

UNIVERSITY OF CALIFORNIA
SANTA BARBARA

Optimal Control of Quantum Systems

A Dissertation submitted in partial satisfaction
of the requirements for the degree of
Doctor of Philosophy

in

Mechanical Engineering

by

Symeon Grivopoulos

Committee in charge:

Professor Bassam Bamieh, Chairperson

Professor Francesco Bullo

Professor Petar Kokotović

Professor Igor Mezić

December 2005

UMI Number: 3202700



UMI Microform 3202700

Copyright 2006 by ProQuest Information and Learning Company.
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

The dissertation of Symeon Grivopoulos is approved:

Professor Francesco Bullo

Professor Petar Kokotović

Professor Igor Mezić

Professor Bassam Bamieh, Committee Chairperson

December 2005

To the memory of my parents

Acknowledgements

I would like to thank my academic advisor professor Bassam Bamieh for his support throughout my graduate studies as well as his excellent advice. I very much enjoyed my interactions with him, both technical and social, and he helped me mature as a researcher and as a person.

I express my gratitude to the late professor Mohammed Dahleh who was my original graduate advisor in the ME department until his death.

I thank the members of my dissertation committee, professors Petar Kokotović, Igor Mezić and Francesco Bullo for their advice on how to improve my dissertation.

Many thanks to the director and the faculty of the CCDC for the excellent research environment they have created and the great classes they have taught.

I am also grateful to my college teachers, professors Dimitris Ghikas and Dimitris Sourlas of the University of Patras for opening my eyes to the beauty of Physics and Mathematics and for their support.

I was fortunate to share my office with Mihailo Jovanović, Aruna Ranaweera, Makan Fardad, Hana El-Samad, Zoran Levnjajić and George Mathew. Not to forget, of course, our “next office neighbors” Mariateresa Napoli, Umesh Vaidya, Thomas John, Vasu Salapaka, Niklas Karlsson, Ove Storset and Erkut Aykutlug. My thanks to all of you for your friendship and support.

Many thanks and my love to my friends from Greece and the friends that I made in the States: Costa, Vassili, Niko, Vaso, Gianni, Christo, Pano, Vangeli,

Ioanna, Soeren and Zaharenia, thanks for making my life richer and my heart bigger.

Finally, my love goes to my family. My beloved parents, Nikos and Anastasia, may not be with me anymore, but their love still burns inside me. My sister Angela, my aunt Eleni and my little nephew Theodoros, the new ray of light in our family, you are always with me.

Curriculum Vitae of Symeon Grivopoulos

Education

PhD in Mechanical Engineering. University of California, Santa Barbara, December 2005.

BSc in Physics. University of Patras, July 1995.

Publications

Journals

1. S. GRIVOPOULOS & B. BAMIEH, *Optimal population transfers in a quantum system for large transfer time*, submitted to the IEEE Transactions on Automatic Control.
2. S. GRIVOPOULOS & B. BAMIEH, *Optimal control of rotational cooling: An efficient approximation based on averaging*, submitted to the Journal of Chemical Physics.
3. S. GRIVOPOULOS & B. BAMIEH, *Optimal feedback control of a monitored spin: A solution*, in preparation.

Refereed Proceedings

1. S. GRIVOPOULOS & B. BAMIEH, *Iterative algorithms for optimal control of quantum systems*, in Proceedings of the 41st IEEE Conference on Decision and Control, Las Vegas, NV, 2002.
2. S. GRIVOPOULOS & B. BAMIEH, *Lyapunov-based control of quantum systems*, in Proceedings of the 42nd IEEE Conference on Decision and Control, Maui, HI, 2003.
3. S. GRIVOPOULOS & B. BAMIEH, *Optimal population transfers for a quantum system in the limit of large transfer time*, in Proceedings of the 2004 American Control Conference, Boston, MA, 2004.

Abstract

Motivated by developments and problems in a number of disciplines including Quantum Chemistry, Information and Optics, the theory of Control of Quantum Systems has emerged. Its goal is to apply the tools and methods of Control Theory in the analysis and design of scientific and engineering applications of Quantum Systems. At the same time, Control Theory itself is been enriched by new models and paradigms.

Our work focuses on the design of control fields that achieve given state transfers with the minimum amount of energy expenditure. Besides their inherent mathematical interest, such optimal designs are closely tied to the dynamics of the underlying system and reveal much about the interplay of dynamics and control.

In the first part of our work, we consider energy-optimal transfers in a general isolated quantum system, for example an atom or a molecule. By examining the large-time limit of these optimal transfer problems, we uncover the general structure of the optimal controls. Moreover, we reduce the computational complexity of the problem significantly.

In the second part, we examine similar problems for open quantum systems, that is, quantum systems that interact with their environment. This interaction creates dissipative effects in the system. Although one usually wants to resist these effects, there are instances, such as the cooling of internal molecular motion, that one can effectively use dissipation mechanisms to one's advantage. We apply techniques similar to those developed for isolated systems to design “cooling” electric (laser) fields for molecular rotations and demonstrate our method in a three-state Lambda system.

The aforementioned designs are open-loop because sensing and feedback are

infeasible for the applications at hand. However, there has been remarkable technological progress in the monitoring of single/small numbers of quantum systems. Also, from the theoretical side, an adequate formalism has been developed for the description of continuously monitored quantum systems along with a theory of optimal feedback. In the third part, we capitalize on these developments and design an optimal feedback control for the stabilization/preparation of a desired state of a continuously measured spin system.

Contents

Acknowledgements	iv
Curriculum Vitae	vi
Abstract	vii
List of Figures	xii
1 Introduction	1
1.1 General introduction	1
1.2 Contributions of the thesis	4
1.3 Organization of the thesis	7
2 Optimal transfers in isolated quantum systems	8
2.1 Introduction to quantum systems	9
2.1.1 Quantum Mechanics at a glance	9
2.1.2 Controlled quantum systems	14
2.1.3 Two model systems	16
2.1.4 Bohr frequencies as resonance frequencies	19
2.2 Overview of problem and results	26
2.3 Optimal population transfers	30

2.4	Optimal population transfers for an averaged system	34
2.5	Main results	41
2.6	Examples	44
2.7	Proof of main results	54
2.8	Summary	65
2.9	Appendices to Chapter 2	66
3	Optimal transfers in open quantum systems	71
3.1	Quantum Mechanics of open systems	72
3.1.1	The density matrix formalism	72
3.1.2	The Lindblad equation	80
3.2	Laser cooling of molecules: The role of dissipation	86
3.3	Optimal control of dissipative dynamics	90
3.4	An efficient approximation based on averaging	94
3.4.1	Separating the two time-scales of the dynamics: An averaged system	94
3.4.2	Optimal transfers in the averaged system	97
3.4.3	Approximate optimality of the resulting control $u(t)$	99
3.5	Example: Numerical solution of the three-level Λ system	103
3.6	Summary	105
4	Lyapunov based control of quantum systems	107
5	Optimal feedback control of a continuously monitored spin	114
5.1	Optimal feedback control of a continuously monitored spin	115
5.2	Solution for the case of perfect detection efficiency	121
5.3	Summary	125

List of Figures

2.1	The Morse potential for the OH bond	18
2.2	Averaged (dashed line) and exact (full line) populations of the two-state system under the approximate optimal control, for $T = 10\pi$ and $T = 6\pi$	46
2.3	Approximate optimal control for the transition $1 \rightarrow 3$ in $T = 20\pi$ and the corresponding frequency profiles.	50
2.4	Averaged (dashed lines) and exact (full lines) populations of the three-state system under the approximate optimal control.	51
2.5	Average populations and profile intensities vs. (scaled) time for a locally optimal transfer $1 \rightarrow 10$	52
2.6	Average populations and profile intensities vs. (scaled) time for a locally optimal transfer $10 \rightarrow 22$	53
3.1	Spontaneous emission in a two-state system.	86
3.2	A three level Lambda system with spontaneous emissions from level 2 to 1 and 2 to 3. There is no direct emission from 3 to 1.	87
3.3	Optimal profile w_{23}	104
3.4	Average populations calculated by our method and optimal populations calculated in Sklarz et al.	105
3.5	Average coherence \bar{r}_{23}	106

4.1	Selective excitation of an intermediate vibrational state of the OH bond using a Lyapunov based feedback control. The transfer is 90% complete in 10 ps.	113
5.1	Schematic of an experiment for continuous measurement and control of a spin. The spin interacts with a probe laser, which is measured continuously using photodetectors. A magnetic field is used for feedback.	115
5.2	Steady-state value function \tilde{V} on the boundary	125
5.3	Optimal feedback on the boundary	125
5.4	Steady-state value function V on the disk	126
5.5	Optimal feedback on the disk	126
5.6	Vector field generated by the feedback control	127
5.7	Typical trajectories of the closed-loop system	127

Chapter 1

Introduction

1.1 General introduction

The subject of Quantum Control was created to address, from a Systems and Control theoretic point of view, a number of problems from disciplines such as Quantum Chemistry, Quantum Information and Quantum Optics. Control Theory lends systematic tools of analysis and design to these disciplines and in return is enriched with new models, applications and paradigms. In the following, we will mention some of these problems trying to give a feeling of the scope and breadth of Quantum Control Theory. Our presentation will by no means be exhaustive. A sense of how much this field has grown and how active it is can be gained from a simple search of the online Physics Preprint Archive at *xxx.lanl.gov*.

Quantum Chemistry was historically a main importer of Control Theory techniques, mainly optimal design. One of the main goals of Quantum Chemistry has been to manipulate reactions, especially between polyatomic molecules, to desired outcomes. In particular, it is desired to selectively break chemical bonds in polyatomic molecules so that new ones can be created. In many cases, traditional techniques (adjusting the temperature and pressure of the reacting system, insert-

ing catalysts, etc.) have no effect. The introduction of high-power lasers along with techniques to manipulate their pulses seems to offer a precision tool for this task. Chemists realized though that the laser pulses need to be carefully designed to ensure collaboration with the complex molecular (vibrational) dynamics. Thus, ideas of optimization and optimal control started to enter the arena [1, 2, 3, 4, 5]. [6, 7] contain accounts of the situation at the beginning and at the end of the 90s both from a theoretical and an experimental perspective, as well as large lists of references. The problem of designing optimal laser pulses for selective excitation of vibrational modes, molecular dissociation (breaking of bonds) and, in general, steering to desired target states has been a major motivator for the introduction of a Control Theory point of view.

Another source of problems amenable to Control Theory techniques is Nuclear Magnetic Resonance (NMR) Spectroscopy. NMR spectroscopy uses the response of the elementary magnetization of nuclei (spin) to magnetic pulses to infer information about the kinds of nuclei present in a molecule and their relative positions. A sequence of magnetic field pulses is applied to a sample of molecules (e.g. a water solution of proteins) and then radiation is collected from the relaxing sample. The spectral analysis of the radiation offers information on the types of chemical groups present in the molecule and their distances. NMR has become a very powerful experimental technique, especially in BioChemistry, for identifying the structure of macromolecules. There are several important problems arising in the design of these pulses that can be posed as optimal transfer problems of spin dynamics [8, 9, 10, 11] (for a Systems theoretic formulation of NMR experiments see [12]). Besides the aforementioned problems, there are many more applications of control-theoretic methods in Quantum Chemistry problems like cooling of molecules, control of molecular alignment, identification of parameters of molecular

Hamiltonians, etc.

A discipline that also promoted the introduction of a control engineering point of view for quantum systems is Quantum Computing. Quantum Computing flourished as a subject during the 90s after the realization that, in many cases, it can offer distinct advantages over classical computing. However, the developments were mainly focused on the theory of computation, rather its implementation via physical systems. The attempt to realize quantum computing devices, created the need that a number of practical design problems be addressed: How will quantum states be prepared in the beginning of a computation, manipulated during one and measured at the end? How can external driving fields (electric or magnetic, depending on the system) be designed that will affect a prescribed unitary transformation on the states (quantum gates)[13, 14]? Even more, how can one protect quantum information from environmental effects, namely decoherence [15, 16, 17]? An introduction to theoretical and practical aspects of Quantum Computing along with an exhaustive list of references is [18].

A lot of the recent impetus in Quantum Control has come from Quantum Optics. Already from the beginning of the 90s there existed in the Optics literature a theory of quantum feedback [19] which, however, contained no measurement step (i.e. some output, usually some photocurrent, would be fed back to the system dynamics through a gain, without its value being “read”). Although many things can be achieved with this limited form of feedback [20, 21, 22, 23], it became clear that a measurement and estimation step would be extremely advantageous [24, 25]. It is interesting that some of the mathematical tools needed for this, namely a quantum filtering equation, had already been developed [26]. Combined with advancing experimental techniques for continuous observation of quantum systems [27], these developments open up a new era for Quantum Control. Ap-

plications to preparation/stabilization of atomic states, precision measurement, feedback cooling of atoms in cavities and quantum information processing are envisioned. This promises to be a very fertile area for developments in the theory and applications of Quantum Control.

1.2 Contributions of the thesis

The topic of this thesis is the design of energy-optimal control laws for state transfers in quantum systems. Though certainly not the only relevant problems in Quantum Control Theory, state transfer problems are central to many applications. They may be about selective excitation of a vibrational mode in a molecule, polarization or coherence transfer in NMR or state preparation/stabilization in a continuously monitored quantum system.

We have concentrated on two types of transfer problems. The first one, transfer of populations, is motivated by applications to Quantum Chemistry such as selective excitation of vibrational eigenstates, molecular dissociation and cooling of internal molecular motions (vibrations and rotations). The systems that one deals with in Quantum Chemistry, namely molecules, have many degrees of freedom and Hamiltonians without any symmetries (usually). This makes the use of constructive design techniques essentially impossible. Optimal control, on the other hand, offers a general setup that can handle large systems as well as the ability to include several constraints in the design from the start. Also important is that optimal designs for appropriate objectives produce control laws that cooperate with the dynamics of the system rather than overpower it. This is essential for systems with very few controls available, as is the case with quantum systems (where usually one has one electric or magnetic field available for control purposes). In particular, we have found that energy-optimal designs have a lot of

interesting structure.

Energy-optimal designs have been considered in the Quantum Control literature for a variety of problems in both isolated and open quantum systems [2, 3, 4, 28, 29, 30, 31, 32, 33, 34, 35]. Except for the few cases where an analytical solution can be found (for two- and three-dimensional systems), one must resort to numerical solution techniques. However, these techniques fail to reveal a general structure of the optimal laws and are computationally expensive. We discovered that, by examining the large time limit of these optimal transfer problems, we can both uncover a general structure of the solutions and significantly reduce the computational cost of their (numerical) solution. In particular, we found that for large times, there is a time-scale separation in the dynamics. There is the short-time natural dynamics and the long-time evolution due to control. The effects of the former can be incorporated in the later using averaging. The original transfer problem is then reduced to a transfer problem for an “averaged” system without drift and with only one time-scale, the long-time scale of the transfer. We prove that there is a one-to-one correspondence between the solutions of the original and the “averaged” transfer problem. Hence, one can solve the “averaged” transfer problem and use its solution to obtain the solution of the original transfer problem. The computational cost reduction comes from the fact that the “averaged” system has only one time-scale, the long-time scale. The solution of the original problem, expressed in terms of that of the “averaged” transfer problem exhibits the time-scale separation and clarifies the effect of the control. The form of the optimal control is especially simple and intuitive: It is a sum of Bohr frequency sinusoids modulated by slow profiles which are calculated from the slow averaged problem. There are many transfer problems which by nature require a transfer time at least 3 to 4 orders of magnitude larger than the time-scale

of their natural dynamics. But, more than that, we believe that this approach is the most effective, computationally and conceptually, for population transfer problems, even for shorter transfer times. We have proved the results above for *exact* population transfers in isolated quantum systems. We believe that they can be extended to other population transfer problems with the same conclusions. We have also extended the technique to open systems and treated the problem of designing cooling laser pulses for molecular rotations in this framework.

The second type of problem is a regulation problem, that of stabilization of a desired spin eigenstate of a continuously monitored spin system. It has attracted much attention in the quantum control community, see, for example, the recent works [36, 37, 38, 39, 40]. In particular, we focus on the feedback stabilization of a spin- $\frac{1}{2}$ system which is a prototypical problem, because the spin- $\frac{1}{2}$ is the simplest quantum mechanical degree of freedom. We use an optimal feedback approach with an objective consisting of a control-energy term and a state-penalty term. Since measurements provide only partial information about the state of the system, this problem is analogous to optimal feedback control of partially observed stochastic systems in classical Control Theory. The solution is similar: The problem is separated into a filtering problem, where the state of the system is continuously estimated based on incoming measurements and an optimal state feedback problem for the dynamics of the filter. We set up the problem of the optimal feedback stabilization of the filter dynamics and derive the corresponding Bellman equation. In the case of perfect detection efficiency the boundary of the state space is a forward invariant set for the dynamics. We pose a “reduced” version of the original stabilization problem on the boundary and solve its Bellman equation. Then, we use this boundary solution as a boundary condition for the full Bellman equation and obtain a complete solution to the problem.

1.3 Organization of the thesis

The thesis is organized as follows:

- In Chapter 2, we introduce the basic elements of Quantum Mechanics and set up the optimal population transfer problem. We review some previous work on it and comment on various conceptual and computational issues. Then, we introduce the “averaged” transfer problem and establish the connection between the solutions of the two problems. We conclude with some analytical and numerical examples of the presented method.
- In Chapter 3, we begin by extending the basic quantum mechanical formalism to open systems. Then, we review the problem of rotational cooling of molecules and set it up as an optimal population transfer problem. We present an approximate solution method based again on an “averaged” transfer problem. We conclude with the example of a three-level Λ system where we compare our solution to the optimal cooling problem with that obtained by a different method.
- Chapter 4 contains some work on Lyapunov-based control design for isolated quantum systems (which is meant to be implemented in *open-loop* fashion).
- In Chapter 5, we introduce the continuously monitored spin - $\frac{1}{2}$ system, present and analyze the associated filtering equation and set up the optimal feedback stabilization problem of a desired eigenstate for the filter dynamics. We conclude with a numerical solution for the optimal feedback law.
- Chapter 6 contains concluding remarks and some thoughts on future research directions.

Chapter 2

Optimal transfers in isolated quantum systems

Our presentation is organized as follows: We begin in section 2.1 with a brief introduction to Quantum Mechanics and introduce relevant notions and facts. We present two simple models of relevance to our work (and to Quantum Control, in general) and build some initial intuition about the dynamics of these systems. In section 2.2, we overview the optimal transfer problem we consider in this chapter along with the main conclusions of our work. In section 2.3, we set the optimal population transfer problem for a finite dimensional quantum system in detail, review approaches to its solution and comment on the difficulties associated with these approaches. In section 2.4, we consider the quantum system with a certain class of controls and show how, for large transfer times T , it is approximated to first order in a $\frac{1}{T}$ perturbation expansion by an “averaged” control system. Then, we set up an optimal population transfer problem for this averaged system. Section 2.5 contains our main result, namely that, for large transfer time, the solution to the original optimal control problem is approximated by the solution of the optimal population transfer problem for the averaged system. This implies

that the optimal control for the original problem belongs to the class of controls used to transform the original system to the averaged one and this provides a useful characterization of it. We demonstrate our approach with some examples in section 2.6. The proofs of our results are contained in section 2.7. Section 2.8 concludes.

2.1 Introduction to quantum systems

2.1.1 Quantum Mechanics at a glance

We collect here a few basic definitions and axioms of Quantum Mechanics. Our intent is to establish terminology rather than offer an introduction to this vast subject that lies in the heart of modern Physics. For a comprehensive introduction, [41, 42] are two excellent and fairly modern references. The four-volume classic work [43] contains a wealth of information on the mathematical structure of Quantum Mechanics.

In general, to define a dynamical system one needs to provide two pieces of information: A set of (dynamical) variables (equivalently, a state space) and a law for their evolution. In the case of quantum systems, the relevant state spaces are complex Hilbert spaces, i.e. complete, separable infinite-dimensional complex Euclidean spaces. All such spaces are equivalent to $l_2(\mathbb{C})$, the space of square-summable sequences of complex numbers, a well-known fact from Functional Analysis. Nevertheless, finite-dimensional state spaces occur as well. For example, all microscopic particles possess a property called **spin**, a kind of internal angular momentum (see section 2.1.3). Ignoring the dynamics of the particle in the physical space, spin dynamics can be described in a finite-dimensional complex state space \mathbb{C}^N . Also, in many cases the interesting dynamics in the system

may take place in a finite-dimensional subspace of the Hilbert space and we wish to look at a reduced model. In this section we will address both cases together and (slightly abusing the established mathematical terminology) we will refer to the state space of a quantum system as its Hilbert space. Its dimension, N , will be either finite or countably infinite. In the rest of the work we concentrate on finite-dimensional systems obtained either in their own, or by infinite-dimensional systems through a reduction process.

Let \mathcal{H} denote the Hilbert space of a quantum system. Traditionally, the state of the system is denoted by ψ . The evolution law is the famous **Schrödinger equation**:

$$i\hbar \dot{\psi} = H\psi, \quad (2.1)$$

where H is a self-adjoint operator defined in a dense domain of \mathcal{H} , referred to as the **Hamiltonian operator** of the system. H might also be a function of time. We will not discuss questions of existence of solutions to (2.1) (in the infinite-dimensional case) and instead refer the reader to [43]. \hbar is a physical constant with dimensions of action (energy \times time) and the numerical value

$$\hbar = 1.054 \times 10^{-34} \text{ Joule.sec.}$$

$h = 2\pi\hbar$ is known in Physics as Planck's constant. This evolution law is linear and unitary (norm preserving) on account of the self-adjointness of H :

$$\begin{aligned} \frac{d}{dt} \|\psi\|^2 &= \frac{d}{dt} (\psi, \psi) = \left(\frac{d\psi}{dt}, \psi \right) + \left(\psi, \frac{d\psi}{dt} \right) = (-iH\psi, \psi) + (\psi, -iH\psi) \\ &= i(H\psi, \psi) - i(\psi, H\psi) = 0. \end{aligned}$$

(The usual notation for inner products and norms is used). Traditionally, states are normalized to 1. As we will see in a moment, a total phase factor in ψ is

physically meaningless, thus all vectors modulo a total phase factor represent the same physical state ($\psi \sim e^{i\phi} \psi$, $\phi \in [0, 2\pi]$). Thus, the actual state space of an N -dimensional quantum system is the complex projective space in N dimensions, $\mathbb{C}P^N$.

In a physical theory, an interpretation of the various mathematical objects in terms of measurable physical quantities must exist. For Quantum Mechanics, this goes as follows: Every physical (measurable) quantity corresponds to a self-adjoint operator in \mathcal{H} . To avoid technicalities, we consider only operators with discrete (point) spectrum. Treatment of operators with continuous spectrum can be found in [43] (see also [44] for a concrete example). Let Q be the operator corresponding to the physical quantity q , e.g. energy. Since Q is self-adjoint, its eigenvalues are real and its normalized eigenvectors form an orthonormal basis for \mathcal{H} :

$$Qv_i = q_i v_i, \quad i = 1, \dots, N.$$

Then, every possible experiment to measure q will yield one of the q_i 's, i.e. the q_i 's are *the only possible* physical values q can take. The reason behind the prominence of self-adjoint operators in Quantum Mechanics is the reality of their spectrum which is interpreted as the set of measurable values of the corresponding physical quantities. Although (2.1) is a deterministic equation, the interpretation of the formalism is statistical, that is, one uses the knowledge of the state ψ of the system to calculate a priori probabilities for the outcome of measurements of physical quantities. The probability of the outcome q_i in a measurement of q when the state of the system is ψ is given by

$$\text{prob}(q = q_i; \psi) = |(v_i, \psi)|^2.$$

ψ is interpreted as a state of knowledge about the quantum system. Hence, right after a measurement of q that yielded the value q_1 , the state of the system is v_1 since q is known with certainty. This discontinuous change in the state after a measurement is often referred to as “collapse of the state”. Notice that completeness of the set $\{v_i\}_{i=1,\dots,N}$ implies that

$$\sum_{i=1}^N \text{prob}(q = q_i; \psi) = 1.$$

Also, for $i \neq j$, $(v_i, v_j) = 0$, which translates to

$$\text{prob}(q = q_i; v_j) = \text{prob}(q = q_j; v_i) = 0.$$

Both of these are required for a consistent probabilistic interpretation. Observe that if one multiplies a state vector ψ by a phase $e^{i\phi}$, nothing changes in its physical properties. This is the reason why one identifies all vectors that differ by an overall phase factor. The expression for the expectation value of Q in the state ψ ,

$$\begin{aligned} \langle Q \rangle &= \sum_{i=1}^N q_i \cdot \text{prob}(q = q_i; \psi) = \sum_{i=1}^N q_i |(v_i, \psi)|^2 = (\psi, (\sum_{i=1}^N q_i v_i v_i') \psi) \\ &= (\psi, Q\psi), \end{aligned} \tag{2.2}$$

is one of the fundamental equations in Quantum Mechanics. Perhaps the most striking consequence of the physical interpretation of the formalism, is the fact that physical observables whose corresponding quantum mechanical operators do not commute, cannot be measured *simultaneously* and hence cannot be assigned values simultaneously. Indeed, only when the self-adjoint operators A and B commute, i.e. $[A, B] \doteq AB - BA = 0$, they have a common set of eigenvectors $\{q_i\}$ and

hence after a measurement that yielded, say, the eigenstate q_1 , the corresponding quantities both can be assigned values, namely a_1 and b_1 .

An important quantity for every physical system is energy. The operator corresponding to energy is the system Hamiltonian H . Its eigenvectors e_i are referred to as energy eigenstates and its eigenvalues E_i as the energy levels of the system ($He_i = E_ie_i$). If we expand a state vector ψ in the basis of the e_i 's,

$$\psi = \sum_{i=1}^N \psi_i e_i,$$

the quantities $|\psi_i|^2 = |(e_i, \psi)|^2$ are called the **populations** of the energy levels i .

We can write down the solution of (2.1) explicitly as

$$\psi(t) = e^{-iHt/\hbar} \psi(0) = \sum_{i=1}^N \psi_i(0) e^{-iE_it/\hbar} e_i.$$

Note that the time evolution is a unitary transformation of the initial state that preserves the populations of the levels, because

$$|\psi_i(t)|^2 = |\psi_i(0) e^{-iE_it/\hbar}|^2 = |\psi_i(0)|^2$$

In particular, if $\psi(0) = e_i$, then $\psi(t) = e^{-iE_it/\hbar} e_i$ and so the energy eigenstates are equilibria of the dynamics.

Finally, we define the notion of **Bohr frequency** which will play an important role in the subsequent development: For every pair of energy levels i and j , we define the Bohr frequency for the transition between them as

$$\omega_{ij} = \frac{E_i - E_j}{\hbar}.$$

$|\omega_{ij}|$ is equal to the frequency of a photon absorbed or emitted by the quantum system when it makes a transition from j to i or vice versa, depending on which state is higher in energy. This is a well-known experimental fact which motivated the development of Quantum Mechanics. In order to put this statement in a sound theoretical framework the development of the quantum-mechanical theory of light and its interaction with matter, Quantum Electrodynamics (QED) was necessary. For our purposes, we will provide a proof of this statement within the semiclassical model of interaction we use, see equation (2.3) in the next section.

2.1.2 Controlled quantum systems

When a quantum system is acted upon by an external field (e.g. an atom in the electric field of a laser), (2.1) is modified as follows:

$$i\hbar \dot{\psi} = (H_0 + \sum_{\alpha} V_{\alpha} u_{\alpha}(t)) \psi. \quad (2.3)$$

The time-independent system Hamiltonian H_0 still describes the internal dynamics of the system. The V_{α} 's are referred to as control Hamiltonians and describe the coupling of the system to the external fields $u_{\alpha}(t)$. We should note that this model is valid in the semiclassical approximation, where the control field influencing the quantum system is taken to be a classical source. We also employ the dipole moment approximation, valid for long wavelength fields, where the spatial features of the system-control field interaction are lumped in the interaction Hamiltonian and the amplitude of the field becomes the control parameter.

Systems of the form (2.3) are called **bilinear** in the Systems Theory literature. If one writes the solution of the (linear) equation (2.3) in the form

$$\psi(t) = U(t)\psi(0), \quad (2.4)$$

the time-dependent unitary operator $U(t)$ (also known as the propagator of the quantum dynamics) satisfies the equation

$$i\hbar \frac{dU(t)}{dt} = (H_0 + \sum_{\alpha} V_{\alpha} u_{\alpha}(t)) U(t), \quad (2.5)$$

with the initial condition $U(0) = I$. Since states are defined modulo total phase factors, one can impose the condition $\det U(t) = 1$, i.e. $U(t)$ is a special unitary transformation. For finite-dimensional systems, (2.5) is known as a right-invariant (control) system on the Lie group of special unitary matrices in N dimensions, $SU(N)$. Right invariant systems on Lie groups have received a lot of attention in the Geometric Control literature, starting with the important works [45, 46, 47], see also [48] for a comprehensive treatment and extensive references. Many properties of system (2.3) can be inferred from (2.5). For example, the controllability properties of (2.3) have been studied extensively for finite dimensional quantum systems ([49, 50, 51, 52]) using results from right invariant systems on Lie groups. We quote the main result and refer to the bibliography for more details:

Necessary and sufficient condition for controllability: The system (2.3) is controllable by locally bounded, measurable controls iff the Lie Algebra generated by H_0 and the V_{α} 's contains $su(N)$, the Algebra of traceless anti-symmetric complex $N \times N$ matrices.

For general infinite-dimensional bilinear systems, [53] is a basic reference for controllability issues while [54] is the first work that undertook the controllability question for infinite-dimensional quantum systems. In general, results concerning controllability of infinite-dimensional quantum systems are far weaker than their finite-dimensional counterparts. Nevertheless, there exist systems for which states with interesting properties can be accessed [55, 56].

2.1.3 Two model systems

We continue this introductory section with the description of two quantum systems which arise often in applications of quantum control ideas. The first one is the one-half spin particle which is the quantum system with the state space of smallest dimension, namely \mathbb{C}^2). Spin one-half systems and two-state systems in general are referred to as **qubits** in the Quantum Information literature and form the building block of many proposed Quantum Computation and Communication technologies. This makes the study of these systems very important. The other system we will briefly describe is the so-called Morse oscillator. This is the quantum version of a mechanical oscillator with non-linear restoring force that models the force between two atoms in a molecule. In other words, it is a quantum mechanical model of a chemical bond.

The spin- $\frac{1}{2}$ particle

Spin is a purely quantum mechanical degree of freedom with no classical counterpart. It corresponds to an “internal” angular momentum of the particle. Like a rotating particle with charge has a magnetic dipole moment proportional to its angular momentum, so does a particle with spin. In some sense, the spin is an elementary inherent magnetization of particles. Spin interacts with magnetic fields, possibly created by other spins.

The electron, the proton, the neutron and many composite nuclei have spin $\frac{1}{2}$. This is the simplest of all quantum mechanical systems. The state space is \mathbb{C}^2 and the Hamiltonian is

$$H = -\gamma \sum_{i=x,y,z} S_i B_i$$

where B_x , B_y , B_z are the components of the magnetic field at the position of the particle and S_x , S_y , S_z are the **spin operators**. They are self-adjoint operators

in \mathbb{C}^2 , so they are 2×2 Hermitian matrices. They are simply expressed in terms of the **Pauli matrices**,

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

as $S_i = \frac{\hbar}{2}\sigma_i$. γ , is a property of the particle called the gyromagnetic ratio, and it may be positive or negative. Here, we take $\gamma > 0$. In most applications, one applies a large uniform, constant magnetic field B_0 in one direction (say z). The reason for this is to overpower any random magnetic fields in the locality of the particle. One then applies uniform time-dependent fields in the x and/or y directions that act as controls. The dynamical equation has the form

$$i\frac{d\psi}{dt} = -\frac{1}{2}\gamma(\sigma_z B_0 + \sigma_x B_x(t) + \sigma_y B_y(t))\psi. \quad (2.6)$$

This equation can be rendered dimensionless by defining the variables

$$\omega_0 = \gamma B_0, \quad \tau = \omega_0 t, \quad u_x = \frac{B_x}{B_0}, \quad u_y = \frac{B_y}{B_0}.$$

ω_0 is sometimes referred to as the Larmor frequency of the spin. The control system takes the simple form

$$i\frac{d\psi}{d\tau} = -\frac{1}{2}(\sigma_z + \sigma_x u_x + \sigma_y u_y)\psi, \quad \psi \in \mathbb{C}^2. \quad (2.7)$$

It is controllable with either control input (u_x or u_y).

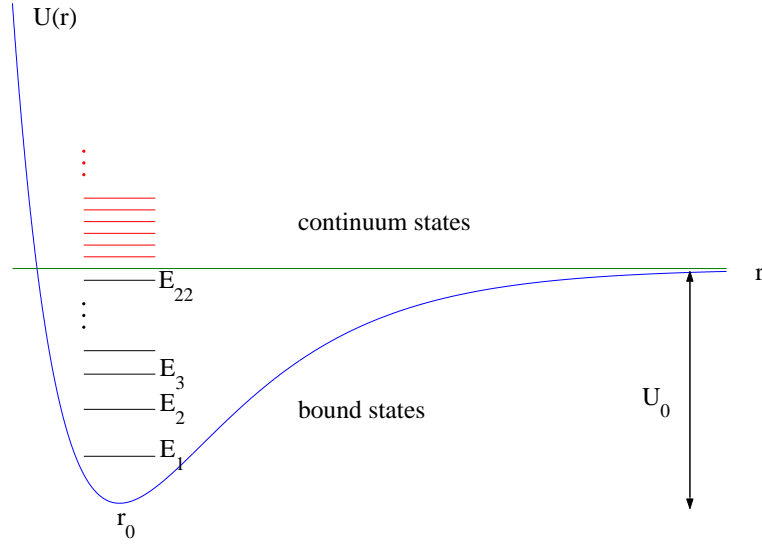


Figure 2.1: The Morse potential for the OH bond

The Morse oscillator

This infinite dimensional system is a model for the vibrational dynamics of two-atom molecules. Its state space is $L^2(\mathbb{R})$. The Hamiltonian operator is

$$H_0 = -\frac{\hbar^2}{2m} \frac{d^2}{dr^2} + U(r),$$

with the Morse potential

$$U(r) = U_0 \{ \exp[-\alpha(r - r_0)] - 1 \}^2 - U_0 ,$$

see figure 2.1, where, r is the intermolecular distance, m is the reduced mass of the molecule and r_0 the equilibrium distance of the molecule. α determines the width of the molecular well and U_0 its depth. The spectrum of the Hamiltonian operator contains a set of discrete points as well as a continuum. The discrete

energy eigenvalues are given by the expression

$$E_n = -\frac{\hbar^2 \alpha^2}{2m} \left(\frac{\sqrt{2mV_0}}{\hbar \alpha} - n - \frac{1}{2} \right)^2,$$

where $n = 0, \dots, \left[\frac{\sqrt{2mV_0}}{\hbar \alpha} - \frac{1}{2} \right]$. Every $E \geq 0$ belongs to the continuous part of the spectrum. The discrete energy levels correspond to bound states of the molecular oscillation while the continuous energies correspond to unbounded (relative) motion of the two atoms. Thus, a transition of the molecule from a bound state to a state in the continuum represents the breaking of the molecular bond. This phenomenon is called the **dissociation** of the molecule and is very important to Quantum Chemistry. Indeed, the role of the laser as a chemical catalyst is exactly that, to break existing bonds, so that new ones may be formed. The interaction of the molecule with the electric field u of a laser is modeled by the interaction Hamiltonian (in the electric dipole approximation)

$$V_{int}(t) = -\mu(r)u(t)$$

and $\mu(r)$ is the molecular dipole function. A common form for it is

$$\mu(r) = \mu_0 r \exp(-r/r^0),$$

(μ_0 and r^0 are parameters of the interaction). In the following, for the purpose of specific examples, we use the model of the OH bond which has 22 bound states.

2.1.4 Bohr frequencies as resonance frequencies

We conclude this section with a calculation that demonstrates the physical significance of the Bohr frequencies and helps us develop some initial intuition about

the control of quantum systems. We begin with the simplest of all quantum systems, the spin- $\frac{1}{2}$. Suppose that the system is initially in its lowest energy state, $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, and we want to transfer it to the higher energy state $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. For the purpose of this illustration, we consider (2.6) with sinusoidal magnetic field inputs:

$$\begin{aligned} B_x(t) &= +B_1 \cos(\omega t), \\ B_y(t) &= -B_1 \sin(\omega t). \end{aligned}$$

Letting $\omega_0 = \gamma B_0$ and $\varepsilon = \frac{B_1}{B_0}$, (2.6) becomes

$$i \frac{d}{dt} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \begin{pmatrix} -\frac{\omega_0}{2} & -\frac{\varepsilon \omega_0}{2} e^{i\omega t} \\ -\frac{\varepsilon \omega_0}{2} e^{-i\omega t} & \frac{\omega_0}{2} \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}.$$

To solve this equation, we perform the transformation

$$\begin{aligned} \psi_1 &= e^{+\frac{i\omega_0 t}{2}} x_1, \\ \psi_2 &= e^{-\frac{i\omega_0 t}{2}} x_2. \end{aligned} \tag{2.8}$$

The resulting system

$$\begin{aligned} i\dot{x}_1 &= -\frac{\varepsilon \omega_0}{2} e^{+i(\omega - \omega_0)t} x_2, \\ i\dot{x}_2 &= -\frac{\varepsilon \omega_0}{2} e^{-i(\omega - \omega_0)t} x_1, \end{aligned}$$

can be solved explicitly because x_1 and x_2 satisfy linear second-order equations with constant coefficients, e.g.

$$\ddot{x}_1 - i(\omega - \omega_0)\dot{x}_1 + \left(\frac{\varepsilon \omega_0}{2}\right)^2 x_1 = 0.$$

With the initial conditions

$$x_1(0) = \psi_1(0) = 1,$$

$$x_2(0) = \psi_2(0) = 0,$$

the solution is

$$\begin{aligned} x_1 &= \left(\cos\left(\frac{\Omega t}{2}\right) - i \frac{\omega - \omega_0}{\Omega} \sin\left(\frac{\Omega t}{2}\right) \right) e^{\frac{i(\omega - \omega_0)t}{2}}, \\ x_2 &= i \frac{\varepsilon \omega_0}{\Omega} \sin\left(\frac{\Omega t}{2}\right) e^{-\frac{i(\omega - \omega_0)t}{2}}, \end{aligned}$$

where $\Omega \doteq \sqrt{(\omega - \omega_0)^2 + \varepsilon^2 \omega_0^2}$. The populations of the two states are given as functions of time by the expressions

$$\begin{aligned} |\psi_1|^2 &= |x_1|^2 = 1 - \frac{\varepsilon^2 \omega_0^2}{\Omega^2} \sin^2\left(\frac{\Omega t}{2}\right), \\ |\psi_2|^2 &= |x_2|^2 = \frac{\varepsilon^2 \omega_0^2}{\Omega^2} \sin^2\left(\frac{\Omega t}{2}\right). \end{aligned}$$

From these expressions we conclude that, in order to achieve the maximum possible transfer of population from state 1 to state 2, we need to use a field of frequency $\omega = \omega_0$. In other words, the external field must be in resonance with the Bohr frequency of the transition we're trying to achieve. In this case, the expressions for the state and the populations are, respectively,

$$x = \begin{pmatrix} \cos\left(\frac{\varepsilon \omega_0 t}{2}\right) \\ i \sin\left(\frac{\varepsilon \omega_0 t}{2}\right) \end{pmatrix},$$

and

$$\begin{aligned} |\psi_1|^2 &= |x_1|^2 = \cos^2\left(\frac{\varepsilon \omega_0 t}{2}\right), \\ |\psi_2|^2 &= |x_2|^2 = \sin^2\left(\frac{\varepsilon \omega_0 t}{2}\right). \end{aligned}$$

The transition can be achieved with arbitrarily small field amplitude and the

transfer time is inversely proportional to this amplitude:

$$T_{\text{transfer}} = \frac{\pi}{\varepsilon\omega_0} = \frac{\pi}{\gamma B_1}.$$

Put differently, given a desired transfer time T_{transfer} , the field amplitude has to be equal to

$$B_1 = \frac{\pi}{\gamma T_{\text{transfer}}}. \quad (2.9)$$

This is an exact result for the two-state system. In the following, we show for a general finite-dimensional quantum system (using averaging) that for small enough amplitude, the frequency of the controlling field must be equal to the Bohr frequency of the transition we 're trying to achieve. From now on, we use units where $\hbar = 1$ and consider (2.3) with only one control:

$$i\dot{\psi} = (H_0 + Vu(t))\psi. \quad (2.10)$$

First, we make a change of variables that generalizes (2.8):

$$\psi = e^{-iH_0t} x$$

This is the transformation to the “**rotating frame**” or “**interaction picture**”.

(2.10) transforms to

$$i\dot{x} = e^{iH_0t} V e^{-iH_0t} u(t)x = F(t)u(t)x,$$

where

$$F(t) \doteq e^{iH_0t} V e^{-iH_0t}.$$

This equation is most easily analyzed in the basis of energy eigenstates. In this

basis, H_0 is diagonal and

$$F_{kl}(t) = V_{kl}e^{i(E_k - E_l)t} = V_{kl}e^{i\omega_{kl}t}$$

Note the appearance of the Bohr frequencies in the matrix elements of $F(t)$. Suppose that we want to induce a transition from level i to level j . For the control field $u(t)$, we take a sinusoid with frequency equal to $\omega > 0$:

$$u(t) = Ae^{i\omega t} + A^*e^{-i\omega t}.$$

Inserting this $u(t)$ in the rotating frame equation and writing the equation in component form, we have:

$$i\dot{x}_k = \sum_l V_{kl}e^{i\omega_{kl}t}[Ae^{i\omega t} + A^*e^{-i\omega t}]x_l. \quad (2.11)$$

The right-hand side of (2.11) is a sum of terms with sinusoidal time dependence. An excellent approximation technique for systems of this sort is averaging. In averaging, one considers equations of the form

$$\dot{w} = \varepsilon f(w, t, \varepsilon), \quad (2.12)$$

where f must be a bounded C^2 function of its arguments with bounded derivatives up to 2nd order such that the limit

$$f_{av}(w) \doteq \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_t^{t+\tau} f(w, t', 0) dt'$$

exists. A standard result (see Chapter 8 of [57] for example) guarantees that, for

sufficiently small ε , the solution of

$$\dot{\bar{w}} = \varepsilon f_{av}(\bar{w})$$

with an initial condition $O(\varepsilon)$ close to the initial condition of (2.12) (i.e. $\bar{w}(0) - w(0) = O(\varepsilon)$) is $O(\varepsilon)$ close to that of (2.12) for a time interval of length $O(\frac{1}{\varepsilon})$. Assuming that $\omega \neq \omega_{kl}$ for all positive Bohr frequencies of the system, the resulting averaged dynamics is just

$$i\dot{\bar{x}}_k = 0, \quad \forall k,$$

i.e. nothing happens. If, however, $\omega = \omega_{ij}$, and there is no other Bohr frequency equal to ω_{ij} , the averaged form of (2.11) is given by

$$\begin{aligned} i\dot{\bar{x}}_k &= 0, \quad k \neq i, j, \\ i\dot{\bar{x}}_i &= A^* V_{ij} \bar{x}_j, \\ i\dot{\bar{x}}_j &= A V_{ji} \bar{x}_i. \end{aligned} \tag{2.13}$$

Using the initial conditions $\bar{x}_k(0) = x_k(0) = \delta_{ik}$, these can be solved immediately to give the average populations

$$\begin{aligned} |\bar{\psi}_k|^2 &= 0, \quad k \neq i, j, \\ |\bar{\psi}_i|^2 &= \cos^2 |AV_{ij}|t, \\ |\bar{\psi}_j|^2 &= \sin^2 |AV_{ij}|t. \end{aligned}$$

For a given transfer time T , the control amplitude must be

$$|A| = \frac{\pi}{2|V_{ij}|T} . \quad (2.14)$$

We can always choose a large enough transfer time T to make $|A|$ as small as needed. Averaging guarantees that the solution of (2.13) is $O(|A|)$ close to the solution of (2.11) for a time of order $O(\frac{1}{|A|})$, which means that the approximation gets better as T becomes larger. This argument substantiates the interpretation of Bohr frequencies as resonance frequencies for any quantum system.

Before leaving this section we want to return to the question of controllability for quantum systems and present a sufficient condition for controllability that assumes more than the necessary and sufficient condition of section 2.1.2 but is extremely easy to check. The reason we present it here rather than in that section is that it becomes very transparent in the light of the physical interpretation of the Bohr frequencies as resonance frequencies. To state it, we need a few simple notions: A quantum system whose energy levels (eigenvalues of H_0) are all different from each other is called non-degenerate. Moreover, a system such that no two Bohr frequencies (differences of energy levels) are the same, is said to have no degenerate transitions. We also define the graph of an N -dimensional quantum system as a planar graph with N nodes, each representing an energy level, such that an edge connects the nodes i and j iff there exists at least one control Hamiltonian V_α in equation (2.3) with $(V_\alpha)_{ij} \neq 0$. The positions of the graph nodes are not important only the connectivity properties of the graph are (e.g. the graph is connected when there exists a sequence of edges connecting any two given nodes).

Sufficient condition for controllability: [51, 52] If the system (2.3) is not degenerate, has no degenerate transitions and its graph is connected, it is control-

lable.

As a matter of fact, this is the generic situation for controllability of (2.10).

2.2 Overview of problem and results

Steering a quantum system from its initial state to a given final state or a set of final states is one of the central problems in the control of such systems. While transfers to specific final states are very important for applications to Quantum Computing and Quantum Chemistry, transfers to given final populations are also important for many applications to Quantum Chemistry and Atomic Physics. Optimal Control is a natural approach to transfer problems: Frequently, one desires to optimize some aspect of the transfer. For example, minimize the transfer time [58, 59], maximize some measure of efficiency of the control in achieving its objective [10] or minimize some measure of the size of the control, for instance its L^2 norm [33, 35, 34, 60]. Moreover, the optimal control(s), singled out of all possible controls that achieve the objective, should have interesting properties tied to the structure of the given system.

We consider here *energy-optimal exact population transfers*. That is, we wish to find a $u \in L^2([0, T])$, that minimizes

$$\|u\|_{L^2([0, T])}^2 = \int_0^T u^2(t) dt, \quad (2.15)$$

and drives an initial state ψ_0 of system (2.10) to a target population distribution $\{|\psi_i(T)|^2 = p_i, i = 1, \dots, N\}$ (N is the dimension of the system and $\psi = \sum_i \psi_i e_i$, where the e_i 's are the orthonormal eigenvectors of H_0). This cost for the control has been used extensively in the literature of optimal control of quantum systems

as part of various objective functionals. It provides a measure of the energy spent to create the controlling field and leads, as we will see, to interesting conclusions.

Relatively little work has been done on the problem of optimal *exact* population transfers. Analytically, it is a hard problem and explicit solutions are known only in few cases: In [33, 34] this is done for transfers between eigenstates in a two-dimensional system with two controls and in [35, 34] it is done for transfers between eigenstates of a three-dimensional system with four controls. In fact, both cases mentioned above are instances of a certain algebraic structure being present (the so-called $K + P$ structure) [35, 61]. It allows one to find analytic expressions for the control and the state in terms of the unknown initial costate but the analytic determination of this unknown initial costate (and hence the complete solution of the problem) is possible only for systems of small dimensionality because one needs to analytically compute matrix exponentials and solve transcendental equations. In dimensions higher than 2, the presence of the $K + P$ structure requires unphysical numbers of controls and very special structure in the control Hamiltonians. For a two-state system with one control [33], it is possible to find an analytic expression for the control (in terms of unknown constants) but not for the state and so the problem has to be solved numerically from that point on. For higher dimensional systems, no general properties of the optimal control and state trajectory are known.

Numerically, the two-point boundary value problem that expresses the necessary conditions of optimality becomes increasingly harder to solve as the dimension of the system or the transfer time grows. Dimension growth dramatically increases the computational cost of the numerical solution. More relevant to our work is the issue of large transfer time. In many typical applications, the transfer time may be a few orders of magnitude larger than the time scale of the free evolution

of the system. This may be necessary for the transfer to be possible or for the amplitude of the control to be small enough so that it is within experimentally feasible limits. The problem here is the presence of two time scales in the solution, i.e. stiffness: There is the fast time scale of the free dynamics of the quantum system and the slow time scale of the transition (which is of the order of T). This creates the need for a very detailed numerical solution in order to guarantee good solution accuracy resulting in large computational times. In applications one often considers relaxations of this optimal transfer problem where one abandons the requirement for *exact* transfers and tries to minimize a combination of the integral cost (2.15) and a weighted sum of the final populations:

$$a \int_0^T u^2(t) dt + \sum_{i=1}^N m_i |\psi_i(T)|^2, \quad (2.16)$$

($a \geq 0$). An advantage of optimal transfer problems like these over the one we are considering in this work is that they lead to two-point boundary value problems with separated boundary conditions and are amenable to iterative solution methods, see for example [2, 31, 62]. Yet, the above remarks hold for these as well. We shall comment on the connections between this work and the later type of optimal transfer problems in section 2.8.

We found that the study of optimal population transfers for large transfer times offers some insight into the nature of optimal control and state trajectory as well as computational advantages in the numerical solution of the problem. The main conclusions of our work are the following:

1. For generic population transfers (this will be defined precisely in the statement of the theorems in section 2.5) and large enough transfer times, the optimal control has the following, physically plausible form: It is a sum

of sinusoids with frequencies equal to the Bohr frequencies of the quantum system multiplied by slowly varying profiles, that is functions of $\frac{t}{T}$:

$$\begin{aligned} u_{opt}(t) &= \frac{i}{T} \text{tr} \left(e^{iH_0 t} V e^{-iH_0 t} L \left(\frac{t}{T} \right) \right) + O\left(\frac{1}{T^2}\right) \\ &= \frac{i}{T} \sum_{k \neq l} V_{kl} e^{i\omega_{kl} t} L_{lk} \left(\frac{t}{T} \right) + O\left(\frac{1}{T^2}\right). \end{aligned}$$

L is an anti-Hermitian matrix with zeros on the diagonal, whose entries are the profiles. This form is explicitly verified in all analytically solvable cases (in the $K + P$ case, there are no $O(\frac{1}{T^2})$ corrections) and it is in fact observed in numerical solutions.

2. Again for generic population transfers and large enough transfer times, the optimal trajectories follow a slow “mean” evolution modulated by fast oscillations. The slow mean evolution and the slow control profiles can be calculated by solving an “averaged” two-point boundary value problem (the term will be explained in section 2.4) in the fixed interval $[0, 1]$, irrespective of how large the transfer time T is. The small deviations of the optimal trajectories from their mean evolution are due to the free dynamics. Quantitatively,

$$\psi(t) = e^{-iH_0 t} \bar{\psi} \left(\frac{t}{T} \right) + O\left(\frac{1}{T}\right),$$

where $\bar{\psi}$ denotes the “mean trajectory” and the oscillatory $e^{-iH_0 t}$ term is responsible for the deviations from the mean evolution. There are also additional corrections, of higher order in a $\frac{1}{T}$ expansion.

Although these results hold for large transfer times, one may use a solution of the optimal control problem obtained this way for a large transfer time T as the first step in a continuation method solution of the original optimization problem

where the continuation parameter is the transfer time. The point is that the large transfer time limit both reveals the structure of the controls and serves as a good starting point for the solution of the problem: Indeed, the associated “averaged” two-point boundary value problem in $[0, 1]$ is much easier to solve numerically than the original problem.

2.3 Optimal population transfers

In this section we derive the necessary conditions of optimality for the optimal population transfer problem described in section 2.2. For simplicity, we consider a system with one control (which is usually the case) and set, as usually, $\hbar = 1$ in Schrodinger’s equation:

$$i\dot{\psi} = (H_0 + V u(t))\psi.$$

Nevertheless, everything we do in this and the following sections can readily be generalized to systems with more controls. We also require the sufficient controllability condition discussed at the end of section 2.1.4:

Controllability assumption: The system (2.10) is non-degenerate, has no degenerate transitions and its graph is connected.

We now proceed with the set up of the optimal control problem. We wish to find a $u \in L^2([0, T])$, that minimizes

$$\|u\|_{L^2([0, T])}^2 = \int_0^T u^2(t) dt,$$

and drives an initial state ψ_0 of system (2.10) to a target population distribution $\{|\psi_i(T)|^2 = p_i, i = 1, \dots, N\}$ (N is the dimension of the system). We will refer to this as optimal transfer problem (I). The Maximum Principle of optimal control

[48, 63] provides necessary conditions for optimality in terms of the *Hamiltonian function* (not to be confused with the Hamiltonian operators of quantum systems!)

$$H(\psi, \lambda, u) = \frac{1}{2}u^2 - i\lambda^*(H_0 + V u)\psi + i\psi^*(H_0 + V u)\lambda,$$

where $\lambda \in \mathbb{C}^n$ is the co-state vector. It is shown in Appendix A that solutions of the optimal control problem satisfy, $\lambda^*(t) \psi(t) = 0$.

The optimal control u and the corresponding state and co-state satisfy the equations

$$\dot{\psi} = \frac{\partial H}{\partial \lambda^*}, \quad \dot{\lambda} = -\frac{\partial H}{\partial \psi^*} \quad \text{and} \quad \frac{\partial H}{\partial u} = 0,$$

which, given the expression for H , have the form

$$i\dot{\psi} = (H_0 + V u)\psi, \tag{2.17}$$

$$i\dot{\lambda} = (H_0 + V u)\lambda, \tag{2.18}$$

$$u = i(\lambda^* V \psi - \psi^* V \lambda). \tag{2.19}$$

For convenience, we rewrite equations (2.17) and (2.18) using u from (2.19):

$$i\dot{\psi} = H_0\psi + i(\lambda^* V \psi - \psi^* V \lambda) V \psi, \tag{2.20}$$

$$i\dot{\lambda} = H_0\lambda + i(\lambda^* V \psi - \psi^* V \lambda) V \lambda. \tag{2.21}$$

To these equations one must append the boundary conditions

$$\psi(0) = \psi_0, \quad |\psi_i(T)|^2 = p_i, \quad \text{Im}(\psi_i^*(T) \lambda_i(T)) = 0, \tag{2.22}$$

$\forall i = 1, \dots, N$. The last of these equations are the transversality conditions at the endpoint. Their proof is also given in Appendix A. We will refer to the two-

point boundary value problem comprised of equations (2.20) - (2.22) as **two-point boundary value problem (I)**.

We briefly discuss now methods for its solution. Analytically, we may proceed as follows: Define the traceless, anti-Hermitian matrix Λ by $\Lambda \doteq \psi \lambda^* - \lambda \psi^*$ and using (2.17) and (2.18) we can show that

$$\dot{\Lambda} = -i [H_0 + V u, \Lambda] = [-iH_0 + \text{tr}(V\Lambda) V, \Lambda] \quad (2.23)$$

where we substituted

$$u = i (\lambda^* V \psi - \psi^* V \lambda) = i \text{tr}(V\Lambda),$$

from (2.19). If we solve (2.23) analytically, we obtain an expression for the control function $u(t)$ in terms of unknown constants (these unknown constants are matrix elements of $\Lambda(0) = \psi(0) \lambda^*(0) - \lambda(0) \psi^*(0)$ and $\lambda(0)$ is what we are after). Then, one may attempt to solve (2.17) and (2.18) using this expression for u and determine these constants by satisfying the boundary conditions (2.22). In practice this program seems impossible to carry out in its entirety except for the few special cases mentioned [33, 34, 35]. [Note: These works consider optimal transfer problems for the system $i\hbar \dot{U} = (H_0 + \sum_{\alpha} V_{\alpha} u_{\alpha}(t)) U$ which, as discussed in subsection 2.1.2, comes from $i\hbar \dot{\psi} = (H_0 + \sum_{\alpha} V_{\alpha} u_{\alpha}(t)) \psi$ by letting $\psi(t) = U(t) \psi(0)$. It turns out that the formulation of the optimal control problem in the two setups is similar and in fact, equation (2.23) in that context is the co-state equation “pulled-back” on the cotangent space at the identity of $SU(N)$, see [47, 48].] Numerically, one may attempt to solve the two-point boundary value problem (2.20) - (2.22) using, for example, some shooting or finite-difference method. This works for systems of small dimensionality and small transfer times but becomes

increasingly harder as the system dimension grows and as larger transfer times are required for the transfer to be possible. As mentioned before, in many typical applications the transfer time has to be a few orders of magnitude larger than the time scale of the free dynamics of the system in order to achieve the desired transfer. Because the free dynamics of the system is oscillatory, the first terms on the right sides of equations (2.20) and (2.21) create small oscillations of the populations around their “mean” evolution towards their final values. This creates the need for a very detailed numerical solution in order to guarantee good solution accuracy. Examples of this can be seen in the graphs of section 2.6. Besides the usual unfavorable scaling of the solution efficiency with dimension (an issue which we do not address), this is the main source of difficulty of the problem. Note that these remarks are quite general and independent of the specific numerical methods used to solve the two-point boundary value problem. We will see in the following how this problem can be overcome.

Before we leave this section we would like to point out that our discussion so far, as well as in the following, will only concern *regular* or *normal* extrema of the optimization problem (I). *Singular* extrema [64, 48] will not be considered. The reason is that the form of the singular extrema does not depend on the exact cost used in the minimization problem, and thus, the same singular extremum can be a local or even a global minimizer to many different cost functionals. In this sense, singular extrema are not particular to one optimal control problem and don’t reflect its particular structure. So, in this work, all the discussion and results concern normal minima of problem (I) (as well as problem (II) to be defined in the next section).

2.4 Optimal population transfers for an averaged system

In this section, we introduce a special form for the control in equation (2.10), a sum of sinusoids with frequencies equal to the Bohr frequencies of the quantum system multiplied by slowly varying profiles, that is functions of $\frac{t}{T}$. We then proceed to “average out” the dynamics in the time scale of the free evolution of the system (this time scale is set by the Bohr frequencies), which is fast compared to the transfer time T . This leaves us with an “averaged” control system whose evolution approximates that of the original under the special form of the control introduced. The motivation for this lies in the following: We set up a corresponding optimal transfer problem for the averaged system, whose cost approximates the cost (2.15). We will show in section 2.7 how solutions to this optimal transfer problem approximate solutions to our original optimal transfer problem (I), to first order in an $O(\frac{1}{T})$ expansion, proving the results described in the introduction. We begin with a transformation to the rotating frame

$$x = e^{iH_0 t} \psi \tag{2.24}$$

in (2.10). The time evolution of the new variable x is due entirely to the control, because the free evolution has been accounted for. In terms of the new variable, (2.10) becomes

$$i\dot{x} = u F(t)x, \tag{2.25}$$

where,

$$F(t) \doteq e^{iH_0 t} V e^{-iH_0 t}. \tag{2.26}$$

Note the appearance of the Bohr frequencies in the matrix elements of F ,

$$F_{ij}(t) = V_{ij}e^{i(E_i-E_j)t} = V_{ij}e^{i\omega_{ij}t}.$$

We adopt the following form for the control $u(t)$:

$$u(t) = \varepsilon \left(u_0(\varepsilon t) + \sum_{i \neq j}^N e^{i\omega_{ij}t} u_{ji}(\varepsilon t) \right), \quad (2.27)$$

where $u_{ji}^* = u_{ij}$ and u_0 is real, so that u is real. u_{ji} is a complex “envelope” that multiplies a sinusoid with frequency equal to the Bohr frequency for the transition i to j . The value of ε will be given shortly. We introduce $u(t)$ from (2.27) in (2.25) and rewrite (2.25) in component form:

$$i\dot{x}_i = \varepsilon \left(u_0(\varepsilon t) + \sum_{k \neq l} e^{i\omega_{kl}t} u_{lk}(\varepsilon t) \right) \sum_j e^{i\omega_{ij}t} V_{ij} x_j. \quad (2.28)$$

We now approximate (2.28) for small ε using averaging. As mentioned before, averaging approximates the solution of an equation of the form

$$\dot{w} = \varepsilon f(w, t, \varepsilon),$$

with that of the equation

$$\dot{\bar{w}} = \varepsilon f_{av}(\bar{w}),$$

where

$$f_{av}(w) \doteq \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_t^{t+\tau} f(w, t', 0) dt'.$$

A standard averaging theorem [57] guarantees that, for sufficiently small ε , the solutions of the two equations remain $O(\varepsilon)$ -close for a time interval of length $O(\frac{1}{\varepsilon})$ as long as their initial conditions are $O(\varepsilon)$ -close. Equation (2.28) involves two

time scales, 1 and $\frac{1}{\varepsilon}$ but we want to average only over the time scale 1 dynamics. This is achieved as follows: Consider, instead of (2.12) the following equation:

$$\dot{w} = \varepsilon f(w, t, \varepsilon t, \varepsilon). \quad (2.29)$$

Define $w_0 = \varepsilon t$ and substitute w_0 for εt in (2.29). Then, consider the system

$$\begin{aligned} \dot{w} &= \varepsilon f(w, t, w_0, \varepsilon), \\ \dot{w}_0 &= \varepsilon \end{aligned}$$

and apply averaging to it. The resulting averaged form of (2.29) is now

$$\dot{\bar{w}} = \varepsilon f_{av}(\bar{w}, \varepsilon t),$$

where

$$f_{av}(w, w_0) \doteq \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_t^{t+\tau} f(w, t', w_0, 0) dt'.$$

We now apply this to equation (2.28). Since the time average of $e^{i\omega t}$ is zero for $\omega \neq 0$ and 1 for $\omega = 0$, only terms with no time dependence will contribute to the averaged equation. Letting \bar{x} be the averaged x , the averaged form of (2.28) is

$$i\dot{\bar{x}}_i = \varepsilon (V_{ii} u_0(\varepsilon t) \bar{x}_i + \sum_{j \neq i} V_{ij} u_{ij}(\varepsilon t) \bar{x}_j). \quad (2.30)$$

Taking $\varepsilon = \frac{1}{T}$ and rescaling time to $s = \varepsilon t = \frac{t}{T}$, (2.30) becomes

$$i \frac{d\bar{x}_i}{ds} = V_{ii} u_0(s) \bar{x}_i + \sum_{j \neq i} V_{ij} u_{ij}(s) \bar{x}_j, \quad (2.31)$$

or, in vector form,

$$i \frac{d\bar{x}}{ds} = \tilde{V}[u_0, u_{ij}] \bar{x},$$

where

$$\tilde{V}[u_0, u_{ij}] = \begin{pmatrix} V_{11}u_0 & V_{12}u_{12} & \cdots \\ V_{21}u_{21} & V_{22}u_0 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} = \begin{pmatrix} V_{11}u_0 & V_{12}u_{12} & \cdots \\ V_{12}^*u_{12}^* & V_{22}u_0 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}.$$

By construction, every solution to (2.31) with controls $u_0(s)$ and $u_{ij}(s)$ and initial state $\bar{x}(0) = \psi(0)$, provides a solution to (2.10) with $u(t)$ given by (2.27) and initial condition $\psi(0)$, correct up to $O(\varepsilon)$ terms for a time interval of size $\frac{1}{\varepsilon} = T$. It is shown in Appendix B that system (2.31) is controllable on account of the controllability assumption on the original system.

Since our goal is to relate optimal transfers of the original system to optimal transfers of the averaged one, we must find an objective for the averaged system that approximates $\|u\|_{L^2([0,T])}^2$. So, we compute $\|u\|_{L^2([0,T])}^2$ for $u(t)$ given by (2.27), with $\varepsilon = \frac{1}{T}$:

$$\begin{aligned} \int_0^T u^2(t) dt &= \frac{1}{T^2} \int_0^T \left(u_0\left(\frac{t}{T}\right) + \sum_{i \neq j} e^{i\omega_{ij}t} u_{ji}\left(\frac{t}{T}\right) \right)^2 dt \\ &= \frac{1}{T^2} \int_0^T \left\{ u_0^2\left(\frac{t}{T}\right) + 2 \sum_{i \neq j} e^{i\omega_{ij}t} u_{ji}\left(\frac{t}{T}\right) u_0\left(\frac{t}{T}\right) \right. \\ &\quad \left. + \sum_{i \neq j} \sum_{k \neq l} e^{i(\omega_{ij} + \omega_{kl})t} u_{ji}\left(\frac{t}{T}\right) u_{lk}\left(\frac{t}{T}\right) \right\} dt \\ &= \frac{1}{T} \int_0^1 \left\{ u_0^2(s) + 2 \sum_{i \neq j} e^{i\omega_{ij}Ts} u_{ji}(s) u_0(s) \right. \\ &\quad \left. + \sum_{i \neq j} \sum_{k \neq l} e^{i(\omega_{ij} + \omega_{kl})Ts} u_{ji}(s) u_{lk}(s) \right\} ds \end{aligned}$$

$$= \frac{1}{T} \int_0^1 \left[\sum_{i \neq j} u_{ij}(s) u_{ji}(s) + u_0^2(s) \right] ds + \frac{1}{T^2} B(T),$$

where $B(T)$ represents terms bounded in T . This last line is the result of separating the integrals into two kinds, these without exponentials, which are explicitly retained, and those with, which can easily be seen to scale like $\frac{1}{T}B(T)$ after a partial integration: Indeed, for any differentiable f ,

$$\begin{aligned} \int_0^1 e^{i\omega T s} f(s) ds &= \frac{1}{i\omega T} \int_0^1 (e^{i\omega T s})' f(s) ds \\ &= \frac{1}{i\omega T} \left\{ e^{i\omega T} f(1) - f(0) - \int_0^1 e^{i\omega T s} f'(s) ds \right\}. \end{aligned}$$

So, we pose the following optimal control problem for the averaged system: Find controls $u_0(s)$, $u_{ij}(s)$, $s \in [0, 1]$, that minimize

$$\int_0^1 \left[\sum_{i \neq j} u_{ij}(s) u_{ji}(s) + u_0^2(s) \right] ds = \int_0^1 \left[\sum_{i \neq j} |u_{ij}(s)|^2 + u_0^2(s) \right] ds \quad (2.32)$$

and drive an initial state $\bar{x}(0) = \psi_0$ of system (2.31) to a target population distribution $\{|\bar{x}_i(1)|^2 = p_i, i = 1, \dots, N\}$. We will refer to this as optimal transfer problem (II).

The necessary conditions for optimality are derived from the Hamiltonian function

$$\begin{aligned} H(\bar{x}_i, \bar{z}_i, u_{ij}) &= \frac{1}{2} u_0^2 + \frac{1}{2} \sum_{i \neq j} |u_{ij}|^2 - i \bar{z}^* \tilde{V}[u_0, u_{ij}] \bar{x} + i \bar{x}^* \tilde{V}[u_0, u_{ij}] \bar{z} \\ &= \frac{1}{2} u_0^2 + \sum_{i \neq j} \frac{1}{2} u_{ij} u_{ji} - i V_{ji} u_{ji} (\bar{x}_i \bar{z}_j^* - \bar{z}_i \bar{x}_j^*) - i u_0 \sum_i V_{ii} (\bar{x}_i \bar{z}_i^* - \bar{z}_i \bar{x}_i^*) \end{aligned}$$

and have the form

$$i \frac{d\bar{x}_i}{ds} = V_{ii} u_0 \bar{x}_i + \sum_{j \neq i} V_{ij} u_{ij} \bar{x}_j \quad (i \frac{d\bar{x}}{ds} = \tilde{V}[u_0, u_{ij}] \bar{x}), \quad (2.33)$$

$$i \frac{d\bar{z}_i}{ds} = V_{ii} u_0 \bar{z}_i + \sum_{j \neq i} V_{ij} u_{ij} \bar{z}_j \quad (i \frac{d\bar{z}}{ds} = \tilde{V}[u_0, u_{ij}] \bar{z}), \quad (2.34)$$

$$u_{ij} = i V_{ji} (x_i z_j^* - z_i x_j^*), \quad u_0 = i \sum_i V_{ii} (\bar{x}_i \bar{z}_i^* - \bar{z}_i \bar{x}_i^*). \quad (2.35)$$

We rewrite equations (2.33) and (2.34) using u_0 and u_{ij} from (2.35):

$$\frac{d\bar{x}_i}{ds} = \sum_{j \neq i} |V_{ij}|^2 (\bar{x}_i \bar{z}_j^* - \bar{z}_i \bar{x}_j^*) \bar{x}_j + \sum_k V_{kk} (\bar{x}_k \bar{z}_k^* - \bar{z}_k \bar{x}_k^*) V_{ii} \bar{x}_i, \quad (2.36)$$

$$\frac{d\bar{z}_i}{ds} = \sum_{j \neq i} |V_{ij}|^2 (\bar{x}_i \bar{z}_j^* - \bar{z}_i \bar{x}_j^*) \bar{z}_j + \sum_k V_{kk} (\bar{x}_k \bar{z}_k^* - \bar{z}_k \bar{x}_k^*) V_{ii} \bar{z}_i. \quad (2.37)$$

The corresponding boundary conditions are given by

$$\bar{x}(0) = \psi_0, \quad |\bar{x}_i(1)|^2 = p_i, \quad \text{Im}(\bar{x}_i^*(1) \bar{z}_i(1)) = 0. \quad (2.38)$$

We will refer to the two-point boundary value problem comprised of equations (2.36) - (2.38) as **two-point boundary value problem (II)**. From (2.35), we see that all u_{ij} with $V_{ij} = 0$ are identically zero, as they should. We can also show that $u_0 = 0$ and simplify the right sides of equations (2.36) and (2.37): In a fashion similar to the construction of Λ in section 2.3, we define L by $L \doteq \bar{x} \bar{z}^* - \bar{z} \bar{x}^*$. We also define another anti-Hermitian matrix $K = K(\bar{x}, \bar{z}) = K(L)$ by,

$$K_{ij} \doteq |V_{ij}|^2 L_{ij} = |V_{ij}|^2 (\bar{x}_i \bar{z}_j^* - \bar{z}_i \bar{x}_j^*), \quad i \neq j, \quad (2.39)$$

$$K_{ii} \doteq V_{ii} \sum_k V_{kk} L_{kk} = V_{ii} \sum_k V_{kk} (\bar{x}_k \bar{z}_k^* - \bar{z}_k \bar{x}_k^*). \quad (2.40)$$

K is the analog of $-iH_0 + \text{tr}(V\Lambda)V$ in section (2.3) and is a linear function of L .

With this definition, (2.36) and (2.37) read simply as

$$\frac{d\bar{x}}{ds} = K(L)\bar{x}, \quad (2.41)$$

$$\frac{d\bar{z}}{ds} = K(L)\bar{z}. \quad (2.42)$$

It is easy to see that L satisfies the differential equation

$$\frac{dL}{ds} = [K(L), L]. \quad (2.43)$$

The ii -th component of this equation reads

$$\begin{aligned} \frac{dL_{ii}}{ds} &= \sum_j K_{ij}L_{ji} - L_{ij}K_{ji} \\ &= K_{ii}L_{ii} + \sum_{j \neq i} K_{ij}L_{ji} - L_{ii}K_{ii} - \sum_{j \neq i} L_{ij}K_{ji} \\ &= \sum_{j \neq i} |V_{ij}|^2 (L_{ij}L_{ji} - L_{ij}L_{ji}) = 0, \end{aligned}$$

and so,

$$L_{ii}(s) = L_{ii}(1) = -2i \operatorname{Im}(\bar{x}_i^*(1) \bar{z}_i(1)) = 0. \quad (2.44)$$

This shows that $K_{ii}(s) = L_{ii}(s) = 0$ and $u_0(s) = 0$ and so, equations (2.36) and (2.37) simplify to

$$\frac{d\bar{x}_i}{ds} = \sum_{j \neq i} |V_{ij}|^2 (\bar{x}_i \bar{z}_j^* - \bar{z}_i \bar{x}_j^*) \bar{x}_j, \quad (2.45)$$

$$\frac{d\bar{z}_i}{ds} = \sum_{j \neq i} |V_{ij}|^2 (\bar{x}_i \bar{z}_j^* - \bar{z}_i \bar{x}_j^*) \bar{z}_j. \quad (2.46)$$

We will refer to this set of equations (along with (2.38)) as **two-point boundary value problem (II)** as well.

2.5 Main results

In this section, we make the connection between solutions of the two-point boundary value problems (II) and (I) in the large T limit.

Theorem 1: Let $(\bar{x}(s), \bar{z}(s))$ be a solution of two-point boundary value problem (II) over $[0, 1]$. Define $\psi(t)$, $\lambda(t)$ and $u(t)$, $t \in [0, T]$, by

$$\psi(t) = e^{-iH_0 t} \bar{x}\left(\frac{t}{T}\right), \quad (2.47)$$

$$\lambda(t) = \frac{1}{T} e^{-iH_0 t} \bar{z}\left(\frac{t}{T}\right), \quad (2.48)$$

$$u(t) = \frac{i}{T} \text{tr}\left(e^{iH_0 t} V e^{-iH_0 t} L\left(\frac{t}{T}\right)\right) = \frac{i}{T} \sum_{kl} V_{kl} e^{i\omega_{kl} t} L_{lk}\left(\frac{t}{T}\right). \quad (2.49)$$

Then, for large enough T , $\psi(t)$ satisfies the necessary conditions of optimality (2.17) - (2.22), up to terms of order $O(\frac{1}{T})$ and $\lambda(t)$ and $u(t)$ up to terms of order $O(\frac{1}{T^2})$. \square

Thus, solutions of the two-point boundary value problem (II) provide *approximate* solutions to the two-point boundary value problem (I) for large transfer times T . A natural question to ask then is, whether these approximate solutions to two-point boundary value problem (I) are in fact *approximations to solutions* of (I), in the large T limit. This is answered positively by the following theorem:

Theorem 2: Let ψ_0 and $\{p_i\}_{i=1,\dots,N}$ be an initial state and a target population of system (2.10), respectively, and let $(\bar{x}(s), \bar{z}(s))$ be a solution of two-point boundary value problem (II) over $[0, 1]$. Then, for almost all pairs $(\psi_0, \{p_i\})$, a solution of two-point boundary value problem (I) exists, for large enough T , of the form

$$\psi(t) = e^{-iH_0 t} \bar{x}\left(\frac{t}{T}\right) + O\left(\frac{1}{T}\right), \quad (2.50)$$

$$\lambda(t) = \frac{1}{T} e^{-iH_0 t} \bar{z}\left(\frac{t}{T}\right) + O\left(\frac{1}{T^2}\right), \quad (2.51)$$

The corresponding control has the form

$$\begin{aligned} u(t) &= \frac{i}{T} \text{tr} \left(e^{iH_0 t} V e^{-iH_0 t} L \left(\frac{t}{T} \right) \right) + O\left(\frac{1}{T^2}\right) \\ &= \frac{i}{T} \sum_{kl} V_{kl} e^{i\omega_{kl} t} L_{lk} \left(\frac{t}{T} \right) + O\left(\frac{1}{T^2}\right). \end{aligned} \quad (2.52)$$

The set of pairs $(\psi_0, \{p_i\})$ of initial states and target populations for which a solution of two-point boundary value problem (II) provides a solution of (I) according to (2.50) and (2.51) is open and full measure in the corresponding product manifold. \square

According to this theorem, solutions to the two-point boundary value problem (II) approximate solutions to (I) for large transfer times in the sense of equations (2.50) - (2.52), for almost every population transfer. In other words, all the local minima of optimal transfer problem (II) approximate local minima for the optimal transfer problem (I) according to (2.50) - (2.52). The question arises naturally: Are *all* local minima of optimal transfer problem (I), for large T , approximated in the sense of equations (2.50) - (2.52) by local minima of (II)? The answer to this question is essentially yes (see comment after theorem 3). We state the following theorem:

Theorem 3: Let ψ_0 and $\{p_i\}_{i=1,\dots,N}$ be an initial state and a target population of system (2.10), respectively. Then, for almost all pairs $(\psi_0, \{p_i\})$ and for large enough T , the globally optimal solution to the optimal transfer problem (I) is approximated by the globally optimal solution of (II) according to theorem 2. As before, the set of pairs $(\psi_0, \{p_i\})$ of initial states and target populations for which this happens is open and full measure in the corresponding product manifold. \square

In fact, we prove that, for large enough T , a number of the lowest energy optima of (I) that depends (in an unknown way) on T , are approximated by the

corresponding lowest energy optima of (II) according to theorem 2. We think, however, that theorem 3 is enough to demonstrate the spirit of our approach.

Theorem 3 precisely states the main results of our work that were delineated in the introduction: We obtain useful, physically plausible properties of the optimal control and state trajectory and at the same time, we reduce the solution of the original optimal control problem to a much easier problem: Indeed, the evolution equations (2.45) and (2.46) of two-point boundary value problem (II) do not contain the free dynamics of the system and thus their solution (which describes the “mean” evolution of the state) is much easier, see section 2.6.

We end this section with two remarks:

1. Theorems 2 and 3 are proven for an open, full measure set of pairs $(\psi_0, \{p_i\})$ of initial states and target populations. Unfortunately, the very important case of the initial state ψ_0 being an eigenstate (or, in general, having some populations equal to 0) is excluded. The reason is that in this case, two-point boundary value problem (II) has non-isolated solutions. Although this property is necessary for our proof of these results, we believe that they can be extended to (at least some) transfers with non-isolated solutions. Nevertheless, theorem (I) which contains all the essential *applicable* aspects of this work still holds.
2. An interesting implication of the theorem above is that, for large T , the locally optimal values of the objective (“energy”) scale like $\frac{1}{T}$. This demonstrates that the quadratic objective we use, cannot correspond to a physical notion of energy.

2.6 Examples

We consider three examples, each involving a different quantum system: A general two-state system with one control, a general three-state system with one control and the Morse oscillator model for the vibrational dynamics of the ground electronic state of the OH bond. We restrict our attention to the 22-dimensional space of bound states for that model and seek to control populations again with one control field. In all of the examples, we consider transfers from one eigenstate of the system, to another. Although in the proof of theorems 2 and 3 we had to exclude such transfers (because in that case two-point boundary value problem (II) has continua of solutions parameterized by angles, see (2.68) and (2.69), and we need isolated solutions to prove the theorems), we believe these theorems to hold for such transfers as well. Perhaps this can be established using different techniques from ours. In any case, theorem 1 still holds and so, every solution of two-point boundary value problem (II) will furnish an approximate solution to (I).

A two-state system

Consider the two-level system $i\dot{\psi} = (H_0 + Vu)\psi$, $\psi \in \mathbb{C}^2$, with

$$H_0 = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix}, \quad \text{and} \quad V = \begin{pmatrix} V_{11} & V_{12} \\ V_{12}^* & V_{22} \end{pmatrix}.$$

By rescaling u , we make $|V_{12}| = 1$. We are interested in the “population inversion” transfer

$$\psi_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \longrightarrow \quad \psi_d = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

In this example, the averaged optimal transfer problem (II) can be solved analytically. To begin, we introduce the anti-Hermitian matrices L and $K(L)$ of section (2.4):

$$L = \begin{pmatrix} 0 & L_{12} \\ -L_{12}^* & 0 \end{pmatrix} \text{ and } K(L) = L.$$

Then, (2.43) implies that $L_{12}(s)$, $s \in [0, 1]$, is constant. Equation (2.41) reads

$$\frac{d\bar{x}}{ds} = \begin{pmatrix} 0 & L_{12} \\ -L_{12}^* & 0 \end{pmatrix} \bar{x}.$$

Its solution with initial condition $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, is given by

$$\bar{x}(s) = \begin{pmatrix} \cos(|L_{12}|s) \\ -i \frac{L_{12}}{|L_{12}|} \sin(|L_{12}|s) \end{pmatrix}.$$

To achieve $\bar{x}(1) = \begin{pmatrix} 0 \\ e^{i*} \end{pmatrix}$, we must have $|L_{12}| = (n + \frac{1}{2})\pi$, with $n \in \mathbb{N}$. The value of the cost (2.32) is $\frac{\pi^2}{2}(2n+1)^2$. Thus, $n = 0$ corresponds to the global minimum of optimal transfer problem (II). The approximate optimal control for problem (I) has the form (2.52)

$$u(t) = -\frac{\pi}{T} \sin(\omega_{21}t + \varphi),$$

where $\varphi \in [0, 2\pi)$ comes from the phases of L_{12} and V_{12} .

Figure 2.2 shows the evolution of populations in a two-state system with $\omega_{21} = 1$ under the approximate optimal control, for $T = 10\pi$ and $T = 6\pi$. Note that even for $T = 6\pi$, the averaged equations are still a good approximation to the full

dynamics.

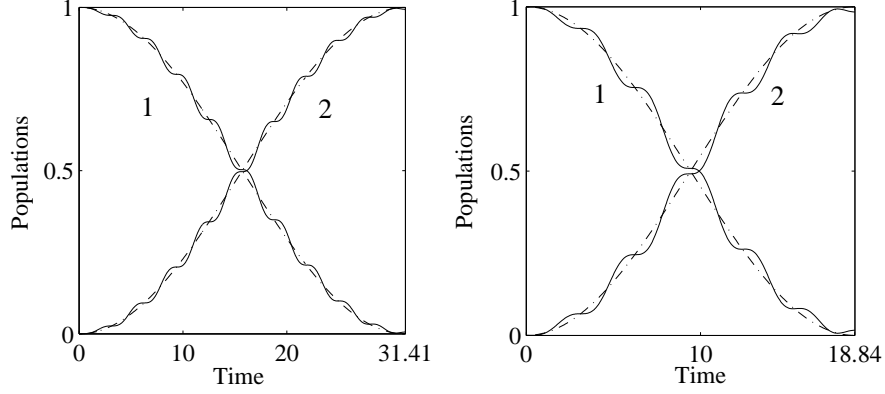


Figure 2.2: Averaged (dashed line) and exact (full line) populations of the two-state system under the approximate optimal control, for $T = 10\pi$ and $T = 6\pi$.

A three-state system

Let us now consider the general three-state system with one control and

$$H_0 = \begin{pmatrix} E_1 & 0 & 0 \\ 0 & E_2 & 0 \\ 0 & 0 & E_3 \end{pmatrix} \text{ and } V = \begin{pmatrix} V_{11} & V_{12} & V_{13} \\ V_{12}^* & V_{22} & V_{23} \\ V_{13}^* & V_{23}^* & V_{33} \end{pmatrix}.$$

We assume that $\omega_{12} = E_1 - E_2$, $\omega_{13} = E_1 - E_3$ and $\omega_{23} = E_2 - E_3$ are all different from each other and from zero. Their exact values are unimportant for the averaged optimal transfer problem as are the values of V_{11} , V_{22} and V_{33} . By rescaling the control, we can take $|V_{12}| = 1$ (we assume $|V_{12}| \neq 0$). Define $p \doteq |V_{23}|^2$ and $r = |V_{13}|^2$. We assume that $1 > p > r \geq 0$, with other cases treated

similarly. We are interested in the transfer

$$\psi_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \longrightarrow \psi_d = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

particularly in the way the “two-photon” transition

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

assists the “direct” (“one-photon”) transition

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

In this example, we are able to calculate the form of the (locally) optimal controls for the averaged problem analytically up to a constant, which has to be computed by solving the state evolution equations numerically and imposing the terminal conditions on the state. For a special case ($p = 1$), we can obtain everything analytically. We begin again with equation (2.43), $\frac{dL}{ds} = [K(L), L]$. The diagonal elements of L are zero and L is anti-Hermitian, so it has only 3 independent

(complex) entries, L_{12} , L_{23} and L_{13} . They satisfy the following equations:

$$\begin{aligned}\frac{dL_{12}}{ds} &= (p-r)L_{13}L_{23}^*, \\ \frac{dL_{23}}{ds} &= (r-1)L_{12}^*L_{13}, \\ \frac{dL_{13}}{ds} &= (1-p)L_{12}L_{23}.\end{aligned}$$

Because $\bar{x}_2(0) = \bar{x}_3(0) = 0$, we have that $L_{32}(0) = 0$. The general solution of the above equations with $L_{32}(0) = 0$ is

$$\begin{aligned}L_{12}(s) &= e^{i\phi_{12}} A \operatorname{cn}(ws), \\ L_{23}(s) &= -e^{i\phi_{23}} B \operatorname{sn}(ws), \\ L_{13}(s) &= e^{i(\phi_{12}+\phi_{23})} C \operatorname{dn}(ws),\end{aligned}$$

where $w > 0$. cn , sn and dn are Jacobi elliptic functions and

$$\begin{pmatrix} A \\ B \\ C \end{pmatrix} = \frac{1}{\sqrt{(1-p)(1-r)(p-r)}} \begin{pmatrix} kw\sqrt{p-r} \\ kw\sqrt{1-r} \\ w\sqrt{1-p} \end{pmatrix}.$$

$0 \leq k \leq 1$ is the modulus of the elliptic functions. Now, from $\bar{x}_1(1) = \bar{x}_2(1) = 0$, we have that $L_{12}(1) = 0$. This allows us to determine w as $w = (2n+1)K(k)$, $n \in \mathbb{N}$, where

$$K(k) \doteq \int_0^{\frac{\pi}{2}} \frac{d\theta}{\sqrt{1-k^2 \sin^2 \theta}},$$

the complete elliptic integral of the first kind, is the quarter-period of the functions cn and sn . The only undetermined parameter is k . It can be solved for numerically

by solving the two-point boundary value problem given by

$$\frac{d\bar{x}}{ds} = \begin{pmatrix} 0 & e^{i\phi_{12}} A \operatorname{cn}(ws) & e^{i(\phi_{12}+\phi_{23})} rC \operatorname{dn}(ws) \\ -e^{-i\phi_{12}} A \operatorname{cn}(ws) & 0 & -e^{i\phi_{23}} pB \operatorname{sn}(ws) \\ -e^{-i(\phi_{12}+\phi_{23})} rC \operatorname{dn}(ws) & e^{-i\phi_{23}} pB \operatorname{sn}(ws) & 0 \end{pmatrix} \bar{x}$$

and the boundary conditions

$$\bar{x}(0) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \bar{x}(1) = \begin{pmatrix} 0 \\ 0 \\ e^{i*} \end{pmatrix}.$$

It is straightforward to see that the phases ϕ_{12} , ϕ_{23} and $\phi_{12} + \phi_{23}$ can be absorbed in the phases of the components of \bar{x} , so k is independent of them and depends only on p and r . Intuitively, one expects a discrete set of solutions for k . The cost of a local minimizer is given by

$$\begin{aligned} J &= 2 \int_0^1 [|u_{12}(s)|^2 + |u_{23}(s)|^2 + |u_{13}(s)|^2] ds \\ &= 2 \int_0^1 (A^2 \operatorname{cn}^2(ws) + rC^2 \operatorname{dn}^2(cs) + pB^2 \operatorname{dn}^2(cs)) ds \\ &= \frac{2(2n+1)^2 K^2(k)}{(1-p)(1-r)(p-r)} [(p-r)k^2 + r(1-p)]. \end{aligned}$$

The expression for the approximate locally optimal controls of optimal transfer problem (I) is

$$\begin{aligned} u(t) = \frac{2}{T} \Big\{ & -A \operatorname{cn}(w \frac{t}{T}) \sin(\omega_{21}t - \alpha_{12} + \phi_{12}) + \\ & \sqrt{p} B \operatorname{sn}(w \frac{t}{T}) \sin(\omega_{32}t - \alpha_{23} + \phi_{23}) - \\ & \sqrt{r} C \operatorname{dn}(w \frac{t}{T}) \sin(\omega_{31}t - \alpha_{13} + \phi_{12} + \phi_{23}) \Big\}, \end{aligned}$$

where $\alpha_{ij} = \arg V_{ij}$.

Figures 2.3 and 2.4 show an approximate locally optimal control, the (slowly-varying) profiles of its Bohr frequency components and the averaged and exact evolution of the three state system with $p = .9$ and $r = .1$, under this control for $T = 20\pi$. Figure 2.4, in particular, demonstrates the point we discussed in the introduction: The ‘averaged’ two-point boundary value problem is non-stiff because the short-time scale natural dynamics of the system has been averaged over and thus, its solutions are much easier to compute compared with those of the original two-point boundary value problem (I). In Appendix C, the special case with $p = 1$ is analyzed in detail and a complete solution to the optimal transfer problem (II) is given for that case.

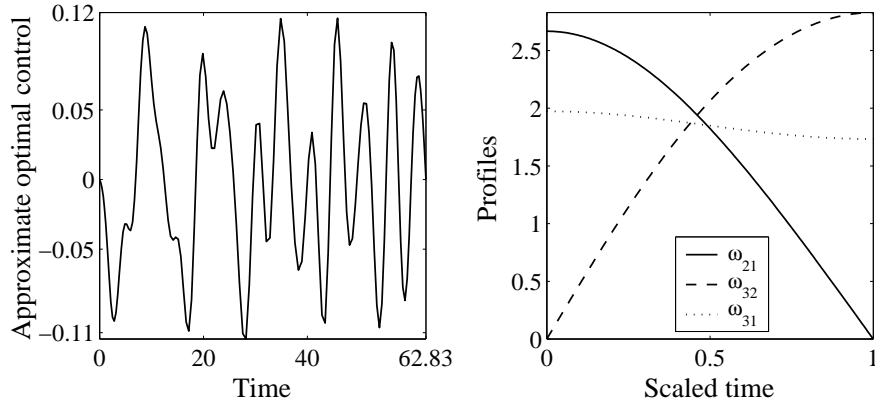


Figure 2.3: Approximate optimal control for the transition $1 \rightarrow 3$ in $T = 20\pi$ and the corresponding frequency profiles.

Bound states of a Morse oscillator

Our final example considers the 22-dimensional space of bound states of the Morse oscillator model for the vibrational dynamics of the ground electronic state of the OH molecule. Here we solve the optimal transfer problem (II) numerically for

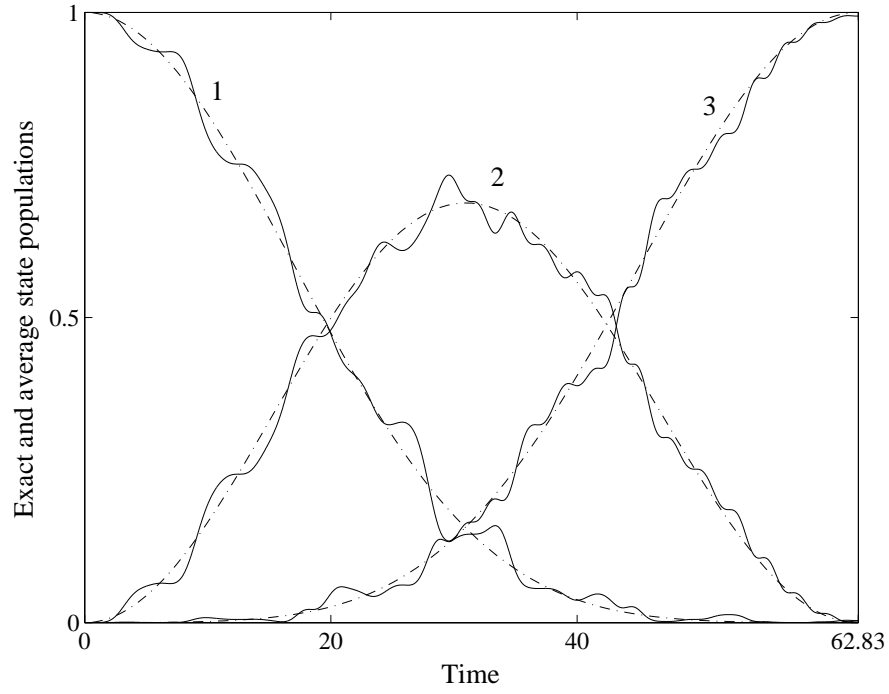


Figure 2.4: Averaged (dashed lines) and exact (full lines) populations of the three-state system under the approximate optimal control.

two transfers, one from the ground vibrational state (state 1) to an intermediate vibrational state (state 10) and one from state 10 to the highest bound state, state 22. Figures 2.5 and 2.6 contain the state populations as well as the intensities (absolute values squared) of the frequency profiles as functions of rescaled time. Note the correspondence between high intensity value for a profile L_{ij} and the transition between states i and j .

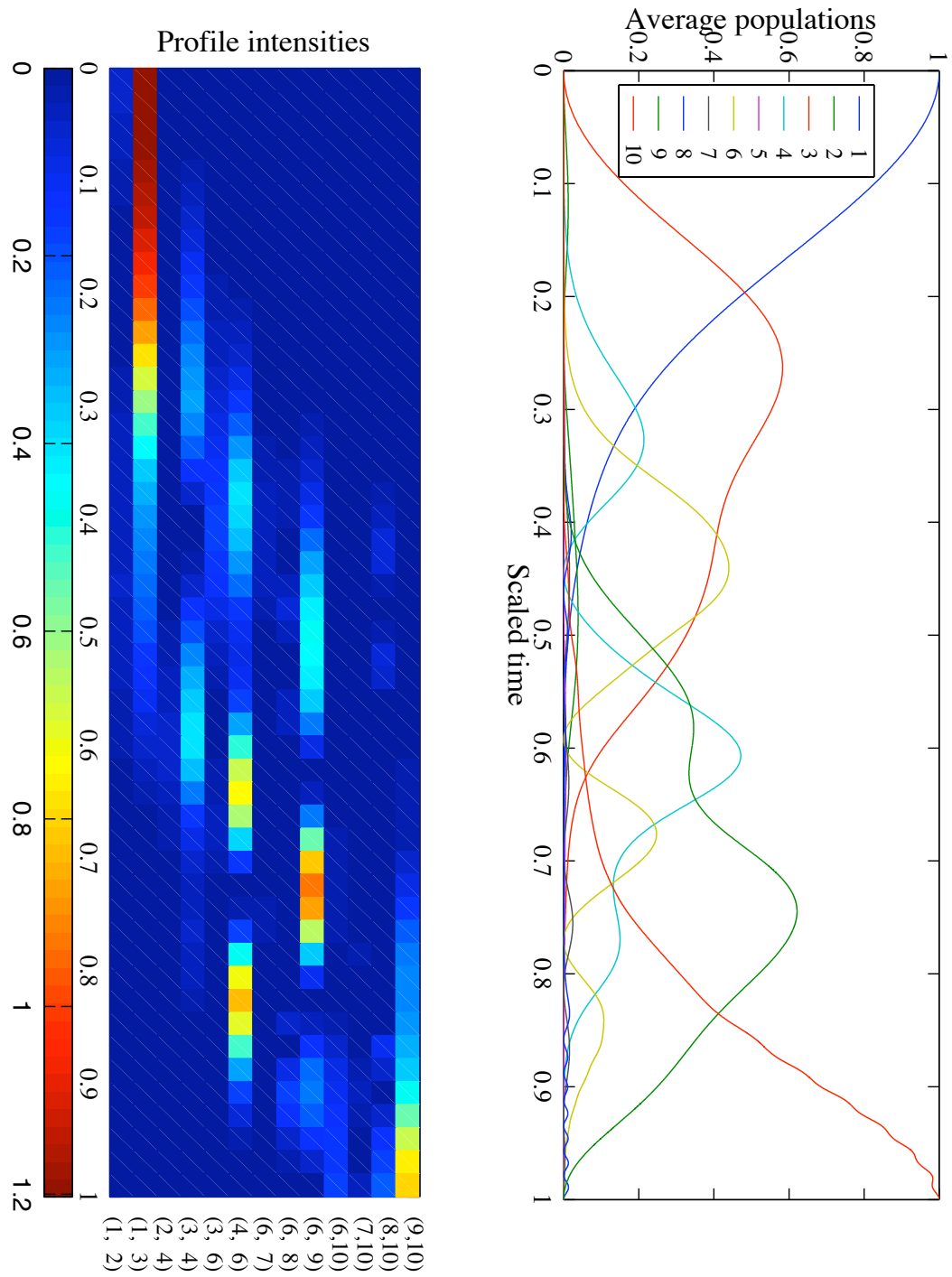


Figure 2.5: Average populations and profile intensities vs. (scaled) time for a locally optimal transfer $1 \rightarrow 10$.

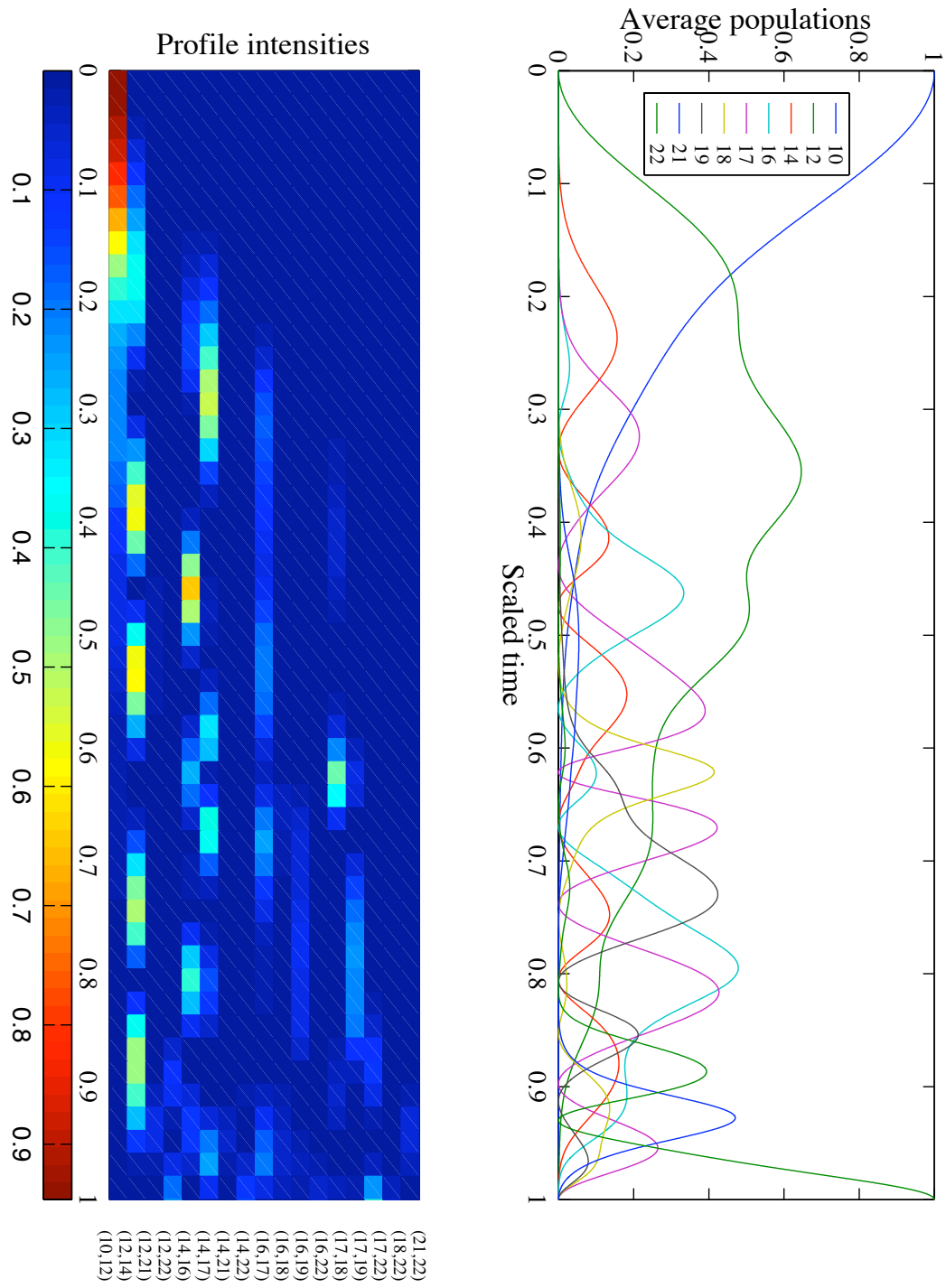


Figure 2.6: Average populations and profile intensities vs. (scaled) time for a locally optimal transfer $10 \rightarrow 22$.

2.7 Proof of main results

Proof of Theorem 1

To begin, we define a new costate variable by $\tilde{\lambda} \doteq T\lambda$ and rewrite (2.20) and (2.21) in terms of $\tilde{\lambda}$:

$$i\dot{\psi} = H_0\psi + \frac{i}{T} (\tilde{\lambda}^* V \psi - \psi^* V \tilde{\lambda}) V \psi, \quad (2.53)$$

$$i\dot{\tilde{\lambda}} = H_0\tilde{\lambda} + \frac{i}{T} (\tilde{\lambda}^* V \psi - \psi^* V \tilde{\lambda}) V \tilde{\lambda}. \quad (2.54)$$

Note that the form of the boundary conditions remains unchanged, as well:

$$\psi(0) = \psi_0, \quad |\psi_i(T)|^2 = p_i, \quad \text{Im}(\psi_i^*(T) \tilde{\lambda}_i(T)) = 0. \quad (2.55)$$

We must show that $e^{-iH_0t} \bar{x}(\frac{t}{T})$ and $e^{-iH_0t} \bar{z}(\frac{t}{T})$ satisfy (2.53) - (2.55) up to terms of order $O(\frac{1}{T})$. To this purpose, we perform one more change of variables in (2.53) - (2.55):

$$y = e^{iH_0t} \psi, \quad l = e^{iH_0t} \tilde{\lambda}.$$

In terms of the new state and costate, the necessary conditions of optimality take the form

$$i\dot{y} = \frac{i}{T} (l^* F(t)y - y^* F(t)l) F(t)y, \quad (2.56)$$

$$i\dot{l} = \frac{i}{T} (l^* F(t)y - y^* F(t)l) F(t)l, \quad (2.57)$$

where, as before, $F(t) = e^{iH_0t} V e^{-iH_0t}$, along with

$$y(0) = \psi_0, \quad |y_i(T)|^2 = p_i, \quad \text{Im}(y_i^*(T) l_i(T)) = 0. \quad (2.58)$$

The boundary conditions retain their form because $|y_i(T)|^2 = |\psi_i(T)|^2$ and, $l_i(T)$ and $y_i(T)$ are rotated by the same amount, $e^{-iE_i T}$, with respect to $\psi_i(T)$ and $\tilde{\lambda}_i(T)$. We will refer to equations (2.56) - (2.58) as two-point boundary value problem (I'). The equivalence of problems (I) and (I') is evident.

Now, we have to show that $\bar{x}(\frac{t}{T})$ and $\bar{z}(\frac{t}{T})$ satisfy (2.56) - (2.58) up to terms of order $O(\frac{1}{T})$. To do this, we average equations (2.56) and (2.57). To make the procedure more transparent, we rewrite equations (2.56) and (2.57) in component form:

$$\begin{aligned}\dot{y}_i &= \frac{1}{T} \left(\sum_{km} V_{km} e^{i\omega_{km}t} (l_k^* y_m - y_k^* l_m) \right) \sum_j V_{ij} e^{i\omega_{ij}t} y_j \\ \dot{l}_i &= \frac{1}{T} \left(\sum_{km} V_{km} e^{i\omega_{km}t} (l_k^* y_m - y_k^* l_m) \right) \sum_j V_{ij} e^{i\omega_{ij}t} l_j\end{aligned}$$

One can see (based on our controllability assumption) that we get non-zero contributions from two groups of terms: Terms with $\omega_{km} \neq 0$ and $\omega_{ij} \neq 0$ such that $m = i$ and $k = j$, and terms with $\omega_{km} = \omega_{ij} = 0$, i.e. $k = m$ and $i = j$. Letting \bar{y} and \bar{l} be the averaged y and l , the averaged state and co-state equations are:

$$\dot{\bar{y}}_i = \frac{1}{T} \left\{ \sum_{j \neq i} |V_{ij}|^2 (\bar{y}_i \bar{l}_j^* - \bar{l}_i \bar{y}_j^*) \bar{y}_j + \sum_k V_{kk} (\bar{y}_k \bar{l}_k^* - \bar{l}_k \bar{y}_k^*) V_{ii} \bar{y}_i \right\}, \quad (2.59)$$

$$\dot{\bar{l}}_i = \frac{1}{T} \left\{ \sum_{j \neq i} |V_{ij}|^2 (\bar{y}_i \bar{l}_j^* - \bar{l}_i \bar{y}_j^*) \bar{l}_j + \sum_k V_{kk} (\bar{y}_k \bar{l}_k^* - \bar{l}_k \bar{y}_k^*) V_{ii} \bar{l}_i \right\}. \quad (2.60)$$

To finish the proof, we rescale time in equations (2.59) and (2.60) to $s = \varepsilon t = \frac{t}{T}$. Letting $\tilde{y}(s) \doteq \bar{y}(t)$ and $\tilde{l}(s) \doteq \bar{l}(t)$, these equations read now:

$$\frac{d\tilde{y}_i}{ds} = \sum_{j \neq i} |V_{ij}|^2 (\tilde{y}_i \tilde{l}_j^* - \tilde{l}_i \tilde{y}_j^*) \tilde{y}_j + \sum_k V_{kk} (\tilde{y}_k \tilde{l}_k^* - \tilde{l}_k \tilde{y}_k^*) V_{ii} \tilde{y}_i, \quad (2.61)$$

$$\frac{d\tilde{l}_i}{ds} = \sum_{j \neq i} |V_{ij}|^2 (\tilde{y}_i \tilde{l}_j^* - \tilde{l}_i \tilde{y}_j^*) \tilde{l}_j + \sum_k V_{kk} (\tilde{y}_k \tilde{l}_k^* - \tilde{l}_k \tilde{y}_k^*) V_{ii} \tilde{l}_i. \quad (2.62)$$

These equations are the same as (2.36) and (2.37) (with the substitution $\tilde{y} \rightarrow \bar{x}$ and $\tilde{l} \rightarrow \bar{z}$). Then, from the sequence of variable changes and the averaging theorem, the conclusion of the theorem follows. ■

Proof of Theorem 2

To prove Theorem 2 we need the following lemma:

Lemma: Let ψ_0 and $\{p_i\}_{i=1,\dots,N}$ be an initial state and a target population of system (2.10), respectively, and let $(\bar{x}(s), \bar{z}(s))$ be a solution of two-point boundary value problem (II) over $[0, 1]$. The set of pairs $(\psi_0, \{p_i\})$ of initial states and target populations for which all solutions of two-point boundary value problem (II) are *isolated* is open and full measure in the corresponding product manifold. \square

Proof of Lemma: We begin by introducing new coordinates for the state and costate of the optimal transfer problem (II) by

$$\bar{x}_i \doteq I_i e^{i\phi_i}, \quad \bar{z}_i \doteq J_i e^{i\theta_i}, \quad (2.63)$$

where $\phi_i \doteq \arg \bar{x}_i \bmod \pi$ and $I_i \doteq |\bar{x}_i|$ for $\text{Im} \bar{x}_i \geq 0$ and $I_i \doteq -|\bar{x}_i|$ for $\text{Im} \bar{x}_i < 0$, and similarly for the θ_i 's and J_i 's. Then, $I_i, J_i \in \mathbb{R}$ and $\phi_i, \theta_i \in [0, \pi)$, $\forall i = 1, \dots, N$. With this definition, the phases ϕ_i and θ_i have discontinuities whenever the signs of the imaginary parts of \bar{x}_i and \bar{z}_i change. We shall see that this will not be a problem for us because ϕ_i and θ_i will turn out to be constant in time. On the other hand, the introduction of these coordinates will prove to be beneficial in the following.

In the new coordinates, (2.45) and (2.46), take the form

$$i\dot{I}_i - I_i\dot{\phi}_i = +iI_i \sum_{j \neq i} |V_{ij}|^2 I_j J_j e^{i(\phi_j - \theta_j)} - iJ_i e^{i(\theta_i - \phi_i)} \sum_{j \neq i} |V_{ij}|^2 I_j^2, \quad (2.64)$$

$$i\dot{J}_i - J_i\dot{\theta}_i = -iJ_i \sum_{j \neq i} |V_{ij}|^2 I_j J_j e^{i(\theta_j - \phi_j)} + iI_i e^{i(\phi_i - \theta_i)} \sum_{j \neq i} |V_{ij}|^2 J_j^2, \quad (2.65)$$

while the boundary conditions at the end become

$$I_i^2(1) = p_i, \quad \theta_i(1) - \phi_i(1) = 0. \quad (2.66)$$

Multiply (2.64) by J_i , (2.65) by $-I_i$, add them and take the real part. The resulting equation reads:

$$I_i J_i (\phi_i - \theta_i)' = \sin(\phi_i - \theta_i) \{ I_i^2 \sum_{j \neq i} |V_{ij}|^2 J_j^2 - J_i^2 \sum_{j \neq i} |V_{ij}|^2 I_j^2 \}. \quad (2.67)$$

Except for the degenerate cases ($I_i(s) = 0$, $J_i(s) = \text{const.}$) and ($J_i(s) = 0$, $I_i(s) = \text{const.}$) $\forall s \in [0, 1]$, which will be excluded, we see that, given the transversality conditions at $s = 1$, this equation implies that $\theta_i(s) = \phi_i(s)$, $\forall s \in [0, 1]$. Using this fact, we see that the right sides of equations (2.64) and (2.65) are purely imaginary. This leads to the simplified equations

$$\begin{aligned} \dot{I}_i &= +I_i \sum_{j \neq i} |V_{ij}|^2 I_j J_j - J_i \sum_{j \neq i} |V_{ij}|^2 I_j^2, \\ \dot{J}_i &= -J_i \sum_{j \neq i} |V_{ij}|^2 I_j J_j + I_i \sum_{j \neq i} |V_{ij}|^2 J_j^2, \\ \dot{\phi}_i(s) &= 0, \quad \dot{\theta}_i(s) = 0. \end{aligned}$$

We see that the two-point boundary value problem (II) separates nicely into two

parts: A trivial problem for the arguments of state and costate components

$$\dot{\phi}_i(s) = 0, \quad \dot{\theta}_i(s) = 0, \quad (2.68)$$

$$\phi_i(0) = \phi_{i0}, \quad \theta_i(1) - \phi_i(1) = 0, \quad (2.69)$$

which has a unique solution as long as $\psi_{0i} \neq 0$, $\forall i = 1, \dots, N$ (so that all $\phi_i(0)$ are unambiguously defined) and a two-point boundary value problem for the (signed) magnitudes of state and costate components

$$\dot{I}_i = +I_i \sum_{j \neq i} |V_{ij}|^2 I_j J_j - J_i \sum_{j \neq i} |V_{ij}|^2 I_j^2, \quad (2.70)$$

$$\dot{J}_i = -J_i \sum_{j \neq i} |V_{ij}|^2 I_j J_j + I_i \sum_{j \neq i} |V_{ij}|^2 J_j^2, \quad (2.71)$$

$$I_i(0) = I_{i0}, \quad I_i^2(1) = p_i. \quad (2.72)$$

We will refer to (2.70) - (2.72) as the real form of problem (II). These two-point boundary value problems can be seen to be necessary conditions for optimal population transfers of the following *real* control system on S^{N-1} :

$$\frac{dI}{ds} = \begin{pmatrix} 0 & |V_{12}|v_{12} & \cdots \\ -|V_{21}|v_{21} & 0 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} I. \quad (2.73)$$

Here the v_{ij} are real controls and (2.73) is controllable on S^{N-1} because of the controllability assumption on the original system, see Appendix B. This separation of the two-point boundary value problem (II) into a trivial problem for the evolution of the arguments of state and costate components and real two-point boundary value problem (II) was inspired by [64], see section 4 of that reference.

We are now ready to prove the lemma. From the discussion so far, it should be

obvious that we need $\psi_{i0} \neq 0$ and $|\psi_{0i}|^2 \neq p_i, \forall i = 1, \dots, N$ so that the (constant) phases of the state components are uniquely defined and certain pathological cases mentioned are excluded. The set of pairs $(\psi_0, \{p_i\})$ of initial states and target populations for which this is the case is open and full measure in the corresponding product manifold.

We define a terminal condition function $\mathcal{G} : \mathbb{R}^{N-1} \rightarrow S^{N-1}$ of the initial costate vector $J(0)$ of real problem (II) by

$$\mathcal{G}(J(0)) \doteq I(1). \quad (2.74)$$

\mathcal{G} is a function of only $N-1$ components of $J(0)$ because one of them is fixed by the condition $0 = \bar{x}^*(0)\bar{z}(0) = I^T(0)J(0)$ (let us decide to fix $J_1(0)$). In the following, when talking about initial costate vectors, we will identify vectors in \mathbb{R}^{N-1} with vectors in \mathbb{R}^N perpendicular to I_0 . \mathcal{G} has two important properties: First, it is a smooth function of its argument. Indeed, the right sides of (2.70) and (2.71) are C^∞ functions of I and J so, their solutions depend smoothly on $I(0)$ and $J(0)$ [65]. Thus, when $I(0)$ is fixed, the terminal conditions depend smoothly on $J(0)$. Second, because of controllability of system (2.73), \mathcal{G} is onto. Then, since \mathcal{G} is a mapping between manifolds of the same dimension, it is a diffeomorphism from (open) neighborhoods of \mathbb{R}^{N-1} to neighborhoods of S^{N-1} . Additionally, Sard's theorem [66] implies that the set of critical values of \mathcal{G} in S^{N-1} (i.e. the set of points around which \mathcal{G} is not a diffeomorphism) is of measure zero.

Any initial costate vector $J(0)$ such that

$$\mathcal{G}(J(0)) = I(1) = (\pm\sqrt{p_1}, \dots, \pm\sqrt{p_N})^T$$

provides a solution to real problem (II). Let's fix a sequence of signs in the com-

ponents of $I(1)$, say all positive. The same conclusions will hold for all such sequences. Because of the properties of \mathcal{G} , the set of p_i 's for which *all* $J(0)$ that satisfy $\mathcal{G}(J(0)) = (\sqrt{p_1}, \dots, \sqrt{p_N})^T$ satisfy also $\mathcal{N}(D\mathcal{G}(J(0))) = \emptyset$, is an open set of full measure in S^{N-1} . For a fixed pair of initial state and final populations which satisfy this, the set of $J(0)$'s must be discrete. Also, it cannot have a limit point. Indeed, if there exists a subsequence $\{J(0)\}_{n_k} \rightarrow J_0$, then

$$\mathcal{G}(J_0) = \lim_{k \rightarrow \infty} \mathcal{G}(\{J(0)\}_{n_k}) = (\sqrt{p_1}, \dots, \sqrt{p_N})^T,$$

and also $\mathcal{N}(D\mathcal{G}(J_0)) \neq \emptyset$, which violates the assumption that for this given transfer all the $J(0)$'s that achieve it satisfy $\mathcal{N}(D\mathcal{G}(J(0))) = \emptyset$. Hence, given an initial state with all populations non-zero, the set of populations that can be achieved by *isolated* locally optimal transfers of the real problem (II) is open and full measure. So, the set of pairs $(\psi_0, \{p_i\})$ of initial states and target populations for which all solutions of two-point boundary value problem (II) are *isolated* is open and full measure in the corresponding product manifold. ■

Proof of Theorem 2: We are going to define terminal condition functions for two-point boundary value problems (II) and (I') in a way similar to that in the proof of the lemma. First, for (II) we define $\mathcal{F} : \mathbb{R}^{2N-2} \longrightarrow \mathbb{R}^{2N-2}$ by

$$\mathcal{F}(\bar{z}(0)) \doteq (|\bar{x}_2^2(1)|, \dots, |\bar{x}_N^2(1)|, \text{Im}(\bar{x}_2^*(1) \bar{z}_2(1)), \dots, \text{Im}(\bar{x}_N^*(1) \bar{z}_N(1)))^T \quad (2.75)$$

(Recall that one complex component of $\bar{z}(0)$ is fixed by $\bar{x}^*(0)\bar{z}(0) = 0$. Here, we again identify initial costate vectors in \mathbb{C}^N perpendicular to ψ_0 with vectors in \mathbb{C}^{N-1} and also identify \mathbb{C}^{N-1} with \mathbb{R}^{2N-2}). \mathcal{F} is a smooth function of its argument. Also, any initial costate vector $\bar{z}(0)$ such that $\mathcal{F}(\bar{z}(0)) = (p_2, \dots, p_N, 0, \dots, 0)^T$ provides a solution to problem (II).

The corresponding terminal condition function \mathcal{F}_1 for problem (I'), is defined exactly the same way:

$$\mathcal{F}_1(l(0); T) \doteq (|y_2(T)|^2, \dots, |y_N(T)|^2, \text{Im}(y_2^*(T) l_2(T)), \dots, \text{Im}(y_N^*(T) l_N(T)))^T \quad (2.76)$$

The second argument of \mathcal{F}_I is just a reminder of the transfer time. \mathcal{F}_1 is also a smooth function of its argument. Again, any initial costate vector $y(0)$ such that $\mathcal{F}_1(y(0)) = (p_2, \dots, p_N, 0, \dots, 0)^T$ provides a solution to problem (I'). From the proof of theorem (I), we know that

$$\mathcal{F}_1(v; T) = \mathcal{F}(v) + O\left(\frac{1}{T}\right). \quad (2.77)$$

From this we also have that

$$D\mathcal{F}_1(v; T) = D\mathcal{F}(v) + O\left(\frac{1}{T}\right).$$

Although \mathcal{F}_1 is not formally defined for $T = \infty$ from (2.76), we can define it from (2.77) as $\mathcal{F}_1(v; 0) \doteq \mathcal{F}(v)$. With this definition, \mathcal{F}_1 is continuous in $\frac{1}{T}$ at $\frac{1}{T} = 0$, with continuous first derivatives in v and $\frac{1}{T}$ there. In particular, $D\mathcal{F}_1(v; \infty) = D\mathcal{F}(v)$.

Consider now an initial costate vector v that solves problem (II), i.e.

$$\mathcal{F}(v) = (p_2, \dots, p_N, 0, \dots, 0)^T$$

We have seen that the set of pairs $(\psi_0, \{p_i\})$ such that *all* v 's that satisfy this are isolated ($D\mathcal{F}(v)$ full rank), is an open set of full measure in the product space. For such a transfer and for large enough T , the implicit function theorem guarantees the existence of a δv such that $\mathcal{F}_1(v + \delta v; T) = (p_2, \dots, p_N, 0, \dots, 0)^T$. Then, $v + \delta v$

provides a solution for problem (I') and $\frac{v+\delta v}{T}$ is a solution for problem (I). Taking T large enough so that $|\delta v| = O(\frac{1}{T})$, the averaging theorem guarantees that this solution $(\psi(t), \lambda(t))$ to two-point boundary value problem (I) is such that

$$\begin{aligned}\psi(t) &= e^{-iH_0 t} \bar{x}\left(\frac{t}{T}\right) + O\left(\frac{1}{T}\right), \\ \lambda(t) &= \frac{1}{T} e^{-iH_0 t} \bar{z}\left(\frac{t}{T}\right) + O\left(\frac{1}{T^2}\right),\end{aligned}$$

where $(\bar{x}(s), \bar{z}(s))$ is the solution to problem (II) we started with. ■

Proof of Theorem 3

We only consider transfers such that problem (II) has isolated solutions. Any initial costate v that satisfies $\mathcal{F}(v) = (p_2, \dots, p_N, 0, \dots, 0)^T$ provides a solution to problem (II) and $D\mathcal{F}(v)$ is full rank. Inside a ball of radius $M > 0$ there can be only a *finite* number of these initial costates v because any discrete set with no limit points inside a compact set must be finite.

From equation (2.77) we may conclude that for a given transfer, we can take T large enough to bound the difference of \mathcal{F} and \mathcal{F}_1 over an open ball around the origin by any $\delta > 0$:

$$\forall M > 0, \forall \delta > 0, \exists T > 0 \text{ such that } |\mathcal{F}_1(v; T) - \mathcal{F}(v)| < \delta, \forall |v| < M.$$

Thus, the only solutions of $\mathcal{F}_1(\tilde{v}; T) = (p_2, \dots, p_N, 0, \dots, 0)^T$ inside the ball of radius M come from perturbing solutions of $\mathcal{F}(v) = (p_2, \dots, p_N, 0, \dots, 0)^T$ by quantities of order $O(\frac{1}{T})$, for T large enough. In particular, they also form a finite set and $D\mathcal{F}_1(\tilde{v}; T)$ is full rank for each such \tilde{v} . We arrive at exactly the same conclusion if, instead of a ball, we define a neighborhood of the origin by an ellipsoid.

We introduce now the following quadratic form in $v \in \{w \in \mathbb{C}^N \text{ s.t. } \psi_0^* w = 0\} \simeq \mathbb{C}^{N-1} \simeq \mathbb{R}^{2N-2}$:

$$E(v) = \sum_{i \neq j}^N |V_{ij}|^2 |\psi_{0i} v_j^* - v_i \psi_{0j}^*|^2.$$

E is non-negative and, furthermore, due to the connectivity of the graph of V (part of the controllability assumption) can be shown to be positive definite. Indeed, suppose that $\exists v \in \{w \in \mathbb{C}^N \text{ s.t. } \psi_0^* w = 0\}$ such that $E(v) = 0$. Then, we have that

$$\psi_{0i} v_j^* - v_i \psi_{0j}^* = 0, \forall (i, j) \text{ s.t. } V_{ij} \neq 0. \quad (2.78)$$

Let us assume for a moment that $V_{12} \neq 0$. Then (recall, all $\psi_{0i} \neq 0$),

$$\frac{v_2}{\psi_{02}} = \frac{v_1^*}{\psi_{01}^*}.$$

It is straightforward to see that the connectivity of the graph of V and the repeated use of the relations (2.78) allows us to show that

$$\frac{v_i}{\psi_{0i}} = \frac{v_1}{\psi_{01}} \text{ or } \frac{v_1^*}{\psi_{01}^*}, \forall i = 2, \dots, N.$$

Recall now that both ψ_0 and v can be defined modulo global phases which we choose such that ψ_{01} and v_1 are real. Then,

$$\frac{v_i}{\psi_{0i}} = \frac{v_1}{\psi_{01}}, \forall i = 2, \dots, N.$$

The relation $\psi_0^* v = 0$ implies that

$$\left(\sum_{i=1}^N |\psi_{0i}|^2 \right) \frac{v_1}{\psi_{01}} = 0,$$

which means that $v_1 = 0$ and hence $v = 0$.

Since E is positive definite, its sub-level sets define ellipsoids in \mathbb{R}^{2N-2} . The significance of our choice for $E(v)$ lies in the following:

$$H(\bar{x}, \bar{z}) = \sum_{i \neq j} |V_{ji}|^2 |\bar{x}_i \bar{z}_j^* - \bar{z}_i \bar{x}_j^*|^2$$

is the Hamiltonian function from which the optimal state and costate equations (2.45) and (2.46) are derived. Thus, H is a constant of motion along the optimal solutions. Note, also, that $H(\bar{x}, \bar{z}) = \sum_{i \neq j} |u_{ij}|^2$. Since E is just H evaluated at $t = 0$, we conclude that E is equal to the cost of a trajectory of system (II) (equations (2.45) and (2.46)) with initial conditions (ψ_0, v) :

$$\int_0^1 \sum_{i \neq j} |u_{ij}(s)|^2 ds = H(0) = \sum_{i \neq j}^N |V_{ij}|^2 |\psi_{0i} v_j^* - v_i \psi_{0j}^*|^2.$$

With a calculation similar to that of section 2.4, one can show that the locally optimal costs for problems (II) and (I/I') coming from the solutions v and $\tilde{v} = v + O(\frac{1}{T})$, respectively are related as follows:

$$\begin{aligned} \int_0^T u^2(t) dt &= \frac{1}{T} \left(\int_0^1 \sum_{i \neq j} |u_{ij}(s)|^2 ds + O\left(\frac{1}{T}\right) \right) \\ &= \frac{1}{T} \left(\sum_{i \neq j}^N |V_{ij}|^2 |\psi_{0i} v_j^* - v_i \psi_{0j}^*|^2 + O\left(\frac{1}{T}\right) \right). \end{aligned}$$

Let us fix a value $E_0 > 0$ such that the initial costate v_0 that achieves the desired transfer with the *minimum* energy for problem (II) satisfies

$$\sum_{i \neq j}^N |V_{ij}|^2 |\psi_{0i} v_j^* - v_i \psi_{0j}^*|^2 < E_0.$$

Then, for large enough T , $\tilde{v}_0 = v_0 + O(\frac{1}{T})$ is the initial costate that achieves the desired transfer with the *minimum* energy for problem (I'). This proves the assertion of the theorem. In fact, we proved a little bit more: Not only the global optimum, but also all local optima of problem (I') with values of energy less than $\frac{E_0}{T}$, come from local optima of problem (II) according to the theorem 2, for T large enough (Note that the solutions \tilde{v} of problem (I/I') *outside* the ellipsoid $E(\tilde{v}) < E_0$ have higher costs than those inside the ellipsoid).■

2.8 Summary

We examined the large transfer time limit of exact, energy-optimal population transfers in a finite dimensional quantum system. The investigation of this problem uncovered useful structure in the optimal control and state trajectory and, moreover, resulted in the much simpler optimal control problem (II) whose solution provides the first order solution to the original optimal transfer problem, in a $\frac{1}{T}$ expansion. The main reason we considered *exact* population transfers is that, we could prove Theorems 2 and 3 which are important for uncovering the structure of solutions of optimal transfer problem (I). On the other hand, solving the two-point boundary value problem (II) becomes extremely difficult as the dimension of the system grows. For practical applications to large-dimensional systems, an optimal transfer problem with an objective like (2.16)

$$a \int_0^T u^2(t) dt + \sum_{i=1}^N m_i |\psi_i(T)|^2,$$

which leads to *separated* boundary conditions would be preferable because iterative techniques exist that converge to local minima [4, 31]. One can obtain an averaged transfer problem here too and prove an analog of Theorem 1 (see Chap-

ter 3, where this is done for an open quantum system) but we haven't been able to do so for Theorems 2 and 3 yet. Nevertheless, we think that this methodology, of considering the long transfer time limit and separating the time-scales of natural and controlled evolution offers a conceptual and a computational simplification in all kinds of population transfer problems for quantum systems.

2.9 Appendices to Chapter 2

Appendix A

First, we prove the transversality conditions $\text{Im}(\psi_i^*(T) \lambda_i(T)) = 0$, $i = 1, \dots, N$. The optimal transfer problem (I) is a standard Bolza problem with (real) terminal state constraints, $|\psi_i(T)|^2 = p_i$, $i = 1, \dots, N$. According to the general theory of such problems [63], the transversality conditions for the costate at the final time are given by

$$\lambda_i(T) = \nu_i \frac{\partial(|\psi_i(T)|^2 - p_i)}{\partial \psi_i(T)^*} = \nu_i \psi_i(T), \quad (2.79)$$

where ν_i are the real Lagrange multipliers that enforce the terminal state constraints. The transversality conditions in the form we state them,

$$\text{Im}(\psi_i^*(T) \lambda_i(T)) = 0, \quad i = 1, \dots, N$$

follow easily.

Next, we prove that $\lambda(t)^* \psi(t) = 0$. One can easily show from equations (2.17) and (2.18) that $\lambda(t)^* \psi(t)$ is constant along any optimal trajectory and so, $\lambda(t)^* \psi(t) = \lambda(T)^* \psi(T)$. From the transversality conditions it follows that $\text{Im}(\psi^*(T) \lambda(T)) = 0$ and thus, $\text{Im}(\psi^*(t) \lambda(t)) = 0$. Let us now decompose λ as follows, $\lambda = c\psi + \lambda_\perp$, with c real and λ_\perp perpendicular to ψ , i.e. $\lambda_\perp^* \psi = 0$. All we

need to show is that $c = 0$. We introduce this representation of λ into equation (2.19). The resulting equation is

$$u = i(\lambda_{\perp}^* V \psi - \psi^* V \lambda_{\perp}) + i(c^* - c)(\psi^* V \psi),$$

and the reality of u forces c to be zero.

Appendix B

We prove that system (2.31) is controllable on account of the controllability assumption on the original system (2.10). Due to the fact that (2.31) has complex controls, every non-zero $V_{ij}, i \neq j$, provides us with two generators of $su(N)$ (traceless anti-Hermitian matrices), $E_{ij} - E_{ji}$ and $i(E_{ij} + E_{ji})$ (E_{ij} denotes a matrix with only one non-zero element, at the position (i, j) , equal to one. So, $(E_{ij})_{kl} = \delta_{ik}\delta_{jl}$. It is easy to see that $E_{ij}E_{mn} = \delta_{jm}E_{in}$). One can easily verify the following commutation relations:

$$[E_{ij} - E_{ji}, E_{jk} - E_{kj}] = E_{ik} - E_{ki}, \quad i \neq k, \quad (2.80)$$

$$E_{ij} - E_{ji}, i(E_{jk} + E_{kj}) = i(E_{ik} + E_{ki}), \quad i \neq k, \quad (2.81)$$

$$E_{ij} - E_{ji}, i(E_{ji} + E_{ij}) = 2i(E_{ii} - E_{jj}). \quad (2.82)$$

Due to the connectivity of the graph of V , there exists a sequence of index pairs that connects any state index i with any other state index j . Thus, starting with the given generators $E_{ij} - E_{ji}$ and $i(E_{ij} + E_{ji})$ for all $V_{ij} \neq 0, i \neq j$, we can generate, with repeated use of (2.80) and (2.81), all missing such generators (corresponding to $V_{ij} = 0, i \neq j$). Finally, using (2.82) we can generate the diagonal generators of $su(N)$ (its Cartan subalgebra). For the system (2.73) which has real controls, every non-zero $V_{ij}, i \neq j$, provides us with a generator of $so(N)$ (anti-symmetric

matrices), $E_{ij} - E_{ji}$. In that case, the repeated use of (2.80) is enough to establish controllability.

Appendix C

In the special case $p = 1$, we can give a complete analytic solution to the optimal transfer problem (II) for the second example of section 2.6. The solution for the matrix elements of L is given by

$$\begin{aligned} L_{12}(s) &= e^{i\phi_{12}} A \cos(ws), \\ L_{23}(s) &= -e^{i\phi_{23}} A \sin(ws), \\ L_{13}(s) &= e^{i(\phi_{12}+\phi_{23})} \frac{w}{1-r}, \end{aligned}$$

where w is determined to be $w = (2n + 1)\frac{\pi}{2}$, $n \in \mathbb{N}$. To determine A , we have to solve the two-point boundary value problem given by (2.41),

$$\frac{d\bar{x}}{ds} = \begin{pmatrix} 0 & e^{i\phi_{12}} A \cos(ws) & e^{i(\phi_{12}+\phi_{23})} \frac{rw}{1-r} \\ -e^{-i\phi_{12}} A \cos(ws) & 0 & -e^{i\phi_{23}} A \sin(ws) \\ -e^{-i(\phi_{12}+\phi_{23})} \frac{rw}{1-r} & e^{-i\phi_{23}} A \sin(ws) & 0 \end{pmatrix} \bar{x},$$

and the boundary conditions

$$\bar{x}(0) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \bar{x}(1) = \begin{pmatrix} 0 \\ 0 \\ e^{i*} \end{pmatrix}.$$

With the change of variables

$$y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \doteq \begin{pmatrix} \cos(ws) & 0 & \sin(ws) \\ 0 & 1 & 0 \\ -\sin(ws) & 0 & \cos(ws) \end{pmatrix} \begin{pmatrix} \bar{x}_1 \\ e^{i\phi_{12}} \bar{x}_2 \\ e^{i(\phi_{12}+\phi_{23})} \bar{x}_3 \end{pmatrix},$$

y satisfies the simpler equation

$$\frac{dy}{ds} = \begin{pmatrix} 0 & A & \frac{w}{1-r} \\ -A & 0 & 0 \\ -\frac{w}{1-r} & 0 & 0 \end{pmatrix} y.$$

The boundary conditions for \bar{x} translate into the following conditions for y :

$$y(0) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad y(1) = \begin{pmatrix} \pm 1 \\ 0 \\ 0 \end{pmatrix}.$$

It is a straightforward calculation to find the solution for $y(s)$ and impose the boundary conditions. We find then that A must satisfy the condition

$$\cos \sqrt{A^2 + \frac{w^2}{(1-r)^2}} = \pm 1$$

which implies that

$$A = A(m, n) = \sqrt{(m\pi)^2 - \frac{w^2}{(1-r)^2}} = \pi \sqrt{m^2 - \frac{(n+1/2)^2}{(1-r)^2}},$$

where $m \in \mathbb{N}$ is such that $m \geq \frac{(n+1/2)}{(1-r)}$. The solution for the state evolution is given by

$$\begin{pmatrix} \bar{x}_1 \\ e^{i\phi_{12}} \bar{x}_2 \\ e^{i(\phi_{12}+\phi_{23})} \bar{x}_3 \end{pmatrix} (s) = \begin{pmatrix} \cos((n + \frac{1}{2})\pi s) & 0 & \sin((n + \frac{1}{2})\pi s) \\ 0 & 1 & 0 \\ -\sin((n + \frac{1}{2})\pi s) & 0 & \cos((n + \frac{1}{2})\pi s) \end{pmatrix} \begin{pmatrix} \cos(m\pi s) \\ -A(m, n) \frac{\sin(m\pi s)}{m\pi} \\ -\frac{n+1/2}{m(1-r)} \sin(m\pi s) \end{pmatrix}.$$

We see that the local minimizers for this optimal transfer problem are parameterized in terms of two integers, $n \in \mathbb{N}$ and $m \in \mathbb{N}$ such that $m \geq \frac{(n+1/2)}{(1-r)}$. The cost of such a minimizer can be easily computed to be equal to

$$J = 2\pi^2 \left[m^2 - \frac{(n + 1/2)^2}{1 - r} \right].$$

The values of m and n and the exact cost for the global minimum are uniquely specified by the value of r .

Chapter 3

Optimal transfers in open quantum systems

Our presentation is organized as follows: We begin in section 3.1 with an introduction to the Quantum Mechanics of open systems, in particular the density matrix formalism and the Lindblad equation. In section 3.2 we discuss the cooling of molecular motions and the role that dissipation mechanisms play in it. We consider the problem of optimal control of cooling of the rotational motion of molecules in section 3.3. This leads us to study an optimal transfer problem for the Lindblad equation with Hamiltonian control. In section 3.4 we propose a method of approximate solution to the two-point boundary-value problem that expresses the necessary conditions of optimality for the transfer. This method takes advantage of the separation between the free Hamiltonian evolution time-scale (fast) and the dissipation time-scale (slow). It substitutes the original two-point boundary-value problem with an “averaged” one, where the fast Hamiltonian free dynamics and control action have been averaged over. The advantages of this approach are twofold: The numerical solution of the averaged problem is much more efficient than that of the original one, and once obtained, it provides the (approximate)

solution of the original problem in a structured form, where the two time-scales are clearly separated. Using this technique, we solve an optimal cooling problem for a three-level Λ system in section 3.5 and we compare our solution with an existing solution for this problem obtained by a different approximation and find excellent agreement.

3.1 Quantum Mechanics of open systems

3.1.1 The density matrix formalism

In chapter 2 we introduced the basic formalism of Quantum Mechanics, how to describe the dynamics of a quantum system in terms of its Hilbert space and the Schrödinger equation. We also incorporated the effects of external *classical* forces, electric or magnetic fields for example, which can be used to control the system. The question of how these external forces are created and manipulated was not addressed. We idealized the physical systems (apparatuses) that create them and ignored their internal dynamics concentrating only on their effect on the quantum system under study.

Here, we'll present the extension of the basic formalism that treats interactions between two quantum systems. We will then show how to pass to an effective description for one of the systems when the other one is ignored, usually due to the fact that it is a much larger system and there is insufficient information about its state. This will lead to the **density matrix** formalism. In the case of interest here, the “ignored” quantum system is the environment of the quantum system under consideration. We emphasize that the density matrix formalism is derived from the Hilbert space/state vector formalism and thus, nothing new is added to the basic structure of Quantum Mechanics.

We begin by considering the interaction of two quantum systems, A and B, with respective Hilbert spaces \mathcal{H}_A and \mathcal{H}_B . The Hilbert space for the composite system, \mathcal{H}_{AB} , is simply the tensor product of \mathcal{H}_A and \mathcal{H}_B :

$$\mathcal{H}_{AB} \doteq \mathcal{H}_A \otimes \mathcal{H}_B.$$

Let $\{e_i\}_{i=1,\dots,N_1}$ and $\{E_I\}_{I=1,\dots,N_2}$ be orthonormal bases for \mathcal{H}_A and \mathcal{H}_B , respectively. Then, $\{e_i \otimes E_I\}$ is an orthonormal basis for \mathcal{H}_{AB} and a general state vector Ψ_{AB} has the representation

$$\Psi_{AB} = \sum_{iI} (\Psi_{AB})_{iI} e_i \otimes E_I.$$

Whenever $(\Psi_{AB})_{iI}$ can be written as

$$(\Psi_{AB})_{iI} = \psi_{Ai} \psi_{BI}, \quad \forall i, I,$$

then, Ψ_{AB} *factorizes* as

$$\Psi_{AB} = \psi_A \otimes \psi_B,$$

with $\psi_A \doteq \sum_i \psi_{Ai} e_i$ and $\psi_B \doteq \sum_I \psi_{BI} E_I$. Such a state of the composite system is called **separable** and for such a state it is meaningful to talk about the individual states of A and B as defined above. The physical meaning of this separation will become more obvious when we talk about operators and measurements for the composite system in a moment. For now, note that this kind of state is rather rare and in general, Ψ_{AB} will not be factorizable. In that case, Ψ_{AB} will be called **entangled**, to emphasize the fact that in the composite system AB, the individual systems “share information about each other”.

We continue with the description of operators in \mathcal{H}_{AB} : They have the general

structure

$$O_{AB} = \sum_{\nu} O_A^{(\nu)} \otimes O_B^{(\nu)},$$

where $O_A^{(\nu)}$ and $O_B^{(\nu)}$ are operators in \mathcal{H}_A and \mathcal{H}_B , respectively. If O_A is an operator in \mathcal{H}_A , then $O_A \otimes Id_B$ represents the same operator in the composite system. The Hamiltonian operator for the total system has the form:

$$H_{AB} = H_A \otimes Id_B + Id_A \otimes H_B + H_{AB}^{int}.$$

$H_{A,B}$ are the Hamiltonians of A and B as isolated systems and

$$H_{AB}^{int} = \sum_{\nu} S_A^{(\nu)} \otimes T_B^{(\nu)}$$

is their interaction Hamiltonian. The Schrödinger equation for the composite system is again

$$i\hbar \frac{d}{dt} \Psi_{AB}(t) = H_{AB} \Psi_{AB}(t).$$

We note the following fact: If there was no interaction term in H_{AB} , i.e. $H_{AB} = H_A \otimes Id_B + Id_A \otimes H_B$, an initial separable state $\Psi_{AB} = \psi_A \otimes \psi_B$ would continue to be separable:

$$\Psi_{AB}(t) = (e^{-iH_A t} \psi_A) \otimes (e^{-iH_B t} \psi_B).$$

The entanglement of states is exactly the effect of the interaction, as one would expect. To physically explain the difference between separable and entangled states, consider the following: Let P be a self-adjoint operator in \mathcal{H}_A that represents a physical quantity p in system A. The expectation value of $P \otimes Id_B$ in the separable

state $\Psi_{AB} = \psi_A \otimes \psi_B$ is equal to

$$\begin{aligned}\langle P \otimes Id_B \rangle &= (\psi_A \otimes \psi_B, (P \otimes Id_B) \psi_A \otimes \psi_B) \\ &= (\psi_A, P\psi_A) \cdot (\psi_B, \psi_B) = \langle P \rangle_A \cdot 1 = \langle P \rangle_A.\end{aligned}$$

Hence, the measurement results of any physical quantity in A are *independent* of B as long as AB is in a separable state. This is not the case in general, i.e. when AB is in an entangled state

$$\begin{aligned}\langle P \otimes Id_B \rangle &= (\Psi_{AB}, P \otimes Id_B \Psi_{AB}) \\ &= \sum_{ijIJ} (\Psi_{AB})_{iI}^* P_{ij} \delta_{IJ} (\Psi_{AB})_{jJ} \\ &= \sum_{ij} P_{ij} \left(\sum_I (\Psi_{AB})_{iI}^* (\Psi_{AB})_{jI} \right),\end{aligned}$$

$\langle P \otimes Id_B \rangle$ depends on the state of the total system. We can, nevertheless, define a quantity that encodes the probabilities for outcomes of measurements of physical quantities in system A. From the last line of the expression for $\langle P \otimes Id_B \rangle$, we define the self-adjoint operator

$$\rho \doteq \sum_{ij} \left(\sum_I (\Psi_{AB})_{iI}^* (\Psi_{AB})_{jI} \right) e_j e_i', \quad (3.1)$$

in \mathcal{H}_A . In the Physics literature it is referred to as **the density matrix** of system A. For the open system A (interacting with its environment B), ρ is the state variable analogous to the vector ψ for an isolated quantum system. In particular, it encodes all information about the system, that is the probabilities for outcomes of measurements of physical quantities. This can be seen from the last line of the

expression for $\langle P \otimes Id_B \rangle$ which can now be written as

$$\langle P \otimes Id_B \rangle = \sum_{ij} P_{ij} \rho_{ji} = \text{tr} P \rho. \quad (3.2)$$

Compare (3.2) with the expectation value for P , if A was an isolated quantum system:

$$\langle P \rangle = (\psi_A, P \psi_A) = \sum_{ij} \psi_i^* \psi_j P_{ij}.$$

In that case, we could still define the density matrix ρ by $\rho = \psi \psi'$ ($\rho_{ji} \doteq \psi_j \psi_i^*$) and every information about the quantum system (probabilities of measurement outcomes) would be encoded in it. But the state vector ψ encodes the same information more efficiently so there is no need to do so. In the case at hand (open quantum systems), the use of the density matrix is absolutely necessary because it encodes information about the composite system necessary for the calculation of probabilities of measurement outcomes in A . Thus ρ becomes the state variable for A and $\langle P \rangle = \text{tr} P \rho$ becomes the analog of the quantum-mechanical expectation formula for closed systems, namely $\langle P \rangle = (\psi, P \psi)$. A basis-independent expression for ρ is

$$\rho \doteq \text{tr}_B \Psi_{AB} \Psi'_{AB}, \quad (3.3)$$

where tr_B is a partial trace with respect to any basis of \mathcal{H}_B . (3.3) easily reduces to (3.1):

$$\rho = \text{tr}_B \Psi_{AB} \Psi'_{AB} = \sum_{ij} \left(\sum_I (\Psi_{AB})_{iI}^* (\Psi_{AB})_{jI} \right) e_j e'_i.$$

From its definition, we can easily infer the following properties of ρ :

1. ρ is self-adjoint.

2. ρ is positive semi-definite. Indeed, for any $v \in \mathcal{H}_A$,

$$(v, \rho v) = \sum_I \left(\sum_i (\Psi_{AB})_{iI} v_i \right)^* \left(\sum_j (\Psi_{AB})_{jI} v_j \right) \geq 0.$$

3. $\text{tr} \rho = 1$. From the definition,

$$\text{tr} \rho = \sum_{iI} (\Psi_{AB})_{iI}^* (\Psi_{AB})_{iI} = \|\Psi_{AB}\|^2 = 1.$$

From these properties it follows that ρ can be diagonalized and all its eigenvalues are real and non-negative and sum to one. Whenever ρ is of rank one, that is there exists a $\psi \in \mathcal{H}_A$ such that $\rho = \psi \psi'$ ($\rho_{ji} = \psi_j \psi_i^*$), the state of the system is called **pure**, otherwise it's called **mixed**. A pure state corresponds to a separable state of the composite system (A + environment) while a mixed one signifies the presence of entanglement. Mathematically, pure states are characterized by the property $\rho^2 = \rho$ while mixed ones satisfy $\rho^2 < \rho$. Any scalar function of ρ that is extremized by pure states only, e.g. $\text{tr} \rho^2$ or the von Neumann entropy $-k \text{tr}(\rho \ln \rho)$, is a measure of the *purity* of the system.

In the case of a two-state system, ρ is a 2×2 Hermitian, positive semi-definite matrix of trace 1. Hence, we need only 3 real numbers to specify it. A useful parametrization of it is the following:

$$\rho = \frac{1}{2}(I_2 + x \sigma_x + y \sigma_y + z \sigma_z) = \frac{1}{2} \begin{pmatrix} 1+z & x-iy \\ x+iy & 1-z \end{pmatrix}, \quad (3.4)$$

where x, y and z are real. Since ρ is 2×2 and has trace 1, the condition $\det \rho \geq 0$ is necessary and sufficient for it being non-negative. We can compute $\det \rho = \frac{1}{4}(1 - x^2 - y^2 - z^2)$ which leads to the condition $x^2 + y^2 + z^2 \leq 1$, i.e the state

space of an open two-state system is the 3-dimensional closed ball of radius one. In the Physics literature, it is often referred to as the **Bloch sphere** and the vector (x, y, z) is referred to as the Bloch vector. The states on the boundary $x^2 + y^2 + z^2 = 1$ have rank one (they have one eigenvalue zero and one equal to one) and thus they are pure. All states in the interior are mixed. For higher dimensional systems, the state space does not have such a simple description, but, due to the properties of ρ , is always convex.

We saw how the need for the introduction of the density matrix arises by considering the interaction of quantum systems with their environment. There is an alternative view of the density matrix formalism. ρ can always be expressed as

$$\rho = \sum_{i=1}^N p_i v_i v_i'. \quad (3.5)$$

The v_i 's need not be orthogonal, although, since ρ is Hermitian, there is always a representation like (3.5) with orthogonal v_i 's, namely it's eigenvectors. Also, the p_i 's must satisfy the relations

$$0 \leq p_i \leq 1, \quad i = 1, \dots, N, \quad \text{and} \quad \sum_{i=1}^N p_i = 1.$$

So ρ can be interpreted as the expectation value

$$\rho = \mathbb{E}[\psi\psi'], \quad (3.6)$$

in a *classical* ensemble of (identical) quantum systems. In this ensemble, the fraction of quantum systems in state v_i is equal to p_i . Equivalently, for one quantum system, there is a probability p_i to find the system in the quantum state v_i . Note that this is classical probability and not quantum probability, as

described in Chapter 2. We want to clarify this important point with an example: Consider the spin- $\frac{1}{2}$ system of Chapter 2. Assume it is in the state

$$\psi = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

The probability of measuring the z component of spin to be equal to $+\frac{1}{2}$ is equal to

$$\left| \begin{pmatrix} 1 \\ 0 \end{pmatrix}' \cdot \psi \right|^2 = \frac{1}{2},$$

and that of being equal to $-\frac{1}{2}$ is

$$\left| \begin{pmatrix} 0 \\ 1 \end{pmatrix}' \cdot \psi \right|^2 = \frac{1}{2},$$

again. Had we've been measuring the corresponding probabilities for the x component of spin, whose eigenvectors are

$$\frac{1}{\sqrt{2}} \begin{pmatrix} +1 \\ \pm 1 \end{pmatrix},$$

we would have gotten the answers 1, for $S_x = +\frac{1}{2}$ and 0 for $S_x = -\frac{1}{2}$. These are quantum-mechanical probabilities. The same answers would come out of the density matrix formalism for the pure state

$$\rho = \psi\psi' = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Consider now a classical ensemble where half the quantum systems have $S_z = +\frac{1}{2}$

and the other half $-\frac{1}{2}$. This state is described by the density matrix

$$\tilde{\rho} = \frac{1}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}' + \frac{1}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}' = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} \neq \rho.$$

Notice that $\tilde{\rho}$ can also be written as

$$\tilde{\rho} = \frac{1}{2} \left(\frac{1}{\sqrt{2}}\right)^2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}' + \frac{1}{2} \left(\frac{1}{\sqrt{2}}\right)^2 \begin{pmatrix} +1 \\ -1 \end{pmatrix} \begin{pmatrix} +1 \\ -1 \end{pmatrix}',$$

and hence it also represents an ensemble with half the systems in the $S_x = +\frac{1}{2}$ eigenstate and the other half in the $S_x = -\frac{1}{2}$ eigenstate, in sharp contrast with the corresponding probabilities computed from ψ above. In the following, both interpretations of ρ will be useful.

3.1.2 The Lindblad equation

So far, we discussed the necessity of the density matrix formalism for the description of open quantum systems along with its physical (statistical) interpretation. The next step is the introduction of an evolution equation for it (ρ is, after all, a state variable). Let Ψ_{AB0} be the initial state of the composite system and $\rho_0 = \text{tr}_B \Psi_{AB0} \Psi_{AB0}'$ the initial density matrix of A. Then, after time t , the state Ψ_{AB0} has evolved to

$$\Psi_{AB}(t) = e^{-iH_{AB}t} \Psi_{AB0},$$

hence,

$$\rho(t) = \text{tr}_B \Psi_{AB}(t) \Psi_{AB}(t)' = \text{tr}_B e^{-iH_{AB}t} \Psi_{AB0} \Psi_{AB0}' e^{iH_{AB}t}.$$

It is fairly obvious from this expression that $\rho(t)$ can not necessarily be simply expressed in terms of ρ_0 alone. From a physical perspective this is to be expected:

ρ contains only part of the information encoded in Ψ_{AB} (the part that is relevant for calculating probabilities of measurement outcomes in A) but to predict its evolution after time t , the evolution of *all* of the state Ψ_{AB} is necessary. Nevertheless, it turns out that one can write an evolution equation for it (the analog of Schrödinger's equation for open systems) under certain simplifying assumptions on the dynamics of system B (the environment). We will comment on these assumptions shortly, but first we introduce the **Lindblad equation** for ρ :

$$\dot{\rho} = -i[H, \rho] + \sum_n K_n \rho K_n' - \frac{1}{2} \left\{ \sum_n K_n' K_n, \rho \right\}. \quad (3.7)$$

We now explain the meaning of the various terms. In Quantum Mechanics, the notations $[S, T]$ and $\{S, T\}$ denote, respectively, the **commutator** and the **anti-commutator** of two operators S and T :

$$\begin{aligned} [S, T] &\doteq ST - TS, \text{ and} \\ \{S, T\} &\doteq ST + TS. \end{aligned}$$

The term $-i[H, \rho]$ represents the reversible (Hamiltonian) evolution of A and H is the Hamiltonian operator of A as an isolated system, usually with some corrections due to the dynamics of B (e.g. some renormalization of constants). The other terms express the effect of the dynamics of B on the evolution of A. As we will see, they represent various irreversible (dissipative) processes. H is self-adjoint while the K_n 's are not necessarily so. In the case of an infinite dimensional Hilbert space, we also assume that H and K_n are bounded and $\sum_n K_n' K_n$ converges strongly if $n \rightarrow \infty$.

(3.7) has some very interesting structure. Denote the (linear) map from the initial condition to the solution at time t as $\Lambda_t : \rho_0 \rightarrow \rho(t)$. Then:

1. Λ_t is a bounded, strongly continuous semigroup, that is

$$\Lambda_t(\Lambda_s(\rho_0)) = \Lambda_{t+s}(\rho_0) \text{ and } \lim_{t \rightarrow 0^+} \Lambda_t(\rho_0) = \rho_0.$$

2. Λ_t preserves the trace, $\text{tr}\Lambda_t(\rho_0) = \text{tr}\rho_0$.
3. Λ_t is positive: If $\rho_0 \geq 0$, then $\Lambda_t(\rho_0) \geq 0$.
4. Λ_t is **completely positive**.

Properties 1 to 3 are all reasonable for an evolution equation that preserves the properties of a density matrix. Property 4, which is a stronger form of property 3, requires some explanation. A positive linear map $\Lambda : \mathcal{T}(\mathcal{H}) \rightarrow \mathcal{T}(\mathcal{H})$ from trace-class operators in \mathcal{H} to trace-class operators in \mathcal{H} is called completely positive if its extension $\Lambda_n : \mathcal{T}(\mathcal{H}) \otimes \mathcal{M}_n \rightarrow \mathcal{T}(\mathcal{H}) \otimes \mathcal{M}_n$ (\mathcal{M}_n is the set of complex $n \times n$ matrices) defined by $\Lambda_n(\rho \otimes M) \doteq \Lambda_n(\rho) \otimes M$ is positive, for all $n \geq 1$. Physically, this requirement means the following: Given any other open quantum system A' , uncoupled to A , their joint state is described by the tensor product of their density matrices. If A evolves but A' doesn't, this evolution must preserve the properties of the density matrix of the total system, its positivity in particular. Complete positivity guarantees exactly that. Furthermore, it guarantees that the independent evolution of the density matrices of A and A' preserves the properties of the density matrix of the total system. Hence, it is an important property.

In the literature on open quantum systems there exist derivations of (3.7) beginning with the Schrödinger equation for the composite system, of varying mathematical rigor and with various simplifying assumptions on the dynamics of the environment. A good entry point to this literature is the recent review [67]. In general, two approximations are involved in these derivations, the *weak coupling limit* between the system and its environment and the *white noise ansatz*

for the dynamics of the environment. The first rigorous derivation of (3.7) was done by Davies in [68]. A fundamental result proved by Gorini, Kossakowski and Sudarshan [69] (who considered (3.7) for finite-dimensional systems) and Lindblad [70] (who considered the general case) is the following:

Theorem: [71] (3.7) is the only evolution equation that has all four properties listed above.

To reinforce the idea that all the properties of the Lindblad equation and in particular complete positivity are absolutely necessary for the evolution of a density matrix and to connect the Lindblad equation with the unitary dynamics of the composite system (system + environment), we quote the following

Theorem: [71] Let $\{\Lambda_t, t \geq 0\}$ be a bounded, norm-continuous, completely positive semigroup in the space $\mathcal{T}(\mathcal{H})$ of trace-class operators in \mathcal{H} . There exist a Hilbert space \mathcal{H}_B , a density matrix ρ_B in \mathcal{H}_B and a strongly continuous one-parameter group of unitary operators $\{U_t, t \geq 0\}$ in $\mathcal{H} \otimes \mathcal{H}_B$ such that

$$\Lambda_t(\rho) = \text{tr}_{\mathcal{H}_B} U_t(\rho \otimes \rho_B) U_t', \quad t \geq 0.$$

These results put the Lindblad equation on a rigorous basis. However, they give no information on the form of H and the K_n 's for specific systems. These have to be determined based on phenomenology and symmetry principles. In contrast, derivations based on models of the interaction with the environment provide forms for H and the K_n 's but the price to pay is either some mathematical rigor sacrificed or certain unphysical approximations/limiting procedures [67].

For the sake of building some intuition about the Lindblad equation and its

structure, we present a short “derivation” of it taken from [67], which is mathematically rigorous but whose starting point we don’t justify. Consider the following Stratonovich stochastic differential equation (SDE):

$$d\psi(t) = -iH\psi(t)dt - i \sum_n K_n \psi(t) \circ dB_n(t). \quad (3.8)$$

H is a self-adjoint operator while the K_n ’s are not necessarily so. The $B_n(t)$ ’s are independent normalized Brownian motions. Heuristically, one may say that (3.8) comes from a Schrödinger equation like

$$i \frac{d\psi(t)}{dt} = H\psi(t) + \sum_n K_n u_n(t) \psi(t),$$

where the functions $u_n(t)$ are “white noises”. Of course, this is just heuristics: K_n are not self-adjoint in general and the “white noise” effect of the environment on the system requires involved and technical approximations. Nevertheless, we will use (3.8) as the starting point for a derivation of the Lindblad equation (3.7). First, we turn (3.8) into an Itô SDE:

$$d\psi(t) = -iH\psi(t)dt - \frac{1}{2} \sum_n K_n' K_n \psi(t)dt - i \sum_n K_n \psi(t) dB_n(t). \quad (3.9)$$

Then, using the Itô rules,

$$\mathbb{E}[dB_n(t)] = 0, \text{ and } \mathbb{E}[dB_n(t) dB_m(t)] = \delta_{nm}dt,$$

we calculate

$$\begin{aligned} d\rho(t) &= d\mathbb{E}[\psi(t)\psi'(t)] = \mathbb{E}[d\psi(t)\psi'(t) + \psi(t)d\psi'(t) + d\psi(t)d\psi'(t)] \\ &= -i[H, \rho(t)]dt - \frac{1}{2}\left\{\sum_n K'_n K_n, \rho(t)\right\}dt + \sum_n K_n \rho(t) K'_n dt. \end{aligned}$$

This coincides with the Lindblad equation (3.7). We conclude this introduction with an example. The phenomenon of **spontaneous emission** in atoms and molecules, where a system transitions from an excited electronic state to the ground electronic state by emitting a photon with frequency equal to the Bohr frequency of the transition (see figure 3.1), is due to the interaction of these systems with the surrounding quantized modes of the electromagnetic field. We will write down the Lindblad equation that describes this phenomenon for a two-state system. Normalizing the energies of the ground and the excited state to 0 and ω , respectively, the Hamiltonian of the system is

$$H = \begin{pmatrix} 0 & 0 \\ 0 & \omega \end{pmatrix}.$$

If the rate of transition from the excited state to the ground state is equal to γ , then the corresponding Lindblad generator is $K_{12} = \sqrt{\gamma}E_{12}$, with

$$E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

(in \mathbb{C}^N the matrices E_{ij} are defined by $(E_{ij})_{kl} \doteq \delta_{ik}\delta_{jl}$, see appendix B of Chapter 2). The Lindblad equation reads

$$\dot{\rho} = -i[H, \rho] + K_{12}\rho K'_{12} - \frac{1}{2}\{K'_{12}K_{12}, \rho\}.$$

Representing $\rho(t)$ as

$$\rho(t) = \begin{pmatrix} p_1(t) & c^*(t) \\ c(t) & 1 - p_1(t) \end{pmatrix},$$

its solution is easily found to be

$$\begin{aligned} p_1(t) &= 1 - (1 - p_1(0)) e^{-\gamma t}, \\ c(t) &= e^{-(\frac{\gamma}{2} + i\omega)t} c(0), \end{aligned}$$

so that $\rho(t)$ is asymptotically approaching the state $\rho_{eq} = E_{11}$, where only the ground state is populated.

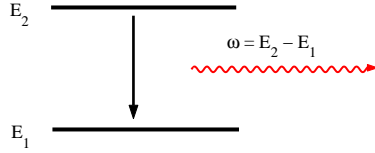


Figure 3.1: Spontaneous emission in a two-state system.

3.2 Laser cooling of molecules: The role of dissipation

Laser cooling of atoms and molecules is about reducing the kinetic energy of their motion. It is a very powerful tool for a number of applications. In the case of atoms, one only needs to reduce their translational kinetic energies. Laser cooling of molecules is much more complicated due to the fact that they have an internal structure. When cooling a sample of molecules, one must cool their translational motion as well as their rotation and internal vibration. In this work we concern ourselves with the laser cooling of the rotational motion of molecules.

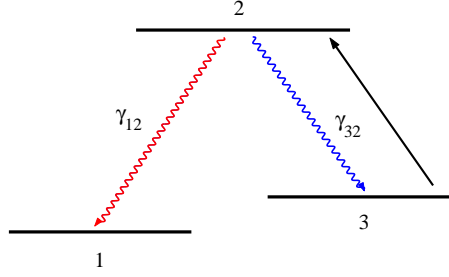


Figure 3.2: A three level Lambda system with spontaneous emissions from level 2 to 1 and 2 to 3. There is no direct emission from 3 to 1.

The essential mechanism and main obstruction are depicted in figure 3.2, which represents a, so-called, three-level Λ system. Molecules cool rotationally by spontaneous emission to the ground electronic-rotational state (state 1) from higher electronic-rotational states (state 2). The problem is that they can also end up at states other than the ground state (state 3) and that from some of these intermediate states, there is no direct spontaneous emission to the ground state. Thus, the molecules are trapped to higher rotational levels. One must then irradiate them with laser pulses to excite them back to states from where they can spontaneously emit and relax to the ground state.

In this work, we will ignore all other dynamics besides rotational and consider a finite-dimensional Hilbert space that contains the electronic-rotational levels involved in the cooling process. As in [72, 73], we assume that the only relaxation mechanisms in the system are spontaneous emissions and therefore we take $K_{ij} = \sqrt{\gamma_{ij}}E_{ij}$ and γ_{ij} is the rate of spontaneous emission from level j to level i . The dynamics of the system, will be given by a Lindblad equation with Hamiltonian control,

$$\dot{\rho} = -i[H_0 + Vu(t), \rho] + L(\rho), \quad (3.10)$$

where

$$L(\rho) = \sum_{ij} K_{ij} \rho K_{ij}^\dagger - \frac{1}{2} \{K_{ij}^\dagger K_{ij}, \rho\}. \quad (3.11)$$

In the three-level Λ system, state 2 spontaneously decays into the states 1 (ground state) and 3 at rates γ_{12} and γ_{32} , respectively, but there is no direct emission from 3 to 1. The corresponding Lindblad generators are $K_{12} = \sqrt{\gamma_{12}} E_{12}$ and $K_{32} = \sqrt{\gamma_{32}} E_{32}$.

Although the technique we'll develop in the following can handle general finite-dimensional quantum systems, we will restrict our analysis to the simple three-level Λ system which captures all the essential elements of the situation without the added complications of real molecular spectra. In [73], an analytic solution is obtained for the optimal cooling strategy in this system under certain simplifying assumptions to be discussed in the next subsection. We want to compare the results of our technique to this known solution.

There are two very important facts about the problem of rotational cooling:

1. The cooling effect is the transfer of an initial state ρ_0 , with $(\rho_0)_{ii} \neq 0$, for some $i > 1$, as close as possible to the ground state. An important feature of the dynamics (3.10) [72, 73] is that only the dissipative term can change the purity of ρ . Indeed, in the absence of the dissipative part, the dynamics would simply be

$$\dot{\rho}(t) = -i[H_0 + Vu(t), \rho(t)].$$

One can immediately establish the fact that $\text{tr} \rho(t)^n$ is a constant of motion, $\forall n \in \mathbb{N}^*$:

$$\frac{d}{dt} \text{tr} \rho(t)^n = n \text{tr}[\rho(t)^{n-1} \dot{\rho}] = -in \text{tr}(\rho(t)^{n-1} [H_0 + Vu(t), \rho(t)]) = 0,$$

from which follows the constancy of the spectrum of $\rho(t)$ and hence of its

purity. In the absence of the dissipative terms, the control $u(t)$ only performs unitary transformations on the state. This is the reason we refer to its as unitary (or Hamiltonian) control. Given an initial state ρ_0 , the most it can do is transfer the state to a diagonal form with the largest eigenvalue of ρ_0 at the $(1, 1)$ position. The cooling effect, which is the increase of ρ_{11} as close to 1 as possible, requires the cooperation of the unitary and the dissipative dynamics. In the case of the three-level Λ system, one can see this from the form of $L(\rho)$: For

$$\rho = \begin{pmatrix} \rho_{11} & \rho_{12} & \rho_{13} \\ \rho_{21} & \rho_{22} & \rho_{23} \\ \rho_{31} & \rho_{32} & \rho_{33} \end{pmatrix},$$

$L(\rho)$ has the form

$$L(\rho) = \begin{pmatrix} \gamma_{12}\rho_{22} & -\frac{1}{2}(\gamma_{12} + \gamma_{32})\rho_{12} & 0 \\ -\frac{1}{2}(\gamma_{12} + \gamma_{32})\rho_{21} & -(\gamma_{12} + \gamma_{32})\rho_{22} & -\frac{1}{2}(\gamma_{12} + \gamma_{32})\rho_{23} \\ 0 & -\frac{1}{2}(\gamma_{12} + \gamma_{32})\rho_{32} & \gamma_{32}\rho_{22} \end{pmatrix}.$$

Looking at the diagonal elements of $L(\rho)$, one can see the basic mechanism: Population from level 3 must be pumped by the laser to level 2 and then it is spontaneous emission from level 2 to level 1 that is responsible for the cooling. At the same time, emission from state 2 to state 3 works against cooling, hence the approach to $\rho = E_{11}$ can only be asymptotic.

2. The dissipation occurs on a much slower time-scale than the unitary (natural and induced) evolution of the system. Indeed, laser control in the time-scale of the free dynamics ($\sim 10^{-13}$ sec) is widely available while typical spontaneous emission times are in the order of a nanosecond.

3.3 Optimal control of dissipative dynamics

We now pose the optimal control problem: Determine the control $u(t)$ in (3.10) that maximizes $\rho_{11}(T)$ (for a desired final time T) given $\rho(0) = \rho_0$. In the following we briefly review two approaches to its solution, a “traditional” approach using a performance objective comprised of the above terminal quantity minus a quadratic penalty term for the control effort [30] and a recent approach pioneered in [72, 73].

In the traditional approach, one seeks to maximize the following objective:

$$J \doteq \rho_{11}(T) - \frac{a}{2} \int_0^T u^2(t) dt. \quad (3.12)$$

$a > 0$ is a weight factor on the size of the control effort which we’ll make very small but non-zero. The general theory of optimal control [63] leads to the Hamiltonian function

$$H(\rho, \lambda, u) = -\frac{a}{2}u^2 + \text{tr}[\lambda^\dagger (-i[H_0 + Vu, \rho] + L(\rho))],$$

where $\lambda = \lambda^\dagger$ is the co-state variable. The Hamiltonian equations satisfied by the local maxima are

$$\dot{\rho} = -i[H_0 + Vu, \rho] + L(\rho), \quad (3.13)$$

$$\dot{\lambda} = -i[H_0 + Vu, \lambda] - L^\dagger(\lambda), \quad (3.14)$$

$$u = -\frac{i}{a} \text{tr}(V[\rho, \lambda]). \quad (3.15)$$

L^\dagger is the adjoint operator of L (acting on the dual of the state space, i.e. on co-state variables). That is, for every Hermitian ρ and λ , $\text{tr}[L(\rho)\lambda] = \text{tr}[\rho L^\dagger(\lambda)]$.

For the three-level Λ system, the action of L^\dagger on

$$\lambda = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{pmatrix},$$

is given by

$$L^\dagger(\lambda) = \begin{pmatrix} 0 & -\frac{1}{2}(\gamma_{12} + \gamma_{32})\lambda_{12} & 0 \\ -\frac{1}{2}(\gamma_{12} + \gamma_{32})\lambda_{21} & \gamma_{12}(\lambda_{11} - \lambda_{22}) + \gamma_{32}(\lambda_{33} - \lambda_{22}) & -\frac{1}{2}(\gamma_{12} + \gamma_{32})\lambda_{23} \\ 0 & -\frac{1}{2}(\gamma_{12} + \gamma_{32})\lambda_{32} & 0 \end{pmatrix}.$$

Finally, as part of the necessary conditions of optimality, one has a terminal condition for λ (transversality condition):

$$\lambda(T) = E_{11}.$$

We insert the expression (3.15) for the optimal control in the state and co-state equations and bundle them with the boundary conditions on ρ and λ :

$$\dot{\rho} = [-iH_0 - \frac{1}{a} \text{tr}(V[\rho, \lambda]) V, \rho] + L(\rho), \quad (3.16)$$

$$\dot{\lambda} = [-iH_0 - \frac{1}{a} \text{tr}(V[\rho, \lambda]) V, \lambda] - L^\dagger(\lambda), \quad (3.17)$$

$$\rho(0) = \rho_0, \quad \lambda(T) = E_{11}. \quad (3.18)$$

We shall refer to equations (3.16) - (3.18) as **two-point boundary value problem III** (in short, TPBVP III). An important feature of this TPBVP is the separation of the boundary conditions for the state and co-state variables, which

makes it amenable to iterative numerical solution methods, of the type developed in [31, 32, 74, 62]. There is, however, a drawback in any attempt to a numerical solution of this TPBVP. As discussed before, cooling is a result of the cooperation between unitary control and dissipation, hence it takes place on the slower of the two time-scales, namely that of the dissipative dynamics. In that time scale, which is several orders of magnitude larger than that of the unitary evolution, the free dynamics terms in (3.16) and (3.17) generate many oscillations on top of the mean evolution of ρ and λ . This means that to achieve good accuracy, numerical solutions must be very detailed, thus increasing computational cost. This is because these numerical techniques do not exploit the separation of the two time-scales. Furthermore, the resulting locally optimal controls obtained by these numerical techniques don't reveal much structure.

In the recent work [73], a different approach is taken to the optimal control of cooling. Given the big separation of the dissipative and unitary time-scales, the authors introduce the following approximation: When the system without dissipation is controllable, i.e. H_0 and H_1 generate the Lie algebra of traceless anti-Hermitian matrices in N dimensions, $su(N)$ [49, 52], then any unitary evolution (similarity transformation) of ρ can be produced *instantaneously* on the dissipation time-scale. Following this, the state ρ is decomposed as $\rho = U\sigma U^\dagger$, with the diagonal σ which contains the spectrum of ρ remaining a state variable, while the unitary U is effectively turned into a control variable. Note that this control variable is not explicitly related with the original control $u(t)$, there may not even exist a $u(t)$ that can produce a desired evolution $U(t)$. This is a generalization of the idea in [75, 10, 11] that certain components (or functions) of the state can be controlled selectively and arbitrarily fast, so one can pass to an effective description of the system where these quantities are treated as control

variables. An evolution equation is derived for σ with U as a control variable and an optimal control problem is posed, namely to maximize $\|\rho\|_\infty$ (the largest eigenvalue of ρ , or equivalently, of σ) at a final time T . The optimal strategy of [73] for steering the density matrix is very simple: Given ρ , instantly “rotate” it by unitary evolution to a $\tilde{\rho}$ such that the rate of change of the objective achieves its maximum value on the orbit $\Phi(\rho) \doteq \{U\rho U^\dagger, U \in SU(N)\}$. The mechanism responsible for the change of the objective is, of course, dissipation.

The advantage of this formulation is that it provides a clear understanding of the roles of control and dissipation in the optimal cooling strategy. A drawback is that, even when an analytic solution exists as in the case of the three-level Λ system, there is no explicit connection between the physical control $u(t)$ and the control variables of the formulation of [73]. Put another way, given an optimal control strategy $U(t)$ it is not obvious how to generate from it a physical control $u(t)$. Moreover, for higher dimensional systems where analytical solutions may not exist, the numerical solution of the associated Hamilton-Jacobi-Bellman partial differential equation will be an extremely hard problem.

Our goal is to combine the best features of the two approaches and directly incorporate the time-scale separation of Hamiltonian and dissipative evolution into the “traditional” optimal control formulation. This addresses the drawbacks of both approaches: It allows us to explicitly calculate (locally) optimal controls $u(t)$ and avoid solving high-dimensional Hamilton-Jacobi-Bellman equations. At the same time, our method provides a structured form for the optimal control and state trajectory where the fast Hamiltonian evolution is clearly separated from the slow dynamics which is the result of the interplay of dissipation and the *average* effect of Hamiltonian control.

3.4 An efficient approximation based on averaging

3.4.1 Separating the two time-scales of the dynamics: An averaged system

In this section, we use the idea of separating the dynamics at the two time-scales. But instead of assuming that the effect of control is instantaneous in the slow (dissipation) time-scale, we take into account its average effect on the slow time-scale evolution. As mentioned before, all dissipation constants in the system can be considered extremely small compared to its Bohr frequencies. We introduce a common scale ε for all these constants γ_{ij} , and define $\tilde{\gamma}_{ij}$'s such that $\gamma_{ij} \doteq \varepsilon \tilde{\gamma}_{ij}$. ε should be defined as $\varepsilon = \frac{\max \gamma_{ij}}{\min \omega_{ij}}$ ($\omega_{ij} > 0$) or any number of the same order of magnitude. We also define $L \doteq \varepsilon \tilde{L}$. It makes sense that the cooling effect which is a result of the dissipation takes place on a time-scale of order at least $\frac{1}{\varepsilon}$. Thus, $S \doteq \varepsilon T$ should be at least of order 1, which is what we assume from now on.

Consider now equation (3.10):

$$\dot{\rho} = -i[H_0 + Vu(t), \rho] + \varepsilon \tilde{L}(\rho).$$

We introduce the change of variables $\rho \doteq e^{-iH_0t} r e^{iH_0t}$ (transformation to the interaction picture). With the definition

$$F(t) \doteq e^{iH_0t} V e^{-iH_0t},$$

it transforms to

$$\dot{r} = -i[F(t), r] u(t) + \varepsilon e^{iH_0 t} \tilde{L}(e^{-iH_0 t} r e^{iH_0 t}) e^{-iH_0 t}.$$

Note the appearance of the Bohr frequencies in the matrix elements of F ,

$$F_{ij}(t) = V_{ij} e^{i(E_i - E_j)t} = V_{ij} e^{i\omega_{ij}t}.$$

For the form of $L(\rho)$ we are using, it is easy to see that

$$\begin{aligned} [L(\rho)]_{kk} &= \sum_i \gamma_{ki} \rho_{ii} - \sum_{ik} \gamma_{ik} \rho_{kk}, \text{ and} \\ [L(\rho)]_{kl} &= -\frac{1}{2} \sum_i (\gamma_{ik} + \gamma_{il}) \rho_{kl}, \quad k \neq l, \end{aligned}$$

from which follows that

$$e^{iH_0 t} L(e^{-iH_0 t} r e^{iH_0 t}) e^{-iH_0 t} = L(r).$$

We thus end up with the transformed system

$$\dot{r} = -i[F(t), r] u(t) + \varepsilon \tilde{L}(r). \tag{3.19}$$

The evolution of the interaction picture density matrix r is entirely due to the control and the dissipation term. Equation (3.19) suggests the following fact: For small control amplitude, if the control is to have any significant effect on the dynamics over a long time, $u(t)$ must contain frequency components whose frequency equals the Bohr frequencies of the system. This is an intuition based on averaging theory. The discussion in section 3.4.3 will make this intuition and the proposed form of the control (3.20) more transparent.

Following this intuition, we introduce the following form for the control:

$$u(t) = \varepsilon \sum_{\omega_{ij} \neq 0} e^{i\omega_{ij}t} u_{ji}(\varepsilon t), \quad (3.20)$$

where the summation is over pairs of indices (i, j) with $\omega_{ij} \neq 0$ (if there is degeneracy in the spectrum, as in the case of rotational dynamics, $\omega_{ij} = 0$ even for $i \neq j$). u_{ji} is a complex “envelope” and $u_{ji}^* = u_{ij}$ so that u is real. By definition, $u_{ij} \doteq 0$ whenever $\omega_{ij} = 0$. We will also use the notation ω_α , with α being an index pair (i, j) , $i > j$, such that i and j are the smallest indices in the set of index pairs (m, n) such that $\omega_{mn} = \omega_{ij}$. Essentially, α runs over the set of independent positive Bohr frequencies of the system. For example, consider a four-state system with $E_4 > E_3 > E_2 > E_1$, such that $\omega_{21} = \omega_{43} \neq \omega_{32}$. Then, $\omega_{42} = \omega_{43} + \omega_{32} = \omega_{32} + \omega_{21} = \omega_{31}$ as well. Hence, the only independent positive Bohr frequencies are ω_{21} , ω_{32} , ω_{31} and ω_{41} and so α can only take the values $(2, 1)$, $(3, 2)$, $(3, 1)$ and $(4, 1)$. Using this convention, we can rewrite (3.20) as

$$u(t) = \varepsilon \sum_{\omega_{ij} \neq 0} e^{i\omega_{ij}t} u_{ji}(\varepsilon t) = \varepsilon \left(\sum_{\alpha} e^{-i\omega_{\alpha}t} w_{\alpha}(\varepsilon t) + e^{i\omega_{\alpha}t} w_{\alpha}^*(\varepsilon t) \right), \quad (3.21)$$

where

$$w_{kl} \doteq \begin{cases} \sum_{\omega_{ij} = \omega_{kl} \neq 0} u_{ij}, & \omega_{kl} \neq 0, \\ 0, & \omega_{kl} = 0. \end{cases}$$

Introduce this form for $u(t)$ in (3.19) and rewrite that equation in component

form:

$$\begin{aligned}\dot{r}_{kl} &= -i\varepsilon \sum_m (V_{km} e^{i\omega_{km}t} r_{ml} - r_{km} V_{ml} e^{i\omega_{ml}t}) \cdot \sum_{i,j} e^{i\omega_{ij}t} u_{ji}(\varepsilon t) \\ &+ \varepsilon [\tilde{L}(r)]_{kl}.\end{aligned}\tag{3.22}$$

We approximate (3.22) for small ε using averaging, as in chapter 2:

$$\dot{\bar{r}}_{kl} = -i\varepsilon \sum_m (V_{km} w_{km}(\varepsilon t) \bar{r}_{ml} - \bar{r}_{km} V_{ml} w_{ml}(\varepsilon t)) + \varepsilon [\tilde{L}(\bar{r})]_{kl}.\tag{3.23}$$

With the definition $(\tilde{V}[w])_{mn} \doteq V_{mn} w_{mn}$, we can write (3.23) in matrix form as

$$\dot{\bar{r}} = \varepsilon (-i [\tilde{V}[w], \bar{r}] + \tilde{L}(\bar{r})).\tag{3.24}$$

Finally, we rescale time to $s \doteq \varepsilon t$ in (3.24):

$$\frac{d\bar{r}}{ds} = -i[\tilde{V}[w], \bar{r}] + \tilde{L}(\bar{r}), \quad s \in [0, \varepsilon T] = [0, S].\tag{3.25}$$

According to the averaging theorem in [57] and the preceding arguments, any control strategy $\{w_\alpha(s)\}$, $s \in [0, S]$, for (3.25), results in a control strategy $u(t)$, $t \in [0, T]$, for (3.10) according to (3.21), such that the state trajectory $\rho(t)$ is $O(\varepsilon)$ -close to $e^{-iH_0t} \bar{r}(\varepsilon t) e^{iH_0t}$, if $\rho(0) = \bar{r}(0)$.

3.4.2 Optimal transfers in the averaged system

We now consider the following optimal transfer problem for the averaged system

(3.25): Maximize

$$\bar{J} \doteq \bar{r}_{11}(S) - a_0 \int_0^S \sum_\alpha |w_\alpha(s)|^2 ds,\tag{3.26}$$

where, as above, $S = \varepsilon T$. The necessary conditions for optimality are derived from the Hamiltonian function [63]

$$H(\bar{r}, \bar{l}, w_\alpha) = -a_0 \sum_{\alpha} |w_\alpha|^2 - i \operatorname{tr} \bar{l} ([\tilde{V}[w], \bar{r}] + \tilde{L}(\bar{r})),$$

and are the following:

$$\frac{d\bar{r}}{ds} = -i [\tilde{V}[w], \bar{r}] + \tilde{L}(\bar{r}), \quad (3.27)$$

$$\frac{d\bar{l}}{ds} = -i [\tilde{V}[w], \bar{l}] - \tilde{L}^\dagger(\bar{l}), \quad (3.28)$$

$$w_\alpha = -\frac{i}{a_0} \sum_{\omega_{ij}=\omega_\alpha} V_{ji} [\bar{r}, \bar{l}]_{ij}, \quad (3.29)$$

along with boundary conditions identical to (3.18),

$$\bar{r}(0) = \rho(0) \text{ and } \bar{l}(S) = E_{11}.$$

We shall refer to this TPBVP as **two-point boundary value problem IV**. The point to be made is that, it is computationally much easier to solve than TPBVP III: The dynamics in the short time-scale (both free and due to fast control action) has been removed and only the dissipative dynamics and the averaged control action on the long time-scale (i.e. of order $\frac{1}{\varepsilon}$) remain. Thus the solution of problem IV requires much less computational burden and accuracy retained. Because of the separated boundary conditions for the state and co-state at the two ends, iterative procedures that use successive forward (in time) integrations of the state equation and backward integrations of the co-state equation can be applied for the solution of TPBVP IV. An example of such an algorithm is the following (see

also [31, 32, 74, 62]):

$$\begin{aligned}
\frac{d\bar{r}^{(k+1)}}{ds} &= -i [\tilde{V}[w^{(k+1)}], \bar{r}^{(k+1)}] + \tilde{L}(\bar{r}^{(k+1)}), \\
w_{ij}^{(k+1)} &= \frac{d - \frac{a_0}{2}}{d + \frac{a_0}{2}} w_{ij}^{(k)} - \frac{i}{d + \frac{a_0}{2}} \sum_{\omega_{mn}=\omega_{ij}} V_{nm} [\bar{r}^{(k+1)}, \bar{l}^{(k)}]_{mn}, \\
\bar{r}^{(k+1)}(0) &= \rho_0, \\
\frac{d\bar{l}^{(k+1)}}{ds} &= -i [\tilde{V}[w^{(k+1)}], \bar{l}^{(k+1)}] - \tilde{L}^\dagger(\bar{l}^{(k+1)}), \\
\bar{l}^{(k+1)}(S) &= E_{11},
\end{aligned}$$

with $d > 0$ and $a_0 \geq 0$. It can be shown that the sequence of objective functionals

$$\bar{J}^{(k)} = \bar{r}_{11}^{(k)}(S) - a_0 \int_0^S \sum_{\alpha} |w_{\alpha}^{(k)}(s)|^2 ds$$

increases monotonically. We note in passing that this algorithm works even when $a_0 = 0$.

Finally, we show in the next section that the resulting control law $u(t)$ for the original system that we get from the optimal envelopes $w_{ij}(s)$ for the averaged system from (3.21), is approximately optimal with respect to J of (3.12), up to terms of $O(\varepsilon)$, for $a = \frac{a_0}{\varepsilon}$.

3.4.3 Approximate optimality of the resulting control $u(t)$

We begin with the change of variables in equations (3.16) - (3.18):

$$\begin{aligned}
\rho &= e^{-iH_0 t} r e^{iH_0 t}, \\
\lambda &= e^{-iH_0 t} l e^{iH_0 t}.
\end{aligned}$$

With the definition $F(t) \doteq e^{iH_0 t} V e^{-iH_0 t}$, they become

$$\dot{r} = -\frac{1}{a} \text{tr}(F(t)[r, l]) [F(t), r] + L(r), \quad (3.30)$$

$$\dot{l} = -\frac{1}{a} \text{tr}(F(t)[r, l]) [F(t), l] - L^\dagger(l), \quad (3.31)$$

$$r(0) = \rho_0, \quad l(T) = E_{11}, \quad (3.32)$$

where the following facts were used:

$$e^{iH_0 t} L(e^{-iH_0 t} r e^{iH_0 t}) e^{-iH_0 t} = L(r),$$

$$e^{iH_0 t} L^\dagger(e^{-iH_0 t} l e^{iH_0 t}) e^{-iH_0 t} = L^\dagger(l).$$

The transformation property of L has already been established and that of L^\dagger follows immediately from its definition. We now take $a = \frac{a_0}{\varepsilon}$ in (3.12). Notice that the choice $a = \frac{a_0}{\varepsilon}$ does not affect the generality of J because the parameter a_0 is chosen independently of ε . Letting $L = \varepsilon \tilde{L}$ and $L^\dagger = \varepsilon \tilde{L}^\dagger$, equations (3.30) and (3.31) become

$$\dot{r} = \varepsilon \left(-\frac{1}{a_0} \text{tr}(F(t)[r, l]) [F(t), r] + \tilde{L}(r) \right),$$

$$\dot{l} = \varepsilon \left(-\frac{1}{a_0} \text{tr}(F(t)[r, l]) [F(t), l] - \tilde{L}^\dagger(l) \right).$$

The final step is to average these equations. To make the procedure a little more transparent, we rewrite the state equation in component form:

$$\begin{aligned} \dot{r}_{mn} &= \varepsilon \left(-\frac{1}{a_0} \sum_{i,j} V_{ji} e^{i\omega_{ji}t} [r, l]_{ij} \cdot \sum_p (V_{mp} e^{i\omega_{mp}t} r_{pn} - V_{pn} e^{i\omega_{pn}t} r_{mp}) \right. \\ &\quad \left. + [\tilde{L}(r)]_{mn} \right). \end{aligned}$$

The averaged form of this equation is

$$\begin{aligned}
\dot{\bar{r}}_{mn} &= \varepsilon \left(\sum_p \underbrace{\left(-\frac{1}{a_0} \sum_{\omega_{ij}=\omega_{mp}} V_{ji} [\bar{r}, \bar{l}]_{ij} \right)}_{-i w_{mp}} V_{mp} \bar{r}_{pn} \right. \\
&\quad \left. - \sum_p \underbrace{\left(-\frac{1}{a_0} \sum_{\omega_{ij}=\omega_{pn}} V_{ji} [\bar{r}, \bar{l}]_{ij} \right)}_{-i w_{pn}} V_{pn} \bar{r}_{mp} + [\tilde{L}(\bar{r})]_{mn} \right) \\
&= \varepsilon \left(-i \sum_p w_{mp} V_{mp} \bar{r}_{pn} - w_{pn} V_{pn} \bar{r}_{mp} + [\tilde{L}(\bar{r})]_{mn} \right),
\end{aligned}$$

or, in compact form,

$$\dot{\bar{r}} = -i\varepsilon ([\tilde{V}[w], \bar{r}] + \tilde{L}(\bar{r})).$$

This is precisely equation (3.27) without the time rescaling, with the envelopes given by (3.29). An identical calculation establishes that (3.28) (without the time rescaling) is the averaged form of (3.31).

The aforementioned averaging theorem and the sequence of variable changes that we introduced guarantees that, for sufficiently small ε , any solution $(\rho(t), \lambda(t))$ of (3.16) and (3.17) with some initial conditions ρ_0 and λ_0 , can be expressed in terms of the solution $(\bar{r}(s), \bar{l}(s))$ of (3.27) and (3.28) with the same *initial* conditions as

$$\rho(t) = e^{-iH_0 t} \bar{r}(\varepsilon t) e^{iH_0 t} + O(\varepsilon), \quad (3.33)$$

$$\lambda(t) = e^{-iH_0 t} \bar{l}(\varepsilon t) e^{iH_0 t} + O(\varepsilon). \quad (3.34)$$

Ideally, we would like to prove a relation identical to (3.33) and (3.34) between the solutions of TPBVPs III and IV with the same *boundary* conditions (3.18). That is, prove that there is a one-to-one correspondence between the solutions of TPBVPs III and IV with the same boundary conditions (3.18) given by (3.33) and

(3.34). We established such a relation in Chapter 2 between TPBVPs I and II for exact population transfers in a closed quantum system with a quadratic penalty on the control. Here, we take a simpler approach: Given a solution $(\bar{r}(s), \bar{l}(s))$ of TPBVP IV, we define $\rho^{apr}(t)$ and $\lambda^{apr}(t)$ by

$$\rho^{apr}(t) \doteq e^{-iH_0 t} \bar{r}(\varepsilon t) e^{iH_0 t}, \quad (3.35)$$

$$\lambda^{apr}(t) \doteq e^{-iH_0 t} \bar{l}(\varepsilon t) e^{iH_0 t}. \quad (3.36)$$

Then, $(\rho^{apr}(t), \lambda^{apr}(t))$ will satisfy the necessary conditions of optimality (3.16)-(3.18), that is TPBVP III, up to terms of order ε . In this sense, $(\rho^{apr}(t), \lambda^{apr}(t))$ furnish *approximate solutions* to the optimal transfer problem of section (3.3). There is a fine distinction to be made here, namely that this is not the same as $(\rho^{apr}(t), \lambda^{apr}(t))$ providing *approximations to solutions* of TPBVP III. Put another way, for any finite ε , no matter how small, there is no guarantee that the approximate solution $(\rho^{apr}(t), \lambda^{apr}(t))$ to TPBVP III defined through (3.35) and (3.36) is an approximation to an actual solution of it, i.e. ε -close to a local maximum of J . Based on the results of the numerical solution of the optimal cooling strategy for the 3-state Λ system (see section 3.5), we believe this to be true. From a practical perspective, what is important is that $\bar{r}_{11}(S)$ achieves a value close to 1 because then, the corresponding approximate solution of TPBVP III defined by (3.35) and (3.36) will furnish a control that achieves the same yield to $O(\varepsilon)$ for the original system.

3.5 Example: Numerical solution of the three-level Λ system

We now present the numerical solution of the optimal cooling problem for the three-level Λ system. We take $\tilde{\gamma}_1 = 1$, $\tilde{\gamma}_2 = .8$ and, for simplicity,

$$V = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

We used the iterative technique outlined at the end of section 3.4.2, starting with the relatively large value 1 for a_0 and gradually decreasing it to 0. Intuitively, one expects that the control will pump population from level 3 to level 2 and then let spontaneous emission do the rest, i.e. dump it on level 1. This intuition is verified, as the profile w_{12} turns out to be zero. Only the real part of the profile w_{23} is nonzero and is represented in figure 3.3. Hence, the optimal control will have the form

$$u(t) = 2\varepsilon \cos \omega_{23} t w_{23}(\varepsilon t) + O(\varepsilon^2).$$

The value of ε should be of the order of magnitude of $\frac{\max\{\gamma_{12}, \gamma_{32}\}}{\omega_{23}}$.

To compare with the results of [73], we plot the average populations \bar{r}_{ii} we calculated numerically with the analytically calculated populations σ_i of [73] in figure 3.4. The agreement between \bar{r}_{11} and σ_1 is excellent. As expected, σ_1 is always larger than \bar{r}_{11} , because the approximation of [73] assumes *instantaneous* unitary control of the quantum system and thus *over-estimates* the response of the system. Still, at the final time, hardly any difference is discernable. The agreement between \bar{r}_{22} and σ_2 and between \bar{r}_{33} and σ_3 is also very good: The populations of levels 2 and 3 become equal after some time (close to the time

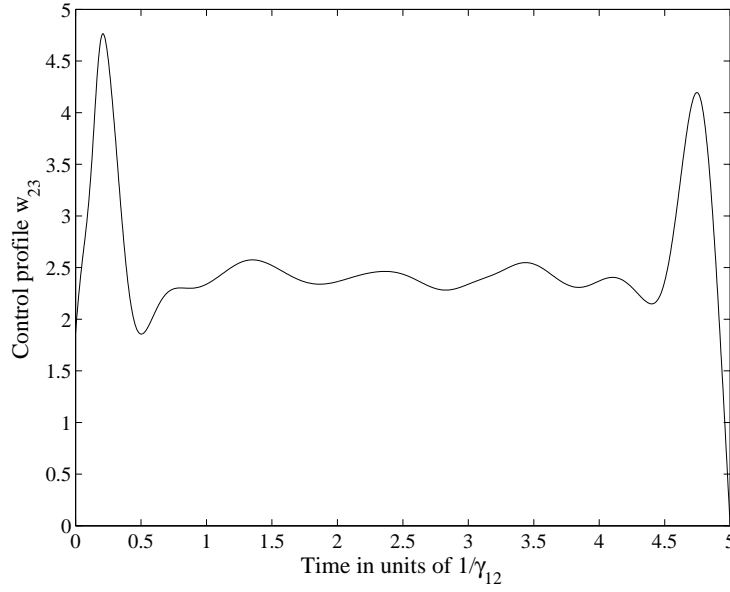


Figure 3.3: Optimal profile w_{23}

predicted in [73]) and stay almost equal for all times after that, as predicted by the analytical solution of [73]. Again, the discrepancy (mostly in the initial stage of the transfer) stems from the assumption of instantaneous unitary control of [73]. And again, the agreement becomes almost perfect close to the final time. Finally, we look at the *coherences* ρ_{12} , ρ_{13} and ρ_{23} . [73] predicts that they are all zero for the optimal trajectory. Our solution finds that \bar{r}_{12} and \bar{r}_{13} are indeed zero but \bar{r}_{23} is non-zero, though small. Once again, the reason for the difference is the fact that the approximation of [73] *over-estimates* the response of the system. We plot $|\bar{r}_{23}|$ in figure 3.5.

We see that the numerical solution of the optimal cooling problem for the three-level Λ system reproduces the features of the solution in [73] with some small but expected deviations.

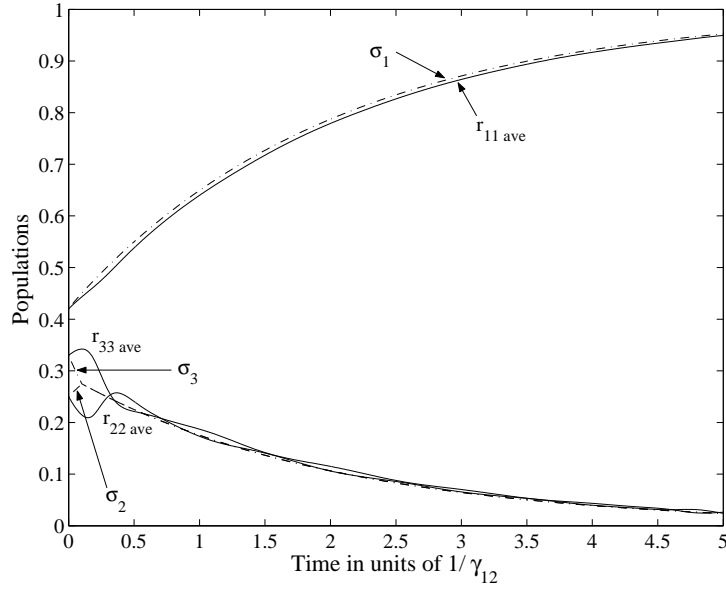


Figure 3.4: Average populations calculated by our method and optimal populations calculated in Sklarz et al.

3.6 Summary

We set up an optimal molecular-rotation cooling problem as an optimal transfer problem for the Lindblad equation with Hamiltonian control. Using the facts that the Hamiltonian natural dynamics is much faster than the dissipative and that it is the latter that controls the time-scale of the cooling process, we introduced an approximation method analogous to that of Chapter 2. We showed that, instead of the original transfer problem, one may consider an averaged version of it, which is computationally much easier to solve since it contains only one time-scale, the slow time-scale of the dissipation. Once solved, it provides an approximate solution to the original transfer problem which exhibits the time-scale separation. Unfortunately, we were not able to obtain as strong results as in Chapter 2 (show the one-to-one correspondence between solutions of the original and the averaged problem) because of the nature of the TPBVP. We did demonstrate, however, in

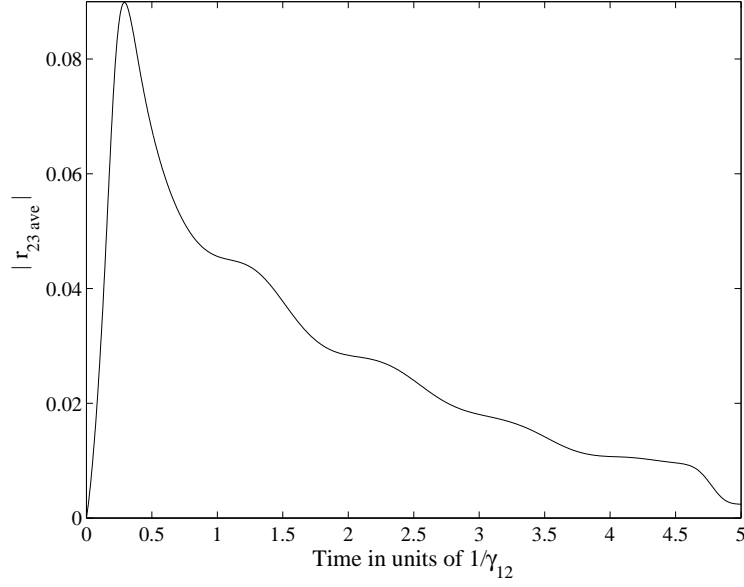


Figure 3.5: Average coherence \bar{r}_{23}

the context of an example, the conceptual and computational simplification that our technique offers. We compared our results for the example of a three-level Λ system with the solution obtained by [73] using a different approximation (that also uses the time-scale separation of the different dynamics in the system) and found excellent agreement. But our technique is able to obtain the form of the (approximate) optimal control and is readily generalizable to higher dimensional systems in contrast to that of [73].

Chapter 4

Lyapunov based control of quantum systems

In this chapter we present some work on Lyapunov-based control of isolated quantum systems. We derive continuous feedback laws that asymptotically stabilize desired energy levels. The obtained feedback laws are to be used to derive *open loop* controls. The inclusion of (continuous) measurement and feedback requires much more complicated models with stochastic dynamics as we'll see in the next chapter. We should point out right away that the best one can do with these Lyapunov-based designs for isolated quantum systems is to create *almost globally stabilizing* feedback laws. This is because of the topology of the phase space of quantum systems: They are (essentially) multidimensional spheres and there does not exist a globally stabilizing continuous vector field on a finite-dimensional sphere. Other works on Lyapunov-based control of isolated quantum systems include [76, 77]. In these works, as in ours, the control Lyapunov functions are conserved quantities of the uncontrolled Hamiltonian dynamics. This idea was proposed in [78] and is a generalization of the “energy” control of [79].

In the following, we consider only finite-dimensional quantum systems with

one control as in (2.10):

$$i\dot{\psi} = (H_0 + Vu(t))\psi.$$

For an N -dimensional quantum system, the state space is $\mathbb{C}P^N$, hence only Lyapunov functions with the property $\Phi(e^{i\phi}\psi) = \Phi(\psi)$, $\phi \in S^1$, are allowed. We use $\Phi(\psi) = \psi'P\psi$, $P' = P$, as a control Lyapunov function. Our first task is to analyze the structure of its critical points on $\mathbb{C}P^N$.

Lemma: The critical points of $\Phi(\psi) = \psi'P\psi$ on $\mathbb{C}P^N$ are the normalized eigenvectors of P . The eigenvector with the largest eigenvalue is the global maximum, the eigenvector with the smallest eigenvalue is the global minimum and all others are saddle points. \square

Proof: We only need to worry about the constraint $\psi'\psi = 1$, all other equations/statements are appropriately phase-invariant or covariant. We consider the “augmented” function

$$\tilde{\Phi}(\psi, \sigma) = \psi'P\psi - \sigma(\psi'\psi - 1),$$

where σ is a Lagrange multiplier. Varying $\tilde{\Phi}$ with respect to ψ we have,

$$0 = \left(\frac{\partial \tilde{\Phi}}{\partial \psi} \right)' = P\psi - \sigma\psi,$$

so, ψ is an eigenvector of P and σ the corresponding eigenvalue. Varying $\tilde{\Phi}$ with respect to σ enforces the normalization of ψ . Denote the normalized eigenvectors

of P by $q_i, i = 1, \dots, N$, and the corresponding eigenvalues by p_i . Then,

$$P = \sum_{i=1}^N p_i q_i q_i'.$$

To determine the structure of Φ around one of its critical points, q_N for example, consider a finite variation $\Delta\psi$ such that $\|q_N + \Delta\psi\|^2 = 1$. Express $\Delta\psi$ in the basis of the eigenvectors of P ,

$$\Delta\psi = \sum_{i=1}^N \Delta\psi_i q_i.$$

The normalization condition $\|q_N + \Delta\psi\|^2 = 1$ implies that

$$(\Delta\psi_N + \Delta\psi_N^* + \Delta\psi_N \Delta\psi_N^*) + \sum_{i=1}^{N-1} \Delta\psi_i \Delta\psi_i^* = 0,$$

so, not all components of the variation can be chosen arbitrarily. Consider the difference

$$\begin{aligned} \Phi(q_N + \Delta\psi) - \Phi(q_N) &= p_N(\Delta\psi_N + \Delta\psi_N^* + \Delta\psi_N \Delta\psi_N^*) + \sum_{i=1}^{N-1} p_i \Delta\psi_i \Delta\psi_i^* \\ &= \sum_{i=1}^{N-1} (p_i - p_N) \Delta\psi_i \Delta\psi_i^*. \end{aligned}$$

Taking the $\Delta\psi_i, i = 1, \dots, N-1$, as independent parameters of the variation, we see that the structure of Φ around q_N depends on the ordering of the eigenvalues: q_N is the global maximum iff p_N is the largest eigenvalue, the global minimum iff p_N is the smallest eigenvalue and a saddle point otherwise. ■

We calculate the rate of change of Φ along a trajectory of (2.10):

$$\begin{aligned}
\frac{d\Phi}{dt} &= -i\psi'P(H_0 + Vu)\psi + i\psi'(H_0 + Vu)P\psi \\
&= i\psi'[H_0 + Vu, P]\psi \\
&= i\psi'[H_0, P]\psi + i\psi'[V, P]\psi u.
\end{aligned}$$

For any matrices A and B , $\text{tr}([A, B]) = \text{tr}(AB) - \text{tr}(BA) = 0$, hence a commutator can never be sign-definite. For this reason, we choose P to commute with H_0 , $[H_0, P] = 0$. This implies that the two matrices must have the same eigenvectors and hence, by the lemma, the critical points of Φ are the eigenvectors of H_0 . Now,

$$\frac{d\Phi}{dt} = i\psi'[V, P]\psi u.$$

To make $\dot{\Phi} \leq 0$, we choose the following (state feedback) control law:

$$u = -f(i\psi'[V, P]\psi), \quad (4.1)$$

with $f : \mathbb{R} \rightarrow \mathbb{R}$, any continuous function such that $xf(x) > 0$, $\forall x \neq 0$. With this choice of u , the controlled Schrödinger's equation becomes

$$i\dot{\psi} = (H_0 - f(i\psi'[V, P]\psi) V) \psi. \quad (4.2)$$

We show now that the set of eigenvectors of P is the largest invariant set of (4.2). For $\bar{\psi}$ to be in the invariant set of (4.2), $\dot{\Phi} = -xf(x) = 0$, with $x = i\bar{\psi}'[V, P]\bar{\psi}$, which means that

$$i\bar{\psi}'(t) [V, P] \bar{\psi}(t) = 0, \quad (4.3)$$

must be satisfied along with

$$i\dot{\bar{\psi}}(t) = (H_0 - f(i\bar{\psi}'(t)[V, P]\bar{\psi}(t))V)\bar{\psi}(t),$$

with initial condition $\bar{\psi}$. The solution is given by

$$\bar{\psi}(t) = e^{-iH_0 t} \bar{\psi} = \sum_{i=1}^N c_i e^{-iE_i t} q_i,$$

where $\bar{\psi} = \sum_{i=1}^N c_i q_i$ and the E_i 's are the eigenvalues of H_0 . Substituting this in (4.3), we obtain the condition

$$-i \sum_{i,j}^N (p_i - p_j) c_i c_j^* e^{-i\omega_{ji}t} (q_j' V q_i) = 0. \quad (4.4)$$

We make the following assumptions:

1. $\omega_{ji} \neq \omega_{lk}, \quad (i, j) \neq (k, l).$
2. $q_j' V q_i \neq 0, \quad i \neq j.$
3. $p_i \neq p_j, \quad i \neq j.$

Assumption 1 guarantees that the exponential functions in (4.4) are linearly independent, while 2 and 3 ensure that the resulting equations,

$$c_i c_j^* = 0,$$

hold $\forall i \neq j$. The only non-trivial solutions of these have only one c_i non-zero. This means that $\bar{\psi}$ is one of the q_i 's. Hence, the largest invariant set of (4.2) is the set of eigenvectors of P . La Salle's invariance principle [57] guarantees its asymptotic stability. This analysis can be generalized in two ways:

1. One may consider a P with multiple eigenvalues. Suppose that P has only $M < N$ distinct eigenvalues. Group the eigenvectors of P in M corresponding eigenspaces. One can show that the conclusions above continue to hold if we substitute ‘eigenspaces’ for ‘eigenvectors’. One still needs assumptions 1-3, but the indices need only belong to levels in *different eigenspaces* of P .
2. For systems with more controls, instead of assumption 2 one needs the weaker condition

$$q_j' V_\alpha q_i \neq 0, \quad i \neq j,$$

for *at least one* V_α .

We summarize our results in the following theorem:

Theorem: Consider system (2.10) with the feedback control law (4.1), where H_0 and P commute. Given the (generalized) assumptions 1-3, this control law asymptotically drives any trajectory of (2.10) to an eigenspace of P . \square

In general, Φ is just a *local* Lyapunov function for the eigenspace corresponding to the minimum element of P . In the special case $p_1 = \dots = p_{N-1} > p_N$, Φ guarantees semi-global convergence to q_N (only points from the subspace spanned by q_1, \dots, q_{N-1} do not converge to q_N but remain in that subspace). Figure 4.1 demonstrates a transfer of population from the ground vibrational state of the OH-bond to its 9th excited state using a Lyapunov based feedback law. The approach to the target state is asymptotic but a good yield (90%) is achieved in a reasonable transfer time.

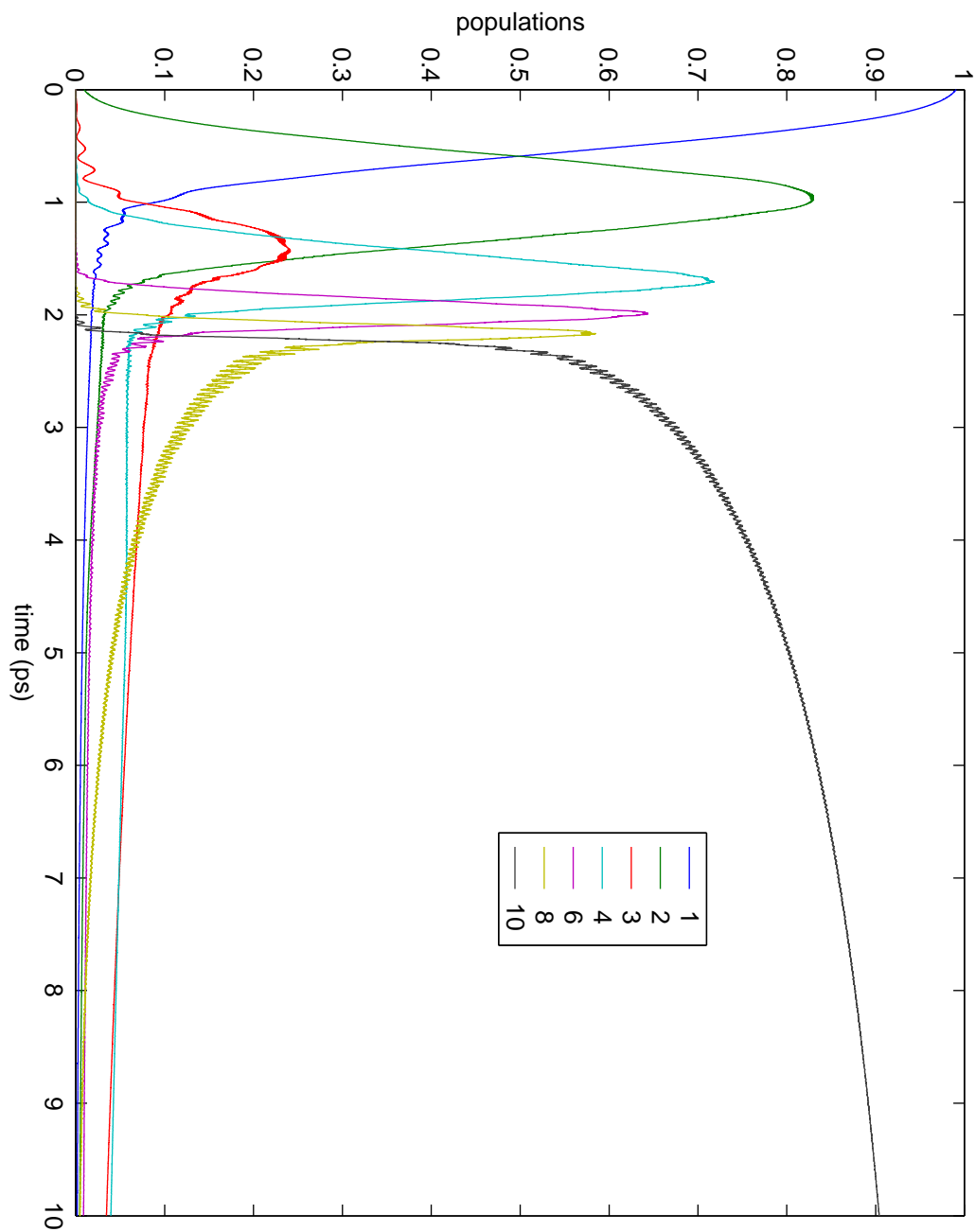


Figure 4.1: Selective excitation of an intermediate vibrational state of the OH bond using a Lyapunov based feedback control. The transfer is 90% complete in 10 ps.

Chapter 5

Optimal feedback control of a continuously monitored spin

Our presentation is organized as follows: In section 5.1 we introduce a spin- $\frac{1}{2}$ system that is continuously monitored and pose an optimal stabilization problem (with partial observation) for that system. The solution invokes the separation principle, according to which the optimal control design is done in two steps: First, one constructs an optimal filter that continuously estimates the state based on incoming measurement results. The second step, involves the solution of an optimal stabilization problem for the filter dynamics with full state information. We set up the relevant Bellman equation for this problem and in section 5.2 we solve the Bellman equation numerically for the special case of perfect detection efficiency.

5.1 Optimal feedback control of a continuously monitored spin

Figure 5.1 is a schematic representation of an experimental set-up for continuous measurement and control of a spin [36, 27]. The spin interacts weakly with an optical mode (laser) along the z -axis. The cavity is used to control the strength of that interaction. Then, a continuous measurement is performed on the laser using a technique known as **homodyne detection** [80]. This provides an indirect continuous measurement of the z -component of the spin angular momentum in such a way that the state of the system does not collapse to an eigenstate of S_z (a so-called **non-demolition measurement** [80, 71]). A magnetic field in the y -direction is applied for control purposes.

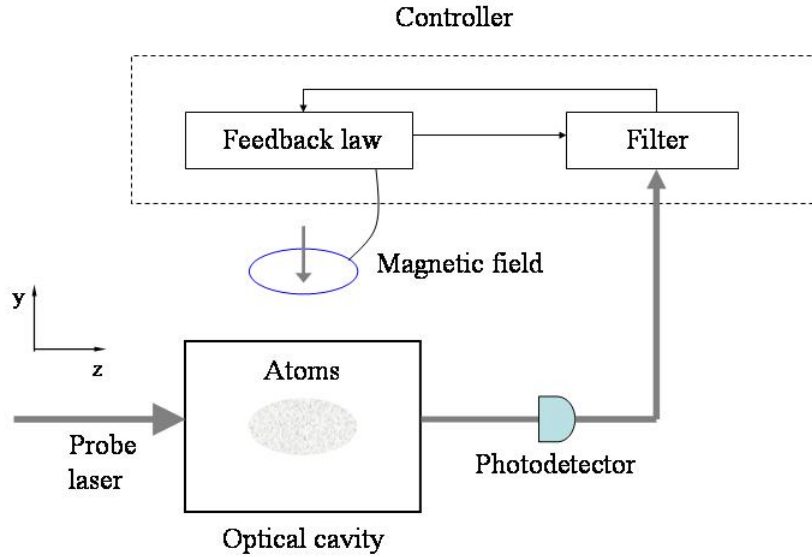


Figure 5.1: Schematic of an experiment for continuous measurement and control of a spin. The spin interacts with a probe laser, which is measured continuously using photodetectors. A magnetic field is used for feedback.

The dynamics of the continuously monitored spin system is described by a **quantum stochastic differential equation** (QSDE). The proper mathematical framework for QSDEs is the non-commutative analog of the Itô stochastic calculus that was developed by Hudson and Parthasarathy [81], see also [82, 71, 83]. There is also a corresponding observation process for a field observable of the laser probe. At this point, we should recall that it is impossible to measure all components of the state because they are non-commuting observables. Hence, any measurement can provide only partial information about the state of the system. The situation is analogous with that of classical partially observed stochastic systems. As in the classical case, one can derive a **quantum filtering equation** [26, 82] which is the equivalent of the Kushner-Stratonovich equation in classical nonlinear filtering theory [84]. The quantum filtering equation is a *classical* Itô SDE for the conditional expectation of the density matrix $\hat{\rho}_t$ (conditioned on the measurement record up to time t). The important property of the filtering equation is that the expectation value of any system observable at time t is the same whether it is computed from the corresponding QSDE or from $\hat{\rho}_t$, even with feedback control that depends on the measurement record [85]. This is a great simplification, since it allows us to circumvent the dynamical QSDE and use the filtering equation (or Belavkin equation) for control design [36, 40, 85]. Moreover, an *output feedback* control problem has been converted into a *state feedback* control problem *for the conditional expectation* of the state $\hat{\rho}_t$.

The starting point of our analysis will be the spin filtering equation. It has the following form [36]:

$$\begin{aligned} d\hat{\rho}_t &= \left(iu(t) [S_y, \hat{\rho}_t] + M(S_z \hat{\rho}_t S_z - \frac{1}{2}(S_z^2 \hat{\rho}_t + \hat{\rho}_t S_z^2)) \right) dt \\ &+ \sqrt{\eta M} (S_z \hat{\rho}_t + \hat{\rho}_t S_z - 2\text{tr}(S_z \hat{\rho}_t) \hat{\rho}_t) dW_t, \end{aligned} \quad (5.1)$$

where the **innovations** process W_t is a Wiener process that describes the difference between the *measured* value of S_z at time t and its *expected* value. The measurement strength M is determined by the properties of the cavity and the probe laser and effectively determines the time-scale of the measurement process. The parameter $\eta \in [0, 1]$ is the detection efficiency of the photodetectors and $u(t)$ is the amplitude of the magnetic field applied in the y -direction. By defining $\tau = Mt$, $V_\tau = \sqrt{M}W_t$ (so that $(dV_\tau)^2 = M(dW_t)^2 = Mdt = d\tau$) and $\tilde{u} = \frac{1}{M}u$, we can set $M = 1$, which we do from now on.

For the case of a spin- $\frac{1}{2}$ system, we use the Bloch vector representation of ρ_t introduced in Chapter 3,

$$\hat{\rho}_t = \frac{1}{2}(I_2 + x_t \sigma_x + y_t \sigma_y + z_t \sigma_z) = \frac{1}{2} \begin{pmatrix} 1 + z_t & x_t - iy_t \\ x_t + iy_t & 1 - z_t \end{pmatrix}$$

in terms of which, the spin filtering equation becomes the following system of SDEs:

$$dx_t = -(u(t)z_t + \frac{1}{2}x_t)dt - \sqrt{\eta}x_tz_t dW_t, \quad (5.2)$$

$$dy_t = -\frac{1}{2}y_t dt - \sqrt{\eta}y_tz_t dW_t, \quad (5.3)$$

$$dz_t = u(t)x_t dt + \sqrt{\eta}(1 - z_t^2) dW_t. \quad (5.4)$$

It is easy to analyze the long time limit of the dynamics of (5.2) - (5.4) in the case $u(t) = 0$. This has been done in [36] and we just cite the conclusion: The state converges either at the equilibrium point $(x, y, z) = (0, 0, +1)$ with probability $\frac{1}{2}(1 + z_0)$ or at $(x, y, z) = (0, 0, -1)$ with probability $\frac{1}{2}(1 - z_0)$. As the authors of that work remark, these are exactly the outcomes and the corresponding probabilities for a measurement of S_z . Hence, they remark, the continuous

measurement of S_z is exactly what its name suggests, a “stretched out” (in time) measurement of S_z . However, this offers the opportunity to *influence* the outcome of this measurement using a feedback control law. For example, one may want to prepare/stabilize the $S_z = +1$ eigenstate of the spin, i.e. $(x, y, z) = (0, 0, +1)$ [Note: The z -direction can be chosen arbitrarily and hence any pure state of the spin system may be prepared]. In [36], Lyapunov-based designs are proposed and analyzed for this objective. In the older work [23], the same objective is considered but in an experimental setup that involves output feedback but no measurement.

We seek to design optimal feedback strategies for this objective. Certainly, one may design feedback laws for the filter dynamics based on some objective but they will not necessarily be optimal for the system. What is needed is an appropriate **Separation Principle**, that is, a statement that under appropriate conditions the optimal feedback control designed based on the filter dynamics remains optimal for the system (with the same objective). Then, the optimal feedback is a *separated strategy*. Such a result has been proven recently in [85], thus establishing a Separation Principle for the optimal control of quantum systems. Hence, from this point on, we will concentrate on optimal designs for the filter, (5.2) - (5.4). Before we do this, we simplify things using the following observation in [36]: The variable y_t does not enter equations (5.2) and (5.4), and (5.3) does not contain the control at all. If our objective is to stabilize the eigenstate with $S_z = +1$, we may restrict our analysis in the system of (5.2) and (5.4) and stabilize $(x, z) = (0, +1)$ and then the geometry of state space will guarantee the stabilization of $y = 0$. So, we will restrict our attention to equations (5.2) and (5.4). The associated state space will now be the disk $0 \leq x^2 + z^2 \leq 1$. In the next section, it will be beneficial to our analysis to use polar coordinates (r, θ) for the state space defined by $x = r \sin \theta$ and $z = r \cos \theta$. It is a simple application of Itô’s rule to obtain the

following SDEs for them:

$$dr_t = \frac{1}{2} \left(\frac{\eta}{r_t} - r_t \right) \sin^2 \theta_t dt + \sqrt{\eta} (1 - r_t^2) \cos \theta_t dW_t, \quad (5.5)$$

$$\begin{aligned} d\theta_t &= \left[-u(t) + \left(\frac{\eta}{r_t^2} - \eta - \frac{1}{2} \right) \sin \theta_t \cos \theta_t \right] dt \\ &\quad - \sqrt{\eta} \frac{\sin \theta_t}{r_t} dW_t. \end{aligned} \quad (5.6)$$

We pose now the following optimization problem: Suppose that at time t the state of the system is (r, θ) . Let $u(s)$, $s \in [t, T]$ (T is the time at which the experiment terminates) be a square integrable function. We define the following *expected cost-to-go*

$$J(t, r, \theta, u(s)) \doteq \mathbb{E}_{(r, \theta)} \left[\int_t^T \left(\frac{1}{2} u^2(s) + U(r_s, \theta_s) \right) ds \right], \quad (5.7)$$

where the expectation value is taken with respect to every possible sample path of (5.5) - (5.6) that starts at (r, θ) at time t . The function U is a measure of the distance of the state from the desired target state $(r, \theta) = (1, 0)$ ($(x, z) = (0, 1)$). For well-posedness we require that $U(1, 0) = 0$ and $U(r, \theta) > 0 \forall (r, \theta) \neq (1, 0)$. For example, $U = 1 - z = 1 - r \cos \theta$. For now, we keep the form of U general and we will specify a form in the following section, where we will also discuss the disadvantages of certain choices. We seek the control law u that minimizes J . The expected cost-to-go of the optimal law is called the value function:

$$V(t, r, \theta) \doteq \min_{u(s)} J(t, r, \theta, u(s)). \quad (5.8)$$

Bellman's principle of optimality and dynamic programming lead to the following

Bellman equation for the value function:

$$\min_u \left\{ \frac{u^2}{2} + U + \frac{\partial V}{\partial t} + \mathcal{L}V \right\} = 0. \quad (5.9)$$

The stochastic generator \mathcal{L} of the SDE system acts on V as follows:

$$\begin{aligned} \mathcal{L}V &= \frac{1}{2} \left(\frac{\eta}{r} - r \right) \sin^2 \theta \frac{\partial V}{\partial r} + \left[-u + \left(\frac{\eta}{r^2} - \eta - \frac{1}{2} \right) \sin \theta \cos \theta \right] \frac{\partial V}{\partial \theta} \\ &+ \frac{1}{2} \eta (1 - r^2)^2 \cos^2 \theta \frac{\partial^2 V}{\partial r^2} + \frac{1}{2} \eta \frac{\sin^2 \theta}{r^2} \frac{\partial^2 V}{\partial \theta^2} \\ &- \eta \left(\frac{1}{r} - r \right) \sin \theta \cos \theta \frac{\partial^2 V}{\partial r \partial \theta}. \end{aligned} \quad (5.10)$$

Besides the penalty term $\frac{u^2}{2}$, the only other term in (5.9) that contains the control is the second term in the expression for $\mathcal{L}V$. The minimization over u is immediate with the result

$$u = \frac{\partial V}{\partial \theta}. \quad (5.11)$$

Substituting (5.10) and (5.11) into (5.9), we end up with the following nonlinear PDE:

$$\begin{aligned} \frac{\partial V}{\partial t} &+ \frac{1}{2} \eta (1 - r^2)^2 \cos^2 \theta \frac{\partial^2 V}{\partial r^2} + \frac{1}{2} \eta \frac{\sin^2 \theta}{r^2} \frac{\partial^2 V}{\partial \theta^2} \\ &- \eta \left(\frac{1}{r} - r \right) \sin \theta \cos \theta \frac{\partial^2 V}{\partial r \partial \theta} \\ &+ \frac{1}{2} \left(\frac{\eta}{r} - r \right) \sin^2 \theta \frac{\partial V}{\partial r} + \left(\frac{\eta}{r^2} - \eta - \frac{1}{2} \right) \sin \theta \cos \theta \frac{\partial V}{\partial \theta} \\ &- \frac{1}{2} \left(\frac{\partial V}{\partial \theta} \right)^2 + U(r, \theta) = 0. \end{aligned} \quad (5.12)$$

In addition to the Bellman equation, the principle of optimality provides a terminal condition for the value function, namely $V(T, r, \theta) = 0$. However, in order to solve the PDE (5.12), one needs boundary conditions for V at $r = 1$. These boundary conditions are not furnished by the principle of optimality and

have to be inferred by other means. This is what makes such optimal control problems so difficult to solve (besides the computational complexity, of course). In the following section, we will see that for the case of perfect detector efficiency ($\eta = 1$), we can provide such a boundary condition and hence obtain a solution to the problem.

5.2 Solution for the case of perfect detection efficiency

In this section we concentrate on the special case $\eta = 1$, i.e. perfect detector efficiency. Letting the initial value of r equal to 1 ($r_0 = 1$) in (5.5),

$$dr_t = \frac{1}{2} \left(\frac{\eta}{r_t} - r_t \right) \sin^2 \theta_t dt + \sqrt{1\eta} (1 - r_t^2) \cos \theta_t dW_t,$$

we see that $dr_0 = 0$ and hence $r_t = 1$, $\forall t \geq 0$. Hence, $r = 1$ is a **forward invariant set** of the stochastic dynamics (5.5) - (5.6). Physically, this means that if the detection is perfect, a pure state of the system will remain pure for all time.

Let us rewrite (5.6) on the invariant set $r = 1$:

$$d\theta_t = -(u(t) + \frac{1}{2} \sin \theta_t \cos \theta_t) dt - \sin \theta_t dW_t. \quad (5.13)$$

This stochastic system is the reduction of the full system (5.5) - (5.6) whose state space is the disk $0 \leq r \leq 1$, on the boundary $r = 1$ which is a forward invariant set of the full dynamics.

In the previous section, we set up an optimal control problem for the control system (5.5) - (5.6) whose objective is to stabilize the target S_z -eigenstate $(x, z) =$

(0, 1). The principle of optimality lead us to the Bellman PDE (5.12) for the value function along with a terminal condition for it. As noted there, in order to solve this PDE one also needs a boundary condition for V at $r = 1$. This boundary condition, however, is not provided automatically by the principle of optimality. In the case of perfect detection efficiency ($\eta = 1$), we can provide such a boundary condition based on the property shown above, that the boundary of the state space is a forward invariant set of the dynamics. Our reasoning is the following: Consider the reduction of the optimal control problem posed in the last section on the boundary $r = 1$. That is, minimize

$$\tilde{J}(t, \theta, u(s)) \doteq \mathbb{E}_{(\theta)} \left[\int_t^T \left(\frac{1}{2} u^2(s) + \tilde{U}(\theta_s) \right) ds \right], \quad (5.14)$$

over all admissible controls ($\tilde{U}(\theta) \doteq U(1, \theta)$). The principle of optimality for this problem leads to the following “boundary” Bellman equation for the “boundary” value function $\tilde{V}(t, \theta) \doteq \min_{u(s)} \tilde{J}(t, \theta, u(s))$:

$$\frac{\partial \tilde{V}}{\partial t} + \frac{1}{2} \sin^2 \theta \frac{\partial^2 \tilde{V}}{\partial \theta^2} - \frac{1}{2} \sin \theta \cos \theta \frac{\partial \tilde{V}}{\partial \theta} - \frac{1}{2} \left(\frac{\partial \tilde{V}}{\partial \theta} \right)^2 + \tilde{U}(\theta) = 0. \quad (5.15)$$

The terminal condition is again $\tilde{V}(T, \theta) = 0$. As before, the (feedback) control is given in terms of the value function by

$$u = \frac{\partial \tilde{V}}{\partial \theta}.$$

Note that one can also obtain this equation from (5.12) by dropping the terms with r -derivatives and setting $r = 1$. Since any state of the full system (5.5) - (5.6) that starts on the boundary remains on the boundary forever, the value function V for the optimal control problem (5.7) *at the boundary* must be equal to the

value function \tilde{V} for the problem (5.14):

$$V(t, r = 1, \theta) = \tilde{V}(t, \theta),$$

$\forall t \leq T$ and $\theta \in S^1$. The solution of (5.15) itself, however, requires no boundary conditions, only periodicity of the solution with respect to θ (so that \tilde{V} is well defined). Hence, one can solve the Bellman PDE (5.15) first, and use its solution, $\tilde{V}(t, \theta)$ as a time-dependent boundary condition for the PDE (5.12).

In the following, we present the solution of an optimal feedback problem of the form (5.7). We make the following remarks:

1. We are interested in the **infinite-horizon** problem, i.e. the $T \rightarrow \infty$ limit of the optimal control problem (5.7). In this limit, the value function (and hence the optimal control) are time-independent. To obtain the solution, we solved (5.15) *backwards in time* with the terminal condition $\tilde{V}(T, \theta) = 0$ until convergence to a steady state was achieved. Then, we used this solution on the boundary as time-dependent boundary condition for (5.12), which was also solved backwards in time (with the terminal condition $V(T, r, \theta) = 0$) until convergence to a steady state.
2. For the numerical solution of the Bellman equations (5.15) and (5.12) the spatial discretization was done as follows: We used a spectral scheme in the θ variable in terms of a Fourier basis $\{1, e^{\pm i\theta}, e^{\pm iN\theta}, \dots\}$ truncated to $N = 80$ and a first order finite difference scheme in the r variable using 50 points. The time stepping was implemented with MATLAB's ODE45 routine.
3. The system (5.2) - (5.4) is invariant under the simultaneous interchange $(x, u) \leftrightarrow (-x, -u)$. If the state-penalty ("potential") term U is symmetric under the interchange of the left and right side of the disk (i.e. $x \leftrightarrow -x$,

or similarly $\theta \leftrightarrow -\theta$), then the value function will have the same symmetry. It will also have corners along $\theta = \pi$, signaling the non-uniqueness of the optimal control at $\theta = \pi$ (recall, $u = \frac{\partial V}{\partial \theta}$). This is a known phenomenon in the solutions of Bellman equations in the presence of symmetry. We observed this in numerical solutions of (5.12). Since sufficient regularity of the value function is a prerequisite for the separation principle to hold [85], to avoid regularity problems in the value function we used a non-symmetric form for U . In particular, we used

$$U = (1 - z)(\alpha + x) = (1 - r \cos \theta)(\alpha + r \sin \theta),$$

with $\alpha = 1.1$. $U(x, z) > 0 \forall (x, z) \neq (0, 1)$ and $U(x = 0, z = 1) = 0$, as it should, for the optimization problem to be well-posed.

Figures 5.2 and 5.3 represent the value function and the optimal feedback for the infinite-horizon problem on the boundary of the state space. Figures 5.4 and 5.5 represent the value function and the optimal feedback on the whole disk. The asymmetry between the left and right side of the disk is obvious in all figures. To elucidate the action of the control on the system, we portray, in figure 5.6, the vector field $(-u(x, z)z, u(x, z)x)$. This vector field is the part of the dynamics of (5.2) - (5.4) that depends on the control. The fact that there are lines on which this field vanishes should not be alarming: The action of the rest of the dynamics, the deterministic dissipative part and the stochastic part, “throws” the system off these points. Finally, figure 5.7 demonstrates some typical trajectories of the closed-loop system.

5.3 Summary

We considered the problem of stabilization of a desired spin eigenstate of a continuously monitored spin- $\frac{1}{2}$ system. We posed an optimal regulation problem for the associated filter, whose solution is also the optimal feedback for the system, according to the Separation Principle. We obtained such a solution in the case of perfect photodetection efficiency.

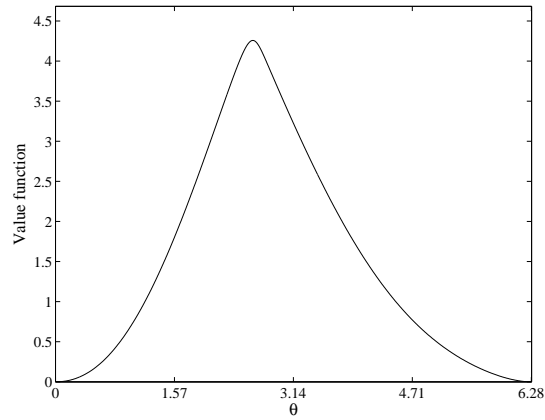


Figure 5.2: Steady-state value function \tilde{V} on the boundary

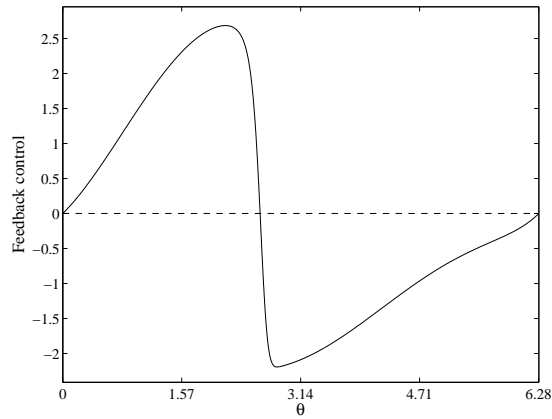


Figure 5.3: Optimal feedback on the boundary

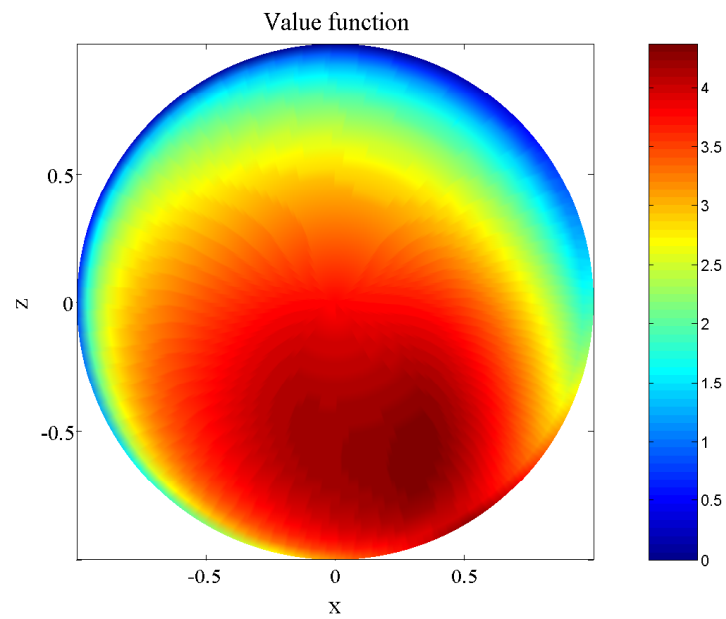


Figure 5.4: Steady-state value function V on the disk

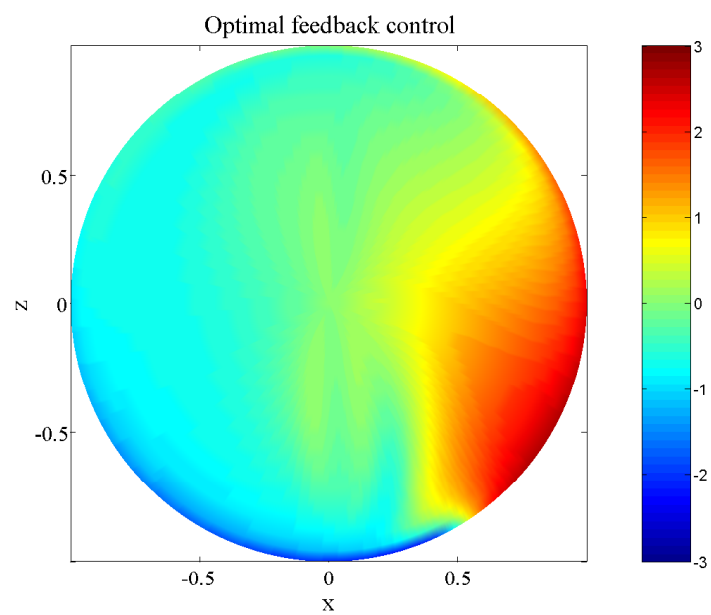


Figure 5.5: Optimal feedback on the disk

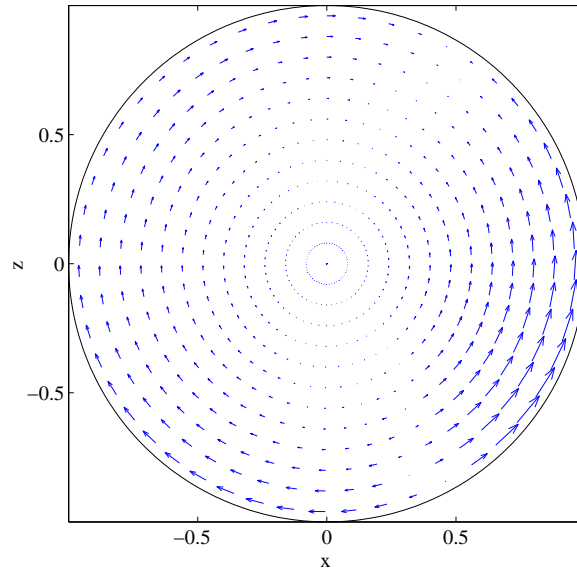


Figure 5.6: Vector field generated by the feedback control

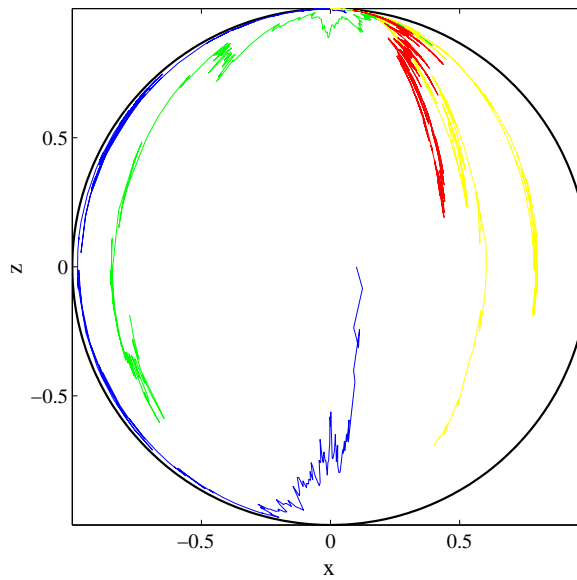


Figure 5.7: Typical trajectories of the closed-loop system

Chapter 6

Conclusions and future directions

In this thesis we presented a general methodology for simplifying the solution of energy-optimal population transfers in (isolated or open) quantum systems. The main point of our approach is that, in the large transfer time limit (much larger than the inverse of the smallest Bohr frequency of the system), the solution has a time-scale separated form. We showed how to construct corresponding “averaged” systems and optimal transfer problems that capture the long time-scale dynamics and in which the contribution of the short time-scale natural dynamics is averaged out. These reduced problems offer great conceptual and computational simplifications to the solution of the original ones. We believe that our methodology is the most useful systematic approach to these problems and can be used to generate very good first approximations to their solutions. We were able to fully prove these assertions for exact population transfer problems in isolated quantum systems but only provided evidence for minimizing-the-distance-from-the-desired-state types of problems. We would like to supply a complete correspondence between original and averaged transfer problems in that case too (the analog of Theorems 2 and 3 of Chapter 2). We would also like to investigate the robustness

properties of the open-loop controls generated through our technique to parameter uncertainty and laser noise.

We also presented the solution to an optimal feedback stabilization problem for an eigenstate of a continuously monitored spin- $\frac{1}{2}$ system in the case of perfect detection. We would like to extend this solution to the case of imperfect detection. We would also like to study optimal stabilization problems for higher dimensional spin systems through the use of effective low-dimensional models.

Bibliography

- [1] D. Tannor and S. Rice, “Control of selectivity of chemical reaction via control of wave packet evolution,” *Journal of Chemical Physics*, vol. 83, no. 10, p. 5013, 1985.
- [2] A. Peirce, M. Dahleh, and H. Rabitz, “Optimal control of quantum-mechanical systems: Existence, numerical approximation, and applications,” *Physical Review A*, vol. 37, no. 12, p. 4950, 1987.
- [3] M. Dahleh, A. Peirce, and H. Rabitz, “Optimal control of uncertain quantum systems,” *Physical Review A*, vol. 42, no. 3, p. 1065, 1990.
- [4] D.J.Tannor, V.A.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.), Plenum, 1992.
- [5] S. Shi and H. Rabitz, “Quantum mechanical optimal control of physical observables in microsystems,” *Journal of Chemical Physics*, vol. 92, no. 1, pp. 364–376, 1989.
- [6] W. Warren, H. Rabitz, and M. Dahleh, “Coherent Control of Quantum Dynamics: The Dream Is Alive,” *Science*, vol. 259, p. 1581, 1993.

- [7] H. Rabitz, R. de Vivie-Riedle, M. Motzkus, and K. Kompa, “Whither the future of controlling quantum phenomena?,” *Science*, vol. 288, no. 5467, 2000.
- [8] S. Glaser, T. Schulte-Herbruggen, M. Sieveking, O. Schedletsky, N. C. Nielsen, O. W. Sørensen, and C. Griesinger, “Unitary control in quantum ensembles: Maximizing signal intensity in coherent spectroscopy,” *Science*, vol. 280, no. 5362, pp. 421–424, 1998.
- [9] R. Brockett and N. Khaneja, “On the stochastic control of quantum ensembles,” in *System Theory: Modeling, Analysis and Control*, ch. 6, Kluwer Academic Publishers, 1999.
- [10] N. Khaneja, T. Reiss, B. Luy, and S. J. Glaser, “Optimal Control of Spin Dynamics in the Presence of Relaxation,” *Journal of Magnetic Resonance*, vol. 162, no. 2, pp. 311–319, 2003.
- [11] D. Stefanatos, N. Khaneja, and S. Glaser, “Optimal control of coupled spins in the presence of longitudinal and transverse relaxation,” *Physical Review A*, vol. 69, no. 022319, February 2004.
- [12] R. Ober and E. Ward, “A system theoretic formulation of NMR experiments,” *Journal of Mathematical Chemistry*, no. 20, pp. 47–65, 1996.
- [13] C. Altafini, “On the generation of sequential unitary gates from continuous time schrodinger equations driven by external fields,” *Quantum Information Processing*, vol. 1, no. 3, pp. 207–224, 2002.
- [14] V. Ramakrishna, K. Flores, H. Rabitz, and R. Ober, “Quantum control by decompositions of $\text{su}(2)$,” *Physical Review A*, vol. 62, no. 053409, 2000.
- [15] L. Viola and E. Knill, “Robust dynamical decoupling of quantum systems with bounded controls,” *Physical Review Letters*, vol. 90, no. 037901, 2003.

- [16] L. Viola and E. Knill, “Random decoupling schemes for quantum dynamical control and error suppression,” *Physical Review Letters*, vol. 94, no. 060502, 2005.
- [17] L. Viola, “Advances in decoherence control,” *Journal of Modern Optics*, no. 51, p. 2357, 2004.
- [18] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, September 2000.
- [19] H. Wiseman, “Quantum theory of continuous feedback,” *Physical Review A*, vol. 49, no. 3, p. 2133, March 1994.
- [20] H. Wiseman and G. Milburn, “Reduction in laser-intensity fluctuations by a feedback-controlled output mirror,” *Physical Review A*, vol. 46, no. 5, p. 2853, September 1992.
- [21] A. Liebman and G. Milburn, “Creating number states in the micromaser using feedback,” *Physical Review A*, vol. 51, no. 1, p. 736, 1994.
- [22] H. Wiseman, “Using feedback to eliminate back-action in quantum measurements,” *Physical Review A*, vol. 51, no. 3, p. 2459, March 1995.
- [23] J. Wang and H. Wiseman, “Feedback-stabilization of an arbitrary pure state of a two-level atom,” *Physical Review A*, vol. 64, no. 063810, November 2001.
- [24] A. Doherty and K. Jacobs, “Feedback-control of quantum systems using continuous state-estimation,” *Physical Review A*, vol. 60, p. 2700, 1999.
- [25] A. Doherty, S. Habib, K. Jacobs, H. Mabuchi, and S. Tan, “Quantum feedback control and classical control theory,” *Physical Review A*, vol. 62, no. 012105, 2000.

- [26] V. Belavkin, “Quantum stochastic calculus and quantum nonlinear filtering,” *Journal of Multivariate Analysis*, no. 42, pp. 171–201, 1992.
- [27] J. Geremia, J. Stockton, and H. Mabuchi, “Real-time quantum feedback control of atomic spin-squeezing,” *Science*, no. 304, 2004.
- [28] A. Bartana, R. Kosloff, and D. Tannor, “Laser cooling of molecular internal degrees of freedom by a series of shaped pulses,” *Journal of Chemical Physics*, vol. 99, no. 1, p. 196, 1993.
- [29] H. Zhang and H. Rabitz, “Robust optimal control of quantum molecular systems in the presence of disturbances and uncertainties,” *Physical Review A*, vol. 49, no. 4, p. 2241, 1994.
- [30] A. Bartana, R. Kosloff, and D. Tannor, “Laser cooling of internal degrees of freedom. II,” *Journal of Chemical Physics*, vol. 106, no. 4, p. 1435, 1997.
- [31] 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*, vol. 109, no. 2, p. 385, 1998.
- [32] Y. Ohtsuki, W. Zhu, and H. Rabitz, “Monotonically convergent algorithm for quantum optimal control with dissipation,” *Journal of Chemical Physics*, vol. 110, no. 20, p. 9825, 1999.
- [33] D. D’Alessandro and M. Dahleh, “Optimal control of two-level quantum systems,” *IEEE Transactions on Automatic Control*, vol. 45, no. 1, June 2001.
- [34] U. Boscain, T. Chambrion, J. Gauthier, S. Guérin, and H. Jauslin, “Optimal control in laser-induced population transfer for two- and three-level quantum systems,” *Journal of Mathematical Physics*, vol. 43, p. 2107, 2002.

- [35] U. Boscain, T. Chambrion, and J. Gauthier, “On the $k+p$ problem for a three-level quantum system: Optimality implies resonance,” *Journal of Dynamical and Control Systems*, no. 8, pp. 547–572, 2002.
- [36] R. van Handel, J. Stockton, and H. Mabuchi, “Feedback control of quantum state reduction,” *IEEE Transactions on Automatic Control*, no. 50, pp. 768–780, 2005.
- [37] J. Stockton, R. van Handel, and H. Mabuchi, “Deterministic Dicke state preparation with continuous measurement and control,” *Physical Review A*, vol. 70, no. 022106, 2004.
- [38] L. Bouten, S. Edwards, and V. Belavkin, “Bellman equations for optimal feedback control of qubit states,” *Journal of Physics B: Atomic, Molecular and Optical Physics*, no. 38, pp. 151–160, 2005.
- [39] C. Altafini and F. Ticozzi, “Almost global stochastic feedback stabilization of conditional quantum dynamics,” *quant-ph/0510222*, 2005.
- [40] M. Mirrahimi and R. van Handel, “Stabilizing feedback controls for quantum systems,” *arXiv:math-ph/0510066*, 2005. preprint.
- [41] J. J. Sakurai, *Modern Quantum Mechanics*. Addison-Wesley, revised ed., 1994.
- [42] C. Cohen-Tannoudji, B. Diu, and F. Laloe, *Quantum Mechanics*. Wiley-Interscience, 1996. 2 volumes.
- [43] M. Reed and B. Simon, *Methods of modern mathematical physics*. Academic Press. 4 volumes.

- [44] R. D. L. Madrid, A. Bohm, and M. Gadella, “Rigged Hilbert space treatment of continuous spectrum,” *Fortschr.Phys.*, vol. 50, no. 2, pp. 185–216, 2002.
- [45] V. Jurdjevic and H. Sussmann, “Control Systems on Lie Groups,” *Journal of Differential Equations*, no. 12, p. 313, 1972.
- [46] R. Brockett, “System Theory on Group Manifolds and Coset Spaces,” *SIAM Journal of Control*, vol. 10, no. 2, p. 265, 1972.
- [47] R. Brockett, “Lie Theory and Control Systems defined on Spheres,” *SIAM Journal of Applied Mathematics*, vol. 25, no. 2, p. 213, 1973.
- [48] V. Jurdjevic, *Geometric control theory*. Cambridge University Press, 1997.
- [49] V. Ramakrishna, M. Salapaka, M. Dahleh, H. Rabitz, and A. Peirce, “Controllability of molecular systems,” *Physical Review A*, vol. 51, no. 2, p. 960, 1995.
- [50] S. Schirmer, H. Fu, and A. Solomon, “Complete controllability of quantum systems,” *arXiv:quant-ph/0010031*, 2000. Preprint.
- [51] G. Turinici and H. Rabitz, “Quantum wavefunction controllability,” *Journal of Chemical Physics*, vol. 267, p. 1, 2001.
- [52] C. Altafini, “Controllability of quantum mechanical systems by root space decomposition of $\mathfrak{su}(N)$,” *Journal of Mathematical Physics*, vol. 43, no. 5, p. 2051, 2002.
- [53] J. Ball, J. Marsden, and M. Slemrod, “Controllability for distributed bilinear systems,” *SIAM Journal on Control and Optimization*, vol. 20, no. 4, p. 575, 1982.

- [54] G. Huang, T.J.Tarn, and J.W.Clark, “On the controllability of quantum-mechanical systems,” *Journal of Mathematical Physics*, vol. 24, no. 11, p. 2608, 1983.
- [55] C. Law and J. Eberly, “Arbitrary control of a quantum electromagnetic field,” *Physical Review Letters*, vol. 76, no. 7, pp. 1055–1058, 1996.
- [56] R. Brockett, C. Rangan, and A. Bloch, “The controllability of infinite quantum systems,” in *Proceedings of the 42nd IEEE Conference on Decision and Control*, December 2003.
- [57] H. Khalil, *Nonlinear Systems*. Prentice Hall, second ed., 1996.
- [58] N. Khaneja, R. Brockett, and S. Glaser, “Time optimal control in spin systems,” *Physical Review A*, vol. 63, no. 032308, February 2001.
- [59] D. D’Alessandro, “The Optimal Control Problem on $so(4)$ and Its Applications to Quantum Control,” *IEEE Transactions on Automatic Control*, vol. 47, no. 1, January 2002.
- [60] L. Shen, S. Shi, and H. Rabitz, “Control of coherent wave functions: A linearized molecular dynamics view,” *Journal of Physical Chemistry*, no. 97, p. 8874, 1993.
- [61] V. Jurdjevic, “Hamiltonian point of view of non-Euclidean geometry and elliptic functions,” *Systems & Control Letters*, no. 43, pp. 25–41, 2001.
- [62] S. Grivopoulos and B. Bamieh, “Iterative algorithms for optimal control of quantum systems,” in *Proceedings of the 41st IEEE Conference on Decision and Control*, December 2002.

- [63] A. P. Sage and C. C. White, III, *Optimum Systems Control*. Prentice-Hall, second ed., 1977.
- [64] U. Boscain and G. Charlot, “Resonance of minimizers for n-level quantum systems with an arbitrary cost,” *Control, Optimisation and Calculus of Variations*, vol. 10, pp. 593–614, 2004.
- [65] V. Arnol’d, *Ordinary Differential Equations*. Cambridge, MIT Press, 1973.
- [66] V. Guillemin and A. Pollack, *Differential Topology*. Prentice-Hall, 1974.
- [67] R. Alicki, “Invitation to quantum dynamical semigroups,” *arXiv:quant-ph/0205188*, 2002. preprint.
- [68] E. Davies, “Markovian master equations,” *Communications in Mathematical Physics*, vol. 39, pp. 91–110, 1974.
- [69] V. Gorini, A. Kossakowski, and E. Sudarshan, “Completely positive dynamical semigroups of n-level systems,” *Journal of Mathematical Physics*, vol. 17, no. 5, p. 821, 1976.
- [70] G. Lindblad, “On the Generators of Quantum Dynamical Semigroups,” *Communications in Mathematical Physics*, vol. 48, pp. 119–130, 1976.
- [71] A. Holevo, *Statistical Structure of Quantum Theory*. Springer, 2001.
- [72] D. Tannor and A. Bartana, “On the Interplay of Control Fields and Spontaneous Emission in Laser Cooling,” *Journal of Physical Chemistry A*, vol. 103, no. 49, pp. 10359–10363, 1999.
- [73] S. Sklarz, D. Tannor, and N. Khaneja, “Optimal Control of Quantum Dissipative Dynamics: Analytic solution for cooling the three level Λ system,” *Physical Review A*, vol. 69, no. 053408, 2004.

- [74] S. Schirmer, M. Girardeau, and J. Leahy, “Efficient algorithm for optimal control of mixed-state quantum systems,” *Phys. Rev. A*, vol. 61, no. 012101, 2000.
- [75] N. Khaneja and S. Glaser, “Constrained Bilinear Systems,” in *Proceedings of the 41st IEEE Conference on Decision and Control*, December 2002.
- [76] P. Vettori, “On the convergence of a feedback control strategy for multilevel quantum systems,” in *Proceedings of the MTNS Conference*, 2002.
- [77] U. Vaidya, D. D’Alessandro, and I. Mezic, “Control of Heisenberg spin systems; Lie algebraic decompositions and action-angle variables,” in *Proceedings of the 42nd IEEE Conference on Decision and Control*, December 2003.
- [78] I. Mezic, “Controllability of Hamiltonians system with drift: Action-Angle variables and ergodic partition,” in *Proceedings of the 42nd IEEE Conference on Decision and Control*, December 2003.
- [79] K. Astrom and K. Furuta, “Swinging up a pendulum by energy control,” *Automatica*, no. 36, p. 287, 2000.
- [80] D. Walls and G. Milburn, *Quantum Optics*. Springer-Verlag, 2nd ed., 1995.
- [81] R. Hudson and K. Parthasarathy, “Quantum Itô’s formula and stochastic evolutions,” *Communications in Mathematical Physics*, no. 93, pp. 301–323, 1984.
- [82] L. Bouten and R. van Handel, “Quantum filtering: a reference probability approach,” *arXiv:math-ph/0508006*, 2005. Preprint.
- [83] K. Parthasarathy, *An Introduction to Quantum Stochastic Calculus*. Birkhauser, 1999.

- [84] A. Bensoussan, *Stochastic control of partially observable systems*. Cambridge University Press, 1992.
- [85] L. Bouten and R. van Handel, “On the separation principle of quantum control,” *arXiv:math-ph/0508006*, 2005. Preprint.