f^* + \epsilon$  and  $t_f^*$  is the smallest time for which this is true. The problem here is to find an expression for the function  $t_f^*(X_f)$  and then to find a control which will drive  $X$  from the identity arbitrarily close to  $X_f$  in time arbitrarily close to  $t_f^*(X_f)$ .

#### 6.4.2.1 The control subgroup

An important role, in this problem, is played by the Lie subgroup of  $e^{\mathcal{L}}$  corresponding to the Lie subalgebra of  $\mathcal{L}$  generated by the control matrices  $B_1, \dots, B_m$ , i.e.,  $\mathcal{K} := \{B_1, \dots, B_m\}_{\mathcal{L}}$ . This subgroup is called the *control subgroup* and the Lie algebra  $\{B_1, \dots, B_m\}_{\mathcal{L}}$  is called the *control subalgebra*. It has the feature that, for every  $K \in e^{\mathcal{K}}$ ,  $t_f^*(K) = 0$ . In order to see this, we show that, for every  $K \in e^{\mathcal{K}}$  there exists a control steering the evolution operator  $X$  of (6.51) arbitrarily close to  $K$  in arbitrarily short time. Consider the system corresponding to (6.51) but without the drift term, i.e.,

$$\dot{X}_c = \sum_{j=1}^m B_j X_c u_j, \quad X_c(0) = \mathbf{1}. \quad (6.52)$$

Every element  $K$  of  $e^{\mathcal{K}}$  can be reached by an appropriate control.<sup>8</sup>

Consider the control  $u = u_j(t)$ ,  $j = 1, \dots, m$ , steering  $X_c$  from the identity to  $K$ , in time  $T$ . Then the control  $Mu_j(Mt)$  steers to  $K$  in time  $\frac{T}{M}$ . In fact, it is easy to verify that the solution of (6.52) with this control in the interval  $[0, \frac{T}{M}]$  is  $X_c(Mt)$ , if  $X_c$  denotes the solution corresponding to the control  $u_j(t)$  in  $[0, T]$ . Consider now the control  $Mu_j(Mt)$  for system (6.51) and for system (6.52) and subtract one from the other to obtain

$$\dot{X} - \dot{X}_c = AX + \left( \sum_{j=1}^m B_j Mu_j(Mt) \right) (X - X_c).$$

<sup>8</sup>Notice that  $e^{\mathcal{K}}$  is a Lie subgroup of a compact subgroup  $e^{\mathcal{L}}$  and Theorem 3.2.1 applies, namely the set of reachable states for system (6.52) is  $e^{\mathcal{K}}$ . In fact, in the proof of Theorem 3.2.1, the property that  $U(n)$  is a compact Lie group was used. The proof goes through the same way for any compact Lie group and in particular it can be applied to this case since we are assuming that  $e^{\mathcal{L}}$  is compact. In fact, this assumption is not even necessary in this case. Since system (6.52) is driftless one can apply results for driftless systems [113] to obtain that the reachable set is  $e^{\mathcal{K}}$  independently of the fact that the ‘ambient’ Lie group ( $e^{\mathcal{L}}$ ) is compact.

Integrating this in the interval  $[0, \frac{T}{M}]$ , we obtain

$$\begin{aligned} & X\left(\frac{T}{M}\right) - X_c\left(\frac{T}{M}\right) \\ &= \int_0^{\frac{T}{M}} AX(t) dt + \int_0^{\frac{T}{M}} \left( \sum_{j=1}^m B_j M u_j(Mt) \right) (X(t) - X_c(t)) dt. \end{aligned}$$

Let  $P$  be a uniform bound on  $\|AX\|$  (which exists because the group  $e^{\mathcal{L}}$  is assumed to be compact) and let  $N$  be a uniform bound on  $\|\sum_{j=1}^m B_j u_j(Mt)\|$ . Then we have

$$\left\| X\left(\frac{T}{M}\right) - X_c\left(\frac{T}{M}\right) \right\| \leq P \frac{T}{M} + MN \int_0^{\frac{T}{M}} \|X(t) - X_c(t)\| dt.$$

By applying (a special case of) the Gronwall-Bellman inequality<sup>9</sup> (cf., e.g., [81], [Appendix A](#)), we obtain

$$\|X\left(\frac{T}{M}\right) - X_c\left(\frac{T}{M}\right)\| \leq P \frac{T}{M} + PTN \int_0^{\frac{T}{M}} e^{MN(\frac{T}{M}-r)} dr,$$

which gives

$$\left\| X\left(\frac{T}{M}\right) - X_c\left(\frac{T}{M}\right) \right\| \leq P \frac{T}{M} + PT \frac{e^{TN}}{M} (1 - e^{-NT}).$$

As  $X_c(\frac{T}{M}) = K$  for every  $M$ , letting  $M \rightarrow \infty$  we obtain the desired result.

With  $K \in e^{\mathcal{K}}$ , for every  $X_f \in e^{\mathcal{L}}$ , we have

$$t_f^*(X_f) = t_f^*(KX_f) = t_f^*(X_fK), \quad (6.54)$$

because after, or before, having driven to  $X_f$  we can produce  $K$  in arbitrarily small time. We summarize this discussion in terms of coset spaces (cf. the definition in [subsection 3.4.2](#)) in the following proposition.

**Proposition 6.4.5** Let  $\pi$  be the natural projection (cf. Definition (3.14))

$$\pi : e^{\mathcal{L}} \rightarrow e^{\mathcal{L}}/e^{\mathcal{K}}.$$

If  $\pi(X_1) = \pi(X_2)$ , then  $t_f^*(X_1) = t_f^*(X_2)$ . In particular, if  $\pi(X_1) = \pi(\mathbf{1})$ , i.e.,  $X_1$  belongs to  $e^{\mathcal{K}}$ , then  $t_f^*(X_1) = 0$ .

---

<sup>9</sup>

$0 \leq m(t) \leq k + h \int_0^t m(s) ds \rightarrow m(t) \leq k + \int_0^t hke^{h(t-r)} dr \quad (6.53)$

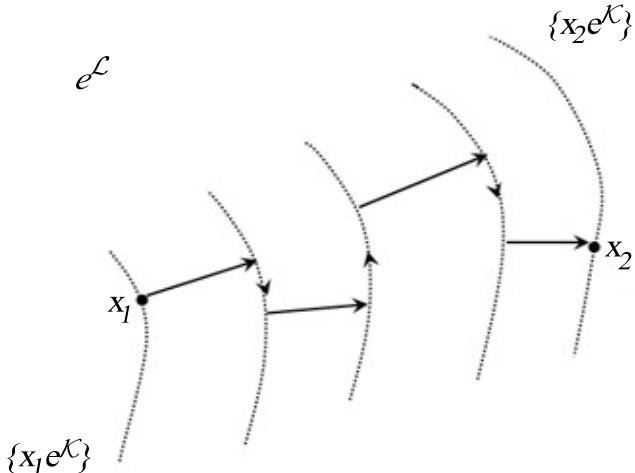
### 6.4.2.2 Motion in the coset space $e^{\mathcal{L}}/e^{\mathcal{K}}$

Consider now a control  $u_j$ ,  $j = 1, \dots, m$  and the corresponding trajectory  $X = X(t)$  solution of (6.51). Consider the projection of such a trajectory onto the coset space  $e^{\mathcal{L}}/e^{\mathcal{K}}$ ,  $\pi(X)$ . Let us in particular consider right cosets. If  $X$  moves for time  $T > 0$  inside a coset, i.e.,  $\pi(X)$  is constant, we can always find another control and trajectory which will give the same resulting state transfer in time less than  $T$ . In other terms, motion inside a coset can be obtained approximately in no time. To look for the minimum time control we have to look for the fastest way to go from the initial coset  $\{1e^{\mathcal{K}}\}$  to the final coset  $\{X_f e^{\mathcal{K}}\}$  (cf. Figure 6.6). So we look at the minimum time problem for the associated system on  $e^{\mathcal{L}}/e^{\mathcal{K}}$ . To study the motion of  $\pi(X)$ , i.e., the motion in the coset space  $e^{\mathcal{L}}/e^{\mathcal{K}}$ , we write the solution  $X = X(t)$  of (6.51) as

$$X(t) := X_c(t)U(t), \quad (6.55)$$

where  $X_c$  is the solution of (6.52) and  $U(t)$  is the solution of the *auxiliary system*:

$$\dot{U} = (X_c^{-1}AX_c)U, \quad U(0) = \mathbf{1}. \quad (6.56)$$



**FIGURE 6.6:** Motion in the Lie Group  $e^{\mathcal{L}}$ : The dotted curves represent cosets. In going from a coset  $\{x_1 e^{\mathcal{K}}\}$  to a coset  $\{x_2 e^{\mathcal{K}}\}$ , motion along dotted curves is arbitrarily fast while motion along the solid arrows is slow.

In fact, by differentiating the right hand side of (6.55) using (6.56) and

(6.52), one obtains that  $X_c(t)U(t)$  satisfies (6.51), and by uniqueness of the solution of a differential equation, it is equal to  $X$ . In the factorization (6.55),  $U(t)$  describes how  $X(t)$  moves from one coset space to the other, while  $X_c$  only contributes motion inside a right coset.

One way to approach the problem of infimum time to  $X_f$  for system (6.51) is to study the minimum time problem for the auxiliary system (6.56) to reach the coset  $\{X_f e^{\mathcal{K}}\}$ . In this problem, the control is the set of matrix functions with values in the *adjoint orbit*  $\mathcal{O}_{\mathcal{K}}(A)$  defined by

$$\mathcal{O}_{\mathcal{K}}(A) := \{X_c^{-1}AX_c | X_c \in e^{\mathcal{K}}\}. \quad (6.57)$$

For this problem, we denote the infimum time to reach the coset  $\{e^{\mathcal{K}}X_f\}$  by  $t_A^*(\{e^{\mathcal{K}}X_f\})$ , analogously to how we have defined the infimum time  $t_f^*(X_f)$  for the original problem. Denote by  $\mathcal{B}$ ,  $\mathcal{B}(T)$  and  $\mathcal{B}(\leq T)$  the reachable sets of the system (6.56), i.e.,  $\mathcal{B}(\leq T) := \cup_{0 \leq t \leq T} \mathcal{B}(t)$  and  $\mathcal{B} := \cup_{T \geq 0} \mathcal{B}(T)$ , with  $\mathcal{B}(t)$  the set of points reachable at time  $t$ . We have

$$t_A^*(\{e^{\mathcal{K}}X_f\}) := \inf\{t \geq 0 | \exists K \in e^{\mathcal{K}}, KX_f \in \bar{\mathcal{B}}(\leq t)\}.$$

In words,  $t_A^*(\{e^{\mathcal{K}}X_f\})$  is the infimum time to drive the state of (6.56) from the identity to arbitrarily close to a coset containing  $X_f$  by using controls with values in  $\mathcal{O}_{\mathcal{K}}(A)$ . The calculation of  $t_f^*(X_f)$  is equivalent to the calculation of  $t_A^*(\{e^{\mathcal{K}}X_f\})$ , as the following equivalence result holds.

#### Proposition 6.4.6

$$t_f^*(X_f) = t_A^*(\{e^{\mathcal{K}}X_f\}).$$

*Proof.* (Sketch) By definition of  $t_A^*(\{e^{\mathcal{K}}X_f\})$ , given  $\epsilon > 0$ , there must exist a  $K \in e^{\mathcal{K}}$ , a  $T \leq t_A^*(\{e^{\mathcal{K}}X_f\}) + \epsilon$  and a piecewise continuous (control) trajectory,  $X_c = X_c(t)$ , such that the corresponding solution  $U(t)$  of (6.56) is arbitrarily close to  $KX_f$  at time  $T$ . From results in geometric control theory (see, e.g., Theorem 3.2. in [128]) we can find a control  $u(t)$  such that the *trajectory* of (6.52) is arbitrarily close (in an appropriate topology) to  $X_c = X_c(t)$  and therefore the corresponding trajectory of (6.56) is arbitrarily close to  $KX_f$  at time  $T$ . The solution of the original system (6.51) is then arbitrarily close to  $K_1X_f$  for some  $K_1 \in e^{\mathcal{K}}$ , and therefore

$$t_f^*(\{K_1X_f\}) = t_f^*(X_f) \leq t_A^*(\{e^{\mathcal{K}}X_f\}).$$

To show

$$t_f^*(X_f) \geq t_A^*(\{e^{\mathcal{K}}X_f\}), \quad (6.58)$$

notice that if there exists a control steering the solution  $X$  of (6.51) arbitrarily close to  $X_f$  in time  $T \leq t_f^*(X_f) + \epsilon$ , then, with the same control, the solution of (6.56) will be arbitrarily close to  $KX_f$  for some  $K \in e^{\mathcal{K}}$ . This gives the inequality (6.58).  $\square$

Proposition 6.4.6 gives a way of studying the problem of minimum time control for system (6.51). This means to study the minimum time problem to steer the state of the auxiliary system (6.56) between two right cosets using controls which are constrained to be in the adjoint orbit (6.57).

#### 6.4.2.3 Determination of the minimum time in the case where $e^{\mathcal{L}}/e^{\mathcal{K}}$ is a symmetric space

We now examine how the problem of minimum time control for system (6.56) specializes in the case where  $e^{\mathcal{L}}/e^{\mathcal{K}}$  is a symmetric space. Let  $\mathcal{L}$  be semisimple and have the Cartan decomposition

$$\mathcal{L} = \mathcal{K} \oplus \mathcal{P},$$

where  $\mathcal{K}$  and  $\mathcal{P} = \mathcal{K}^\perp$  satisfy the commutation relations (5.24), (5.25), (5.26). Consider  $A$  in (6.51) and assume  $A \in \mathcal{P}$ . Let  $\mathcal{A}$  be a Cartan subalgebra containing  $A$  (it always exists because of Theorems 5.3.4, 5.3.5) and define the *Weyl orbit*  $\mathcal{W}(A)$  to be the set of elements in the adjoint orbit  $\mathcal{O}_{\mathcal{K}}(A)$  that also belong to  $\mathcal{A}$ , i.e.,

$$\mathcal{W}(A) := \mathcal{A} \cap \mathcal{O}_{\mathcal{K}}(A).$$

The *positive span* of  $\mathcal{W}(A)$ ,  $\text{span}_{\geq}(\mathcal{W}(A))$ , is defined as

$$\text{span}_{\geq}(\mathcal{W}(A)) := \left\{ \sum_j \beta_j A_j \mid A_j \in \mathcal{W}(A), \beta_j \geq 0 \right\}.$$

We make the following assumption

**Assumption 1:**

$$\text{span}_{\geq}(\mathcal{W}(A)) = \mathcal{A}.$$

This assumption is certainly satisfied if  $\tilde{A} \in \mathcal{W}(A)$  implies  $-\tilde{A} \in \mathcal{W}(A)$  and  $\text{span}(\mathcal{W}(A)) = \mathcal{A}$ , as it will be the case for the one and two spin examples below.

If Assumption 1 is verified, then, from the Cartan decomposition (5.31) of Theorem 5.3.7, we have that, for every element  $X_f$ , there exist elements  $K_1$  and  $K_2$  in  $e^{\mathcal{K}}$  and commuting elements in the adjoint orbit  $\mathcal{O}_{\mathcal{K}}$ ,  $A_j$ ,  $j = 1, 2, \dots, p$ , along with positive coefficients  $\beta_j = 1, \dots, p$  so that

$$X_f = K_1 e^{\sum_{j=1}^p \beta_j A_j} K_2. \quad (6.59)$$

The following theorem and remark solve the problem of calculating  $t_f^*(X_f)$  and giving the minimum time control in the case of a Riemannian symmetric space when Assumption 1 is verified. The proof of the theorem is given in [118], [119].

**Theorem 6.4.7** Under the above assumptions,  $t_f^*(X_f)$  is the minimum value of  $\sum_{j=1}^p \beta_j$  for which it is possible to write (6.59), with commuting  $A_j \in \mathcal{O}_K(A)$ .

**Remark 6.4.8** As  $A_j = X_j^{-1}AX_j$ ,  $j = 1, \dots, p$ , for some  $X_j \in e^K$ , a time minimum (infimum) control and trajectory can be obtained by driving  $X$  first to  $K_2$  with a very large, very short, pulse, then to  $X_1$  with the same method, then setting the control equal to zero for a time  $\beta_1$ , so as to produce  $e^{A\beta_1}$ . Then a new large and short pulse is used to drive the solution  $X$  of (6.51) from the identity to arbitrarily close to  $X_2X_1^{-1} \in e^K$ . Then the control is set equal to zero for time  $\beta_2$  and so on. The last large and short pulse is used to generate  $K_1$  in (6.59). This control achieves the transfer of the solution  $X$  of (6.51) from the identity  $\mathbf{1}$  to arbitrarily close to  $X_f$  in time arbitrarily close to  $t_f^*(X_f)$  given in Theorem 6.4.7.

The main idea behind Theorem 6.4.7 is as follows. Given a decomposition (6.59) of  $X_f$ , we have from (6.54) that  $t_f^*(X_f) = t_f^*(e^{\sum_{j=1}^p \beta_j A_j})$ , and by the equivalence result of Proposition 6.4.6 we can calculate the infimum time  $t_A^*(\{e^K e^{\sum_{j=1}^p \beta_j A_j}\})$  for the auxiliary system (6.56) to reach the coset containing  $e^{\sum_{j=1}^p \beta_j A_j}$ . We can, without loss of generality, assume for the system (6.56) ‘controls’  $X_c^{-1}AX_c$ , piecewise constant. Moreover we need to have all the control matrices  $X_c^{-1}AX_c$  commuting for the trajectory to be time optimal. In fact, if two of them, say  $A_k$  and  $A_l$ , do not commute, then write, using (for example) the exponential formula (E.5) of Appendix E,

$$e^{A_k \frac{t}{n}} e^{A_l \frac{t}{n}} = e^{\frac{t}{n}(A_k + A_l) + \frac{t^2}{2n^2}[A_k, A_l] + o(\frac{1}{n^3})}.$$

We now use the assumption that  $e^\mathcal{L}/e^K$  is a Riemannian symmetric space. From (5.26), we obtain that  $[A_k, A_l] \in \mathcal{K}$  and therefore there is a component of the motion in  $e^K$  which we could have obtained in a much quicker way by using large and short pulses. Therefore, in this case, the control cannot be optimal. The candidate optimal final points for system (6.56) are all of the form  $e^{\sum_{j=1}^p \beta_j A_j}$ , with commuting  $A_j$ ’s, satisfying (6.59), and the corresponding time is  $\sum_{j=1}^p \beta_j$ . Among those, the infimizing time is the minimum  $\sum_{j=1}^p \beta_j$  so that (6.59) holds.

#### 6.4.2.4 One and two spin example

Consider now the model of a spin- $\frac{1}{2}$  particle in a magnetic field studied in Chapter 2, subsection 2.3.2, in section 5.1.4 of the previous chapter, and in section 6.3. In particular, let us assume that we have a constant magnetic field in the  $y$  direction and a varying control magnetic field in the  $z$  direction, without any a priori bound on its magnitude. After a re-scaling of the control and of time we can write the Schrödinger operator equation (cf. (5.19)) as

$$\dot{X} = (\bar{\sigma}_y + u(t)\bar{\sigma}_z)X,$$

with  $\bar{\sigma}_{x,y,z}$  defined in (5.5). The Cartan decomposition of the dynamical Lie algebra  $\mathcal{L} = su(2)$  is  $\mathcal{L} = \mathcal{K} \oplus \mathcal{P}$  with the Lie algebra  $\mathcal{K} := \text{span}\{\bar{\sigma}_z\}$  and  $\mathcal{P} := \text{span}\{\bar{\sigma}_y, \bar{\sigma}_x\}$ . It induces a decomposition of the Lie group  $SU(2)$  which is the Euler decomposition of subsection 5.1.2. The rank of the corresponding symmetric space is one. The Weyl orbit  $\mathcal{W}(\bar{\sigma}_y)$  is given by

$$\mathcal{W}(\bar{\sigma}_y) = \{\bar{\sigma}_y, -\bar{\sigma}_y\}. \quad (6.60)$$

Applying Theorem 6.4.7, the minimum time to reach a final state  $X_f$  is the smallest nonnegative  $t_2$  (Euler angle) so that we can write (cf. (5.13))

$$X_f = e^{\bar{\sigma}_z t_3} e^{\bar{\sigma}_y t_2} e^{\bar{\sigma}_z t_1}.$$

Methods to calculate the Euler angles were discussed in subsection 5.1.3. It follows from the discussion in that subsection that the infimum time  $t_f^*(X_f)$  is to be taken in the interval  $[0, 2\pi]$ .

The situation is only slightly more complicated in the case of two spin- $\frac{1}{2}$  particles interacting through Ising interaction (cf. subsection 2.3.2). We assume that it is possible to address each one of the two spins separately with magnetic fields  $u$  and  $v$  in independent directions so that the model is written as

$$\dot{X} = (i\sigma_z \otimes \sigma_z)X + \left( i \left( \sum_{j=x,y,z} u_j \sigma_j \right) \otimes \mathbf{1} + i\mathbf{1} \otimes \left( \sum_{j=x,y,z} v_j \sigma_j \right) \right) X,$$

where  $\sigma_{x,y,z}$  are the Pauli matrices (1.20). In this case, the relevant Cartan decomposition is the one described in Example 5.3.10 and the rank of the associate symmetric space is 3. The Weyl orbit  $\mathcal{W}(i\sigma_z \otimes \sigma_z)$  is given by

$$\mathcal{W}(i\sigma_z \otimes \sigma_z) = \{\pm i\sigma_z \otimes \sigma_z, \pm i\sigma_x \otimes \sigma_x, \pm i\sigma_y \otimes \sigma_y\}. \quad (6.61)$$

Theorem 6.4.7 applies and it says that  $t_f^*(X_f)$  is the minimum value of  $\beta_x + \beta_y + \beta_z$  for which we can write

$$X_f = L_1 e^{\sum_{j=x,y,z} i\beta_j \sigma_j \otimes \sigma_j} L_2,$$

with  $L_{1,2}$  of the form  $L_{1,2} = K_{1,2} \otimes R_{1,2}$  with  $K_{1,2}$  and  $R_{1,2}$  in  $SU(2)$ . Methods to calculate such a decomposition were discussed in subsection 5.3.6.

## 6.5 Numerical Methods for Optimal Control of Quantum Systems

There exists a wealth of numerical methods to solve optimal control problems. Some of them directly aim at finding the optimal control in a specific

problem, some others are useful in the solution of auxiliary problems, as for example the integration of the adjoint equations of the maximum principle. We summarize in the following subsections some of the main ideas with emphasis on application to quantum systems. The following two sections describe, respectively, methods to find the optimal control by discretization and by iterative algorithms. These iterative methods generate a sequence of control functions converging to the optimal control or to an optimal candidate. [Subsection 6.5.3](#) deals with the numerical solution of two points boundary value problems arising in the application of the PMP.

### 6.5.1 Methods using discretization

Given the (realification of a) quantum system

$$\dot{x} = \tilde{H}(u)x, \quad (6.62)$$

and a cost of the Bolza type

$$J(u) = \phi(x(T)) + \int_0^T L(x(t), u(t))dt,$$

we can choose a time step  $\Delta t$  and a positive integer  $N$  so that  $N\Delta t = T$ .  $\Delta t$  has to be chosen small if the relevant dynamics of (6.62) is fast. Defining  $u_k := u(k\Delta t)$ ,  $x_k := x(k\Delta t)$ ,  $k = 0, \dots, N$ , we write

$$x_{k+1} \approx x_k + (\Delta t)\tilde{H}(u_k)x_k, \quad (6.63)$$

$$J(u) \approx \phi(x_N) + \sum_{k=0}^{N-1} L(x_k, u_k)\Delta t. \quad (6.64)$$

If  $n$  is the dimension of the state vector  $x$  and  $m$  the dimension of the control vector  $u$ , there are  $Nn +Nm$  variables,  $x_1, \dots, x_N$ ,  $u_0, \dots, u_{N-1}$ , in (6.64), (6.63). We wish to find the values that minimize the function on the right hand side of (6.64) subject to the constraint in (6.63). This constraint can be eliminated and the problem transformed into an unconstrained optimization problem by defining  $G(u_k) := I + (\Delta t)\tilde{H}(u_k)$  and setting, from (6.63),

$$x_k = G(u_{k-1})G(u_{k-2}) \cdots G(u_1)G(u_0)x_0.$$

Placing this into the function in (6.64) we only have a function of  $u_0, \dots, u_{N-1}$  to be minimized. Standard numerical methods of unconstrained optimization, such as *steepest descent method* (see, e.g., [65]) can be used for this problem.

### 6.5.2 Iterative methods

Iterative methods for optimal control generate sequences of control laws that converge to extremals satisfying the necessary conditions of optimality

(the PMP). Several proposals for such algorithms have been put forward and theoretical and numerical studies exist on the convergence properties of such algorithms (see, e.g., [90], [140], [186], [228], [229]).

Let us consider the case of the cost (6.13) and let us assume the common situation in molecular control where the available control is only one laser field. The dynamics can be written as

$$\frac{d}{dt}\vec{\psi} = (A + uB)\vec{\psi},$$

where  $\vec{\psi}$  is the pure state,  $u$  the control field and  $A$  and  $B$  are matrices in  $su(n)$ . Transforming the problem into a real one (cf. [subsection \(6.1.2\)](#)), the cost takes the form

$$\frac{1}{2}x^T \tilde{O}x + \frac{k}{2} \int_0^T u^2(t) dt,$$

where  $x$  is defined in (6.9), and  $\tilde{O}$  is the symmetric negative definite matrix

$$\tilde{O} = \begin{pmatrix} O_R & -O_I \\ O_I & O_R \end{pmatrix},$$

with  $O = O_R + iO_I$ ,  $O_R = O_R^T$ ,  $O_I^T = -O_I$ . The dynamics take the form

$$\dot{x} = \tilde{A}x + \tilde{B}xu, \quad x(0) = x_0 \quad (6.65)$$

with (cf. (6.10))

$$\tilde{A} = \begin{pmatrix} A_R & -A_I \\ A_I & A_R \end{pmatrix}, \tilde{B} = \begin{pmatrix} B_R & -B_I \\ B_I & B_R \end{pmatrix}, A = A_R + iA_I, B = B_R + iB_I,$$

and  $x_0$ , the appropriate initial condition. The necessary conditions of [subsection 6.2.2](#) give, in this case, that there exists a nonzero costate vector  $\lambda$  satisfying

$$\dot{\lambda} = \tilde{A}\lambda + \tilde{B}\lambda u, \quad (6.66)$$

with terminal condition

$$\lambda(T) = -\tilde{O}x(T), \quad (6.67)$$

and so that, given  $x = x(t)$  and  $\lambda = \lambda(t)$ ,

$$\lambda^T(\tilde{A}x + \tilde{B}xu) - \frac{k}{2}u^2 \geq \lambda^T(\tilde{A}x + \tilde{B}xv) - \frac{k}{2}v^2,$$

for every admissible value  $v$ . If we do not assume any bound on the value of the control  $u$ , we can calculate the maximum by taking the derivative of the right hand side with respect to  $v$  and setting it equal to zero. This gives

$$u(t) = \frac{1}{k}\lambda^T(t)\tilde{B}x(t). \quad (6.68)$$

Iterative methods generate sequences of functions  $\{u^k(t), x^k(t), \lambda^k(t)\}$  which converge to a triple  $\{u(t), x(t), \lambda(t)\}$  satisfying the equations of the maximum principle, namely (6.65), (6.66), with initial condition given by (6.67).

The most natural and simple iterative method is as follows: Starting with a control  $u^0 = u^0(t)$ , one integrates (6.65) forward to obtain  $x^1 = x^1(t)$ . Then using  $x^1$  instead of  $x$  in the expression of the terminal condition (6.67), one integrates (6.66) backwards to obtain  $\lambda^1 = \lambda^1(t)$ . Then  $x^1$  and  $\lambda^1$  are used in (6.68) to calculate the new estimate for the control  $u^1 = \frac{1}{k} \lambda^{1T} \tilde{B} x^1$ . This process is then iterated. Unfortunately, such a simple and intuitive iterative scheme does not converge in several situations. This has stimulated much research to design iterative algorithms which are guaranteed to converge. The algorithm given in [140] contains as special cases the Krotov algorithm [201] and its modification, the Zhu-Rabitz algorithm [228]. We describe it here as an example of an iterative algorithm for quantum optimal control. For applications and results of numerical simulations we refer the reader to the original papers.

### Algorithm 1 ([140])

The iteration formulas are as follows. Choose two constants  $\delta, \eta \in [0, 2]$ ,

$$\dot{x}^k = \tilde{A}x^k + \tilde{B}x^k u^k, \quad x^k(0) = x_0, \quad (6.69)$$

$$u^k = (1 - \delta)v^{k-1} + \frac{\delta}{k}\lambda^{k-1T} \tilde{B}x^k, \quad (6.70)$$

$$\dot{\lambda}^k = (\tilde{A} + \tilde{B}v^k)\lambda^k, \quad \lambda^k(T) = -\tilde{O}x^k(T), \quad (6.71)$$

$$v^k = (1 - \eta)u^k + \frac{\eta}{k}(\lambda^{kT} \tilde{B}x^k). \quad (6.72)$$

One starts with a guess for an auxiliary control  $v^0$  and a guess for the costate  $\lambda^0$ . Using (6.70) in (6.69) and integrating forward the resulting nonlinear equation, one obtains  $x^1$ , which in turn, placed in (6.70), gives  $u^1$ . Now,  $u^1$  and  $x^1$ , with (6.72), are used to integrate backwards (6.71). This produces  $\lambda^1$  and using (6.72)  $v^1$ , which are then used in place of  $\lambda^0$  and  $v^0$  to run the iteration. This scheme has the following convergence property [140]:

**Theorem 6.5.1** *Algorithm 1 is such that*

$$J(u^{k+1}) \leq J(u^k).$$

In particular, since  $J(u)$  is bounded from below,  $J(u^k)$  will converge. Notice that, in principle, there is no guarantee that the sequence of control laws will converge to a *global* minimum. Experience has however shown that in many cases the result of a numerical optimal control algorithm turns out to be the actual optimal control [166].

### 6.5.3 Numerical methods for two points boundary value problems

As we have observed in the discussion after Theorem 6.2.1, in subsection 6.2.1, the application of the maximum principle requires the solution of a system of differential equations with conditions at the initial time as well as at the final time. In general, we would like to have conditions at the initial time only so that we can integrate the equation forward. The search for an appropriate initial condition can be cast in the following class of problems.

Consider a differential equation

$$\dot{y} = f(y(t), t), \quad (6.73)$$

and assume we want to find an initial condition  $y(0) = y_0$  that, at the final time  $y(T, y_0)$ , satisfies

$$\tilde{F}(y_0) := F(y(T, y_0)) = 0, \quad (6.74)$$

where  $F$  has the same dimension as  $y$ . This algebraic equation can be solved numerically. For example, a Newton iteration (see, e.g., [65]) gives

$$y_0^{k+1} = y_0^k - \left[ \frac{\partial \tilde{F}}{\partial y_0} \Big|_{y_0=y_0^k} \right]^{-1} \tilde{F}(y_0^k). \quad (6.75)$$

To obtain  $\frac{\partial \tilde{F}}{\partial y_0} \Big|_{y_0=y_0^k}$ , we calculate

$$\frac{\partial \tilde{F}}{\partial y_0} \Big|_{y_0=y_0^k} = \frac{\partial F}{\partial y} \frac{\partial y(T)}{\partial y_0} \Big|_{y_0=y_0^k}.$$

In this formula  $\frac{\partial y(t)}{\partial y_0} \Big|_{y_0=y_0^k}$  describes the dependence of the solution  $y(t)$  of (6.73) on the initial condition when the initial condition is  $y_0^k$ . It is the solution of the *variational equation* associated with (6.73)

$$\frac{d}{dt} \frac{\partial y(t)}{\partial y_0} \Big|_{y_0=y_0^k} = f_y(y(t, y_0^k, t)) \frac{\partial y(t)}{\partial y_0} \Big|_{y_0=y_0^k}. \quad (6.76)$$

This suggests the following algorithm to find the vector  $y_0$  satisfying (6.74).

#### Algorithm 2

- Step 1: Choose the initial guess  $y_0^0$  for  $y_0$  and set  $k = 0$ .
- Step 2: Integrate (6.73) with the current value of  $y_0$ .
- Step 3: Check (6.74). If it is satisfied, up to some a priori tolerance STOP. Otherwise go to Step 4.

- Step 4: Integrate the variational equation (6.76).

- Step 5: Use  $\frac{\partial y(T)}{\partial y_0}|_{y_0=y_0^k}$  calculated in the previous step to update the estimate of  $y_0$  according to (6.75). Also, increment  $k$  by 1 and go to Step 2.
- 

## 6.6 Notes and References

Optimal control of quantum mechanical systems has been considered in many papers, especially in connection with application to molecular dynamics. Early work was done by A. P. Peirce, M. A. Dahleh, and H. Rabitz in [159], and many studies have been presented since then. The books [174], [192] present several examples of applications to molecular control. There exist more general versions of the maximum principle which include, as special cases, Theorems 6.2.1 and 6.4.2. These can be found in [81] or other advanced books on optimal control such as [132], [135], [161]. The problem of optimal control for two level systems was treated in [62], and several generalizations have been given (see, e.g., [27]). A general result of [27] states that minimizers are in resonance for a class of systems, which includes as a special case the one treated in subsection 6.4.1. Moreover this is true also for the control problem of minimizing the energy in a given interval subject to a terminal condition on the state as well as for the problem of minimizing the time of transfer subject to an energy constraint. The geometric treatment of minimum time problems for quantum systems with no a priori bound on the magnitude of the control in subsection 6.4.2 follows [118], [119]. Similar ideas were used for the treatment of systems other than the one and two spin case, for example for a system with three spin- $\frac{1}{2}$  particles in [121]. There are many papers on the subject of iterative algorithms for optimal control of quantum systems, and the subject still poses many theoretical and practical problems. In particular, there are several unsolved mathematical questions concerning not only the convergence of the given algorithms, but also the analysis of the speed of convergence, as well as the properties of the sets of limit points. For the subject of iterative numerical algorithms for quantum control we refer to [69], [90], [140], [166], [174], [186], [201], [228], [229]. For the discussion of numerical methods for two point boundary value problems of subsection 6.5.3, we followed [124].

## 6.7 Exercises

**Exercise 6.1** Consider the three level problem

$$\frac{d}{dt}\vec{\psi} = A\vec{\psi} + Bu\vec{\psi},$$

with

$$A := \begin{pmatrix} i & 0 & 0 \\ 0 & -i & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix},$$

with cost

$$J(u) = ||\vec{\psi}(T) - \vec{\psi}_f||^2 + \int_0^T u^2(t) dt,$$

with  $T = 2$  and  $\vec{\psi}_f = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$  and initial condition  $\vec{\psi}(0) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ .

- a) Transform this problem into an equivalent real problem of Mayer.
- b) Write the associate adjoint equation for the 6-dimensional costate  $\lambda$  and use the maximum principle to obtain an expression of the optimal control in terms of state and costate.
- c) Specialize Algorithm 1 to this problem.

**Exercise 6.2** Consider the example of section 6.3 and notice that the form of the controls (6.40)-(6.41) does not depend on the nonintegral part of the cost in (6.31). Therefore we can replace the cost with

$$J'(u_x, u_y) = -||\vec{\psi}^\dagger(T)\vec{\psi}_f|| + \eta \int_0^T u_x^2(t) + u_y^2(t) dt. \quad (6.77)$$

Set the desired state  $\vec{\psi}_f = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  and the initial state  $\vec{\psi}(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ . Notice that the cost does not change if we replace  $\vec{\psi}_f$  with  $\vec{\psi}_f e^{ir}$  for some real value  $r$ . By plugging the controls (6.40)-(6.41) into the differential equations (6.30) the latter equations can be integrated explicitly. Find the parameters  $\omega$ ,  $\gamma$  and  $M$  in (6.40)-(6.41) which minimize the cost (6.77). Discuss the dependence of the minimum cost on  $T$  and  $\eta$ .

**Exercise 6.3** Verify formulas (6.60) and (6.61). How does the treatment of the two spin case go if we replace Ising interaction with Heisenberg interaction (2.67)?

**Exercise 6.4** Prove that the map  $\phi : u(n) \rightarrow so(2n) \cap sp(n)$ , defined by

$$\phi(R + iI) := \begin{pmatrix} R & -I \\ I & R \end{pmatrix},$$

where  $I$  and  $R$  are real matrices is a Lie algebra isomorphism.

**Exercise 6.5** How does Proposition 6.4.1 modify in the case of infimum (rather than minimum) time?

# Chapter 7

---

## More Tools for Quantum Control

In this chapter, we discuss some more tools for the control of quantum systems. Some of these are commonly considered in the physics and chemistry literature. Some others, like the Lyapunov control method, have been very well studied in the control literature and have then been extended to quantum systems.

To every level of energy of a quantum system there corresponds an eigenstate (or an eigenspace). Transferring population between two levels means transferring the state from one eigenstate to the other or, more generally, increasing the probability of one value of energy as compared to the other. We start by explaining the fact that to induce a population transfer between two levels, we need to use a field (approximately) at the frequency corresponding to the energy difference between the two levels. This way, selective population transfer is obtained between two energy levels. This is discussed in section 7.1. In section 7.2, we briefly review *time varying perturbation theory*, which is a powerful tool to analyze quantum dynamics. Time varying perturbation theory also gives an alternative illustration of the above described behavior of the solution of the Schrödinger equation in terms of the frequency of the control. In section 7.3, we study the technique of *adiabatic control* and the *adiabatic approximation* on which it is based. The main idea here is that, if the Hamiltonian varies very slowly, the evolution of the state vector will approximately follow an eigenstate of the Hamiltonian. Both the ideas of selective population transfer via frequency tuning and adiabatic control are used in the technique called STIRAP (*STImulated Raman scattering involving Adiabatic Passage*), which can be used to induce population transfer between two levels which are not directly coupled. This is discussed in section 7.4. Section 7.5 summarizes the technique of Lyapunov control as applied to quantum control systems.

---

### 7.1 Selective Population Transfer via Frequency Tuning

A common technique in the control of quantum systems is to transfer population between two energy levels using a field whose frequency is tuned at the

energy difference between the selected levels. This idea is at the foundation of quantum optics and the theory of absorption and emission of radiation by atoms (see, e.g., [138]). When one considers the electromagnetic field itself as a quantum system, its energy can assume a discrete set of values. For a monochromatic field the possible energy values differ by integer multiples of a constant quantity called a *photon* which is a quantum of available energy. It classically corresponds (modulo the Plank constant  $\hbar$ ) to the frequency of the field. Population transfer happens because the controlled quantum system absorbs one (or more) quantum of energy. However, the system can only absorb (or emit) quantities of energy corresponding to the difference between its allowed energy levels. Therefore, there must be a correspondence between the frequency of the controlling field and the energy gap between the two levels considered. In a semiclassical treatment, a control at a frequency equal to the energy difference between two levels only affects the population in these two levels and leaves untouched populations in other levels separated by very different energy gaps. This can also be seen as a justification of the multi-level approximation (cf. [Chapter 2](#)). In the rest of the section we illustrate and mathematically justify more rigorously such an idea in the semiclassical setting.

For simplicity of notation, we work in units where  $\hbar = 1$  so that we shall deal with controlled Schrödinger equation in the bilinear form<sup>1</sup>

$$\frac{d}{dt}\vec{\psi} = (-iH_0 - iH_1 u(t))\vec{\psi}. \quad (7.1)$$

We shall assume  $H_0$  diagonal,

$$H_0 := \text{diag}(E_1, E_2, \dots, E_n),$$

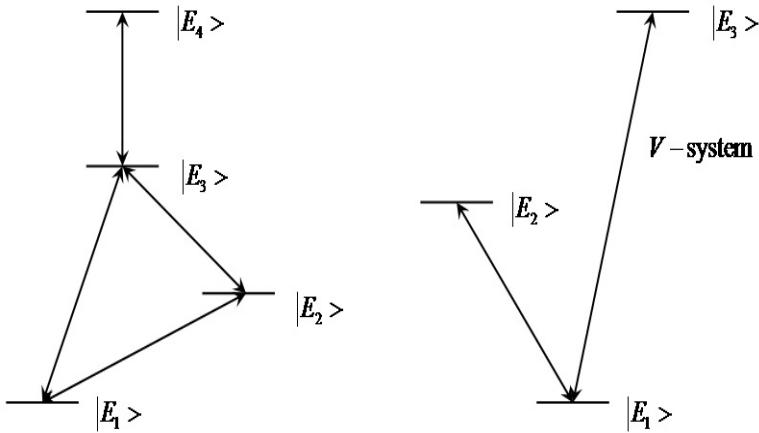
and assume we want to control within the subspace identified by the first two energy levels,  $E_1$  and  $E_2$ , so that, ideally, the other portion of the  $n$ -dimensional Hilbert space associated with the system is not affected by the dynamics. The matrix  $H_1$  is of the form

$$H_1 = \{h_{jk}\}, \quad h_{jk} = h_{jk}^*, \quad h_{jj} = 0, \quad j, k = 1, 2, \dots, n.$$

The constants  $h_{jk}$  represent the coupling between the  $j$ -th and the  $k$ -th energy level (eigenstate). The dynamics of quantum control systems may be represented by diagrams of the type in [Figure 7.1](#) where every horizontal segment represents an energy level, and levels  $j$  and  $k$  are connected if  $h_{jk}$  is different from zero. One example of a possible configuration is a  $V$ -system which is a three level system where the lowest energy level is coupled to the remaining two levels but these are not coupled to each other. Another typical three level configuration is a  $\Lambda$  system, where the coupling is only between the highest energy level and the remaining two levels (cf. [Figure 7.2](#)).

---

<sup>1</sup>Only notational changes are needed to handle the case of more than one control.



**FIGURE 7.1:** Level representation of quantum systems for control. The configuration on the right represents a *V* system.

We shall assume a control field  $u = u(t)$  with the functional dependence

$$u(t) = v(t) \sin(\omega_L t), \quad \omega_L \geq 0, \quad (7.2)$$

where  $v = v(t)$  is a smooth *slow varying envelope* which we shall refer to as the **pulse**, and which has the property

$$\lim_{t \rightarrow 0} v(t) = \lim_{t \rightarrow +\infty} v(t) = 0. \quad (7.3)$$

The magnitudes of the various components of  $\psi$  are called the **populations** in the various levels of energy. Going to the interaction picture by defining (cf. subsection 1.3.2)  $\vec{\phi} = e^{iH_0 t} \vec{\phi}$  which does not modify the populations, we obtain the differential equation for  $\vec{\phi}$

$$\frac{d}{dt} \vec{\phi} = -i \tilde{H}_1(t) \vec{\phi}, \quad (7.4)$$

where  $\tilde{H}_1$  is defined by

$$(\tilde{H}_1)_{jk} := h_{jk} u(t) e^{i\Delta_{jk} t}, \quad j < k,$$

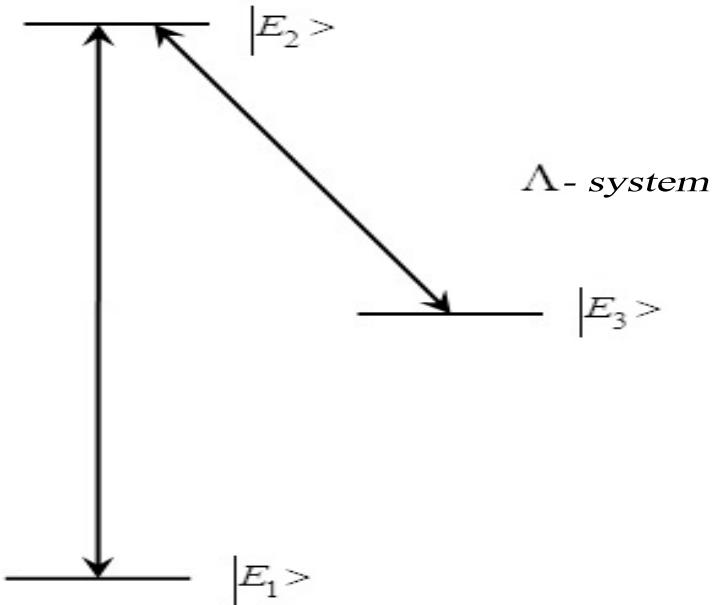
where the difference in energy levels  $j$  and  $k$ ,  $\Delta_{jk}$ , is defined by  $\Delta_{jk} := E_j - E_k$ .

Using expression (7.2) for the control  $u$  and separating the terms containing a *difference* of frequencies from those containing a *sum*, we can write  $\tilde{H}_1$  as

$$\tilde{H}_1(t) = \tilde{D}_1(t) + \tilde{A}_1(t). \quad (7.5)$$

Here  $\tilde{D}_1$  and  $\tilde{A}_1$  are both Hermitian and, for  $j < k$ ,

$$(\tilde{D}_1)_{jk} := \pm v(t) \frac{h_{jk}}{2i} e^{\pm i(\omega_L - |\Delta_{jk}|)t}, \quad (\tilde{A}_1)_{jk} := \mp v(t) \frac{h_{jk}}{2i} e^{\mp i(\omega_L + |\Delta_{jk}|)t},$$



**FIGURE 7.2:** Level representation of a  $\Lambda$  system.

where the + (−) sign for  $\tilde{D}_1$  and the − (+) sign for  $\tilde{A}_1$  is used if  $\Delta_{jk} \leq 0$  ( $\Delta_{jk} > 0$ ).

Our goal is now to show that terms of the type  $v(t)\mu e^{i\omega t}$ , for a complex constant  $\mu$  and a real frequency  $\omega$ , when  $\omega$  is large, and slow pulse  $v$ , do not significantly affect the solution of the differential equation (7.4). The intuitive idea is that fast oscillating terms average out to zero when integrated. Consider  $\tilde{H}_1$  in (7.4), and write it as  $\tilde{H}_1(t) := v(t)S(t) + v(t)F(t)$ , where  $F(t)$  is zero everywhere except in entries  $(j, k)$  and  $(k, j)$ , for one selected pair  $(j, k)$ , which are of the form  $\mu e^{i\omega t}$  and  $\mu^* e^{-i\omega t}$ , respectively. We assume  $j = 1$ ,  $k = 2$ , to have a simpler notation. We rewrite (7.4) as

$$\vec{\phi}(t) = \vec{\phi}(0) + \int_0^t -iv(s)S(s)\vec{\phi}(s) ds + \int_0^t -iv(s)F(s)\vec{\phi}(s) ds. \quad (7.6)$$

Because of the fast dependence of  $F$  on  $s$  the last integral gives a small contribution. Consider the diagonal matrix

$$G_\omega := \text{diag}(i\omega, -i\omega, 1, \dots, 1).$$

We can write the last integral in (7.6), using integration by parts, as

$$\begin{aligned} \int_0^t -iv(s)F(s)\vec{\phi}(s) ds &= G_\omega^{-1} \int_0^t -iv(s)\frac{dF(s)}{ds}\vec{\phi}(s) ds = \\ &-iG_\omega^{-1} \left( \left[ F(s)v(s)\vec{\phi}(s) \right]_0^t - \int_0^t F(s)\frac{d}{ds}(v(s)\vec{\phi}(s)) ds \right). \end{aligned}$$

Because of the shape of the function  $v$ , if  $t$  is sufficiently large,

$$\left[ F(s)v(s)\vec{\phi}(s) \right]_0^t \approx 0,$$

so that we can write (7.6) as

$$\begin{aligned} \vec{\phi}(t) &= \vec{\phi}(0) - i \int_0^t v(s)S(s)\vec{\phi}(s) ds \quad (7.7) \\ &+ iG_\omega^{-1} \left( \int_0^t \dot{v}(s)F(s)\vec{\phi}(s) ds - i \int_0^t v^2(s)(F(s)S(s) + F^2(s))\vec{\phi}(s) ds \right). \end{aligned}$$

Notice that  $F^2(s)$  is a constant matrix, independent of  $s$ , and that only the first two components of the vector multiplying  $G_\omega^{-1}$  in (7.7) are possibly different from zero. Assuming that  $\dot{v}$  is small, this shows that, as  $\omega \rightarrow \infty$ , these terms go to zero. By comparing the solution of (7.4) with the one of the ideal case

$$\frac{d}{dt}\vec{\phi}_d = -iS(t)v(t)\vec{\phi}_d, \quad \vec{\phi}_d(0) = \vec{\phi}(0),$$

we obtain

$$\|\vec{\phi}(t) - \vec{\phi}_d(t)\| \leq \int_0^t \|v(s)S(s)\| \|\vec{\phi}(s) - \vec{\phi}_d(s)\| ds + \frac{1}{\omega}m.$$

We have denoted by  $m$  a uniform bound for  $\|\int_0^t \dot{v}F\vec{\phi} ds - i \int_0^t Fv^2(S+F)\vec{\phi} ds\|$ , which we assume exists.<sup>2</sup> We can obtain an explicit bound by using the *Gronwall-Bellman inequality* (see (6.53) of Chapter 6). We have

$$\|\vec{\phi}(t) - \vec{\phi}_d(t)\| \leq \frac{m}{\omega} + \int_0^t \frac{m}{\omega} \|v(s)S(s)\| e^{\int_s^t \|v(r)S(r)\| dr} ds,$$

which goes to zero as  $\omega$  goes to infinity.

We may carry out this procedure for other frequencies in  $S(t)$  and obtain various estimates for the error resulting from neglecting the high frequency elements, in the Schrödinger equation. In particular, by considering the decomposition of  $\tilde{H}_1$  (7.5), we can neglect all the elements in  $\tilde{A}_1$ , as these contain sums of the energy gaps and the laser frequency  $\omega_L$  which give a fast

---

<sup>2</sup>Notice that all the quantities inside the integral can be bounded independently of  $\omega$ . We assume that  $v(s)$  and  $\dot{v}(s)$  go to zero sufficiently fast. This is done only to simplify notations. Otherwise we would write  $m = m(t)$ .

varying contribution. This is the **rotating wave approximation**. Moreover, elements corresponding to energy gaps which are very different from the frequency of the laser  $\omega_L$  can also be neglected. In an ideal situation where  $H_0$  has eigenvalues very much spaced from each other, a laser frequency  $\omega_L$  close to the energy gap between two particular levels is such that the system can be effectively considered a two level system. This idea is used to selectively induce population transfer between two given levels. The method used in this section, which is based on the Gronwall-Bellman inequality, can be used to estimate the error.

---

## 7.2 Time Dependent Perturbation Theory

We shall discuss **time dependent perturbation theory** for the evolution operator. Let us write the Schrödinger equation in the interaction picture (7.4) for the evolution operator  $X = X(t)$ .

$$\dot{X} = -i\tilde{H}_1(t)X(t), \quad X(0) = \mathbf{1}. \quad (7.8)$$

Written as an integral equation, this becomes

$$X(t) = \mathbf{1} - i \int_0^t \tilde{H}_1(t_1)X(t_1)dt_1. \quad (7.9)$$

By using this expression for  $X(t_1)$  and plugging it into (7.9) again, we obtain

$$X(t) = \mathbf{1} - i \int_0^t \tilde{H}_1(t_1)dt_1 + (-i)^2 \int_0^t \int_0^{t_1} \tilde{H}_1(t_1)\tilde{H}_1(t_2)X(t_2)dt_2 dt_1.$$

Continuing this way, we obtain the *Dyson series* expansion of the solution of the Schrödinger operator equation in the interaction picture (7.8)

$$X(t) = \mathbf{1} + \sum_{n=1}^{\infty} D_n(t), \quad (7.10)$$

with

$$D_n(t) := (-i)^n \int_0^t \int_0^{t_1} \cdots \int_0^{t_{n-1}} \tilde{H}_1(t_1)\tilde{H}_1(t_2) \cdots \tilde{H}_1(t_n)dt_n \cdots dt_1.$$

For finite  $t$  and finite dimensional (bounded)  $\tilde{H}_1$  this series converges and the error truncating the series (7.10) after  $m$  terms is given by

$$\left| X(t) - (\mathbf{1} + \sum_{n=1}^m D_n(t)) \right| =$$

$$\left| \int_0^t \int_0^{t_1} \cdots \int_0^{t_m} \tilde{H}_1(t_1) \tilde{H}_1(t_2) \cdots \tilde{H}_1(t_{m+1}) X(t_{m+1}) dt_{m+1} dt_m \cdots dt_1 \right| \\ \leq \frac{t^{m+1}}{(m+1)!} \|\tilde{H}_1\|_{[0,t]}^{m+1},$$

where we have denoted by  $\|\tilde{H}_1\|_{[0,t]}$  the infinity norm of  $\tilde{H}_1$  in the interval  $[0, t]$ .

In time varying perturbation theory, one replaces the solution of Schrödinger operator equation with the first  $m$  terms of the Dyson expansion, for some  $m$ . This is called *m-th order perturbation theory*. It gives a good approximation when the norm of  $\tilde{H}_1$  is small, e.g., for a controlled Schrödinger equation (7.1), when the control pulse has small amplitude. The most common situation is first order perturbation theory where one considers

$$X(t) \approx \mathbf{1} - i \int_0^t \tilde{H}_1(t_1) dt_1. \quad (7.11)$$

This first order expansion is used to analyze dynamics in several contexts (see, e.g., [185]). First order perturbation theory and formula (7.11) can also be used to give another (elementary) proof that the high frequency terms can be neglected in the dynamics.<sup>3</sup> For  $t \rightarrow \infty$  a treatment based on residue calculus is given in [192] section 2.2.

There has been a lot of study on the solutions of linear time varying equations such as (7.8). An alternative method of approximating the solution is the *Magnus expansion formula* [141] where one expresses the solution  $X$  as

$$X(t) = e^{\Omega(t)}, \quad (7.12)$$

where  $\Omega$  is given by a series expansion

$$\Omega(t) = \sum_{k=0}^{\infty} \Omega_k(t), \quad (7.13)$$

where  $\Omega_k(t)$  are multiple integrals of nested commutators of  $-i\tilde{H}_1$ . For example,

$$\Omega_1 = \int_0^t -i\tilde{H}_1(t_1) dt_1, \quad \Omega_2(t) = (-i)^2 \int_0^t \int_0^{t_1} [\tilde{H}_1(t_1), \tilde{H}_1(t_2)] dt_2 dt_1.$$

<sup>3</sup>If  $A_1 = A_1(t)$  is a high frequency term in  $\tilde{H}_1$ , i.e., having a  $2 \times 2$  block of the form  $\begin{pmatrix} 0 & \mu e^{i\omega t} \\ \mu^* e^{-i\omega t} & 0 \end{pmatrix}$ , then integration of  $A_1$  gives

$$\int_0^t A_1(t_1) dt_1 = F_{\omega}^{-1}(A(t) - \mathbf{1}),$$

with  $F_{\omega}^{-1}$  going to zero as  $\omega \rightarrow \infty$ .

The study of Magnus expansion (existence, computation and properties of the Magnus series (7.13) ) is more complicated than of Dyson series (7.10). However, the main advantage is that, as all the elements  $\Omega_k$  (7.13) are in the Lie algebra  $u(n)$ , so are their sums. Truncation at any order of the series (7.13) gives a matrix which is in  $u(n)$  and the resulting approximation of  $X(t)$ , according to (7.12) is unitary, a property which is not shared by the Dyson expansion. For studies comparing the two expansions, we refer to [157] and the references therein.

---

### 7.3 Adiabatic Control

Consider a finite dimensional, quantum control system in the general (not necessarily bilinear) form

$$\frac{d}{dt}\vec{\psi} = -iH(u(t))\vec{\psi}, \quad \vec{\psi}(0) = \vec{\psi}_0. \quad (7.14)$$

At every time  $t$  the Hamiltonian can be diagonalized via a unitary matrix  $T = T(t)$ , i.e.,

$$T(t)H(u(t))T^\dagger(t) = \Lambda(t), \quad (7.15)$$

with  $\Lambda(t)$  diagonal and assume that  $T(t)$  is at least of class  $C^1$  in the interval of control. Define the vector

$$\vec{\phi}(t) := T(t)\vec{\psi}(t).$$

Then the differential equation for  $\vec{\phi}$  reads, using (7.15), as

$$\frac{d}{dt}\vec{\phi} = \frac{d}{dt}(T\vec{\psi}) = (\dot{T}T^\dagger - i\Lambda)\vec{\phi}, \quad \vec{\phi}(0) = T(0)\vec{\psi}_0. \quad (7.16)$$

Notice that the norm of  $\dot{T}T^\dagger$  depends only on  $\dot{T}$ , as  $T^\dagger$  is unitary. If  $\dot{T}$  is small, which means that the control  $u = u(t)$  is *slow*, we may neglect the term  $\dot{T}T^\dagger$  in equation (7.16) and write the solution of (7.16) as

$$\vec{\phi}(t) = e^{-i\int_0^t \Lambda(r)dr}T(0)\vec{\psi}_0,$$

and the solution of (7.14) as

$$\vec{\psi}_a(t) = T^\dagger(t)e^{-i\int_0^t \Lambda(r)dr}T(0)\vec{\psi}_0. \quad (7.17)$$

This is the **adiabatic approximation**, whose validity depends on the fact that the control, and therefore the Hamiltonian, varies slowly with time. In particular, assume  $\vec{\psi}_0$  is an eigenstate of  $H(u(0))$  corresponding to a nondegenerate eigenvalue  $\lambda_0$ , and assume that the time evolution of  $H(u(t))$  is such

that the eigenvalue function  $\tilde{\lambda} = \tilde{\lambda}(t)$ , with  $\tilde{\lambda}(0) = \lambda_0$  remains a nondegenerate eigenvalue of  $H(u(t))$ , for all times  $t$ . Then the corresponding eigenvector  $\vec{\phi} = \vec{\phi}(t)$  can be taken as a column of  $T^\dagger(t)$  and the *adiabatic solution*  $\vec{\psi}_a$  in (7.17) is equal to  $\vec{\phi}$  modulo a phase factor, i.e.,

$$\vec{\psi}_a(t) = e^{-i \int_0^t \tilde{\lambda}(r) dr} \vec{\phi}(t).$$

We shall assume, from now on, to be in the above mentioned scenario, namely that the initial state is an eigenvector corresponding to a nondegenerate eigenvalue and that this situation is preserved during the time evolution.

The relevant question is how much the actual solution  $\vec{\psi}(t)$  differs from the adiabatic solution  $\vec{\psi}_a(t)$ , and therefore to obtain *estimates* of the norm of the difference  $\vec{\psi}(t) - \vec{\psi}_a(t)$  in terms of the norm of  $\dot{T}$ . This topic has been the subject of many papers which have proved such estimates in the so called *adiabatic theorems* (see, e.g., [18], [93], [109], [111], [115], [153], [173]).<sup>4</sup> Qualitative common wisdom in quantum mechanics states that the adiabatic approximation is good when the *eigenvalue gap*,  $\gamma = \gamma(t)$ , is large. The eigenvalue gap is defined as the minimum distance between the eigenvalue  $\tilde{\lambda}$  and the other eigenvalues of  $H = H(u(t))$ , i.e., denoting by  $\lambda_j = \lambda_j(t)$ , the other eigenvalues<sup>5</sup>,

$$\gamma(t) := \min_j |\lambda_j(t) - \tilde{\lambda}(t)|. \quad (7.18)$$

To make precise this statement, we shall illustrate a theorem of [109] which gives simple estimates. The theorem is more general but we shall restrict ourselves to the finite dimensional, matrix case (see also [173]), which is the case of interest to us.

To study the norm of the difference  $\|\vec{\psi} - e^{ir}\vec{\psi}_a\|$  for every real  $r$ , and in particular the minimum over real  $r$ , we equivalently study the norm  $\|\vec{\psi} - \vec{\psi}_a \vec{\psi}_a^\dagger \vec{\psi}\|$ , i.e., the norm<sup>6</sup> of the difference between the true solution  $\vec{\psi}$  and its orthogonal projection onto the one dimensional subspace spanned by  $\vec{\psi}_a$  (cf. Exercise 7.6). Defining  $P_a(t) := \vec{\psi}_a(t) \vec{\psi}_a^\dagger(t)$ , we look for estimates of the quantity  $\|(\mathbf{1} - P_a(t))\vec{\psi}(t)\|$ . The following adiabatic Theorem (cf. Theorem 3 in [109]) gives an easily computable bound for this error.<sup>7</sup>

<sup>4</sup>The subject of adiabatic theorems is in fact much more general in particular dealing with linear operators in infinite dimensional spaces.

<sup>5</sup>Here the minimum is taken over the  $j$  for every  $t$ .

<sup>6</sup>Recall we are using the Euclidean norm  $\|\vec{\psi}\| = \sqrt{\vec{\psi}^\dagger \vec{\psi}}$ .

<sup>7</sup>The operator norm used in this theorem and in the discussion that follows is the induced norm, i.e.,  $\|A\| = \max_{\|\vec{\psi}\|=1} \|A\vec{\psi}\|$ , which depends on the norm chosen for the vectors.

In the case of the Euclidean vector norm  $\|\vec{\psi}\| := \sqrt{\vec{\psi}^\dagger \vec{\psi}}$ , this is equal to the maximum singular value of  $A$ , namely the largest nonnegative square root of the eigenvalues of  $A^\dagger A$ .

**Theorem 7.3.1** Assume the Hamiltonian  $H = H(u(t))$  is of class at least  $C^2$  in the interval  $[0, T]$ . Then, for every  $t \in [0, T]$ , we have

$$\|(\mathbf{1} - P_a(t))\vec{\psi}(t)\| \leq \frac{\|H'(0)\|}{\gamma^2(0)} + \frac{\|H'(t)\|}{\gamma^2(t)} + \int_0^t \frac{\|H''(r)\|}{\gamma^2(r)} + 7 \frac{\|H'(r)\|^2}{\gamma^3(r)} dr. \quad (7.19)$$

**Remark 7.3.2** Adiabatic theorems usually give estimates of the error which depend on three quantities: The interval of control  $[0, T]$ , which we would like to be large; the minimum eigenvalue gap  $\gamma$  as defined in (7.18); and the *normalized* derivatives of the Hamiltonian, which give a rate of change of the Hamiltonian as compared to the interval of control. More precisely, one defines a variable  $s := \frac{t}{T}$  and defines a function of  $s$ ,  $\hat{H} = \hat{H}(s)$  by  $H(t) = H(Ts) := \hat{H}(s)$ . It is clear that

$$\frac{d^l \hat{H}}{ds^l} \Big|_{s=\bar{s}} = T^l \frac{d^l H}{dt^l} \Big|_{t=T\bar{s}},$$

for every  $l = 0, 1, \dots$ . Using this in the bound given in [109], we have stated Theorem 7.3.1 eliminating the dependence on  $T$  in the formula given there. In many instances  $T$  defines a time scale of the Hamiltonian, and the analysis is done in that time scale, so that  $\hat{H}$  is considered instead of  $H$ . As originally written in [109], formula (7.19), defining for every (matrix) function  $f := f(t)$ ,  $\hat{f}(s) := f(Ts)$ ,  $s \in [0, 1]$ , reads as

$$\|(\mathbf{1} - \hat{P}_a(s))\hat{\vec{\psi}}(s)\| \leq \frac{1}{\hat{\gamma}^2(0)T} \|\hat{H}'_{s=0}\| + \frac{1}{\hat{\gamma}^2(s)T} \|\hat{H}'(s)\| \quad (7.20)$$

$$+ \frac{1}{T} \int_0^s \frac{1}{\hat{\gamma}^2(r)} \|\hat{H}''(r)\| + 7 \frac{\|H'(r)\|^2}{\hat{\gamma}^3(r)} dr.$$

The general wisdom, in the study of adiabatic evolutions, says the total evolution time  $T$  should be of the order of  $\frac{1}{\hat{\gamma}_{min}^2}$ , where  $\hat{\gamma}_{min}$  is  $\min_{0 \leq s \leq 1} \hat{\gamma}(s)$ . This does not follow immediately from the bound in (7.20). However, in several cases, additional knowledge of the function  $\hat{\gamma}(s)$  allows us to extract this behavior from the bound in (7.20) [109].

Clearly (7.20) implies the adiabatic theorem in the form

$$\lim_{T \rightarrow \infty} \|(\mathbf{1} - \hat{P}_a(1))\hat{\vec{\psi}}(1)\| = \lim_{T \rightarrow \infty} \min_{r \in \mathbf{R}} \|\hat{\vec{\psi}}(1) - e^{ir} \hat{\vec{\psi}}_a(1)\| = 0.$$

**Remark 7.3.3** In **adiabatic quantum computation** (see, e.g., [77], [78], [187], [212]), the solution of a computational problem is encoded in an eigenstate (typically the ground state) of a Hamiltonian  $H_1$ . As this eigenstate is difficult to prepare, one lets a quantum system evolve according to an *interpolating Hamiltonian*

$$(1 - u(s))H_0 + u(s)H_1, \quad s \in [0, 1] \quad (7.21)$$

where  $u(s)$  is a slowly varying control function with  $u(0) = 0$ ,  $u(1) = 1$ .  $H_0$  is a Hamiltonian whose ground state is known and easy to prepare. According to the adiabatic theorem, under adiabatic evolution, the state of a quantum system evolving according to the Hamiltonian (7.21), starting at the ground state of  $H_0$ , will be close to the (unknown) ground state of  $H_1$  at time  $s = 1$ . From this state it is then possible to extract the solution of the original computational problem. The paper [77] contains several examples of this idea applied to quantum computations of interest in applications.

---

## 7.4 STIRAP

A population transfer between two levels which are not directly coupled can be obtained with a technique called **STImulated Raman scattering involving Adiabatic Passage (STIRAP)**, which combines control via frequency tuned pulses (section 7.1) with adiabatic control (section 7.3). The method is best explained in the simplest case of a  $\Lambda$  type system, although extensions exist to different types of configurations (see, e.g., [23], [52], [174], [192]). Let a system be in the configuration depicted in Figure 7.2. There is a direct coupling between the eigenstates  $|E_1\rangle$  and  $|E_2\rangle$  corresponding to energy level  $E_1$  and  $E_2$ , respectively, and between the eigenstates  $|E_2\rangle$  and  $|E_3\rangle$  corresponding to the energy levels  $E_2$  and  $E_3$ , respectively. There is no direct coupling between the eigenstates  $|E_1\rangle$  and  $|E_3\rangle$  and the problem consists of transferring population from  $|E_1\rangle$  to  $|E_3\rangle$ .

Two laser pulses are used to couple, respectively, levels  $|E_1\rangle$  and  $|E_2\rangle$  and levels  $|E_2\rangle$  and  $|E_3\rangle$ . We denote them by

$$u_{12}(t) := 2v_{12}(t) \cos(\omega_{12}t), \quad u_{23}(t) := 2v_{23}(t) \cos(\omega_{23}t), \quad (7.22)$$

where  $v_{12}$  and  $v_{23}$  are slow varying pulse functions and  $\omega_{12}$  and  $\omega_{23}$  are the corresponding laser frequencies. The pulses  $v_{12}$  and  $v_{23}$  are called *pump pulse* and *Stokes pulse*, respectively. The Hamiltonian writes as

$$H = H_0 + H_1(t),$$

with

$$H_0 := \text{diag}(E_1, E_2, E_3), \quad H_1(t) = \begin{pmatrix} 0 & h_{12}u_{12}(t) & 0 \\ h_{12}^*u_{12}(t) & 0 & h_{23}u_{23}(t) \\ 0 & h_{23}^*u_{23}(t) & 0 \end{pmatrix},$$

where  $h_{12}$  and  $h_{23}$  are constants coupling the levels 1 and 2 and 2 and 3, respectively. The initial state is  $\vec{\psi}(0) = [1, 0, 0]^T$ , i.e., all the initial population is assumed to be in level  $|E_1\rangle$ . Let  $D$  be given by

$$D := \text{diag}(E_2 - \omega_{12}, E_2, E_2 - \omega_{23}).$$

The Schrödinger equation

$$\frac{d}{dt}\vec{\psi} = -i(H_0 + H_1(t))\vec{\psi},$$

is transformed, by using the transformation of coordinates  $\vec{\phi} = e^{iDt}\vec{\psi}$ , into

$$\frac{d}{dt}\vec{\phi} = -i\tilde{H}_1(t)\vec{\phi}, \quad (7.23)$$

where

$$\tilde{H}_1(t) := -i \begin{pmatrix} -(E_2 - E_1 - \omega_{12}) & h_{12}u_{12}e^{-i\omega_{12}t} & 0 \\ h_{12}^*u_{12}e^{i\omega_{12}t} & 0 & h_{23}u_{23}e^{i\omega_{23}t} \\ 0 & h_{23}^*u_{23}e^{-i\omega_{23}t} & -(E_2 - E_3 - \omega_{23}) \end{pmatrix}.$$

By using the expression of  $u_{12}$  and  $u_{23}$  in (7.22) along with a rotating wave approximation (cf. section 7.1) and defining

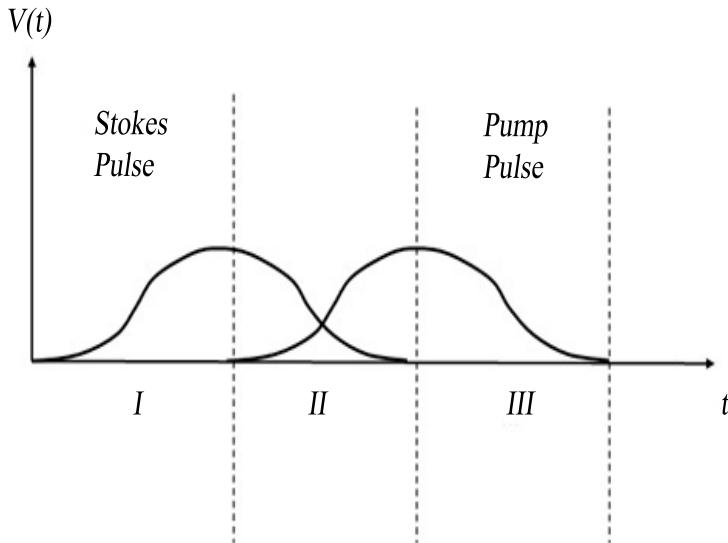
$$\Delta_{12} := E_2 - E_1 - \omega_{12}, \quad \Delta_{23} := E_2 - E_3 - \omega_{23},$$

we replace  $\tilde{H}_1$  by

$$\tilde{H}_2 := \begin{pmatrix} -\Delta_{12} & h_{12}v_{12}(t) & 0 \\ h_{12}^*v_{12}(t) & 0 & h_{23}v_{23} \\ 0 & h_{23}^*v_{23} & -\Delta_{23} \end{pmatrix}. \quad (7.24)$$

In order to transfer population from level  $|E_1\rangle$  to level  $|E_3\rangle$ , the *intuitive strategy* would be to use the pump pulse first at (nearly) the resonance frequency,  $E_2 - E_1$ , i.e., with  $\Delta_{12} = 0$ , so as to couple level  $|E_1\rangle$  and  $|E_2\rangle$  and transfer population from level  $|E_1\rangle$  to level  $|E_2\rangle$ . Then, one would use a Stokes pulse at (nearly) the resonance frequency  $E_2 - E_3$ , i.e., with  $\Delta_{23} = 0$ , to couple levels  $|E_2\rangle$  and  $|E_3\rangle$  and to transfer population to level  $|E_3\rangle$ . This intuitive strategy has the drawback of populating the intermediate level  $|E_2\rangle$ . The interaction with the external environment, i.e., with energy levels which are not included in the model, makes some of the population transfer from level  $|E_2\rangle$  to these levels. In probabilistic terms, this means that there is an increasing probability to find the system at an energy level lower than  $E_2$ . This phenomenon of *spontaneous decay* is particularly important as  $|E_2\rangle$  is a higher energy level and open systems tend to the lowest energy configuration. In this setting one could pose an interesting optimal control problem and try to shape the pulses  $v_{12}$  and  $v_{23}$  so as to minimize  $\int_{-\infty}^{+\infty} |\psi_2(t)|^2 dt$ , where  $|\psi_2|$  represents the population in level  $|E_2\rangle$ .

The STIRAP method uses a *counterintuitive pulse sequence* in which the Stokes pulse  $v_{23}$  precedes the pump pulse  $v_{12}$ . The two pulses overlap for some time in which the magnitude of the Stokes pulse slowly decreases while the magnitude of the pump pulse slowly increases. The typical situation is described in Figure 7.3. If the pulses are slow enough, the resulting evolution



**FIGURE 7.3:** Counterintuitive pulse sequence in the STIRAP control.

can be considered adiabatic and the resulting trajectory analyzed using the ideas in section 7.3.

Assume

$$\Delta_{12} = \Delta_{23} := \Delta.$$

The eigenvalues of  $\tilde{H}_2$  in (7.24) are

$$\lambda_1 = -\Delta,$$

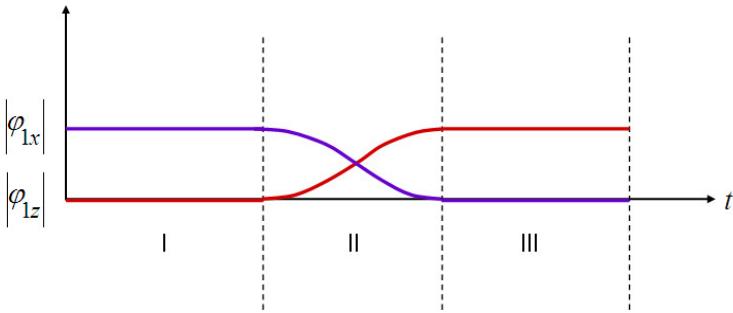
$$\lambda_2(t) = -\frac{\Delta}{2} + \frac{1}{2}\sqrt{\Delta^2 + 4(|h_{23}|^2v_{23}^2 + |h_{12}|^2v_{12}^2)},$$

$$\lambda_3(t) = -\frac{\Delta}{2} - \frac{1}{2}\sqrt{\Delta^2 + 4(|h_{23}|^2v_{23}^2 + |h_{12}|^2v_{12}^2)}.$$

Of particular interest to us is the normalized eigenvector corresponding to the eigenvalue  $\lambda_1$ , i.e.,

$$\vec{\phi}_1 := \frac{1}{\sqrt{|h_{12}v_{12}|^2 + |h_{23}v_{23}|^2}} \begin{pmatrix} -h_{23}v_{23} \\ 0 \\ h_{12}^*v_{12} \end{pmatrix}. \quad (7.25)$$

We now follow the trajectory of the state  $\vec{\phi}$  solution of (7.23), assuming that we start with an initial state  $\vec{\phi}(0) = [1, 0, 0]^T$ . During the interval  $I$  in Figure 7.3, such a state is not changed except for an irrelevant phase factor, since  $v_{12} \equiv 0$  in this interval. For the whole interval, the state coincides with



**FIGURE 7.4:** Time dependence of the  $x$  and  $z$  component ( $|\phi_{1x}|$  and  $|\phi_{1z}|$ , respectively) of the state vector  $\vec{\phi}_1$  in adiabatic approximation for the counterintuitive pulse sequence in the STIRAP control.

the eigenstate  $\vec{\phi}_1$  (7.25) corresponding to  $\lambda_1$ . In the interval  $II$ , assuming adiabatic evolution, the trajectory follows  $\vec{\phi}_1(t)$ . As the magnitude of  $v_{23}$  slowly decreases while the one of  $v_{12}$  slowly increases,  $\vec{\phi}_1$  and therefore  $\vec{\phi}$  moves from a state parallel to  $[1, 0, 0]^T$  to a state parallel to  $[0, 0, 1]^T$ . The magnitude of the pump pulse  $v_{23}$  is then slowly decreased to zero in the interval  $III$ , and this does not modify the final state. Notice that since  $\vec{\phi}_1$  does not have any component along the state  $|E_2\rangle$ , the process ideally does not place any population in  $|E_2\rangle$  so as to avoid the problems of the intuitive strategy. The behavior of the adiabatic vector  $\vec{\phi}_1$  during the interval of control  $I + II + III$  is summarized in Figure 7.4. The errors of the adiabatic approximation can be evaluated using one of the adiabatic theorems quoted in the previous section. In particular (cf. Theorem 7.3.1) the error is smaller if the gap (7.18) between the eigenvalues is large. In our case, the gap is given by

$$\gamma(t) := \min\{|\lambda_1 - \lambda_2|, |\lambda_1 - \lambda_3|\} = \frac{1}{2} \left| \sqrt{\Delta^2 + 4(|h_{12}|^2 v_{12}^2 + |h_{23}|^2 v_{23}^2)} - |\Delta| \right|.$$

In the interval of interest, where the actual state transfer occurs, i.e., the interval  $II$ , the quantity measuring the strength of the coupling with the laser field, namely  $|h_{12}|^2 v_{12}^2 + |h_{23}|^2 v_{23}^2$ , is always larger than a positive lower bound. If the strength of the coupling with the laser field is sufficiently large,  $\gamma$  is also large, which makes the adiabatic scheme valid. A *resonant* STIRAP algorithm is such that  $\Delta = 0$ . However, resonance is not necessary for the validity of the scheme.

## 7.5 Lyapunov Control of Quantum Systems

In Lyapunov control of quantum systems,<sup>8</sup> one specifies a function of the state and designs the control so that the value of the function decreases to a desired value. Such a function is called a **Lyapunov function**. The Lyapunov method of control is a powerful tool in nonlinear control systems design [117].

### 7.5.1 Quantum control problems in terms of a Lyapunov function

The main form of a Lyapunov function we shall consider, written in terms of the density matrix  $\rho$ , is

$$V(\rho) := \text{Tr}(P\rho), \quad (7.26)$$

where  $P$  is a Hermitian matrix, which, without loss of generality, we can assume positive semi-definite, so as to satisfy the standard requirement for a Lyapunov function,  $V(\rho) \geq 0$ .<sup>9</sup> This has the meaning of the expectation value of the observable  $P$  (cf. 1.37). Several quantum control problems can be formulated in terms of finding a control function  $u$  so that the solution of

$$\dot{\rho} = -i[H_0 + H_1 u, \rho], \quad \rho(0) = \rho_0, \quad (7.27)$$

converges to the minimum of  $V(\rho)$  in (7.26).<sup>10</sup> In the case where the system is operator controllable the state  $\rho$  which minimizes  $V(\rho)$  has to be sought among the states which are unitarily equivalent to  $\rho_0$ , or, equivalently, the ones that have the same spectrum as  $\rho_0$ , that is matrices in the orbit

$$\mathcal{O}_{u(n)}(\rho_0) := \{X\rho_0X^\dagger | X \in U(n)\}.$$

We shall assume operator controllability in the following.

**Example 7.5.1** Consider the problem of driving the pure state of (7.27) to a state collinear to  $\vec{\psi}_f := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ . The problem is equivalent to driving the state  $\rho$  to the minimum of the Lyapunov function (7.26) with  $P$  given by

<sup>8</sup>During the computations in this section, we often use the identity  $\text{Tr}(A[B, C]) = \text{Tr}(B[C, A])$ .

<sup>9</sup>We can always replace  $P$  with  $P + k\mathbf{1}$ , with  $k$  sufficiently large. This will just shift the value of  $V$  by a quantity  $k$ , having no effect on the following analysis.

<sup>10</sup>We shall treat the case of a general density matrix and then specialize to pure states in Remarks 7.5.4, 7.5.7, 7.5.8. We assume a bilinear quantum control system (7.27) with one control. The generalization to the multi-input case is left to the reader.

$P := \text{diag}(0, 1, 1)$ . The minimum is obtained for  $\text{Tr}(P\rho) = 0$ . This condition is equivalent to  $\rho = \vec{\psi}_f \vec{\psi}_f^\dagger$  and therefore the problem can be posed as driving the state  $\rho$  to the minimum of  $\rho$ .

Both the minimum and maximum of the function  $V(\rho)$  correspond to matrices that commute with  $P$ . In fact, we have the following result.

**Proposition 7.5.2** For any stationary point  $\rho$  of  $V(\rho)$  (which includes maximum and minimum value)

$$[\rho, P] = 0.$$

Conversely, if  $\rho$  is such that  $[\rho, P] = 0$ , then  $\rho$  is a stationary point of  $V(\rho)$ .

*Proof.* Assume  $\rho$  is a stationary point in the orbit

$$\mathcal{O}_{u(n)}(\rho_0) = \{X\rho X^* | X \in U(n)\}.$$

Let  $\{A_1, \dots, A_{n^2}\}$  be a basis of  $u(n)$ . A neighborhood of  $\rho$  is given by the set of matrices  $F = F(t_1, \dots, t_{n^2})$ , defined as

$$F(t_1, t_2, \dots, t_{n^2}) := e^{A_1 t_1} e^{A_2 t_2} \dots e^{A_{n^2} t_{n^2}} \rho e^{-A_{n^2} t_{n^2}} \dots e^{-A_2 t_2} e^{-A_1 t_1},$$

with  $t_1, t_2, \dots, t_{n^2}$  in an appropriate open neighborhood of the origin of  $\mathbf{R}^{n^2}$ . For a stationary point,

$$\left. \frac{d}{dt_j} \text{Tr}(PF(t_1, \dots, t_{n^2})) \right|_{t_1=t_2=\dots=t_{n^2}=0} = 0, \quad j = 1, \dots, n^2.$$

This gives

$$\text{Tr}(P[\rho, A_j]) = \text{Tr}(A_j[P, \rho]) = 0, \quad j = 1, \dots, n^2. \quad (7.28)$$

As  $[P, \rho]$  is a skew-Hermitian matrix whose component along any element of a basis of  $u(n)$  is zero,  $[P, \rho]$  is equal to zero, i.e.,  $\rho$  commutes with  $P$ . Conversely, if  $[P, \rho] = 0$ , then from (7.28) all the derivatives of  $\text{Tr}(PF(t_1, \dots, t_{n^2}))$  at  $t_1 = t_2 = \dots = t_{n^2} = 0$  are zero, and  $\rho$  is a stationary point.  $\square$

From the previous proposition, we have the following (see [Exercise 7.4](#)).

**Proposition 7.5.3** Assuming, without loss of generality, that  $P$  is diagonal, i.e., of the form

$$P := \text{diag}(p_1 \mathbf{1}_{n_1 \times n_1}, p_2 \mathbf{1}_{n_2 \times n_2}, \dots, p_r \mathbf{1}_{n_r \times n_r}),$$

---

<sup>11</sup>The Lyapunov control for a pure state will be discussed in the following. One way to see that the minimum is obtained for  $\rho = \vec{\psi}_f \vec{\psi}_f^\dagger$  is to write  $P$  as  $P = \mathbf{1} - \vec{\psi}_f \vec{\psi}_f^\dagger$  and  $\rho$  as  $\rho = (\alpha \vec{\psi}_f + \vec{\psi}_o)(\alpha^* \vec{\psi}_f^\dagger + \vec{\psi}_o^\dagger)$ , for  $\vec{\psi}_o$  a vector perpendicular to  $\vec{\psi}_f$ . We have  $\text{Tr}(P\rho) = \|\vec{\psi}_o\|^2$ , which is zero if and only if  $\vec{\psi}_o = 0$  and  $|\alpha| = 1$ . We could also apply Proposition 7.5.3 as in Remark 7.5.4.

with increasing  $p_1 < \dots < p_r$ , and  $n_1 + n_2 + \dots + n_r = n$ , then the *maximum* (*minimum*) has the block diagonal form  $\rho = \text{diag}(B_{n_1 \times n_1}, \dots, B_{n_r \times n_r})$ . The eigenvalues (possibly coinciding)  $\lambda_1, \lambda_2, \dots, \lambda_n$  of  $\rho$  are eigenvalues of the blocks  $B_{n_1 \times n_1}, \dots, B_{n_r \times n_r}$ , in *increasing* (*decreasing*) order.

**Remark 7.5.4** The Lyapunov functions considered in the literature are typically special cases of the general Lyapunov function in (7.26). In particular, for the case of a pure state  $\vec{\psi}$ ,

$$V(\vec{\psi}) = \vec{\psi}^\dagger P \vec{\psi} = \text{Tr}(P \vec{\psi} \vec{\psi}^\dagger) \quad (7.29)$$

is considered (cf. [91]). By taking

$$P := \mathbf{1} - \vec{\psi}_f \vec{\psi}_f^\dagger \quad (7.30)$$

for some desired final state  $\vec{\psi}_f$ , one obtains the *Hilbert Schmidt distance*

$$V(\vec{\psi}) = 1 - |\vec{\psi}^\dagger \vec{\psi}_f|^2, \quad (7.31)$$

which is used, for example, in [215]. In agreement with Proposition 7.5.3, the minimum of this function is obtained for  $\vec{\psi} = e^{ir}\vec{\psi}_f$  for any  $r \in \mathbb{R}$ , the maximum for any vector perpendicular to  $\vec{\psi}_f$ . The Lyapunov function  $V(\vec{\psi})$  in (7.29) has the advantage that it is independent of a phase factor, which does not have physical meaning. In other words,  $V(\vec{\psi}) = V(e^{ir}\vec{\psi})$  for any  $r \in \mathbb{R}$ . Another Lyapunov function, used for example in [146], is the square of the error, i.e.,

$$V_1(\vec{\psi}) = (\vec{\psi} - \vec{\psi}_f)^\dagger (\vec{\psi} - \vec{\psi}_f) = 2(1 - \text{Re}(\vec{\psi}_f^\dagger \vec{\psi})).$$

The objective of the control is to drive  $\vec{\psi}$  to the minimum of  $V_1(\vec{\psi})$  up to a phase factor. The relation

$$\min_{r \in \mathbb{R}} V_1(e^{ir}\vec{\psi}) = 2(1 - |\vec{\psi}_f^\dagger \vec{\psi}|),$$

along with (7.31), establishes the connection with the Hilbert-Schmidt Lyapunov function  $V(\vec{\psi})$ . The approach of [146] uses the Lyapunov function  $V_1(\vec{\psi})$  and treats the phase degree of freedom by adding a fictitious control  $\omega$  in the Schrödinger equation

$$\frac{d}{dt} \vec{\psi} = -i(H_0 + H_1 u) \vec{\psi}, \quad (7.32)$$

which becomes

$$\frac{d}{dt} \vec{\psi} = -i(H_0 + \omega + H_1 u) \vec{\psi}. \quad (7.33)$$

The equivalence between the treatment with the Lyapunov function  $V(\vec{\psi})$  with Schrödinger equation (7.32) and the one with Lyapunov function  $V_1(\vec{\psi})$

with Schrödinger equation (7.33) is explained in the following FACT, whose proof is left as an exercise (see [Exercise 7.5](#)).

**FACT** Assume  $u_d$  is such that the corresponding solution  $\vec{\psi}$  of (7.32) satisfies  $\lim_{t \rightarrow \infty} V(\vec{\psi}) = 0$ , then there exists a function  $\omega_d$  such that the solution  $\vec{\psi}$  of (7.33), with  $(u, \omega) = (u_d, \omega_d)$ , satisfies  $\lim_{t \rightarrow \infty} V_1(\vec{\psi}) = 0$ . Conversely, if, with an augmented control  $(u, \omega) := (u_d, \omega_d)$ , the solution (7.33) satisfies  $\lim_{t \rightarrow \infty} V_1(\vec{\psi}) = 0$ , then, with the control  $u = u_d$ , the solution  $\vec{\psi}$  of (7.32) satisfies  $\lim_{t \rightarrow \infty} V(\vec{\psi}) = 0$ .

### 7.5.2 Determination of the control function

As we wish to minimize the value of the Lyapunov function  $V(\rho)$  in (7.26), we choose the control  $u$  to make the time derivative of  $V(\rho)$ , along the corresponding trajectory, negative at all times. We calculate, using (7.27),

$$\frac{d}{dt}V(\rho) = -i \operatorname{Tr}(P[H_0 + H_1 u, \rho]) = -i \operatorname{Tr}(\rho[P, H_0]) - i \operatorname{Tr}(\rho[P, H_1])u.$$

Since  $\operatorname{Tr}([P, H_0]) = 0$  and therefore  $[P, H_0]$  is not sign defined we assume  $[P, H_0] = 0$ , and therefore, we assume that both  $P$  and  $H_0$  are diagonal. We state this formally.

**Assumption 1** The matrices  $P$  and  $H_0$  are both diagonal.

With this assumption, we have

$$\frac{d}{dt}V(\rho) = -i \operatorname{Tr}(\rho[P, H_1])u,$$

and the natural choice for the control is

$$u = iK \operatorname{Tr}(\rho[P, H_1]), \quad (7.34)$$

for some  $K > 0$ . With this, we obtain

$$\frac{d}{dt}V(\rho) = -K |\operatorname{Tr}(\rho[P, H_1])|^2 \leq 0.$$

Therefore, the value of  $V(\rho)$  will decrease towards a given value.

### 7.5.3 Study of the asymptotic behavior of the state $\rho$

With the above control law, the value of the state  $\rho$  will tend to a limit set which can be characterized using **La Salle's invariance principle** (see, e.g., [117], Theorem 4.4, p. 128).<sup>12</sup>

---

<sup>12</sup>We state the theorem in a weaker form which is sufficient for our purposes.

**Theorem 7.5.5 (La Salle invariance principle)** Consider a system of differential equations

$$\dot{x} = f(x), \quad (7.35)$$

with smooth  $f$ , and  $x \in \mathbf{R}^n$ . Given a smooth function  $V : \mathbf{R}^n \rightarrow \mathbf{R}$  such that  $\frac{d}{dt}V = (\frac{d}{dt}V)(x) = (\nabla V \cdot f)(x) \leq 0$ , for every  $x \in \mathbf{R}^n$ , let  $M$  be the largest invariant set in  $\mathbf{R}^n$  where  $(\frac{d}{dt}V)(x) = 0$ . Then, every solution of (7.35) converges to  $M$  as  $t \rightarrow \infty$ .

The set  $M$  for the Lyapunov function  $V$  in (7.26), with dynamics (7.27) and control (7.34), is the set of all the Hermitian matrices  $\rho_1$  with the same spectrum as  $\rho_0$  which satisfy

$$\text{Tr}(e^{-iH_0t}\rho_1e^{iH_0t}[P, H_1]) = 0, \quad \forall t \in \mathbf{R}, \quad (7.36)$$

i.e.,

$$M := \{\rho_1 \in \mathcal{O}_{u(n)}(\rho_0) \mid \text{Tr}(e^{-iH_0t}\rho_1e^{iH_0t}[P, H_1]) = 0, \quad \forall t \in \mathbf{R}\}.$$

This set is clearly invariant under the differential equation (7.27) with control (7.34). That is, if we start with an initial condition in  $M$ , the corresponding trajectory remains in  $M$ . Moreover  $M$  is included in the set where  $(\frac{d}{dt}V)(\rho) = 0$ . Furthermore, this is the *largest* invariant set having the property of being included in  $\{\rho \mid (\frac{d}{dt}V)(\rho) = 0\}$ . In fact, if  $\tilde{\rho}(t)$  is such that  $(\frac{d}{dt}V)(\tilde{\rho}(t)) = 0$ , then  $\tilde{\rho}(t) = e^{-iH_0t}\tilde{\rho}(0)e^{iH_0t}$ , and therefore, from (7.36),  $\tilde{\rho}(0) \in M$ .

We now want to characterize the set  $M$  of density matrices  $\rho_1$  satisfying (7.36), which we rewrite using Assumption 1,

$$\text{Tr}(e^{iH_0t}H_1e^{-iH_0t}[\rho_1, P]) = 0, \quad \forall t \in \mathbf{R}.$$

Taking the  $n$ -th derivative with respect to  $t$  at time  $t = 0$  and recalling the *ad* notation (3.38), this is equivalent to

$$\text{Tr}((ad_{iH_0}^n H_1)[\rho_1, P]) = 0, \quad n = 0, 1, 2, \dots \quad (7.37)$$

By writing  $H_0$  as

$$H_0 := \text{diag}(\lambda_1, \dots, \lambda_n),$$

and

$$H_1 := \{h_{jk}\}, \quad h_{jk} = h_{kj}^*, \quad j, k = 1, \dots, n,$$

we have

$$(ad_{iH_0}^n H_1)_{lm} = (ad_{iH_0}^n H_1)_{ml}^* = (i)^n \omega_{lm}^n h_{lm}, \quad l < m,$$

where  $\omega_{lm} := \lambda_l - \lambda_m$ . Let us denote the  $(j, k)$ -th element of  $[\rho_1, P]$  by  $[\rho_1, P]_{jk} := s_{jk} = -s_{kj}^*$ . For a fixed  $n$ , in the condition (7.37),  $s_{jk}$ , with  $j < k$ , appears in the trace only two times:  $-s_{jk}^*$  multiplies  $(i)^n \omega_{jk}^n h_{jk}$  and

$s_{jk}$  multiplies  $(-i)^n \omega_{jk}^n h_{jk}^*$ . Therefore, for a fixed  $n$ , condition (7.37) writes as

$$\sum_{j < k} s_{jk} (-i)^n \omega_{jk}^n h_{jk}^* - s_{jk}^* (i)^n \omega_{jk}^n h_{jk} = 0.$$

For  $n$  even, this gives

$$\sum_{j < k} \omega_{jk}^n \operatorname{Im}(s_{jk} h_{jk}^*) = 0, \quad n = 0, 2, 4, \dots \quad (7.38)$$

For  $n$  odd, this gives

$$\sum_{j < k} \omega_{jk}^n \operatorname{Re}(s_{jk} h_{jk}^*) = 0, \quad n = 1, 3, 5, \dots \quad (7.39)$$

These are the general conditions that have to be satisfied by  $[\rho_1, P]$  and, therefore by  $\rho_1$ , according to the La Salle invariance principle (7.5.5). As we want  $\rho_1$  to be a minimum for  $V(\rho)$ , we would like, according to Proposition 7.5.3,  $[\rho_1, P] = 0$ , i.e.,  $s_{jk} = 0$  for all  $j < k$ .<sup>13</sup> This is implied by (7.38) and (7.39) under the following two additional conditions.

**Assumption 2** The transition frequencies  $\omega_{jk}$  of the natural Hamiltonian  $H_0$  are such that

$$\omega_{jk} \neq \omega_{lm}, \quad j, k \neq l, m.$$

**Assumption 3** For every pair  $j, k$ ,  $j < k$ ,  $h_{jk} \neq 0$ .

Assumption 2 says that the transition energies between two different levels are clearly identified. In principle, it is possible to use tailored radio-frequency pulses to selectively transfer population between two levels (cf. section 7.1). Assumption 3 says that all the levels are directly coupled. Under Assumption 2, a Vandermonde determinant argument on the equations (7.38) and (7.39) shows that both  $\operatorname{Im}(s_{jk} h_{jk}^*)$  and  $\operatorname{Re}(s_{jk} h_{jk}^*)$  are equal to zero. Therefore, since  $s_{jk} h_{jk}^* = 0$ , for all  $j < k$ , Assumption 3 implies that  $s_{jk} = 0$ . Therefore, we can conclude with the following result

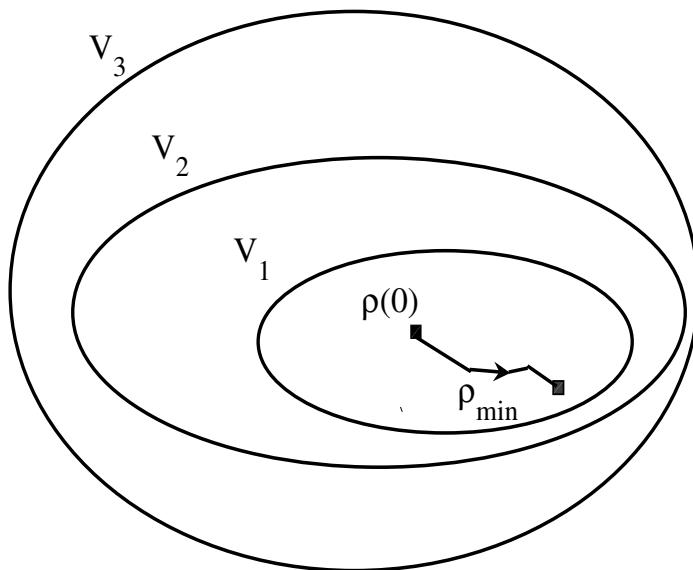
**Theorem 7.5.6** *Under Assumptions 1, 2 and 3, the control (7.34) asymptotically drives the state  $\rho$  of (7.27) to a stationary point of  $V(\rho) = \operatorname{Tr}(P\rho)$ .*

The nature of the stationary limit point in the above theorem has to be studied for the particular situation at hand. For example, if  $P$  has  $n$  distinct eigenvalues, all of them nondegenerate, then there exists a finite number of matrices with a given spectrum (the spectrum of the initial condition  $\rho_0$ ) which commute with  $P$ . Call them  $\tilde{\rho}_1, \tilde{\rho}_2, \dots, \tilde{\rho}_r, \rho_{\min}$ , for some integer  $r$

---

<sup>13</sup>Notice that, as  $P$  is diagonal, all the diagonal elements of  $[\rho_1, P] = 0$  are automatically zero.

and  $\rho_{min}$  corresponds to the minimum of  $V(\rho)$ . Since  $V(\rho)$  is decreasing, if  $V(\rho(0)) < V(\tilde{\rho}_j)$ ,  $j = 1, 2, \dots, r$ , then the solution will necessarily converge to  $\rho_{min}$ . In other words, the condition  $V(\rho(0)) < V(\tilde{\rho}_j)$ ,  $j = 1, 2, \dots, r$ , determines the *region of attraction* for the given design (cf. Figure 7.5). This observation can be used to shape  $P$  in problems where the goal is to drive the state to a determined value.



**FIGURE 7.5:** The matrices commuting with  $P$ ,  $\tilde{\rho}_1, \tilde{\rho}_2, \dots, \tilde{\rho}_r$ ,  $\rho_{min}$  determine level lines in the space of density matrices for the Lyapunov function  $V = \text{Tr}(P\rho)$ . The level lines corresponding to  $\tilde{\rho}_1, \tilde{\rho}_2, \dots, \tilde{\rho}_r$  are labeled in this example by  $V_1, V_2$  and  $V_3$ , with  $V_1 < V_2 < V_3$ . If  $\text{Tr}(P\rho(0)) < V_1$  then the trajectory will converge to  $\rho_{min}$ . The innermost curve is the boundary of the region of attraction.

The case of a rank one density matrix  $\rho$ , namely of a pure state, is treated in the following remark.

**Remark 7.5.7** In the special case where  $\rho$  is a pure state, i.e., of the form  $\rho = \vec{\psi}\vec{\psi}^\dagger$ , and the objective is to drive the state asymptotically to a given eigenstate  $\vec{\psi}_f$  of  $H_0$ , the Lyapunov function takes the form  $V(\vec{\psi})$  in (7.29) with  $P$  of the form in (7.30). Assumption 1 does not give any factual constraint

since  $P$  and  $H_0$  commute if  $\vec{\psi}_f$  is an eigenstate of  $H_0$ . Under Assumptions 2 and 3, the state will converge to a value  $\vec{\psi}_1$  such that  $\vec{\psi}_1 \vec{\psi}_1^\dagger$  commutes with  $\vec{\psi}_f \vec{\psi}_f^\dagger$ . Writing this as

$$(\vec{\psi}_f^\dagger \vec{\psi}_1) \vec{\psi}_f \vec{\psi}_1^\dagger = (\vec{\psi}_1^\dagger \vec{\psi}_f) \vec{\psi}_1 \vec{\psi}_f^\dagger,$$

and multiplying on the right by  $\vec{\psi}_f$ , we obtain that the limit state  $\vec{\psi}_1$  is collinear with  $\vec{\psi}_f$  unless  $\vec{\psi}_f^\dagger \vec{\psi}_1 = 0$ , which is the situation corresponding to the maximum for  $V(\psi)$ . If this is not the situation at the beginning it never is since  $V$  is always not increasing. Therefore, for almost all initial conditions the state will converge to the desired state parallel to  $\vec{\psi}_f$ .

In general, if we do not assume Assumptions 2 and 3, all we can say, applying La Salle's invariance principle of Theorem 7.5.5, is that the state  $\rho$  will converge to a set of values  $\rho_1$  such that  $[\rho_1, P]$  is perpendicular to the space  $\mathcal{V}$  spanned by  $ad_{iH_0}^n H_1$ ,  $n = 0, 1, \dots$ . In the pure state case, however, Assumptions 2 and 3 can be weakened to a condition equivalent to local controllability of the Schrödinger equation about the final target eigenstate. This is explained in the following remark.

**Remark 7.5.8** Consider system (7.33) and a control  $\omega \equiv -\lambda$ ,  $u \equiv 0$ , in the state  $\vec{\phi}$ . If  $\vec{\phi}$  is an eigenvector of  $H_0$  corresponding to the eigenvalue  $\lambda$ , then  $\vec{\phi}$  is an equilibrium point for system (7.33). To linearize this system about  $\vec{\phi}$ , set  $\vec{\psi} = \vec{\phi} + \Delta\vec{\psi}$ ,  $\omega := -\lambda + \Delta\omega$ ,  $u = \Delta u$ . From (7.33), we obtain

$$\frac{d}{dt}(\Delta\vec{\psi}) = (-iH_0 + i\lambda)\Delta\vec{\psi} - i\vec{\phi}\Delta\omega - iH_1\vec{\phi}\Delta u - i\Delta\omega\Delta\vec{\psi} - iH_1\Delta u\Delta\vec{\psi}.$$

By assuming  $\Delta\omega$ ,  $\Delta u$  and  $\Delta\vec{\psi}$  small, we neglect the last two terms. We obtain the linear system

$$\frac{d}{dt}(\Delta\vec{\psi}) = (-iH_0 + i\lambda)\Delta\vec{\psi} - i\vec{\phi}\Delta\omega - iH_1\vec{\phi}\Delta u.$$

According to the Kalman controllability condition (see, e.g., [114]) this system is controllable if and only if the rank of the controllability matrix<sup>14</sup>

$$C := \quad (7.40)$$

$$\left( \begin{array}{cccccc} (-iH_0 + i\lambda)^{n-1}(-iH_1\vec{\phi}) & \dots & (-iH_0 + i\lambda)(-iH_1\vec{\phi}) & (-iH_1\vec{\phi}) & -i\vec{\phi} \end{array} \right),$$

<sup>14</sup>Recall the controllability matrix for a linear system  $\dot{x} = Ax + Bu$  of dimension  $n$ , with  $A$  an  $n \times n$  matrix and  $B$  an  $n \times m$  matrix with  $m$  the number of controls, is given by  $C := [A^{n-1}B, A^{n-2}B, \dots, AB, B]$ . It is an  $n \times nm$  matrix. In our case  $B = [(-iH_1\vec{\phi}), -i\vec{\phi}]$  and  $A = (-iH_0 + i\lambda)$ . Since  $(-iH_0 + i\lambda)\vec{\phi} = 0$  all the columns of  $C$  in (7.40) corresponding to  $(-iH_0 + i\lambda)^k(-i\vec{\phi})$ , with  $k \geq 1$ , will be zero. We have omitted these columns in (7.40).

is equal to  $n$ . This condition is independent of the coordinates. So we can assume  $\vec{\phi} := (1, 0, \dots, 0)$  and  $H_0 := \text{diag}(0, \lambda_2 - \lambda, \dots, \lambda_n - \lambda)$ . The controllability matrix [114] takes the form

$$C = \begin{pmatrix} 0 & \cdot & \cdot & \cdot & 0 & -ih_{11} & -i \\ (-i)^{n-1}(\lambda_2 - \lambda)^{n-1}h_{12}^* & \cdot & \cdot & \cdot & -i(\lambda_2 - \lambda)h_{12}^* & 0 & 0 \\ (-i)^{n-1}(\lambda_3 - \lambda)^{n-1}h_{13}^* & \cdot & \cdot & \cdot & -i(\lambda_3 - \lambda)h_{13}^* & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ (-i)^{n-1}(\lambda_n - \lambda)^{n-1}h_{1n}^* & \cdot & \cdot & \cdot & -i(\lambda_n - \lambda)h_{1n}^* & 0 & 0 \end{pmatrix}.$$

If Assumptions 2 and 3 are verified, then all the  $h_{1j}$ ,  $j = 2, \dots, n$  are different from zero and the factors  $\lambda_n - \lambda$  are all different from each other. Therefore, a Vandermonde determinant argument shows that the rank of  $C$  is equal to  $n$ . Conversely, it is easily seen that if controllability holds for *all* the eigenvectors, then Assumptions 2 and 3 are verified.

In [146], only the pure state case is considered and convergence is proved under conditions weaker than Assumption 2 and 3. In particular, these conditions are that  $\omega_{1j} \neq \omega_{1k}$  and that only the level corresponding to the target eigenstate is directly coupled to all the others. Here the index 1 refers to the target eigenstate. This condition is equivalent to controllability of the linearized system about the target eigenstate. In our treatment, we have considered the more general case of density matrix control, which led to the stronger Assumptions 2 and 3.

In the general case of density matrices control, as treated here, Assumptions 2 and 3 guarantee convergence to a stationary point of the Lyapunov function  $V(\rho) = \text{Tr}(P\rho)$  with the control (7.34), and the value of such a function will be decreasing along the resulting trajectory. If the state reached is not the desired minimum, then in principle one could try to switch the control to a function which moves the state away from the reached stationary point so that the value of  $V(\rho)$  decreases. Then one switches again to the Lyapunov control to reach the next stationary point. Methods to design such a strategy are an interesting research problem. In the case of a pure state, as discussed in Remarks 7.5.7 and 7.5.8, assumptions weaker than 2 and 3 and equivalent to local controllability about a given eigenstate guarantee the convergence to a desired eigenstate of  $H_0$ , for almost all the initial conditions. There have been several attempts in the literature to overcome the limitations of such assumptions. The technique used in [146] uses an adiabatic trajectory which starts at the desired eigenstate  $\vec{\psi}_f$  and terminates at the same eigenstate. The control is designed via a Lyapunov function so as to track such an adiabatic trajectory. More precisely, one first considers a *slow* control  $u_T(t)$  in an interval  $[0, T]$ , with  $u_T(0) = u_T(T) = 0$ . Given the Hamiltonian  $H(t) = H_0 + u_T(t)H_1$ , denote by  $\vec{\psi}_T := \vec{\psi}_T(t)$  the eigenvector of  $H(t)$  corresponding to initial condition  $\vec{\psi}_T(0) = \vec{\psi}_T(T) = \vec{\psi}_f$ . If  $\vec{\psi}_T$  was constant,

we would use a control corresponding to (7.34) for the pure state case (cf. Remark 7.5.7)

$$\begin{aligned} u &= -iK \operatorname{Tr}(\vec{\psi}\vec{\psi}^\dagger[(\mathbf{1} - \vec{\psi}_f\vec{\psi}_f^\dagger), H_1]) \\ &= iK \operatorname{Tr}(\vec{\psi}\vec{\psi}^\dagger[\vec{\psi}_f\vec{\psi}_f^\dagger, H_1]) = iK \operatorname{Tr}(\vec{\psi}\vec{\psi}^\dagger[\vec{\psi}_T\vec{\psi}_T^\dagger, H_1]). \end{aligned}$$

The fact that  $\vec{\psi}_T$  varies slowly suggests using the control

$$u = u_T + iK \operatorname{Tr}(\vec{\psi}\vec{\psi}^\dagger[\vec{\psi}_T\vec{\psi}_T^\dagger, H_1]),$$

i.e., the sum of a term inducing the adiabatic trajectory and a term which is the one suggested by the Lyapunov design. For this control, assume there exists a time  $\bar{t} \in (0, T)$  such that the system (7.33) linearized about the eigenvector  $\vec{\psi}_T(\bar{t})$  and the controls  $u = u_T(\bar{t})$  and  $\omega = -\lambda(\bar{t})$  (with  $\lambda(\bar{t})$  the corresponding eigenvalue of  $H_0 + u_T(\bar{t})H_1$ ) is controllable. Then, given an initial distance between the initial state  $\vec{\psi}(0)$  and the desired state  $\vec{\psi}_f$  and a desired (small) final distance  $\epsilon$  between the two states, there exists a  $\bar{T}$  such that

$$1 - |\vec{\psi}^\dagger(T)\vec{\psi}_f|^2 < \epsilon,$$

for every  $T > \bar{T}$ . The proof of this result is given in [146], and it is based on an adiabatic theorem given in [75].

## 7.6 Notes and References

We have surveyed some more tools to achieve the control of finite dimensional quantum systems. The list of the techniques indicated here along with the ones of the previous two chapters is not complete. The books [174] and [192], for example, describe some more techniques in the setting of molecular control. One can combine the techniques described here and the ones in the previous two chapters. For example, in [211] a Lie Group decomposition was used to decompose a control problem into lower dimensional problems and each of them was treated using a Lyapunov approach.

Our main references for the adiabatic theorems have been [109] and [173], which can be used as a starting point to explore the literature on adiabatic theorems. It should be mentioned that Theorem 7.3.1 is much more general in [109] than in the version we have given here. In particular, the authors do not consider the adiabatic trajectory corresponding to a single eigenvector but one corresponding to an element in a given eigenspace with spectrum separated by a gap from the rest of the spectrum of the time varying Hamiltonian.

The discussion of the STIRAP technique in section 7.4 is focused on the case of  $\Lambda$  systems. The technique can be extended to other configurations.

We refer to the books [174] and [192] for these extensions as well as for several practical applications of STIRAP.

The main references in the treatment of Lyapunov control have been [91], [146] and [215]. Although these papers treat the case of pure state, we have presented the treatment in the general case of density matrix dynamics and then considered the special case of pure states in Remarks 7.5.7, 7.5.8. For example, Proposition 7.5.2 is an extension of Lemma 1 in [91] to the density matrix setting. We believe that the treatment with the density matrix is not only more general but also more compact. However, convergence is proved at the price of stronger assumptions (Assumptions 2 and 3) which can be weakened for the pure state case to conditions equivalent to the controllability of the system linearized about the target eigenstate (cf. Remark 7.5.8). The proof of convergence in the pure state case can be adapted from the one presented here for density matrices.

All the Lyapunov based controls are in feedback form, i.e., the control is a function of the state. For quantum systems the state is not available for measurement and, in fact, it follows from the measurement postulate that a measurement will modify the state and therefore (abruptly) change the dynamics of the system. Therefore, such closed loop controls have to be first simulated. From the simulation one obtains a control function which is then, in practice, applied in open loop. In the setting of quantum feedback, one considers stochastic dynamics for the quantum state and the Lyapunov function approach, in the stochastic version, has proved a useful tool to design control laws for these models (see, e.g., [214]).

## 7.7 Exercises

**Exercise 7.1** Consider an interpolating 1-qubit Hamiltonian of the form

$$H_T(t) := \left(1 - \frac{t}{T}\right)H_0 + \frac{t}{T}H_p,$$

with  $H_0 = \frac{1}{2} + \frac{1}{2}\sigma_z$  (cf. (1.20)).

- a) Give necessary and sufficient conditions on  $H_p$  so that it is possible to analyze the dynamics in terms of adiabatic evolution in the interval  $[0, T]$ , i.e., no level crossing occurs in this interval ( $\gamma_{\min} \neq 0$ ).
- b) Pick one such  $H_p$  Hamiltonian and estimate the error of the adiabatic approximation  $\|(\mathbf{1} - P_a(t))\vec{\psi}(t)\|$  using Theorem 7.3.1 or Remark (7.3.2), for  $T = 10$ ,  $T = 100$  and  $T = 1000$ .

**Exercise 7.2** Consider two spin  $\frac{1}{2}$  particles interacting via Ising interaction (cf. subsection 6.4.2). Assume the interaction is weak. Consider two constant magnetic fields in the  $z$  direction different from each other, and a common

control field in the  $y$ -direction. The Hamiltonian may be written, after some normalization, as

$$H = J\sigma_z \otimes \sigma_z + (u_{z1}\sigma_z \otimes \mathbf{1} + u_{z2}\mathbf{1} \otimes \sigma_z) + u_y(\mathbf{1} \otimes \sigma_y + \sigma_y \otimes \mathbf{1}),$$

where  $J$  is the *small* coupling constant. By making  $u_{z1}$  very much different from  $u_{z2}$ , we can create two different energy gaps and, this way, we can effectively address the two spins separately.

- a) Using this fact and the strategy of section 7.1 devise a methodology to address the two spins separately.
- b) Combine this, with the methodology of section 6.4.2, to obtain a (almost) minimum time control.
- c) Consider the passage from the state  $|00\rangle$  (both spin down), to the ‘entangled’ state  $\frac{1}{2}(|00\rangle + |11\rangle)$ . Use this strategy to perform this control and the methods of section 7.1 and 7.2 to obtain estimates for the error.

**Exercise 7.3** In section 7.1 we have assumed that the diagonal elements of  $H_1$  in (7.1) are all zero. How does the treatment modify if this is not the case?

**Exercise 7.4** Prove Proposition 7.5.3.

**Exercise 7.5** Prove the **FACT** of Remark 7.5.4.

**Exercise 7.6** Prove that

$$\min r \in \mathbf{R} \|\vec{\psi} - e^{ir}\vec{\psi}_a\| = \sqrt{2(1 - |\vec{\psi}^\dagger \vec{\psi}_a|)},$$

and

$$\|\vec{\psi} - \vec{\psi}_a \vec{\psi}_a^\dagger \vec{\psi}\| = \sqrt{1 - |\vec{\psi}_a^\dagger \vec{\psi}|^2}.$$

From this deduce that the two measures of error in the adiabatic approximation are equivalent.

# Chapter 8

---

## ***Analysis of Quantum Evolutions; Entanglement, Entanglement Measures and Dynamics***

Entanglement of quantum systems refers to the situation where two or more systems interact and the state of the whole system cannot be seen as a superposition of states of the single systems. This feature of quantum dynamics is crucial in many algorithms of quantum communication and quantum information processing and it is a resource for many fascinating applications. Moreover entanglement, as a feature of quantum mechanics, has stimulated in-depth investigation on the value of quantum mechanics itself as a physical theory.

In their famous paper [74] on the *EPR paradox*, Einstein, Podolski and Rosen proposed an experiment to argue that quantum mechanics is not a complete physical theory. In the Bohm simplified version of this experiment, a pair of spin- $\frac{1}{2}$  particles are placed in a *singlet state*  $|\psi_s\rangle$ , defined as

$$|\psi_s\rangle := \frac{1}{\sqrt{2}}(|10\rangle - |01\rangle), \quad (8.1)$$

where we have denoted by  $|0\rangle$  ( $|1\rangle$ ) the spin down (up) state. This state is **entangled** in that it cannot be written in the form  $|a\rangle \otimes |b\rangle$ , for two states  $|a\rangle$  and  $|b\rangle$  of the two systems. The two particles are sent to two different locations  $A$  and  $B$ . If a measurement is performed on the spin of the particle at location  $A$  and the result is a positive value of the spin angular momentum (say in the  $z$  direction), then the state of the system of two spins will collapse, according to the measurement postulate, into  $|\psi\rangle := |10\rangle$ , which means that a measurement of the spin at location  $B$  will, with certainty, give a negative value of the spin (corresponding to the  $|0\rangle$  eigenvector). It appears therefore that the result of an experiment at one site has immediately affected the result of an experiment at another (possibly very distant) site. This appears to contradict a principle of *locality*, i.e., that experiments at two different locations are independent of each other, a fact consistent with the prescription from *special relativity* that information cannot be transmitted at a speed faster than the speed of light. The EPR paradox has generated many theoretical and experimental studies. We shall not be concerned here with the attempts to solve the paradox, whether by providing different interpretations of quantum

mechanics, alternative theories or experimental evidence. Within the standard interpretation of quantum mechanics, we shall be concerned with the mathematical aspects of entanglement, i.e., its precise definition, detection, and quantification, as well as the aspects concerning dynamics, in particular its evolution. We shall see how Lie group decompositions of the type used in [Chapter 5](#) for control purposes can be used to analyze how entanglement evolves.

This chapter is organized in three sections. The first one, section 8.1, is devoted to the definition of entanglement both for pure states and mixed states. We also discuss methods to detect entanglement in [subsection 8.1.2](#) and to quantify it in [subsection 8.1.3](#). The second section, [section 8.2](#), is devoted to the study of the dynamics of entanglement. Several decompositions are given there which allow us to factorize a unitary evolution into factors which produce entanglement and factors that do not. The last section, [section 8.3](#), is devoted to the problem of local equivalence between states, i.e., how to determine when, given two states, it is possible to evolve from one to the other by transformations which act on systems separately (local transformations). This can be seen as a problem of controllability.

---

## 8.1 Entanglement of Quantum Systems

### 8.1.1 Basic definitions and notions

#### 8.1.1.1 Separable and entangled pure states

As discussed in Chapter 1, [subsection 1.1.3](#), the state of a quantum system composed of  $N$  subsystems is described by a vector  $|\psi\rangle$  in the tensor space

$$\mathcal{H} := \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_N,$$

where  $\mathcal{H}_j$ ,  $j = 1, 2, \dots, N$ , is the Hilbert space associated with the  $j$ -th system. A state  $|\psi\rangle \in \mathcal{H}$  is called **separable** if it can be written as

$$|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes \cdots \otimes |\psi_N\rangle,$$

where  $|\psi_j\rangle$  are states of  $\mathcal{H}_j$ ,  $j = 1, \dots, N$ . States that are not separable are called **entangled states**.

**Example 8.1.1** Consider two two-level systems so that the Hilbert space of the total system is 4-dimensional. Let  $|1\rangle := \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  and  $|0\rangle := \begin{pmatrix} 0 \\ 1 \end{pmatrix}$  form a basis for both systems. Therefore the Hilbert space of the total system is

spanned by  $|j\rangle \otimes |k\rangle$ ,  $j, k = 0, 1$ . The state

$$|1\rangle \otimes |0\rangle := \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

is separable, while the singlet state (8.1),

$$\frac{1}{\sqrt{2}}(|1\rangle \otimes |0\rangle - |0\rangle \otimes |1\rangle) := \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \quad (8.2)$$

is entangled. In fact, if we try to express the state in (8.2) as a tensor product  $\begin{pmatrix} a \\ b \end{pmatrix} \otimes \begin{pmatrix} c \\ d \end{pmatrix}$ , we find that  $a \times c$  should be zero, i.e., at least one between  $a$  and  $c$  should be zero, which contradicts the fact that both  $a \times d$  and  $b \times c$  should be different from zero.

### 8.1.1.2 Separable and entangled mixed states

If the quantum system is an ensemble, a density matrix description of the state must be used. In this context, a state  $\rho$  is said to be **separable** if

$$\rho := \sum_j \alpha_j \rho_{j1} \otimes \rho_{j2} \otimes \cdots \otimes \rho_{jN}, \quad (8.3)$$

$$\alpha_j > 0, \quad \sum_j \alpha_j = 1,$$

where all the  $\rho_{jk}$ ,  $k = 1, \dots, N$ , are themselves states (i.e., self-adjoint, positive semidefinite, trace 1 operators) for the single subsystems. We call a **product state** a state of the form

$$\rho_{Prod} := \rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_N,$$

where the  $\rho_j$ 's,  $j = 1, \dots, N$ , are themselves states. Therefore, we can say that a separable state is a statistical (convex) superposition of product states. Notice that there is no loss of generality in assuming that each factor in the product states in (8.3) is a pure state. In fact, by the definition (1.17) every term  $\rho_{jk}$ ,  $k = 1, \dots, N$  in (8.3) has the form

$$\rho_{jk} = \sum_l c_{jk,l} |\psi_{jk,l}\rangle \langle \psi_{jk,l}|, \quad c_{jk,l} > 0, \quad \sum_l c_{jk,l} = 1,$$

and we can re-arrange the sum (8.3) by possibly redefining the coefficients  $\alpha_j$ , and define a separable state as a state of the form

$$\rho := \sum_j \alpha_j |\psi_{j1}\rangle \langle \psi_{j1}| \otimes |\psi_{j2}\rangle \langle \psi_{j2}| \otimes \cdots \otimes |\psi_{jN}\rangle \langle \psi_{jN}|, \quad (8.4)$$

with

$$\alpha_j > 0, \quad \sum_j \alpha_j = 1.$$

Notice that each term in the sum (8.4) represents a pure separable state since

$$|\psi_{j1}\rangle\langle\psi_{j1}| \otimes |\psi_{j2}\rangle\langle\psi_{j2}| \otimes \cdots \otimes |\psi_{jN}\rangle\langle\psi_{jN}| = \\ (|\psi_{j1}\rangle \otimes |\psi_{j2}\rangle \otimes \cdots |\psi_{jN}\rangle)(\langle\psi_{j1}| \otimes \langle\psi_{j2}| \otimes \cdots \langle\psi_{jN}|).$$

Therefore, we can say that a mixed separable state is the statistical superposition of pure separable states. In particular, rank one mixed separable states are pure separable states.

A mixed state which is not separable is called **entangled**.

Notice that, being a Hermitian operator,  $\rho$  can *always* be written in the form (8.3), with general real coefficients  $\alpha_j$  and general Hermitian matrices  $\rho_{j1}, \dots, \rho_{jN}$ . However only when the  $\alpha_j$  satisfy the conditions in (8.3) and the operators  $\rho_j, \dots, \rho_j$  represent states can we say that  $\rho$  is a separable state.

### 8.1.1.3 Partial trace

Given a general mixed state  $\rho$  describing the state of a multipartite quantum system, we would like to have a way to extract the state of one of its subsystems, say the first one. ‘To extract’ means here that when we measure any observable which concerns only the first subsystem, i.e., it is of the type  $A \otimes \mathbf{1} \otimes \mathbf{1} \otimes \cdots \otimes \mathbf{1}$ , we are able to assess the probability distribution of the possible results, just as if we had the state of the first system alone. The appropriate tool is the **partial trace** of the density operator  $\rho$ . The definition can be given for a bipartite system first and then extended (see [Property 3 in Proposition 8.1.3](#)).

**Definition 8.1.2** Consider a bipartite system with Hilbert space<sup>1</sup>  $\mathcal{H} := \mathcal{H}_S \otimes \mathcal{H}_P$ . Let  $\rho$  be a Hermitian operator on the total Hilbert space  $\mathcal{H}$ . Write  $\rho$  in the form  $\rho = \sum_l \alpha_l \sigma_l^S \otimes \sigma_l^P$ , for some real coefficients  $\alpha_l$  and Hermitian operators  $\sigma_l^S$  and  $\sigma_l^P$  on  $\mathcal{H}_S$  and  $\mathcal{H}_P$ , respectively. The partial trace  $\text{Tr}_P$  of a state  $\rho$  with respect to (the second system)  $P$  is defined as

$$\rho_S := \text{Tr}_P(\rho) := \sum_l \alpha_l \sigma_l^S \text{Tr}(\sigma_l^P).$$

Analogously one defines the partial trace with respect to the system  $S$ .

The main properties of the partial trace are listed in the following proposition whose proof is left as an exercise (see [Exercise 8.1](#)).

---

<sup>1</sup>The letters  $S$  and  $P$  are used here to indicate the main object system  $S$  and an auxiliary system  $P$  (probe).

**Proposition 8.1.3** The partial trace  $\text{Tr}_P$  has the following properties

1. It is the *unique* linear operator from the space of linear operators on  $\mathcal{H}_S \otimes \mathcal{H}_P$  to the space of linear operators on  $\mathcal{H}_S$  having the property

$$\text{Tr}_P(A \otimes B) = A \text{Tr}(B).$$

2. For general operators  $F$ , and unitary operator  $G \in U(n_P)$ ,

$$\text{Tr}_P(F \otimes G\rho F^\dagger \otimes G^\dagger) = F \text{Tr}_P(\rho)F^\dagger \quad (8.5)$$

3. Consider a multipartite system on  $\mathcal{H} := \mathcal{H}_S \otimes \mathcal{H}_{P_1} \otimes \mathcal{H}_{P_2} \otimes \cdots \otimes \mathcal{H}_{P_{N-1}}$  and denote by  $P_1 P_2 \cdots P_{N-1}$  the subsystem corresponding to  $\mathcal{H}_{P_1} \otimes \cdots \otimes \mathcal{H}_{P_{N-1}}$ . Then

$$\text{Tr}_{P_1 P_2 \cdots P_{N-1}}(\rho) = \text{Tr}_{P_1}(\text{Tr}_{P_2}(\cdots(\text{Tr}_{P_{N-1}}(\rho)))).$$

4. For every observable  $M$  on  $\mathcal{H}_S$

$$\text{Tr}(M \text{Tr}_P(\rho)) = \text{Tr}(M \otimes \mathbf{1}\rho). \quad (8.6)$$

5. Consider an orthonormal basis of  $\mathcal{H}_S$ ,  $\{|s_j\rangle\}$ ,  $j = 1, 2, \dots, n_S$  and an orthonormal basis of  $\mathcal{H}_P$ ,  $\{|p_k\rangle\}$ ,  $k = 1, 2, \dots, n_P$ . The corresponding basis of  $\mathcal{H}_S \otimes \mathcal{H}_P$  is  $\{|s_j, p_k\rangle\}$  and, in this basis, a Hermitian operator  $A$  can be written as

$$A = \sum_{jk,lm} a_{jk,lm} |s_j, p_k\rangle \langle s_l, p_m|,$$

with  $a_{jk,lm}^* = a_{lm,jk}$ . Then

$$\text{Tr}_P(A) = \sum_{j,l} b_{j,l} |s_j\rangle \langle s_l|,$$

with  $b_{j,l} := \sum_{k=1}^{n_P} a_{jk,lk}$ .

Property (8.6) is the reason to choose the partial trace to describe the state of a subsystem of a quantum system. Given  $M$ , an observable on system  $\mathcal{H}_S$ , the corresponding observable on the overall subsystem is  $M \otimes \mathbf{1}$ . The state of the system  $S$  has to be chosen so that the statistics of  $M$  on  $S$  is the same as the one of  $M \otimes \mathbf{1}$  on the system  $S + P$ . In particular, the expectation value of  $M$  should be the same, which gives equation (8.6). The partial trace is the only map which has this property (see, e.g., [155]). In fact, consider any other map  $f$  from the space of linear Hermitian operators on  $\mathcal{H}_S \otimes \mathcal{H}_P$  to the space of linear operators on  $\mathcal{H}_S$ , satisfying, for every  $M$ ,

$$\text{Tr}(Mf(\rho)) = \text{Tr}(M \otimes \mathbf{1}\rho).$$

Since this is true for every element  $M_j$  of an orthonormal basis in the space of linear operators on  $\mathcal{H}_S$ , comparing with (8.6) we obtain that the components of  $f(\rho)$  and  $\text{Tr}_P(\rho)$  along every  $M_j$  coincide, which means that  $f(\rho) = \text{Tr}_P(\rho)$ .

#### 8.1.1.4 Von Neumann entropy

In probability theory, given a probability distribution  $\mathcal{P} := \{p_1, \dots, p_n\}$ ,  $p_j \geq 0$ ,  $\sum_{j=1}^n p_j = 1$ , describing the results of an experiment, one defines **entropy** as a measure of the uncertainty on the outcome of the experiment. Entropy,  $E(\mathcal{P})$ , is defined as

$$E(\mathcal{P}) := - \sum_{j=1}^n p_j \log(p_j),$$

where the log is usually taken in base 2, and it is understood that  $0 \log(0) = 0$ .  $E(\mathcal{P})$  is minimum and equal to zero when only one of the probabilities  $p_j$  is different from zero (i.e., one of the events has probability one). It is maximum and equal to  $\log(n)$  when all the outcomes have the same probability.

In analogy with probability theory, in quantum mechanics one defines **Von Neumann entropy** as a measure of the uncertainty of a quantum state. In particular, the state of an ensemble  $\rho$  can always be written as

$$\rho := \sum_{j=1}^n \lambda_j |\psi_j\rangle\langle\psi_j|,$$

for some orthonormal states  $|\psi_j\rangle$ ,  $\lambda_j \geq 0$ ,  $\sum_j \lambda_j = 1$ . This can be interpreted as an ensemble of systems where a fraction  $\lambda_j$  is in the pure state  $|\psi_j\rangle$ . The Von Neumann entropy of the state  $\rho$  is defined as

$$S(\rho) := - \sum_{j=1}^n \lambda_j \log(\lambda_j) = - \text{Tr}(\rho \log(\rho)),$$

with the understanding that  $0 \log(0) = 0$ . It is minimum and equal to zero for a pure state. It is maximum and equal to  $\log(n)$  for the perfectly mixed state  $\rho := \frac{1}{n}\mathbf{1}$ . Von Neumann entropy is a fundamental quantity in quantum mechanics and in particular in quantum information theory. We refer to [156] for an exposition on the use of entropy in quantum mechanics. Von Neumann entropy can be used as a test for a quantum state to be pure. We have (Exercise 8.2)

**Proposition 8.1.4** The state  $\rho$  is a pure state if and only if  $S(\rho) = 0$ .

#### 8.1.1.5 Bipartite and multipartite systems

Entanglement theory is very different if we consider bipartite systems and general multipartite systems and much fewer results are known in the second

case. Even for the bipartite case, many problems are still open and only the simplest case of two coupled two-level systems (*qubits*) seems to be sufficiently understood at the moment (see, e.g., [1], [98], [103], [227] and references therein). From an application point of view, the bipartite is by far the most interesting one as it seems very difficult in practice to entangle (i.e., to create entangled states of) more than two quantum systems. We shall devote our discussion in the following mainly to the bipartite case and then indicate ideas and generalizations for the multipartite case.

### 8.1.2 Tests of entanglement

A fundamental and practical problem in entanglement theory is to give mathematical criteria to decide whether a quantum state is entangled or not. In the following we shall discuss some of the most powerful criteria for bipartite pure and mixed states, and, more briefly, for multipartite states.

#### 8.1.2.1 Tests of entanglement for bipartite pure states

The partial trace itself provides a test for a pure state of a bipartite system to be entangled or separable. In fact, when the state of system  $S$  is entangled with the state of system  $P$ , the partial trace will give a mixed state. In information theoretic terms, some information on the state of system  $S$  is lost in the interaction with system  $P$ , so the Von Neumann entropy of the partial trace will be strictly positive (cf. Proposition 8.1.4). Formally, we have the following test of entanglement for bipartite pure systems.

**Proposition 8.1.5** Let  $\rho := |\psi\rangle\langle\psi|$  represent a pure state for a bipartite system  $S + P$ . Then  $\rho$  is separable if and only  $\text{Tr}_P(\rho)$  is pure, that is (cf. Proposition 8.1.4) if and only if the entropy of the partial trace is equal to zero, i.e.,

$$S(\text{Tr}_P(\rho)) = 0.$$

Notice that there is nothing special about  $P$  in the above statement and we could have replaced the partial trace  $\text{Tr}_P$  with the partial trace  $\text{Tr}_S$ . This also means that  $\text{Tr}_P(\rho)$  is pure if and only if  $\text{Tr}_S(\rho)$  is pure.

In the proof of 8.1.5 we use an important and very useful result for bipartite systems, known as **Schmidt decomposition**, given in the following lemma.

**Lemma 8.1.6** Let  $|\psi\rangle$  be a state of the tensor product space  $\mathcal{H}_S \otimes \mathcal{H}_P$  of dimensions  $n_S \times n_P$ . Assume without loss of generality  $n_S \leq n_P$ . Then there exists an orthonormal basis  $\{|s_j\rangle\}$ ,  $j = 1, \dots, n_S$ , of  $\mathcal{H}_S$  and an orthonormal basis  $\{|p_k\rangle\}$ ,  $k = 1, \dots, n_P$ , of  $\mathcal{H}_P$  such that

$$|\psi\rangle = \sum_{l=1}^{n_S} r_l |s_l\rangle \otimes |p_l\rangle. \quad (8.7)$$

Moreover the coefficients  $r_l$  can be chosen real and nonnegative.

*Proof.* The proof follows from the *singular value decomposition* (see, e.g., [99]). Consider an orthonormal basis  $\{|e_j\rangle\}$ ,  $j = 1, \dots, n_S$ , of  $\mathcal{H}_S$  and an orthonormal basis  $\{|f_k\rangle\}$ ,  $k = 1, \dots, n_P$ , of  $\mathcal{H}_P$ . Then every vector  $|\psi\rangle \in \mathcal{H}_S \otimes \mathcal{H}_P$  can be written as

$$|\psi\rangle = \sum_{j=1}^{n_S} \sum_{k=1}^{n_P} a_{jk} |e_j\rangle \otimes |f_k\rangle. \quad (8.8)$$

The  $n_S \times n_P$  matrix  $A := \{a_{jk}\}$  has the singular value decomposition  $A = U\Lambda V$ , where  $U := \{u_{jl}\}$  and  $V := \{v_{lk}\}$  are unitary matrices of dimensions  $n_S$  and  $n_P$ , respectively and  $\Lambda$  is an  $n_S \times n_P$  matrix, where the last  $n_P - n_S$  columns are zero while the first  $n_S$  columns are occupied by the diagonal matrix  $\text{diag}(\lambda_{11}, \dots, \lambda_{n_S n_S})$ . Using this, we can write (8.8) as

$$|\psi\rangle := \sum_{j=1}^{n_S} \sum_{k=1}^{n_P} \sum_{l=1}^{n_S} u_{jl} \lambda_{ll} v_{lk} |e_j\rangle \otimes |f_k\rangle = \sum_l \lambda_{ll} \left( \sum_{j=1}^{n_S} u_{jl} |e_j\rangle \right) \otimes \left( \sum_{k=1}^{n_P} v_{lk} |f_k\rangle \right).$$

This gives

$$|\psi\rangle = \sum_{l=1}^{n_S} \lambda_{ll} |s_l\rangle \otimes |p_l\rangle,$$

if one defines  $|s_l\rangle := (\sum_{j=1}^{n_S} u_{jl} |e_j\rangle)$  and  $|p_l\rangle := (\sum_{k=1}^{n_P} v_{lk} |f_k\rangle)$ . Using the fact that  $U$  and  $V$  are unitary and that the bases  $\{|e_j\rangle\}$  and  $\{|f_k\rangle\}$  are orthonormal, it is easily seen that  $\{|s_l\rangle\}$  and  $\{|p_l\rangle\}$ ,  $l = 1, \dots, n_S$  are orthonormal sets. One obtains (8.7) defining  $r_l := \lambda_{ll}$  and recalling that, from the properties of the singular value decomposition  $\lambda_{ll} \geq 0$ .<sup>2</sup>  $\square$

The main advantage of Schmidt decomposition is that it replaces a double sum (8.8) with a single sum (8.7). This allows us to give a straightforward proof of the entanglement test of Proposition 8.1.5.

*Proof of Proposition 8.1.5* It is obvious that if  $|\psi\rangle$  is a separable pure state, i.e.,

$$|\psi\rangle = |\psi_S\rangle \otimes |\psi_P\rangle,$$

with  $|\psi_S\rangle \in \mathcal{H}_S$  and  $|\psi_P\rangle \in \mathcal{H}_P$ , then

$$\text{Tr}_P(|\psi\rangle\langle\psi|) = \text{Tr}_P(|\psi_S\rangle\langle\psi_S| \otimes |\psi_P\rangle\langle\psi_P|) = |\psi_S\rangle\langle\psi_S|$$

is a pure state. Conversely, assume that  $\text{Tr}_P(|\psi\rangle\langle\psi|)$  is a pure state. Write  $|\psi\rangle$  according to the Schmidt decomposition (8.7). We have

$$\text{Tr}_P(|\psi\rangle\langle\psi|) = \text{Tr}_P\left(\sum_{lj} r_l r_j |s_l\rangle\langle s_j| \otimes |p_l\rangle\langle p_j|\right) =$$

---

<sup>2</sup>Alternatively by writing  $\lambda_{ll} := r_l e^{i\phi_l}$  with  $r_l := |\lambda_{ll}|$  and  $\phi_l := \arg(\lambda_{ll})$  and redefining  $e^{i\phi_l} |s_l\rangle \rightarrow |s_l\rangle$ , one obtains that  $r_l \geq 0$ .

$$\sum_{lj} r_l r_j |s_l\rangle\langle s_j| \delta_{lj} = \sum_l r_l^2 |s_l\rangle\langle s_l|.$$

As this is a pure state, all the  $r_l$ 's must be equal to zero except one, which is equal to one. Therefore from (8.7),  $|\psi\rangle$  is a separable state.  $\square$

### 8.1.2.2 Tests of entanglement for bipartite mixed states

For general mixed states the problem of finding a test to check whether the state is entangled or separable has not received a complete answer and it is one of the unsolved problems in quantum information theory. There have been, however, several important results in this area, some of which are of fundamental value.

A fundamental result is the necessary condition for states to be separable in terms of **positive maps**, which we now describe. Let  $\mathcal{H}$  be a finite dimensional Hilbert space and the set of linear operators on  $\mathcal{H}$  be denoted by  $\mathcal{B}(\mathcal{H})$ . This is also a Hilbert space with the inner product  $\langle A, B \rangle := \text{Tr}(B^\dagger A)$ . A linear operator  $\mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$  is called a *superoperator* or simply a *map*. We shall use the definition of positive and completely positive maps which is given in [Appendix A](#). We have the following fundamental result proved in [103].

**Theorem 8.1.7** *A density matrix  $\rho$  on  $\mathcal{H}_S \otimes \mathcal{H}_P$  is separable if and only if for any positive map  $\Gamma : \mathcal{B}(\mathcal{H}_S) \rightarrow \mathcal{B}(\mathcal{H}_S)$ ,*

$$(\Gamma \otimes \mathbf{1}_{n_P})(\rho) \geq 0.$$

From the above Theorem, if we had all the positive maps  $\Gamma$  we could in principle test separability of any state  $\rho$ . However, finding the structure of all the positive maps is an extremely difficult problem. It is however possible that a quorum of maps could be sufficient to discern separable and entangled states. In general if we find a positive map  $\Gamma$  such that

$$\Gamma \otimes \mathbf{1}_{n_P}(\rho) \not\geq 0, \tag{8.9}$$

then we can say that  $\rho$  is entangled. Several entanglement criteria for mixed states are of this type. They use specific positive maps  $\Gamma$ , for which condition (8.9) is satisfied for some entangled states  $\rho$ . In that case one says that the map  $\Gamma$  *detects* entanglement in  $\rho$ . A suitable map  $\Gamma$  cannot be chosen among the completely positive maps otherwise (8.9) will never be satisfied. More generally, if a map  $\Gamma_0$  detects a certain set of entangled states, any map  $F + G \circ \Gamma_0$ , with  $F$  and  $G$  completely positive, will recognize a subset of the states recognized by  $\Gamma_0$ . In fact, if  $\rho$  is not detected by  $\Gamma_0$ , i.e.,  $\Gamma_0 \otimes \mathbf{1}_{n_P}(\rho) \geq 0$ , then

$$(F \otimes \mathbf{1}_{n_P})(\rho) + (G \otimes \mathbf{1}_{n_P}) \circ (\Gamma_0 \otimes \mathbf{1}_{n_P})(\rho) \geq 0.$$

Notice that if we consider the completely positive map  $G$  given by the unitary evolution  $G(\rho_S) := U\rho_SU^\dagger$ , with  $U$  unitary,  $\Gamma_0$  and  $G \circ \Gamma_0$  are able to recognize exactly the same entangled states. The tests based on  $\Gamma_0$  and  $G \circ \Gamma_0$  are therefore equivalent (they are called *unitarily equivalent*).

We present three important tests of entanglement for bipartite mixed states, in the following three theorems. The first two correspond to the choice of positive maps in Theorem 8.1.7 which are not completely positive.

**Theorem 8.1.8 Peres-Horodecki Positive Partial Transposition criterion [103], [160].** *Let  $\Gamma_T : \mathcal{B}(\mathcal{H}_S) \rightarrow \mathcal{B}(\mathcal{H}_S)$  be given by transposition, i.e.,  $\Gamma_T(A) := A^T$ . Then, if  $\rho$  is separable,  $(\Gamma_T \otimes \mathbf{1}_{n_P})(\rho) \geq 0$ . Assume  $n_S = 2$  and  $n_P = 2$  or  $n_S = 2$  and  $n_P = 3$ . Then  $\rho$  is separable if and only if  $\Gamma_T \otimes \mathbf{1}_{n_P}(\rho) \geq 0$ .*

**Example 8.1.9** Consider the state

$$\rho := \frac{1}{4} \begin{pmatrix} 1 & i & 0 & 0 \\ -i & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

which can be written as

$$\rho = \frac{1}{2} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \frac{1}{4} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{4} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

From this expression, it is clear that  $(\Gamma_T \otimes \mathbf{1}_2)(\rho) = \rho \geq 0$  and this along with the fact that the system is such that  $n_S = n_P = 2$  shows that the state is separable. In fact,  $\rho$  can be written as the convex combination of two product states,

$$\rho = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} \frac{1}{2} & \frac{i}{2} \\ \frac{-i}{2} & \frac{1}{2} \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Consider now the state

$$\rho = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

which can be written as

$$\begin{aligned} \rho = & \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \\ & \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}. \end{aligned} \quad (8.10)$$

From this expression, one can easily check that

$$(\Gamma_T \otimes \mathbf{1}_2)(\rho) = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix},$$

which has an eigenvalue equal to  $-1$  and therefore is not positive semidefinite. This shows that this state is not separable. It is entangled. In fact this state is the singlet state (8.2).

Although it gives a complete answer only for the low-dimensional cases  $2 \times 2$  and  $2 \times 3$ , the Positive Partial Transposition (PPT) criterion of Theorem 8.1.7 has proved a very powerful one in that it detects a large class of entangled states. Any interesting new criterion will have to be able to recognize entanglement in states that satisfy the necessary condition of separability of the PPT criterion, which are called *PPT states*. The following two additional criteria are able to recognize some entangled PPT states.

The criterion of [31], [32] is defined in terms of a different positive but not completely positive map. This is done with the help of the time reversal symmetry which is the **AII** Cartan symmetry,  $\tilde{\theta}_{II}$  defined in [Table 5.1](#) of Chapter 5 (with  $T = \mathbf{1}$ ). The map  $\tilde{\theta}_{II}$  is given, for any element  $A$  in  $i\mathfrak{u}(n_S)$ , with  $n_S$  even, by

$$\tilde{\theta}_{II}(A) = J\bar{A}J^\dagger = JA^T J^\dagger,$$

with the matrix  $J$  defined in (3.18). This map is a positive map and, in particular, it is unitarily equivalent to the transposition. Therefore a criterion based on Theorem 8.1.7 and the map  $\Gamma := \tilde{\theta}_{II}$  does not detect any new state other than the ones already detected by the PPT criterion of Theorem 8.1.8. However, let us modify the map  $\tilde{\theta}_{II}$  as follows

$$\Phi(A) := \text{Tr}(A)\mathbf{1} - A - \tilde{\theta}_{II}(A). \quad (8.11)$$

$\Phi$  is a positive map. To check this, it is enough to check that  $\Phi(|\psi\rangle\langle\psi|)$  is positive for any normalized state  $|\psi\rangle$ . We have, using the definition (8.11),

$$\Phi(|\psi\rangle\langle\psi|) = \mathbf{1} - \Pi, \quad \Pi := |\psi\rangle\langle\psi| + J|\bar{\psi}\rangle\langle\bar{\psi}|J^\dagger,$$

with  $\Pi \geq 0$  as it is an orthogonal projection onto the space spanned by  $|\psi\rangle$  and  $J|\bar{\psi}\rangle$ .<sup>3</sup> Therefore, applying the fundamental Theorem 8.1.7, we have the following criterion of separability which was proposed by H-P. Breuer in [31], [32].

---

<sup>3</sup>The main step in the proof that  $\Pi^2 = \Pi$  is to recognize that  $\langle\psi|J|\bar{\psi}\rangle = 0$ . Here  $|\bar{\psi}\rangle$  denotes the complex conjugate of  $|\psi\rangle$ .

**Theorem 8.1.10 (Breuer's criterion of separability)** *If  $\rho$  is separable then*

$$(\Phi \otimes \mathbf{1})(\rho) \geq 0.$$

Therefore if  $(\Phi \otimes \mathbf{1})(\rho) \not\geq 0$ ,  $\rho$  is entangled. By studying a special family of states, it is possible to show that Breuer's criterion detects some PPT entangled states. We refer for this as well as for several other (optimality) properties of Breuer's criterion to [31], [32].

The **computable cross norm (CCN)**, or **realignment criterion**, [47], [180], [181], [182], involves the calculation of a particular norm of a linear operator  $\Gamma : \mathcal{B}(\mathcal{H}_P) \rightarrow \mathcal{B}(\mathcal{H}_S)$ . This criterion has several equivalent formulations [181]. Consider the density matrix  $\rho$  as an operator on  $\mathcal{H}_S \otimes \mathcal{H}_P$ , i.e., as an element of  $\mathcal{B}(\mathcal{H}_S \otimes \mathcal{H}_P)$ . We can write  $\rho$  as

$$\rho = \sum_{j,k=1}^{n_S} \sum_{l,m=1}^{n_P} \rho_{jk,lm} E_{jk} \otimes E_{lm},$$

where  $E_{jk} := e_j e_k^T$  ( $E_{lm} = e_l e_m^T$ ), where  $e_j$  denotes the element of the standard orthonormal basis in  $\mathcal{H}_S$  or  $\mathcal{H}_P$ . This representation of  $\rho$  induces a map  $\Gamma_\rho : \mathcal{B}(\mathcal{H}_P) \rightarrow \mathcal{B}(\mathcal{H}_S)$  as follows,

$$\Gamma_\rho(A) := \sum_{j,k=1}^{n_S} \sum_{l,m=1}^{n_P} \rho_{jk,lm} E_{jk} \text{Tr}(E_{lm}^\dagger A),$$

which is clearly a linear map. If  $T_{\Gamma_\rho}$  is the matrix representation of the map  $\Gamma_\rho$ , we can define the *trace class norm*  $\|T_{\Gamma_\rho}\|_1$  as the sum of the magnitudes of the singular values of  $T_{\Gamma_\rho}$ . We have the following criterion.

**Theorem 8.1.11 CCN or Realignment Criterion ([47], [180], [181], [182])** *If a state  $\rho$  is separable then*

$$\|T_{\Gamma_\rho}\|_1 \leq 1.$$

It follows that if  $\|T_{\Gamma_\rho}\|_1 > 1$ , then  $\rho$  is entangled. Contrary to the PPT criterion, the realignment criterion is not, in general, necessary and sufficient for  $2 \times 2$  (two qubits) systems. However the criterion has proved strong in detecting many PPT entangled states. We refer to [47], [180], [181], [182] for a study of several properties of this criterion. We illustrate the calculations involved with a simple two qubits example taken from [180].

**Example 8.1.12** Consider the class of 2 qubits states

$$\rho_p := p|00\rangle\langle 00| + \frac{(1-p)}{2}(|01\rangle + |10\rangle)(\langle 01| + \langle 10|), \quad (8.12)$$

depending on the parameter  $p$ ,  $0 \leq p \leq 1$ . The matrix  $\rho_p$  can be written as

$$\rho_p := pE_{00} \otimes E_{00} + \frac{1-p}{2}(E_{00} \otimes E_{11} + E_{11} \otimes E_{00} + E_{10} \otimes E_{01} + E_{01} \otimes E_{10}),$$

where  $E_{jk} := |j\rangle\langle k|$ . The corresponding linear operator  $\Gamma_{\rho_p}$  is such that

$$\begin{aligned}\Gamma_{\rho_p}(E_{00}) &= pE_{00} + \frac{1-p}{2}E_{11}, \\ \Gamma_{\rho_p}(E_{01}) &= \frac{1-p}{2}E_{10}, \\ \Gamma_{\rho_p}(E_{10}) &= \frac{1-p}{2}E_{01}, \\ \Gamma_{\rho_p}(E_{11}) &= \frac{1-p}{2}E_{00}.\end{aligned}$$

Therefore, the matrix representation of  $\Gamma_{\rho_p}$  is

$$T := \begin{pmatrix} p & 0 & 0 & \frac{1-p}{2} \\ 0 & 0 & \frac{1-p}{2} & 0 \\ 0 & \frac{1-p}{2} & 0 & 0 \\ \frac{1-p}{2} & 0 & 0 & 0 \end{pmatrix}.$$

The trace class norm  $\|T\|_1$  is calculated by calculating the singular values of  $T$ . The eigenvalues of  $T^T T$  are

$$\lambda_{1,2} = \left(\frac{1-p}{2}\right)^2, \quad \lambda_{3,4} := \frac{1}{2} \left(p^2 + \frac{(1-p)^2}{2} \pm \sqrt{p^4 + p^2(1-p)^2}\right).$$

Therefore

$$\begin{aligned}\|T\|_1 &= \sum_{l=1}^4 \sqrt{\lambda_l} = 1-p + \sqrt{\frac{1}{2} \left(p^2 + \frac{(1-p)^2}{2} + \sqrt{p^4 + p^2(1-p)^2}\right)} + \\ &\quad \sqrt{\frac{1}{2} \left(p^2 + \frac{(1-p)^2}{2} - \sqrt{p^4 + p^2(1-p)^2}\right)}.\end{aligned}$$

From this it follows that

$$\|T\|_1 \geq 1-p + p\sqrt{1 + \frac{(1-p)^2}{2p^2}} \geq 1,$$

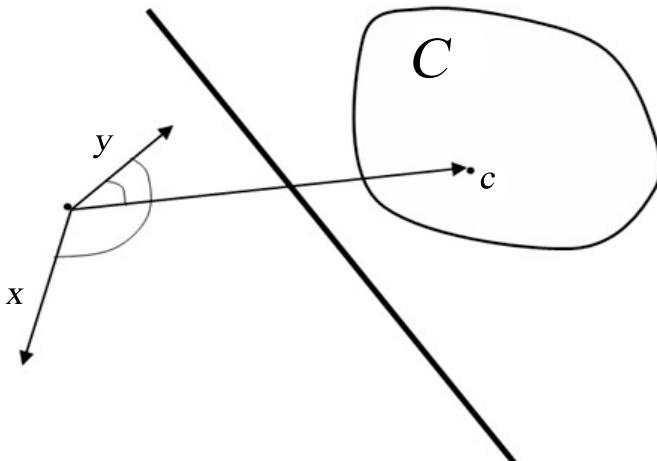
with equality if and only if  $p = 1$ . It follows from the criterion that the states (8.12) are entangled for every value of  $p$  except for  $p = 1$ . In this case, clearly  $\rho_p$  is a separable state (in fact it is a product state).

### 8.1.2.3 Tests of entanglement for multipartite systems

For pure states of multipartite systems, one may think of extending the partial trace test of Proposition 8.1.5. It is clear in fact that if a pure state is separable, the partial trace with respect to *any* subset of subsystems, taken as a subsystem, must be a pure state. The proof of the converse does not go through in the same way because the Schmidt decomposition does not extend simply to multipartite systems although several generalizations and methods to express states of multipartite systems in a canonical way exist [44]. Nevertheless, this test should detect generic entangled states. More specific results exist for pure quantum states with more structure. For example, tests for separability of pure states on  $n$  quantum bits are given in [110].

For mixed states, the tests described above for the bipartite case do not directly extend, although some other tests do extend naturally from the bipartite to the multipartite case. This is the case for the numerical tests based on semi-definite programming presented in [70]. In particular in this paper a sequence of tests is proposed that would eventually detect if a state is entangled.

One important concept which extends from the bipartite to the multipartite case is the existence of *entanglement witnesses*. An entanglement witness is an observable  $W$  such that  $\text{Tr}(W\rho) \geq 0$  for all separable states  $\rho$  and  $\text{Tr}(W\rho_e) < 0$  for some entangled state  $\rho_e$ . If this is the case, one says that  $W$  detects  $\rho_e$ . The existence of a witness which detects an entangled state follows from a classical theorem of convex analysis [175], the *separating hyperplane theorem*, noticing that the set of separable states is a convex and compact set in the space of Hermitian matrices. The separating hyperplane theorem states that for a compact convex set  $\mathcal{C}$  in a real inner product space  $\mathcal{H}$  and an element  $x \notin \mathcal{C}$  there exists a hyperplane separating  $x$  from  $\mathcal{C}$ . This means that there exists a  $y := y(x) \in \mathcal{H}$  such that  $(y, c) \geq 0$ , for all  $c \in \mathcal{C}$  and  $(y, x) < 0$ , where  $(\cdot, \cdot)$  denotes the inner product in  $\mathcal{H}$ . This can be applied to the set of Hermitian matrices with inner product  $(A, B) := \text{Tr}(AB)$ . Noticing that the set of separable density matrices is compact and convex *both* in the bipartite and in the multipartite case, the theorem says that if a state is entangled, there must exist an entanglement witness which detects it. Therefore checking whether a state is entangled amounts to checking whether an entanglement witness exists. If a constructive proof of existence is provided, then the witness  $W$  also gives, in principle, a method to test entanglement experimentally. In fact  $\text{Tr}(W\rho)$  is the expectation value of the observable  $W$ . Therefore, a measurement of  $W$  over many copies of  $\rho$  allows to detect the entanglement of  $\rho$ . Some further information concerning entanglement tests and in particular the use of entanglement witnesses can be found in the survey paper [202].



**FIGURE 8.1:** The Separating Hyperplane Theorem in  $\mathbb{R}^2$ . The angle between the vector  $y$  and the vector  $x$  is larger than 90 degrees while for any vector  $c$  whose endpoint is in the compact and convex set  $C$  the angle is less or equal than 90 degrees.

### 8.1.3 Measures of entanglement and concurrence

Measures of entanglement are defined in the *finite regime*, i.e., on a single system, or the *asymptotic regime*, i.e., on a sequence of systems. Their definitions are dictated by *abstract*, i.e., mathematical or *operational* considerations. In the operational definition one state is more or less entangled if it is more or less utilizable for some task where entanglement is a resource [101]. We are here interested in properties of single states and their evolution; therefore we shall focus on the finite regime.

#### 8.1.3.1 General axioms for entanglement measures

A **measure of entanglement** is a positive real valued function  $E$  on the set of density matrices<sup>4</sup> which satisfies the following axioms:

1.  $E$  is zero on separable states.

---

<sup>4</sup>Sometimes measures are defined on the set of *all* density matrices so that one can compare for example the entanglement of a four level system with one of a two level system. In this setting, it is reasonable to require that, for every measure of entanglement  $E$ ,  $E(\rho \otimes \rho) = 2E(\rho)$ , i.e., entanglement does not increase by simply tensoring two states  $\rho$  and the amount of entanglement present in two copies of  $\rho$  is double the one presented in one copy. In particular, this, along with the convexity property (8.14), implies that the entanglement on a separable state is zero as in property 1. It will be clear from the context whether we are talking about a measure of entanglement defined on the set of all states or for a specific bipartite or multipartite system with a given structure.

2.  $E$  does not increase under LOCC operations, i.e., if  $\Phi$  is an LOCC operation

$$E\left(\frac{1}{\text{Tr}(\Phi(\rho))}\Phi(\rho)\right) \leq E(\rho). \quad (8.13)$$

3.  $E$  is convex, i.e.,

$$E\left(\sum_j p_j \rho_j\right) \leq \sum_j p_j E(\rho_j), \quad (8.14)$$

for arbitrary states  $\rho_j$  and  $p_j > 0$ , with  $\sum_j p_j = 1$ , i.e., entanglement does not increase on average under mixing of states.

LOCC operations are *Local Operations along with Classical Communication*. Recall (cf. [Appendix A](#)) that an operation  $\Phi$  is a completely positive linear map from  $\mathcal{B}(\mathcal{H}_1)$  (the set of bounded linear operators on  $\mathcal{H}_1$ , where  $\mathcal{H}_1$  is the Hilbert space under consideration) to  $\mathcal{B}(\mathcal{H}_2)$  which is trace-nonincreasing, i.e.,  $\text{Tr}(\Phi(\rho)) \leq \text{Tr}(\rho)$ . By the Kraus representation theorem [127] every operation has the *operator sum representation*

$$\Phi(\rho) = \sum_k \Omega_k \rho \Omega_k^\dagger.$$

Local operations are the ones where the  $\Omega_k$  are of the form

$$\Omega_k := \mathbf{1} \otimes \cdots \otimes \mathbf{1} \otimes \tilde{\Omega}_k \otimes \mathbf{1} \otimes \cdots \otimes \mathbf{1},$$

i.e., they act only on one of the subsystems. LOCC allow for classical communication, i.e., one local operation may depend on the result of another at a different location. Therefore the composition of two local operations along with classical communication might have the form  $\tilde{\Phi}_\alpha \circ \Phi(\rho)$ , where  $\alpha$  is the result of the operation  $\Phi$ . The combination of local operations and classical communication always has the form

$$\Phi_{sep}(\rho) = \sum_j \Omega_{1j} \otimes \Omega_{2j} \otimes \cdots \otimes \Omega_{nr,j} \rho \Omega_{1j}^\dagger \otimes \Omega_{2j}^\dagger \otimes \cdots \otimes \Omega_{nr,j}^\dagger,$$

for some operators  $\Omega_{jl}$ . Operators of the form  $\Phi_{sep}$  are called *separable* operators. The set of separable operators contains *properly* (cf. [22]) the set of LOCC operators. There is no explicit characterization of LOCC's whose composition can be very complicated as they allow exchange of classical information in various ways as well as for different types of actions following classical communication. Therefore the set of separable operations, being a super-set of the set of LOCC, is useful in characterizing several properties of LOCC's and, in particular, to study entanglement measures. Notice that the actual state is obtained after *normalization* once an operation is applied.

It should be mentioned that there is no total agreement on the requirements that an entanglement measure should satisfy. Several variations of the above axioms are of interest in various situations [101]. For example, the convexity axiom (8.14) is sometimes considered unnecessary (cf. [163]). Requirement 2. is often replaced by the weaker requirement that a measure of entanglement does not increase *on average* under LOCC, namely

$$\sum p_j E\left(\frac{\Phi_j(\rho)}{\text{Tr}(\Phi_j(\rho))}\right) \leq E(\rho),$$

where  $p_j$  is the probability for the operation  $\Phi_j$  to occur. We shall, however, refer to the above three axioms in the following.

The fact that entanglement measures should not increase under LOCC operations implies that if a state  $\rho$  is such that all the states can be obtained from it using LOCC operations then any measure attains a global maximum at  $\rho$ , that is  $\rho$  is *maximally entangled*. For example, it is possible to prove (see, e.g., [163]) that for a bipartite system  $S + P$  with  $n_S = n_P := d$ , every state (pure or mixed) can be obtained using *LOCC* starting from the *generalized singlet state*

$$|\psi_d\rangle := \frac{|00\rangle + |11\rangle + \cdots + |d-1, d-1\rangle}{\sqrt{d}},$$

which is therefore maximally entangled. Here the orthogonal bases for the two  $d$ -dimensional systems  $S$  and  $P$  are chosen coinciding and given by  $|0\rangle, |1\rangle, \dots, |d-1\rangle$ . It follows that any measure of entanglement must have its maximum value (which is often normalized to 1) at  $|\psi_d\rangle$ . For other cases the situation is more complicated and more or less entangled is a partial order in that there are states that are not comparable.

Local unitary (LU) operations act on single subsystem, i.e., locally, according to  $\rho_S \rightarrow X\rho_S X^\dagger$ , with  $X$  unitary. If  $\Phi$  is LU then  $\Phi$  is LOCC and invertible and such that  $\Phi^{-1}$  is also local unitary and therefore LOCC. Therefore, from the axiom (8.13), we obtain

$$E(\rho) \geq E(\Phi(\rho)) \geq E(\Phi^{-1} \circ \Phi(\rho)) = E(\rho).$$

Therefore any measure of entanglement  $E$  must be constant under LU operations.

### 8.1.3.2 Entanglement measures for bipartite pure states

For bipartite pure states the Von Neumann entropy of the partial trace, i.e.,  $E(\rho) = S(\text{Tr}_P(\rho))$  is the most widely used measure of entanglement. Notice that it follows directly from Schmidt decomposition Lemma 8.1.6 that, for a pure state  $\rho$ ,  $\text{Tr}_P(\rho)$  and  $\text{Tr}_S(\rho)$  have the same spectrum. Therefore  $S(\text{Tr}_P(\rho)) = S(\text{Tr}_S(\rho))$ . It is possible to show [71], [163] that the Von Neumann entropy is the only measure of entanglement which satisfies two

additional requirements. These are *additivity* on pure states and *asymptotic continuity*. Additivity for a general entanglement measure  $E$  means that  $E(\rho^{\otimes n}) = nE(\rho)$ , for any state  $\rho$  and any positive integer  $n$ , where  $\rho^{\otimes n}$  is the tensor product of  $n$  copies of  $\rho$ . Asymptotic continuity on pure states means that for every pair of sequences of states  $\{|\psi_n\rangle, |\phi_n\rangle\}$  on a sequence of pairs of coinciding Hilbert spaces  $\{\mathcal{H}_n, \mathcal{H}_n\}$ , if  $\lim_{n \rightarrow \infty} |||\psi_n\rangle\langle\psi_n| - |\phi_n\rangle\langle\phi_n||_1 = 0$ , then

$$\lim_{n \rightarrow \infty} \frac{E(|\psi_n\rangle) - E(|\phi_n\rangle)}{1 + \log(\dim(\mathcal{H}_n))} = 0.$$

### 8.1.3.3 Convex roof extensions to bipartite mixed states; Entanglement of formation

For mixed bipartite states there are two important entanglement measures known as *distillable entanglement* and *entanglement cost* (see [163] for their definitions). These measures are of the operational type, namely they are defined in terms of the amount of operations and resources needed for a given task. Other measures for mixed states can be obtained by **convex roof extension** of measures defined on pure states. This means that if  $E_p$  is a measure defined on pure states, then a measure on general mixed states  $\rho$ ,  $E(\rho)$  is defined as

$$E(\rho) = \inf \sum_j p_j E_p(|\psi_j\rangle), \quad \sum_j p_j = 1, \quad p_j \geq 0,$$

where the infimum is taken among all the possible sets of  $\{p_j, |\psi_j\rangle\}$  such that  $\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|$ . If  $E_p$  is continuous it is possible to prove [209] that the infimum is in fact a minimum, namely it is reached for a given ensemble  $\{p_j, |\psi_j\rangle\}$ . Convex roofs are the largest functions  $E$  which are convex and give  $E(\rho) = E_p(|\psi\rangle)$  if  $\rho := |\psi\rangle\langle\psi|$  is a pure state. The **entanglement of formation**  $E_f$  is defined as the convex roof associated with the Von Neumann entropy of the partial trace. It is therefore the minimum average entropy of the partial trace among all the possible ensemble realizations of a given state  $\rho$ . Its relevance as a measure of entanglement derives from its relation (in some cases proved to be true and in some other cases conjectured, cf. [163]) with operational measures such as the entanglement cost, as well as its own operational meaning [21], [224]. If we have a method to calculate the entanglement of formation we also have a method, alternative to the entanglement tests, to check whether a state is entangled. In fact, it is easily proven that  $\rho$  is separable if and only if  $E_f(\rho) = 0$ .<sup>5</sup> There have been several analytical results and numerical methods towards the solution of the optimization problem in the definition of  $E_f$ . In particular, a result due to Uhlmann [208] says that in the infimization problem it is sufficient to consider ensembles with at most

---

<sup>5</sup>This is a simple consequence of the continuity of the entropy and the fact that the entropy of the partial trace is zero if and only if the state is separable.

$r^2$  terms where  $r$  is the rank of  $\rho$ . Estimates have also been found (see, e.g., [32]). However there is no general method for the explicit calculation.

#### 8.1.3.4 Concurrence

The **concurrence**  $C := C(|\psi\rangle)$  is a function of the state which was first defined [98], [223] as instrumental to the calculation of entropy of formation. It has then become of interest as a measure of entanglement on its own. Such a quantity has a standard definition only for the case of a pair of qubits and different types of generalizations exist for other classes of systems. We shall consider this next.

The concurrence  $C(|\psi\rangle)$  is defined in terms of a specific type of symmetry, the time reversal symmetry, which is a  $\Theta_{II}$  type of symmetry with  $T = \mathbf{1}$  (cf. [Table 5.1](#) of Chapter 5), on a single two level system. In particular, we have the following definition

$$C(|\psi\rangle) := |\langle\psi|\Theta_{II} \otimes \Theta_{II}|\psi\rangle|. \quad (8.15)$$

If  $|\psi\rangle$  is a product state  $|\psi\rangle := |\psi_1\rangle \otimes |\psi_2\rangle$ , we have

$$C(|\psi\rangle) = |\langle\psi_1|\Theta_{II}|\psi_1\rangle\langle\psi_2|\Theta_{II}|\psi_2\rangle| = 0.$$

In fact, if we express  $|\psi_j\rangle$ ,  $j = 1, 2$  as a unit vector  $\vec{\psi}_j := \begin{pmatrix} x_j \\ y_j \end{pmatrix}$ , then, with  $J$  defined in (3.18),

$$\Theta_{II}\vec{\psi}_j := J \begin{pmatrix} \bar{x}_j \\ \bar{y}_j \end{pmatrix} = \begin{pmatrix} -\bar{y}_j \\ \bar{x}_j \end{pmatrix}. \quad (8.16)$$

Therefore

$$\vec{\psi}_j^* \vec{\psi}_j = -\bar{x}_j \bar{y}_j + \bar{y}_j \bar{x}_j = 0.$$

In general,  $0 \leq C \leq 1$ . There exists a simple functional relation between the Von Neumann entropy of the partial trace  $S \circ \text{Tr}_P$  and the concurrence  $C$ . It is given by (cf. [Exercise 8.4](#))

$$S(\text{Tr}_P(|\psi\rangle\langle\psi|)) = H(C(|\psi\rangle)) := h\left(\frac{1 + \sqrt{1 - C^2}(|\psi\rangle)}{2}\right), \quad (8.17)$$

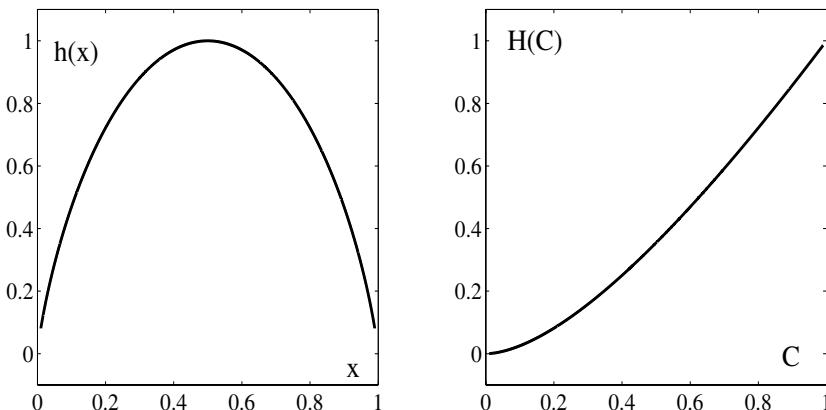
where  $h = h(x)$  is the function<sup>6</sup>

$$h(x) := -x \log(x) - (1 - x) \log(1 - x),$$

where the logarithm is considered base 2 and we have set  $0 \log(0) = 0$ .

---

<sup>6</sup>It is used in the proof of several properties of the concurrence mentioned below that, as a function of  $C$ ,  $H(C) := h(\frac{1+\sqrt{1-C^2}}{2})$  is convex (cf. [Figure 8.2](#)).



**FIGURE 8.2:** The function  $h(x) := -x \log(x) - (1-x) \log(1-x)$  and the function  $H(C) := h(\frac{1+\sqrt{1-C^2}}{2})$  which relate the entanglement of formation to the concurrence.

The entropy of the partial trace is a monotonically increasing function of the concurrence which has its minimum (maximum) at zero (one). Therefore the entropy and the concurrence agree on which state is more or less entangled and one can take the concurrence itself rather than the entropy as a measure of entanglement.

For the mixed state of two qubits, the concurrence is defined by convex roof extension of the concurrence for pure states, i.e., we have

$$C(\rho) := \inf \sum_k p_k C(|\psi_k\rangle),$$

with the infimum taken over all the possible ensembles  $\{p_k, |\psi_k\rangle\}$  with  $\rho = \sum_k p_k |\psi_k\rangle\langle\psi_k|$ ,  $p_k \geq 0$ ,  $\sum_k p_k = 1$ . It is a remarkable result [223] that (cf. (8.17))

$$E_f(\rho) = H(C(\rho)) := h\left(\frac{1 + \sqrt{1 - C^2(\rho)}}{2}\right),$$

and therefore knowing the concurrence of a given state is equivalent to knowing the entanglement of formation. Moreover, there exists an explicit formula for the concurrence of a state  $\rho$  given by [223]

$$C(\rho) = \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\}, \quad (8.18)$$

where  $\lambda_{1,2,3,4}$  are, in descending order, the square roots of the eigenvalues of the matrix  $\rho\tilde{\theta}_{II} \otimes \tilde{\theta}_{II}(\rho)$  where  $\tilde{\theta}_{II}$  is the quantum symmetry associated with  $\Theta_{II}$  (cf. Table 5.1) (cf. Exercise 8.5).

### 8.1.3.5 Generalized concurrences

All we have said from the definition of concurrence on concerns the case of a system of two qubits. The definition of concurrence (8.15) can be formally extended by replacing the specific quantum symmetry  $\Theta_{II} \otimes \Theta_{II}$  with a general quantum symmetry  $\Theta$  acting on the overall Hilbert space of the total (possibly multipartite) system. One defines a  $\Theta$ -concurrence  $C_\Theta(|\phi\rangle)$  as

$$C_\Theta(|\psi\rangle) := |\langle\psi|\Theta|\psi\rangle|, \quad (8.19)$$

and the  $\Theta$ -concurrence of a mixed state,  $C_\Theta(\rho)$  is defined as the convex roof of  $C_\Theta(|\psi\rangle)$ , i.e., as

$$C_\Theta(\rho) = \inf \sum_k p_k C_\Theta(|\psi_k\rangle), \quad (8.20)$$

with the infimum taken over all the possible pure state decompositions of  $\rho$ . The following property proved in [210] directly generalizes (8.18) to the case of a general antiunitary Cartan involutory symmetry  $\Theta$ , with  $\Theta^2 = \mathbf{1}$ ,

$$C_\Theta(\rho) = \max\{0, \lambda_1 - \sum_{j>1} \lambda_j\},$$

where  $\lambda_j$  are the square roots in descending order of the eigenvalues of  $\rho\tilde{\theta}(\rho)$ , with  $\tilde{\theta}$  the quantum symmetry associated with  $\Theta$ . The question arises whether such generalized concurrences are of interest in the quantification of entanglement. We consider this next.

Dealing with multipartite systems it is natural to consider quantum symmetries  $\Theta$  which are tensor products of quantum symmetries on the single subsystems. Moreover, as we want  $\tilde{\theta}^2 = \mathbf{1}$  where  $\tilde{\theta}(\rho) := \Theta\rho\Theta^{-1}$ , we require  $\Theta^2 = \pm\mathbf{1}$ . Choosing

$$\Theta := \Theta_{II} \otimes \Theta_{II} \otimes \cdots \otimes \Theta_{II}, \quad (8.21)$$

with  $N$  factors for a system of  $N$ -qubits, where  $\Theta_{II}$  is as from [Table 5.1](#) with  $T = \mathbf{1}$ , we obtain the  $N$ -concurrence on a system of  $N$  qubits, defined in [222], which was proved in [222] to be a viable measure of entanglement for pure states, i.e., a measure satisfying the axioms 1-3 above. This measure is identically zero for odd values of  $N$ . However, for  $N$  even, it gives another meaningful measure of entanglement which is called the  $N$ -tangle (cf. [222] and references therein). In the case of a bipartite system  $S + P$ , with  $n_S = 2$  and  $n_P \geq 2$  and even, a formula exists relating the entanglement of formation to the generalized concurrences  $C_\Theta$ . This formula [210] is (cf. (8.17))

$$E_f(\rho) \geq \sup_\Theta h \left( \frac{1 + \sqrt{1 - C_\Theta^2(\rho)}}{2} \right), \quad (8.22)$$

where the sup is taken over the set of all antiunitary quantum symmetries of the form<sup>7</sup>  $\Theta := \Theta_S \otimes \Theta_P$ , with  $\Theta_S^2 := -\mathbf{1}$  and  $\Theta_P^2 := -\mathbf{1}$ . This means in particular that, in this case, if there exists a  $\Theta$  such that  $C_\Theta(\rho) > 0$ , then the state  $\rho$  is entangled. This gives an alternate test of entanglement. The ‘sup’ in formula (8.22) is, in general, a function of the state  $\rho$  under consideration. This raises the question whether a certain number of generalized concurrences would be sufficient to capture the entanglement of even more general systems than the bipartite  $S + P$ ,  $n_S = 2$ ,  $n_P$  even. In particular one can consider a vector of concurrences. This is the approach used in [17] where the operators  $\Theta$  are replaced by more general operators  $\Theta_\alpha$  for some parameters  $\alpha$ . The operator  $\Theta_\alpha$  for a bipartite system on  $\mathcal{H}_S \oplus \mathcal{H}_P$  map to zero all the vectors in  $\mathcal{H}_S \oplus \mathcal{H}_P$  except the ones which are linear combinations of tensor products of vectors in two 2-dimensional subspaces (parametrized by  $\alpha$ ). On these vectors  $\Theta_\alpha$  acts as  $\Theta_{II} \otimes \Theta_{II}$ . The authors of [17] use the generalized concurrences associated with  $\Theta_\alpha$  to formulate (and conjecture) several separability conditions for bipartite quantum systems.

Another generalization of the concurrence for general (not necessarily two qubits) bipartite pure states is [183]

$$C(|\psi\rangle) := \sqrt{2(1 - \text{Tr}[(\text{Tr}_P(|\psi\rangle\langle\psi|))^2])},$$

which can then be extended by convex roof to mixed states. Estimates for this concurrence on mixed states have been obtained in [32], [46].

## 8.2 Dynamics of Entanglement

Cartan decompositions of the type introduced in [Chapter 5](#) can be used in the analysis of entanglement dynamics in two ways. First, they relate to generalized  $\Theta$ -concurrences as defined in the previous subsection, in that they factorize unitary evolutions in a term that modifies concurrence and a term that does not. Secondly, they give expressions of every unitary evolution in terms of local and entangling evolutions. These expressions are useful in the design of evolutions performing a given task. In fact, they allow us to identify exactly the entangling evolutions which are the ones which are usually more difficult to perform in a laboratory. The most complete example of treatment of entanglement dynamics using Lie group decomposition is the two qubits system which we shall describe in the next subsection. Several aspects of the treatment extend to general multipartite systems of arbitrary dimensions, as

<sup>7</sup>Notice these antiunitary quantum symmetries can be all parametrized as they are all of the form  $\Theta_{II}$  of [Table 5.1](#).

we shall see. In [section 8.2.3](#) we shall describe several decompositions of unitary evolutions which recursively use the fundamental Cartan decomposition described in [section 5.3](#) to give decompositions in entangling and local parts for every evolution.

Consider a  $\Theta$ -concurrence  $C_\Theta$  as defined in (8.19), (8.20), the Cartan symmetry  $\tilde{\theta}$  associated with  $\Theta$ , the associated Cartan involution  $\theta$  of  $u(\bar{n})$ <sup>8</sup>, and the associated Cartan decomposition of  $u(\bar{n})$

$$u(\bar{n}) = \mathcal{K} \oplus \mathcal{P},$$

with  $\mathcal{K}$  and  $\mathcal{P}$  the +1 and -1 eigenspaces, respectively, of  $u(\bar{n})$ . Cartan decompositions and generalized concurrences are related via the following result.

**Proposition 8.2.1** If  $K \in e^{\mathcal{K}}$ , then, for every  $|\psi\rangle$ ,

$$C_\Theta(K|\psi\rangle) = C_\Theta(|\psi\rangle), \quad (8.23)$$

and

$$C_\Theta(K\rho K^\dagger) = C_\Theta(\rho). \quad (8.24)$$

*Proof.* Formula (8.24) follows from (8.23) (see [Exercise 8.6](#)). To show (8.23), notice that since the Lie group  $e^{\mathcal{K}}$  is always compact, the exponential map is always surjective (see, e.g., [184]) and we can write every element in  $e^{\mathcal{K}}$  as  $K = e^A$ , with  $A \in \mathcal{K}$ . To show that  $C_\Theta(e^A|\psi\rangle) = C_\Theta(|\psi\rangle)$ , we show that (cf. (8.19))

$$e^{-A}\Theta e^A = \Theta,$$

for every  $A \in \mathcal{K}$ . Equivalently we show

$$\Theta e^A \Theta^{-1} = e^A. \quad (8.25)$$

Since  $\Theta$  is antiunitary, from the general expression (5.38)

$$\Theta = L \circ \text{Conj},$$

where  $L$  denotes a unitary matrix and  $\text{Conj}$  denotes conjugation of the components of a vector once a basis is fixed, i.e.,  $\text{Conj}(|\psi\rangle) := |\bar{\psi}\rangle$ .  $\Theta^{-1}$  is given by  $\Theta^{-1} = \text{Conj} \circ L^\dagger$ . A straightforward calculation using these definitions gives

$$\Theta e^A \Theta^{-1} = L e^{\bar{A}} L^\dagger = \sum_{k=0}^{\infty} \frac{L \bar{A}^k L^\dagger}{k!}. \quad (8.26)$$

Write  $A$ , skew-Hermitian, as  $A = iB$ , with  $B$  Hermitian in the space  $i\mathcal{K}$ . We have

$$L \bar{A}^k L^\dagger = (-i)^k L \bar{B}^k L^\dagger = (-i)^k \Theta B^k \Theta^{-1} = (-i)^k \tilde{\theta}(B^k) = (-i)^k (-1)^k B^k = A^k,$$

---

<sup>8</sup> $\bar{n}$  denotes the dimension of the overall system, namely, as we are dealing with multipartite systems,  $\bar{n} = n_1 \times n_2 \times \cdots \times n_N$ , where  $n_j$ ,  $j = 1, \dots, r$  is the dimension of the  $j$ -th subsystem.

where we have used the fact that  $B$  is in the  $-1$  eigenspace of  $\tilde{\theta}$  and the definition of  $B$ . Plugging this into (8.26), we obtain (8.25).  $\square$

In conclusion, in the factorization (5.29) (or (5.31)) only the factor  $P$  (or the factor  $A$ ) possibly changes the generalized concurrence.

### 8.2.1 The two qubits example

A special case of the above situation is for the two qubits concurrence (8.15) for which  $\Theta := \Theta_{II} \otimes \Theta_{II}$  is defined in (8.16). The associated decomposition is the decomposition of  $u(4)$  described in example 5.3.10,<sup>9</sup> i.e.,

$$u(4) = \mathcal{K} \oplus \mathcal{P}, \quad (8.27)$$

with

$$\mathcal{K} := \text{span}\{i\sigma \otimes \mathbf{1}, i\mathbf{1} \otimes \sigma\}, \quad \mathcal{P} := \text{span}\{i\sigma \otimes \sigma, i\mathbf{1} \otimes \mathbf{1}\}, \quad (8.28)$$

where  $\sigma$  is a generic Pauli matrix (1.20). If we choose as Cartan subalgebra in  $\mathcal{P}$ ,

$$\mathcal{A} := \text{span}\{i\sigma_x \otimes \sigma_x, i\sigma_y \otimes \sigma_y, i\sigma_z \otimes \sigma_z, i\mathbf{1} \otimes \mathbf{1}\},$$

then every unitary matrix  $X_f$  can be written as

$$X_f := e^{i\phi} L_1 \otimes L_2 e^{ic_x \sigma_x \otimes \sigma_x} e^{ic_y \sigma_y \otimes \sigma_y} e^{ic_z \sigma_z \otimes \sigma_z} L_3 \otimes L_4, \quad (8.29)$$

with real parameters  $c_x, c_y, c_z, \phi$ , and matrices  $L_1, L_2, L_3, L_4 \in SU(2)$ . According to Proposition 8.2.1, the factors  $L_1 \otimes L_2$  and  $L_3 \otimes L_4$  do not modify the concurrence. Clearly the phase factor  $e^{i\phi}$  does not modify the concurrence either. The only terms that create entanglement in (8.29) are the terms  $e^{ic_j \sigma_j \otimes \sigma_j}$ ,  $j = x, y, z$ . Therefore the *entanglement capability* of a unitary transformation only depends on these terms.

Two unitary evolutions  $X_1$  and  $X_2$  on a system of two qubits are called **locally equivalent** if there exist  $K_1$  and  $K_2$ ,  $K_1, K_2 \in SU(2) \otimes SU(2)$ , such that

$$X_1 = K_1 X_2 K_2.$$

This means that it is possible to obtain the evolution  $X_1$  by having a local evolution  $K_2$  followed by the evolution  $X_2$  then followed by another local evolution  $K_1$ . It is clear that local equivalence is an equivalence relation and two unitary transformations with the same values of  $c_x, c_y$  and  $c_z$  in (8.29) are locally equivalent. A *local invariant* on  $SU(4)$  is a function  $f : SU(4) \rightarrow \mathbb{C}$  such that, if  $X_1$  and  $X_2$  are locally equivalent,  $f(X_1) = f(X_2)$ . A *complete set of local invariants* is a set of local invariants  $f_1, f_2, \dots, f_m$ , such

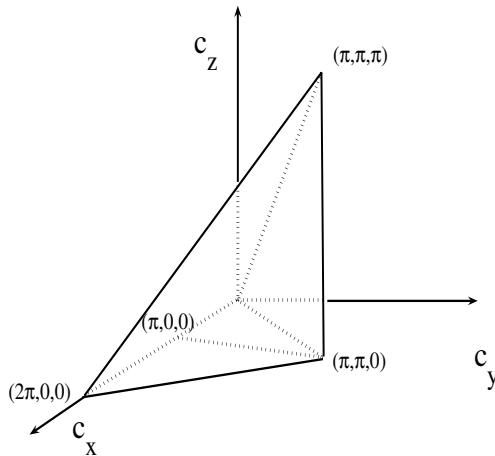
---

<sup>9</sup>This decomposition was also used in the time optimal control problem for two spin- $\frac{1}{2}$  particles in subsection 6.4.2.

that two matrices in  $SU(4)$ ,  $X_1$  and  $X_2$ , are locally equivalent if and only if  $f_j(X_1) = f_j(X_2)$ ,  $j = 1, \dots, m$ .<sup>10</sup> According to a result of [142], [227], once the decomposition of an element in  $SU(4)$  is given (i.e., (8.29), with  $\phi = 0$ ), a complete set of invariants is given by the spectrum of  $A^2$  where  $A$  is defined as (cf. (8.29))

$$A := e^{ic_x\sigma_x\otimes\sigma_x}e^{ic_y\sigma_y\otimes\sigma_y}e^{ic_z\sigma_z\otimes\sigma_z}. \quad (8.30)$$

In fact, it is possible to parametrize the set of locally equivalent transformations in  $SU(4)$  according to the values of the parameters  $c_x$ ,  $c_y$  and  $c_z$  [227] as illustrated in Figure 8.3



**FIGURE 8.3:** According to a result of [227], there exists a one to one correspondence between the points of this tetrahedron and the equivalence classes of locally equivalent transformations in  $SU(4)$ . This correspondence is such that the point  $(c_x, c_y, c_z)$  corresponds to the equivalence class containing  $e^{ic_x\sigma_x\otimes\sigma_x}e^{ic_y\sigma_y\otimes\sigma_y}e^{ic_z\sigma_z\otimes\sigma_z}$ . The only exception concerns points on the base where for each point in the triangle with vertices  $(2\pi, 0, 0)$ ,  $(\pi, 0, 0)$ ,  $(\pi, \pi, 0)$ , there corresponds an equivalent point in the triangle with vertices  $(\pi, 0, 0)$ ,  $(0, 0, 0)$ ,  $(\pi, \pi, 0)$ .

For general elements of  $U(4)$  ( $\phi$  possibly  $\neq 0$  in (8.29)) local equivalence is of interest up to the common phase factor. To check this type of local equivalence for two matrices in  $U(4)$ , one writes the two matrices in the form

---

<sup>10</sup>Similar definitions will be given in the next section for invariants on the set of density matrices.

(8.29) where  $e^{i\phi} = \det(X_f)$  and then checks the local equivalence of the factors in  $SU(4)$ , according to the mentioned results of [142], [227].

**Definition 8.2.2** A unitary evolution  $X_f \in U(4)$  is a **perfect entangler** if there exists a separable pure state  $|\psi\rangle$ , i.e.,  $C(|\psi\rangle) = 0$ , such that  $C(X_f|\psi\rangle) = 1$ , where  $C$  is the concurrence defined in (8.15).<sup>11</sup>

It is clear that with the expression of  $X_f$  in (8.29) and  $A$  in (8.30),  $X_f$  is a perfect entangler if and only if  $A$  is a perfect entangler. The entangling capabilities of  $X_f$  are the same as those of  $A$ . It is possible to characterize perfect entanglers exactly. To do this we need the following definition.

**Definition 8.2.3** Given  $r$  complex numbers  $\lambda_1, \dots, \lambda_r$ , the *convex hull*  $\mathcal{C}$  is the set

$$\mathcal{C} := \left\{ \sum_{j=1}^r c_j \lambda_j \middle| c_j \geq 0, \quad j = 1, \dots, r, \quad \sum_{j=1}^r c_j = 1 \right\}.$$

We notice that all the eigenvalues of  $A^2$  are on the unit circle in the complex plane. We have the following characterization of perfect entanglers in terms of the eigenvalues of  $A^2$ .

**Proposition 8.2.4** [142] [227]  $A$  in (8.30) is a perfect entangler if and only if the convex hull of the eigenvalues of  $A^2$  contains the origin.

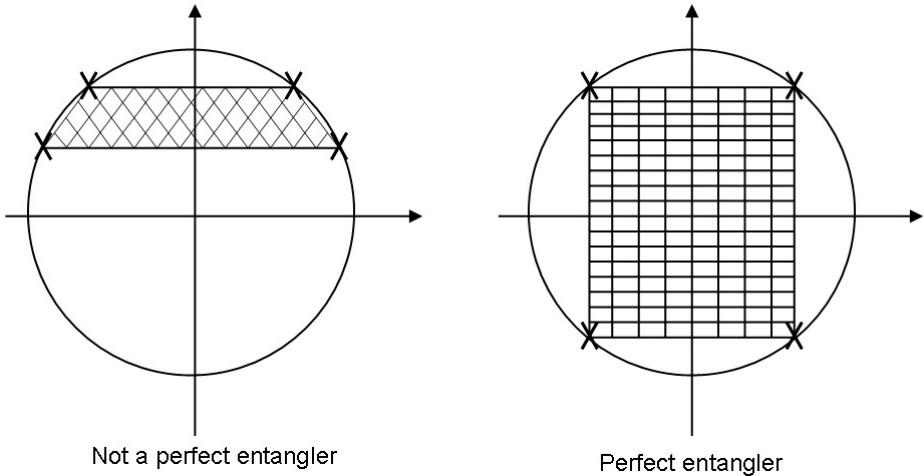
Another consequence of the decomposition (8.29) is that only one entangling Hamiltonian (say  $\sigma_z \otimes \sigma_z$ ), along with local transformations, is sufficient to implement every unitary transformation. In fact, every Hamiltonian  $\sigma_j \otimes \sigma_k$ ,  $j, k = x, y, z$  (and therefore in particular  $\sigma_x \otimes \sigma_x$  and  $\sigma_y \otimes \sigma_y$ ) is similar, via a local similarity transformation, to  $\sigma_z \otimes \sigma_z$  (a consequence of the similarity among Pauli matrices).

Notice that not all the nonlocal transformations  $X_f$  are entanglers in the sense that there exists a separable state  $|\psi\rangle$  with  $C(X_f|\psi\rangle) > 0$ . For example every evolution locally equivalent to the *SWAP operator*  $X_S$ , which is defined as  $X_S(|\psi_1\rangle \otimes |\psi_2\rangle) := |\psi_2\rangle \otimes |\psi_1\rangle$ , has this property.

In conclusion, every transformation on the system of two qubits can be decomposed, according to Cartan decomposition, into local factors and factors that are responsible for the entanglement of the two qubits. The local factors do not modify the concurrence but the entangling factor possibly does. The analysis of the entanglement capability of a given unitary transformation can be performed by considering its entangling part only. In particular, studying the eigenvalues of this factor, it is possible to deduce whether or not the overall evolution is a perfect entangler.

---

<sup>11</sup>The definition could have been given equivalently in terms of the entropy of the partial trace.



**FIGURE 8.4:** Illustration of the characterization of perfect entanglers of Proposition 8.2.4. Given that the eigenvalues of  $A^2$  are represented by the symbol  $\times$  on the unit sphere, the figure on the left represents an evolution which is **not** a perfect entangler. The figure on the right represents an evolution which is a perfect entangler.

### 8.2.2 The odd-even decomposition and concurrence dynamics

The Cartan decomposition of  $u(4)$  (8.27) (8.28) is an example of a general procedure for a multipartite system to generate a decomposition of the total system in terms of decompositions on the single subsystems [61]. We shall describe this next.

Consider a multipartite quantum system composed of  $N$  subsystems of dimensions  $n_1, \dots, n_N$ . The Jordan algebra of the Hamiltonians on the  $j$ -th subsystem is  $iu(n_j)$ ,  $j = 1, \dots, N$ . Assume we perform a Cartan decomposition on each  $u(n_j)$  of type **AI** or **AI<sup>12</sup>**,  $j = 1, \dots, N$  and write

$$u(n_j) := \mathcal{K}_j \oplus \mathcal{P}_j, \quad (8.31)$$

with  $\mathcal{K}_j$  a subalgebra (isomorphic to  $so(n_j)$  or  $sp(\frac{n_j}{2})$ ) and  $\mathcal{P}_j$  its orthogonal complement. To the decomposition (8.31) there corresponds a decomposition of the Jordan algebra  $iu(n_j)$ ,

$$iu(n_j) = i\mathcal{K}_j \oplus i\mathcal{P}_j,$$

a quantum mechanical symmetry  $\Theta_j$  ( $\tilde{\theta}_j$ ) and the Cartan involution  $\theta_j$  (cf. section 5.3.5). Consider a (orthonormal) basis of  $i\mathcal{K}_j$  and denote its generic

---

<sup>12</sup>Recall that  $\text{span}\{i\mathbf{1}\}$  is included in  $\mathcal{P}_j$  in the decomposition (8.31) (cf. Tables 5.1, 5.2 in Chapter 5).

element by  $\sigma$  and a (orthonormal) basis of  $i\mathcal{P}_j$  and denote its generic element by  $S$ . A (orthonormal) basis of  $iu(n_1 \times n_2 \times \cdots \times n_N)$  is given by elements of the form

$$F := F_1 \otimes F_2 \otimes \cdots \otimes F_N, \quad (8.32)$$

with  $F_j$ ,  $j = 1, \dots, N$ , of the type  $\sigma$  or  $S$ . Denote now by  $\mathcal{I}_o$  the span of elements of the form  $F$  in (8.32) with an *odd* number of  $\sigma$ 's and by  $\mathcal{I}_e$  the span of elements of the form  $F$  in (8.32) with an *even* number of  $\sigma$ 's. Construct a quantum symmetry  $\Theta_{o,e}$ , as tensor product of all the symmetries  $\Theta_j$ ,  $j = 1, \dots, N$ ,

$$\Theta_{o,e} := \Theta_1 \otimes \Theta_2 \otimes \cdots \otimes \Theta_N,$$

and the associated symmetry on observables  $\tilde{\theta}_{o,e}$ ,

$$\tilde{\theta}_{o,e} := \tilde{\theta}_1 \otimes \tilde{\theta}_2 \otimes \cdots \otimes \tilde{\theta}_N,$$

where, for all  $A \in iu(n_j)$ ,  $\tilde{\theta}_j(A) = \Theta_j A \Theta_j^{-1}$ .<sup>13</sup> Since  $\tilde{\theta}_j(\sigma) = -\sigma$  and  $\tilde{\theta}_j(S) = S$   $\tilde{\theta}_{o,e}(A) = A$  if  $A \in \mathcal{I}_e$  and  $\tilde{\theta}_{o,e}(A) = -A$  if  $A \in \mathcal{I}_o$ . The involution associated with  $\tilde{\theta}_{o,e}$  (cf. (5.40)),  $\theta_{o,e}$ , is such that, for  $A \in iu(n_1 \times n_2 \times \cdots \times n_N)$ ,

$$\theta_{o,e}(iA) = -i\tilde{\theta}_{o,e}(A),$$

and therefore if  $A \in \mathcal{I}_o$ ,  $\theta_{o,e}(iA) = iA$  and if  $A \in \mathcal{I}_e$   $\theta_{o,e}(iA) = -iA$ . Therefore  $\theta_{o,e}$  is a Cartan involution on  $u(n_1 \times n_2 \times \cdots \times n_N)$  with  $i\mathcal{I}_o$  and  $i\mathcal{I}_e$  the +1 and -1 eigenspaces, and

$$u(n_1 \times n_2 \times \cdots \times n_N) := i\mathcal{I}_o \oplus i\mathcal{I}_e \quad (8.33)$$

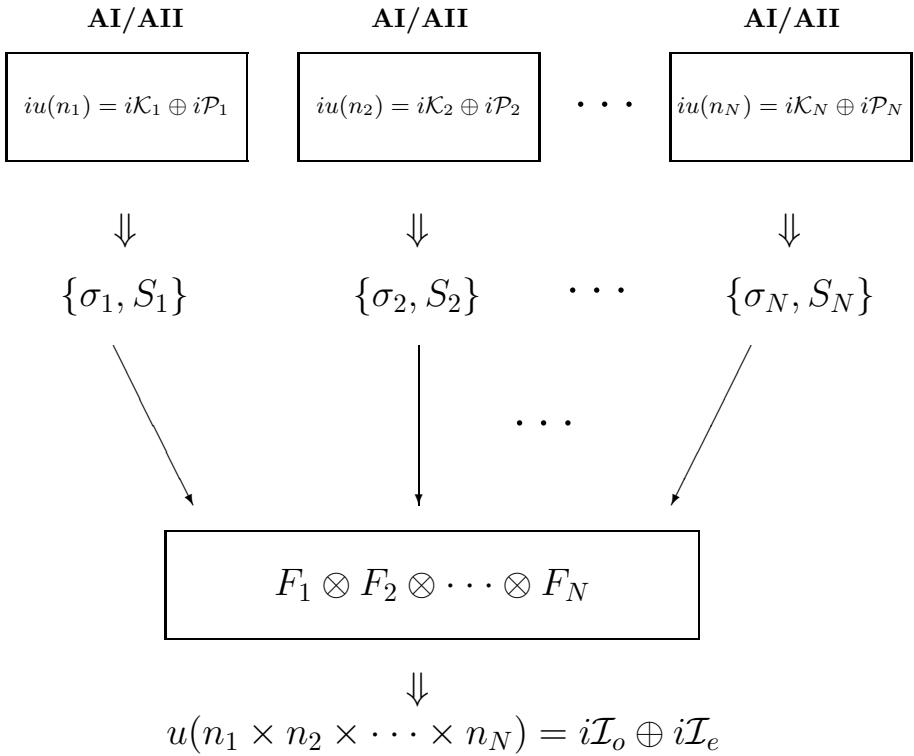
is a Cartan decomposition of  $u(n_1 \times n_2 \times \cdots \times n_N)$  with  $i\mathcal{I}_o$  and  $i\mathcal{I}_e$  being the subalgebra and the complementary orthogonal subspace respectively. This is called the **odd-even decomposition (OED)**. Figure 8.5 summarizes the procedure to obtain this decomposition.

**Remark 8.2.5** The decomposition (8.27) used in subsection 8.2.1 is a special case of the OED. This special case is obtained when  $N = 2$  and both decompositions on  $u(n_1) = u(n_2) = u(2)$  are of type **AII**. In fact  $sp(1) = su(2)$  and the relevant decomposition of  $u(2)$  is  $u(2) = \mathcal{K} \oplus \mathcal{P}$ , with  $\mathcal{K} = su(2) = sp(1)$  and  $\mathcal{P} := \text{span}\{i\mathbf{1}\}$ . In this case, the decomposition separates local and entangling contributions of every unitary evolution.

The interest of the odd-even decomposition in entanglement theory follows from the fact that such a decomposition of the dynamics is naturally obtained

---

<sup>13</sup>The definition of the tensor product of two antilinear operators (on the space of Hermitian matrices) is discussed in subsection 5.3.6, Remark ??, and extends naturally to the case of  $N$  antilinear operators.



**FIGURE 8.5:** Scheme for the OED decomposition. Decompositions of the types **AI** or **AII** are performed for each subsystem. This determines a set of matrices  $\{\sigma_j, S_j\}$ ,  $j = 1, \dots, N$ , which span orthonormal bases for the associated pairs of subspaces  $i\mathcal{K}_j$  and  $i\mathcal{P}_j$ . By taking the tensor products  $F_1 \otimes F_2 \otimes \cdots \otimes F_N$  with  $F_j$  of the form  $\sigma_j$  or  $S_j$  we obtain the basis of  $i\mathcal{U}(n_1 \times n_2 \times \cdots \times n_N)$ . By splitting this basis in the odd and even part, we obtain the OED decomposition.

in terms of tensor products of Hamiltonians. Local and entangling Hamiltonians for general dimensions and decompositions are, however, both in the Lie subalgebra part  $i\mathcal{I}_o$  and in the complement part  $i\mathcal{I}_e$  contrary to what happened for the special case of the previous remark and subsection 8.2.1. For example, the local Hamiltonian  $H_{local} := A \otimes \mathbf{1} \otimes \cdots \otimes \mathbf{1}$  belongs to  $\mathcal{I}_o$  or  $\mathcal{I}_e$  according to whether  $A$  belongs to  $i\mathcal{K}_1$  or not (cf. (8.31)), respectively. Therefore, consider the associated generalized concurrence  $C_{\Theta_{o,e}}$ . Proposition 8.2.1 does not guarantee that  $C_{\Theta_{o,e}}(\lvert\psi\rangle) = C_{\Theta_{o,e}}(L\lvert\psi\rangle)$ , for  $L$  local, a requirement necessary for any measure of entanglement. Nevertheless, we can obtain invariance if we restrict the local operations to fewer, physically motivated ones. We illustrate this in the following Remark.

**Remark 8.2.6** Consider a system of  $N$  particles with spin, with variable values of the spin (which can be integer or half-integer). For every spin system of dimension  $n_j$ , there exists a three-dimensional Lie subalgebra of  $u(n_j)$  isomorphic to  $su(2)$ , which is given by the infinitesimal rotations. Expressions for the matrices representing these operators in an appropriate basis can be calculated explicitly (see, e.g., [185] section 3.5.). Their exponentials represent rotations of the spins and therefore we call this subalgebra here the *rotation Lie subalgebra*. It is possible to show [8] that, if  $n_j$  is even (half integer spin), there exists a subalgebra of  $u(n_j)$ , conjugate to  $sp(\frac{n_j}{2})$  and containing as subalgebra the rotation Lie subalgebra. Analogously, it is possible to show that if  $n_j$  is odd (integer spin) there exists a subalgebra of  $u(n_j)$ , conjugate to  $so(n_j)$  and containing as subalgebra the rotation Lie subalgebra.<sup>14</sup> Let us perform an odd-even decomposition using a decomposition of type **AI** on odd dimensional systems and of type **AII** on even dimensional systems and let us use as Lie algebras  $\mathcal{K}_j$  in (8.31) the ones containing the rotation Lie subalgebra in all cases. Then all the infinitesimal generators of local rotations will be in the  $i\mathcal{I}_o$  part of the decomposition (8.33). Therefore the concurrence  $C_{\Theta_{o,e}}$  will be left invariant by local rotations as from Proposition 8.2.1.

**Remark 8.2.7** Another special case of the OED is for the case of  $N$  qubits when a decomposition of type **AII** is performed on every subsystem. The resulting decomposition is known as **Concurrence Canonical Decomposition (CCD)** [36] [37].<sup>15</sup> In this case, the local unitary evolutions all belong to  $e^{i\mathcal{I}_o}$  since the corresponding Hamiltonians are of the type  $\mathbf{1} \otimes \mathbf{1} \otimes \cdots \otimes \sigma \otimes \mathbf{1} \otimes \mathbf{1} \otimes \cdots \otimes \mathbf{1}$ , with only one  $\sigma$  which is a linear combination of Pauli matrices. The generalized concurrence  $C_{\Theta_{o,e}}$  is the same as the  $N$ -concurrence

<sup>14</sup>In fact the converse of these statements also holds. If  $n_j$  is even there is no Lie subalgebra of  $u(n_j)$  conjugate to  $so(n_j)$  and containing the rotation Lie subalgebra. See [8].

<sup>15</sup>In an independent study [7] the CCD was used to characterize networks of spin- $\frac{1}{2}$  particles which are *input-output equivalent* in the sense that they give the same value of the total average spin when subject to the same (input) magnetic field. The generalization of this result to general spin uses the OED and is provided in [8].

defined after formula (8.21), which is, as we have said, a viable measure of entanglement (which is, however, trivial for odd  $N$ ).

### 8.2.2.1 Computation of the OED decomposition

The OED is a Cartan decomposition of  $u(n_1 \times n_2 \times \cdots \times n_N)$  and therefore it has to be of the type **AI**, **AII** and **AIII** as defined in subsection 5.3.4. It is of interest to provide a change of coordinates to put the decomposition in the standard basis so as to be able to calculate it using the methods in subsection 5.3.6. The odd-even decomposition cannot be of type **AIII** because the associated quantum symmetry  $\Theta_{o,e}$  is the tensor product of antiunitary symmetries and therefore also antiunitary, while for decompositions of the type **AIII** the associated symmetry is unitary. To discover whether the decomposition is of type **AI** or **AII** we can do a simple dimension count. This is done in [61] and the result is the following.

**Proposition 8.2.8** Consider an odd-even decomposition on  $N$  subsystems obtained by performing **AII** decompositions on  $r$  subsystems and **AI** decompositions on  $N - r$  subsystems. Then the resulting decomposition is of type **AII** if  $r$  is odd and of type **AI** if  $r$  is even.

To put the decomposition in the standard basis so as to apply computational techniques of subsection 5.3.6 we have to compute the matrix  $T$  in Table 5.2 so as to make

$$i\mathcal{I}_o = T \text{so}(n_1 \times n_2 \times \cdots \times n_N) T^\dagger,$$

or

$$i\mathcal{I}_o = T \text{sp}\left(\frac{n_1 \times n_2 \times \cdots \times n_N}{2}\right) T^\dagger,$$

according to whether we have an **AI** or **AII** total decomposition. Assume, for simplicity of notation, that all the decompositions are performed in the standard basis. Then we have for all  $A$  in  $u(n_1 \times n_2 \times \cdots \times n_N)$

$$\theta_{o,e}(A) := i\tilde{\theta}_{o,e}(iA) = (\mathbf{1} \otimes J \otimes \mathbf{1} \cdots \otimes J) \bar{A} (\mathbf{1} \otimes J^{-1} \otimes \mathbf{1} \cdots \otimes J^{-1}),$$

where  $\mathbf{1}$  or  $J$ , of appropriate dimensions, appear whether we perform a decomposition of the type **AI** or **AII**, respectively on the corresponding subsystem. In the case where the total decomposition is **AI** (but analogous reasoning is valid for the decomposition **AII**) comparing with Table 5.1, we must have

$$(\mathbf{1} \otimes J \otimes \mathbf{1} \cdots \otimes J) \bar{A} (\mathbf{1} \otimes J^{-1} \otimes \mathbf{1} \cdots \otimes J^{-1}) = TT^T \bar{A}(TT^T)^\dagger.$$

As this equation has to hold for every  $A \in u(n_1 \times n_2 \times \cdots \times n_N)$  we must have<sup>16</sup>

$$TT^T = \mathbf{1} \otimes J \otimes \mathbf{1} \cdots \otimes J,$$

---

<sup>16</sup>This argument uses Schur's lemma (see, e.g., [213]). Since  $(TT^T)^\dagger \mathbf{1} \otimes J \otimes \mathbf{1} \cdots \otimes J$  commutes with every element in  $u(n_1 \times n_2 \times \cdots \times n_N)$  it must be a multiple of the identity.

from which one obtains the unitary  $T$ . In this fashion, one can for example obtain the transformation of coordinates given by the magic basis in Example 5.3.10.

### 8.2.3 Recursive decomposition of dynamics in entangling and local parts

The theorem of Cartan decomposition of section 5.3 can be used recursively to obtain finer decompositions of dynamics in entangling and local parts. There are several proposals in the literature in this direction. We shall summarize some of them in the following.

In [54], M. Dagli, J. D. H. Smith and the author considered the case of  $N$  qubits and applied the CCD recursively as follows. First the CCD is applied to write

$$u(2^N) = i\mathcal{I}_o \oplus i\mathcal{I}_e,$$

where  $i\mathcal{I}_o$  ( $i\mathcal{I}_e$ ) is the span of elements  $i\sigma \otimes \mathbf{1} \otimes \sigma \otimes \cdots \otimes \sigma$  (i.e., tensor products of  $\mathbf{1}$ 's and  $\sigma$ 's), where  $\sigma$  is a general Pauli matrix, with an odd (even) number of elements  $\sigma$ . Therefore a general element  $X$  of  $U(2^n)$  can be written, according to the  $KAK$  decomposition (5.31), as

$$X = K_1 A K_2, \quad (8.34)$$

with  $K_1, K_2 \in e^{i\mathcal{I}_o}$  and  $A \in e^{\mathcal{A}}$  where  $\mathcal{A}$  is the Cartan subalgebra associated with the decomposition. From Proposition 8.2.8, according to whether  $N$  is even or odd,  $i\mathcal{I}_o$  is conjugate to  $so(2^N)$  or  $sp(2^{N-1})$ . In the two cases the Cartan subalgebra<sup>17</sup> has dimension  $2^N$  or  $2^{N-1}$ , respectively. In the first case, a maximal set of commuting elements of  $i\mathcal{I}_e$  can be obtained considering the set

$$\mathcal{S} := \{\sigma_x \otimes \sigma_x, \sigma_y \otimes \sigma_y, \sigma_z \otimes \sigma_z, \mathbf{1} \otimes \mathbf{1}\}. \quad (8.35)$$

Then a Cartan subalgebra is spanned by elements of the form  $iA_1 \otimes A_2 \otimes \cdots \otimes A_{\frac{N}{2}}$ , where every element  $A_j$ ,  $j = 1, \dots, \frac{N}{2}$ , belongs to  $\mathcal{S}$  defined in (8.35). This way one obtains  $4^{\frac{N}{2}} = 2^N$  elements, which is the dimension of the Cartan subalgebra. In the second case, a Cartan subalgebra is spanned by elements of the form  $iA_1 \otimes A_2 \otimes \cdots \otimes A_{\frac{N-1}{2}} \otimes \mathbf{1}$  where every element  $A_j$ ,  $j = 1, \dots, \frac{N-1}{2}$  belongs to  $\mathcal{S}$  defined in (8.35).

To refine the decomposition (8.34) one needs to decompose the elements  $K_1$  and  $K_2$ . This is obtained by performing a Cartan decomposition of  $i\mathcal{I}_o$  as

$$i\mathcal{I}_o := \mathcal{K}_o \oplus \mathcal{P}_o,$$

where

$$\mathcal{K}_o = \text{span}\{A \otimes \mathbf{1}, B \otimes \sigma_z | A \in i\mathcal{I}_o^1, B \in i\mathcal{I}_e^1\}, \quad (8.36)$$

---

<sup>17</sup>We are including also multiples of the identity as we are dealing with decompositions of  $u(2^N)$ .

$$\mathcal{P}_o := \text{span}\{C \otimes \sigma_{x,y} | C \in i\mathcal{I}_e^1\}, \quad (8.37)$$

where  $\mathcal{I}_o^1$  ( $\mathcal{I}_e^1$ ) is the set of tensor products  $\sigma \otimes \mathbf{1} \otimes \sigma \otimes \cdots \otimes \mathbf{1}$  with an odd (even) number of  $\sigma$ 's and  $N - 1$  places.<sup>18</sup> Therefore  $K_1$  (and analogously  $K_2$ ) can be decomposed as

$$K_1 = K_{1,1} \tilde{A} K_{1,2},$$

where  $K_{1,1}$  and  $K_{1,2}$  belong to the Lie group  $e^{\mathcal{K}_o}$  and  $\tilde{A}$  belongs to  $e^{\mathcal{A}_1}$  with  $\mathcal{A}_1$  the Cartan subalgebra ( $\subseteq \mathcal{P}_o$ ) associated with the decomposition. This is a decomposition of  $so(2^N)$ , for  $N$  even or a decomposition of  $sp(2^{N-1})$  for  $N$  odd. The crucial observation to iterate the procedure is to notice that  $\mathcal{K}_o$  is isomorphic to  $u(2^{N-1})$ . In fact the associated decomposition of  $i\mathcal{I}_o$  is of type **DIII** and **CI**, respectively (see [96]) and this gives us the information about the dimensions of the associated Cartan subalgebras which are  $2^{N-2}$  and  $2^{N-1}$ , for  $N$  even or  $N$  odd, respectively. The Cartan subalgebra in the case **DIII** ( $N$  even) is spanned by elements of the type  $iA_1 \otimes A_2 \otimes \cdots \otimes A_{\frac{N-2}{2}} \otimes \mathbf{1} \otimes \sigma_x$ , for  $A_j$ ,  $j = 1, \dots, \frac{N-2}{2}$  in the set  $\mathcal{S}$  defined in (8.35). In the case **CI** ( $N$  odd), the Cartan subalgebra is spanned by elements of the type  $iA_1 \otimes A_2 \otimes \cdots \otimes A_{\frac{N-1}{2}} \otimes \sigma_x$ , where the  $A_j$ 's,  $j = 1, \dots, \frac{N-1}{2}$ , are defined as above.

The next step is to factorize the elements  $K_{1,1}$  and  $K_{1,2}$  which belong to  $e^{\mathcal{K}_o}$ . However  $\mathcal{K}_o$  is the same as  $u(2^{N-1})$  and one performs a decomposition equivalent to the CCD on  $N - 1$  qubits, given by

$$\mathcal{K}_o := \mathcal{K}_o^1 \oplus \mathcal{P}_o^1,$$

where  $\mathcal{K}_o^1 := \text{span}\{A \otimes \mathbf{1} | A \in i\mathcal{I}_o^1\}$  and  $\mathcal{P}_o^1 := \text{span}\{B \otimes \sigma_z | B \in i\mathcal{I}_e^1\}$ . The procedure repeats itself with  $N - 1$  replacing  $N$ , and so on. A scheme is summarized in [Figure 8.6](#). At the end of the recursive procedure a decomposition of the unitary evolution in elementary factors is obtained. This decomposition is in terms of tensor products of matrices. We refer to [54] for details and examples.

Another recursive decomposition for  $N$  qubits was presented by N. Khaneja and S. Glaser in [120]. One starts with  $su(2^N)$  and decompose it as

$$su(2^N) = \mathcal{K} \oplus \mathcal{P},$$

where

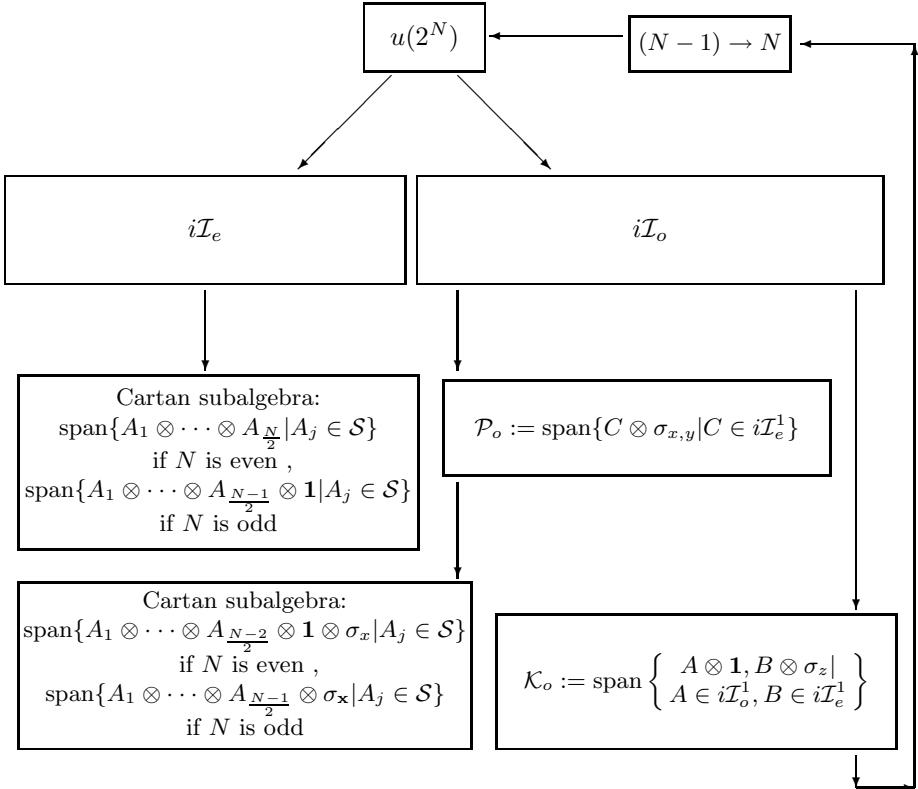
$$\mathcal{K} := \text{span}\{\sigma_z \otimes A, \mathbf{1} \otimes B, i\sigma_z \otimes \mathbf{1} | A, B \in su(2^{N-1})\},$$

$$\mathcal{P} := \text{span}\{\sigma_{x,y} \otimes A | A \in u(2^{N-1})\}.$$

It is easy to verify that this is a Cartan decomposition of  $su(2^N)$ . It is in fact an **AIII** type of decomposition with  $p = q = 2^{N-1}$  (see [subsection 5.3.4](#)

---

<sup>18</sup>Notice we might as well have replaced  $\sigma_z$  with  $\sigma_x$  (or  $\sigma_y$ ) in (8.36) and accordingly  $\sigma_{x,y}$  with  $\sigma_{y,z}$  (or  $\sigma_{x,z}$ ) in (8.37).



**FIGURE 8.6:** Scheme of the recursive decomposition of M. Dagli, J.H.D. Smith and the author [54]. The last iteration is with  $N = 1$ . The set  $\mathcal{S}$  is defined in (8.35) and  $\mathcal{I}_o^1$  and  $\mathcal{I}_e^1$  are the same as  $\mathcal{I}_o$  and  $\mathcal{I}_e$  but with  $N - 1$  factors in the tensor product rather than  $N$ . Also, notice that  $\mathcal{K}_o$  is isomorphic to  $u(2^{N-1})$ .

and formula (5.33)) as the subalgebra  $\mathcal{K}$  is spanned by all the elements of the form  $\begin{pmatrix} C & 0 \\ 0 & D \end{pmatrix}$ ,  $C, D \in u(2^{N-1})$ ,  $\text{Tr}(C) + \text{Tr}(D) = 0$ . Therefore an element  $X_f \in SU(2^N)$  can be decomposed in the form

$$X_f = K_1 A K_2, \quad (8.38)$$

with  $K_1, K_2 \in e^{\mathcal{K}}$ , and  $A \in e^{\mathcal{A}}$  with  $\mathcal{A}$  the Cartan subalgebra in  $\mathcal{P}$ . The dimension of  $\mathcal{A}$  is the rank of the decomposition which, in this case, is  $2^{N-1}$  ( $= \min\{p, q\}$ , cf. subsection 5.3.4). A maximal set of linearly independent commuting matrices in  $\mathcal{P}$  can be obtained as  $\sigma_x \otimes F$ , where  $F$  is in a maximal set of linearly independent commuting matrices in  $su(2^{N-1})$ , for example diagonal matrices.

The next step of the procedure is to factorize the elements  $K_1$  and  $K_2$  in (8.38). To do this one performs a Cartan decomposition of  $\mathcal{K}$ ,

$$\mathcal{K} := \mathcal{K}_1 \oplus \mathcal{P}_1,$$

where

$$\begin{aligned}\mathcal{K}_1 &= \text{span}\{\mathbf{1} \otimes B | B \in su(2^{N-1})\}, \\ \mathcal{P}_1 &:= \text{span}\{\sigma_z \otimes A | A \in u(2^{N-1})\}.\end{aligned}$$

This is again a Cartan decomposition where  $\mathcal{K}_1$  plays the role of the Lie subalgebra. The element  $K_1$  (and analogously  $K_2$ ) in (8.38) can be written as

$$K_1 = K_{1,1} \tilde{A} K_{1,2}.$$

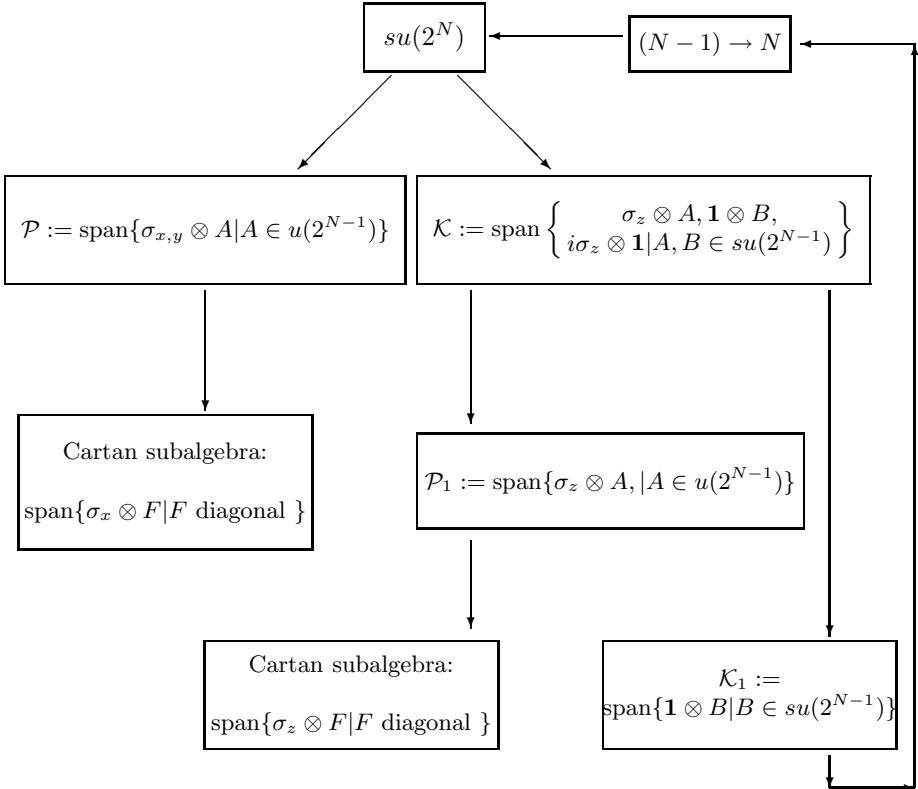
Here  $K_{1,1}$  and  $K_{1,2}$  are in  $e^{\mathcal{K}_1}$  and  $\tilde{A} \in e^{\mathcal{A}_1}$ , with  $\mathcal{A}_1$  a maximal Abelian subalgebra in  $\mathcal{P}_1$ . Since  $\mathcal{P}_1$  is spanned by matrices of the form  $\begin{pmatrix} B & 0 \\ 0 & -B \end{pmatrix}$ , with  $B \in u(2^{N-1})$ , a maximal Abelian subalgebra has dimension  $2^{N-1}$ . It is spanned by matrices of the form  $\sigma_z \otimes F$  where  $F$  is in a maximal commuting set in  $u(2^{N-1})$ , for example diagonal matrices. The key observation, at this point, is that  $\mathcal{K}_1$  is the same as  $su(2^{N-1})$ , and therefore the procedure can be applied recursively starting from the decomposition of  $K_{1,1}$ ,  $K_{1,2}$ . A scheme of the procedure for this decomposition is given in [Figure 8.7](#).

In [120] it is shown how to choose the Cartan subalgebras so that, in the final decomposition, all the nonlocal transformations can be obtained by local transformations and two qubit interactions, which are more easily implemented than multi-body interactions. This fact makes this procedure appealing for the direct applications of decompositions in the generation of given evolutions. Of course one could combine the techniques of [54], [120] and decompose some of the factors obtained with [54] with the technique of [120], or vice versa.

The recursive decompositions in [54], [120] only deal with multipartite systems of qubits. For the case of systems of arbitrary dimensions, a recursive decomposition was given by R. Romano and the author in [64] for a bipartite system. Such a decomposition uses an odd-even decomposition at the first step and recursive decompositions of  $so(n)$  in the following recursive procedure.

### 8.3 Local Equivalence of States

A problem of great interest in entanglement theory is to find methods to determine whether or not two states are **locally equivalent**, which means that it is possible to transfer the state of the total system from one value to



**FIGURE 8.7:** Procedure for the decomposition of  $su(2^N)$  of Khaneja and Glaser [120]. The last iteration is for  $N = 1$ .

the other by only local transformations. Any measure of entanglement will give then the same value on the two states. This problem can be seen as a special case of a general controllability problem (cf. Chapter 3). Given two density matrices  $\rho_1$  and  $\rho_2$  and a (dynamical) Lie algebra  $\mathcal{L}$ , how can we check whether  $\rho_1$  is in the orbit (cf. (3.28))

$$\mathcal{O}_{\mathcal{L}}(\rho_2) := \{X\rho_2 X^\dagger | X \in e^{\mathcal{L}}\} ? \quad (8.39)$$

In the case where  $\mathcal{L}$  is the direct sum of Lie algebras  $su(n_j)$ , where  $n_j$  is the dimension of the  $j$ -th system in a multipartite system, we have the problem of local equivalence of states. This problem has a vast literature (e.g., [9], [11], [89], [136], [142], [167]), and to a large extent is still an open problem.

### 8.3.1 General considerations on dimensions

The space of density matrices for a general multipartite system with dimensions  $n_1, \dots, n_N$  is a manifold with boundary with dimension  $(n_1 n_2 \cdots n_N)^2 - 1$ ,

i.e., the dimension of the space of Hermitian matrices modulo the constraint  $\text{Tr}(\rho) = 1$ . Any orbit (8.39) (with  $\mathcal{L} = \bigoplus_{j=1}^N su(n_j)$ ) is a sub-manifold. The structure of the orbits in the set of density matrices is in general quite complicated although some low dimensional cases have been studied in detail [129]. The dimension of an orbit  $\mathcal{O}_{\mathcal{L}}(\rho_2)$  in (8.39) can be calculated explicitly by noticing (cf. section 3.6) that  $\mathcal{O}_{\mathcal{L}}(\rho_2)$  is diffeomorphic to  $e^{\mathcal{L}}/e^{\mathcal{C}_{\rho_2} \cap \mathcal{L}}$  where  $\mathcal{C}_{\rho_2}$  is the subspace of matrices in  $u(n_1 n_2 \cdots n_N)$  which commute with  $\rho_2$ . The tangent space of  $\mathcal{O}_{\mathcal{L}}(\rho_2)$  at  $\rho_2$  can be identified with  $[i\rho_2, \mathcal{L}]$  and therefore

$$\dim \mathcal{O}_{\mathcal{L}}(\rho_2) = \dim [i\rho_2, \mathcal{L}] = \dim \mathcal{L} - \dim \mathcal{C}_{\rho_2} \cap \mathcal{L}.$$

If we consider the linear operator  $ad_{i\rho_2}$  acting on  $\mathcal{L}$  as  $ad_{i\rho_2}(A) := [i\rho_2, A]$ ,  $\mathcal{C}_{\rho_2} \cap \mathcal{L} = \ker(ad_{i\rho_2})$ . For example, if  $\rho_2 = \frac{1}{n_1 n_2 \cdots n_N} \mathbf{1}$ ,  $\ker(ad_{i\rho_2}) = \mathcal{L}$ , and the number on the right hand side is zero. This is the generic case. In fact, we have the following proposition [136].

**Proposition 8.3.1** Assume the number of subsystems  $N$  is  $\geq 2$ . For almost every density matrix  $\rho_2$ ,  $\ker(ad_{i\rho_2}) = \{0\}$ . Therefore, since  $\mathcal{L}$  is equal to the direct sum of  $su(n_j)$ ,  $j = 1, \dots, N$ , for almost every  $\rho_2$  the dimension of the orbit (8.39) is  $\sum_{j=1}^N n_j^2 - N$ .

We shall not completely prove this result here and refer to [136] which uses a somewhat different language. However, we shall indicate an alternative proof and illustrate this proof for the  $2 \times 2$  case.<sup>19</sup>

We have to prove that, except for a set of measure zero of density matrices, the rank of the linear transformation  $ad_{i\rho_2}$  is equal to  $\dim \mathcal{L}$ . Since the determinant is an analytic function, the set of its zeros is a set of measure zero. Therefore, the rank of the matrix corresponding to  $ad_{i\rho_2}$  is maximum except at a set of measure zero. The proof is obtained if we find a density matrix  $\rho_2$  for which  $\text{rank}(ad_{i\rho_2}) = \sum_{j=1}^N n_j^2 - N$ , which is the maximum possible, given the dimension of the Lie algebra  $\mathcal{L}$ . Consider for example the case of two qubits. Recall the definition of the Pauli matrices  $\sigma_{x,y,z}$  in (1.20). For  $\alpha$  sufficiently small the matrix

$$\rho_2 := \frac{1}{4} \mathbf{1} + \alpha \sigma_z \otimes \mathbf{1} + \alpha \mathbf{1} \otimes \sigma_z + \alpha \sigma_z \otimes \sigma_z$$

is a density matrix, i.e., it has trace 1, it is Hermitian, and positive semidefinite. If we choose as basis in  $\mathcal{L}$ ,  $\{i\sigma_{x,y,z} \otimes \mathbf{1}, i\mathbf{1} \otimes \sigma_{x,y,z}\}$ , in that order and as basis in  $su(4)$   $\{i\sigma_{x,y,z} \otimes \mathbf{1}, i\mathbf{1} \otimes \sigma_{x,y,z}, i\sigma_x \otimes \sigma_{x,y,z}, i\sigma_y \otimes \sigma_{x,y,z}, i\sigma_z \otimes \sigma_{x,y,z}\}$ ,

---

<sup>19</sup>The dimension of the generic orbit under local transformations for the case of pure states is calculated in [44].

in that order, the matrix  $ad_{i\rho_2}$  is given by

$$ad_{i\rho_2} := \begin{pmatrix} 0 & -\alpha & 0 & 0 & 0 & 0 \\ \alpha & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\alpha & 0 \\ 0 & 0 & 0 & \alpha & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\alpha \\ 0 & 0 & 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & -\alpha & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \alpha & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

which has rank 6 for  $\alpha \neq 0$ .

We notice that the above considerations already suggest a test for local equivalence of states. In fact, if two states  $\rho_1$  and  $\rho_2$  are equivalent the corresponding orbits  $\mathcal{O}_L(\rho_1)$  and  $\mathcal{O}_L(\rho_2)$  coincide by definition and therefore they have the same dimensions. If we add this to the obvious requirement that  $\rho_1$  and  $\rho_2$  must have the same spectrum, we have

**Proposition 8.3.2** If two density matrices  $\rho_1$  and  $\rho_2$  are locally equivalent, then

1. They have the same spectrum
2.  $\dim \ker(ad_{i\rho_1}) = \dim \ker(ad_{i\rho_2})$ .

Proposition 8.3.2 then reduces the problem to checking local equivalence for iso-spectral density matrices whose orbits have the same dimension. Given the dimension of a *generic* orbit as above, there will be  $\sum_{j=1}^N n_j^2 - N$  parameters to identify the position of a density matrix on that orbit and

$$D := (n_1^2 n_2^2 \cdots n_N^2 - 1) - \left( \sum_{j=1}^N n_j^2 - N \right) \quad (8.40)$$

parameters to identify the orbit itself. Therefore we will need to check at least  $D$  parameters to verify that two density matrices are on the same orbit, assuming that they are on orbits with maximal dimension. If they are on orbits with smaller dimensions we shall need to verify a larger number of parameters.

### 8.3.2 Invariants and polynomial invariants

Much of the study of the problem of local equivalence of states has focused on the search for local invariants. Let  $\mathcal{B}$  denote the set of density matrices for a general multipartite system and  $\mathcal{L}$ , as above, be the Lie algebra corresponding to the Lie group of local transformations. A function  $f : \mathcal{B} \rightarrow \mathbf{R}$  (or  $\mathbb{C}$ ) is called a **local invariant** if for every  $\rho \in \mathcal{B}$  and  $X \in e^{\mathcal{L}}$

$$f(\rho) = f(X\rho X^\dagger).$$

If  $f$  is a local invariant, for every two local equivalent matrices  $\rho_1$  and  $\rho_2$ ,  $f(\rho_1) = f(\rho_2)$ , in particular every measure of entanglement must be a local invariant. An example of a local invariant for a system of  $N$  qubits is the  $N$ -qubits concurrence  $C_{\Theta_{o,e}}$  of Remark 8.2.7. To distinguish two density matrices which are on generic (maximal dimension) orbits one needs (at least) a number  $D$  in (8.40) of local invariants. A set of local invariants  $\mathcal{S} := \{f_1, \dots, f_m\}$  is *redundant* if there are two different subsets of  $\mathcal{S}$ ,  $\mathcal{S}_1$  and  $\mathcal{S}_2$ , such that the fact that  $\mathcal{S}_1$  is a set of invariants implies that  $\mathcal{S}_2$  is also a set of invariants. If it is not redundant it is called *independent*. A set of invariants  $\mathcal{S} := \{f_1, \dots, f_m\}$  is called **complete** when  $\rho_1$  and  $\rho_2$  are locally equivalent *if and only if*  $f_j(\rho_1) = f_j(\rho_2)$ ,  $j = 1, \dots, m$ . One would like to be able to determine the local equivalence of density matrices using the minimum possible number of invariants and therefore a set of independent invariants. The crucial problem is however to provide a *complete* set of invariants.

Among functions, polynomials have remarkable algebraic properties. In particular, invariant polynomials with values in  $\mathbf{R}$  or  $\mathbb{C}$  form an algebra over  $\mathbf{R}$  or  $\mathbb{C}$  since the linear combination or product of invariant polynomials is still an invariant polynomial. It follows from results in invariant theory [194] that the algebra of invariant polynomials is generated by a *finite* set of homogeneous invariant polynomials.<sup>20</sup> Therefore the task is to find a set of homogeneous invariant polynomials. Examples are given by the coefficients of the characteristic polynomial of the matrix  $\rho$ .

In [89], [167] a method is described to construct systematically all the homogeneous invariant polynomials of degree 1, 2, ... and so on. Also methods are described to eliminate, in the process, some polynomials which are not independent. Notice in this regard that if  $p_1$  and  $p_2$  are homogeneous invariant polynomials of degree  $d_1$  and  $d_2$ ,  $p_1 p_2$  is a homogeneous invariant polynomial of degree  $d_1 + d_2$ . Let  $\bar{n}$  be the dimension of the total system  $\bar{n} := n_1 n_2 \cdots n_r$ . The starting point of the procedure of [89], [167] for the construction of polynomial invariants is the observation that, to every homogeneous polynomial  $p = p(\rho)$  of degree  $k$ , there corresponds an  $\bar{n}^k \times \bar{n}^k$  matrix  $F$ , such that

$$p(\rho) = \text{Tr}(F\rho^{\otimes k}), \quad (8.41)$$

---

<sup>20</sup>That is every invariant polynomial is obtained by linear combination and-or products of a finite set of homogeneous invariant polynomials.

where  $\rho^{\otimes k}$  denotes the tensor product of  $k$  copies of  $\rho$ . Moreover,  $p$  is an invariant polynomial if and only if (8.41) is satisfied for an  $F$  which commutes with  $(e^{\mathcal{L}})^{\otimes k}$ . In order to see this, assume (8.41) holds for some  $F$  which commutes with  $(e^{\mathcal{L}})^{\otimes k}$ . We have, with  $X \in e^{\mathcal{L}}$ ,

$$\begin{aligned} p(X\rho X^\dagger) &= \text{Tr}(F(X\rho X^\dagger)^{\otimes k}) = \text{Tr}(FX^{\otimes k}\rho^{\otimes k}(X^\dagger)^{\otimes k}) = \\ &\text{Tr}((X^\dagger)^{\otimes k}FX^{\otimes k}\rho^{\otimes k}) = \text{Tr}(F\rho^{\otimes k}) = p(\rho), \end{aligned}$$

where we used the commuting property of  $F$ . This shows that  $p$  is an invariant polynomial. Conversely, assume that  $p$  is an invariant polynomial and write  $p$  as in (8.41). From the invariance property of  $p$ , we have

$$p(\rho) = \text{Tr}(U^\dagger FU\rho^{\otimes k}), \quad (8.42)$$

for every  $U \in (e^{\mathcal{L}})^{\otimes k}$ . If  $d\mu$  denotes the invariant measures on the compact Lie group  $(e^{\mathcal{L}})^{\otimes k}$  (see, e.g., [150], [220]), then we can integrate (8.42) with this measure over the compact Lie group  $(e^{\mathcal{L}})^{\otimes k}$  and write

$$p(\rho) = \text{Tr}(\tilde{F}\rho^{\otimes k}),$$

with

$$\tilde{F} := \int_{(e^{\mathcal{L}})^{\otimes k}} U^\dagger FU d\mu.$$

Because of the invariance of the measure  $d\mu$ ,  $\tilde{F}$  commutes with all of  $(e^{\mathcal{L}})^{\otimes k}$ .

The next observation is that there is no loss of generality in looking for invariant homogeneous polynomial  $p$  such that the corresponding matrix  $F$  is Hermitian. In fact, given that  $F$  commutes with every  $U \in (e^{\mathcal{L}})^{\otimes k}$ , we also have

$$U(F + F^\dagger)U^\dagger = F + F^\dagger,$$

and

$$U(iF - iF^\dagger)U^\dagger = iF - iF^\dagger.$$

Therefore there are two Hermitian (linearly independent) matrices  $M_1 := F + F^\dagger$  and  $M_2 := iF - iF^\dagger$  which correspond to homogeneous invariant polynomials of degree  $k$ . Notice, moreover, that these polynomials are real.

The last observation is that if (and only if) the  $\bar{n}^k \times \bar{n}^k$  Hermitian matrix  $M$  commutes with the elements of the Lie group  $e^{\mathcal{L}^{\otimes k}}$ , then it commutes with all the elements of the Lie algebra corresponding to  $e^{\mathcal{L}^{\otimes k}}$ . Such a Lie algebra is a  $\bar{n}^k$  representation (see, e.g., [106]) of  $\mathcal{L}$ . It is spanned by matrices of the form

$$G := A \otimes \mathbf{1} \otimes \cdots \otimes \mathbf{1} + \mathbf{1} \otimes A \otimes \mathbf{1} \otimes \cdots \otimes \mathbf{1} + \cdots + \mathbf{1} \otimes \cdots \otimes \mathbf{1} \otimes A, \quad (8.43)$$

with  $A$  in  $\mathcal{L}$ , where the tensor products are taken  $k$  times.

In conclusion, one has the following procedure to calculate all the homogeneous polynomial invariants of degree  $k$ : First, one calculates a basis of  $\mathcal{L}$  and therefore a maximal set of linearly independent matrices  $G$  in (8.43). Then one finds all the Hermitian,  $\bar{n}^k \times \bar{n}^k$  matrices which commute with all the elements  $G$ . This involves the solution of a linear system of equations.

Although the above procedure only involves eventually linear algebra operations and allows us to calculate all the homogeneous invariant polynomials of a degree  $k$ , the dimension of the system of equations grows exponentially with  $k$ . Moreover, the procedure does not take into account the fact some homogeneous polynomials of degree  $k$  can be obtained from polynomials of degree smaller than  $k$ . Therefore their computation is a waste of computational resources. We refer to [89] and [167] for a treatment of these problems.

Finally, there is always the problem, once a certain number of invariants has been found, to show that they form a complete set of invariants.

### 8.3.3 Some solved cases

For some particular situations, a complete set of local invariants have been found. These invariants allow us to decide whether two density matrices are locally equivalent. They are not necessarily polynomials. We mention here some cases presented in the literature without going into details of the specific results. In [142] the case of two qubits is solved and a complete set of 18 invariants is found. The case of  $N \geq 2$  qubits is considered in [136] and generic (i.e., maximal dimensional) orbits are considered. In this case the invariants are the parameters of a canonical point on the orbit. These invariants allow us to distinguish density matrices belonging to two different generic orbits. The case of three qubits is considered in detail in [136] and generalizations are indicated for a larger number of qubits. A complete set of invariants is found for bipartite generic states in [11] while a complete set of invariants for a class of tripartite pure states is given in [9]. A special class of bipartite quantum systems is also considered in [10] while nondegenerate tripartite mixed states are considered in [83] using a method developed in [79].

## 8.4 Notes and References

In this chapter, we have given some of the main ideas concerning the quantum theory of entanglement from the point of view of the properties of states and their dynamics. There are several review papers (see, e.g., [101], [163], [202], [224]) on quantum entanglement and textbooks (see, e.g., [125], [143], [155], [164]). We refer to these references for more details. These references also present a more operational point of view in the sense that they consider

quantum entanglement as a resource in quantum communication, teleportation and computation. A neighboring field is the study of quantum states and their classification [20]. It must be said that the theory is in rapid progress and several mathematical problems have received only a partial solution. In our treatment we have illustrated three clear interrelated mathematical problems:

- The problem of finding tests to check whether the state of a multipartite system is entangled or not.
- The problem of defining meaningful and useful measures of entanglement.
- The problem of testing whether it is possible to transfer a quantum state from one value to another by only local transformations.

In the study of entanglement several new concepts are being introduced and relations with other concepts are discovered. As for controllability, a concept of *indirect controllability* was introduced in [177], [178] which concerns the possibility of driving the state of a system by manipulating the state of an ancillary system which then interacts with the target system. In this scheme of control (as opposed to the schemes treated in this book) the evolution is fixed and determined by a (total) Hamiltonian  $H$ , while the control degree of freedom is the initial state of the ancillary system. If  $H$  only gives local evolutions, clearly there is no possibility of controlling the target system with this method. In general, it is reasonable to expect that the controllability properties of this scheme will depend on the entangling properties of the Hamiltonian  $H$ . In fact, in [177] an explicit relation between controllability and entanglement was discovered in the simplest case where both target system and ancillary system are two level systems.

The three entanglement criteria described in subsection 8.1.2 are among the strongest available at the moment of writing of this book (November 2006). For example, the CCN criterion is proved in [180] to be not weaker than the reduction criterion in [102]. However, the list is obviously not complete.

Our main references have been [163] for measures of entanglement and [224], in particular for entanglement of formation and concurrences, to which we refer for further references. The main reference for the OED Decomposition was [61] and for the treatment of local invariant polynomial [89].

We believe that a treatment of entanglement based on dynamics, i.e., on the study of the unitary group action on the set of density matrices, along with application of the results on the structure of the unitary group is very promising. We have seen some evidence of this in the treatment of concurrences and generalized concurrences in subsections 8.2.1 and 8.2.2.

## 8.5 Exercises

**Exercise 8.1** Prove Proposition 8.1.3. Use the last part of this Proposition to give an algorithm in terms of matrix elements for the calculation of the partial trace in the case of two spin  $\frac{1}{2}$  systems, where, for each subsystem, we consider the natural basis  $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ .

**Exercise 8.2** Prove Proposition 8.1.4.

**Exercise 8.3** Analyze the entanglement of the two qubits system in Example 8.1.12 using the PPT criterion of Theorem 8.1.7.

**Exercise 8.4** Prove formula (8.17). Use this formula to prove that the entropy of the partial trace is a monotonically increasing function of the concurrence which has its minimum (maximum) at zero (one) and that the minimum (maximum) is equal to zero (one).

**Exercise 8.5** Prove that  $\tilde{\theta}_{II} \otimes \tilde{\theta}_{II}(\rho) = J \otimes J \bar{\rho} J \otimes J$  and that all the eigenvalues of  $\rho \tilde{\theta}_{II} \otimes \tilde{\theta}_{II}(\rho)$  are real and nonnegative.

**Exercise 8.6** Prove that (8.24) follows from (8.23) (Hint: can use contradiction by assuming that one quantity is strictly larger than the other in (8.24). Use (8.20)).

**Exercise 8.7** Notice that the procedure described in subsection 8.3.2 to generate homogeneous invariant polynomials does not depend on the nature of the Lie algebra  $\mathcal{L}$  (which in the case treated was the Lie algebra corresponding to local transformation on a multipartite system). Consider a two level system and calculate all the homogeneous polynomials of degree 1 and 2 which are invariant with respect to the Lie sub-algebra  $so(2)$  of  $su(2)$ .

# Chapter 9

---

## *Applications of Quantum Control and Dynamics*

In this chapter, we shall discuss some practical applications of quantum control and dynamics. Our goal is not to give an exhaustive introduction to areas of application, but to point out the role of control and dynamics so as to further motivate the field as well as the techniques described in the previous chapters. The basics of mathematical modeling of quantum mechanical control systems were described in [Chapter 2](#). Here we focus on more practical aspects in the areas of nuclear magnetic resonance and control of molecular and atomic systems. We also discuss a possible implementation of quantum information processing, namely the one with trapped ions. The dynamics here is used to implement quantum logic gates. We shall show how, changing the physical set-up, it is possible to create different Hamiltonians. The control is obtained by switching among these different Hamiltonians.

---

### 9.1 Nuclear Magnetic Resonance Experiments

#### 9.1.1 Basics of NMR

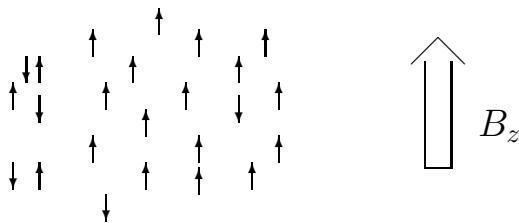
The basic set-up of an experiment in nuclear magnetic resonance (NMR) consists of an ensemble of spin- $\frac{1}{2}$  particles in a magnetic field  $\vec{B}$ , which is constant and pointing in a fixed direction, say the  $z$ -direction. If all the other degrees of freedom are neglected, the energy of one of the spin- $\frac{1}{2}$  particles is proportional to the dot product between the spin angular momentum and the magnetic field (cf. [subsection 2.3.1](#)).

Liouville's equation describing the dynamics of this system is given in (2.65) of Chapter 2. In this case, ( $B_x = 0$ ,  $B_y = 0$ ), this equation takes the form<sup>1</sup>

$$i \frac{d}{dt} \rho = \gamma \frac{1}{2} [\sigma_z B_z, \rho].$$

---

<sup>1</sup>For simplicity, we are using units such that  $\hbar = 1$ . We also drop the index  $e$  in  $\vec{B}_e$  used in Chapter 2 to denote an ‘external’ field.



**FIGURE 9.1:** The initial equilibrium set-up of an NMR experiment consists of an ensemble of spin  $\frac{1}{2}$  particles in a constant magnetic field  $B_z$  along the  $z$  direction. Most of the spins are aligned in the lowest energy configuration in the same direction as the field. Some of them are aligned in the opposite direction.

This equation has, as equilibrium points, the density matrices commuting with  $\sigma_z$ . For a single spin,  $\rho_+ := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ ,  $\rho_- := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$  correspond to spin angular momentum aligned in the same direction as the field, the  $z$  direction, or in the opposite direction, respectively. These correspond to two values of the energy which are given by  $\pm \frac{\gamma}{2}B_z$ . In the absence of any other external field, the ensemble of spin particles will be in one of the two equilibrium states, with a majority in the state of lower energy, which corresponds to spin in the same direction as the magnetic field (while the state with higher energy corresponds to the spin in the opposite direction). Denoting by  $N_L$  the number of spins in the lower energy state and by  $N_H$  the number of spins in the higher energy value, the density matrix describing the ensemble will be

$$\rho_{EQ} := \begin{pmatrix} \frac{N_L}{N_H + N_L} & 0 \\ 0 & \frac{N_H}{N_H + N_L} \end{pmatrix}. \quad (9.1)$$

The ratio between the number  $N_L$  of spins in the lower energy state and the number  $N_H$  of spins in the higher energy state can be calculated using *Maxwell-Boltzmann statistics*. It is given by

$$\frac{N_H}{N_L} = e^{-\frac{\Delta}{kT}},$$

where  $\Delta := |\gamma B_z|$  is the difference between the two values of the energy,  $k$  is the Boltzmann constant and  $T$  the absolute temperature. This ratio is always less than one although at room temperature it is almost one. As the temperature decreases, there are few spins with the high energy value.

In NMR experiments, the spin system is perturbed from the equilibrium state using the  $x$  and  $y$  components of the magnetic field ( $B_x$  and  $B_y$  in equation (2.65)) as controls. Once the state  $\rho$  is in a nonequilibrium position  $\rho_0$  and the  $x$  and  $y$  components of the magnetic field have been switched back to zero,  $\rho$  will evolve according to

$$\rho(t) = e^{-i\frac{\omega_0}{2}\sigma_z t} \rho_0 e^{i\frac{\omega_0}{2}\sigma_z t}, \quad (9.2)$$

where the *Larmor frequency*  $\omega_0$  is given by  $\omega_0 := \gamma B_z$ . Define a *magnetization vector*,  $\vec{M} := M_x \vec{i} + M_y \vec{j} + M_z \vec{k}$ , with  $M_{x,y,z} := \text{Tr}(\sigma_{x,y,z} \rho)$ . According to (9.2), the vector  $\vec{M}$  will revolve around the  $z$  axis in a periodic motion with frequency given by the Larmor frequency  $\omega_0$ . Such a motion of the spin system is called **precession**.

The system of spins has to be regarded as the quantum mechanical version of a magnet whose associated magnetic field is the magnetization vector  $\vec{M}$ . According to Faraday's induction law (2.6) the precession motion will induce a current in a nearby coil.<sup>2</sup> The signal detected is a sinusoid with magnitude proportional to the component of  $\vec{M}(0)$  = in the plane determined by the normal to the coil inner surface; see Figure 9.2. The frequency of the sinusoid is given by the Larmor frequency.

In reality, the signal detected by the coil is a damped sinusoid (rather than a perfect sinusoid) of the type in Figure 9.3. This signal is called *Free Induction Decay* (FID). The damping is due to the *relaxation* of the spin, which is the interaction of the spin with its surroundings. The relaxation process eventually takes the spin system back to its original state  $\rho_{EQ}$  in (9.1).<sup>3</sup>

---

<sup>2</sup>In order to see this consider the integral version of Faraday's law, i.e., the one obtained from (2.6) by taking the surface integral on the inner surface of the coil  $S_c$ . This gives

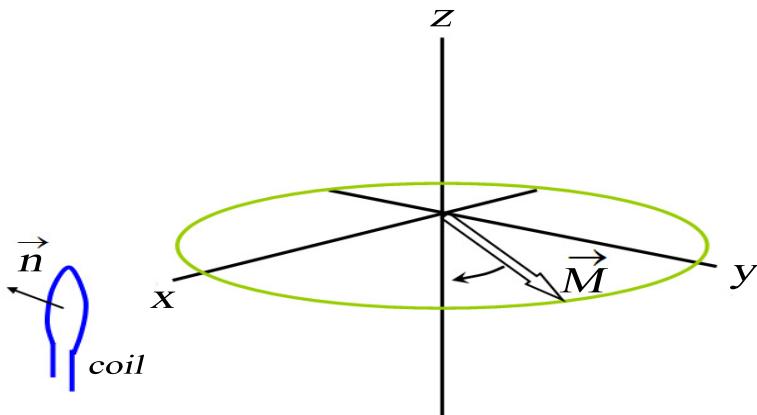
$$\int_{S_c} (\nabla \times \vec{E}) \cdot \vec{n} dS = -\frac{\partial}{\partial t} \int_{S_c} \vec{B} \cdot \vec{n} dS,$$

and applying Stokes' theorem of vector calculus we transform the first integral into a line integral of the electric field along the border of  $S_c$ ,  $\partial S_c$ . We have

$$\int_{\partial S_c} \vec{E} \cdot d\vec{l} = -\frac{\partial}{\partial t} \int_{S_c} \vec{B} \cdot \vec{n} dS.$$

The electric field along the coil produces the current.

<sup>3</sup>A phenomenological analysis of the relaxation process [25], [41] shows that it is mainly due to two processes: The *relaxation spin-lattice* is due to the fluctuations of the magnetic field in the lattice and takes the value of  $M_z$  back to its equilibrium value. The *spin-spin relaxation* is due to interactions with other spins as well as inhomogeneity of the field  $B_z$ , which causes the particles to rotate at slightly different Larmor frequencies. It is the process responsible for making the transversal components of the magnetization  $M_x$  and  $M_y$  go to zero. The characteristic times of the relaxation process depend on the external environment of the spin particle.



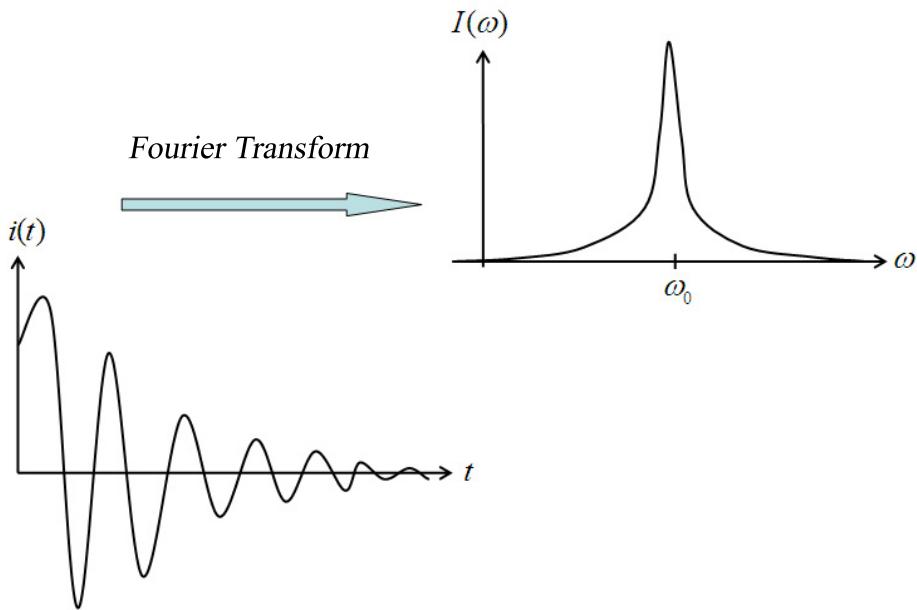
**FIGURE 9.2:** In NMR experiments, a precessing magnetization vector induces a sinusoidal current in a coil. This is a consequence of Faraday's induction law (2.6).

The FID signal is Fourier transformed and shows a peak at the Larmor frequency  $\omega_0$ . This Larmor frequency depends on the type of spin particle which has produced the FID signal. For example, it is in ratio 10 : 1 : 2.5 for the nuclei of the Hydrogen isotope  $^1\text{H}$ , the Nitrogen isotope  $^{15}\text{N}$  and the Carbon  $^{13}\text{C}$ , respectively. Therefore, the evaluation of the Larmor frequency allows one to determine the type of nucleus that has generated the signal and therefore it is a tool for the determination of the structure of molecules.<sup>4</sup>

The motion of the electrons around the nucleus creates a magnetic field which (in most cases) opposes the constant external magnetic field. This magnetic field will be different according to the environment of the nucleus under consideration and the chemical bonds with the neighboring nuclei. It will result in a change in the Larmor frequency which is called *chemical shift*. The Larmor frequency is always proportional to the external field in the  $z$

---

<sup>4</sup>This mode of operation is sometimes called *Fourier spectroscopy* NMR. Another method of operation is *continuous wave NMR*. There are two ways to perform the NMR experiments in continuous wave NMR. In one of them, in addition to the magnetic field in the  $z$ -direction which is kept constant, the sample is subject to a rotating electromagnetic field in the  $x$ - $y$  plane for a certain time. The frequency of this field is varied. At a frequency equal to the Larmor frequency the magnetization in the  $z$ -direction will reverse direction. In terms of energy, the magnetic field will induce population transfer from the lower to the higher energy level (cf. section (7.1)). The frequency at which this state transfer occurs gives information on the Larmor frequency of the spin particle. In another scheme, the magnetic field in the  $z$ -direction is varied while the frequency of the transversal field is kept constant. Analogously to the previous case a transition will occur when the frequency of the incoming field is equal to the Larmor frequency. Recall from Chapter 6 and in particular section 6.3 that a rotating field in the  $x$ - $y$  plane (6.40) (6.41) at the resonance frequency optimizes the energy necessary for the given state transfer.

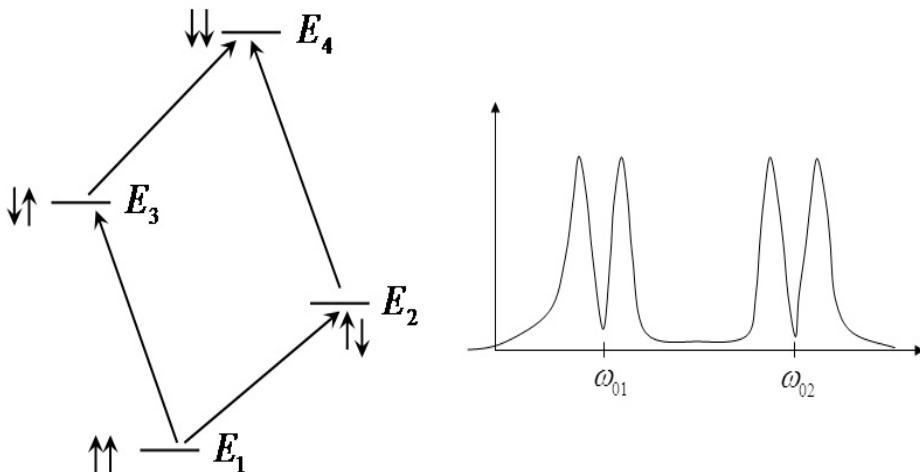


**FIGURE 9.3:** Free induction decay signal  $i = i(t)$  in the time domain and its Fourier transform  $I(\omega)$ .

direction, but the proportionality constant changes because of the chemical shift. Thanks to the chemical shift, it is possible to distinguish nuclei of the same type but in different positions in a given molecule and therefore to assign various peaks in a spectrum to the various nuclei in a molecule. This process is called **frequency labeling**.

A different way of looking at an NMR experiment is in terms of energy levels and transitions among different levels. A spin- $\frac{1}{2}$  particle can be in two distinguished states with energy  $E = \pm \frac{\omega_0}{2}$ . When stimulated with an external time-varying field the system undergoes some population transfer from the lower to the higher energy level. The energy so accumulated is then emitted during the precession-relaxation process and therefore the frequency detected in the Fourier analysis corresponds to the difference in energy levels. This picture extends to the case of several noninteracting spin particles in that every peak in the spectrum, corresponding to a precession-relaxation process of a single spin, is associated to a transition frequency between two different energy levels in the energy diagram. This picture is particularly useful in the case where the spin particles cannot be considered isolated and the coupling among them is not negligible. The principle that the lines in the spectrum correspond to transitions among different energy levels still holds and from these transitions it is possible to study the structure of a molecule. A typical

energy diagram and spectrum for a two spin systems is presented in Figure 9.4.

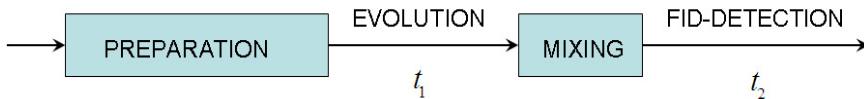


**FIGURE 9.4:** Energy diagram for a system of two spins with different Larmor frequencies  $\omega_{0,1}$  and  $\omega_{0,2}$  in the presence of a weak coupling. The transition from the level  $E_1$  to the level  $E_2$  as well as the transition from  $E_3$  to  $E_4$  correspond to a ‘flip’ of the second spin. In absence of coupling we would have  $E_4 - E_3 = E_2 - E_1 = \omega_{0,2}$ . This would correspond to a peak in the energy spectrum at  $\omega_{0,2}$ . In the presence of a weak coupling between the two spins this peak slightly splits to form a doublet centered at  $\omega_{0,2}$  whose width depends on the size of the coupling (which is (in the weak coupling case) assumed much smaller in magnitude than the Larmor frequency). A similar doublet is formed around the Larmor frequency  $\omega_{0,1}$  corresponding to transitions involving the first spin.

### 9.1.2 2-Dimensional NMR

In principle, in NMR experiments, one can analyze the geometry of a molecule by exciting one nucleus and studying the effect on neighboring nuclei. This assumes that the frequency labeling is performed accurately. For NMR of large proteins, [45], [225], the spectra are typically poorly resolved and the frequency labeling process is quite difficult if not impossible. This is due both to the larger number of nuclei, which make the spectrum overly crowded, and to the fact that larger relaxation rates make the peaks wider in the Fourier transform of the FID signal. A way to overcome this problem is to

use 2-dimensional NMR spectroscopy in which two nuclei are distinguished by studying their interaction with a third spin. A typical 2-D NMR experiment follows the scheme of Figure 9.5.



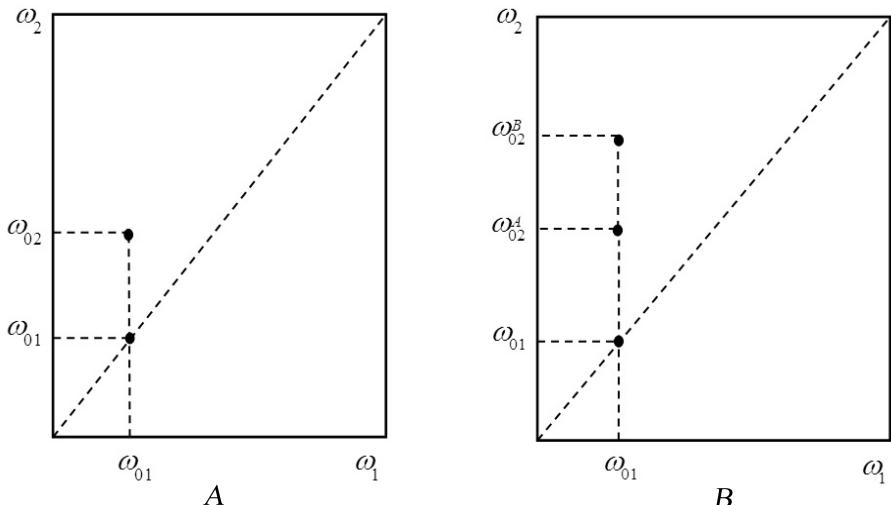
**FIGURE 9.5:** Various Stages of a 2-Dimensional NMR Experiment.

In the period called *preparation*, a sample is perturbed from equilibrium with one or more control pulses which may cause the spin of one of two nuclei, say spin 1, to transfer from being aligned along the  $z$ -axis to a position in the  $x$ - $y$  plane. During the *evolution* time  $t_1$  the system evolves for example by allowing the magnetization of spin 1  $M_{x,y,z} = \text{Tr}(\sigma_{x,y,z}\rho)$  to precess in the  $x$ - $y$  plane according to (9.2). During the *mixing* period more control pulses are applied in order to modify the state of spin 2 by interaction with spin 1. This process is sometimes called *polarization transfer* or *magnetization transfer*. After the mixing period the FID signal  $f(t_2)$  is detected as a function the time  $t_2$  elapsed from the end of the mixing period. This signal will, however, depend on the time  $t_1$ , i.e., it will be  $f(t_2) := f_{t_1}(t_2)$ . The experiment is repeated for different values of  $t_1$ , starting with  $t_1 = 0$  and with small increments of  $t_1$ . Typically between 50 and 500 repetitions are performed so as to obtain a function  $s(t_1, t_2) := f_{t_1}(t_2)$ . This function is then Fourier transformed with respect to  $t_1$  and  $t_2$ , with a two dimensional Fourier transform

$$S(\omega_1, \omega_2) := \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(t_1, t_2) e^{-i2\pi(t_1\omega_1 + t_2\omega_2)} dt_1 dt_2.$$

The function  $S(\omega_1, \omega_2)$  is represented as a two dimensional function, in particular as a contour plot, and it will present various (two dimensional) peaks. A peak at a frequency pair  $(\omega_{01}, \omega_{02})$  is interpreted as due to a portion of the magnetization for spin 1 which was precessing with Larmor frequency  $\omega_{01}$  which is transferred to spin 2 and therefore precesses at Larmor frequency  $\omega_{02}$ . Therefore the mixing time has the role of transferring signal from one spin to another. Part A of Figure 9.6 shows two peaks in a two-dimensional spectrum, one diagonal peak at  $(\omega_{01}, \omega_{01})$  corresponds to signal which is not transferred from one spin to the other while the cross peak  $(\omega_{01}, \omega_{02})$  corresponds to a signal transferred to a different frequency.

Using 2-D NMR it is possible to distinguish spins that are indistinguishable in conventional NMR because their Larmor frequencies are too similar, say  $\approx \omega_{01}$ . If they interact with two different nuclei having different Larmor



**FIGURE 9.6:** Plots in 2-D NMR spectroscopy. In part A two peaks reveal that part of the signal has been transferred during the mixing time to a second spin. In part B two off diagonal peaks allow to distinguish two different spins with approximately equal Larmor frequency.

frequencies equal to  $\omega_{02}^A$  and  $\omega_{02}^B$ , respectively, the two dimensional Fourier transform will have plots as in Part B of Figure 9.6, which allows to distinguish the two spins.

### 9.1.3 Control problems in NMR

Already in basic NMR spectroscopy several manipulation tasks of nuclear spins can be studied from the point of view of control theory. For example, assume we want to maximize the signal to noise ratio in the detection of the FID signal. Given a certain orientation for the coil detecting the *FID* signal, we require that the component of the magnetization vector in the plane orthogonal to the coil be maximum at the beginning of the precession. In fact, we have seen that the amplitude of the sinusoidal signal detected is proportional to such a component. This is a control problem of choosing the  $B_x$  and  $B_y$  functions in Liouville's equation (2.65) in the period preceding the precession.

There are many important control problems in 2-*D* NMR concerning the construction of control pulses to transfer magnetization from one spin to the other. Methods to obtain this by minimizing the effects of relaxation, maximizing the efficiency of transfer, and minimizing the time have been only recently rigorously tackled using tools of control theory. We refer in particular the work of N. Khaneja and co-workers ([122], [123], and references

therein) which also treat specific applications to structure determination for large proteins. The mathematical models in the case of closed systems are bilinear control systems as described in [Chapter 2](#).

---

## 9.2 Molecular Systems Control

The mathematical modeling of quantum mechanical molecular and atomic control systems was discussed in Chapter 2. Our goal here is to indicate applications as well as to describe practical issues related to the laboratory implementation.

### 9.2.1 Pulse shaping

Molecular and atomic quantum control is achieved through appropriately shaped laser pulses. A laser emits electromagnetic radiation which may have different shapes. These shapes can be modified using *pulse shaping* techniques. Modern pulse shapers use optical materials whose refractive indexes depend (in a nonlinear fashion) on the frequency and/or amplitude of the light which crosses them. By placing in parallel and-or in sequence several optical materials with different refractive indexes and-or refocusing the light in a given point, one can create various shapes, which match the controls designed according to the techniques described in the previous chapters. An example of a pulse shaper is given in [219] to which we refer for references and alternative schemes. The shaper of [219] includes 128 pixels whose refractive index can be changed by modifying the voltages applied to each of them. The pixels form a panel and they are connected to a computer which has encoded the control algorithm to be applied. The scheme of [219] is particularly useful in cases where the control scheme is changed in different experiments because it allows one to obtain different control laws with the same set-up.

In principle, there is no restriction to the light shape which can be obtained with pulse shaping except that the electromagnetic field, once treated itself as a quantum mechanical system, has to obey the Heisenberg uncertainty relation for any two noncommuting observables (cf. 1.2.2.3). For a single sinusoidal pulse with energy content  $\Delta H$  and duration  $\Delta t$ , the Heisenberg uncertainty principle gives

$$\Delta H \Delta t \geq \frac{\hbar}{2}.$$

This, using the relation between energy  $\Delta H$  and frequency  $\omega$ , i.e.,  $\Delta H = \hbar\omega$ ,

gives that it is not possible to engineer very short pulses with small frequency. In general, a pulse will be the superposition of several Fourier components. The Heisenberg uncertainty principle says that very short pulses must have a very high frequency content. This places a restriction in cases where the target of control evolves with characteristic times which are very short.

### 9.2.2 Objectives and techniques of molecular control

The general objective of molecular control is to selectively break or make chemical bonds to achieve desired products in a reaction. In traditional methods, control is often obtained by varying the external conditions of a chemical reaction, by varying for example temperature or pressure, or by introducing a catalyst in the reaction. However, control by electromagnetic pulses appropriately shaped may be the best way to proceed when more traditional methods fail. This is the case for example in the synthesis of molecules whose existence has been predicted theoretically but that cannot be obtained with conventional chemistry. One example might be a ring configuration of Ozone  $O_3$  which is considered a high energy compound of great interest. At this time it has been produced only on a solid support [162] but it would be of interest to obtain it in the gas phase.

An example of a finite dimensional molecular system was given in [203]. The goal of that paper was to demonstrate an optimal control algorithm. The system under consideration is a five level system whose Hamiltonian is given by

$$H = H_0 + H_1 u,$$

with control  $u$ , where the internal Hamiltonian  $H_0$  is given by

$$H_0 := \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1.2 & 0 & 0 & 0 \\ 0 & 0 & 1.3 & 0 & 0 \\ 0 & 0 & 0 & 2.0 & 0 \\ 0 & 0 & 0 & 0 & 2.15 \end{pmatrix},$$

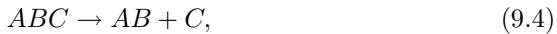
and an interaction Hamiltonian  $H_1$ ,

$$H_1 := \begin{pmatrix} 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{pmatrix}. \quad (9.3)$$

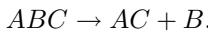
There are two groups of eigenstates of the internal Hamiltonian. The first three levels correspond to vibrational levels of a ground state. These are levels with slightly different energies as compared to the lowest energy level, the

ground state. The remaining two energy levels correspond to excited levels. In absence of a driving field there is no coupling between these levels. Moreover, it follows from (9.3) that the external field only couples eigenstates in the first group with eigenstates in the second group. For this model, Lie algebraic calculations as in Chapter 3 show that one has complete controllability over the possible state transfers. This was done in [168]. One can then apply a method among the ones described in the previous chapters to obtain constructive control. In particular the method described in [203] is a numerical method based on optimal control theory.

The models arising in molecular control are often more complicated than the one described above and in several cases they are infinite dimensional. For this reason, many of the proposed schemes in the literature use physical intuition and take into account laboratory constraints along with the mathematical model. Many examples of schemes in molecular control are presented in the books [174], [192]. Several schemes in the infinite dimensional case use the *combined action* of two or more laser fields. One of the first ideas in this direction was put forward by Brumer and Shapiro [34] and it is developed in depth in the book [192], with a comprehensive review given in [191]. The main idea is as follows. Among the eigenstates in the continuum of energies there are some that evolve into states corresponding to the desired outcome. For example, the photo-dissociation of a three atomic molecule  $ABC$  can proceed according to two different ‘channels’



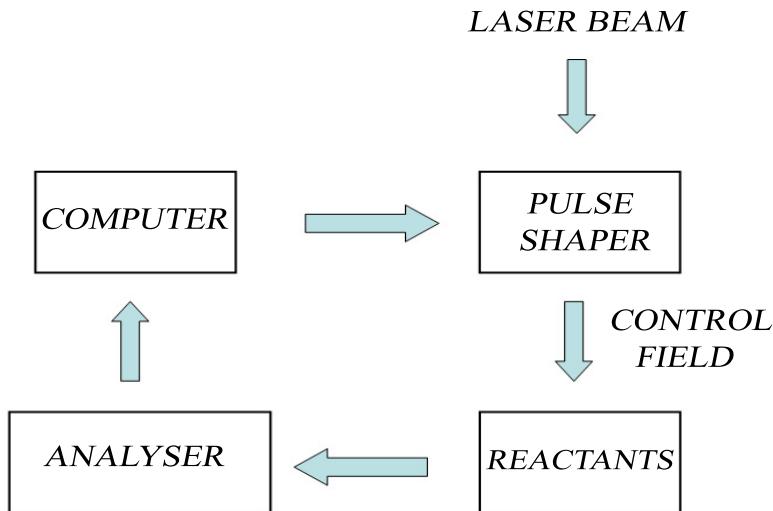
or



There will be, at a given energy  $E$ , states  $|E, n, 1^-\rangle$  which evolve according to the channel (9.4) and states  $|E, n, 2^-\rangle$  which evolve according to the channel 2. Here  $n$  represents additional degrees of freedom and  $E$  is the energy of both types of states  $|E, n, 1^-\rangle$  and  $|E, n, 2^-\rangle$ . The technique of [34] consists of exciting the system to energy  $E$  with *two* (or more) lasers. In first order perturbation theory, the probability of being in a state of the type  $|E, n, 1^-\rangle$  will be the sum of two terms due to the two laser fields and an interference term depending on the relative parameters of the two fields. The control consists of adjusting these parameters so as to maximize (or minimize) the ratio between the probability of having a state of the type  $|E, n, 1^-\rangle$  and the one of having a state of the type  $|E, n, 2^-\rangle$ .

One of the major obstacles in applications of control algorithms to molecular systems is that the models are not only very complex but also have unknown parameters. For this reason, a *learning approach* to the control of molecular processes has been recently introduced. This approach has shown great promise [165]. Several samples of the same reactants are used and in each experiment the result is recorded. On the basis of this result the control law

is modified for the following experiment, and the process is repeated several times. Algorithms of genetic type [85] are often used for this purpose. Figure 9.7 (cf. [16]) presents a learning scheme for molecular control.



**FIGURE 9.7:** Scheme for the laboratory implementation of learning molecular control with pulse shaping.

### 9.3 Atomic Systems Control; Implementations of Quantum Information Processing with Ion Traps

Atomic systems such as *atoms in gas phase*, *atoms in optical cavities*, *trapped ions* provide a wealth of models for the application of techniques of control and dynamical analysis. These systems are relatively simple and typically have large coherence times (i.e., large intervals of time where they behave quantum mechanically). There are several applications of these systems which motivate the use of control techniques. These include *quantum metrology*, *quantum communication*, and *quantum information processing*. Using

control methods, it is possible to achieve particular states of interest which are not available with other means. Moreover, control of atomic systems can be used to test physical theories and models. We refer to [139] for a discussion of the applications of control of atomic systems and for a list of references.

In the sequel of this section, we shall discuss a specific proposal of implementation of quantum information processing [48], [108], [195]. In particular we shall be concerned with the proposal using *trapped ions*. We use this as an example of the modeling techniques of Chapter 2 as well as of practical implementation of different Hamiltonians for the same system. It should be emphasized that this is only one of many proposals for the implementation of quantum computation which include, among others, the use of NMR experiments.

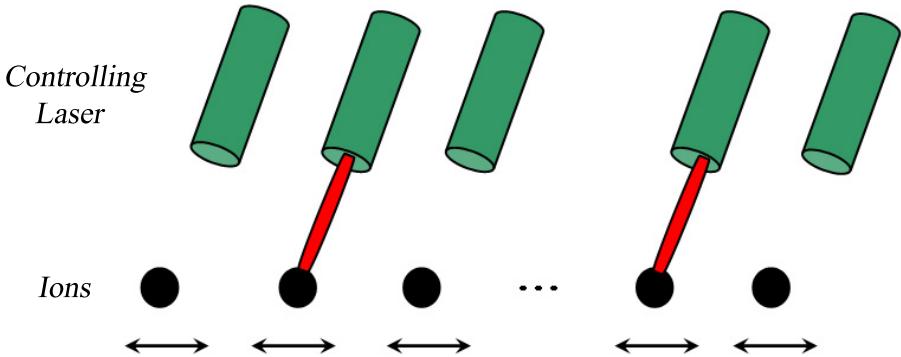
### 9.3.1 Physical set-up of the trapped ions quantum information processor

Trapped ions can be controlled by laser fields and used to perform quantum information processing. The **ion trap quantum information processor** consists of a number  $N$  of controlled ions which are confined in a region by an appropriate potential. Typically, the ions are placed along a line (linear trap) which we take as the  $x$  axis. Each ion is used to implement a quantum bit. In particular, two of the energy levels of each ion are used as representative of 0 and 1 in quantum information. Each ion has internal degrees of freedom (such as electron spin and nuclear spin) which contribute to the internal energy levels. Moreover, there exists an external *vibrational* degree of freedom given by the motion of the ions along the  $x$  axis. By various techniques, it is possible to *cool* the ions to the point where only collective vibrations of the ions are significant and they are small. With a laser field it is possible to control the single ions and therefore perform one-quantum-bit operations. Two-quantum-bit operations can be performed by using the extra degree of freedom given by the position of the center of mass of the ions as an auxiliary qubit. One first performs a quantum logic operation by shining a laser on the  $j$ -th ion. This mathematically corresponds to an operation on the space spanned by the eigenvectors corresponding to the internal degrees of freedom of the  $j$ -th ion and the eigenvectors of the position of center of mass. There is a coupling between these two ‘subsystems’<sup>5</sup> because the electromagnetic field seen by the  $j$ -th ion depends on its position. Similarly, a quantum evolution can be performed between the  $k$ -th ion and the subsystem corresponding to the position of center of mass. The net result is a two qubit operation between the  $j$ -th and the  $k$ -th ions and it is possible to show that the operations so

---

<sup>5</sup>They do not correspond physically to two quantum systems but mathematically are treated as such.

obtained are universal for quantum computation. A scheme of the ion trap quantum information processor is described in Figure 9.8 (cf. [48]).



**FIGURE 9.8:** Ion Trap quantum information processor consisting of a string of collectively vibrating ions in a linear trap controlled through laser sources. In the figure two laser sources control two ions. These can then be switched off and other laser sources can be activated to control other ions.

### 9.3.2 Classical Hamiltonian

The (classical) Hamiltonian describing the motion of the ions is given by (2.33) where the external potential  $V_e$  (2.32) takes the form of a harmonic potential term, that is

$$V_e := \sum_{j=1}^N \frac{m_j}{2} (\omega_x^2 x_j^2 + \omega_y^2 y_j^2 + \omega_z^2 z_j^2),$$

imposed by the external set up. This is the Hamiltonian describing the *collective motion* of the ions. It does not take into account the internal dynamics of each ion and the interaction of the ions with the laser fields. These will be considered in the next two subsections. Assuming that the ions all have the same mass  $M$  and the same charge  $e$  and neglecting the terms in (2.45) corresponding to the energy of the field only, we rewrite the Hamiltonian describing the motion of the  $N$  ions as

$$H = \sum_{j=1}^N \frac{1}{2} M \left( \left( \frac{d\vec{r}_j}{dt} \right)^2 + \omega_x^2 x_j^2 + \omega_y^2 y_j^2 + \omega_z^2 z_j^2 \right) + \frac{1}{8\pi\epsilon_0} \sum_{k \neq j} \frac{q_k q_j}{|\vec{r}_k - \vec{r}_j|}. \quad (9.5)$$

The design is such that  $\omega_y, \omega_z \gg \omega_x$ . Moreover, the ions are cooled to very low temperatures so that the term  $M(\frac{d\vec{r}_j}{dt})^2$  representing kinetic energy in (9.5)

is small. As the energy  $H$  is constant, we can assume that  $z_j \approx y_j \approx 0$ , so that the ions are aligned along the  $x$  axis and therefore  $|\vec{r}_k - \vec{r}_j| \approx |x_k - x_j|$  in (9.5). Under these circumstances, the motion in the  $y$ - $z$  plane and the motion along the  $x$  axis of the ions can be separated,<sup>6</sup> and we can study the motion of the ions along the  $x$  axis using the Hamiltonian

$$H_x := \sum_{j=1}^N \left( \frac{1}{2} M \dot{x}_j^2 + \frac{1}{2} M \omega_x^2 x_j^2 \right) + \frac{1}{8\pi\epsilon_0} \sum_{k \neq j} \frac{q_k q_j}{|\vec{x}_k - \vec{x}_j|}. \quad (9.6)$$

An analysis based on classical Hamiltonian mechanics [195] shows that the motion of the ions is a superposition of oscillations about an equilibrium position  $x_1 = x_{1,eq}$ ,  $x_2 = x_{2,eq}, \dots, x_N = x_{N,eq}$ . The frequencies of the various modes of oscillation can be calculated explicitly and they are nearly independent of  $N$ . The lowest frequency is  $\omega_x$ . It corresponds to the frequency of the oscillation of the center of mass of the system of ions. In this mode, the ions oscillate together back and forth. This frequency is also clearly separated from other frequencies so that the effect of an external control tuned at this frequency could be decoupled from the other oscillation modes. These considerations suggest treating the whole system as a unique quantum harmonic oscillator with mass  $NM$  and frequency  $\omega_x$ . In this setting the analysis is not different from the case of a *single* trapped ion except that the mass is  $NM$  rather than  $M$ . A rigorous analysis and justification of this assumption is given in ([108] (section 5)).

### 9.3.3 Quantum mechanical Hamiltonian

With the above approximations, the quantum mechanical system for the trapped ions consists of a harmonic oscillator subsystem corresponding to the motion of the center of mass of the ions and  $N$  subsystems each describing the internal degree of freedom of each ion. The associated Hilbert space will be therefore the tensor product of the Hilbert space for a harmonic oscillator and  $N$  Hilbert spaces for the internal degrees of freedom of the ions. The quantum mechanical Hamiltonian of the system of  $N$  ions under the control of  $N$  electromagnetic fields (lasers) has the form

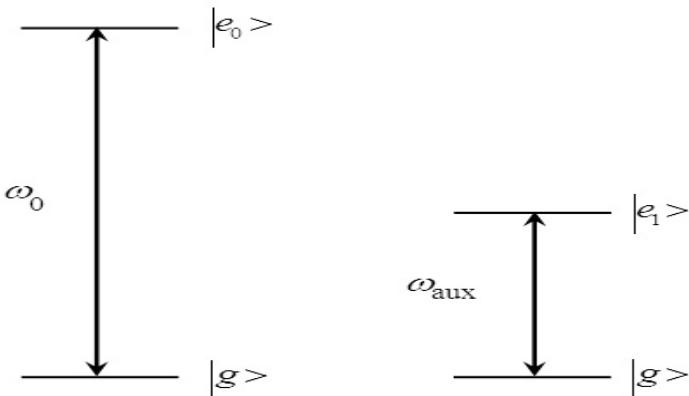
$$H_{TOT} := \sum_{j=1}^N H_{0,j} + H_{ho} + \sum_{j=1}^N H_{I,j}.$$

---

<sup>6</sup>Taking as canonical variables  $q := \{x_j, y_j, z_j\}$ ,  $j = 1, \dots, N$  and canonical momentum  $p := M\dot{q}$ , and applying the Hamilton-Jacobi equations of Hamiltonian dynamics (B.13), (B.14) one easily sees that, to study the motion in the  $x$ -direction, one can take as Hamiltonian  $H_x$  in (9.6).

In this expression,  $H_{0,j}$  is the internal Hamiltonian of the  $j$ -th ion,  $H_{ho}$  is the Hamiltonian corresponding to the center of mass motion modeled as a harmonic oscillator and  $H_{Ij}$  represents the interaction Hamiltonian of the  $j$ -th ion with the external control laser field.

Following the scheme of [48], each ion has three distinguished levels. The levels  $|g\rangle$  and  $|e_0\rangle$  (ground state  $|g\rangle$  and excited state  $|e_0\rangle$ ) are chosen to represent the 0 and 1 value of the qubit, respectively, while an auxiliary level  $|e_1\rangle$  is used in the implementation of quantum logic gates.<sup>7</sup> The energy level diagram for the internal degrees of freedom of a single ion is of the type reported in Figure 9.9.



**FIGURE 9.9:** Energy levels for the internal degrees of freedom of a single ion. In absence of control and coupling it is a three level system with Hamiltonian given in (9.7).

Transitions between two different levels are induced by the controlling laser field as we shall see in the following. The internal Hamiltonian  $H_{0,j}$  for a single ion, using  $\hbar = 1$ , and omitting energy levels which are of no interest, has therefore the form

$$H_{0,j} = \frac{\omega_0}{2}(|e_0\rangle\langle e_0| - |g\rangle\langle g|) + \left(-\frac{\omega_0}{2} + \omega_{aux}\right)|e_1\rangle\langle e_1|, \quad (9.7)$$

for some given transition energies  $\omega_0$  and  $\omega_{aux}$ .

The harmonic oscillator Hamiltonian  $H_{ho}$ , which corresponds to the lowest mode of oscillation where all the ions oscillate together, can be written in

---

<sup>7</sup>Methods to perform quantum logic operations without the auxiliary level exist [147].

terms of the destruction and creation operators  $\hat{a}$  and  $\hat{a}^\dagger$  defined in [subsection 2.1.3](#). Eliminating the shift in energy  $\hbar \frac{\omega_x}{2}$  in (2.44), we can write the quantum harmonic oscillator Hamiltonian as (again,  $\hbar = 1$ )

$$H_{ho} := \omega_x \hat{a}^\dagger \hat{a}.$$

### 9.3.4 Practical implementation of different interaction Hamiltonians

The interaction Hamiltonian for the  $j$ -th ion  $H_{Ij}$  has the form of the electric dipole interaction discussed in 2.2.1.2. Consider only operators concerning the  $j$ -th ion and omit the index  $j$  for simplicity of notation. A control laser field (acting on the  $j$ -th ion) in *standing wave configuration* has the form

$$\vec{E} = \vec{\epsilon} E_0 \sin(kd(t)) \cos(\omega_L t + \phi).$$

In this expression,  $\vec{\epsilon}$  is the polarization vector which gives the direction of the field,  $E_0$  the maximum magnitude of the field and  $k$  is the wave-number given by  $k := \frac{2\pi}{\lambda}$ , with  $\lambda$  the (spatial) wave-length.  $d = d(t)$  is the distance of the ion from the laser source. It depends on time since the ion oscillates about its equilibrium position.  $\omega_L$  is the temporal frequency.

There are two schemes that are used, the one where the equilibrium position of the ion corresponds to the *antinode* (i.e., a point of maximum) of the standing wave and the one where it corresponds to a *node* of the standing wave (i.e., a point where the field is always zero). It is possible to go from a node to an antinode configuration, for example, by changing the inclination of a mirror. The antinode and node configurations are described in [Figures 9.10](#) and [9.11](#), respectively.

In both cases, the distance  $d$  from the laser source is written as

$$d = d_{eq} + \cos(\theta)\delta x,$$

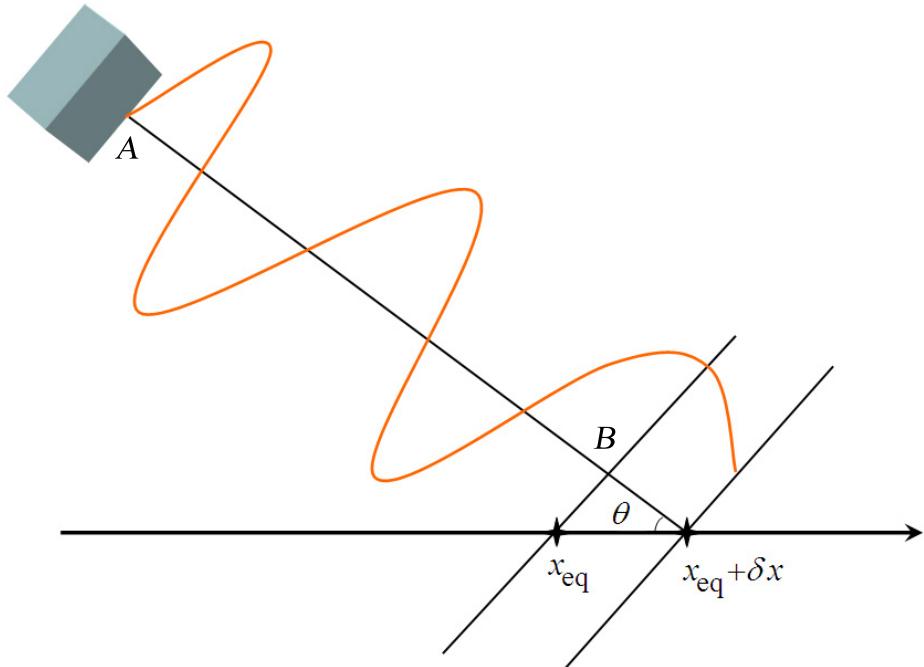
where  $d_{eq}$  is the distance of the equilibrium position from the laser source, while  $\theta$  is the inclination angle of the laser field with respect to the  $x$  axis and  $\delta x$  is the displacement of the ion from its equilibrium position.

In the *antinode configuration*  $d_{eq} = \frac{4l+1}{4}\lambda$ , for some integer  $l$ , so that

$$\begin{aligned} \sin(kd) &= \sin\left(k \frac{4l+1}{4}\lambda + k \cos(\theta)\delta x\right) \\ &= \sin\left(\frac{\pi}{2} + k \cos(\theta)\delta x\right) = \cos(k \cos(\theta)\delta x). \end{aligned}$$

In the so-called *Lamb-Dicke limit* the displacement is much smaller than the wavelength  $\lambda$  and therefore  $|k\delta x| \ll 1$  and omitting terms quadratic and higher in  $k \cos(\theta)\delta x$ , we can write  $\cos(k \cos(\theta)\delta x) \approx 1$ . Therefore, in the

## LASER SOURCE



**FIGURE 9.10:** Scheme of laser control of a single ion in the **antinode** configuration. The distance between the two points  $A$  and  $B$  is  $d_{eq}$ .

antinode configuration, the interaction Hamiltonian  $H_{Ia}$  is obtained by quantization of the following classical Hamiltonian where  $\vec{d}$  represents the electric dipole (cf. the last term in (2.52))

$$\tilde{H}_{Ia} := -\vec{d} \cdot \vec{\epsilon} E_0 \cos(\omega_L t + \phi).$$

In the *node configuration*,  $d_{eq} = l\lambda$  for some integer  $l$ , and therefore

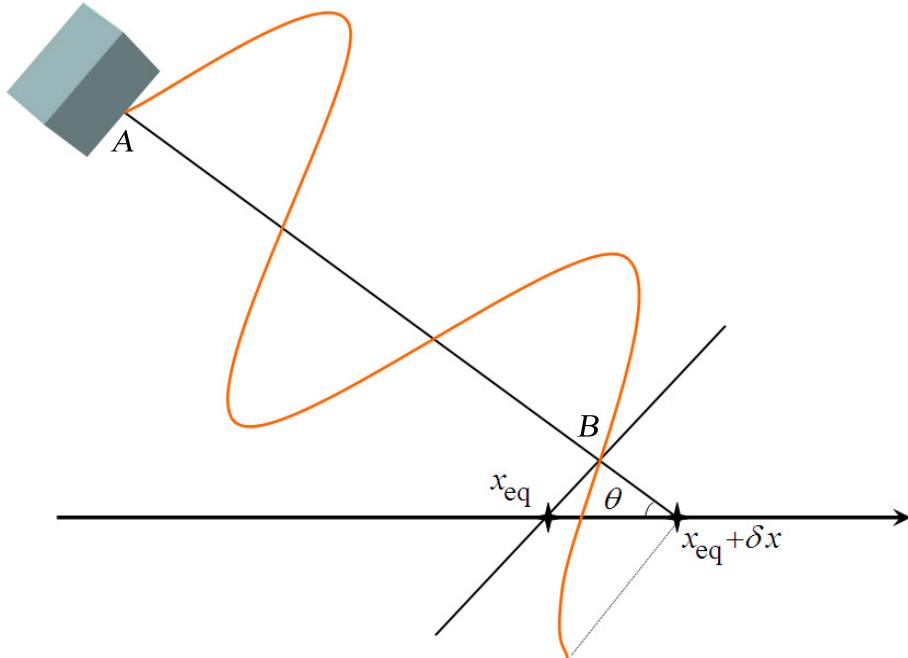
$$\sin(kd) = \sin(kl\lambda + k \cos(\theta)\delta x) = \sin(k \cos(\theta)\delta x).$$

Again, in the Lamb-Dicke limit,  $\sin(k \cos(\theta)\delta x) \approx k \cos(\theta)\delta x$  so that the interaction Hamiltonian is obtained by quantization of the classical Hamiltonian

$$\tilde{H}_{In} := -\vec{d} \cdot \vec{\epsilon} E_0 k \cos(\theta)\delta x \cos(\omega_L t + \phi). \quad (9.8)$$

The quantum mechanical Hamiltonian in the antinode configuration is obtained by denoting by  $\hat{d}_e$  the operator (acting on the Hilbert space of the internal degrees of freedom of the ion) corresponding to the projection of the

## LASER SOURCE



**FIGURE 9.11:** Scheme of laser control of a single ion in the **node** configuration. The distance between the two points  $A$  and  $B$  is  $d_{eq}$ .

dipole vector  $\vec{d}$  along the polarization vector  $\vec{\epsilon}$ .<sup>8</sup> Therefore, in the antinode configuration and with polarization  $\vec{\epsilon}$ , the *total* quantum mechanical Hamiltonian (including the internal, the harmonic oscillator and interaction parts) can be written (for one ion) as

$$\begin{aligned} H_{a,\epsilon} = & \frac{\omega_0}{2}(|e_0\rangle\langle e_0| - |g\rangle\langle g|) + (\omega_{aux} - \frac{\omega_0}{2})|e_1\rangle\langle e_1| + \\ & \omega_x \hat{a}^\dagger \hat{a} - \frac{1}{2} \hat{d}_\epsilon E_0 \cos(\omega_L t + \phi). \end{aligned}$$

This Hamiltonian determines the evolution of one ion when a laser field acts on it.

It is possible to choose the polarization  $\vec{\epsilon}$  so that  $\langle e_1 | \hat{d}_\epsilon | g \rangle = 0$ . In this case, the Hamiltonian  $H_{a,\epsilon}$  is effectively a two level ( $|g\rangle$  and  $|e_0\rangle$ ) Hamiltonian. The third level  $|e_1\rangle$  and the harmonic oscillator levels are decoupled in this case. By appropriately choosing the phase  $\phi$ , the laser frequency  $\omega_L$  and possibly

<sup>8</sup>If  $\vec{\epsilon} := \epsilon_x \vec{i} + \epsilon_y \vec{j} + \epsilon_z \vec{k}$ ,  $\vec{d} = d_x \vec{i} + d_y \vec{j} + d_z \vec{k}$ , and if  $\hat{d}_{x,y,z}$  denote the Hermitian operators corresponding to  $d_x$ ,  $d_y$  and  $d_z$ , then  $\hat{d}_\epsilon := \hat{d}_x \epsilon_x + \hat{d}_y \epsilon_y + \hat{d}_z \epsilon_z$ .

the amplitude  $E_0$ , it is possible to perform any rotation in the space of the qubit spanned by  $|g\rangle$  and  $|e_0\rangle$ .

The quantization of the node configuration Hamiltonian (9.8) is slightly more complicated because one not only introduces the operator  $\hat{d}_\epsilon$  but also replaces the displacement  $\delta x$  with the corresponding quantum mechanical operator  $\hat{\delta}x$ . The operator  $\hat{\delta}x$  can be written in terms of the creation and destruction operators  $\hat{a}^\dagger$  and  $\hat{a}$ . In particular, using (2.41) and (2.42) with  $\hat{\delta}x$  in place of  $\hat{x}$ ,  $NM$  in place of  $m$  and  $\omega_x$  in place of  $\omega$ , with  $\hbar = 1$ , we have

$$\hat{\delta}x := \frac{1}{\sqrt{2NM\omega_x}}(\hat{a} + \hat{a}^\dagger).$$

In the node configuration, and with polarization  $\vec{\epsilon}$ , the *total* quantum mechanical Hamiltonian can be written (for one ion) as

$$\begin{aligned} H_{n,\epsilon} = & \frac{\omega_0}{2}(|e_0\rangle\langle e_0| - |g\rangle\langle g|) + (\omega_{aux} - \frac{\omega_0}{2})|e_1\rangle\langle e_1| + \\ & \omega_x \hat{a}^\dagger \hat{a} - \frac{k \cos(\theta) E_0}{\sqrt{2NM\omega_x}} \hat{d}_\epsilon (\hat{a} + \hat{a}^\dagger) \cos(\omega_L t + \phi). \end{aligned}$$

This Hamiltonian determines the evolution of one ion when a laser field acts on it. With respect to the corresponding Hamiltonian in the antinode configuration the evolution of the ion is now coupled (through the external field) to the harmonic oscillator.

It is possible to choose the polarization  $\vec{\epsilon}$  so that  $\langle e_0 | \hat{d}_\epsilon | g \rangle = 0$  or  $\langle e_1 | \hat{d}_\epsilon | g \rangle = 0$ . Let us consider the second case in more detail. The level  $|e_1\rangle$  is virtually decoupled from the rest of the system so that we write  $H_{n,\epsilon}$  in a simplified form as

$$\begin{aligned} H_{n,\epsilon} = & \frac{\omega_0}{2}(|e_0\rangle\langle e_0| - |g\rangle\langle g|) + \\ & \omega_x \hat{a}^\dagger \hat{a} - \frac{1}{2} \frac{k \cos(\theta) E_0}{\sqrt{2NM\omega_x}} \hat{d}_\epsilon (\hat{a} + \hat{a}^\dagger) (e^{i(\omega_L t + \phi)} + e^{-i(\omega_L t + \phi)}). \end{aligned}$$

To analyze the dynamics under this Hamiltonian, we now perform a change of coordinates  $|\psi\rangle \rightarrow |\tilde{\psi}\rangle = e^{iH_0 t} |\psi\rangle$ , with

$$H_0 := \frac{\omega_0}{2}(|e_0\rangle\langle e_0| - |g\rangle\langle g|) + \omega_x \hat{a}^\dagger \hat{a}$$

and go to the interaction picture (cf. subsection 1.3.2).  $|\tilde{\psi}\rangle$  evolves according to the Hamiltonian  $H'_I(t) := e^{iH_0 t} H_I e^{-iH_0 t}$ , with  $H_I$  given by

$$H_I := -\frac{1}{2} \frac{k \cos(\theta) E_0}{\sqrt{2NM\omega_x}} \hat{d}_\epsilon (\hat{a} + \hat{a}^\dagger) (e^{i(\omega_L t + \phi)} + e^{-i(\omega_L t + \phi)}).$$

We define  $S_0 := |e_0\rangle\langle e_0| - |g\rangle\langle g|$  and calculate

$$H'_I(t) = e^{i\frac{\omega_0}{2}S_0 t} e^{i\omega_x \hat{a}^\dagger \hat{a} t} H_I e^{-i\frac{\omega_0}{2}S_0 t} e^{-i\omega_x \hat{a}^\dagger \hat{a} t}.$$

This gives

$$H'_I(t) = -\frac{1}{2} \frac{k \cos(\theta) E_0}{\sqrt{2NM\omega_x}} e^{i\frac{\omega_0}{2}S_0t} \hat{d}_\epsilon e^{-i\frac{\omega_0}{2}S_0t} \times \\ (e^{i\omega_x \hat{a}^\dagger \hat{a} t} \hat{a} e^{-i\omega_x \hat{a}^\dagger \hat{a} t} + e^{i\omega_x \hat{a}^\dagger \hat{a} t} \hat{a}^\dagger e^{-i\omega_x \hat{a}^\dagger \hat{a} t}) (e^{i(\omega_L t + \phi)} + e^{-i(\omega_L t + \phi)}).$$

Using the fact (see [Exercise 9.1](#)) that

$$e^{i\omega_x \hat{a}^\dagger \hat{a} t} \hat{a} e^{-i\omega_x \hat{a}^\dagger \hat{a} t} = \hat{a} e^{-i\omega_x t}, \quad (9.9)$$

we have

$$H'_I(t) = -\frac{1}{2} \frac{k \cos(\theta) E_0}{\sqrt{2NM\omega_x}} e^{i\frac{\omega_0}{2}S_0t} \hat{d}_\epsilon e^{-i\frac{\omega_0}{2}S_0t} \times \\ (\hat{a} e^{-i\omega_x t} + \hat{a}^\dagger e^{i\omega_x t}) (e^{i\omega_L t} e^{i\phi} + e^{-i\omega_L t} e^{-i\phi}).$$

Given that  $\omega_L > 0$  and  $\omega_x > 0$ , we do a first *rotating wave approximation* (cf. [section 7.1](#)) by neglecting the fast oscillating terms. Moreover, by defining

$$\alpha := \langle e_0 | \hat{d}_\epsilon | g \rangle,$$

we obtain

$$H'_I(t) := -\frac{1}{2} \frac{k \cos(\theta) E_0}{\sqrt{2NM\omega_x}} (\alpha |e_0\rangle\langle g| e^{i\omega_0 t} + \alpha^* |g\rangle\langle e_0| e^{-i\omega_0 t}) \times \\ (\hat{a} e^{i(\omega_L - \omega_x)t} e^{i\phi} + \hat{a}^\dagger e^{-i(\omega_L - \omega_x)t} e^{-i\phi}).$$

At this point, by choosing

$$\omega_L = \omega_x - \omega_0,$$

and doing another rotating wave approximation, we finally obtain the Hamiltonian<sup>9</sup>

$$H'_I(t) = -\frac{1}{2} \frac{k \cos(\theta) E_0}{\sqrt{2NM\omega_x}} (\alpha |e_0\rangle\langle g| \hat{a} e^{-i\phi} + \alpha^* |g\rangle\langle e_0| \hat{a}^\dagger e^{i\phi}). \quad (9.10)$$

Although this is an infinite dimensional Hamiltonian, the two-dimensional subspace spanned by  $|g\rangle|1\rangle$  and  $|e_0\rangle|0\rangle$  is invariant ([Exercise 9.3](#)). Here  $|0\rangle$  and  $|1\rangle$  are the first two eigenstates of the harmonic oscillator corresponding to no phonon and one phonon respectively. Therefore the dynamics given by this Hamiltonian can be studied as a finite dimensional dynamics. In an analogous manner a similar Hamiltonian ([Exercise 9.2](#)) can be obtained for a coupling between  $|e_1\rangle|0\rangle$  and  $|g\rangle|1\rangle$ .

---

<sup>9</sup>Note that the product between  $|e_0\rangle\langle g|$  and  $\hat{a}$  is a tensor product as these two operators act on different spaces.

### 9.3.5 The control problem: Switching between Hamiltonians

Summarizing the above discussion we are able to implement three different types of Hamiltonians:

1. A Hamiltonian in the antinode configuration which provides rotations for a single qubit, i.e., rotations in the space spanned by the eigenvectors  $|e_0\rangle$  and  $|g\rangle$  associated to the internal degree of a ion.
2. A Hamiltonian in the node configuration which couples the levels  $|e_0\rangle$  and  $|g\rangle$  to the harmonic oscillator motion of the center of mass of the ions.
3. A Hamiltonian in the node configuration which couples the levels  $|e_1\rangle$  and  $|g\rangle$ , where  $|e_1\rangle$  is the auxiliary level of the internal degree of one ion, to the harmonic oscillator motion of the center of mass of the ions.

With the above Hamiltonians in the node and antinode configuration and with the various polarizations of the field, it is possible to perform local operations on the qubits (in the antinode case) or entangling operations between the qubits system whose Hilbert space is spanned by  $|e_0\rangle$  and  $|g\rangle$  and the harmonic oscillator system with Hilbert space spanned by  $|0\rangle$  and  $|1\rangle$ . For a two qubit operation between two ions  $j$  and  $m$ , one uses entangling operations first for the  $j$ -th ion and the harmonic oscillator and then between the harmonic oscillator and the  $m$ -th ion. This way, it is possible to perform evolutions corresponding to quantum logic operations between the two quantum bits which, along with single rotations of the qubits, are universal for quantum computation [48].

## 9.4 Notes and References

In this chapter, we have discussed several applications of quantum control and dynamics. These physical examples, along with the modeling considerations in Chapter 2, should provide physical motivation for the analysis presented in the remaining chapters. A survey on physical applications of quantum control along with a long list of references can be found in [139]. This paper also contains a discussion of NMR experiments from the control perspective. An introductory book on spin dynamics and NMR is [134]. Another popular text on NMR is [193], which uses the density matrix approach followed in our description.

Molecular control is treated in the books [174], [192] as well as the review papers [165], [226].

The original proposal for quantum information processing using ion traps was put forward in [48]. After that, several papers have analyzed and improved that scheme. In our treatment, we have greatly benefited from [108] and [133]. The paper [171] presents results on the controllability of the scheme of ion trap quantum information processor, in particular in reference to earlier results on controllability of infinite dimensional quantum mechanical systems [205], [206]. We refer to [155] for a survey of methods to implement quantum information and comparison of the trapped ion scheme with other schemes.

---

## 9.5 Exercises

**Exercise 9.1** Prove formula (9.9) using the commutation relation (2.43).

**Exercise 9.2** Consider the interaction Hamiltonian in the node configuration in [subsection 9.3.4](#) and assume the polarization vector  $\vec{\epsilon}$  is chosen so that  $\langle e_0 | \hat{d}_\epsilon | g \rangle = 0$ . Then obtain the Hamiltonian corresponding to (9.10), coupling the levels  $|e_1\rangle$  and  $|g\rangle$  to the harmonic oscillator.

**Exercise 9.3** Verify the invariance of the vector space spanned by  $|e_0\rangle|0\rangle$  and  $|g\rangle|1\rangle$  under the Hamiltonian (9.10).

# Appendix A

---

## Positive and Completely Positive Maps, Quantum Operations and Generalized Measurement Theory

---

### A.1 Positive and Completely Positive Maps

Consider a finite dimensional<sup>1</sup> Hilbert space  $\mathcal{H}$  and the set of linear operators on  $\mathcal{H} \rightarrow \mathcal{H}$ ,  $\mathcal{B}(\mathcal{H})$ , which is a Hilbert space with the inner product  $\langle A, B \rangle := \text{Tr}(B^\dagger A)$ . A linear operator  $\mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$  is called a *super-operator* although we shall more often use the word *map*.

**Definition A.1.1** A map  $\Gamma : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$  is called **positive** (or **positivity preserving**) if and only if  $A \geq 0$  implies  $\Gamma(A) \geq 0$ .<sup>2</sup>  $\Gamma$  is called  **$n$ -positive** if and only if the map  $\Gamma \otimes \mathbf{1}_n : \mathcal{B}(\mathcal{H} \otimes \mathcal{H}_n) \rightarrow \mathcal{B}(\mathcal{H} \otimes \mathcal{H}_n)$ , where  $\mathbf{1}_n$  is the identity map over the  $n$ -dimensional Hilbert space  $\mathcal{H}_n$ , is positive. A map  $\Gamma$  is **completely positive** if it is  $n$ -positive for any positive integer  $n$ .

In particular notice that 1-positivity is equivalent to positivity and therefore clearly a completely positive map is also positive. A simple example of a completely positive map is the unitary evolution  $\Gamma(\rho_S) := U\rho_SU^\dagger$ . In fact we have for any  $\rho \geq 0$ ,  $\rho \in \mathcal{B}(\mathcal{H} \otimes \mathcal{H}_n)$  and any  $n$ ,  $(\Gamma \otimes \mathbf{1}_n)(\rho) = (U \otimes \mathbf{1}_n)\rho(U^\dagger \otimes \mathbf{1}_n) \geq 0$ .

Historically positive and complete positive maps were introduced in quantum mechanics in the study of the dynamics of open systems. When a system evolves interacting with the environment, a ‘reduced’ dynamics is specified for the system [12],[33]. This dynamics is different from the unitary evolution of closed systems studied in this book. It is a one parameter ( $t \in \mathbb{R}$ ) family of maps  $\{\Gamma_t\}$  acting on the set of density matrices  $\rho$ . Each map necessarily has to be positive since, if  $\rho$  is a density matrix,  $\Gamma_t(\rho)$  has to be another density matrix, in particular it must have nonnegative eigenvalues. However

---

<sup>1</sup>These definitions can be extended to the infinite dimensional case (see, e.g., [176]).

<sup>2</sup> $A \geq 0$  means that  $\langle \psi | A | \psi \rangle \geq 0$ , for all  $|\psi\rangle$  in  $\mathcal{H}$ , or, equivalently, that all the eigenvalues of  $A$  are nonnegative.

sometimes the stronger requirement of complete positivity is imposed so that the system coupled with an auxiliary system (e.g., an environment), which has no dynamics, still gives a density matrix as a result of the evolution. This topic is discussed in depth in the thesis [176] both in theoretical terms and with physical examples.

The state of quantum systems is modified not only by evolutions but also by measurements. All the possible physical transformations on a state  $\rho$  of a quantum system can be described introducing, in an abstract manner, the concept of quantum operations.

---

## A.2 Quantum Operations and Operator Sum Representation

**Definition A.2.1** A **quantum operation**  $\Gamma$  is a linear map  $\Gamma : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$ , which satisfies the following axioms:

1.  $\Gamma$  is *trace-nonincreasing*, i.e.,

$$\text{Tr}(\Gamma(\rho)) \leq \text{Tr}(\rho).$$

2.  $\Gamma$  is *completely positive*.

Examples of quantum operations are the unitary evolution  $\rho \rightarrow X\rho X^\dagger$  for  $X \in U(n)$ , or the transformation (cf. (1.36)) due to a Von Neumann-Lüders measurement with result  $j$ ,

$$\rho \rightarrow P_j \rho P_j,$$

where  $P_j$  is the associated projection. Another operation is obtained by letting the system (of dimension  $n_S$ ) in state  $\rho$  first couple with an auxiliary system (of dimension  $n_P$ ) in state  $\sigma$  and then evolve according to a unitary evolution  $X \in U(n_S n_P)$ . Then the state of the main system is extracted via the partial trace operation with respect to the second system  $P$ . Therefore, the operation  $\Gamma$  is

$$\Gamma : \rho \rightarrow \text{Tr}_P(X\rho \otimes \sigma X^\dagger).$$

**Remark A.2.2** It is possible to extend Definition A.2.1 to maps  $\Gamma$  between two different spaces, that is  $\Gamma : \mathcal{B}(\mathcal{H}_1) \rightarrow \mathcal{B}(\mathcal{H}_2)$  with  $\mathcal{H}_1 \neq \mathcal{H}_2$ . In this case, one more example of operation  $\Gamma$  is the addition of an uncorrelated system, i.e.,

$$\Gamma : \rho \rightarrow \rho \otimes \sigma.$$

A general and very useful *representation of operations* is given by the **Kraus operator sum representation theorem**.

**Theorem A.2.3** ([127], [155]/(Theorem 8.1)) *Every quantum operation  $\Gamma$  as defined in Definition A.2.1 can be written as*

$$\Gamma(\rho) := \sum_k \Omega_k \rho \Omega_k^\dagger, \quad (\text{A.1})$$

for a countable set of operators  $\Omega_k$  on  $\mathcal{H}$  satisfying<sup>3</sup>

$$\sum_k \Omega_k \Omega_k^\dagger \leq \mathbf{1}.$$

### A.3 Generalized Measurement Theory

As it is mentioned at the beginning of section 1.2, there exist several types of measurement and the Von Neumann-Lüders measurement is just one special case which however describes correctly most experimental situations. The formalism of operations allows us a unified treatment of measurements which goes under the name of **generalized measurement theory**. Central to this theory are the concepts of operations and effects.

Given a measurement scheme and a measurable set of possible outcomes  $\mathcal{M}$ , with every result  $m \in \mathcal{M}$  is associated a positive operator  $F_m$ , called an **effect**. If  $\rho$  is the current state of the system, the probability of obtaining the result  $m$  (or of an event  $m$  to occur) is

$$\Pr(m) = \text{Tr}(F_m \rho).$$

After a result  $m$  (or, more generally an event  $m$ ) has occurred, the state is modified according to

$$\rho \rightarrow \Pr(m)^{-1} \Phi_m(\rho).$$

The positive maps  $\Phi_m$  are an example of *operations* and  $\text{Tr}(\Phi_m(\rho)) = \Pr(m) = \text{Tr}(F_m \rho)$ . They can be expressed in operator sum representation according to Theorem A.2.3. The Von Neumann-Lüders measurement with a discrete set of outcomes is a special case where the effects are given by the projections  $P_m$  and the associated operations  $\Phi_m$  are given by  $\Phi_m(\rho) := P_m \rho P_m$ .

Another example is the *indirect Von Neumann-Lüders measurement*, where a system  $S$  with state  $\rho$  is first coupled with an auxiliary system  $P$  in state  $\sigma$ . Then the whole system evolves with evolution  $X$  and an observable is

---

<sup>3</sup>Recall that, for two operators  $A$  and  $B$ ,  $A \leq B$  means that  $B - A$  is a positive operator.

measured on  $S$ . Assume this measurement gives a result  $m$  whose associated projection is  $P_m$ . The operation  $\Gamma$  is given by

$$\Gamma(\rho) := \text{Tr}_P (P_m \otimes \mathbf{1} X \rho \otimes \sigma X^\dagger P_m \otimes \mathbf{1}). \quad (\text{A.2})$$

To obtain the operation sum representation of this operation, we use the property (8.5) of the partial trace along with its definition 8.1.2. Given an orthonormal basis  $\{|p_k\rangle\}$ ,  $k = 1, \dots, n_P$  of the Hilbert space associated to system  $P$ , and writing  $\sigma$  as

$$\sigma := \sum_{l=1}^{n_P} \alpha_l |p_l\rangle \langle p_l|, \quad \alpha_l \geq 0,$$

we can write

$$\Gamma(\rho) := P_m \left( \sum_{k=1}^{n_P} \sum_{l=1}^{n_P} \alpha_l \mathbf{1}_{n_S} \otimes \langle p_k | X \rho \otimes |p_l\rangle \langle p_l | X^\dagger \mathbf{1}_{n_S} \otimes |p_k\rangle \right) P_m,$$

which we can write

$$\begin{aligned} & \Gamma(\rho) \\ &= P_m \left( \sum_{k=1}^{n_P} \sum_{l=1}^{n_P} \alpha_l \mathbf{1}_{n_S} \otimes \langle p_k | X \mathbf{1}_{n_S} \otimes |p_l\rangle (\rho \otimes \mathbf{1}_1) \mathbf{1}_{n_S} \otimes \langle p_l | X^\dagger \mathbf{1}_{n_S} \otimes |p_k\rangle \right) P_m. \end{aligned}$$

Since  $\rho \otimes \mathbf{1}_1 = \rho$ , the operator sum representation (A.2.1) holds for  $\Gamma$  with the  $\Omega_k$ 's given by the  $\Omega_{kl}$

$$\Omega_{kl} := P_m \sqrt{\alpha_l} \mathbf{1}_{n_S} \otimes \langle p_k | X \mathbf{1}_{n_S} \otimes |p_l\rangle.$$

For more details on the general measurement theory in terms of effects and operations we refer to [33].

# Appendix B

---

## Lagrangian and Hamiltonian Formalism in Classical Electrodynamics

---

### B.1 Lagrangian Mechanics

#### Euler-Lagrange equations

Consider a system with  $n$  degrees of freedom  $x_1, \dots, x_n$ . The state of the system at every instant and its future evolution is determined by  $x_1, \dots, x_n$  and the time derivatives  $\dot{x}_1, \dots, \dot{x}_n$ . The *equations of motion* are differential equations involving  $x_1, \dots, x_n$ . From their integration it is possible to obtain the state of the system at every time.

In **Lagrangian mechanics** (see, e.g., [131]) one derives the equations of motion from a function  $L := L(x_1, \dots, x_n, \dot{x}_1, \dots, \dot{x}_n, t)$  called the **Lagrangian**. The *principle of least action* states that the real path connecting initial  $(x_j(t_0))$  and final  $(x_j(t_f))$ ,  $j = 1, \dots, n$ , configuration is an extremum for the integral

$$S := S(x_j) := \int_{t_0}^{t_f} L(x_j(t), \dot{x}_j(t), t) dt. \quad (\text{B.1})$$

The integral  $S$  is called the *action*. The fact that  $x_j$  is an extremum means that

$$\frac{d}{d\epsilon} S(x_j + \epsilon \delta x_j) \Big|_{\epsilon=0} = 0,$$

for every admissible perturbation of the trajectory  $\delta x_j$  with  $\delta x_j(t_0) = \delta x_j(t_f) = 0$ . From this variational principle, one obtains<sup>1</sup> that the true trajectory  $x_j$ ,

---

<sup>1</sup>Consider

$$\frac{d}{d\epsilon} S(x_j + \epsilon \delta x_j) = \int_{t_0}^{t_f} \frac{d}{d\epsilon} L(x_j + \epsilon \delta x_j, \dot{x}_j + \epsilon \delta \dot{x}_j, t) dt.$$

Applying the chain rule we obtain

$$\frac{d}{d\epsilon} S(x_j + \epsilon \delta x_j) = \int_{t_0}^{t_f} L_x \delta x + L_{\dot{x}} \delta \dot{x} dt.$$

$j = 1, \dots, n$ , has to satisfy the **Euler-Lagrange equations**

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}_j} = \frac{\partial L}{\partial x_j}, \quad j = 1, \dots, n. \quad (\text{B.2})$$

The choice of the Lagrangian is crucial for the correct determination of the equations of motion. For a system of  $N$  particles, let the positions  $\vec{r}_j$ ,  $j = 1, \dots, N$ , be functions of the *independent* degrees of freedom  $x_1, \dots, x_n$  and time  $t$ . The  $l$ -th component of the *generalized force*,  $Q_l$ ,  $l = 1, \dots, n$ , is defined as

$$Q_l := \sum_{j=1}^N \vec{F}_j \cdot \frac{\partial \vec{r}_j}{\partial x_l},$$

where  $\vec{F}_j$  is the force applied to the  $j$ -th particle. If there exists a function  $V = V(x_j, \dot{x}_j)$  such that

$$Q_l := -\frac{\partial V}{\partial x_l} + \frac{d}{dt} \frac{\partial V}{\partial \dot{x}_l}, \quad (\text{B.3})$$

then the Lagrangian  $L$  can be chosen as

$$L = L(x_j, \dot{x}_j, t) := T - V, \quad (\text{B.4})$$

where  $T$  is the kinetic energy of the system of particles and  $V$  is the same as in (B.3) and it is called *generalized potential*. A special but important case is when  $V$  does not depend on  $\dot{x}_j$ . In this case, it is called *potential* and the generalized forces  $Q_l$ ,  $l = 1, \dots, n$ , are given by

$$Q_l := -\frac{\partial V}{\partial x_l}.$$

The general situation of equation (B.3) is however relevant in several situations as in the following example in electrodynamics [86], [185].

**Example B.1.1** Consider a particle with charge  $q$  moving in an electromagnetic field. As there are no constraints to the motion of the particle, the degrees of freedom  $x_1, x_2, x_3$  are taken equal to the components of the position vector  $\vec{r}$ . The generalized forces  $Q_1, Q_2, Q_3$  are taken to be the components of the applied force  $\vec{F} := [F_1, F_2, F_3]^T$  which is equal to the Lorentz force in (2.10). The Lorentz force (2.10) can be derived as in (B.3), i.e.,

$$F_j = \frac{\partial V}{\partial x_j} + \frac{d}{dt} \frac{\partial V}{\partial \dot{x}_j}, \quad j = 1, 2, 3, \quad (\text{B.5})$$

---

Using integration by part on the second term in the integral we have

$$\frac{d}{d\epsilon} S(x_j + \epsilon \delta x_j) = [L_{\dot{x}} \delta x]_{t_0}^{t_f} + \int_{t_0}^{t_f} L_x \delta x - \left( \frac{d}{dt} L_{\dot{x}} \right) \delta x dt.$$

Since  $\delta x(t_0) = \delta x(t_f) = 0$  the term outside the integral is zero. Moreover, since the integral has to be zero for every  $\delta x$ , one obtains (B.2).

if we use ( $\vec{v} := \dot{\vec{r}}$ )

$$V := q\phi - q\vec{A} \cdot \vec{v}, \quad (\text{B.6})$$

where  $\phi$  and  $\vec{A}$  are the scalar and vector potential, respectively.<sup>2</sup>

Therefore, the Lagrangian for a particle in an electromagnetic field can be written as

$$L := T - V = \frac{1}{2}m\vec{v} \cdot \vec{v} - q\phi + q\vec{A} \cdot \vec{v}.$$

**Remark B.1.2** The Lagrangian  $L$  giving the correct equations of motion is *not* unique. Given a Lagrangian  $L$ , consider a function  $L'$  obtained by adding to  $L$  the total derivative of a function  $f$  of  $x_j$  and  $t$ , i.e.,

$$L' := L + \frac{d}{dt}f(x_j(t), t) = L + \sum_j \frac{\partial f}{\partial x_j} \dot{x}_j + \frac{\partial f}{\partial t}. \quad (\text{B.9})$$

The values of the action  $S$  in (B.1) calculated for  $L'$  and  $L$  differ only by a constant term independent of the trajectory. Therefore the trajectories corresponding to the extrema in the two cases are the same. Equivalently one can verify that if a trajectory  $x_j$  is such that Euler-Lagrange equations (B.2) are verified with  $L$  then these equations are verified with  $L'$  given in (B.9).

## Conjugate variables, Hamiltonian and Hamilton-Jacobi equations

Given the Lagrangian of a particular problem with degrees of freedom  $x_1, \dots, x_n$ , one defines a *conjugate canonical momentum*  $\vec{p} := [p_1, \dots, p_n]^T$  by

$$p_j := \frac{\partial L}{\partial \dot{x}_j}, \quad j = 1, \dots, n. \quad (\text{B.10})$$

---

<sup>2</sup>This can be easily verified for each component of  $\vec{F}$  (cf. [86] Chpt. 1). For example, for the first component  $F_1$ , we have from (2.10)

$$F_1 = qE_1 + q(\vec{v} \times \vec{B})_1, \quad (\text{B.7})$$

while calculating  $F_1$  using (B.5) and (B.6), we obtain

$$F_1 = -q\frac{\partial \phi}{\partial x_1} - q\frac{dA_1}{dt} - q\frac{\partial}{\partial x_1}(\vec{A} \cdot \vec{v}). \quad (\text{B.8})$$

Expanding the second term on the right hand side, we obtain

$$F_1 = -q\frac{\partial \phi}{\partial x_1} - q\frac{\partial A_1}{\partial t} - q \sum_{j=1,2,3} \frac{\partial A_1}{\partial x_j} \dot{x}_j - q\frac{\partial}{\partial x_1}(\vec{A} \cdot \vec{v}).$$

Using (2.15), the first two terms are equal to  $qE_1$  while a straightforward calculation shows that

$$q \sum_{j=1,2,3} \frac{\partial A_1}{\partial x_j} \dot{x}_j - q\frac{\partial}{\partial x_1}(\vec{A} \cdot \vec{v}) = (\vec{v} \times (\nabla \times \vec{A}))_1$$

which along with (2.12) shows that  $F_1$  in (B.8) is equal to  $F_1$  in (B.7).

With this definition, the Euler-Lagrange equations (B.2) can be written as

$$\frac{d}{dt} p_j = \frac{\partial L}{\partial x_j}, \quad j = 1, \dots, n. \quad (\text{B.11})$$

With the introduction of the conjugate momentum  $\vec{p}$ , the Euler Lagrange equations, which are  $n$  second order differential equations, can be transformed into a set of  $2n$  first order ordinary differential equations. One introduces the **Hamiltonian function**

$$H := \sum_{j=1}^n \dot{x}_j p_j - L(x_j, \dot{x}_j, t). \quad (\text{B.12})$$

By taking the partial derivative with respect to  $x_j$  of both sides of equation (B.12) and using (B.11), we obtain

$$\dot{p}_j = -\frac{\partial H}{\partial x_j}, \quad j = 1, \dots, n. \quad (\text{B.13})$$

Taking the partial derivative with respect to  $p_j$  in both sides of (B.12), we obtain

$$\dot{x}_j = \frac{\partial H}{\partial p_j}, \quad j = 1, \dots, n. \quad (\text{B.14})$$

Equations (B.13),(B.14) are the **Hamilton-Jacobi equations**.

**Remark B.1.3** Refer to the situation of Remark B.1.2. If  $\vec{p}$  is the conjugate momentum for the Lagrangian  $L$ , the conjugate momentum for the Lagrangian  $L'$  is, using (B.9),

$$p'_j = \frac{\partial L'}{\partial \dot{x}_j} = \frac{\partial L}{\partial \dot{x}_j} + \frac{\partial f}{\partial x_j} = p_j + \frac{\partial f}{\partial x_j}, \quad j = 1, \dots, n. \quad (\text{B.15})$$

The corresponding Hamiltonian  $H'$  is by definition (B.12)

$$H' := \sum_{j=1}^n \dot{x}_j p'_j - L' = \sum_{j=1}^N \dot{x}_j (p_j + \frac{\partial f}{\partial x_j}) - L - \sum_{j=1}^N \frac{\partial f}{\partial x_j} \dot{x}_j - \frac{\partial f}{\partial t}.$$

This formula shows that, if  $f$  does not depend explicitly on time, we have

$$H'(x_j, p_j + \frac{\partial f}{\partial x_j}) = H(x_j, p_j),$$

i.e., by replacing the definition (B.15) in the Hamiltonian  $H'$  we obtain the Hamiltonian  $H$ .

In the case where the Hamiltonian  $H = H(x_j, p_j)$  does not depend explicitly on time, we have, using (B.13) and (B.14),

$$\frac{d}{dt}H(x_j, p_j) = \sum_j \frac{\partial H}{\partial x_j} \dot{x}_j + \sum_j \frac{\partial H}{\partial p_j} \dot{p}_j = \sum_j \frac{\partial H}{\partial x_j} \frac{\partial H}{\partial p_j} - \sum_j \frac{\partial H}{\partial p_j} \frac{\partial H}{\partial x_j} = 0.$$

Therefore  $H$  is a constant of motion which is usually equal to the *energy* of the system.

**Example B.1.4** Consider a set of particles with coordinates  $x_1, \dots, x_n$  in a potential  $V = V(x_1, \dots, x_n)$ . The Lagrangian can be written as

$$L = T - V = \frac{1}{2} \sum_{j=1}^n m_j \dot{x}_j^2 - V,$$

where  $m_j$  is the mass of the particle with coordinate  $x_j$ . The conjugate momentum from equation (B.10) coincides with the *kinematical momentum*, i.e.,  $p_j = m_j \dot{x}_j$ . Using the definition of Hamiltonian (B.12) we have  $H = \frac{1}{2} \sum_j m_j \dot{x}_j^2 + V$  namely the Hamiltonian is the sum of the kinetic and potential energy.

There are however examples of Lagrangian functions which give the correct equations of motions but whose corresponding Hamiltonian is not equal to the energy (see Exercise 1 in Complement  $D_{II}$  in [50]).

**Remark B.1.5** The optimal control problem of Lagrange, with no bound on the control, can be seen as the problem to find vector functions  $x$  and  $u$  which minimize

$$J(u) = \int_0^T L(x, u, t) dt,$$

subject to the differential constraint

$$\dot{x} = f(x, u).$$

If  $\{x, u\}$  is an optimal then there exists a vector function  $\lambda = \lambda(t)$  of *Lagrange multipliers* such that  $\{\lambda, x, u\}$  is an optimal for the unconstrained problem of minimizing

$$J'(u) = \int_0^T L + \lambda^T (\dot{x} - f) dt.$$

The augmented problem has Lagrangian

$$L' := L + \lambda^T (\dot{x} - f).$$

The Euler-Lagrange equations give

$$L'_\lambda = \dot{x}^T - f^T = 0,$$

$$\frac{d}{dt}L'_x = \dot{\lambda}^T = L_x - \lambda^T f_x,$$

and

$$L'_u = L_u - \lambda^T f_u = 0,$$

where the last two correspond to (6.27) (with  $\mu = -1$ ) and minimization with respect to  $u$  of  $h$  in (6.28). In particular, only the canonical moment conjugate to  $x$  is not identically zero and it is equal to  $L'_x = \lambda$ . Using the definition (B.12), we obtain

$$H = \lambda^T \dot{x} - L' = \lambda^T f - L,$$

which is equal to the optimal control Hamiltonian  $h$  in (6.28). More on a unified study of Lagrangian mechanics, optimal control and calculus of variations can be found in textbooks such as [81], [112], [135].

---

## B.2 Extension of Lagrangian Mechanics to Systems with Infinite Degrees of Freedom

An electromagnetic field is specified by the value of the field at any point in space. The degrees of freedom of this system are uncountably infinite being given by the values of the field at every point in space. The coupled system of particles and field is what interests us. However, in this section, we outline the general formalism. Our treatment in this and the following two sections follows at times [216] and is largely based on [50] to which we refer for a complete treatment.

### Euler-Lagrange equations

In the case of an infinite set of variables, the generalized coordinates (replacing  $x_1, \dots, x_n$  above) will be denoted by  $X(j, \vec{r}, t)$ . Here  $j$  varies on a countable set, which we shall assume finite, and  $\vec{r}$  varies in an uncountable set. In view of the application to electrodynamics we shall assume this set to be the space  $\mathbf{R}^3$  so that  $\vec{r}$  denotes a point in space. In  $X(j, \vec{r}, t)$ ,  $j$  and  $\vec{r}$  have to be seen as a couple of indexes one varying in a discrete set and one (uncountable) varying in a continuous set  $\mathbf{R}^3$ .

In analogy with what was done for the case of finite degrees of freedom, we consider a *Lagrangian*  $L$  depending on all of the  $X(j, \vec{r}, t)$ 's, their (partial) derivatives with respect to time  $\frac{\partial}{\partial t}X(j, \vec{r}, t)$  and possibly explicitly on time, i.e.,

$$L := L \left( X(j, \vec{r}, t), \frac{\partial}{\partial t}X(j, \vec{r}, t), t \right).$$

Given a path, i.e., specified for every  $X(j, \vec{r}, \cdot)$  the functional dependence on the third argument  $t$ ,  $L$  will be a function of  $t$  only which we can integrate between initial and final time to obtain the *action* (cf. (B.1))

$$S := \int_{t_0}^{t_f} L \left( X(j, \vec{r}, t), \frac{\partial}{\partial t} X(j, \vec{r}, t), t \right) dt.$$

The principle of least action still holds true for a system with an infinite number of degrees of freedom. It states that the true path will be an extremum for the action.

In applications to electrodynamics, it is sufficient to restrict ourselves to Lagrangian functions which satisfy the following conditions.

1. The Lagrangian  $L$  has the partial derivatives with respect to the spatial degrees of freedom explicitly appearing as arguments so that

$$L := L \left( X(j, \vec{r}, t), \frac{\partial X(j, \vec{r}, t)}{\partial t}, \frac{\partial X(j, \vec{r}, t)}{\partial(x, y, z)}, t \right).$$

These partial derivatives are not new independent variables as they can be derived from  $X(j, \vec{r}, t)$  but we only assume that they appear explicitly in the expression of the Lagrangian.

2. The Lagrangian can be written as the triple integral of a *Lagrangian density function*

$$\mathcal{L} := \mathcal{L} \left( X(j, \vec{r}, t), \frac{\partial X(j, \vec{r}, t)}{\partial t}, \frac{\partial X(j, \vec{r}, t)}{\partial(x, y, z)}, \vec{r}, t \right),$$

i.e.,

$$L = \int_{\mathbf{R}^3} \mathcal{L} d\vec{r}.$$

Moreover, it is assumed that the variables  $X(j, \vec{r}, t)$ , which in electrodynamics represent potentials, vanish as  $|\vec{r}| \rightarrow \infty$ .

Under these assumptions, from the principle of least action, one can derive the equations corresponding to the Euler-Lagrange equations (B.2) [50].<sup>3</sup> Define  $X_j := X_j(\vec{r}, t) := X(j, \vec{r}, t)$ ,  $\dot{X}_j := \dot{X}_j(\vec{r}, t) := \frac{\partial X(j, \vec{r}, t)}{\partial t}$ ,  $X_{j(x,y,z)} := X_{j(x,y,z)}(\vec{r}, t) := \frac{\partial}{\partial(x, y, z)} X(j, \vec{r}, t)$ , so that

$$\mathcal{L} := \mathcal{L} \left( X_j, \dot{X}_j, X_{jx}, X_{jy}, X_{jz}, \vec{r}, t \right),$$

---

<sup>3</sup>The derivation follows the same steps as that of Euler-Lagrange equation (B.2) except that one needs to introduce the functional derivative, in order to deal with functions of an uncountable number of variables (cf. [50]).

with  $j = 1, \dots, n$ , the Euler-Lagrange equations are

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{X}_j} = \frac{\partial \mathcal{L}}{\partial X_j} - \sum_{l=x,y,z} \frac{\partial}{\partial l} \frac{\partial \mathcal{L}}{\partial (X_{jl})}. \quad (\text{B.16})$$

**Remark B.2.1** A property analogous to the one discussed in Remark B.1.2 holds in the case of infinite degrees of freedom. We can modify the Lagrangian density by adding the total derivative of a function  $f = f(X_j, \vec{r}, t)$  without modifying the resulting equations of motion (B.16) and the corresponding trajectory. A new Lagrangian density  $\mathcal{L}'$  is obtained from  $\mathcal{L}$  as

$$\mathcal{L}' = \mathcal{L} + \frac{d}{dt} f(X_j, \vec{r}, t).$$

### Conjugate variables, Hamiltonian and Hamilton-Jacobi equations

As in the discrete case, one associates to every degree of freedom  $X_j$  a conjugate momentum  $P_j := P_j(\vec{r}, t) := P(j, \vec{r}, t)$  which is defined using the Lagrangian density function as

$$P_j := \frac{\partial \mathcal{L}}{\partial \dot{X}_j}, \quad j = 1, \dots, n. \quad (\text{B.17})$$

With this definition, the Euler-Lagrange equations (B.16) are written as

$$\frac{d}{dt} P_j = \frac{\partial \mathcal{L}}{\partial X_j} - \sum_{l=x,y,z} \frac{\partial}{\partial l} \frac{\partial \mathcal{L}}{\partial (X_{jl})}, \quad j = 1, \dots, n. \quad (\text{B.18})$$

The derivative  $\frac{d}{dt}$  is written as a total derivative to emphasize that in the equations (B.18)  $j$  and  $\vec{r}$  need to be treated as fixed indexes.

To extend the definition of Hamiltonian function (B.12) we need to ‘sum’ over all the indexes  $j$  and  $\vec{r}$ , the products of  $P_j = P(j, \vec{r}, t)$  and  $\dot{X}_j := \frac{\partial}{\partial t} X(j, \vec{r}, t)$ . Since  $\vec{r}$  is a continuous index, we replace the sum with an integral and define

$$H := \int_{\mathbf{R}^3} \sum_j P_j \dot{X}_j d\vec{r} - L = \int_{\mathbf{R}^3} \sum_j P_j \dot{X}_j - \mathcal{L} d\vec{r}. \quad (\text{B.19})$$

This suggests to define an *Hamiltonian density*

$$\mathcal{H} := \sum_j P_j \dot{X}_j - \mathcal{L}.$$

By taking the partial derivative of both sides with respect to  $X_j$  and using (B.18), we obtain

$$\frac{d}{dt} P_j = -\frac{\partial \mathcal{H}}{\partial X_j} + \sum_{l=x,y,z} \frac{\partial}{\partial l} \frac{\partial \mathcal{H}}{\partial (X_{jl})}, \quad j = 1, \dots, n. \quad (\text{B.20})$$

This equation corresponds to the first Hamilton-Jacobi equation (B.13). By differentiating the definition of the Hamiltonian (B.19) with respect to  $P_j$ , one obtains the Hamilton-Jacobi equation corresponding to (B.14), namely

$$\frac{d}{dt}X_j = \frac{\partial \mathcal{H}}{\partial P_j}, \quad j = 1, \dots, n. \quad (\text{B.21})$$

Equations (B.20) and (B.21) are the Hamilton-Jacobi equations for a system with an infinite uncountable number of degrees of freedom. They are not formally identical to (B.13) and (B.14). This can be obtained by introducing the notion of functional derivative (see [50]).

**Remark B.2.2** In this and the previous section we have assumed the dynamical variables (indexed by a discrete or a continuous set) used to describe the system form an *independent* set of variables. The case of dependent and redundant variables appears naturally in electrodynamics and will be further discussed in the following Remarks B.3.1, B.3.2.

### B.3 Lagrangian and Hamiltonian Mechanics for a System of Interacting Particles and Field

Consider a system of  $N$  particles in an electromagnetic field in free space. This system is a combination of the two types of systems in the previous two sections, i.e., a system with a discrete set of degrees of freedom (the positions of the particles) and a system with a continuous set of degrees of freedom (the value of the field at every point in space). We choose a set of  $3N$  variables given by the positions  $\vec{r}_j$ , of the  $j$ -th particle  $j = 1, \dots, N$  for the degrees of freedom of the particles. Moreover the fact that in the equation of motion (2.17) the second derivative with respect to time appears just as in (B.16) suggests to use the potentials  $A_{x,y,z}(\vec{r})$ , and  $\phi(\vec{r})$  as the dynamical variables of the field. The Lagrangian associated to the system has to be such that the Euler-Lagrange equations (B.2) and (B.16) give the correct equations of motion which are the Lorentz equation (2.10) (with  $\vec{B}$  and  $\vec{E}$  given by (2.12) and (2.15)) and the equations (2.16), (2.17).

#### Lagrangian and equations of motion

In terms of the variables  $A_{x,y,z}$  and  $\phi$  the correct Lagrangian is given by

$$L = \sum_{j=1}^N \frac{1}{2} m_j \dot{\vec{r}}_j^2 + \int_{\mathbf{R}^3} \mathcal{L} d\vec{r}, \quad (\text{B.22})$$

where the Lagrangian density  $\mathcal{L}$  is given by

$$\mathcal{L} := \frac{\epsilon_0}{2} \left( \left( \nabla \phi + \frac{\partial \vec{A}}{\partial t} \right)^2 - c^2 (\nabla \times \vec{A})^2 \right) + \vec{J}(\vec{r}) \cdot \vec{A}(\vec{r}) - \rho(\vec{r})\phi(\vec{r}),$$

where  $J(\vec{r})$  and  $\rho(\vec{r})$  depend on the coordinates  $\vec{r}_j$  because of (2.8) and (2.9). Notice that this Lagrangian density satisfies the conditions set at the beginning of the previous section. In particular, it is an explicit function of the partial spatial derivatives of the vector potential  $\vec{A}$ .

To show that this Lagrangian gives the correct equations of motion, let us consider the Euler-Lagrange equations corresponding to the discrete variables  $\vec{r}_j$ . To simplify notations, fix  $j$  and consider the equation corresponding to the  $x$  coordinate of the  $j$ -th particle,  $x_j$ . Using (2.8) and (2.9) we can write  $\int_{\mathbf{R}^3} \mathcal{L} d\vec{r}$  as

$$\begin{aligned} \int_{\mathbf{R}^3} \mathcal{L} d\vec{r} &= \sum_{j=1}^N (q_j \dot{\vec{r}}_j \cdot \vec{A}(\vec{r}_j) - q_j \phi(\vec{r}_j)) \\ &\quad + \frac{\epsilon_0}{2} \int_{\mathbf{R}^3} \left( \left( \nabla \phi + \frac{\partial \vec{A}}{\partial t} \right)^2 - c^2 (\nabla \times \vec{A})^2 \right) d\vec{r}. \end{aligned} \quad (\text{B.23})$$

Noticing that the integral in the expression on the right hand side does not depend on  $\vec{r}_j$ , nor on  $\vec{r}_j$ , and using (B.22), the right hand side of (B.2) is

$$\frac{\partial L}{\partial x_j} = q_j \dot{\vec{r}}_j \cdot \frac{\partial \vec{A}(\vec{r}_j)}{\partial x_j} - q_j \frac{\partial \phi(\vec{r}_j)}{\partial x_j},$$

while the left hand side gives

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}_j} = \frac{d}{dt} (m_j \dot{x}_j + q_j A_x(\vec{r}_j)) = m_j \ddot{x}_j + q_j \nabla A_x(\vec{r}_j) \cdot \dot{\vec{r}}_j + q_j \frac{\partial A_x}{\partial t}(\vec{r}_j).$$

Equating these two expressions, we obtain

$$m_j \ddot{x}_j = -q_j \left( \frac{\partial}{\partial x_j} \phi(\vec{r}_j) + \frac{\partial}{\partial t} A_x(\vec{r}_j) \right) + q_j \dot{\vec{r}}_j \cdot \left( \frac{\partial \vec{A}(\vec{r}_j)}{\partial x_j} - \nabla A_x(\vec{r}_j) \right). \quad (\text{B.24})$$

It can be verified directly that for any vector  $\vec{v}$  and vector field  $\vec{A}$ ,

$$\left( \vec{v} \times (\nabla \times \vec{A}) \right)_1 = \vec{v} \cdot \left( \frac{\partial \vec{A}}{\partial x} - \nabla A_x \right),$$

where  $(\cdot)_1$  denotes the first component. Using this and (2.12) and (2.15), we see that (B.24) is the same as the first component of Lorentz equation (2.10).

Using the Euler-Lagrange equations (B.16) one can obtain analogously (2.16), (2.17) (cf. [50] II.B.2).

## Conjugate variables and Hamiltonian

Consider the variables  $\vec{r}_j$ . The conjugate momentum  $\vec{p}_j$  associated with  $\vec{r}_j$  can be calculated from the definition (B.10) using (B.22), (B.23). We obtain

$$\vec{p}_j = m_j \dot{\vec{r}}_j + q_j \vec{A}(\vec{r}_j). \quad (\text{B.25})$$

The Lagrangian (B.22) does not depend on the derivative of  $\phi$  with respect to time. Therefore the conjugate momentum associated with  $\phi$  is zero. This shows that the variable  $\phi$  is not an independent variable but it can be expressed in terms of the other variables and their time derivatives. This dependence is given by equation (2.16) which is obtained from the Euler Lagrange equations (B.16) with the Lagrangian (B.22), (B.23). In principle therefore, using equation (2.16) we can eliminate the variable  $\phi$  from the Lagrangian and reduce ourselves to only the variables  $A_{x,y,z}$ . This is not straightforward however as (2.16) is a differential constraint. One way to deal with this problem is to go to *reciprocal space* where all the variables are replaced by their spatial Fourier transforms and differential operators are replaced by algebraic operations (see [50]).

The variables  $A_{x,y,z}$  still form however a redundant set of variables as the choice of the gauge is not specified. This further redundancy can be solved by imposing the Coulomb gauge constraint

$$\nabla \cdot \vec{A} = 0. \quad (\text{B.26})$$

$A_{x,y,z}$  with the constraint  $\nabla \cdot \vec{A} = 0$  form a nonredundant number of independent variables describing the field.

**Remark B.3.1** Recall that  $\vec{r}$  in  $\vec{A}(\vec{r})$  has to be considered as an *index* for the dynamical variables. A constraint such as (B.26) relates the value of  $\vec{A}(\vec{r}_0)$  to the value of  $\vec{A}(\vec{r})$  for values of  $\vec{r}$  (infinitesimally) near  $\vec{r}_0$ . Lagrangian and Hamiltonian mechanics have to be modified in order to deal with problems with constraints. There are several approaches to do that. One of them is to use Lagrange multipliers (see, e.g., [130]). The Euler-Lagrange equations are not valid in the presence of constraints because in deriving them one assumes that the variation to the trajectory is unconstrained. Therefore a different approach followed in [95], Chapter 3, consists of incorporating the constraints in the derivation of the Euler-Lagrange equations by imposing that the variation from the actual path also satisfies the constraint. This gives modified equations. A third approach is to reduce the number of variables to the truly independent ones by expressing the redundant variables in terms of the independent ones or their time derivatives. The Lagrangian is then rewritten in terms of the minimum number of independent variables. If the constraints are of differential type this can be done only in an implicit form. However it can be done explicitly transforming the variables into their spatial Fourier transforms. The spatial Fourier transform  $\vec{\mathcal{A}}(\vec{k})$  of a vector field  $\vec{A}(\vec{r})$  defined

in Exercise 2.4 can be seen as a *re-parametrization* of the field variables, where the parameter  $\vec{r}$  is replaced by the parameter  $\vec{k}$ . For Fourier transforms *differential* constraints transform into *algebraic* constraints. For instance, the constraint  $\nabla \cdot \vec{A} = 0$  transforms into the algebraic constraints  $\vec{k} \cdot \vec{A}(\vec{k}) = 0$ .<sup>4</sup> Therefore adopting the description in terms of Fourier transform the only two degrees of freedom for every  $\vec{k}$  are the two components of the vector  $\vec{A}(\vec{k})$  perpendicular to  $\vec{k}$  while the component parallel to  $\vec{k}$  is zero (cf. Figures 2.1 and 2.2).

**Remark B.3.2** We have given the form of the Lagrangian (B.22), without considering the dependence among the dynamical variables, only based on the fact that the Euler-Lagrange equations give the correct equations of motion. Then we have found that  $\phi$  is not an independent variable and that we could make a change in the gauge without changing the equations of motion. A more rigorous approach would have been to first identify the truly independent variables. By adopting the Coulomb gauge, these can be taken in reciprocal space as the transversal components of the spatial Fourier transform of  $\vec{A}$ , while the longitudinal component is set equal to zero. The Fourier transform of  $\phi$  is obtained from the Fourier transform of equation (2.16). The Lagrangian in (B.22) can be written in reciprocal space in terms of the Fourier transform of the various variables and then the number of variables reduced by eliminating the redundant variables. The resulting Lagrangian, called the *reduced Lagrangian*, is written in terms of only the truly independent variables. It gives the correct equation of motion, which correspond to (2.17). The conjugate momentum of the Fourier transform  $\vec{A}(\vec{k})$ ,  $\vec{P}(\vec{k})$  can then be calculated using the definition (B.17) (opportunely adapted to accommodate complex variables). It is given by

$$\vec{P}(\vec{k}) = \epsilon_0 \dot{\vec{A}}(\vec{k}).$$

If  $\vec{P}(\vec{r})$  is the inverse transform of  $\vec{P}$ , we have

$$\vec{P}(\vec{r}) = \epsilon_0 \dot{\vec{A}}(\vec{r}). \quad (\text{B.27})$$

From Parseval-Plancherel theorem<sup>5</sup> it follows that

$$\int_{\mathbf{R}^3} \vec{P}^*(\vec{k}) \cdot \dot{\vec{A}}(\vec{k}) d\vec{k} = \int_{\mathbf{R}^3} \vec{P}^*(\vec{r}) \cdot \dot{\vec{A}}(\vec{r}) d\vec{r}. \quad (\text{B.28})$$

---

<sup>4</sup>In general,  $\nabla$  is replaced by  $i\vec{k}$ . The Fourier transform  $\vec{A}(\vec{k})$  is a complex vector. However this does not lead to an increase in the number of variables since, from the fact that  $\vec{A}(\vec{r})$  is real, it follows that

$$\vec{A}(\vec{k}) = \vec{A}^*(-\vec{k}).$$

<sup>5</sup>

$$\int_{\mathbf{R}^3} \vec{F}^*(\vec{r}) \vec{G}(\vec{r}) d\vec{r} = \int_{\mathbf{R}^3} \vec{F}^*(\vec{k}) \cdot \vec{G}(\vec{k}) d\vec{k},$$

where  $\vec{F}$  and  $\vec{G}$  are the spatial Fourier transforms of the fields  $\vec{F}$  and  $\vec{G}$ , respectively.

In the Coulomb gauge, the definition of the momentum conjugate to  $\vec{A}$  can be given rigorously in reciprocal space according to the discussion in the previous remark. The Hamiltonian is then written in reciprocal space according to the definition (B.19) and then can be written in real space ( $\vec{r}$ ) using (B.28), where  $\vec{P}(\vec{r})$  is given in (B.27). We have in real space,

$$H = \sum_{j=1}^N \vec{p}_j \cdot \dot{\vec{r}}_j + \int_{\mathbf{R}^3} P(\vec{r}) \cdot \dot{\vec{A}}(\vec{r}) d\vec{r} - L, \quad (\text{B.29})$$

where  $L$  is given in (B.22), (B.23). By replacing the expressions for the conjugate momenta (B.25) and (B.27) we obtain<sup>6</sup>

$$H = \frac{1}{2} \sum_{j=1}^N m_j \dot{\vec{r}}_j^2 + \frac{\epsilon_0}{2} \int_{\mathbf{R}^3} \dot{\vec{A}}^2 + c^2 (\nabla \times \vec{A})^2 d\vec{r} + \int_{\mathbf{R}^3} \rho \phi - \frac{\epsilon_0}{2} (\nabla \phi)^2 d\vec{r}. \quad (\text{B.30})$$

The last term has been calculated in Remark 2.1.1 in [Chapter 2](#). It is equal to  $V_{coul}$  in (2.26). Using (2.12) and (2.15) and the fact that we have adopted the Coulomb gauge, we obtain

$$H = \frac{1}{2} \sum_{j=1}^N m_j \dot{\vec{r}}_j^2 + \frac{\epsilon_0}{2} \int_{\mathbf{R}^3} \vec{E}_T^2 + c^2 \vec{B}^2 d\vec{r} + V_{coul}. \quad (\text{B.31})$$

This is the energy of the system of particles and electromagnetic field. The first term represents the kinetic energy of the particles while the remaining two terms represent the energy needed to establish the field.

## Lagrangian and Hamiltonian formulation in the presence of an external field

In the presence of an external field the Lagrangian  $L$  in (B.22), (B.23) has to be modified as follows

$$L_e := L + \int_{\mathbf{R}^3} \mathcal{L}_e d\vec{r},$$

with  $\mathcal{L}_e$  defined by

$$\mathcal{L}_e = J(\vec{r}) \cdot \vec{A}_e(\vec{r}, t) - \rho(\vec{r}) \phi_e(\vec{r}, t), \quad (\text{B.32})$$

---

<sup>6</sup>In applying the expression for the Lagrangian (B.23), we have used the fact that

$$\int_{\mathbf{R}^3} \dot{\vec{A}} \cdot \nabla \phi d\vec{r} = 0,$$

since  $\nabla \phi$  is purely longitudinal and  $\dot{\vec{A}}$  is purely transversal because of the Coulomb gauge assumption (cf. (2.25)).

where  $\vec{A}_e$  and  $\phi_e$  are the vector and scalar external potential. This Lagrangian is justified by the fact that it gives the correct equations of motion. In particular (cf. the diagram in 2.1.1.5) these are 1) the Lorentz equation (2.10) with  $\vec{E}$  replaced by  $\vec{E} + \vec{E}_e$  and  $\vec{B}$  replaced by  $\vec{B} + \vec{B}_e$ , with  $\vec{E}_e$  and  $\vec{B}_e$  the external electric and magnetic field and 2) equations (2.16) and (2.17) (where the external field does not appear). Starting with this Lagrangian, the discussion in the previous subsection can be repeated with minor modifications. In particular the moment conjugate to  $\vec{r}_j$ ,  $j = 1, \dots, N$ , is given by (2.36) and the Hamiltonian is given by (2.33).

# Appendix C

---

## Cartan Semisimplicity Criterion and Calculation of the Levi Decomposition

We present here some more notions of Lie algebra theory, including Cartan semisimplicity criterion for a Lie algebra. These notions should be sufficient for the calculation of the Levi decomposition in the simplest case, when the solvable radical is Abelian. These also give an idea of the algorithm in the general case, since this algorithm is an extension of the one for the Abelian case.

In all we shall say, we assume that we are dealing with subalgebras of  $su(n)$ , since these are the Lie algebras we are interested in. However, the results are valid much more in general, e.g. for Lie algebras over fields different from  $\mathbb{R}$ . We refer to [213] for a complete treatment.

---

### C.1 The Adjoint Representation

Consider a Lie algebra  $\mathcal{L}$  and an element  $X$  of  $\mathcal{L}$ .  $X$  acts on  $\mathcal{L}$  according to the linear map  $ad_X$ , defined by

$$ad_X Y := [X, Y].$$

Given a basis of  $\mathcal{L}$ ,  $ad_X$  can be represented by a matrix. The map  $X \rightarrow ad_X$  is called the **adjoint representation** of  $\mathcal{L}$ . It is a *representation* as it is a homomorphism from  $\mathcal{L}$  to the Lie algebra of real matrices of dimension equal to  $\dim(\mathcal{L})$ . In fact we have

$$ad_{[X,Y]} = ad_X ad_Y - ad_Y ad_X.$$

The *Killing form* associated with a Lie algebra  $\mathcal{L}$  is the bilinear form

$$\langle X, Y \rangle_K = \text{Tr}(ad_X ad_Y).$$

A bilinear form  $\langle X, Y \rangle$  over a vector space  $\mathcal{L}$  is said to be *non-degenerate* if  $\langle X, Y \rangle = 0$  for every  $Y \in \mathcal{L}$  implies  $X = 0$ .

## C.2 Cartan Semisimplicity Criterion

The following fundamental theorem due to Cartan allows us to check whether a Lie algebra is semisimple or not.

**Theorem C.2.1** *A Lie algebra  $\mathcal{L}$  is semisimple if and only if its Killing form  $\langle \cdot, \cdot \rangle_K$  is non-degenerate.*

An alternative test for semisimplicity says that  $\mathcal{L}$  is semisimple if and only if

$$[\mathcal{L}, \mathcal{L}] = \mathcal{L}. \quad (\text{C.1})$$

The following result allows us to calculate the solvable radical  $\mathcal{R}$  in equation (5.48) (cf. [213]).

**Theorem C.2.2** *The solvable radical  $\mathcal{R}$  of a Lie algebra  $\mathcal{L}$  is given by the set of elements  $X \in \mathcal{L}$  such that*

$$\langle X, Y \rangle_K = 0, \quad \forall Y \in [\mathcal{L}, \mathcal{L}].$$

In the special case when the Lie algebra is semisimple, (C.1) and Theorem C.2.1 give  $\mathcal{R} = \{0\}$ .

---

## C.3 Quotient Lie Algebras

Consider a Lie algebra  $\mathcal{L}$  and an ideal  $\mathcal{R}$  of  $\mathcal{L}$  and write  $\mathcal{L}$  as

$$\mathcal{L} = \mathcal{S} \oplus \mathcal{R}.$$

Consider the quotient vector space  $\mathcal{L}/\mathcal{R}$  defined as the set of equivalence classes  $\{(V)\}$  of elements  $V$  in  $\mathcal{L}$  which differ by an element in  $\mathcal{R}$ .<sup>1</sup>  $\mathcal{L}/\mathcal{R}$  can be given the structure of a Lie algebra by defining

$$[(V_1), (V_2)] := ([V_1, V_2]).$$

It is easily seen that, since  $\mathcal{R}$  is an ideal, this definition is well posed as  $([V_1, V_2])$  does not depend on the choice of the representative  $V_1$  and  $V_2$  in the equivalence classes  $(V_1)$  and  $(V_2)$ .

The following property of Levi decomposition is useful in the calculation of the Levi subalgebra.

---

<sup>1</sup>The basic vector space operation is defined by

$$\alpha(V_1) + \beta(V_2) = (\alpha V_1 + \beta V_2).$$

**Proposition C.3.1** The Levi subalgebra of  $\mathcal{L}$ ,  $\mathcal{S}$  in (5.48), is isomorphic to  $\mathcal{L}/\mathcal{R}$ .

---

## C.4 Calculation of the Levi Subalgebra in the Levi Decomposition

Assume now that a basis of  $\mathcal{R}$  is known,  $\{R_1, R_2, \dots, R_r\}$ . We make the additional assumption that  $\mathcal{R}$  is Abelian. We choose a complementary set of linearly independent elements  $\{X_1, \dots, X_s\}$  so that  $\{X_1, X_2, \dots, X_s, R_1, \dots, R_r\}$  is a basis of  $\mathcal{L}$ . Then, a basis of  $\mathcal{S}$  can be written as

$$S_l = X_l + \sum_{j=1}^r \alpha_{lj} R_j, \quad l = 1, \dots, s. \quad (\text{C.2})$$

As there exists an isomorphism from  $\mathcal{S}$  to  $\mathcal{L}/\mathcal{R}$  by Proposition C.3.1, we, in particular, consider the basis corresponding to the equivalence classes  $(X_l)$ ,  $l = 1, \dots, s$ , i.e., the isomorphism is such that

$$(X_l) \leftrightarrow S_l, \quad l = 1, \dots, s. \quad (\text{C.3})$$

Since the  $X_l$ 's are known, the coefficients  $c_{jk}^l$ , in the commutation relations

$$[(X_j), (X_k)] = \sum_{l=1}^s c_{jk}^l (X_l), \quad j < k, \quad j, k = 1, \dots, s, \quad (\text{C.4})$$

are known. Moreover, by isomorphism, the elements of the basis  $\{S_l\}$  in (C.2) must satisfy the same commutation relations as in (C.4), i.e., we have,

$$[S_j, S_k] = \sum_{l=1}^s c_{jk}^l S_l, \quad j < k, \quad j, k = 1, \dots, s. \quad (\text{C.5})$$

By placing (C.2) into (C.5) we obtain a system of  $\frac{s(s-1)}{2}$  equations in the  $rs$  unknowns  $\alpha_{lj}$ ,  $l = 1, \dots, s$ ,  $j = 1, \dots, r$ . The assumption that  $\mathcal{R}$  is Abelian is used in making this system linear, as the commutator of the two sums of (C.2) is zero. The solution of such a system exists, because of Levi's Theorem 5.4.4, and gives the desired basis of  $\mathcal{S}$ .

---

## C.5 Algorithm for the Levi Decomposition

Summarizing the above discussion, the algorithm for the computation of Levi decomposition is as follows.

1. Calculate a basis of  $[\mathcal{L}, \mathcal{L}]$ .
2. Use this basis and Theorem C.2.2 to calculate a basis,  $\{R_1, R_2, \dots, R_r\}$ , of the solvable radical  $\mathcal{R}$ .
3. Choose a complementary linearly independent set in  $\mathcal{L}$ ,  $\{X_1, \dots, X_s\}$  so that  $\{X_1, X_2, \dots, X_s, R_1, \dots, R_r\}$  is a basis of  $\mathcal{L}$ .
4. Calculate the coefficients  $c_{jk}^l$  in (C.5).
5. Solve the linear system of equations obtained by plugging (C.2) into (C.5). Using the solution  $\alpha_{lj}$ ,  $l = 1, \dots, s$ ,  $j = 1, \dots, r$ , in (C.2) we obtain a basis of  $\mathcal{S}$ .

# Appendix D

---

## ***Proof of the Controllability Test of Theorem 3.2.1***

We need to prove the formula

$$\mathcal{R} = e^{\mathcal{L}},$$

where  $\mathcal{R}$  is the reachable set defined in (3.2). We shall consider the class of piecewise constant controls.

The inclusion  $\mathcal{R} \subseteq e^{\mathcal{L}}$  is obvious since for a piecewise constant control the solution of (3.1) is the product of exponentials  $e^{-iH(u_j)t_j}$  and since  $-iH(u_j) \in \mathcal{L}$  and the product belongs to  $e^{\mathcal{L}}$ , because of the way  $e^{\mathcal{L}}$  was defined. The proof of  $e^{\mathcal{L}} \subseteq \mathcal{R}$  is longer. This inclusion means that, given an element  $X_f$  of  $e^{\mathcal{L}}$ , there exist piecewise constant controls such that the solution of (3.1) attains the value  $X_f$ . The topology of  $e^{\mathcal{L}}$  is the one induced by the one of  $U(n)$ . The proof presented here uses arguments borrowed from [56], [113], [218]. Related papers are [105], [168], [198]. We first state two facts and then conclude.

### **$\mathcal{R}$ is dense in $e^{\mathcal{L}}$**

Let  $\mathcal{S} := \{A_1, \dots, A_s\}$  be a set of linearly independent generators of  $\mathcal{L}$  obtained by setting the controls equal to appropriate values in  $-iH(u)$ . Clearly, all the elements in the one dimensional semigroups

$$\{X \in e^{\mathcal{L}} | X = e^{A_j t}, t \geq 0\}, \quad j = 1, \dots, s,$$

as well as their products, are in  $\mathcal{R}$ . It follows from Lemma 1, proved below, that every matrix in  $e^{\mathcal{L}}$  can be expressed as a finite product of matrices of the form  $e^{At}$ , with  $A \in \{A_1, \dots, A_s\}$  and  $t \in \mathbf{R}$ . Moreover, given a matrix of the form  $e^{At}$ , with  $t \in \mathbf{R}$ , we can always choose a *positive*  $\alpha$  such that  $e^{A\alpha}$  is arbitrarily close to  $e^{At}$ . In fact, for each element  $e^{-A|t|}$ , there exists a sequence of positive values  $t_k \geq 0$  such that

$$\lim_{k \rightarrow \infty} e^{At_k} = e^{-A|t|}. \quad (\text{D.1})$$

To construct such a sequence, pick the sequence  $e^{nA|t|}$ , which by compactness of  $U(n)$  has a converging subsequence  $e^{n(k)A|t|}$ . By setting  $e^{At_k} := e^{(n(k+1)-n(k)-1)A|t|}$  we obtain (D.1). In conclusion, we can always construct a sequence of elements in  $\mathcal{R}$  which converges to an element of  $e^{\mathcal{L}}$ .

### Finite generation of $e^{\mathcal{L}}$

**Lemma 1** Every matrix  $X_f$  in  $e^{\mathcal{L}}$  can be expressed as a finite product of matrices of the form  $e^{At}$ , with  $A \in \{A_1, \dots, A_s\}$  and  $t \in \mathbb{R}$ . The number of factors depends on  $X_f$ .

The previous Lemma is often expressed by saying that  $e^{\mathcal{L}}$  is *generated* by  $\{A_1, \dots, A_s\}$ . In other words, if  $A_1, \dots, A_s$  generate the Lie algebra  $\mathcal{L}$  we say that they (or the one dimensional subgroups associated with them) also generate its Lie group  $e^{\mathcal{L}}$ .

*Proof.* First notice that if  $\{A_1, A_2, \dots, A_m\}$  is a basis for  $\mathcal{L}$ , a neighborhood of the identity in  $e^{\mathcal{L}}$  can be obtained by varying  $t_1, \dots, t_m$  in a neighborhood of the origin in  $\mathbb{R}^m$  in the expression

$$K := e^{A_1 t_1} e^{A_2 t_2} \cdots e^{A_m t_m}. \quad (\text{D.2})$$

This follows from the Inverse Function Theorem (see e.g. [196], Lemma 3.1. Chp. 5, for a statement of this result).

We now show how to obtain a basis of  $\mathcal{L}$  by using similarity transformations involving only elements in the set  $\mathcal{S}$ .

There exist two elements  $A_k, A_j$  and a time (arbitrarily small)  $\tilde{t}$  such that  $e^{A_k \tilde{t}} A_j e^{-A_k \tilde{t}}$  is linearly independent from  $A_1, \dots, A_s$ . If this were not the case, we would have

$$e^{A_k t} A_j e^{-A_k t} = \sum_{l=1}^s a_l(t) A_l, \quad (\text{D.3})$$

for some functions  $a_l(t)$  and for every  $t$ . Differentiating (D.3) at  $t = 0$ , we obtain

$$[A_k, A_j] = \sum_{l=1}^s \dot{a}_l(0) A_l,$$

for every  $k, j = 1, 2, \dots, s$ , which (if  $\dim(\mathcal{L}) > m$ ) contradicts the fact that  $A_1, \dots, A_s$  are generators of  $\mathcal{L}$ . Set

$$A_{s+1} := e^{A_k \tilde{t}} A_j e^{-A_k \tilde{t}}. \quad (\text{D.4})$$

Of course,  $A_1, \dots, A_{s+1}$  still form a set of linearly independent generators and therefore, as above, there exist two elements  $A_k, A_j$  in the set  $\{A_1, \dots, A_s, A_{s+1}\}$  and a time  $\tilde{t}$  (arbitrarily small) such that  $e^{A_k \tilde{t}} A_j e^{-A_k \tilde{t}}$  is linearly independent from  $A_1, \dots, A_{s+1}$ . As in (D.4), we obtain a new element of  $\mathcal{L}$  that we denote by  $A_{s+2}$ , such that  $\{A_1, A_2, \dots, A_{s+2}\}$  is an  $s + 2$ -dimensional subspace of  $\mathcal{L}$ . Proceeding this way we obtain a basis  $A_1, \dots, A_s, A_{s+1}, \dots, A_m$  of the Lie algebra  $\mathcal{L}$  where the first  $s$  elements are the generators we started with and the elements  $A_{s+1}, \dots, A_m$  are obtained via similarity transformations with the

iterative procedure we have described. Every element  $A_j$ ,  $j = 1, \dots, m$ , can be written in the form

$$A_j = e^{\tilde{A}_r t_r} e^{\tilde{A}_{r-1} t_{r-1}} \cdots e^{\tilde{A}_1 t_1} \tilde{A}_k e^{-\tilde{A}_1 t_1} \cdots e^{-\tilde{A}_{r-1} t_{r-1}} e^{-\tilde{A}_r t_r},$$

where the indeterminates  $\tilde{A}_1, \dots, \tilde{A}_r, \tilde{A}_k$  belong to  $\mathcal{S}$ , and  $r$  is finite.

Replacing now in (D.2) the expressions of  $e^{A_j t}$ ,  $j = 1, \dots, m$ , in terms of the matrices in  $\mathcal{S}$  we obtain that all the elements in a neighborhood  $N$  of the identity in  $e^{\mathcal{L}}$  can be written as

$$K := e^{\tilde{A}_1 t_1} e^{\tilde{A}_2 t_2} \cdots e^{\tilde{A}_p t_p}. \quad (\text{D.5})$$

The indeterminates  $\tilde{A}_j$  are in the set  $\mathcal{S}$ ,  $t_j \in \mathbf{R}$ ,  $j = 1, \dots, p$ . Now, every element  $X_f$  in  $e^{\mathcal{L}}$  can be written as a product of elements of the type  $e^B$ , with  $B \in \mathcal{L}$ , but each one of these can be written as  $K$  in (D.5) because so can  $e^{\frac{B}{n}}$  for a sufficiently large positive integer  $n$ . This concludes the proof of the lemma.  $\square$

### $\mathcal{R}$ contains a neighborhood of the Identity in $e^{\mathcal{L}}$

Recall that  $m$  denotes the dimension of  $\mathcal{L}$ . From the proof of Lemma 1, we can choose  $m - s$  elements in  $e^{\mathcal{L}}$ ,  $U_1, \dots, U_{m-s}$ , and  $m - s$  elements in the set  $\{A_1, \dots, A_s\}$ , say  $\bar{A}_1, \dots, \bar{A}_{m-s}$ , such that,

$$\{A_1, A_2, \dots, A_s, U_1 \bar{A}_1 U_1^{-1}, \dots, U_{m-s} \bar{A}_{m-s} U_{m-s}^{-1}\},$$

form a basis of  $\mathcal{L}$ . Given  $m$  elements  $V_1, \dots, V_m$  in  $\mathcal{R}$ , consider the function  $\Phi(j_1, \dots, j_m)$  from a neighborhood of the point  $(1, \dots, 1)$  in  $\mathbf{R}^m$  to  $e^{\mathcal{L}}$ ,

$$\Phi(j_1, \dots, j_m) := e^{A_1 j_1} V_1 e^{A_2 j_2} V_2 \cdots e^{A_s j_s} V_s e^{\bar{A}_1 j_{s+1}} V_{s+1} \cdots e^{\bar{A}_{m-s} j_m} V_m.$$

All the elements obtained by  $\Phi(j_1, \dots, j_n)$  with  $j_1, \dots, j_n \geq 0$  are in  $\mathcal{R}$ . We can show that, if we choose appropriately  $V_1, \dots, V_m$ , the set of these elements contains a neighborhood of the point  $U_0 := e^{A_1} V_1 e^{A_2} V_2 \cdots e^{A_s} V_s e^{\bar{A}_1} V_{s+1} \cdots e^{\bar{A}_{m-s}} V_m$  by showing that the Jacobian at  $(1, \dots, 1)$  is not zero and applying the Implicit Function Theorem (cf. [218]).

Differentiating at  $(1, \dots, 1)$  the function  $\Phi(j_1, \dots, j_m)$ , we obtain

$$\begin{aligned} \frac{\partial \Phi(j_1, \dots, j_m)}{\partial j_1} \Big|_{j_1, \dots, j_m=1, \dots, 1} &= A_1 e^{A_1} V_1 e^{A_2} V_2 \cdots e^{A_s} V_s e^{\bar{A}_1} V_{s+1} \cdots e^{\bar{A}_{m-s}} V_m, \\ \frac{\partial \Phi(j_1, \dots, j_m)}{\partial j_2} \Big|_{j_1, \dots, j_m=1, \dots, 1} &= e^{A_1} V_1 A_2 e^{A_2} V_2 \cdots e^{A_s} V_s e^{\bar{A}_1} V_{s+1} \cdots e^{\bar{A}_{m-s}} V_m, \\ &\vdots \\ \frac{\partial \Phi(j_1, \dots, j_m)}{\partial j_{s+1}} \Big|_{j_1, \dots, j_m=1, \dots, 1} &= e^{A_1} V_1 e^{A_2} V_2 \cdots e^{A_s} V_s \bar{A}_1 e^{\bar{A}_1} V_{s+1} \cdots e^{\bar{A}_{m-s}} V_m, \end{aligned}$$

$$\frac{\partial \Phi(j_1, \dots, j_m)}{\partial j_m} \Big|_{j_1, \dots, j_m=1, \dots, 1} = e^{A_1} V_1 e^{A_2} V_2 \cdots e^{A_s} V_s e^{\bar{A}_1} V_{s+1} \cdots \bar{A}_{m-s} e^{\bar{A}_{m-s}} V_m.$$

Choosing

$$\begin{aligned} V_1 &\approx e^{-A_1}, \\ &\vdots \\ V_{s-1} &\approx e^{-A_{s-1}}, \\ V_s &\approx e^{-A_s} U_1, \\ V_{s+1} &\approx e^{-\bar{A}_1} U_1^{-1} U_2, \\ &\vdots \\ V_{m-1} &\approx e^{-\bar{A}_{m-s-1}} U_{m-s-1}^{-1} U_{m-s}, \\ V_m &\approx e^{-\bar{A}_{m-s}} U_{m-s}^{-1}, \end{aligned}$$

we obtain

$$\begin{aligned} \frac{\partial \Phi(j_1, \dots, j_m)}{\partial j_1} \Big|_{j_1, \dots, j_m=1, \dots, 1} &\approx A_1, \\ &\vdots \\ \frac{\partial \Phi(j_1, \dots, j_m)}{\partial j_s} \Big|_{j_1, \dots, j_m=1, \dots, 1} &\approx A_s, \\ \frac{\partial \Phi(j_1, \dots, j_m)}{\partial j_{s+1}} \Big|_{j_1, \dots, j_m=1, \dots, 1} &\approx U_1 \bar{A}_1 U_1^{-1}, \\ &\vdots \\ \frac{\partial \Phi(j_1, \dots, j_m)}{\partial j_m} \Big|_{j_1, \dots, j_m=1, \dots, 1} &\approx U_{m-s} \bar{A}_{m-s} U_{m-s}^{-1}, \end{aligned}$$

which are linearly independent and this proves our claim.

Now, from the fact that the open ball with radius  $\epsilon$  and center at  $U_0$ ,  $B_\epsilon(U_0)$  is a subset of  $\mathcal{R}$ , it follows that  $\mathcal{R}$  contains a neighborhood of the identity. In particular, using the fact that  $\mathcal{R}$  is dense in  $e^{\mathcal{L}}$ , choose  $\bar{U} \in \mathcal{R}$ , such that

$$\|\bar{U} - U_0^{-1}\| < \frac{\epsilon}{2}.$$

Then if  $F \in B_{\frac{\epsilon}{2}}(I)$ , writing  $F = \bar{U}X$  we have

$$\frac{\epsilon}{2} > \|F - I\| \geq \|X - U_0\| - \|\bar{U} - U_0^{-1}\|,$$

and therefore

$$\|X - U_0\| < \epsilon,$$

which implies  $X \in \mathcal{R}$  and therefore  $F = \bar{U}X \in \mathcal{R}$ , since  $\mathcal{R}$  is a semigroup.

### Conclusion of the argument

The semigroup  $\mathcal{R}$  contains a neighborhood of the identity in  $e^{\mathcal{L}}$ . Since  $e^{\mathcal{L}}$  is connected, it follows that  $\mathcal{R} = e^{\mathcal{L}}$ . Alternatively, for every  $A \in \mathcal{L}$ ,  $e^{\frac{A}{n}}$  belongs to an arbitrarily small neighborhood of the identity and therefore is in the reachable set  $\mathcal{R}$  and so is  $(e^{\frac{A}{n}})^n = e^A$ .

# Appendix E

---

## The Baker-Campbell-Hausdorff Formula and Some Exponential Formulas

There are several versions of the **Baker-Campbell-Hausdorff** (BCH) formula. One version, particularly useful, is as follows. Consider two matrices  $A$  and  $B$ , we have

$$e^A B e^{-A} = \sum_{k=0}^{\infty} \frac{1}{k!} ad_A^k B, \quad (\text{E.1})$$

where  $ad_A^k B$  is defined recursively by  $ad_A^0 B := B$ ,  $ad_A^k B := [A, ad_A^{k-1} B]$ , if  $k \geq 1$ . This expression can be obtained by calculating the McLaurin series of the matrix function  $F(t) := e^{At} B e^{-At}$  at  $t = 1$ . Exponentiating both sides of (E.1) we obtain the formula

$$e^A e^B e^{-A} = e^{\sum_{k=0}^{\infty} \frac{1}{k!} ad_A^k B}. \quad (\text{E.2})$$

If  $A$  and  $B$  commute, i.e.,  $[A, B] = 0$ , we have

$$e^A e^B = e^B e^A = e^{A+B}. \quad (\text{E.3})$$

This formula has a generalization to the case where both  $A$  and  $B$  commute with  $[A, B]$ , i.e.,

$$[B, [A, B]] = [A, [A, B]] = 0. \quad (\text{E.4})$$

In this case, (E.2) reduces to

$$e^A e^B e^{-A} = e^{B + \frac{1}{2}[A, B]},$$

which using the commutativity properties (E.4) becomes

$$e^A e^B = e^B e^A e^{\frac{1}{2}[A, B]},$$

that is property (E.3) is modified by adding the extra factor  $e^{\frac{1}{2}[A, B]}$ .

We now consider ways to estimate how  $e^{A+B}$  differs from  $e^A e^B$ , in particular when  $A$  and  $B$  are small. Let us introduce a parameter  $t \in \mathbb{R}$ . We have that there exists an  $\epsilon > 0$  such that for  $|t| < \epsilon$

$$e^{tA} e^{tB} = e^{tA+tB+\frac{t^2}{2}[A, B]+O(t^3)}. \quad (\text{E.5})$$

Moreover we can approximately move in the direction of  $[A, B]$  by moving back and forth along the directions of  $A$  and  $B$  since we have

$$e^{-tA} e^{-tB} e^{tA} e^{tB} = t^{t^2[A, B] + O(t^3)}. \quad (\text{E.6})$$

In (E.5) and (E.6)  $O(t^3)$  is a matrix function of  $t$  such that  $\frac{O(t^3)}{t^3}$  is bounded for  $|t| < \epsilon$ . Another useful formula is

$$e^{A+B} = \lim_{n \rightarrow \infty} (e^{\frac{A}{n}} e^{\frac{B}{n}})^n. \quad (\text{E.7})$$

The proofs of formulas (E.5), (E.6) and (E.7) can be found in [184] along with several other exponential formulas.

# Appendix F

---

## Proof of Theorem 6.2.1

We shall follow [81] (section II-10, II-11) to which we refer for details.

We rewrite the definition of strong variation  $u^\epsilon$ . Fix  $\tau \in (0, T]$ . A strong variation of  $u = u(t)$  is a function  $u_\epsilon = u_\epsilon(t)$ , such that  $u^\epsilon(t) = u(t)$ , for  $t \in [0, \tau - \epsilon]$  and for  $t \in (\tau, T]$ ,  $u^\epsilon(t) \equiv v$ , for  $t \in (\tau - \epsilon, \tau]$ , where  $v$  is a value in the set of admissible controls,  $\mathcal{U} \subset \mathbb{R}^m$ .

We denote by  $x^\epsilon$  the trajectory corresponding to  $u^\epsilon$ . Condition (6.14) gives

$$\frac{d}{d\epsilon} J(u^\epsilon) \Big|_{\epsilon=0} \geq 0.$$

We have

$$\frac{d}{d\epsilon} J(u^\epsilon) \Big|_{\epsilon=0} = \frac{d}{d\epsilon} \phi(x^\epsilon(T)) \Big|_{\epsilon=0} = \phi_x(x(T)) \frac{d}{d\epsilon} x^\epsilon(T) \Big|_{\epsilon=0}, \quad (\text{F.1})$$

where  $\phi_x$  is the Jacobian with respect to  $x$  of the function  $\phi$ , i.e.,

$$\phi_x := \left[ \frac{\partial \phi}{\partial x_1}, \dots, \frac{\partial \phi}{\partial x_n} \right].$$

In the next subsection, we calculate  $\frac{d}{d\epsilon} x^\epsilon(T) \Big|_{\epsilon=0}$ . Defining  $x_\epsilon := x^\epsilon(\tau)$ , we can write

$$\frac{d}{d\epsilon} x^\epsilon(T) \Big|_{\epsilon=0} = \frac{\partial x^\epsilon(T)}{\partial x_\epsilon} \frac{\partial x_\epsilon}{\partial \epsilon} \Big|_{\epsilon=0}. \quad (\text{F.2})$$

On the right hand side of the above expression, the term  $\frac{\partial x_\epsilon}{\partial \epsilon} \Big|_{\epsilon=0}$  describes how  $x(\tau)$  varies because of the variation of the control, while the term  $\frac{\partial x^\epsilon(T)}{\partial x_\epsilon}$  describes the propagation of this change of the state in the interval from time  $\tau$  to time  $T$ .

### Effect of a strong variation of the control

We write (cf. (6.1))

$$x^\epsilon(\tau) = x^\epsilon(\tau - \epsilon) + \int_{\tau-\epsilon}^{\tau} f(x^\epsilon(s), v) ds,$$

$$x(\tau) = x(\tau - \epsilon) + \int_{\tau-\epsilon}^{\tau} f(x(s), u(s)) ds.$$

Since  $x^\epsilon(\tau - \epsilon) = x(\tau - \epsilon)$ , we have

$$x^\epsilon(\tau) = x(\tau) + \int_{\tau-\epsilon}^{\tau} (f(x^\epsilon(s), v) - f(x(s), u(s))) ds.$$

Adding and subtracting  $f(x(s), v)$  inside the integral, we obtain

$$\begin{aligned} x^\epsilon(\tau) &= x(\tau) + \int_{\tau-\epsilon}^{\tau} (f(x(s), v) - f(x(s), u(s))) ds \\ &\quad + \int_{\tau-\epsilon}^{\tau} (f(x^\epsilon(s), v) - f(x(s), v)) ds. \end{aligned}$$

As  $\epsilon \rightarrow 0$ , the second integral goes to zero faster than  $\epsilon$  since both the integrand and the interval of integration go to zero. Therefore, using a Taylor series for the first integral, with initial point  $\epsilon = 0$ , we can write

$$x_\epsilon := x^\epsilon(\tau) = x(\tau) + (f(x(\tau), v) - f(x(\tau), u(\tau))) \epsilon + o(\epsilon).$$

Differentiating this with respect to  $\epsilon$ , at  $\epsilon = 0$ , we obtain

$$\frac{\partial x_\epsilon}{\partial \epsilon} \Big|_{\epsilon=0} = f(x(\tau), v) - f(x(\tau), u(\tau)). \quad (\text{F.3})$$

**Variational equation** On the interval  $(\tau, T]$ ,  $x^\epsilon$  and  $x$  satisfy the same differential equation (6.1) i.e.

$$\dot{x} = f(x(t), u(t)), \quad (\text{F.4})$$

however with different initial conditions. For a general differential equation of the form (F.4), write  $x$  as  $x := x(t, y)$  where the dependence on the initial condition  $y$  is made explicit. Differentiating both sides of (F.4) with respect to  $y$  at the actual value of the initial condition  $x_0$ , and changing the order of differentiation, we obtain

$$\frac{d}{dt} \frac{\partial}{\partial y} x(t, y) \Big|_{y=x_0} = f_x(x(t, x_0), u(t)) \frac{\partial}{\partial y} x(t, y) \Big|_{y=x_0}.$$

If  $V$  is defined as  $V(t) := \frac{\partial}{\partial y} x(t, y)|_{y=x_0}$ , it measures the *sensitivity* of the solution  $x(t, x_0)$  to variations of the initial condition  $x_0$ . It satisfies the **variational equation**

$$\dot{V} = f_x(x(t, x_0), u(t))V, \quad (\text{F.5})$$

with initial condition equal to the identity matrix.

### The Pontryagin maximum principle

In (F.2),  $\frac{\partial x^\epsilon(T)}{\partial x_\epsilon}|_{\epsilon=0}$  is equal to  $V(T, \tau)$ , where  $V$  is the solution of the variational equation (cf. F.5)

$$\frac{\partial V(t, \tau)}{\partial t} = f_x(x(t, x(\tau)), u(t))V(t, \tau), \quad (\text{F.6})$$

with initial condition equal to the identity matrix. The expression for  $\frac{\partial x_\epsilon}{\partial \epsilon}|_{\epsilon=0}$  is given in (F.3). Therefore we obtain

$$\left. \frac{d}{d\epsilon} x^\epsilon(T) \right|_{\epsilon=0} = V(T, \tau) (f(x(\tau), v) - f(x(\tau), u(\tau))).$$

Plugging this into (F.1), we obtain

$$\left. \frac{d}{d\epsilon} J(u^\epsilon) \right|_{\epsilon=0} = \left. \frac{d}{d\epsilon} \phi(x^\epsilon(T)) \right|_{\epsilon=0}$$

$$= \phi_x(x(T)) V(T, \tau) (f(x(\tau), v) - f(x(\tau), u(\tau))) \geq 0.$$

We summarize this in the following theorem.

**Theorem F.0.1** *Assume  $u$  is the optimal control and  $x$  is the corresponding trajectory, both of them defined in  $[0, T]$ . Then, for every  $\tau \in (0, T]$  and  $v$  in the set of admissible values of the control, if  $V(\cdot, \tau)$  is the solution of the variational equation (F.6), starting at  $\tau$ , with  $V(\tau, \tau)$  equal to the identity, we have*

$$\phi_x(x(T)) V(T, \tau) (f(x(\tau), v) - f(x(\tau), u(\tau))) \geq 0. \quad (\text{F.7})$$

The necessary conditions for optimality are usually not expressed in the form stated in Theorem F.0.1 but they are reformulated as follows. Given an optimal control  $u$  and corresponding trajectory  $x$ , define an  $n$ -dimensional time varying vector function  $\lambda$ , called the *costate*, satisfying the linear time varying differential equation

$$\dot{\lambda}^T = -\lambda^T f_x(x(t), u(t)). \quad (\text{F.8})$$

These equations are called *adjoint equations*. If  $z$  is a solution of the linear time varying equation

$$\dot{z} = f_x(x(t), u(t))z, \quad (\text{F.9})$$

a straightforward calculation shows that  $\frac{d}{dt} \lambda^T z = 0$  so that

$$\lambda^T(t)z(t) = \text{constant}.$$

Now, fix  $\tau$  and consider the solution of (F.9) with initial condition

$$z(\tau) = f(x(\tau), v) - f(x(\tau), u(\tau)).$$

We have

$$z(t) = V(t, \tau) (f(x(\tau), v) - f(x(\tau), u(\tau)))$$

and

$$\lambda^T(t)V(t, \tau) (f(x(\tau), v) - f(x(\tau), u(\tau))) = \text{constant}.$$

In particular, if we choose

$$\lambda^T(T) = -\phi_x(x(T)), \quad (\text{F.10})$$

we have

$$\begin{aligned} & \lambda^T(t)V(t, \tau) (f(x(\tau), v) - f(x(\tau), u(\tau))) \\ &= -\phi_x(x(T)) V(T, \tau) (f(x(\tau), v) - f(x(\tau), u(\tau))) \leq 0, \end{aligned} \quad (\text{F.11})$$

for every  $t \in [\tau, T]$ , where we have used (F.7). Writing (F.11) for  $t = \tau$ , using the fact that  $V(\tau, \tau)$  is equal to the identity, we obtain

$$\lambda^T(\tau)(f(x(\tau), v) - f(x(\tau), u(\tau))) \leq 0. \quad (\text{F.12})$$

Equation (F.12), along with (F.8) and (F.10) give equations (6.17), (6.15) and (6.16) of Theorem 6.2.1, respectively.

---

## References

- [1] A. F. Abouraddy, B. E. A. Saleh, A. V. Sergienko and M. C. Teich, Degree of entanglement for two qubits, *Phys. Rev. A* 64, 050101 (2001).
- [2] A. Abragam, *Principles of Nuclear Magnetism*, International Series of Monographs on Physics, Vol. 32, Oxford Science Publications, New York, Reprint 2002.
- [3] N. I. Akhiezer and I. M. Glazman, *Theory of Linear Operators in Hilbert Spaces*, Pitman, Boston, 1981.
- [4] F. Albertini and D. D'Alessandro, The Lie algebra structure and controllability of spin systems, *Linear Algebra and its Applications*, 350, 213-235 (2002).
- [5] F. Albertini and D. D'Alessandro, Notions of controllability for bilinear multilevel quantum systems, *IEEE Transactions on Automatic Control*, 48, No. 8, 1399-1403 (2003).
- [6] F. Albertini and D. D'Alessandro, Control of the evolution of Heisenberg spin systems, *European Journal of Control*, Special issue on Lagrangian and Hamiltonian Methods for Nonlinear Control, 10/5, 497-504, (2004).
- [7] F. Albertini and D. D'Alessandro, Model identification for spin networks, *Linear Algebra and its Applications*, 394, 237-256, (2005).
- [8] F. Albertini and D. D'Alessandro, Analysis and identification of quantum dynamics using Lie algebra homomorphisms and Cartan decompositions, arXiv:quant-ph/0606057.
- [9] S. Albeverio, L. Cattaneo, S-M Fei and X-H Wang, Equivalence of tripartite quantum states under local unitary transformations, arXiv:quant-ph/0512246
- [10] S. Albeverio, S-M Fei, D. Goswami, Local invariants for a class of mixed states, *Phys. Lett. A* 340, 37-42, (2005).
- [11] S. Albeverio, S-M Fei, P. Parashar and W-L Yang, Nonlocal properties and local invariants for bipartite systems, *Phys. Rev. A* 68, 010313 (R), (2003).
- [12] R. Alicki and K. Lendi, *Quantum Dynamical Semigroups and Applications*, Lecture Notes in Physics Vol. 286 (Springer-Verlag, 1987).

- [13] C. Altafini, Controllability of quantum mechanical systems by root space decompositions of  $su(n)$ , *J. Math. Phys.*, 43, 2051-2062, (2002).
- [14] G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists*, 4-th edition, Academic Press, San Diego, 1995.
- [15] V. I. Arnold, *Mathematical Methods of Classical Mechanics*, Springer-Verlag, New York, 1978.
- [16] A. Assion, T. Baumert, M. Bergt, T. Brixner, B. Kiefer, V. Seyfried, M. Strehle and G. Gerber, Control of chemical reactions by feedback-optimized phase-shaped femtosecond laser pulses, *Science*, 282 No. 30, 919-922, (1998).
- [17] K. Audenaert, F. Verstraete and B. De Moor, Variational characterisation of separability and entanglement of formation, *Phys. Rev. A* 64, 052304, (2001).
- [18] J.E. Avron, R. Seiler and L. G. Yaffe, Adiabatic theorems and applications to the quantum Hall effect, *Commun. Math. Phys.*, 110, 33-49, (1987)
- [19] G. Benenti, G. Casati and G. Strini, *Principles of Quantum Computation and Information*, Volume I: Basic Concepts, World Scientific Publishing Co., River Edge, NJ, 2004.
- [20] I. Bengtsson and K. Zyczkowski, *Geometry of Quantum States : an Introduction to Quantum Entanglement*, Cambridge University Press, Cambridge, New York, 2006.
- [21] C. H. Bennett, H. J. Bernstein, S. Popescu and B. Schumacher, Concentrating partial entanglement by local operations, *Phys. Rev. A*, 53, 2046, (1996).
- [22] C. H. Bennett, D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. Smolin and W. K. Wootters, *Physical Review A*, 59, 1070, (1999).
- [23] K. Bergmann, H. Theuer and B.W. Shore, Coherent population transfer among quantum states of atoms and molecules, *Rev. Mod. Phys.*, 70, 1003, (1998).
- [24] A. M. Bloch, R. W. Brockett and C. Rangan, The controllability of infinite quantum systems and closed subspace criteria, xxx.lanl.gov, arXiv:quant-ph/0608075.
- [25] F. Bloch, W. Hansen and M. Packard, Nuclear induction, *Phys. Rev.*, 69:127, (1946).
- [26] W. M. Boothby, *An Introduction to Differentiable Manifolds and Riemannian Geometry*, Orlando, FL, Academic Press, 1986.

- [27] U. Boscain, Resonance of minimizers for  $n$ -level quantum systems, Proceedings of the 42-nd Conference on Decision and Control, Maui, Hawaii, Dec 9-12, 2003, Vol. 4, pp. 416-421.
- [28] U. Boscain, T. Chambrion and G. Charlot, Nonisotropic 3-level quantum systems: Complete solutions for minimum time and minimum energy, *Discrete and Continuous Dynamical Systems-B*, 5, No. 4, 957-990, (2005).
- [29] U. Boscain and B. Piccoli, *Optimal Syntheses for 2-D Manifolds*, Springer SMAI, Vol 43, (2004).
- [30] D. W. Boukhvalov, M. Al-Seqer, E. Z. Kurmaev, A. Moewes, V. R. Galakhov, L. D. Finkelstein, S. Chiuzbaian, M. Neumann, V.V. Dobrovitski, M. I. Katsnelson, A. I. Lichtenstein, B. N. Harmon, K. Endo, J. M. North, and N. S. Dalal, Electronic structure of a  $Mn_{12}$  molecular magnet: Theory and experiments, *Physical Review B*, 75, 014419, (2007).
- [31] H-P. Breuer, Optimal entanglement criterion for mixed quantum states, *Phys. Rev. Lett.*, 97, 080501, (2006).
- [32] H-P. Breuer, Separability criteria and bounds for entanglement measures, *J. Phys. A, Math. Gen.*, 39, 11847, (2006).
- [33] H-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems*, Oxford; New York, Oxford University Press, 2002.
- [34] P. Brumer and M. Shapiro, Control of unimolecular reactions using coherent light, *Chemical Physics Letters*, 126, 541, (1986).
- [35] R. K. Brylinski and G. Chen eds., *Mathematics of Quantum Computation*, Chapman and Hall, CRC, Boca Raton FL, 2002.
- [36] S. S. Bullock and G. K. Brennen, Canonical decompositions of  $n$ -qubits quantum computations and concurrence, *Journal of Math. Phys.*, 45, No. 6, 2447, (2004).
- [37] S. S. Bullock, G. K. Brennen and D. P. O'Leary, Time reversal and  $n$ -qubit canonical decompositions, *Journal of Math. Phys.*, 46, 062104, (2005).
- [38] A. Bunse-Gerstner, R. Byers and V. Mehrmann, A chart of numerical methods for structured eigenvalue problems, *SIAM J. Matrix Anal. Appl.*, 13, 419, (1992).
- [39] A. G. Butkovskiy and Y. I. Samoilenco, *Control of Quantum Mechanical Processes and Systems*, Mathematics and its Applications, Kluwer Academic Publishers, Dordrecht, 1990.

- [40] R. N. Cahn, *Semi-simple Lie algebras and their representations*, Frontiers in Physics, Lecture notes series 59, The Benjamin/Cummings Publishing Company, Inc., Menlo Park, California, 1984.
- [41] H. Y. Carr and E. M. Purcell, Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Phys. Rev.*, 94, 630-638, (1954).
- [42] E. Cartan, Sur une class remarkable d'espaces de Riemann, *Bull. Soc. Math. France*, 54, 214-264, (1926).
- [43] E. Cartan, Sur une class remarkable d'espaces de Riemann, *Bull. Soc. Math. France*, 55, 114-134, (1927).
- [44] H.A. Carteret, A. Higuchi and A. Sudbery, Multipartite generalisation of the Schmidt decomposition, *Journal of Mathematical Physics*, 41, 7932, (2000).
- [45] J. Cavanagh, W. J. Fairbrother, A. G. Palmer and N. J. Skelton, *Protein NMR Spectroscopy, Principles and Practice*, Academic Press, San Diego, 1996.
- [46] K. Chen, S. Albeverio and S-M Fei, Concurrence of arbitrary dimensional bipartite quantum states, *Phys. Rev. Lett.* 95, 040504, (2005).
- [47] K. Chen and L. A. Wu, A matrix realignment method for recognizing entanglement, *Quantum Information and Computation*, 3, 193, (2003).
- [48] J. I. Cirac and P. Zoller, Quantum computation with cold trapped ions, *Physical Review Letters*, 74, No. 20, 4091-4094, (1995).
- [49] C. Cohen-Tannoudji, B. Diu and F. Laloe, *Quantum Mechanics*, Wiley, New York, 1977.
- [50] C. Cohen-Tannoudji, J. Dupont-Roc and G. Grynberg, *Photons and Atoms; Introduction to Quantum Electrodynamics*, Wiley Professional Paperback Series, New York, 1997.
- [51] J. B. Conway, *A Course in Functional Analysis*, Graduate Texts in Mathematics, 96, Springer-Verlag, New York, 1990.
- [52] G. W. Coulston and K. Bergmann, Population transfer by stimulated Raman scattering with delayed pulses: Analytical results for multilevel systems, *J. Chem. Phys.*, 96, 3467 (1992).
- [53] M. L. Curtis, *Matrix Lie Groups*, Springer, New York, 1979.
- [54] M. Dagli, D. D'Alessandro and J. D. H. Smith, A recursive decomposition of unitary operators on N qubits, arXiv:quantum-ph/0701193
- [55] D. D'Alessandro, Algorithms for quantum control based on decompositions of Lie groups, in *Proceedings of the 39-th Conference on Decision and Control*, Sidney, Dec. 2000, pp. 967-968

- [56] D. D'Alessandro, Uniform finite generation of compact Lie groups, *Systems and Control Letters*, 47, 87-90, (2002).
- [57] D. D'Alessandro, On quantum state observability and measurement, *Journal of Physics A: Math. Gen.* 36, 9721-9735, (2003).
- [58] D. D'Alessandro, Controllability of one spin and two interacting spins, *Math. Control Signals Systems*, 16, 1-25, (2003).
- [59] D. D'Alessandro, On the observability and state determination of quantum mechanical systems, *Proceedings of the 43-nd Conference on Decision and Control*, Paradise Island, Bahamas, Dec. 2004.
- [60] D. D'Alessandro, Optimal evaluation of generalized Euler angles with application to control, *Automatica*, 40, 1997-2002, (2004).
- [61] D. D'Alessandro and F. Albertini, Quantum symmetries and Cartan decompositions in arbitrary dimensions, *Journal of Physics A: Mathematical and Theoretical* 40, 2439-2453, (2007).
- [62] D. D'Alessandro and M. Dahleh, Optimal control of two-level quantum systems, *IEEE Transactions on Automatic Control*, 45, No.1, (2001).
- [63] D. D'Alessandro and R. Romano, Further results on the observability of quantum systems under general measurement, *Quantum Information Processing*, 5, No. 3, 139-160, (2006).
- [64] D. D'Alessandro and R. Romano, Decompositions of unitary evolutions and entanglement dynamics of bipartite quantum systems, *J. Math. Phys.* 47, 082109, (2006).
- [65] R. W. Daniels, *An introduction to numerical methods and optimization techniques*, North-Holland, New York, 1978.
- [66] G. M. D'Ariano, M. G. A. Paris, and M. F. Sacchi, Quantum tomographic methods, *Lect. Notes Phys.*, 649, 7-58, (2004).
- [67] G. M. D'Ariano, L. Maccone and M. Paini, Spin tomography, *J. Opt. B: Quantum Semiclass. Opt.*, 5, 77-84, (2003).
- [68] J. H. Davies, *The Physics of Low-Dimensional Semiconductors*, Cambridge University Press, Cambridge, U.K., 1998.
- [69] M. Demiralp and H. Rabitz, Optimally controlled quantum molecular dynamics: A perturbation formulation and the existence of multiple solutions, *Physical Review A*, 47, No. 2, 809, (1992).
- [70] A. C. Doherty, P. A. Parrilo and F. M. Spedalieri, Distinguishing separable and entangled states, *Phys. Rev. Lett.*, 88, 187904, (2002).
- [71] M. J. Donald, M. Horodecki and O. Rudolph, The uniqueness theorem for entanglement measures, *Journal of Mathematical Physics*, 43, 4252, (2002).

- [72] J. Dongarra, J. Gabriel, D. Koelling and J. Wilkinson, The eigenvalue problem for Hermitian matrices with time reversal symmetry, *Linear Algebra App.*, 60, 27, (1984).
- [73] E. B. Dynkin, Semi-simple subalgebras of semi-simple Lie algebras, *Am. Math. Soc. Trans. Ser. 2*, 6, 111-244, (1957).
- [74] A. Einstein, B. Podolsky and N. Rosen, Can quantum mechanical description of physical reality be considered complete?, *Physical Review A*, 47, 777, (1935).
- [75] A. Elgart and J. E. Avron, Adiabatic theorem without a gap condition, *Communications in Mathematical Physics*, 203, 445-463. (1999).
- [76] R. R. Ernst, G. Bodenhausen and A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Oxford University Press, 1987.
- [77] E. Farhi, J. Goldstone, S. Gutmann and M. Spiser, Quantum computation by adiabatic evolution, xxx.lanl.quant-ph/0001106.
- [78] E. Farhi, J. Goldstone, G. Sam, J. Lapan, A. Lundgren and D. Preda, A quantum adiabatic evolution algorithm applied to random instances of an *NP*-complete problem, *Science* 292, 472-474, (April 20-th 2001).
- [79] Shao-Ming Fei and Naihuan Jing, Equivalence of Quantum States under Local Unitary Transformations, *Phys. Lett. A*, 342, 77-81, (2005).
- [80] R. P. Feynman, R. B. Leighton and M. Sands, Volume III of *The Feynman Lectures on Physics*, Addison-Wesley, Reading, Mass, 1965.
- [81] W. Fleming and R. Rishel, *Deterministic and Stochastic Optimal Control*, Applications of Mathematics, Springer-Verlag, New York, 1975.
- [82] A. Galindo and P. Pascual, *Quantum Mechanics I, Texts and Monographs in Physics*, Springer-Verlag, Heidelberg, 1990.
- [83] X.H. Gao, S. Alberverio, S.M. Fei, Z.X. Wang, Matrix tensor product approach to the equivalence of multipartite states under local unitary transformations, *Commun. Theor. Phys.*, 45, 267-270, (2006).
- [84] S. Gasiorowicz, *Quantum Physics* 3-rd edition, Wiley, Hoboken, N.J., 2003.
- [85] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Reading, UK, 1993.
- [86] H. Goldstein, *Classical Mechanics*, 2nd Edition, Addison-Wesley, Reading, MA, 1980.
- [87] W. Gordy, *Theory and Applications of Electron Spin Resonance*, Wiley, New York, 1980.

- [88] K. Gottfried, *Quantum Mechanics:Fundamentals*, Graduate texts in contemporary physics, Springer, New York, 2003.
- [89] M. Grassl, M. Rötteler and T. Beth, Computing local invariants of qubit systems, *Phys. Rev. A*, 58, 1833, (1998).
- [90] S. Grivopoulos and B. Bamieh, Iterative algorithms for optimal control of quantum systems, *Proceedings of the 41-st IEEE Conference on Decision and Control*, Las Vegas, Nevada USA, December 2002, pp. 2687-2691.
- [91] S. Grivopoulos and B. Bamieh, Lyapunov-based control of quantum systems, *Proceedings of the 42-nd Conference on Decision and Control*, Maui, Hawaii USA, December 2003, pp. 434-437.
- [92] S. Gudder, Quantum computation, *Amer. Math. Monthly*, 110, No.3, 181-201, (2003).
- [93] G.A. Hagedorn and A. Joye, Elementary exponential error estimates for the adiabatic approximation, *J. Math. Anal. Appl.*, 267, 235-246, (2002).
- [94] B. Hall, *Lie Groups, Lie Algebras, and Representations, An Elementary Introduction*, Graduate Texts in Mathematics, 222, Springer-Verlag, 2003.
- [95] W. P. Healy, *Non-Relativistic Quantum Electrodynamics*, Academic Press, New York, 1982.
- [96] S. Helgason, *Differential geometry, Lie groups and symmetric spaces*, Academic Press, New York, 1978.
- [97] R. Hermann, *Lie Groups for Physicists*, Benjamin, New York, 1966.
- [98] S. Hill and W. Wootters, Entanglement of a pair of quantum bits, *Physical Review Letters*, 78, 5022, (1997).
- [99] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, New York, 1985.
- [100] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, New York, 1991.
- [101] M. Horodecki, Entanglement measures, *Quantum Information and Computation*, 1, No. 1, 3-26, (2001).
- [102] M. Horodecki and P. Horodecki, Reduction criterion of separability and limits for a class of protocols of entanglement distillation, *Physical Review A*, 59, 4206-4216, (1999).
- [103] M. Horodecki, P. Horodecki and R. Horodecki, Separability of mixed states: necessary and sufficient conditions, *Physics Letters A*, 223, 1-8, (1996).

- [104] S. Howison, *Practical Applied Mathematics. Modelling, Analysis, Approximation*, Cambridge Texts in Applied Mathematics, No. 38, Cambridge U.K., New York, 2005.
- [105] G. M. Huang, T. J. Tarn and J. W. Clark, On the controllability of quantum mechanical systems, *Journal of Mathematical Physics*, 24 No. 11, 2608-2618, (1983).
- [106] J. E. Humphreys, *Introduction to Lie algebras and Representation Theory*, Springer-Verlag, New York, 1972.
- [107] J. D. Jackson, *Classical Electrodynamics*, 3rd edition, Wiley, New York, 1999.
- [108] D.F.V. James, Quantum dynamics of cold trapped ions with application to quantum computation, *Appl. Phys. B*, 66, 181-190, (1998).
- [109] S. Jansen, M. B. Ruskai and R. Seiler, Bounds for the adiabatic approximation with applications to quantum computation, xxx.lanl.quant-ph/0603175
- [110] P. Jorrand and M. Mhalla, Separability of pure N-qubit states: two characterizations, *International Journal of Foundations of Computer Sciences (IJFCS)*, 14, No. 5, 797-814, (2003).
- [111] A. Joye and C. Pfister, Full asymptotic expansion of transition probabilities in the adiabatic limit, *J. Phys. A*, 24, 753-766, (1991).
- [112] V. Jurdjević, *Geometric Control Theory*, Cambridge University Press, 1997.
- [113] V. Jurdjević and H. Sussmann, Control systems on Lie groups, *Journal of Differential Equations*, 12, 313-329, (1972).
- [114] T. Kailath, *Linear Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1980.
- [115] T. Kato, On the adiabatic theorem of quantum mechanics, *J. Phys. Soc. Japan*, 5, 435-439, (1950).
- [116] K. Kawakubo, *The Theory of Transformation Groups*, Oxford University Press, New York, 1991.
- [117] H. K. Khalil, *Nonlinear Systems*, Third Edition, Prentice Hall, Upper Saddle River, NJ, 2002.
- [118] N. Khaneja, Geometric Control in Classical and Quantum Systems, Ph. D. Thesis, Division of Engineering and Applied Sciences, Harvard University, Cambridge, MA, June 2000.
- [119] N. Khaneja, R. Brockett and S. J. Glaser, Time optimal control of spin systems, *Phys. Rev. A*, 63, 032308, (2001).

- [120] N. Khaneja and S. J. Glaser, Cartan decomposition of  $SU(n)$ ; constructive controllability of spin systems and universal quantum computing, *Chem. Physics*, 267, 11, 2001.
- [121] N. Khaneja, S. J. Glaser and R. W. Brockett, Sub-Riemannian geometry and time optimal control of three spin systems: Coherence transfer and quantum gates, *Phys. Rev. A*, 65, 032301, 2002.
- [122] N. Khaneja and S. Glaser, Optimal control of coupled spin dynamics under cross-correlated relaxation, *Proceedings of the 42-nd IEEE Conference on Decision and Control*, Maui, Hawaii USA, December 2003, pp. 422-427.
- [123] N. Khaneja, F. Kramer and S. Glaser, Optimal experiments for maximizing coherence transfer between coupled spins, *Journal of Magnetic Resonance*, 173, 116-124, (2005).
- [124] D. E. Kirk, *Optimal Control Theory: An Introduction*, Prentice-Hall, Englewood Cliffs, NJ, 1970.
- [125] A.Y. Kitaev, A. H. Shen and M. N. Vyalyi, *Classical and Quantum Computation*, Graduate Studies in Mathematics. Volume 47, American Mathematical Society Translations, 2002.
- [126] R. Kock and F. Lowenthal, Uniform finite generation of three-dimensional linear Lie groups, *Canad. J. Math.*, 27, 396-417, (1975).
- [127] K. Kraus, *States, Effects, and Operations*, Vol. 190 of Lecture Notes in Physics, Springer-Verlag, Berlin, 1983.
- [128] H. Kunita, Supports of diffusion processes and controllability problems, *Proc. of Intern. Symp. Stochastic Differential Equations*, Kyoto 1976, pp. 163-185.
- [129] M. Kus and K. Zyczkowski, Geometry of entangled states, *Phys. Rev. A*, 63, 032307-13, (2001).
- [130] C. Lanczos, *The Variational Principles of Mechanics*, University of Toronto Press, 1970.
- [131] L. D. Landau and E. M. Lifshitz, *Mechanics*, Pergamon, New York, 1976.
- [132] B. Lee and Markus, *Foundations of Optimal Control Theory*, Wiley, New York, 1967.
- [133] D. Leibfried, R. Blatt, C. Monroe and D. Wineland, Quantum dynamics of single trapped ions, *Reviews of Modern Physics*, 75, 281-324, (2003).
- [134] M.H. Levitt, *Spin Dynamics: Basics of Nuclear Magnetic Resonance*, John Wiley and Sons, New York, 2001.

- [135] F. L. Lewis and V. L. Syrmos, *Optimal Control*, 2nd Edition, John Wiley and Sons, New York 1995.
- [136] N. Linden, S. Popescu and A. Sudbery, Non-local properties of multi-particle density matrices, *Physical Review Letters*, 83, 243, (1999).
- [137] S. Lloyd, Almost any quantum logic gate is universal, *Physical Review Letters*, 75, No. 2, 346, (1995).
- [138] R. Loudon, *The Quantum Theory of Light*, Third edition, Oxford University Press, New York, 1973.
- [139] H. Mabuchi and N. Khaneja, Principles and applications of control in quantum systems, *International Journal of Robust and Nonlinear Control*, 15, 647-667, (2005).
- [140] Y. Maday and G. Turinici, New formulations of monotonically convergent quantum control algorithms, *Journal of Chemical Physics*, 118, No. 18, 8191-8196, 2003.
- [141] W. Magnus, On the exponential solution of differential equations for a linear operator, *Commun. Pure Appl. Math.* 7, 649-673, (1954).
- [142] Y. Makhlin, Nonlocal properties of two-qubit gates and mixed states and optimization of quantum computation, *Quant. Inf. Proc.*, 1, 243-252, (2002).
- [143] H. Masahito, *Quantum Information. An Introduction*, Springer-Verlag, Berlin, 2006.
- [144] E. Merzbacher, *Quantum Mechanics*, John Wiley and Sons, New York, 1961.
- [145] A. Messiah, *Quantum Mechanics*, North-Holland Pub. Co., Interscience Publishers, Amsterdam, New York, 1961-62.
- [146] M. Mirrahimi, P. Rouchon and G. Turinici, Lyapunov control of bilinear Schrödinger equations, *Automatica*, 41, 1987-1994, (2005).
- [147] C. Monroe, D. Leibfried, B.E. King, D.M. Meekhof, W. M. Itano and D. J. Wineland, Simplified quantum logic with trapped ions, *Physical Review A*, 55, No. 4, R2489-R2491, (1997).
- [148] D. Montgomery and H. Samelson, Transformation groups of spheres, *Ann. Math.*, 44, 454-470, (1943).
- [149] D. Montgomery and L. Zippin, *Topological Transformation Groups*, New York, Vol. 1, Interscience Tracts in Pure and Applied Mathematics, 1955.
- [150] F. D. Murnaghan, *The Theory of Group Representations*, Johns Hopkins Press, Baltimore, 1938.

- [151] F. D. Murnaghan, *The Orthogonal and Symplectic Groups*, Communications of the Dublin Institute for Advanced Studies. Series A, No. 13, 1958.
- [152] F. D. Murnaghan, *The Unitary and Rotation Groups*, Lecture Notes in Applied Mathematics, Vol. 3, Spartan Books, Washington, D.C., 1962.
- [153] G. Nenciu, On the adiabatic theorem of quantum mechanics, *J. Phys. A*, 13, L15-L18, (1980).
- [154] D. Nguyen and T. Odagaki, Quantum and classical electrons in a potential well with uniform electric field, *Am. J. Phys.*, 55, 466-469, (1987).
- [155] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, U.K., New York, 2000.
- [156] M. Ohya and D. Petz, *Quantum Entropy and its Use*, Springer-Verlag, Series in Theoretical and Mathematical Physics, Berlin, New York, 2004.
- [157] J. A. Oteo and J. Ros, From time-ordered products to Magnus expansion, *Journal of Mathematical Physics*, 41, No. 5, 3268-3277, (2000).
- [158] G. K. Pedersen, *Analysis Now*, Springer-Verlag, New York, 1989.
- [159] A. Peirce, M. Dahleh and H. Rabitz, Optimal control of quantum mechanical systems: Existence, numerical approximations and applications, *Physical Review A*, 37, No.12, 4950, 1987.
- [160] A. Peres, Separability criterion for density matrices, *Phys. Rev. Lett.*, 77, No. 8, 1413-1415, (1996).
- [161] E. Pinch, *Optimal Control and the Calculus of Variations*, Oxford University Press, Oxford, U.K., New York, 1993.
- [162] R. Plass, K. Egan, C. Collazo-Davila, D. Grozea, E. Landree, L. D. Marks, and M. Gajdardziska-Josifovska, Cyclic ozone identified in magnesium oxide (111) surface reconstructions, *Physical Review Letters*, 81, 4891, (1998).
- [163] M. B. Plenio and S. Virmani, An introduction to entanglement measures, *Quant. Inf. Comp.* 7, 1, (2007).
- [164] J. Preskill, Lecture Notes on Quantum Computation, <http://www.theory.caltech.edu/~preskill/ph229/>
- [165] H. Rabitz, R. de Vivie-Riedle, M. Motzkus and K. Kompa, Whither the future of controlling quantum phenomena, *Science*, 288, 824-828, (2000).
- [166] H. Rabitz, M. Hsieh and C. Rosenthal, Quantum optimally controlled transition landscapes, *Science*, 303, 998, (2004).

- [167] E. M. Rains, Polynomial invariants of quantum codes, *IEEE Transactions on Information Theory*, 46, 54, (2000).
- [168] V. Ramakrishna, M. V. Salapaka, M. Dahleh, H. Rabitz and A. Peirce, Controllability of molecular systems, *Physical Review A*, 51, No. 2, 960-966, (1995).
- [169] V. Ramakrishna, K.L. Flores, H. Rabitz and R. J. Ober, Quantum control by decompositions of  $SU(2)$ , *Phys. Rev. A*, 62, 053409, (2000).
- [170] V. Ramakrishna, R. J. Ober, K. L. Flores and H. Rabitz, Control of a coupled two spin systems without hard pulses, *Physical Review A*, 65, 063405, (2002).
- [171] C. Rangan, A.M.Bloch, C.Monroe and P. H. Bucksbaum, Control of trapped-ion quantum states with optical pulses, *Physical Review Letters*, 92, No. 11, 113004, (2004).
- [172] K.N.S.Rao, *Linear Algebra and Group Theory for Physicists*, New Age International, New Delhi, 1996.
- [173] B. Reichardt, The quantum adiabatic optimization algorithm and local minima, *Proceedings of the 36-th Annual ACM Symposium on Theory of Computing (STOC)*, Chicago, IL, USA, June 13-16, pp. 502-510, 2004.
- [174] S. A. Rice and M. Zhao, *Optical Control of Molecular Dynamics*, New York, Wiley, 2000.
- [175] R. T. Rockafellar, *Convex Analysis*, Princeton Landmarks in Mathematics and Physics, Princeton University Press, Princeton, NJ, 1970.
- [176] R. Romano, *Dissipative Dynamics in Particle Physics*, Ph. D. Thesis, Universita' degli Study di Trieste, Italy, 2002, arXiv:hep-ph/0306164.
- [177] R. Romano and D. D'Alessandro, Incoherent control and entanglement for two dimensional coupled systems, *Physical Review A*, 73, 022323, (2006).
- [178] R. Romano and D. D'Alessandro, Environment mediated control of a quantum system, *Physical Review Letters*, 97, No. 8, 080401, (2006).
- [179] W. Rudin, *Real and Complex Analysis*, 3-rd edition, Mc Graw-Hill, New York, 1987.
- [180] O. Rudolph, Further results on the cross norm criterion for separability, arXiv:quant-ph/0202121.
- [181] O. Rudolph, A note on ‘A matrix realignment method for recognizing entanglement’, arXiv:quant-ph/0205017 v1.
- [182] O. Rudolph, Some properties of the computable cross-norm criterion for separability, *Physical Review A*, 67, 032312, (2003).

- [183] P. Rungta, V. Buzek, C. M. Caves, M. Hillery and G. J. Milburn, Universal state inversion and concurrence in arbitrary dimensions, *Physical Review A*, 64, 042315, (2001).
- [184] A. A. Sagle and R. E. Walde, *Introduction to Lie Groups and Lie Algebras*, Academic Press, New York, 1973.
- [185] J. J. Sakurai, *Modern Quantum Mechanics*, Addison-Wesley Pub. Co., Reading MA, c1994.
- [186] J. Salomon, Limit points of the monotonic schemes in quantum control, *Proceedings of the 44-st IEEE Conference on Decision and Control*, Seville, December 2005.
- [187] G. Schaller, S. Mostame and R. Schützold, General error estimate for adiabatic quantum computing, *Physical Review A*, 73, 062307, (2006).
- [188] S. G. Schirmer, Quantum control using Lie group decompositions, in *Proceedings of the 40-th IEEE Conference on Decision and Control*, 4-7 Dec. 2001, Pages 298-303, vol. 1.
- [189] S. G. Schirmer, A. D. Greentree, V. Ramakrishna and H Rabitz, Constructive control of quantum systems using factorization of unitary operators, *J. Phys. A* 35, 8315-8339, (2002).
- [190] S. G. Schirmer, J. V. Leahy and A. I. Solomon, Degrees of controllability for quantum systems and applications to atomic systems, *J. Phys. A*, 35, 4125-4141, (2002).
- [191] M. Shapiro and P. Brumer, Quantum control of chemical reactions, *J. Chem. Soc., Faraday Trans.*, 93, (No.7), 1263-1277, (1997).
- [192] M. Shapiro and P. Brumer, *Principles of the Quantum Control of Molecular Processes*, John Wiley and Sons, Inc., Hoboken, NJ, 2003.
- [193] C. P. Slichter, *Principles of Magnetic Resonance*, Springer-Verlag, New York, 1996.
- [194] T.A. Springer, *Invariant Theory*, Lecture Notes in Mathematics, vol. 585, Springer, Berlin, 1997.
- [195] A. Steane, The ion trap quantum information processor, *Appl. Phys. B*, 64, 623-642, (1997).
- [196] S. Sternberg, *Lectures on Differential Geometry*, Prentice Hall, Englewood Cliffs, NJ, 1964.
- [197] H. J. Sussmann, Lie brackets, real analyticity and geometric control, in *Differential Geometric Control Theory*, R. W. Brockett, R. S. Millman and H. J. Sussmann eds., pp. 1-116, Birkhauser, Boston, 1983.
- [198] H. J. Sussmann and V. Jurdjevic, Controllability of nonlinear systems, *Journal of Differential Equations*, 12, 95-116, 1972.

- [199] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, McGraw-Hill Publishing Company, New York, 1989.
- [200] P. I. Tamborenea and H. Metiu, Localization and entanglement of two interacting electrons in a quantum-dot molecule, *Europhys. Lett.*, 53, No. 6, 776-782, (2001).
- [201] D. J. Tannor, V. Kazakov and V. Orlov, Control of Photochemical branching. Novel procedures for finding optimal pulses and global upper bounds, in *Time dependent quantum molecular dynamics*, J. Broeckhove and L. Lathouwers eds., 347-360, Plenum, 1992.
- [202] B. Terhal, Detecting quantum entanglement, *Theoretical Computer Science*, 287, No.1, 313, (2002).
- [203] S. H. Tersigni, P. Gaspard and S. Rice, On using shaped light pulses to control the selectivity of product formation in a chemical reaction: An application to a multiple level system, *J. Chem. Phys.*, 93, No. 3, 1670-1680, (1990).
- [204] W. Tung, *Group Theory in Physics*, World Scientific, Singapore, 1985.
- [205] G. Turinici, Controlabilite exacte de la population des etats propres dans les systemes quantiques bilineaires. (French), *C. R. Acad. Sci. Paris Sr. I Math.* 330, No. 4, 327–332, (2000).
- [206] Gabriel Turinici and Herschel Rabitz. Quantum wave function controllability. *Chem. Phys.*, 267, 1-9, (2001).
- [207] G. Turinici and H. Rabitz, Wavefunction controllability for finite-dimensional bilinear quantum systems, *J. Phys. A: Math. Gen.*, 36, 2565-2576, (2003).
- [208] A. Uhlmann, Entropy and optimal decompositions of states relative to a maximal commutative subalgebra, *Open Systems Inf. Dyn.*, 5, No. 3, 209-228, (1998).
- [209] A. Uhlmann, Optimizing entropy relative to a channel or a subalgebra, *Open Systems Inf. Dyn.*, arxiv:quant-ph/9701014.
- [210] A. Uhlmann, Fidelity and concurrence of conjugated states, *Phys. Rev. A*, 62, 032307, (2000).
- [211] U. Vaidya, D. D'Alessandro and I. Mezić, Control of Heisenberg spin systems; Lie algebraic decompositions and action-angle variables, in *Proceedings 42-nd Conference on Decision and Control* Vol. 4, pp 4174-4178, Maui HI, Dec. 2003.
- [212] W. van Dam, M. Mosca and U. Vazirani, How powerful is adiabatic quantum computation?, in *Proceedings 42-nd IEEE Symposium on Foundations of Computer Science* (Las Vegas, NV, 2001), 279-287.

- [213] W. Van der Graaf, *Lie Algebras; Theory and Algorithms*, North-Holland, 2000.
- [214] R. Van Handel, J. K. Stockton and H. Mabuchi, Feedback control of quantum state reduction, *IEEE Trans. Automat. Control*, 50, No. 6, 768-780, (2005).
- [215] P. Vettori, On the convergence of a feedback control strategy for multi-level quantum systems, in 15-th International symposium MTNS 2002, University of Notre Dame U.S.A.
- [216] Y. Yamamoto and A. Imamoglu, *Mesoscopic Quantum Optics*, Wiley-Interscience Publication, New York, 1999.
- [217] F. W. Warner, *Foundations of Differentiable Manifolds and Lie Groups*, Springer, New York, 1983.
- [218] N. Weaver, On the universality of almost every quantum logic gate, *J. Math. Phys.*, 41, (No. 1), 240-243, (2000).
- [219] A. M. Weiner, D. E. Leaird, J. S. Patel and J. R. Wullert II, Programmable shaping of femtosecond optical pulses by use of 128-element liquid crystal phase modulator, *IEEE Journal of Quantum Electronics*, 28, No. 4, 908-920, (1992).
- [220] H. Weyl, *The Classical Groups*, Princeton University Press, Princeton, NJ, 1939.
- [221] H. M. Wiseman, *Quantum Trajectories and Feedback*, Ph. D. Thesis, Physics Department, University of Queensland, 1994.
- [222] A. Wong and N. Christensen, A potential multipartite entanglement measure, *Physical Review A*, 63, 044301, (2001).
- [223] W. K. Wootters, Entanglement of formation of an arbitrary state of two qubits, *Physical Review Letters*, 80, 2245-2248, (1998).
- [224] W. K. Wootters, Entanglement of formation and concurrence, *Quantum Information and Computation*, 1, No. 1, 27-44, (2001).
- [225] K. Wüthrich, *NMR in Biological Research: Peptides and Proteins*, North-Holland, Amsterdam, 1976.
- [226] R. Zare, Laser control of chemical reactions, *Science*, 279, 1875-1879, (1998).
- [227] J. Zhang, J. Vala, S. Sastry and K. Whaley, Geometric theory of non-local two-qubit operations, *Physical Review A*, 67, 042313, (2003).
- [228] W. Zhu and H. Rabitz, A rapid monotonically convergent iteration algorithm for quantum optimal control over the expectation value of a positive definite operator, *Journal of Chemical Physics*, 109, No. 2, 385-401, (1998).

- [229] W. Zhu and H. Rabitz, Attaining optimal controls for manipulating quantum systems, *International Journal of Quantum Chemistry*, 93, 50-58, (2003).