# Geometric Control in Classical and Quantum Systems

A thesis presented

by

Navin Khaneja

to

The Division of Engineering and Applied Sciences
in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy
in the subject of

Applied Mathematics

Harvard University
Cambridge, Massachusetts

June 2000

Dedicated to my parents, Darshan and Saroj.

# Abstract

In this thesis, we address problems in control and stabilization of nonholonomic control systems arising in the areas of quantum physics, robotics, and locomotion systems. We provide a control theoretic framework for problems involving manipulation of quantum systems. The problem of design of pulse sequences in coherent spectroscopy is treated as a problem of constructive controllability in geometric control. We derive time optimal control laws for a class of control problems with drift, evolving on compact Lie groups. It is shown that these results find applications in design of pulse sequences that minimize decoherence effects in spectroscopic experiments and maximize signal to noise ratio. We also analyze in detail the problem of feedback stabilization of nonholonomic control systems. For nonholonomic systems, smooth state feedback control laws do not exist. In this thesis, we show how this topological obstruction can be overcome by embedding the system in a higher-dimensional manifold and constructing dynamic controllers. The choice of higher dimensional space is dictated by the symmetries of the system and can be interpreted as a gauge in the system. We finally draw a bridge between gauge theories in physics and control of nonholonomic control systems.

# Acknowledgements

First, I would like to express my sincere thanks to my thesis advisor, Prof. Roger Brockett. His ideas of a broad, scientific, and mathematical training have guided my education and research for last three years, and I have cherished every moment of it. His zest for creative work will always be a source of inspiration. I am also grateful to Prof. Steffen Glaser, whose collaboration has been so valuable in creating a bridge between the fields of control theory and NMR. I would also like to express my appreciation towards the members of my doctoral committee, Prof. Roger Brockett, Prof. Sanjoy Mitter, Prof. Y.C. Ho, and Prof. Michael Brandstein, for evaluation of this work. I am also indebted to Prof. Ulf Grenander, whose collaboration instilled in me a love for Applied Mathematics.
I must also acknowledge the help and support of fellow students and friends including:

Anuj Srivastava, Dimitiris Hristu and Kristi Morgansen, who have always been generous with advice and help.

Manuela Pereira, for being such a great and supportive friend.

Kathleen LaFrance, Hongyi Li, Magnus Egerstedt, Stan Jurga, Mark Hoefer and Sam Pfister, for creating a great environment in the Robotics Lab.

Dan, Kommu, Amit, Rob Grothe, Shantanu, Joy, Abed, Nimmi, Sudhanshu, Pallavi and Madhura for sharing some great times in and outside school.

Sarah Morelli, who has been so supportive of my efforts, whose friendship and support was so important in navigating the ups and downs of graduate school.

Finally, I would like to thank my family, whose love and support have accompanied me throughout my life. My grandparents, whose conviction in hard work and simple living have always been such a source of inspiration. Rakhi, for all the great times we shared together while growing up. My father, who has always managed to convince me that all my failures are indeed such great opportunities. My mother, who loves me so much and always makes me feel so special. Mom and Dad this thesis is dedicated to you.

# Contents

# List of Figures

# Chapter 1

# Introduction

The central theme of the thesis is control and stabilization of nonholonomic systems arising in classical and quantum mechanics, with applications to coherent spectroscopy, robotics, and locomotion systems. Our concern will be with situations where one exploits geometry to do control. A common scenario where this geometric structure occurs is in presence of non-integrable constraints in the system, which cannot be written as time derivative of some functions of the generalized coordinates. The constraints force the instantaneous velocity of the bodies to live in a restricted set of directions, but this does not restrict the reachable states of the system, as the movement in other directions can be achieved by cyclical motions in directly accessible directions. A familiar example of this geometric structure is exploited in parallel parking of an automobile, where periodic motion of the driving speed and the steering angle can be used to achieve a net sideways motion.

These systems are examples of nonholonomic systems and have been studied in classical mechanics for more than a century, but only recently has the study of control of such systems been initiated. These nonlinear systems are not generally amenable to the methods of linear control theory and cannot be approximated by linear systems in any meaningful way. Hence, these problems require fundamentally nonlinear approaches. Examples of nonholonomic control systems have been studied in context of robotics manipulation, mobile robots, wheeled vehicles, and space robotics. Specific examples of nonholonomic control systems include sledges or knife edge systems that slide on a plane (A. Bloch and M.Reyhanoglu (1992)), a simple wheel rolling without slipping on a plane, and spheres rolling without slipping on a plane (Brockett and L.Dai (1992)). There is now an extensive literature on control of mobile robots and wheeled vehicles (kolmanovsky and McClamroch (1995)).

Another common setting where nonholonomic control systems arise is in mechanical systems with symmetries. If the motion of a mechanical system exhibits certain symmetries then there are always associated conserved quantities. If these conserved quantities are not integrable then a nonholonomic system is obtained. Examples of such nonholonomic control systems include actuated multi-body spacecraft and space robotics. A popular example of this phenomenon, is the example of a falling cat (KANE AND SHUR (1969); MONTGOMERY (1990); ENOS (1993)), which can execute a 180°, rotation in space, even though it has nothing to push against. This is achieved by executing a sequence of internal shape changes. This phenomenon is ubiquitous in biological and robotic locomotion and has been well studied. The general idea being that when one variable in the system moves in a cyclical motion, other variables that could not be directly actuated can be effected. Astronauts who wish to reorient themselves in space can similarly do so by means of shape change. Similarly microorganisms swimming at low Reynolds numbers (SHAPERE AND WILCZEK (1987)) can generate translations by specific cyclic manipulation of their internal shape. The key geometric concept underlying all these examples is the notion of holonomy of a connection also known in physics by the name of geometric phase. To get an intuitive feel for the concept consider the following familiar Figure 1.1 often used to introduce the concept of holonomy. We now interpret this from a control theory perspective. Suppose the control task is to rotate the vector pointing down at the start position in the Figure 1.1, counterclockwise by 90°. Notice there exist no actuation mechanism for effecting this change. However, if we translate this vector parallel to itself along a loop as shown in the figure, on return we find the vector has rotated by 90°. This is what we mean by using geometry to do control.

These rich geometric phenomenon are not just confined to classical domain. Since the early days of quantum mechanics, it has always been man's dream to manipulate phenomenon at the molecular and atomic scale. Since the original inception of control over quantum systems as a goal, the potential applications have grown. Over the last 50 years quantum mechanical effects have come to be applied in very sophisticated ways, involving the control and observation of quantum systems using subtle non-commutative effects. Advances in the areas of nuclear and electron magnetic resonance spectroscopy, microwave and optical spectroscopy, laser coherent control, solid state physics and quantum computing involve active control of quantum dynamics (BROCKETT AND KHANEJA (1999), WARREN ET AL. (1993), RABITZ (1993), SCHWEIGER (1990), LLOYD (1996), TAUBES (1997), CORY ET AL. (1997), GLASER ET AL. (1998)). Analysis and study of control of quantum systems,

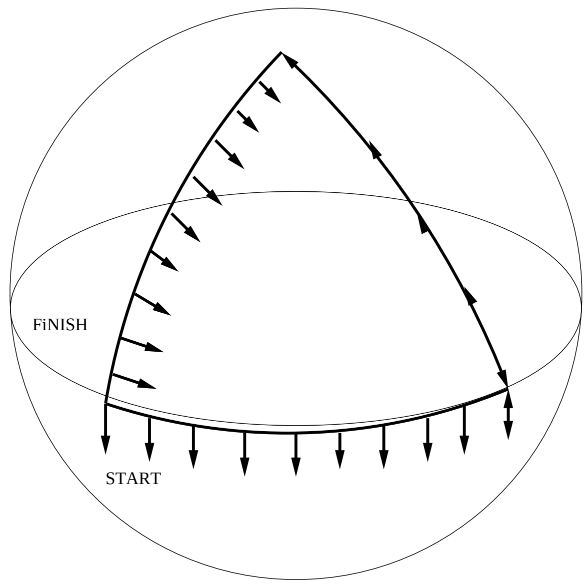


Figure 1.1: The panel illustrates the concept of holonomy. The vector experiences a rotation by 90 degrees, when moved around in a loop.

which form a major contribution of this thesis, provides an excellent opportunity to exercise ideas from geometric control. Several experiments in optics have been performed, that exhibit interference effects attributable to geometric phase or holonomy (Shapere and Wilczek (1987)). Though our focus in this thesis is on geometric control problems arising in the control of quantum systems, we give here a synopsis of the nature of problems arising in control of quantum systems. In principal, quantum phenomenon of any sort can be manipulated and controlled by optical or other means. These range from controlling electrons in solid state multiple quantum wells. Encoding and decoding of information in the wavepacket of moving electrons in semiconductors. Active control of molecular dynamics to extract information about the underlying Hamiltonian of the system, resulting in wide range of spectroscopy methods.

In particular, we are interested in looking at the control theoretic aspects of the phenomenon characterized by atom radiation interactions. The interaction between atoms and radiation have fascinated physicists for a long time. In fact, the interaction of light with matter, in particular blackbody radiation and photo-electric effect, were among the major experimental discoveries that initiated the development of quantum mechanics. In this introductory chapter, we have made an attempt to give a synopsis of the nature of problems arising in control of quantum systems. Towards this goal, we classify the problems we are interested in into four broad categories. This list is by no means exhaustive. Our attempt has been to call out for attention some interesting problems involving control and manipulation of phenomenon at molecular scale.

**Mechanical Effects of Light**

The idea that light carries momentum and can therefore exert pressure was proposed by Kepler and later elaborated by Newton. But it was Maxwell's 1873 theory of light that consistently implied the existence of a radiation pressure. Indeed, Maxwell's theory of electromagnetism provided the first classical expression for the momentum carried by radiation field. Over the years, the concept of radiation pressure was used successfully in connection with the physical problems as diverse as the internal stability of massive stars and orbital motion of satellites. In quantum optics, the momentum carried by light becomes of central importance. The possibility of using light for controlling the motion of atoms was realized early and even observed experimentally in 1933. For visible light, this effect is too small as the momentum transfer from a visible photon to an atom is orders of magnitude smaller than the momentum of a thermal atom. The situation has changed with the introduction of modern laser sources of high brightness. The use of laser sources leads to such high pressure on macroscopic bodies that they can be utilized for actual manipulation of the particles. It has been demonstrated that small neutral particles, dielectric balls can be accelerated and trapped using the radiation pressure force. The repeated absorption of photons from an intense laser beam can induce a force more than a million times as large as earth's gravity. The realization that these forces might be used to reduce atomic velocity of atoms, thus cooling them to temperatures in milli and micro-kelvin scale range (S. Chu and Ashkin (1985),C.Cohen-Tannoudji (1992)), has opened a new field in quantum electronics. Besides the cooling, it is possible to store the atoms in traps whose wall consist of potential energy from interaction with the electromagnetic field (W.D. Phillips and Metcalf (1985)). To summarize, these problems exploit the mechanical effects of radiation to control phenomenon at atomic scale (Seto and Brockett (1999)).

**Spectroscopy**

Chemists use a variety of spectroscopic methods to characterize and study their compounds, to follow reactions, and to understand bonding. In each case, as indeed with any form of spectroscopy, a system having various energy levels available to it, is probed. Some form of electromagnetic radiation is provided which has energy in the range needed to excite transitions between the energy levels, following the normal resonance absorption conditions $\Delta E = h\nu$, where $h$ is the Planck's constant and $\nu$ is the frequency of the radiation. The design of these electromagnetic pulse sequences that are needed to probe, the structure of matter, provides us with many control theoretic problems. We will discuss the design of pulse sequences in high-resolution Nuclear Magnetic Resonance Spectroscopy (NMR) in detail in this thesis. Similar ideas find applications in the fields of:

1. Magnetic Resonance Imaging ;
2. Nuclear Quadrupole Resonance ;
3. Microwave Spectroscopy ;
4. Ultraviolet and visible Resonance ;
5. Infrared Spectroscopy ;
6. Electron Paramagnetic Resonance ;
7. Laser and Atomic Spectroscopy.

From a control theory perspective, we can classify these problems as system identification or observability problems, where the task is to excite the system in an optimal way so as to get maximum information out of it.

**Reducing noise in quantum systems**

A very interesting control problem of great practical importance is controlling noise level in a quantum systems (Bloch and Rojo (2000)). Squeezing has been suggested as a mechanism for reducing uncertainty in quantum systems below the standard quantum limit. The realizations of sources of squeezed light opens exciting possibilities for precession measurement, beyond the vacuum state or shot noise limit, including applications in interferometers, optical communication, and ultrasenstive laser spectroscopy. Squeezed states of light have phase dependent quantum fluctuations. In one quadrature, the quantum fluctuations are reduced below the vacuum level, while the fluctuations in the other quadrature are increased. Theoretical work in this area has appeared since 1960's, however only recently has the experimental realization of squeezed light with fewer quantum fluctuations than vacuum achieved.

**Coherent Control of Chemical Reactions**

Another exciting theoretical possibility of control of quantum phenomenon occurs in coherent control of chemical reactions (Rabitz (1993)). The main goal here is to selectively break bonds in a polyatomic molecule, thereby giving the chemist the ability to alter and control chemical reactions with high specificity. This is achieved by identifying the mode frequency of the targeted bond and then irradiating the molecule with a high energy laser at that frequency, till the bond breaks. Calculations suggest that appropriately shaped laser pulses can break strong bonds, change reaction pathways and force molecules to climb an harmonic ladders (Warren et al. (1993)). From a control theory perspective, the goal is to manipulate quantum interferences at a molecular scale. The major challenges in this area being lack of precise knowledge of the molecular Hamiltonian and robustness issues associated with the errors in laboratory implementation of controls. Nonetheless, the field is a very active area of research from both chemistry and control theory perspective.

## 1.1 Goals of the thesis

The main contribution of the thesis is two-fold. We first study some typical problems arising in the control of quantum systems, in the areas of Nuclear Magnetic Resonance, Electron Magnetic Resonance, and Solid State Physics. These resemble the nonlinear control problems often studied in the context of nonlinear control as one only has very limited control degree of freedom as compared to the state space of the system. However, as control problems, these are non-standard because one is controlling an ensemble of (nearly) identical systems, using (nearly) the same control. Only certain aspects of the distribution of the initial states are known and the goal is to temporarily alter this distribution. The evolution of the individual elements of the ensemble is governed by Schrödinger's equation and the control is achieved by manipulating the potential energy term in the Hamiltonian. In practical applications, the potential energy term is altered by irradiating the system with sequences of radio frequency pulses of appropriate shape and frequency. Because the potential energy term enters the Schrödinger equation as a multiplicative factor, the system has the character of a bilinear system and non-commutative effects are important. Thus, one might characterize these problems as involving the control of the distribution associated with an ensemble of bilinear systems. Another important feature of such problems is that the time window in which one can execute a control is very small because the interaction with external environment perturbs the system and forces it to return to the equilibrium state. Hence, time optimal control are absolutely essential requirement for such problems. We will treat in detail the problem of design of pulse sequences in NMR as a problem of control of systems on compact Lie groups. We will derive optimal control laws for such system and show that they improve the sensitivity of NMR experiments as compared with the one obtained using current techniques. We then look at the problem of feedback stabilization in nonholonomic systems. Recently there has been lot of interest in doing feedback control in quantum systems (WISEMAN AND MILBURN (1993); LLOYD AND SLOTINE (1998)). The problem of feedback stabilization is one of the most frequently studied problems in automatic control. Here, one has a desired value of a variable, say the temperature in a room, and the task is to keep the temperature level constant in wake of changing weather. This is a special case of the problem of tracking a desired signal, e.g. keeping a camera focussed on a moving target. The design of stable regulators is one of the oldest problems in control theory. Following the work of BROCKETT (1983), we will show that the nonlinear control problems we are interested in, where not all degrees of freedom are directly actuated and one needs to produce cycles in certain directions to effect control in other directions, there are inherent problems with smooth feedback stabilization. We will elaborate on all this in much more detail, however to fix ideas for now consider the following system called the nonholonomic integrator:

$$
\begin{aligned}
\dot{x_1} &= u_1 \\
\dot{x_2} &= u_2 \\
\dot{x_3} &= x_1 u_2 - x_2 u_1.
\end{aligned}
$$

Although the above system is controllable, i.e there exist control laws $u_1 : [0,1] \to \mathbb{R}$ and $u_2 : [0,1] \to \mathbb{R}$ which transfer the system state from given initial state $(x_1(0),\ x_2(0),\ x_3(0))$ to some specified final state $(x_1(1),\ x_2(1),\ x_3(1))$, there exists no continuous feedback (BROCKETT (1983)) $u_1(x_1, x_2, x_3)$ and $u_2(x_1, x_2, x_3)$ which asymptotically stabilize the system trajectories to $(x_1, x_2, x_3) = 0$. The nonexistence of a continuous feedback control law, is a topological phenomenon. We will show how this problem can be circumvented by constructing dynamic regulators. By adding additional degrees of freedom into system dynamics, we will show that these topological problems can be taken care of. We will look at some nonholonomic systems with symmetry and show how the additional degrees of freedom introduced in the system dynamics can be interpreted as a gauge. This will bring us to an interesting mix of ideas between gauge theories in physics and stabilization of nonholonomic systems.

## 1.2 Organization of the thesis

The thesis is organized as follows

- In chapter 2 we recapitulate the fundamentals of quantum statistical mechanics and look at the structure of a typical problem in control of quantum ensembles. We then cast the design of pulse sequences in NMR as a problem of geometric control. Our main goal in this chapter is to introduce and motivate the control theoretic aspects of problems in coherent spectroscopy.

- In chapter 3 we focus on the design of shortest pulse sequences which accomplish desired transfers in NMR spectroscopy. We will treat this problem as the problem of time optimal control of systems on compact Lie groups. We give analytical characterization of the time optimal trajectories. We show how the optimal control formulation improves the signal to noise ratio in coherence transfer experiments, obtained using current techniques.

- In chapter 4 we introduce the problem of feedback stabilization in nonholonomic systems. It is shown how the topological obstruction encountered in smooth stabilization of nonholonomic systems can be overcome by embedding the state space of the system in a higher dimensional manifold. To demonstrate our approach we construct dynamic feedback controllers for the first bracket controllable systems.

- In chapter 5 we derive smooth feedback control laws for stabilizing nonholonomic systems as solutions to variational problems. In the process, we introduce the concept of gauge extension, where additional controls are introduced in the system dynamics by making the global symmetries of the system time varying.

In this, thesis we have assumed that the reader is familiar with the basic background and notation from differential geometry. However for sake of completion, we have included the essential material as an appendix chapter.

# Chapter 2

# Control in Quantum Systems

Over the last 50 years quantum mechanical effects have come to be applied in very sophisticated ways. Some of these applications involve the control and observation of quantum systems using subtle non-commutative effects. Advances in the areas of nuclear and electron magnetic resonance spectroscopy, microwave and optical spectroscopy, laser coherent control, solid state physics and quantum computing involve active control of quantum dynamics (Warren et al. (1993), Rabitz (1993), Schweiger (1990), Lloyd (1996), Taubes (1997), Cory et al. (1997), Glaser et al. (1998)). However, only recently has there been any attempt to look at these from a control theory perspective. As control problems, these are nonstandard because one is controlling an ensemble of (nearly) identical systems, using (nearly) the same control. Only certain aspects of the distribution of the initial states are known and the goal is to temporarily alter this distribution. The evolution of the individual elements of the ensemble is governed by the Schrödinger's equation and the control is achieved by manipulating the potential energy term in the Hamiltonian. In practical applications, the potential energy term is altered by irradiating the system with a sequence of electromagnetic pulses of appropriate shape and frequency. Because the potential energy term enters the Schrödinger equation as a multiplicative factor, the system has the character of a bilinear system and non-commutative effects are important. Thus, one might characterize these problems as involving the control of the (empirical) probability distribution associated with an ensemble of bilinear systems.

In this chapter we cast some of the main ideas from NMR spectroscopy in a system theoretic framework. For example, NMR spectroscopy is taken to be a system identification problem. Many key aspects of high-resolution NMR spectroscopy involve manipulating and controlling nuclear spins using radio frequency pulse sequences in such a way as to generate a suitable signal for the identification problem. This active control of nuclear spin is presented as a problem in the control of nonlinear systems. In the early work of Hahn and Purcell (HAHN (1950b) and CARR AND PURCELL (1954)) on pulse sequences, it was possible to obtain significant results using physical intuition to specify the nature of the pulses. Subsequently, as more difficult experiments with more demanding performance specifications have emerged, there have been attempts to model the problem using ideas from optimal control theory. These problems are usually formulated as fixed time, state to state transfer problems. Computer-controlled pulse shaping tools with resolution on femtosecond time scales are becoming practical making theoretical designs closer to practice. In many practical situations of interest, the target state may not be accessible as the evolution in quantum systems is necessarily unitary, then the question of interest is that of finding a transfer between a initial and the final state that is as close to the desired target as possible (GLASER ET AL. (1998)). Another important issue of great relevance is the robustness of the control as all the internal parameters of the system encoded in the Hamiltonian are not known completely and also the implementation of the control field is not exact (WARREN ET AL. (1993)). This active control of nuclear spin is presented as a problem in the control of nonlinear systems.

The chapter is organized as following. We begin with a brief introduction to the phenomenon of NMR. This is followed by recapitulation of essential notions in the nonrelativistic quantum mechanics which are most pertinent to the NMR problem. A full quantum mechanical description of NMR leading to operator formalism and density matrix theory will be given. Design of pulse sequences for NMR experiments is treated an optimal control problem of bilinear systems evolving on Lie groups. Algorithms based on gradient flows on Lie Groups are presented for optimizing signal-to-noise ratio in NMR experiments. The need for designing shorter pulse sequences which accomplish desired transfers cannot be overemphasized, especially in spectroscopy of macromolecules where relaxation effects might be important. Naturally, this will lead us to the subject of time optimal control in spin systems, which is developed in detail in the Chapter 3.

## 2.1 Nuclear Spin and Resonance

NMR in bulk condensed phase were detected in 1946 by BLOCH ET AL. (1946) and PURCELL ET AL. (1946). Nuclear Magnetism is a manifestation of the quantum mechanical property of nuclear spin angular momentum, a phenomenon that has no classical counterpart. We will treat the spin angular momentum is great detail later on. At present, we just assume that some atomic nuclei have an intrinsic spin which makes this nuclei behave as little bar magnets. This angular momentum is characterized by an integer or half integer spin quantum number $I$. Nuclei with odd mass number have half integral spin quantum number, on the other hand nuclei with odd charge number and even mass number have integral spin quantum numbers. The nuclear spin angular momentum $\mathbf{I}$ is a vector quantity with the three components along the $(x, y, z)$ direction specified by $(I_x, I_y, I_z)$. Because of the Heisenberg's uncertainty principle in quantum mechanics, only one of the three Cartesian components of $\mathbf{I}$ can be specified. By convention, the value of the $z$ component of $\mathbf{I}$ is specified by the equation

$$I_z = \hbar m \ ,$$

where $m = (-I,\ -I+1,\ \ldots\ ,\ I-1,\ I)$ and $2\pi\hbar$ is the Planck's constant. Thus, a nucleus with spin $I$ has $(2I+1)$ possible orientations specified by the value of magnetic quantum number $m$. Nuclei that have nonzero spin angular momentum possess a magnetic moment. The nuclear magnetic moment, $\mu$, is collinear with the vector representing the nuclear spin angular momentum vector and is defined by $\mu = \gamma\mathbf{I}$, where $\gamma$ is the gyromagnetic ratio and is a characteristic for the given nucleus. In absence of an external magnetic field, all the possible $2I+1$ orientations have the same energy and the spin angular momentum does not have a preferred orientation. If a magnetic field $B_0$ is applied say in $z$-direction, then the energy of the various states are given by $E = -\mu B_0$ or

$$\begin{aligned} E &= -\gamma I_z B_0 \\ E &= -\gamma\hbar m B_0. \end{aligned}$$

The energy separation between the levels is constant and is given by $\gamma\hbar B_0$. Thus, energy is lower if the nuclear magnetic moment is aligned with the magnetic field and increases as the magnetic moment is aligned opposite to the vector field. Analogous to other spectral phenomenon, the presence of various states differing in energy provides for a situation where interactions can take place with the electromagnetic radiations of the correct frequency and

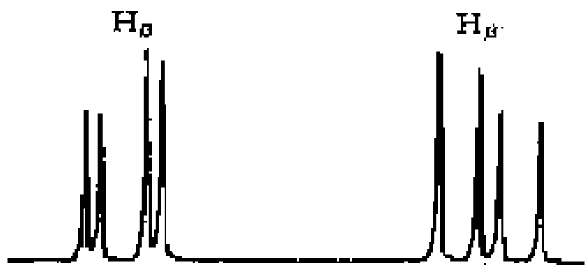

Figure 2.1: The panel shows the NMR spectrum of hydrogen nuclei. The different peaks in the frequency spectrum correspond to hydrogen nuclei resonating at slightly different frequencies as they have different chemical environment.

can excite excitations between these states. The frequency is obtained from Bohr's relation

$$h\nu = \Delta E.$$

For NMR, the energy separation is $h\nu = \gamma\hbar B_0$. Due to the selection rules of quantum mechanics, transitions are only allowed between levels $\Delta m = \pm 1$. Thus, the nucleus interacts with radiation whose frequency depends only on the applied magnetic field and the nature of the nucleus. From now on, we will only consider spin $\frac{1}{2}$ nuclei. We conclude from our previous discussion that, for spin $\frac{1}{2}$ nuclei, there are two spin states with spin quantum number $+\frac{1}{2}$ and $-\frac{1}{2}$ which we call $\alpha$ and $\beta$ states, respectively. In the absence of external magnetic field, both the states are equally populated but, in the presence of a magnetic field, the ratio between nuclei in low energy state $N_l$ and high energy state $N_h$ is given by the Maxwell-Boltzmann distribution

$$\frac{N_h}{N_l} = exp(\frac{-\Delta E}{kT}),$$

where $\Delta E$ is the energy separation between the levels, $k$ is the Boltzmann constant, and $T$ is the absolute temperature. For a hydrogen nuclei in a field of 1.4 Tesla at room temperature $\frac{N_h}{N_l} \approx 10^{-5}$. The excess population in the low energy state is extremely small. If this system is pumped with an electromagnetic radiation of the right frequency, then the excess spin in the lower energy state absorb energy and is excited to a higher energy state. However, this excited state is unstable and the system relaxes to the original state, in the process gives out radiations having a characteristic frequency which depends on the applied magnetic field and the nature of the nuclei. This forms the basis of NMR spectroscopy.

Figure 2.1 shows the Fourier transform of a signal given out by the hydrogen nuclei. The peaks in the spectrum correspond to the frequency corresponding to the energy gap between the two states of the hydrogen. However, in practice, the signal we observe is not from an individual hydrogen spin but from an ensemble. Hence, what one observes can then be described in a oversimplified way by a bulk magnetic moment, M and the bulk angular momentum J, of a macroscopic sample. Theses are given by vector sum of the corresponding quantities for individual nuclei, $\mu$ and $\mathbf{I}$, respectively. BLOCH ET AL. (1946) formulated a simple semi-classical vector model to describe the behavior of a sample of noninteracting spin $\frac{1}{2}$ nuclei in a static magnetic field. We give a brief overview of the model here as it captures the essential physics in a single spin-case.

### 2.1.1 Bloch Model and One Pulse NMR Experiment

We now describe the most basic of NMR experiments: the one pulse experiment in Bloch's framework. To understand spin dynamics in molecules with more that one spin, we will need to resort to a quantum mechanical description, which we will develop in a subsequent section. The evolution of a bulk magnetic moment vector $M(t)$ is central to the Bloch formalism. In the presence of a magnetic field $B(t)$, which may include components in addition to the static field $B_0$, the magnetic moment vector $M(t)$ experiences a torque, which is equal to the time derivative of angular momentum

$$\dot{J}(t) = M(t) \times B(t).$$

Using $M(t) = \gamma J(t)$, we can rewrite the above equation as

$$\dot{M}(t) = M(t) \times \gamma B(t). \tag{2.1}$$

If we make a change of coordinates to a frame rotating with respect to the fixed axis with angular velocity represented by the vector $\vec{\omega}$, the equation of motion in the rotating frame takes the form

$$\dot{M}(t) = M(t) \times (\gamma B(t) + \vec{\omega}).$$

Thus, we observe that the motion of the magnetization in the rotating frame has the same form as the laboratory frame if we replace the field $B(t)$ by an effective field,

$$B_{eff} = B(t) + \frac{\vec{\omega}}{\gamma}.$$

Notice that for $\vec{\omega} = -\gamma B(t)$, the effective field is zero so that $M(t)$ is time independent in the rotating frame. As seen from the laboratory frame, $M(t)$ precesses around $B(t)$ with frequency $\omega = -\gamma B$. Precession of the magnetization about the effective field in the rotating field is illustrated by in Figure 2.2. For a static field of strength $B_0$ say along the $z$-axis, the precessional frequency, or the *Larmor frequency*, is given by

$$\omega_0 = -\gamma B_0. \tag{2.2}$$

This frequency is precisely the resonant frequency needed to excite transitions between spin states. Precession of the bulk magnetic moment around a static magnetic field constitutes a time-varying magnetic field. By Faraday's laws of induction, it will induce a current in a nearby placed coil. This is the signal that is observed in a NMR experiment. In this chapter, we will focus on pulsed NMR. This is characterized by a short burst of radiofrequency (rf) electromagnetic radiation, typically of the order of several microseconds, which perturbs the bulk magnetization from equilibrium. Such a burst of electromagnetic radiation is called a pulse. The rf field or pulse takes then the following form

$$B_{rf}(t) = B_1\{cos(\omega_{rf}t + \phi)\mathbf{i} + sin(\omega_{rf}t + \phi)\mathbf{j}\}.$$

In equilibrium and in the presence of a static magnetic field $B_0$, the net bulk magnetization points in the $z$-direction. Recall that the bulk magnetic moment is the vector sum of magnetic moment of individual nuclei. Since there are excess of nuclei pointing in the direction of magnetic field, we have a net $z$-component of bulk magnetic moment. Also, because the $x$ and $y$ component of the angular momentum of these nuclei are uncorrelated, they all add up to zero, leaving the net bulk magnetization pointing in the $z$ direction.
If we transform to a rotating frame with angular velocity $\omega_{rf}$ about the $z$ axis, the equation of motion for the magnetization in the rotating frame, $M^r(t)$, is given by

$$\frac{dM^r(t)}{dt} = M^r(t) \times \gamma B^r,$$

where the effective field, $B^r$, in the rotating frame takes the form

$$B^r = B_1 cos\phi\mathbf{i} + B_1 sin\phi\mathbf{j} + (\omega_0 - \omega_{rf})\mathbf{k}.$$

Where $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are the unit vectors in $x, y, z$ direction. By convention, we choose $\phi = 0$. If the transmitter frequency $\omega_{rf}$ is equal to $\omega_0$, then the irradiation is said to be applied on resonance and the effective magnetic field $B^r$ is $B_1\mathbf{i}$. This implies that the bulk magnetization $M^r$ will precess around a magnetic field pointing in $x$-direction with frequency $\omega^r = -\gamma B_1 = \omega_1$. If the duration of the pulse is such that the magnetization which was initially pointing in the $z$ direction is brought into the $x - y$ plane, the pulse is called a $90°$ or $\frac{\pi}{2}$ pulse. A $\frac{\pi}{2}$ pulse equalizes the populations of the $\alpha$ and $\beta$ spin states.

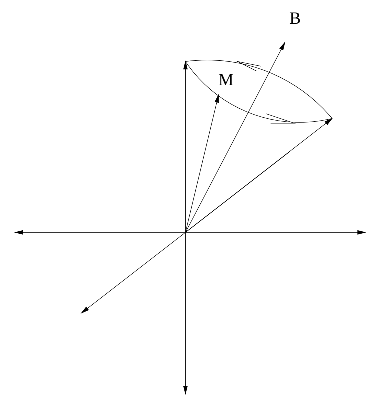


Figure 2.2: The panel shows the precession of the bulk magnetic moment around the net magnetic field. NMR spectrum of hydrogen nuclei. The different nuclei in the spectrum correspond to different chemical environment and scalar couplings the nuclei see.

Precession of the magnetization about the effective field in the rotating field is illustrated by in Figure 2.2. Following this rf pulse, the bulk magnetization precesses about the static magnetic field with a Larmor frequency $\omega_0$. This precessing magnetization during the so called acquisition period t generates the signal that is recorded by the NMR spectrometer. The signal is referred to as a *free-induction decay (FID)*. The FID is not just a purely oscillating signal as the magnetization will not evolve freely forever. Instead, due to the relaxation phenomenon, returns to the equilibrium state. BLOCH ET AL. (1946) provided two phenomological processes to account for this relaxation phenomenon. The first relaxation mechanism which accounts for the return of population difference back to the Boltzmann distribution level is called the *spin-lattice relaxation* and is characterized by

$$\frac{dM_z(t)}{dt} = R_1[M_0(t) - M_z(t)], \tag{2.3}$$

and a second process which is responsible for decay of transverse magnetization in x-y plane following a pulse is called the *transverse or spin-spin relaxation* and is also characterized by a first-order rate expression

$$\frac{dM_x(t)}{dt} = -R_2 M_x(t) \tag{2.4}$$

$$\frac{dM_y(t)}{dt} = -R_2 M_y(t), \tag{2.5}$$

where $R_2$ is the spin-spin relaxation rate constant and the corresponding time constant $T_2 = \frac{1}{R_2}$. Combining the relaxation equations with equations of free magnetic precession

2.1, we obtain the following equations also called the *Bloch's Model*

$$\frac{dM_z(t)}{dt} = \gamma(M(t) \times B(t))_z - R_1[M_0(t) - M_z(t)] \tag{2.6}$$

$$\frac{dM_x(t)}{dt} = \gamma(M(t) \times B(t))_x - R_2 M_x(t) \tag{2.7}$$

$$\frac{dM_y(t)}{dt} = \gamma(M(t) \times B(t))_y - R_2 M_y(t). \tag{2.8}$$

The free-precession Bloch equations in the rotating frame show that the FID can be described in terms of two components

$$\begin{aligned} M_x(t) &= M_0 \sin\theta \cos(\Omega t) \exp(-R_2 t) \\ M_y(t) &= M_0 \sin\theta \sin(\Omega t) \exp(-R_2 t), \end{aligned}$$

which can be combined in complex notation as

$$M^+(t) = M_x(t) + iM_y(t) = M_0 \sin\theta \exp(i\Omega t - R_2 t).$$

As a consequence of relaxation, the components of the bulk magnetization vector precessing in the transverse plane following a rf pulse are damped by the exponential factor $exp(-R_2 t)$. In practice, both parts of the complex signal are detected simultaneously by the NMR spectrometer as $s^+(t) = \lambda M^+(t)$. This complex time domain signal can be Fourier-transformed to produce a complex frequency-domain spectrum

$$S(\omega) = \int_0^\infty s^+(t) exp(-i\omega t) dt \tag{2.9}$$

$$= v(\omega) + iu(\omega), \tag{2.10}$$

where

$$v(\omega) = \lambda M_0 \frac{R_2}{R_2^2 + (\Omega - \omega)^2} \tag{2.11}$$

$$u(\omega) = \lambda M_0 \frac{\Omega - \omega}{R_2^2 + (\Omega - \omega)^2}. \tag{2.12}$$

The function $v(\omega)$ represents an absorptive Lorentzian signal, and the function $u(\omega)$ represents a dispersive Lorentzian signal. The real part of the complex spectrum $v(\omega)$ is normally displayed as the NMR spectrum. The one pulse experiment is displayed schematically in Figure 2.3.

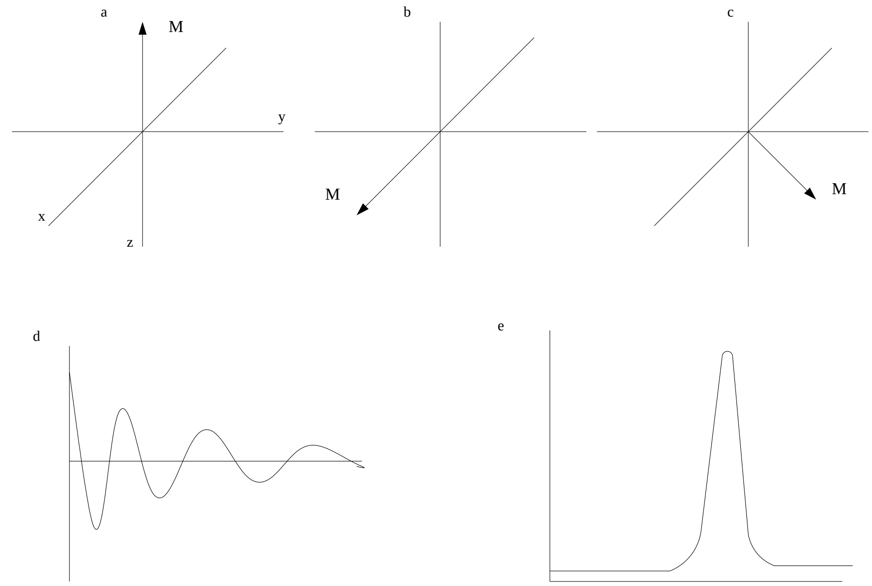


Figure 2.3: The panel shows the one pulse experiment. (a) At equilibrium the bulk magnetization is oriented in the $z$ direction. (b) The magnetization points along x axis following a 90 degree pulse with y phase (c) The magnetization precesses in the x-y plane (d) The FID signal (e) Real Part of the Fourier Transform of FID Signal.

### 2.1.2 Chemical Shift

The NMR spectrum not only depends on the applied magnetic field but also on the local environments of individual nuclei and hence the resonance frequency $\Omega$ in the FID spectrum (2.11) differs from the one predicted by (2.2). It is this difference in resonance frequencies that helps us distinguish between spin in different environment and is called *chemical shift*. The phenomenon of chemical shift arises because of motion of electrons induced by an external magnetic field generate secondary magnetic fields. Thus, the net magnetic field at the nucleus site is the result of applied magnetic field and the secondary fields produced by electron currents. This effect of secondary fields called nuclear shielding can enhance or oppose the main field. In general, the electronic charge distribution in a molecule is not isotropic and this shielding effect is described by a second rank tensor. However, in isotropic liquid solutions, collisions lead to rapid reorientation of the molecule, and consequently, of the shielding tensor. Under these circumstances, the effect of shielding on a particular nucleus can be accounted for by modifying (2.2) as

$$\Omega = -\gamma(1 - \sigma)B_0, \tag{2.13}$$

where $\sigma$ is a measure of the shielding. Until now, we described the dynamics of ensemble of isolated nuclear spins in external magnetic fields using the formalism developed by Bloch. However, nuclei in an molecule are not isolated but interact with each other. To study the evolution of coupled spin system one has to resort to tools from quantum statistical mechanics, which involves the theory of density matrices. In the next few sections, we will develop this theory. This formalism is also very important from control theoretic point of view as the final control problems will involve steering the density matrix and related questions. In the following section, we will review the basics of non-relativistic quantum mechanics with special consideration to the spin dynamics.

## 2.2 Postulates of Quantum Mechanics

In quantum mechanics, the state of the system is defined by a vector in a Hilbert space **H**. The vectors are denoted by the well known Dirac notation, $|\psi>$, and $<,>$ is used to denote the scalar product. The evolution in time of a quantum mechanical system is governed by the non-relativistic Schrödinger equation.

$$i\hbar\frac{d|\psi(t)>}{dt} = \mathcal{H}|\psi(t)> . \tag{2.14}$$

The operator $\mathcal{H}$ is the Hamiltonian of the system and incorporates the essential physics determining the evolution of the system. The Hamiltonian may be time dependent or independent. Units in which $\hbar = 1$ will be assumed and factors of $\hbar$ will not be written; thus

$$i\frac{d|\psi(t)>}{dt} = \mathcal{H}|\psi(t)> .$$

The solution to the above Schrödinger equation is called the wavefunction and it contains all the knowledge about the state of the system of interest. If the wavefunction is known, then all the observable properties of the system can be deduced. In language of quantum mechanics, every physically observable quantity A, has associated with it a Hermitian operator **A** which satisfies the eigenvalue equation

$$\mathbf{A}|\psi >= \lambda|\psi > .$$

The values taken on by the observable quantity then corresponds to the eigenvalues of the operator corresponding to the observable. Since the operators are Hermitian, the eigenvalues are always real and the eigenfunctions form a complete orthonormal set. Given the operator $\mathbf{A}$, let $\{|\psi_n>\}$ denote its eigenfunctions. Therefore, any vector in the Hilbert space can be expanded in terms of these orthonormal eigenfunctions by

$$|\psi(t)>=\sum_{n=1}^{N} c_n(t)|\psi_n>,$$

in which the $c_n(t)$ are complex numbers and may depend on time. The expectation value of the observable $A$ is denoted $<A>$. It is defined as the scalar product of $|\psi>$ and $\mathbf{A}|\psi>$ and is written as

$$<A>=<\psi|\mathbf{A}|\psi>.$$

Expanding the wavefunction in terms of eigenfunctions of $\mathbf{A}$, $|\psi>=\sum_{n=1}^{N} c_n|\psi_n>$, we have that,

$$<A>=\sum_{n=1}^{N} \|c_n\|^2\lambda_n, \tag{2.15}$$

where $\lambda_n$ is the eigenvalue corresponding to the eigenfunction $|\psi_n>$. In deriving (2.15), we have used the orthonormality of the eigenvectors $|\psi_n>$. The interpretation of the expectation value of an operator is the following. Given an ensemble of quantum system, with every element in state $|\psi>$, when a measurement is made on a member, the observed value corresponds to one of the eigenvalues $\lambda_k$ with probability $\|c_k\|^2$. Therefore, when we perform this measurement a sufficient number of times, the expected value will coincide with the expectation of the operator as defined in (2.15).
Now, consider two observables $A$ and $B$ represented by operators $\mathbf{A}$ and $\mathbf{B}$. Let the eigenfunctions and eigenvalues corresponding to $A$ be denoted by $\{|\psi_i>\}$ and $\{a_i\}$, respectively, and let the eigenfunctions and eigenvalues corresponding to $B$ be denoted by $\{|\phi_i>\}$ and $\{b_i\}$, respectively. If the state of the system is denoted by the vector $|\psi>=\sum_{i=1}^{N} c_i|\psi_i>$, then a measurement of the observable $A$ will yield one of the eigenvalues $a_k$ and, after the measurement the state of the system will collapse to the value $|\psi_k>$. We can now expand this state in terms of eigenfunctions of $B$, i.e. $|\psi_k>=\sum_{i=1}^{N} d_i|\phi_i>$. Now a measurement of the observable $B$ will yield one of the eigenvalues $b_k$ and after the measurement the state of the system will collapse to the value $|\phi_k>$. Thus, a measurement of $B$ disturbs the eigenstate of $A$. This forms the basis of the Heisenberg's uncertainty principle. Simultaneous measurement of $A$ and $B$ are only possible if they have the same eigenfunctions, that is if $\mathbf{AB}=\mathbf{BA}$ or equivalently $[\mathbf{A},\ \mathbf{B}]=0$.

Until now, we focussed attention on description and evolution of a general quantum state. In the next section, we describe the Hilbert state corresponding to spin $\frac{1}{2}$ particles and the associated quantum mechanical operators for the angular momentum.

## 2.3 Spin, Magnetic Moment and Angular Momentum

Angular momentum plays an important role in many control problems that are adequately described using classical dynamics. Bicycles and spinning projectiles are examples. Angular momentum also plays a major role in the description of NMR spectroscopy, where the central theme is the manipulation and detection of nuclear magnetic moments. It was discovered early on the development of quantum mechanics that the nuclear magnetic moments can be associated with operators occuring in the representation of the Lie algebra corresponding to the group of two-by-two unitary operators of unit determinant. The purpose of this section is to introduce and motivate the concept of spin operators and present various techniques for their manipulation.

Before introducing the quantum mechanical operators for the characterization of the internal angular momentum of elementary particles, we review the quantum analog of classical angular momentum and its operator description in quantum mechanics. For a particle of mass $m$ and position vector $r = (x, y, z)$ with respect to a fixed origin and having linear momentum $p = mv$, the angular momentum $L$ about the fixed origin is $L = r \times p$. The components of the angular momentum are therefore related to the linear momentum by

$$\begin{aligned} L_x &= yp_z - zp_y \\ L_y &= zp_x - xp_z \\ L_z &= xp_y - yp_x. \end{aligned}$$

In classical mechanics, angular momentum is a measure of rate of rotation with respect to the origin, thus angular momentum can also be thought of as generator of the rotation of the particle. In quantum mechanics, this angular momentum takes the form of an operator, the three components of which are described by

$$\begin{aligned} l_x &= -i(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}) \\ l_y &= -i(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}) \\ l_z &= -i(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}). \end{aligned}$$

This operator formulation of angular momentum in quantum mechanics arose out of the study of atoms which, in a simplistic view, can be described as a cloud of negatively charged electrons orbiting the positively charged nucleus. We now show this operator definition of angular momentum has close connections with rotations and explain how it can be viewed as a generator of rotation in the Hilbert space for the quantum mechanical system. If we denote the state of the particle in three dimensional space by the wave function $\psi(x, y, z)$, then the Hilbert state of the particle is an infinite-dimensional function space where $\psi(x, y, z)$ is a element. Let $U(R)$ denote the operator that corresponds to a rotation $R$ in this space. $U(R)$ can also be thought of as a representation of the rotation group in this space. One would want $U(R)$ to satisfy the following properties

$$U(R_1)U(R_2) = U(R_1R_2) \tag{2.16}$$
$$U(1) = I \tag{2.17}$$
$$(U(R)\psi)(r) = \psi(R^{-1}r), \tag{2.18}$$

where $\psi$ is an element of the Hilbert space and $I$ is the identity operator. Let us consider a infinitesimal rotation $R_z(\epsilon)$ around the $z$-axis which transforms $\hat{r} = (\hat{x}, \hat{y}, \hat{z})$ to $r = (x, y, z)$ given by

$$\begin{aligned} \hat{x} &\approx x + \epsilon y \\ \hat{y} &\approx y - \epsilon x \\ \hat{z} &\approx z. \end{aligned}$$

Thus, from equation (2.18), $U(R_z(\epsilon))\psi(r) = \psi(\hat{r})$, which leads to

$$\begin{aligned} U(R_z(\epsilon))\psi\ (x, y, z) &= \psi(\hat{x}, \hat{y}, \hat{z}) \\ U(R_z(\epsilon))\psi\ (x, y, z) &\approx \psi(x, y, z) + \epsilon(x\frac{\partial\psi(x, y, z)}{\partial y} - y\frac{\partial\psi(x, y, z)}{\partial x}) \\ U(R_z(\epsilon))\psi\ (x, y, z) &\approx (1 - i\epsilon l_z)\psi(x, y, z), \end{aligned}$$

where $l_z = -i(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x})$ is the angular momentum operator. Similarly $l_x$ and $l_y$ are generators for representation of rotations around $x$-and $y$-axis. Having motivated the operator definition for the classical angular momentum we now come to the concept of spin. Besides possessing an angular momentum associated with their orbital motions in space, most of the elementary particles including electrons, protons, and neutrons have an intrinsic angular momentum called *spin* in analogy with a spinning top. By convention, spin is measured in units of $\hbar$, for example electrons have spin $\frac{1}{2}$ and hence an intrinsic angular momentum of $\frac{\hbar}{2}$. The right way to think about spin is by associating the particles with internal degrees of freedom which comprise the spin space. The state of a particle with spin $\frac{1}{2}$ is characterized by a vector in a two-dimensional Hilbert space. The spin angular momentum can then be thought of as generator of rotations in this spin space. The particles of most interest to us are the spin $\frac{1}{2}$ particles, so we look in some depth at the form the angular momentum operators take for this situation. By analogy with the orbital angular momentum we will try to construct generators for irreducible representation of the rotation group on this two dimensional Hilbert space.

The construction process is a standard one found in physics text. Identify the two dimensional Hilbert space with the stereographic projection of a 2-sphere in $R^3$. Let $(x, y, z)$ denote the coordinates of the point on the sphere in $R^3$. Let us write its stereographic projection on the plane as a complex number $\zeta = a + ib$. Express $\zeta$ as ratio of two complex numbers $\zeta = \frac{\eta}{\xi}$ obeying the constraint $\|\eta\|^2 + \|\xi\|^2 = 1$. Thus, we have have the pair $(\eta, \xi)$ representing a unit vector in the two-dimensional Hilbert space. A easy calculation shows that, under the infinitesimal rotation $R_z(\epsilon)$ around the $z$-axis , $(x, y, z) \to (\hat{x}, \hat{y}, \hat{z})$,

$$\begin{aligned} \hat{x} &\approx x - \epsilon y \\ \hat{y} &\approx y + \epsilon x \\ \hat{z} &= z \end{aligned}$$

induces a transformation on the stereographic coordinates in accordance with

$$\begin{pmatrix} \hat{\eta} \\ \hat{\xi} \end{pmatrix} \approx (\mathbf{1} \; - i\frac{\epsilon}{2} I_z) \begin{pmatrix} \eta \\ \xi \end{pmatrix},$$

where

$$I_z = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

and $\mathbf{1}$ is the two-dimensional identity matrix. Thus, we see that $\sigma_z$ is the generator of the representation of rotations around the $z$-axis on a two-dimensional Hilbert space. Similarly, the generators for rotation around $x$ and $y$ axis are denoted by

$$I_x = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{2.19}$$

$$I_y = \frac{1}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{2.20}$$

The matrices ($I_x$, $I_y$, $I_z$) are the famous Pauli matrices, a set of generators for rotation in the two dimensional Hilbert space and basis for the Lie algebra of traceless skew Hermitian matrices $\mathfrak{su}(2)$. They obey the following commutation relations

$$[I_x\ I_y] = iI_z \ \ ; \ \ [I_y\ I_z] = iI_x \ \ ; \ \ [I_z\ I_x] = iI_y. \tag{2.21}$$

By convention the eigenvectors of $I_z$ are chosen to be the basis for the two dimensional spin space and correspond to spin $\frac{1}{2}$ and $-\frac{1}{2}$ respectively. These are the column vectors

$$|\alpha >= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \ \ ; |\beta >= \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{2.22}$$

We are now ready to give a quantum mechanical treatment of NMR spectroscopy. First notice that NMR is not performed on isolated quantum state but on ensembles consisting of the order of Avagadro's number. If all elements in that system have the same wavefunction, the ensemble is said to be in pure state. However, in practical applications, not all elements of the ensemble have the same wavefunction and, in that case, one says that the system is in a mixed state. Study of such an ensemble is of primary interest to us and, in the following section, we will develop the machinery of density matrices needed to treat such ensembles.

## 2.4 Quantum Ensembles

The phenomena of interest in NMR involve both stochastic and quantum effects in an essential way. Roughly speaking, stochastic effects, with limitations coming from quantum mechanics, determine the steady-state distribution of the energies for the individual subsystems, while the time evolution equations of quantum mechanics determine the short term transient effects seen when the distribution is not an equilibrium distribution. In this section, we provide a brief introduction to certain aspects of quantum statistical mechanics, culminating in the concept of a density matrix.

Suppose that we have an ensemble of identical, or nearly identical, systems satisfying the same Schrödinger equation, and suppose we choose an orthonormal basis in terms of which we expand the solutions of these equations. Then we can write the wave function for the j$^{th}$ element of the ensemble as

$$|\psi_j >= \sum_{k=1}^{M} c_{jk}|\phi_k > . \tag{2.23}$$

This same basis can be used for each element of the collection and we can take an empirical average over the ensemble. Define the *density matrix operator* by

$$\rho = \frac{1}{\mathcal{N}} \sum_{j=1}^{\mathcal{N}} |\psi_j >< \psi_j|^{\dagger},$$

where $\mathcal{N}$ represents the number of elements in the ensemble. If we express the operator in the chosen orthonormal basis, it takes the form

$$\rho_{kl} = \frac{1}{\mathcal{N}} \sum_{j=1}^{\mathcal{N}} c_{jk} c_{jl}^{\dagger}.$$

The off-diagonal elements of the density matrix

$$\rho_{rs} =< r|\rho|s >= c_r(t) c_s^*(t)$$

represent the coherent superposition of the eigenstate $|r>$ and $|s>$ in the sense that the time dependence and phase of the elements of the ensemble are correlated with respect to the orthogonal states $|r>$ and $|s>$. It is obvious by definition that $\rho$ is Hermitian. In particular, the value of the diagonal element $\rho_{jj}$ is the probability that the system will be found in state state $j$, consistent with the observation tr$\rho = 1$. A value of $\rho$ is said to represent a *pure state* if $\rho$ is of rank one. In this case, $\rho$ is unitarily equivalent to a matrix which is zero everywhere except for a one in the one-one entry and we may say that each element of the ensemble is in the same state. This may also be stated as saying that the system is in pure state if and only if

$$tr \rho^2 = tr \rho = 1.$$

The density matrix for a pure state is like a projection operator. For a mixed state, we have

$$tr \rho^2 < 1$$

and the density matrix is no more a projection operator.
Starting from the evolution equation for $\psi$, we see immediately that the evolution equation for $\rho$ takes the isospectral form

$$\dot{\rho} = -i[H, \rho]. \tag{2.24}$$

This differential equation, called *Liouville-von Neumann* equation or simply the density matrix equation, is of central importance for studying the dynamics of the quantum mechanical systems. Its formal solution is

$$\rho(t) = U(t)\rho(0)U(t)^{-1}, \tag{2.25}$$

where $U(t)$ is the unitary matrix that specifies the evolution equation

$$\dot{U}(t) = -iHU(t). \tag{2.26}$$

We now compute the expected value of an arbitrary observable operator for an ensemble. The expectation value $< A >$ of an operator $\mathbf{A}$ is given by

$$< A >= \frac{1}{\mathcal{N}} \sum_{j=1}^{\mathcal{N}} < \psi_j|A|\psi_j >,$$

which can also be written in terms of density matrix as

$$< A >= \sum_{r\ s} \rho_{rs} < \psi_s|\mathbf{A}|\psi_r >,$$

leading to the important expression

$$< A >= tr(A\rho). \tag{2.27}$$

Thus, the expectation value of an operator is found by taking the trace of the product of the observable operator and the density operator.

The density matrix in thermal equilibrium, at a temperature $T$, is given by

$$\rho_0 = \frac{1}{Z} exp \frac{-H\hbar}{kT},$$

where

$$Z = tr(exp \frac{-H\hbar}{kT})$$

is the partition function of the system. Because the Maxwell-Boltzmann analysis predicts the distribution of energies that would be found in an ensemble, it is most informative to organize the quantum superposition in terms of eigenstates associated with specific energies. That is, we focus on the operator associated with the Hamiltonian and identify the orthogonal basis as the orthonormal eigenfunctions $\psi_k$ satisfying

$$H\psi_k = E_k\psi_k,$$

where $E_k$ represents the observed eigenvalues. By evaluating $\rho_0$ in the eigenbasis of the Hamiltonian, one easily verifies that the probability distribution of the energy eigenstates $|\psi_k>$, $P_k = \rho_{kk}$., correctly describes the Boltzmann distribution

$$P_k = \frac{1}{Z} exp \frac{-E_k \hbar}{kT} \tag{2.28}$$

Having reviewed the essentials of quantum statistical mechanics, we now start a quantum mechanical treatment of the spin systems. Our final goal will be to view the problem of control of spin systems as a problem in nonlinear control theory. We begin with a single spin-system and present quantum mechanical version of Bloch equations and then we generalize to the case of more than one mutually coupled spins.

## 2.5 Quantum Mechanical NMR Spectroscopy

We saw in section 2.1.1 that, classically, the Hamiltonian of a massless particle with magnetic moment $\mu$ in a magnetic field $\vec{B}$ is

$$H = -\mu \cdot \vec{B}.$$

A consequence of this form of energy is that the particle experiences a torque given by

$$\tau = \mu \times \vec{B}$$

and this torque is equal to the rate of change of angular momentum

$$\frac{dL}{dt} = \mu \times \vec{B},$$

where $L$ is the angular momentum vector. The magnetic moment $\mu = \gamma L$ where $\gamma$ is the *gyromagnetic ratio* for the particle

$$\frac{dL}{dt} = \gamma L \times B$$

We saw in the section 2.1 that if we have a magnetic field $B_0$ pointing in the $z$-direction, let the magnetic moment make an angle $\theta$ with $B_0$ and $\theta \neq 0$, then the above equation describes the precession of $\mu$ around the $z$ axis with frequency $\omega_0 = -\gamma B_0$, also called the Larmor frequency. In quantum mechanics, a particle with spin in presence of a magnetic field evolves under the Hamiltonian operator

$$H = -\gamma \hbar \vec{B} \cdot I \tag{2.29}$$

$$H = -\gamma \hbar \left( B_x \cdot I_x + B_y \cdot I_y + B_z \cdot I_z \right), \tag{2.30}$$

where $\vec{B} = B_x\hat{i} + B_y\hat{j} + B_z\hat{k}$ the externally applied magnetic field and $I_x,\ I_y,\ I_z$ are the Pauli spin matrices. The usual setup for a NMR experiment consists of an ensemble of atomic nuclei with spin subjected to a stationary field (longitudinal field) $B_0$ in the $z$-direction and an oscillating field $B_1$ (transverse field) in the $x-y$ plane, where

$$\vec{B}_1 = B_1 u = B_1(\cos(\omega t + \phi)\ \hat{x} + \sin(\omega t + \phi)\ \hat{y}).$$

This term is to be thought of as a control term giving rise to a Hamiltonian operator of the form

$$H = -\gamma\ (B_0 I_z + B_1 \cos(\omega t + \phi) I_x + B_1 \sin(\omega t + \phi) I_y),$$

which can be written as

$$H = \omega_0 I_z + \omega_1(\cos(\omega t + \phi) I_x + \sin(\omega t + \phi) I_y),$$

where $\omega_0 = -\gamma B_0$ and $\omega_1 = -\gamma B_1$. All together, then viewing the magnetic field as a control in the system, we have

$$H = (\omega_0 I_z + u_1 I_x + u_2 I_y) = (\omega_0 I_z + \omega_1 (I_x \cos(\omega t + \phi) + I_y\ \sin(\omega t + \phi))). \tag{2.31}$$

The statistical description of the system given by the density matrix then evolves as

$$\dot{\rho} = -i[H, \rho].$$

Until now, we have only described the Hamiltonian operator for a single $\frac{1}{2}$ spin nuclei in an external magnetic field. We will now describe the evolution equation for a $N$-spin system, where these spins are coupled together. We present a quick review of the spin dynamics as applied to coupled spin systems. For more details see the standard references Ernst et al. (1987), Goldman (1988), Cavanagh et al. (1996).

In most applications of interest, NMR experiments involve molecules with more than one nuclei. Consider a molecule consisting of $N$ nuclei each with spin $\frac{1}{2}$. Then the joint wavefunction of the $N$ spins takes the form

$$\Psi = |m_1> \otimes |m_2> \cdots \otimes |m_N>$$

where each $|m_i>$ can be in one of 2 states $\alpha$ or $\beta$, the eigenstates of the Pauli $I_z$ matrix corresponding to spin $\frac{1}{2}$ and $-\frac{1}{2}$. This results in $2^N$ possible wavefunctions for the $N$ spin $\frac{1}{2}$ nuclei. We will use the following notation for the direct product of the two matrices

$$A \otimes B = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \otimes \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11}B & A_{12}B \\ A_{21}B & A_{22}B \end{bmatrix}$$

The wavefunction for two spin system in the product basis is

$$\begin{aligned}
\psi_1 &= |\alpha\alpha>= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\
\psi_2 &= |\alpha\beta>= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\
\psi_3 &= |\beta\alpha>= \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\
\psi_4 &= |\beta\beta>= \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix}.
\end{aligned}$$

Thus, the spin state space of a molecule with $N$ nuclei is a complex Hilbert space of dimension $2^N$. The state of the molecule is represented by a vector $|\psi>$ in this Hilbert space that evolves unitarily $|\psi(t)>= U(t)|\psi(0)>$ according to the Schrödinger equation, where

$$\dot{U} = -iHU, \; U \in SU(2^N).$$

Observe that $iH \in \mathfrak{su}(2^n)$, the tangent space of $SU(2^n)$. There is a very natural choice of basis for the $\mathfrak{su}(2^N)$ algebra, for describing the evolution in NMR ( Sørensen et al. (1983)). This is the set $\{iB_s\}$ where

$$B_s = 2^{q-1} \prod_{k=1}^{n} (I_{k\alpha})^{a_{ks}} \tag{2.32}$$

$\alpha = x, y, or\ z$ and

$$I_{k\alpha} = \mathbf{1} \otimes \cdots \otimes I_\alpha \otimes \mathbf{1} , \tag{2.33}$$

where $I_\alpha$ the Pauli matrix appears in the above expression only at the $k^{th}$ position, and $\mathbf{1}$ the two-dimensional identity matrix appears everywhere except at the $k^{th}$ position. $a_{ks}$ is 1 for $q$ of the indices and 0 for the remaining. Note that $q \geq 1$ as $q = 0$ corresponds to the identity matrix and is not a part of the algebra. As an example, for $n = 2$, the basis for $\mathfrak{su}(4)$ takes the form

$$\begin{array}{ll}
q = 1 & I_{1x}, I_{1y}, I_{1z}, I_{2x}, I_{2y}, I_{2z} \\
q = 2 & 2I_{1x}I_{2x}, 2I_{1x}I_{2y}, 2I_{1x}I_{2z} \\
 & 2I_{1y}I_{2x}, 2I_{1y}I_{2y}, 2I_{1y}I_{2z} \\
 & 2I_{1z}I_{2x}, 2I_{1z}I_{2y}, 2I_{1z}I_{2z}.
\end{array}$$

It is important to note that these operators are only normalized for $n = 2$ as

$$tr(B_r B_s) = \delta_{rs} 2^{n-2}.$$

The Hamiltonian operator for the $n$-spin system can then be expressed in terms of this set $\{B_s\}$. There are two parts to the Hamiltonian. The Zeeman part, which in the control theory corresponds to a drift in the system due to internal spin dynamics, and the effect of a static magnetic field. We write it as

$$H_d = \sum_{k=1}^{n} \omega_k I_{kz} + \sum_{k<l}^{n} J_{kl}(I_{kz}I_{lz} + I_{ky}I_{ly} + I_{kx}I_{lx}), \tag{2.34}$$

where $\omega_k = -\gamma_k B_0$ is the Larmor frequency for the $k^{th}$ nucleus, $\gamma_k$ the gyromagnetic ratio of the $k^{th}$ nucleus and $J_{kl}$ represent *scalar couplings* between the $k^{th}$ and $l^{th}$ spin. These are interactions between nuclear spins that are mediated through the electron cloud. The magnetic field that influences an individual nucleus is partly the result of an external magnetic field and partly the result of the magnetic fields that are generated indirectly by the other nuclei in the molecule, through a cloud of electrons. The indirect interaction process occurs when the spin of the cloud of electrons is influenced by other nuclei and whose own magnetic fields then influence the nucleus in question. There is also a direct interaction between magnetic nuclei called the *dipolar interaction* which are relatively less important in liquid phase NMR as these are averaged out to zero to first order in isotropic solutions.

The other part of the Hamiltonian is the *Rf Hamiltonian* expressed as

$$H_{rf} = \sum_{k=1}^{N} \omega_k^{rf} \{I_{kx} \cos(\omega t + \phi) + I_{ky} \sin(\omega t + \phi)\}, \tag{2.35}$$

where $\omega_k^{rf} = -\gamma_k B_1$ where $B_1$ is the magnitude of the applied transverse field. By appropriate choice of $B_1$, $\phi$ and $\omega$, this part of the Hamiltonian constitutes the control in the system by appropriate choice of $B_1$, $\phi$ and $\omega$. Thus writing this part of the Hamiltonian in a more control theoretic notation

$$H_{rf} = \sum_{k=1}^{n} u_{1k} I_{kx} + u_{2k} I_{ky}. \tag{2.36}$$

The state of the molecule evolves according to the Schrödinger equation

$$|\dot{\psi} > = -i\,(H_d + H_{rf})\,|\psi > . \tag{2.37}$$

If we consider the ensemble statistics described by the density matrix $\rho$, the evolution of $\rho$ is

$$\dot{\rho} = -i[H_d + H_{rf}, \rho]. \tag{2.38}$$

Let us take an example and see how the various terms of the Hamiltonian look like. The molecular systems we look at are the $I_nS$ spin systems. Each $I_nS$ spin system consists of $n$ similar nuclei of spin $\frac{1}{2}$ denoted, $I_1, \cdots, I_n$ (e.g $^1$H nuclear spins) and one different nuclei, also of spin $\frac{1}{2}$, denoted S (e.g $^{13}$C carbon nuclei). We now consider the concrete example, of the $I_2S$ system ($CH_2$ group). By abuse of notation, we use $I$ to denote spin operators for hydrogen and $S$ to denote spin operators for carbon. From Equation 2.34, the Hamiltonian $H_d$ for such a molecule is of the form

$$H_d = \omega_I \sum_{i=1}^{2} I_{kz} + \omega_s S_z + J \sum_{k=1}^{2} I_{kz} S_z,$$

where the following simplifying assumptions have been made: all coupling constants $J_{ij}$ have been assumed to be equal to $J$, a characteristic of isotropic liquids and coupling between equivalent spins $I_1$, and $I_2$ have been neglected. The state space of this three spin system is a $2^3$-dimensional Hilbert space. As before, we choose the basis of this space the elements $e^i \otimes e^j \otimes e^k$ where $i, j, k \in \{1, 2\}$ and $e^1, e^2$ are the eigenbasis of $\sigma_z$. In the above chosen basis

$$I_{1z} = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

where

$$I_{1z} = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}.$$

The part of the Hamiltonian which comes from the rf pulses takes the form $H_{rf} =$

$$\omega_I^{rf} \sum_{k=1}^{2} (I_{kx} \cos(\omega t + \phi_I) + I_{ky}(\sin \omega t + \phi_I)) + \omega_S^{rf} (S_x \cos(\omega t + \phi_S) + S_y \sin(\omega t + \phi_S)),$$

where for example

$$I_{2x} = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$I_{2x} = \frac{1}{2}\begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

A typical control task is to steer the density matrix from a given initial state $\rho(0)$ to some final state $\rho(T)$. The controls available for manipulating the the transition probabilities are the electromagnetic fields. The observations are the ensemble statistics of spins. The particular control pattern is the one that can control the Hamiltonian to the following extent. There is a drift $H_d$ and the chance to alter it as $H \to H_d \ + u \ H_u$. The Lie algebra generated by $H_u$ and $H_d$ defines the reachable state space.
There are a number of interesting system theoretic problems involving identification and control of systems described by this model. We can lump these into problems of the *spectroscopic type* in which the goal is to apply inputs that will allow the identification of some aspects of $H_d, H_1, ..., H_r$, and problems of the *transfer type* in which the object is to apply an input which drives the mean of the ensemble from one value to another. The general plan is to use $H_d$ in equation (2.31) to create a nonzero average for the magnetic moment and then to use $H_u$ to steer this average to a desired value.

We are now equipped with the necessary background to study the control problems arising in NMR spectroscopy. We first begin with the problem of using ideas from gradient flows on Lie Groups to compute bounds on best achievable signal to noise ratio in NMR spectroscopy.

## 2.6 Maximizing the Signal-to-Noise Ratio

The most important issue for an experimental spectroscopist is that of getting the system to generate a revealing signal having adequate signal-to-noise ratio. The starting point is a sample in thermal equilibrium. Using a suitable excitation, the sample is forced out of equilibrium so that it will emit a useful signal during a subsequent relaxation process. A measurement (observation), in situations described by basic quantum mechanics, results in an evaluation of a terms of the form

$$y(t) = < \psi | C | \psi >,$$

with $C$ being the Hermitian operator that corresponds to the variable being measured. One can not simultaneously measure both

$$y_1(t) = < \psi | C_1 | \psi >$$

and

$$y_2(t) = < \psi | C_2 | \psi >,$$

unless the operators $C_1$ and $C_2$ commute. On the other hand, when dealing with quantum ensembles, the problems associated with noncommutivity largely go away. This means, for example, it makes sense to measure simultaneously both the $x$-component and the y-component of the angular momentum of a spin ensemble. This can be interpreted as saying that one can measure the complex number

$$y(t) = \mathrm{tr}(I_x + iI_y)\rho.$$

Observe that the matrix $I_x + iI_y$ is not Hermitian. The problem of maximizing the signal strength is then the problem of steering $\rho$ to a value that maximizes the absolute value of this trace. In NMR spectroscopy this is the problem of interest where operator $I_x + iI_y$ corresponds to the measurement obtained using a procedure called quadrature detection. This problem is discussed at length in GLASER ET AL. (1998).

The process of forcing the sample out of its equilibrium state and into some state that will generate a measurable signal can then be thought of as an control problem and, if one is willing to be precise about the definition of the signal-to-noise ratio, an optimal control problem. As such, it is a problem of control of density matrix with evolution equation

$$\dot{\rho} = i[\rho, H_d + H_{rf}], \tag{2.39}$$

where $H_d$ is the the same as in equation (2.34)

$$H_d = \sum_{k=1}^{n} \omega_k I_{kz} + \sum_{k<l}^{n} J_{kl}(I_{kz}I_{lz} + I_{ky}I_{ly} + I_{kx}I_{lx}),$$

similarly $H_{rf}$ from (2.35)

$$H_{rf} = \sum_{k=1}^{n} \{ I_{kx}\ u_{1k} + I_{ky}\ u_{2k} \},$$

where $u_{1k}$ and $u_{2k}$ are the controls which are generally chosen to be sinusoids of appropriate frequency. In the language of control theory, $H_d$ is the drift field and $H_{rf}$ is the part of Hamiltonian one has complete control. A typical task is to drive the system from initial density matrix $\rho(t_0)$ to a final density matrix $\rho(t_f)$ so that the expectation value $< A > (t) = tr\,(A\rho(t))$ of some observable $A$ is maximized. This can be stated as following

**Problem Statement 1** *Find $U^*$ such that $\| tr\left(A^\dagger \rho(t)\right) \|$ is maximized, i.e.*

$$U^* = \text{argmax}_U\ \| tr(A^\dagger U \rho(0) U^\dagger) \|.$$

There are three basic problems to solve here.

1. We need to determine the $U^*$, also known as *optimal propogator*, that achieves the maximum transfer.

2. Given $U^*$, its need to be shown that that there exists controls in (2.39), which can generate $U^*$. This corresponds to the problem of controllability on the Lie Group $SU(2^n)$.

3. Finally, find control laws that accomplish these transfers and are optimal under a desired cost function.

   The most pertinent cost function for NMR problems is the time of transfer because, as we saw in Bloch equations, the system has relaxation and, if transfers are not accomplished rapidly, the decoherence effects due to relaxation become predominant. Thus, the need for shorter pulse sequences or in language of control theory, time optimal control control laws cannot be overemphasized. Sometimes, the energy spent in rf pulses during these transfers is also an important consideration, hence we will also derive optimal control laws with energy as the cost.

In the following section, we show how ideas from gradient flows on Lie groups can be used to solve the first problem. We will show later that, under appropriate conditions, it is possible to generate any unitary $U$ using the controls we have, The proof brings out the non-commutative aspects of the problem with direct connections to non-linear control theory.

## 2.7 Optimal Propogator

The most general problem relevant to NMR spectroscopy is the following.

**Problem Statement 2** Let $A = A_x + iA_y$ and $B = B_x + iB_y$ be $n \times n$ complex matrices, where $A_x, A_y, B_x, B_y$ are Hermitian. Find $U \in U(n)$ such that $\|tr(AUBU^\dagger)\|$ is maximized.

Problem 2 can be solved exactly if $A$ and $B$ are Hermitian operator (VON NEUMANN (1937), BROCKETT (1991), STOUSTRUP ET AL. (1995)). The solution is given by the following theorem.

**Theorem 1** Let $A$ and $B$ be $n \times n$ Hermitian operators, and let $U \in U(n)$ the space of $n \times n$ unitary matrices. Let

$$A = V \cdot \Sigma_1 \cdot V^\dagger \tag{2.40}$$

$$B = W \cdot \Sigma_2 \cdot W^\dagger, \tag{2.41}$$

where $\Sigma_1$ and $\Sigma_2$ are the diagonal matrices with eigenvalues arranged in descending order. Let $U^* = \operatorname{argmax} tr(AUBU^\dagger)$. Then $U^* = V \cdot W^\dagger$ and $tr(AU^*BU^{*\dagger}) = tr(\Sigma_1 \cdot \Sigma_2)$.

We now present gradient flows on unitary group that maximizes the expression

$$f(U) = tr(AUBU^\dagger). \tag{2.42}$$

For more details see BROCKETT (1991).
The gradient ascent flow $\dot{U} = \nabla f(U)$ for the above function takes the following form. Observe that along the curve $\dot{U} = \Omega U$, a change in $f$ is given by

$$\begin{aligned} \frac{df}{dt} &= tr(A\Omega UBU^\dagger - AUBU^\dagger\Omega) \\ &= tr(\Omega^T[A, UBU^\dagger]) \end{aligned}$$

and, therefore, $\nabla f(U) = [A, UBU^\dagger]\, U$. The flow

$$\dot{U(t)} = [A, UBU^\dagger]\, U$$

is the then the gradient ascent flow for the function (2.42). The qualitative features of the flow is summarized by the following theorem

**Theorem 2** Let $A$ and $B$ be $n \times n$ Hermitian matrices with eigenvalues $\lambda_1 > \lambda_2 > \ldots > \lambda_n$ and $\mu_1 > \mu_2 > \ldots > \mu_n$. Then the flow $\dot{U(t)} = [A, UBU^\dagger]\, U$ has $2^n n!$ equilibrium points of which exactly $2^n$ are local maxima, at which points the function $f(U) = tr(AUBU^\dagger)$ takes on the value $\lambda_1\mu_1 + \lambda_2\mu_2 + \ldots + \lambda_n\mu_n$.

We now look at the more general case of this problem when $A$ and $B$ in theorem are not Hermitian. This is of most interest to us in the framework of NMR spectroscopy. Let us

$$g(U) = \|tr(A^\dagger UBU^\dagger)\|^2, \tag{2.43}$$

where as before $U$ is unitary. We first present the gradient ascent equation for this function.

**Theorem 3** *Let $U \in U(n)$ the space of $n \times n$ unitary matrices, $g(U) = \|tr(A^\dagger UBU^\dagger)\|^2$, and $\rho = UBU^\dagger$. The gradient ascent flow for the function to within a scale factor is given by*

$$\dot{U} = \{tr(A^\dagger\rho)\ [\rho, A^\dagger] + tr(A\rho^\dagger)\ [\rho^\dagger, A]\}\ U.$$

**Proof:** Observe that

$$g(U) = tr(A^\dagger UBU^\dagger) tr(AUB^\dagger U^\dagger).$$

Along the curve $\dot{U} = \Omega U$, where $\Omega$ is a skew Hermitian matrix, we have

$$\begin{aligned} \frac{d\ g(U)}{dt} &= tr(A^\dagger UBU^\dagger)\ \{\ tr(A^\dagger\Omega UBU^\dagger) - tr(A^\dagger UBU^\dagger\Omega)\ \} \\ &\quad + tr(AUB^\dagger U^\dagger)\ \{tr(A\Omega UB^\dagger U^\dagger) - tr(AUB^\dagger U^\dagger\Omega)\ \} \\ \frac{d\ g(U)}{dt} &= tr(\Omega^\dagger, \{tr(A^\dagger\rho)\ [\rho, A^\dagger] + tr(A\rho^\dagger)\ [\rho^\dagger, A]\}). \end{aligned}$$

Therefore the gradient flow for $g(U)$ is

$$\dot{U} = \Omega U \tag{2.44}$$

$$\Omega = \{tr(A^\dagger\rho)\ [\rho, A^\dagger] + tr(A\rho^\dagger)\ [\rho^\dagger, A]\}, \tag{2.45}$$

where $[\cdot,\cdot]$ represents the commutator.

A concrete example of controlling density matrices is in coherence transfer in NMR (Weitekamp et al. (1982), Glaser and Quant (1996)) for the $I_nS$, where a typical control task is to transfer the initial density matrix

$$\rho(0) = S_x - iS_y$$

to the final state

$$\rho(t_f) = \sum_{k=1}^{n} I_{kx} - iI_{ky},$$

where the operators $S_x$, $I_{kx}$, etc have the usual meaning as described earlier. Of course the gradient flows will necessarily converge to a stationary value of the given function. Moreover, given the realities of roundoff noise, the flow will converge to a local minimum. Because there can be no assurance that any given local minimum occur, it may be necessary to use some stochastic technique (such as simulated annealing) to find the true optimal $\hat{U}$.

## 2.8 Unitary Controllability of Spin Systems

The fundamental issue of controllability of quantum mechanical systems has been of great interest (Ramakrishna et al. (1995) Butkovskiy and Samoilenko (1990) Huang et al. (1983)). In this section, we address the problem of controllability of spin systems in NMR. The question we want to answer is the following.
**Question** Given the unitary evolution of the density matrix $\rho(t) = U\rho(0)U^T$, where

$$\dot{U} = -i[H_d + H_{rf}]U, \; U \in SU(2^n), \tag{2.46}$$

is it possible to steer $U$ from $U(0) = I$ to some specified $U_F$ in finite time, i.e. is it possible to generate any unitary transformation?
It is shown that, under mild conditions, $\frac{1}{2}$ spin systems have the nice property that any unitary transformation can be generated or that the system (2.46) is controllable on the group $SU(2^n)$. These conditions are given in the following theorem which appears in the recent theses (Herbrüggen (1998)).

**Proposition 1** Given a network of $n$ mutually weakly coupled $\frac{1}{2}$ spins, if the Hamiltonians

$$\begin{aligned} H_d &= u_0 \sum_{k=1}^{n} \omega_k I_{kz} + \sum_{k<l}^{n} J_{kl} I_{kz} I_{lz} \\ H_{rf} &= \sum_{k=1}^{n} \omega_k^{rf} I_{kx} \, u_k \end{aligned}$$

in Equation (2.46) satisfy the constraints

- all $\omega_k$ are nondegenerate;
- all couplings $J_{kl}$ are resolved;
- each spin $k$ is accessible to selective rf-pulses;

then the system (2.46) is controllable on the group $SU(2^n)$.

The above proposition is a direct corollary of the following theorem. We firstpoint a few well known facts

- First of all, observe that the Pauli matrices, $iI_x$, $iI_y$, $iI_z$ form an orthogonal basis for $\mathfrak{su}(2)$ with respect to scalar product, $\langle \cdot, \cdot \rangle = tr\ (\cdot^\dagger, \cdot)$.

- Let $\{\alpha, \beta, \gamma\} \in \{x, y, z\}$. The following relations hold for the Pauli spin matrices

$$I_\alpha I_\beta = \frac{1}{4} I, \quad if \alpha = \beta$$
$$I_\alpha I_\beta = \epsilon_{\alpha\beta\gamma} I_\gamma, \quad if \alpha \neq \beta,$$

  where $\epsilon_{\alpha\beta\gamma} = \pm 1$ if $\alpha\beta\gamma$ are all different and are obtained from $xyz$ by a even permutation.

- Observe that if $A,\ B,\ C,\ D$ are arbitrary matrices of the same dimension, then $[A \otimes B,\ C \otimes D] = [A,\ C] \otimes (\ B \cdot\ D) + (\ C \cdot\ A) \otimes [B,\ D]$. This follows immediately from the identity

$$(A \otimes B)\ \ (C \otimes D) = (\ A\ C) \otimes (\ B\ D).$$

**Theorem 4** *Let $g_n$ denote the set of $\frac{n(n+3)}{2}$ skew-Hermitian matrices $\{iI_{kx},\ iI_{kz},\ iI_{kz}I_{lz} |\ k < l,\ l = 1,\ 2, \cdots,\ n\}$. The smallest Lie algebra containing the set $g_n$ is $su(2^n)$, the Lie algebra of all traceless skew-Hermitian $2^n \times 2^n$ matrices.*

**Proof:** We show that all the $4^n - 1$ basis elements (2.32) of $\mathfrak{su}(2^n)$ can be generated using from the set

$$g_n = \{I_{kx},\ I_{kz},\ I_{kz}I_{lz} | k < l,\ l = 1,\ 2, \cdots,\ n\}$$

by commutation. The proof is inductive and proceeds as follows

- For $n = 1$, $I_{kz}I_{lz}$ is absent. The proof is obvious and follows directly from the commutation relations for Pauli matrices given above.

- The set of $4 \times 4$ traceless skew-Hermitian matrices defines a 15 dimensional space. The "one spin" coherence terms (i.e. the elements realizable as one Pauli matrix tensored with identity) are clearly present–four of them are already present in $g_2$ and the remaining two of this type can be generated with a single bracket. The remaining nine terms can also be generated with a single bracket but now involving $I_{1z}I_{2z}$. Here is an example

$$[I_{1z}I_{2z}, I_{1x}] = iI_{1y}I_{2z}.$$

- By a $r$ spin term we mean elements in the algebra which can be expressed as a tensor product using exactly $r$ Pauli matrices and $n-r$ identity matrices. Our induction hypothesis is that, for n=s, we can generate all $4^s - 1$ basis terms and that of these there are
$$P_s = 3^r \binom{s}{r}$$
terms of $r$ spin type. We need to show that for $n = s+1$, we can generate all $4^{s+1} - 1$ basis terms and that of these $3^r \binom{s+1}{r}$ are of the $r$ spin type. Observe the new basis elements we add at this new stage are
$$g_{s+1} - g_s = \{I_{(s+1)x},\ I_{(s+1)z},\ I_{kz}I_{lz}|k < l,\ l = s+1\ \}.$$
Thus, again using the same argument as for $n = 2$, we can generate all the terms $3^r\ \binom{s}{r}$ of $r$ spin type from $3^{r-1}\binom{s}{r-1}$ terms of $r-1$ spin type already present in the $su(2^s)$ algebra
$$\sum_{r=1}^{s+1} 3^r \binom{s+1}{r} = 3 \cdot 4^s$$
adding these newly generated terms to terms previously present $3{\cdot}4^s + 4^s - 1 = 4^{s+1} - 1$, we see we obtain all the basis elements.

## 2.9 Conclusions

In this chapter, we have sought to bring the reader to the point he/she can appreciate the control theoretic aspects in NMR. Our hope is that ideas from system theory can shed light on the basic questions of interest when refining the current state of the art. Many interesting techniques and phenomenon in NMR such as spin echoes, spin-spin decoupling, homonuclear and heteronuclear transfers in multidimensional NMR spectroscopy (Ernst et al. (1987), Glaser and Quant (1996)) have very interesting system theoretic interpretations. The recent burst of interest in quantum computing involving delicate noncommutative effects suggest the potential for developing a unified framework for thinking about more general forms of computation as the output of a dynamical system.
Having demonstrated the unitary controllability of the spin $\frac{1}{2}$ spin systems, we now address the important question of synthesising a time-varying Hamiltonian which brings about the desired change in the initial state of the system. From a control theory perspective, this is a constructive controllability problem. During the last decade, the design of pulse sequences

for coherence or polarization transfer in pulsed coherent spectroscopy has received lot of attention. Algorithms for determining unitary bounds quantifying the maximum possible efficiency of transfer between Hermitian and non-Hermitian operators have been determined (GLASER ET AL. (1998)). There is therefore utmost need for design strategies for pulse sequences that can achieve these bounds. It is also desirable that the pulse sequences should be optimal with respect to some cost functions. The most pertinent cost function for NMR problems is the time of transfer because, as we saw in Bloch equations, the system has relaxation and if transfers are not accomplished rapidly, the decoherence effects due to relaxation become predominant. Sometimes the energy spent in rf pulses during these transfers is also an important consideration. Hence, we will also derive optimal control laws with energy as the cost.

# Chapter 3

# Optimal Control in NMR

In this chapter, we study the design of pulse sequences in NMR spectroscopy as a time optimal control problem on a compact Lie group. Though our primary interets lies in the unitary group $U(n)$, as the evolution in quantum mechanics is unitary, we have kept the discussion general enough to include other compact groups, including $SO(n)$, as these results are of interest from general control theory perspective. The need for shorter pulse sequences which accomplish desired transfers cannot be overemphasized especially when relaxation times are very short. The problems we want to look at have the following character. Suppose we are given a controllable right-invariant system on a Lie group. The question that we want to answer is, what is the minimum time required to steer the system from some initial point to a specified final point. In NMR spectroscopy and quantum computing, this translates to, what is the minimum time required to produce a unitary propogator. In particular, we assume that our controls are unrestricted. This is a good approximation for NMR spectroscopy as we can use high power hard NMR pulses. In some cases, we will try to explicitly specify the shape of the reachable set.

Our main results have to do with the time optimal coherence transfer for multiple spin systems. We give an analytical characterization of such time optimal pulse sequences in variety of transfers involving two spin-systems and some generalizations. We show, for example, what is the best possible inphase and antiphase coherence transfer possible in a given time. We also demonstrate the optimal transfer sequence improves the efficiency of the known isotropic mixing sequences by 40 percent. Also, we demonstrate the optimality of some known pulse sequences.

In non-relativistic quantum mechanics, the time evolution of a quantum system is defined through the time dependent Schroedinger equation

$$\dot{U(t)} = -iH(t)U(t), \quad U(0) = I,$$

where $H(t)$ and $U(t)$ are the Hamiltonian and the unitary displacement operators, respectively. Recall, we can split the Hamiltonian

$$H = H_d + \sum_{i=1}^{m} v_i(t)H_i,$$

where $H_d$ is the part of Hamiltonian that is internal to the system, and we call it the *drift Hamiltonian* and $\sum_{i=1}^{m} v_i(t)H_i(t)$ is the part of Hamiltonian that can be externally changed and is called the *control or rf Hamiltonian.* Equation for $U(t)$ dictates then the evolution of the density matrix according to

$$\rho(t) = U(t)\rho(0)U^{\dagger}(t).$$

The problem we are ultimately interested in is to find the minimum time required to transfer the initial state $\rho_0$ to a final state $\rho_F$. Thus we will be interested in computing the minimum time required to steer the system

$$\dot{U} = -i(H_0 + \sum_{i=1}^{m} u_i H_i)\ U \tag{3.1}$$

from identity, $U(0) = I$, to a final state $U_F$.

In the following section we establish a framework for studying such problems.

**Preliminaries**

Throughout this chapter, $G$ will denote a compact Lie group with a bi-invariant metric $<,>_G$ and $e$ its identity element. Let $K$ be a compact closed subgroup of $G$. We will denote by $L(G)$ the Lie algebra of the right-invariant vector fields on $G$ and similarly $L(K)$ the Lie algebra of the right-invariant vector fields on $K$. There is a one-to-one correspondence between these vector fields and the tangent spaces $T_e(G)$ and $T_e(K)$, which we denote by $\mathfrak{g}$ and $\mathfrak{k}$ respectively. Consider the direct sum decomposition $\mathfrak{g} = \mathfrak{m} + \mathfrak{k}$ such that $\mathfrak{m} = \mathfrak{k}^{\perp}$ with respect to the metric.

To fix ideas, let $G = SU(n)$ and $\mathfrak{g} = \mathfrak{su}(n)$ its associated Lie algebra of $n \times n$ traceless skew-Hermitian matrices. Then $< A, B >_G = tr(A^{\dagger}B),\ \ A, B \in \mathfrak{su}(n)$, represents a bi-invariant metric on $SU(n)$.

If we consider the (right) coset space $G/K = \{KU : U \in G\}$, then it is know that the $G/K$ (homogeneous space), admits the structure of a differentiable manifold. Define the origin of $G/K$ by $o = \pi(e)$. Given the decomposition $\mathfrak{g} = \mathfrak{m} + \mathfrak{k}$, there exists a neighborhood of $0 \in \mathfrak{m}$ which is mapped homeomorphically onto a neighborhood of the origin $o \in G/K$ by the mapping $\pi \circ exp|_{\mathfrak{m}}$. The tangent space $T_o(G/K)$ can be then identified with the vector subspace $\mathfrak{m}$. The geometry of homogeneous space will play an essential part in determining the shortest possible times for coherence transfers.

The Lie group $G$ acts on its Lie algebra $\mathfrak{g}$ by conjugation $Ad_G : \mathfrak{g} \to \mathfrak{g}$. This is defined as follows. Given $U \in G,\ X \in \mathfrak{g}$, then

$$Ad_U(X) = \frac{d\ U^{-1} exp(tX) U}{dt}|_{t=0}.$$

Once again to fix ideas if $G = SU(n)$ and $U \in G,\ A \in \mathfrak{su}(n)$, then $Ad_U(A) = U^T AU$.

The homogeneous space $G/K$ is a Riemannian Symmetric space, if the Lie algebra decomposition $\mathfrak{g} = \mathfrak{m} + \mathfrak{k}$ satisfies the commutation relation,

$$[\mathfrak{k}, \mathfrak{k}] \subset \mathfrak{k},\ \ [\mathfrak{m}, \mathfrak{k}] \subset \mathfrak{m}, [\mathfrak{m}, \mathfrak{m}] \subset \mathfrak{k}.$$

If $\mathfrak{h}$ is a sub-algebra of $\mathfrak{g}$, which is contained in $\mathfrak{m}$, then $\mathfrak{h}$ is abelian because $[\mathfrak{m}, \mathfrak{m}] \in \mathfrak{g}$. It is well known (Wolf (1984)) that

**Theorem 5** If $\mathfrak{h}$ and $\mathfrak{h}'$ be two maximal abelian sub-algebras of $\mathfrak{m}$. Then

1. There exists an element $X \in \mathfrak{h}$ whose centralizer in $\mathfrak{m}$ is just $\mathfrak{h}$.

2. There is an element $k \in K$ such that $ad_k(\mathfrak{h}) = \mathfrak{h}'$.

3. $\mathfrak{m} = \bigcup_{k \in K} ad_k(\mathfrak{h})$

Thus the maximal abelian sub-algebras of $\mathfrak{m}$ are all $ad_K$ conjugate, and in particular they have the same dimension. The dimension will be called the rank of the symmetric space $G/K$, and the maximal abelian sub-algebras of $\mathfrak{m}$ is called the *Cartan sub-algebras* of the pair $(\mathfrak{g}, \mathfrak{k})$.

**Assumptions:** Let $U \in G$, and let the control system

$$\dot{U} = [H_d + \sum_{i=1}^{m} v_i H_i]U,\ U(0) = e, \tag{3.2}$$

be given. We will assume that $\{H_d, H_1, \ldots, H_m\}_{LA} = \mathfrak{g}$, and since $G$ is compact, it follows that the system (3.2), is controllable (Jurdjevic and Sussmann (313-329)). Let $\mathfrak{k} = \{H_i\}_{LA}$, and $K = exp\{H_i\}_{LA}$ be the closed compact group generated by $\{H_i\}$. Given the direct decomposition $\mathfrak{g} = \mathfrak{m} + \mathfrak{k}$ where $\mathfrak{m} = \mathfrak{k}^{\perp}$ with respect to the bi-invariant metric $<,>$, let $H_d \in \mathfrak{m}$. We will assume that $Ad_K(\mathfrak{m}) \subset \mathfrak{m}$ and we say that the homogeneous space $G/K$ is reductive. All our examples will fall into this category.

**Notation:** Let $\mathcal{C}$ denote the class of all locally bounded measurable functions defined on the interval $[0, \infty)$ and taking value in $\mathbb{R}^m$. $\mathcal{C}[0, T]$ denotes its restriction on the interval $[0, T]$. We will assume throughout that in equation (3.2), $v = (v_1, v_2, \ldots, v_m) \in \mathcal{C}$. Given $v \in \mathcal{C}$, we denote the solution of equation (3.2)by $U(t)$ such that $U(0) = e$. If, for some time $t \geq 0$, $U(t) = U'$, we say that the control $v$ steers $U$ into $U'$, in $t$ units of time and $U'$ is attainable or reachable from $U$ at time $t$.

**Definition 1** *(***Reachable Set:***)* The set of all $U' \in G$ attainable from $U_0$ at time $t$ will be denoted by $R(U_0, t)$. Also we use the following notation

$$\begin{aligned} \mathbf{R}(U_0, T) &= \bigcup_{0 \leq t \leq T} R(U_0, t) \\ \mathbf{R}(U_0) &= \bigcup_{0 \leq t \leq \infty} R(U_0, t). \end{aligned}$$

We will refer to $\mathbf{R}(U_0)$ as the *reachable set* of $U_0$.

**Remark 1** From the right-invariance of control systems it follows that $R(U_0, T) = R(e, T)U_0$, $\mathbf{R}(U_0, T) = \mathbf{R}(e, T)U_0$ and $\mathbf{R}(U_0) = \mathbf{R}(e)U_0$. Note $\mathbf{R}(U_0, T)$ need not be a closed set, we use $\overline{\mathbf{R}(U_0, t)}$ to denote its closure.

**Definition 2** *(***Infimizing Time:***)* Given $U_F \in G$, we will define

$$\begin{aligned} t^*(U_F) &= \inf_t \{U \in \overline{\mathbf{R}(e, t)}\} \\ t^*(KU_F) &= \min_{k \in K} t^*(kU_F) \end{aligned}$$

and $t^*(U)$ is called the *infimizing time.*

We have two goals in this chapter, to characterize $\overline{\mathbf{R}(e, t)}$ and hence compute $t^*(U_F)$, the infimizing time for $U_F \in G$. To characterize the infimizing control sequence $v^i$ in (3.2), which in the limit achieve the transfer time $t^*(U_F)$ of steering the system (3.2) from identity $e$ to $U_F$.

The Lie Group which we will be most interested is $SU(2^n)$ the unitary group describing the evolution of $n$ coupled $\frac{1}{2}$ spins. Its Lie algebra $\mathfrak{su}(2^n)$ is the $4^n - 1$ dimensional space of traceless $n \times n$ skew-Hermitian matrices. Recall that the orthonormal basis which we will use for this space are expressed as tensor products of Pauli spin matrices. The basis for $\mathfrak{su}(2^n)$ takes the form $\{iB_s\}$, where

$$B_s = 2^{q-1} \prod_{k=1}^{n} (I_{k\alpha})^{a_{ks}}, \tag{3.3}$$

$\alpha = x, y, \text{or } z$, and

$$I_{k\alpha} = \mathbf{1} \ \otimes \cdots \otimes I_\alpha \otimes \mathbf{1} \tag{3.4}$$

where $I_\alpha$ the Pauli matrix appears in the above expression only at the $k^{th}$ position, and $\mathbf{1}$ the two dimensional identity matrix, appears everywhere except at the $k^{th}$ position. $a_{ks}$ is 1 for $q$ of the indices and 0 for the remaining. Note $q \geq 1$ as $q = 0$ corresponds to the identity matrix and is not a part of the algebra. With this background we discuss the general optimal time control problem.

## 3.1 Time Optimal Control

Given the system

$$\dot{U} = [H_d + \sum_{i=1}^{m} v_i H_i]U, \ U(0) = e.$$

Let $\mathfrak{k} = \{H_i\}_{LA}$ denote the Lie Algebra generated by $H_i$, and $K = exp\ \mathfrak{k}$ the compact closed subgroup of $G$ generated by the algebra $\mathfrak{k}$. The key observation is the following. If $U_F \in K$, then $t^*(U_F) = 0$. To see this, note that by letting $v$ in (3.2), be large, we can move on the subgroup $K$, as fast as we wish. In the limit as $v$ approaches infinity, we can come arbitrarily close to any point in $K$ in arbitrarily small time, with almost no effect from the term $H_d$. By same reasoning $t^*(U) = t^*(kU)$ for $k \in K$. Thus finding $t^*(U_F)$ boils down to finding the minimum time to steer the system (3.2) between the cosets $Ke$ and $KU_F$.

This is illustrated in the Figure 3.1, where the cosets $KU$ and $KV$ are depicted as arcs and the infimizing time path is shown with the dashed part of the curve depicting the fast motion within the coset and solid curve showing the drift part of the flow, also known as the evolution of couplings in NMR literature. The minimum time problem then corresponds to finding shortest path between these cosets or, in other words, the shortest path in the space $G/K$.

With this intuitive picture in mind, we now state a well known result.

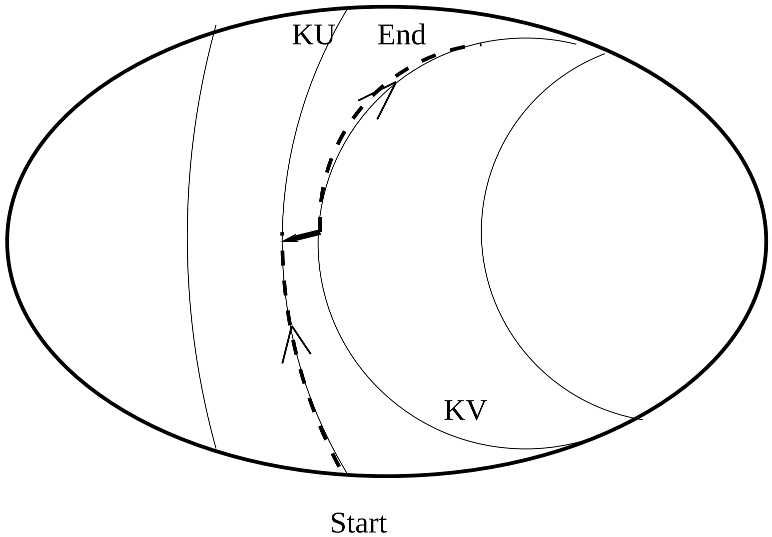


Figure 3.1: The panel shows the time optimal path between elements $U$ and $V$ belonging to $G$. The dashed line depicts the fast portion of the path corresponding to movement within the coset $KU$ and in traditional NMR language corresponds to the pulse and the solid line corresponds to the slow portion of the curve connecting different cosets and corresponds to evolution of the couplings.

**Lemma 1** For the right-invariant control system (3.2), the mapping $(v,t) \to U(t)$ from $\mathcal{C}(T) \times [0,T]$ into $G$ is continuous for each $U$ and each $T \geq 0$, if $\mathcal{C}(T)$ is given the topology of weak convergence.

We use this to show

**Lemma 2** For the right invariant control system in equation (3.2), $t^*(U_F) = t^*(KU_F)$.

**Proof:** All that needs to be proved here is that, if $U_F \in K$ then $t^*(U_F) = 0$. Since the Lie algebra $exp\ \{H_1, \ldots, H_m\}_{LA} = K$, given any $T > 0$, $\exists$, $\bar{v} \in \mathcal{C}(T)$ such that the solution $U(t)$ to

$$\dot{U} = [\sum_1^m \bar{v}_i H_i]U,\ U(0) = e$$

satisfies $U(T) = U_F$. Now consider the family of control system

$$\dot{U} = [H_d + \alpha \sum_{i=1}^m \bar{v}_i H_i]U,\ U(0) = e.$$

Rescaling time $\tau = \alpha t$, we obtain

$$\frac{dU}{d\tau} = [\frac{H_d}{\alpha} + \sum_{i=1}^{m} \bar{v}_i H_i]U, \ U(0) = e.$$

Observe that by Lemma (3.1) as $\alpha \to \infty$, $U(\tau)|_{\tau=T} = U_F$ or $\lim_{\alpha\to\infty} U(t)|_{t=\frac{T}{\alpha}} = U_F$. Therefore $U_F \in \overline{\mathbf{R}(e,T)}$, for all $T > 0$, implying $t^*(U_F) = 0$. **Q.E.D**

**Remark 2** The above observation will help us make a bridge between the problem of computing $t^*(U)$ and the problem of computing minimum length paths for a related problem which we now explain. Let $P \in G$. Associated with the control system (3.2), is the right invariant control system

$$\dot{P} = HP \tag{3.5}$$

where the control $H$ no longer belongs to the vector space but is restricted to an adjoint orbit i.e. $H \in Ad_K(H_d) = \{k^{-1}H_d k | k \in K\}$.

For the control system (3.5), we say that $KU_F \in B(U_0, t')$, if there exists a control $H[0, t']$, which steers $P(0) = U_0$ to $P(t') \in KU_F$ in $t'$ units of time. We use the notation

$$\mathbf{B}(U,T) = \bigcup_{0 \le t \le T} B(U,t).$$

From lemma , we see that $\mathbf{B}(U,T)$ is closed.

**Definition 3** Define the Minimum coset time

$$L^*(KU_F) = \inf_t \{KU_F \in \mathbf{B}(U,t)\}$$

**Theorem 6** *(***Equivalence theorem***)* The infimizing time $t^*(U_F)$ of steering the system

$$\dot{U} = [H_d + \sum_{i=1}^{m} u_i H_i]U$$

from $U(0) = e$ to $U_F$ is same as the minimum time $L^*(KU_F)$of steering the system

$$\dot{P} = HP, \ H \in Ad_K(H_d)$$

from $P(0) = e$ to $KU_F$.

**Proof:** Let $Q \in K$ and let us consider the flow

$$\dot{Q} = [\sum_{i=1}^{m} v_i H_i]Q, \ Q(0) = e. \tag{3.6}$$

Furthermore, let $P \in G$ such that $P$ evolves according to the equation

$$\dot{P} = (Q^{-1} H_d Q)\ P,\ \ P(0) = e. \tag{3.7}$$

Then, observe that

$$\frac{d(Q\ P)}{dt} = [H_d + \sum_1^m v_i H_i](QP), \quad Q(0)P(0) = e.$$

which is the same evolution equation as that of $U$ and since $U(0) = Q(0)P(0) = e$, therefore by the uniqueness theorem for the differential equations, $U(t) = Q(t)P(t)$. Therefore, given a solution $\hat{U}(t)$ of equation (3.2) with the initial condition $\hat{U}(0)$, there exists a unique curve $P_{\hat{U}}(t)$ defined through equations (3.6) and (3.7) as above. Observe that, if $\hat{U}(T) = U_F$ then it follows that $P_{\hat{U}}(T) \in KU_F$. Thus, if $U_F \in \overline{\mathbf{R}(e,T)}$, then $KU_F \in \mathbf{B}(e,T)$ which implies that $t^*(U_F) \geq L^*(KU_F)$.

To prove the equality observe that if $KU_F \in \mathbf{B}(e,T)$ , there exists a control $\bar{H}[0,T]$ such that the solution $P(t)$ to (3.5) satisfies $P(T) \in KU_F$. Let $v^k(t)$ be a family of control laws such that, for

$$\dot{Q}_k = [\sum_{i=1}^m v_i^k H_i]Q_k,\ \ Q_k(0) = e,$$

$\lim_{k\to\infty} \int_0^T \|\bar{H} - Q_k^{-1} H_d Q_k\| dt = 0$. Hence, using Lemma 3.1, we claim that the family of differential equations

$$\dot{P}_k = [Q_k^{-1}(t) H_d Q_k(t)]P_k,\ \ P_k(0) = e$$

satisfies $\lim_{k\to\infty} P_k(T) \in KU_F$. Therefore, $t^*(KU_F) \leq T$. Since the choice of $T$ was arbitrary, it follows $t^*(KU_F) \leq L^*(KU_F)$. Hence the proof. **Q.E.D**

**Remark 3** Since $\|H\| = 1$ in (3.5), we can also define $L^*(KU_F)$ as the infimizing value of $\int_0^T < \dot{P}, \dot{P} >^{\frac{1}{2}} dt$ of steering the system

$$\dot{P} = \gamma HP,\ \ \gamma > 0,$$

from $P(0) = e$ to $P(1) \in KU_F$.

We will now compute the expression $t^*$ based on the properties of the set $Ad_K(H_d)$. To recapitulate, we have a right invariant system (3.2) on the compact Lie group $G$ which admits a bi-invariant metric $<,>_G$. $K$ is a compact closed subgroup of $G$ and $\mathfrak{g}$ and $\mathfrak{k}$ represent the Lie-algebras of $G$ and $K$, respectively. Let $\mathfrak{g} = \mathfrak{m} + \mathfrak{k}$ be a direct sum decomposition, where $\mathfrak{m} = \mathfrak{k}^{\perp}$ with respect to the the metric $<,>_G$. The action $Ad_K(\mathfrak{m}) \subset \mathfrak{m}$ and $H_d \in \mathfrak{m}$. Let $S_{\mathfrak{m}}$ denote a unit sphere in $\mathfrak{m}$. We consider the following classification based on the qualitative nature of time optimal control laws.

1. *Riemannian Symmetric Case* In addition to above assumptions if we have the restriction $[\mathfrak{m}, \mathfrak{m}] \subset \mathfrak{k}$, then we are in the Riemannian symmetric case as described in the section 3. We make two distinctions here based on the qualitative nature of optimal time trajectories.

   - *Pulse-drift-pulse sequence*: This is the most simple case and a characteristic of single spin systems. In this case, the rank of the symmetric space is $G/K$ is 1.

   - *Chained Pulse-drift-pulse sequence*: In this case, the rank of the symmetric space $G/K$ is more than 1. This is a characteristic of two spin systems.

2. **Chatter sequence** In this case we no more have $G/K$ as a Riemannian symmetric case, i.e. $[\mathfrak{m}, \mathfrak{m}] \not\subset \mathfrak{k}$. This is a characteristic of more that two spin systems.

**Pulse-drift-pulse sequence**

We begin with the first case where rank of the symmetric space $G/K$ is one. Let $S_{\mathfrak{m}}$ denote a unit sphere in $\mathfrak{m}$, then it follows from theorem 5 that $Ad_K(H_d) = S_{\mathfrak{m}}$. In this case computing $t^*(U)$ reduces to finding geodesic distance on the homogeneous space $G/K$. To see this, consider the homogeneous space $P = G/K$ and let $\pi$ denote the natural projection map $\pi : G \rightarrow P$ such that $o = \pi(e)$. Given the bi-invariant metric $<,>$ on $G$, there is a corresponding left invariant metric, called the normal metric $<,>_n$, on the homogeneous space $G/K$ arising from the restriction of $<,>$ to $\mathfrak{m}$. Let $L_n(\gamma)$ represent the length of a curve $\gamma \in G/K$ under the normal metric. In case $Ad_K(H_d) = \mathfrak{m}$, there is one-to-one correspondence between the curves $\{\gamma(t) \in G/K | \gamma(0) = 0,\ \gamma(1) = \pi(U_F),\ L_n(\gamma[0,1]) = T\}$ and the trajectories of system 3.5, satisfying $\{\dot{P} = HP | P(0) = e, P(T) \in KU_F\}$. Therefore $L^*(KU)$ is the Riemannian distance between $o$ and $\pi(U)$ under the normal metric $<,>_n$. This is then computed in the following theorem, which is essentially a restatement about geodesics on the homogeneous space $G/K$, under the normal metric.

**Theorem 7** Let $G$, be a compact Lie group with a bi-invariant metric $<,>$, and $K$ be a closed subgroup. Let $\mathfrak{g}$ and $\mathfrak{k}$ denote their Lie algebras with the direct decomposition

$\mathfrak{g} = \mathfrak{m} + \mathfrak{k}$, $\mathfrak{m} = \mathfrak{k}^{\perp}$. Given the right invariant control system

$$\dot{U} = [H_d + \sum_1^m v_i H_i]U, \; U \in G, \; U(0) = e$$

where $v_i \in \mathbb{R}$, $H_d \in \mathfrak{m}$, $\{H_i\}_{LA} = \mathfrak{k}$ and $\{H_d, H_i\}_{LA} = \mathfrak{g}$. Suppose $G/K$ is a Riemannian symmetric space of rank one, then $t^*(U_F)$ is the smallest value of $\alpha > 0$ such that we can solve for $U_F = Q_1 exp(\alpha H_d) Q_2$ with $Q_1, Q_2 \in K$.

**Proof:** By the equivalence theorem $t^*(U_F) = L^*(KU_F)$, the minimum time for steering the system

$$\dot{P} = HP, \; H \in Ad_K(H_d),$$

from $P(0) = e$ to $KU_F$. Because $G/K$ is a Riemannian symmetric space of rank one, $Ad_K(\gamma H_d)\mathfrak{m}, \gamma \geq 0$. Therefore $L^*(KU_F)$, is the Riemannian distance between $o$ and $\pi(U)$ under the normal metric $<,>_n$. From KOBAYASHI AND NOMIZU (1969), the geodesics under the normal metric $<,>_n$ originating from $o$, take the form $\pi(exp(\tau H))$ for $H \in \mathfrak{m}$. If $U = Q_1 exp(tH_d) Q_2$ for $Q_1, \; Q_2 \in K$, then $\pi(U) = \pi(exp(t \; Q_2^{-1} H_d Q_2))$. To see this note that $U = (Q_1 Q_2) Q_2^{-1} exp(tH_d) Q_2 = (Q_1 Q_2) exp(t \; Q_2^{-1} H_d Q_2)$. Thus , from the above assertion, the geodesic connecting $o$ to $\pi(U)$ takes the form $\pi(exp(\tau \; Q_2^{-1} H_d Q_2))$ and has the length $L = t$. Hence the proof. **Q.E.D**

We now use illustrate these ideas through some examples

**Corollary 1** Let $U \in G = SU(2)$ and let $I_x = \frac{1}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $I_z = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ represent the Pauli spin matrices. Consider the unitary evolution of single spin system

$$\dot{U} = -i[I_z + uI_x]U$$

where the control $u \in \mathbb{R}$. Let $U_x = exp(-iI_x t)$ represent the one parameter subgroup generated by $I_x$ and $U_F = U_1 exp[-i\alpha I_x] U_2$ where $U_1, U_2 \in U_x$, then

$$t^*(U_F) = \|\alpha \; mod \; [-\pi, \pi]\|.$$

**Proof:** First note that the Lie algebra $\mathfrak{g} = \mathfrak{su}(2)$ has the decomposition $\mathfrak{m} = \{iI_y, iI_z\}$, $\mathfrak{k} = \{iI_x\}$, and $Ad_{U_x}(I_z) = \mathfrak{m}$. Observe that, if $U_F = U_1 exp[\alpha \Omega_z] U_2$, where $U_1, U_2 \in U_x$, then $U_F = U_1 exp[\beta \Omega_z] U_2$, where $\beta = \alpha \; mod \; [-\pi, \pi]$. The proof then follows directly from the Theorem 7 **Q.E.D.**

**Corollary 2** Let $\Theta \in G = SO(3)$, and let $\Omega_x = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $\Omega_z = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$ represent the generators of rotation around the $x-$ and $z-$axis. Consider the control system

$$\dot{\Theta} = [\Omega_z + u\Omega_x]\Theta,$$

where the control $u \in \mathbb{R}$. Let $\Theta_x = exp(\Omega_x t)$ represent the one parameter subgroup generated by $\Omega_x$ and let $\Theta_f = \Theta_1 exp[\alpha\Omega_x]\Theta_2$ where $\Theta_1, \Theta_2 \in \Theta_x$, then

$$t^*(U_F) = \|\alpha \; mod \; [-\pi, \pi]\|.$$

**Proof:** First note that the Lie algebra $\mathfrak{g} = \mathfrak{so}(3)$ has the decomposition $\mathfrak{m} = \{\Omega_y, \Omega_z\}$, $\mathfrak{k} = \{\Omega_x\}$ and $Ad_{\Theta_x}(\Omega_z) = \mathfrak{m}$. Observe that if $\Theta_f = Q_1 exp[\alpha\Omega_z]Q_2$, where $Q_1, Q_2 \in \Theta_x$, then $\Theta_f = \Theta_1 exp[\beta\Omega_z]\Theta_2$ where $\beta = \alpha \; mod \; [-\pi, \pi]$. The proof then follows directly from the Theorem 7. **Q.E.D**

We now generalize the above example to the following case where $G = SO(n)$, the group of $n \times n$ orthogonal matrices. The Lie algebra $\mathfrak{g} = \mathfrak{so}(n)$, is the the set of $n \times n$ skew symmetric matrices. The bi-invariant metric on $G$ is $< \Omega, \Omega >= tr(\Omega^T\Omega)$. Consider the following decomposition of $\mathfrak{g} = \mathfrak{m} + \mathfrak{k}$, where $\mathfrak{m}$ consists of skew-symmetric matrices which are zero except the first row and column and $\mathfrak{k}$ consists of skew-symmetric matrices which are zero in the first row and column. Observe that $\mathfrak{k}$ generates the subgroup $SO(n-1)$. Then we have the following corollary

**Corollary 3** Let $\Theta \in G = SO(n)$ and let the control system

$$\dot{\Theta} = [\Omega_d + \sum_1^m v_i\Omega_i]\Theta, \; \Theta(0) = I$$

be given, where $\Omega_d \in \mathfrak{m}$, $\Omega_i \in \mathfrak{k}$, and $u_i \in \mathbb{R}$. Suppose that $K = exp\{\Omega_i\}_{LA} = SO(n-1)$. Given $\Theta_f = \Theta_1 exp(\alpha\Omega_d)\Theta_2$, where $\Theta_1, \; \Theta_2 \; \in K$, then

$$t^*(\Theta_f) = \|\alpha \; mod \; [-\pi, \pi]\|.$$

**Proof:** Observe that $Ad_K(\Omega_d) = \mathfrak{m}$ and hence the proof is on the same lines as Corollary 2.

### Chained Pulse-drift-pulse sequence

Let us now consider the second case in our classification scheme. Now we assume that the rank of the Riemannian symmetric space $G/K$ is greater than one.

**Notation:** Given the decomposition $\mathfrak{g} = \mathfrak{m} + \mathfrak{k}$, let $\mathfrak{h} \subset \mathfrak{m}$ be the maximal abelian subalgebra containing $H_d$. We use the notation $\Delta_{H_d} = \mathfrak{h} \bigcap \in Ad_K(H_d)$ to denote the maximal commuting set contained in the adjoint orbit of $H_d$. We denote the span of this set by $Sp(\Delta_{H_d}) = \{\sum_{i=1}^{n} \beta_i X_i | \beta_i \geq 0, X_i \in \Delta_{H_d}\}$.

**Theorem 8** Let $G$, be a compact matrix Lie group and $K$ be a closed subgroup with $\mathfrak{g}$ and $\mathfrak{k}$ their Lie algebras with the direct decomposition $\mathfrak{g} = \mathfrak{m} + \mathfrak{k}$, $\mathfrak{m} = \mathfrak{k}^{\perp}$. Given the right invariant system

$$\dot{U} = [H_d + \sum_{1}^{m} v_i H_i] U, \; U \in G, \; U(0) = e$$

where $v_i \in \mathbb{R}$, $H_d \in \mathfrak{m}$, $\{H_i\}_{LA} = \mathfrak{k}$. Suppose $G/K$ is a Riemannian symmetric space and $Sp(\Delta_{H_d}) = \mathfrak{h}$, then $t^*(U_F)$ is the smallest value of $\sum_i \beta_i$, for $\beta_i \geq 0$, such that we can solve

$$U_F = Q_1 \; exp(\sum_{i=1}^{m} \beta_i X_i) Q_2,$$

where $Q_1, \; Q_2 \in K$ and $X_i \in \Delta_{H_d}$.

**Proof:** Recall, from Theorem 6, that if we consider the control system $\dot{P} = \gamma H P, \; P \in G$, where $H \in Ad_K(H_d)$, and let $L^*(KU_F)$ be the optimal cost $\int_0^1 < \dot{P}, \; \dot{P} >^{\frac{1}{2}} dt$ of steering the system from $P(0) = e$ to $P(1) \in KU_F$, then $t^*(U_F) = L^*(KU_F)$. To compute $L^*(KU_F)$, we derive the first order necessary conditions by applying the maximum principle of Pontryagin. We represent the linear functional on $\dot{P}$ as $\phi_Y(\dot{P}) = tr(Y\dot{P})$ with $PY \in \mathfrak{m}$. The Hamiltonian is then

$$h(P, Y, H, \gamma) = \gamma tr(YHP) + \frac{1}{2}\gamma^2 tr(H^T(t)H(t)).$$

Since $\|H\| = 1$, the Hamiltonian takes the form

$$h(P, Y, H, \gamma) = \gamma tr(YHP) + \frac{1}{2}\gamma^2.$$

Then $Y$ satisfies the equation

$$\begin{aligned} \dot{Y}(t) &= -Y(t)(H(t)) \\ tr(dH \; PY) &= 0 \\ \gamma &= -tr(YHP). \end{aligned}$$

Observe that $H = Q^{-1} H_d Q$, where $Q \in K$ and, therefore $dH = [A, H]$, where $A \in \mathfrak{k}$, implying that

$$tr(A[H, PY]) = 0.$$

Since $A \in \mathfrak{k}$, is arbitrary we have

$$[H, PY] \in \mathfrak{m}. \tag{3.8}$$

Let $M = PY$. The differential equation for $M$ is

$$\dot{M} = \gamma^2[H,\ M]. \tag{3.9}$$

Since $H \in \mathfrak{m}$ and $M \in \mathfrak{m}$ the condition $[\mathfrak{m}, \mathfrak{m}] \in \mathfrak{k}$ implies that if (3.8) holds then $[H, M] = 0$ and, from (3.9), $\dot{M} = 0$. Therefore the optimal $H(t)$ satisfies $[H(t), M(0)] = 0$ with $\gamma = tr(H(t)M(0))$ a constant. Let us now characterize the optimal $H(t)$.

Observe from the theorem 5 that $U_F = Q_1 exp(\sum_{i=1}^{m} a_i Y_i)$, where $Q_1 \in K$ and $Y_i \in Ad_Q(\Delta_{H_d})$, for some $Q \in K$. Note for $M(0) = \sum_{i=1}^{m} Y_i$ and $H(t) = \sum_{i=1}^{m} b_i(t) Y_i$, where at any time $t$, all but one $b_i$ (nonzero $b_i$ is equal to one), are zero, all the optimality conditions are met. The corresponding time optimal trajectory takes the form $P(t) = \Pi_{i=1}^{m} exp(\int_0^t b_i(t) Y_i)$. For $T = \sum_{i=1}^{n} a_i$, and $\int_0^t b_i(t) = a_i$, the time optimal curve satisfies $P(T) \in KU_F$. Hence the proof follows. **Q.E.D**

We now consider the application of the above theorem to find optimum time optimal pulse sequence for generating a unitary propogator in the 2 spin system. We consider the heteronuclear two spin case where by going in a rotating frame we can reduce the free precession part of the Hamiltonian to a coupling evolution.

**Theorem 9** Let $U \in SU(4)$. Consider the evolution of heteronuclear $IS$ spin system described by the following equation

$$\dot{U} = -i(\ H_d + \sum_{i=1}^{4} u_i H_i\ )U,$$

where

$$\begin{aligned} H_d &= 2\pi J I_z S_z \\ H_1 &= 2\pi I_x \\ H_2 &= 2\pi I_y \\ H_3 &= 2\pi S_x \\ H_4 &= 2\pi S_y. \end{aligned}$$

Let $K = SU(2) \times SU(2)$ be the subgroup generated by $\{H_i\}_{i=1}^{4}$. Given $U_F \in SU(4)$, if

$$\tilde{T}(U_F) = \arg \min_{\sum_{i=1}^{3} \alpha_i} \{U_F = Q_1 exp(\alpha_1 IzSz + \alpha_2 IxSx + \alpha_3 IySy) Q_2 |\ Q_1,\ Q_2 \in K, \alpha_i > 0\}.$$

Then $t^*(U_F) = \tilde{T}(U_F)$.

**Proof:** Consider the decomposition of $\mathfrak{g} = \mathfrak{m}+\mathfrak{k}$, where $\mathfrak{m} = span\{I_\alpha S_\beta\}$, $\mathfrak{k} = span\{I_\alpha,\ S_\beta\}$, and $(\alpha, \beta) \in (x, y, z)$. Then, observe that $[\mathfrak{m}, \mathfrak{m}] \in \mathfrak{k}$, $[\mathfrak{m}, \mathfrak{k}] \in \mathfrak{m}$, and $[\mathfrak{k}, \mathfrak{k}] \in \mathfrak{k}$. Furthermore, since $\Delta_{IzSz} = \{IzSz, IxSx, IySy\}$, and also $Ad_K(Sp(\Delta_{IzSz})) = \mathfrak{m}$. Thus the above example satisfies all the conditions of the theorem, and hence the proof follows. **Q.E.D**

Now we address the question of maximum possible achievable transfer in some given time $T$. We will look at the inphase and antipahse transfers in the two-spin systems and give expressions for maximum transfer efficiencies.

## 3.2 Optimal Transfer Efficiency in Two-Spin Systems

In this section, we will focus on the following question. Given a spin system with a known internal Hamiltonian, what is the maximum transfer possible between a given initial and final state, in some specified time $T$ and what is the pulse sequence that accomplishes this transfer. Our results will be the consequence of previous theorems. In the following theorems we give fundamental bounds on transfer efficiencies for inphase and antiphase transfer experiments in heteronuclear two-spin system. We start by some definitions

**Definition 4 Transfer Efficiency**Consider the evolution of the density matrix

$$\rho(t) = U(t)\rho(0)U^\dagger(t),$$

where

$$\dot{U} = -i(\ H_d + \sum_{i=1}^{m} u_i H_i\ )U,\ U(0) = I.$$

Define the transfer efficiency $\eta(t)$ to some given target operator $F$ as

$$\eta(t) = \|tr(F^\dagger U(t)\rho(0)U^\dagger(t))\|.$$

The questions we are interested in answering are the following

**Problem Statement 3 Maximum Inphase Transfer**Given the unitary evolution of the heteronuclear two spin system,

$$\dot{U} = -i(\ H_d + \sum_{i=1}^{4} u_i H_i\ )U,\ U(0) = I,$$

where $H_d = 2\pi J I_z S_z,\ H_1 = 2\pi I_x,\ H_2 = 2\pi I_y,\ H_3 = 2\pi S_x,\ H_4 = 2\pi S_y$. Let $\rho(0) = S_x+iS_y$ and the target operator be $F = I_x + iI_y$. Find the maximum achievable transfer efficiency $\eta^*(t)$.

**Problem Statement 4 Maximum Antiphase Transfer** Consider the unitary evolution of the heteronuclear two spin system,

$$\dot{U} = -i(\ H_d + \sum_{i=1}^{4} u_i H_i\ )U, \quad U(0) = I.,$$

where $H_d = 2\pi J I_z S_z,\ H_1 = 2\pi I_x,\ H_2 = 2\pi I_y,\ H_3 = 2\pi S_x,\ H_4 = 2\pi S_y$. Let $\rho(0) = S_x + iS_y$ and the target operator be $F = S_z(I_x + iI_y)$. Find the maximum achievable transfer efficiency $\eta^*(t)$.

**Lemma 3** Given the vector $p = \begin{bmatrix} 1 \\ i \\ 0 \end{bmatrix}$ and let $\Sigma$ be a real diagonal matrix

$$\Sigma = \begin{bmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{bmatrix},$$

If $a_i \geq a_j \geq a_k$, where $\{i, j, k\} \in \{1, 2, 3\}$ and let $U, V \in SO(3)$. The maximum value of

$$\|p^\dagger U \Sigma V p\|^2 = a_i^2 + a_j^2.$$

**Proof:** Let $\Sigma V p = b + ic$ where $b, c \in R^3$. First observe that the maximum value of $\|p^\dagger U(b + ic)\|^2$ is obtained when $b \perp c$ and the maximum value is $\|b\|^2 + \|c\|^2$. Now observe that maximum value of $\|\Sigma V p\|^2 = \|b\|^2 + \|c\|^2$ is $a_i^2 + a_j^2$ and when this is obtained $b \perp c$. Hence the proof follows **Q.E.D.**

**Theorem 10** Consider the density matrix evolution for the heteronuclear $IS$, spin system

$$\dot{\rho} = -i[\ H_d + \sum_{i=1}^{4} u_i H_i, \rho],$$

where $H_d = 2\pi J I_z S_z,\ H_1 = 2\pi I_x,\ H_2 = 2\pi I_y,\ H_3 = 2\pi S_x,\ H_4 = 2\pi S_y$. Let $\rho(0) = S_x + iS_y$ and $F = I_x + iI_y$ then the maximum achievable transfer

$$\eta^*(t) = \|tr(F^\dagger U(t)\rho(0)U^\dagger(t))\| = sin(J\pi\alpha_1)sin(J\pi\alpha_2),$$

where $\alpha_1 + 2\alpha_2 = t$ and $\tan(J\pi\alpha_1) = 2\tan(J\pi\alpha_2)$.

**Proof:** Let

$$\Lambda(t) = exp(i2\pi J(\alpha_1 IzSz + \alpha_2 IxSx + \alpha_3 IySy).$$

From theorem 8

$$U(t) \in \{Q_1 \Lambda(t) Q_2 | \ Q_1, \ Q_2 \in K \ \ \alpha_i > 0, \ \sum_{i=1}^{3} \alpha_i \leq t\}.$$

Let $S = exp\{iS_x, iS_y, iS_z\}$ and $I = exp\{iI_x, iI_y, iI_z\}$. By definition $K = S \times I$. In the expression for

$$\eta^*(t) = \|tr(F^\dagger U(t)\rho(0)U^\dagger(t))\|,$$

since $\rho(0)$ commutes with $I$ and $F$ commutes with $S$, it suffices to restrict $Q_1$ and $Q_2$ to $S$ and $I$ respectively. Let $P_I$ denote the projection on the subspace generated by $\{I_x, I_y, I_z\}$,then a simple computation yields that

$$\begin{aligned} P_I(\Lambda S_x) &= \sin(J\pi\alpha_2)\sin(J\pi\alpha_3)I_x \\ P_I(\Lambda S_y) &= \sin(J\pi\alpha_1)\sin(J\pi\alpha_3)I_y \\ P_I(\Lambda S_z) &= \sin(J\pi\alpha_2)\sin(J\pi\alpha_3)I_z \end{aligned}$$

Since $\{Ix, I_y, I_z\}$ forms an orthogonal pair, we can rewrite

$$\eta(t) = \|tr(F^\dagger Q_1 \Lambda Q_2 \rho(0) Q_2^\dagger \Lambda^\dagger Q_1^\dagger)\|.$$

as

$$\eta(t) = \|p^\dagger U \Sigma V p\|^2,$$

where $p = [1 \ i \ 0]^T$, and

$$\Sigma = \begin{bmatrix} \sin(J\pi\alpha_2)\sin(J\pi\alpha_3) & 0 & 0 \\ 0 & \sin(J\pi\alpha_1)\sin(J\pi\alpha_3) & 0 \\ 0 & 0 & \sin(J\pi\alpha_1)\sin(J\pi\alpha_2) \end{bmatrix}$$

and $U$ and $V$ are real orthogonal matrices. Using the result of lemma (3) we get that if $\sin(J\pi\alpha_1) \geq \sin(J\pi\alpha_2) \geq \sin(J\pi\alpha_3)$, then the maximum value of

$$\eta(t) = \sin(J\pi\alpha_1)\sin(J\pi\alpha_2) + \sin(J\pi\alpha_1)\sin(J\pi\alpha_3).$$

Now maximize the above expression with respect to $\alpha_1, \alpha_2, \alpha_3$ to get the above result. **Q.E.D**

The optimal transfer curve is plotted in comparison with the transfer achieved using the isotropic mixing Hamiltonian in the figure 3.2

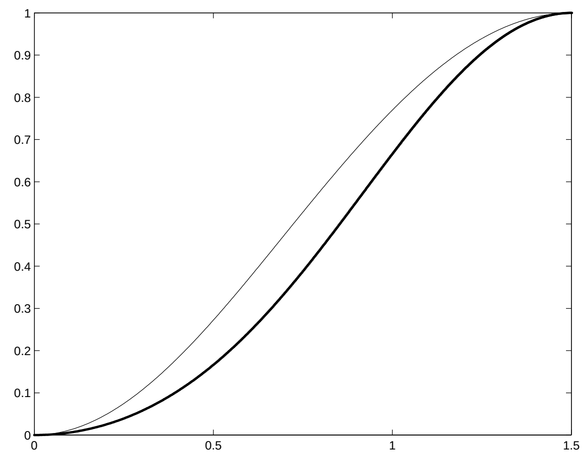


Figure 3.2: The panel shows the comparison between the best achievable transfer and the transfer achieved using the isotropic mixing Hamiltonian (bold curve) for the inphase transfer in 2 spin case. On X axis is plotted time in units of $1/J$.

**Theorem 11** Consider the density matrix evolution for the heteronuclear $IS$, spin system

$$\dot{\rho} = -i[\ H_d + \sum_{i=1}^{4} u_i H_i, \rho],$$

where $H_d = 2\pi J I_z S_z,\ H_1 = 2\pi I_x,\ H_2 = 2\pi I_y,\ H_3 = 2\pi S_x,\ H_4 = 2\pi S_y$. Let $\rho(0) = F_z S^+ = F_z(S_x + iS_y)$ and $F = I^+ = I_x + iI_y$ then the maximum achievable transfer for $t \leq 1/2J$, is

$$\eta^*(t) = \|tr(F^\dagger U(t)\rho(0)U^\dagger(t))\| = sin(J\pi t).$$

**Proof:** Let

$$\Lambda(t) = exp(i2\pi J(\alpha_1 IzSz + \alpha_2 IxSx + \alpha_3 IySy).$$

From theorem 8

$$U(t) \in \{Q_1\Lambda(t)Q_2 |\ Q_1,\ Q_2 \in K\ \alpha_i > 0,\ \sum_{i=1}^{3}\alpha_i \leq t\}.$$

Let $S = exp\{iS_x, iS_y, iS_z\}$ and $I = exp\{iI_x, iI_y, iI_z\}$. By definition $K = S \times I$. In the expression for

$$\eta(T) = \|tr(F^\dagger Q_1 \Lambda Q_2 \rho(0) Q_2^\dagger \Lambda Q_1^\dagger)\|$$

let $Q_2 = Q_{2I} \times Q_{2S}$, where $Q_{2I} \in I$ and $Q_{2S} \in S$. Let the optimum $Q_{2S}{}^*$, be such that

$$Q_{2S}{}^* \rho(0) Q_{2S}{}^{*\dagger} = Q_{2S}{}^* I_z S^+ Q_{2S}{}^{*\dagger} = a_z I_z S^+ + a_y I_y S^+ + a_x I_x S^+.$$

Denote

$$\begin{aligned}
\eta_z &= \|tr(F^\dagger Q_1 \Lambda Q_2 (I_z S^+) Q_2^\dagger \Lambda^\dagger Q_1^\dagger)\| \\
\eta_y &= \|tr(F^\dagger Q_1 \Lambda Q_2 (I_y S^+) Q_2^\dagger \Lambda^\dagger Q_1^\dagger)\| \\
\eta_x &= \|tr(F^\dagger Q_1 \Lambda Q_2 (I_x S^+) Q_2^\dagger \Lambda^\dagger Q_1^\dagger)\|
\end{aligned}$$

Then observe

$$\eta(t) \leq a_z \eta_z + a_y \eta_y + a_x \eta_x.$$

We now claim that maximum value of $\eta_z(t)$ is $sin(J\pi t)$ for $t \leq 1/(2J)$. Let $P_{I_z S}$ denote the projection on the subspace generated by $\{I_z S_x, I_z S_y, I_z S_z\}$,then a simple computation yields that

$$\begin{aligned}
P_{I_z S}(\Lambda S_x) &= -\cos(J\pi\alpha_2)\sin(J\pi\alpha_3) I_z S_y \\
P_{I_z S}(\Lambda S_y) &= \sin(J\pi\alpha_3)\cos(J\pi\alpha_1) I_z S_x \\
P_{I_z S}(\Lambda S_z) &= 0
\end{aligned}$$

Since $\{I_z S_x, I_z S_y, I_z S_z\}$ forms an orthogonal pair, we can rewrite

$$\|tr(F^\dagger Q_1 \Lambda Q_2 (I_z S^+) Q_2^\dagger \Lambda^\dagger Q_1^\dagger)\|.$$

as

$$\eta(t) = \|p^\dagger U \Sigma V p\|^2,$$

where $p = [1 \ i \ 0]^T$, and

$$\Sigma = \begin{bmatrix} -\sin(J\pi\alpha_3)\cos(J\pi\alpha_2) & 0 & 0 \\ 0 & \sin(J\pi\alpha_3)\cos(J\pi\alpha_1) & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and $U$ and $V$ are real orthogonal matrices. Using the result of lemma (3) we get that the maximum value of $\eta_z$ is

$$\sqrt{sin(J\pi\alpha_3)^2(cos(J\pi\alpha_2)^2 + cos(J\pi\alpha_1)^2)},$$

for $\alpha_1 + \alpha_2 + \alpha_3 = t \leq 1/(2J)$, which is maximized for $\alpha_1 = \alpha_2 = 0$ and $\alpha_3 = t$ and the value is $\sin(J\pi t)$. Similarly the maximum value of $\eta_x$ and $\eta_y$ is $\sin(J\pi t)$. Since $a_1^2 + a_2^2 + a_3^2 = 1$, we get $\eta(t) \leq \sin(J\pi t)$.

The optimal transfer curve for the antiphase transfer plotted as a function of time measured in units of $1/J$ is shown in the figure 3.3

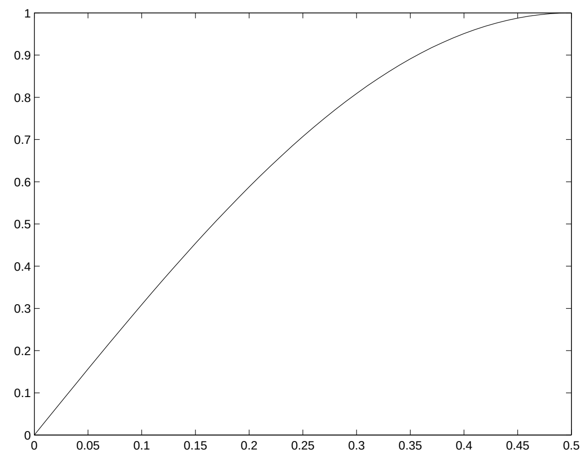


Figure 3.3: The panel shows the best achievable transfer as a function of time measured in units of $1/J$ for the antiphase transfer in 2 spin case.

## 3.3 Optimal Control in NMR

In this section we treat the problem of design of a time varying Hamiltonian in a a more general setting. Our goal is the design pulse sequences which not only produce efficient transfers but are economical in terms of the energy they spent and robust with respect to uncertainties about the complete knowledge of the system. Thus we would like to design control laws, which optimize more general cost functions.

To fix ideas we set up the following prototypical control problem. Consider the control system

$$\dot{\rho} = -i[H_d + \sum_{i=1}^{n} u_i H_i, \rho].$$

Where $H_d$ *and* $H_i$, belong to the space of $n \times n$ hermitian matrices denotes by $HE(n)$. Consider the cost function

$$J = \|tr(\rho^\dagger(T)F)\|^2 + \alpha \int_0^T \sum_{i=1}^{n} u_i^2.$$

The task is to find measurable control laws $\hat{u}_i : [0, T] \to R$, which minimize the cost $J$. We treat this as an optimal control problem. We will use maximum principle to find control laws which render the cost function $J$ stationary along the system trajectories.

**Theorem 12** Given the control system $\dot{\rho} = -i[H_d + \sum_{i=1}^{n} u_i H_i, \rho]$. where $H_d$ *and* $H_i \in$

$HE(n)$, Let $\lambda = \lambda_x + i\lambda_y$, where $\lambda_x, \lambda_y \in HE(n)$ and let

$$\begin{aligned}\dot{\lambda} &= -i[H_d + \sum_{i=1}^{n} u_i H_i, \lambda] \\ \lambda(T) &= 2tr(\rho^\dagger F)F\end{aligned}$$

The control laws $u_i : [0,T] \to R$ given by $u_i(t) = \alpha^{-1} Re\{tr(\lambda^\dagger(t), [-iH_i, \rho(t)])\}$ renders the cost function $J = \|tr(\rho^\dagger(T)F)\|^2 + \alpha \int_0^T \sum_{i=1}^n u_i^2$, stationary along the system trajectories.

**Proof:** Observe we can decompose $\rho$ as $\rho = \rho_x + i\rho_y$, where $\rho_x,\ \rho_y \in HE(n)$ and

$$\begin{aligned}\dot{\rho_x} &= -i[H_d + \sum_{i=1}^{n} u_i H_i, \rho_x] \\ \dot{\rho_y} &= -i[H_d + \sum_{i=1}^{n} u_i H_i, \rho_y].\end{aligned}$$

Consider the Hamiltonian function

$$h(\rho, \lambda, u_i) = tr(\lambda_x [H_d + \sum_{i=1}^{n} u_i H_i, -i\rho_x]) + tr(\lambda_y\ [H_d + \sum_{i=1}^{n} u_i H_i, -i\rho_y]) + \alpha \sum_{i=1}^{n} u_i^2.$$

Using the maximum principle we obtain that optimal control equations are given by

$$\begin{aligned}\lambda_x &= -\frac{\partial h}{\partial \rho_x} \\ \lambda_x(T) &= \frac{\partial J}{\partial \rho_x(T)} \\ \lambda_y &= -\frac{\partial h}{\partial \rho_y} \\ \lambda_y(T) &= \frac{\partial J}{\partial \rho_y(T)} \\ \frac{dh}{du_i} &= 0\end{aligned}$$

Substituting for $h$ and $J$ in the above equation gives us the required equations as in the statement of the theorem. **Q.E.D**

**Remark 4** Observe the variational equations for the optimal control constitute a two point boundary value problem. The system state equation given by the density matrix evolution has the boundary condition specified at the initial time $t = 0$ and the evolution of the co-state equation given by the evolution of $\lambda$ has its boundary value specified at time $t = T$.

It is difficult to find a closed form solutions to these equations. We therefore suggest here two algorithms to compute the solution to the optimal control problem. We first present the algorithm based on gradient flows in control space.

**Algorithm 1 Descent in Function Space**

1. Guess a starting value for $u_i(t)$.
2. Integrate the equation for $\rho(t)$ to get $\rho(T)$.
3. Evaluate th value of $\lambda(T)$ from $\rho(T)$.
4. Integrate the equation for $\lambda(t)$ backward.
5. Evaluate $\frac{\partial H}{\partial u_i}$ for all $t$.
6. Update $u_i(t) \to u_i(t) + \epsilon \frac{\partial H}{\partial u_i}(t)$.
7. Repeat till all $\|\frac{\partial H}{\partial u_i}(t)\| < \epsilon$, for some prescribed tolerance $\epsilon$.

We know present some concrete examples, where the above algorithm can be used to find optimal control laws.

**Example 1 Inphase Coherence Transfer**Consider the heteronuclear $I_nS$, spin system described by the following equation

$$\dot{\rho} = -i[H_d + \sum_{i=1}^{4} u_i H_i, \rho].$$

where

$$\begin{aligned} H_d &= 2\pi \sum_{k=1}^{n} I_{kz} S_z \\ H_1 &= 2\pi \sum_{k=1}^{n} I_{kx} \\ H_2 &= 2\pi \sum_{k=1}^{n} I_{ky} \\ H_3 &= 2\pi S_x \\ H_4 &= 2\pi S_y \end{aligned}$$

The cost $J = \|tr(\rho(T)^\dagger F)\|^2$, $\rho(0) = S_x - iS_y$ and $F = \sum_{k=1}^{n} I_{kx} - iI_{ky}$.

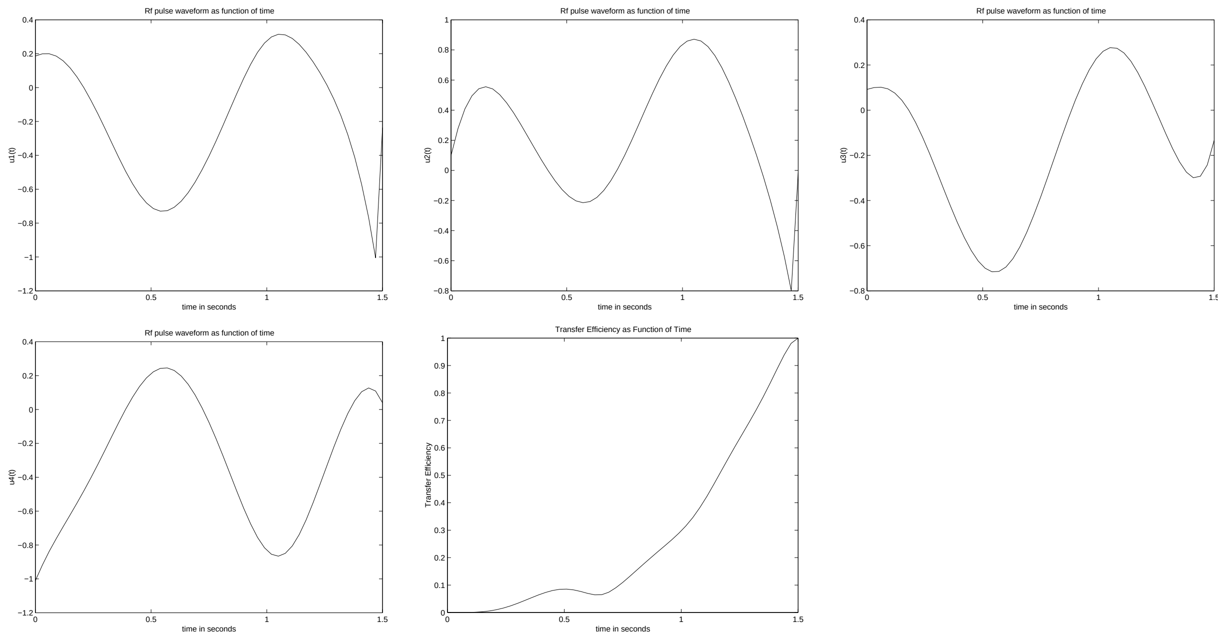


Figure 3.4: The top panels show the optimal control function $u_1(t)$, $u_2(t)$, $u_3(t)$ respectively for the example 1 for $n = 1$. The bottom panel show $u_4(t)$, and the cost $J(t)$ as function of time. The time is plotted in seconds.

The result of the algorithm for $n = 1$, $n = 2$ is shown in figure 3.4 and 3.5 respectively.

**Example 2 : Antiphase Coherence Transfer** *Consider the heteronuclear $I_nS$, spin system described by the following equation*

$$\dot{\rho} = -i[H_d + \sum_{i=1}^{4} u_i H_i, \rho].$$

*where*

$$\begin{aligned} H_0 &= 2\pi \sum_{k=1}^{n} I_{kz} S_z \\ H_1 &= 2\pi \sum_{k=1}^{n} I_{kx} \\ H_2 &= 2\pi \sum_{k=1}^{n} I_{ky} \\ H_3 &= 2\pi S_x \\ H_4 &= 2\pi S_y \end{aligned}$$

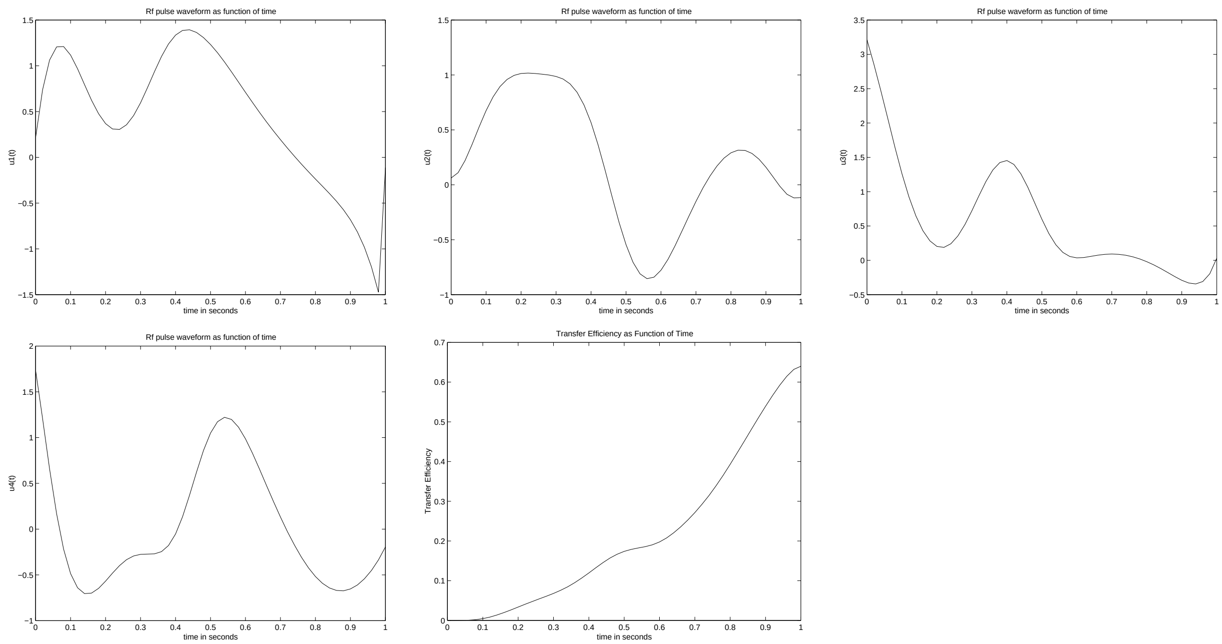


Figure 3.5: The top panels show the optimal control function $u_1(t)$, $u_2(t)$, $u_3(t)$ respectively for the example 1 for $n = 1$. The bottom panel show $u_4(t)$, and the cost $J(t)$ as function of time. The time is plotted in seconds.

*and* $F_z = \sum_{k=1}^{n} I_{kz}$. *The cost* $J = \|tr(\rho(T)^\dagger F)\|^2$, $\rho(0) = 2F_z(S_x - iS_y)$ *and* $F = \sum_{k=1}^{n} I_{kx} - iI_{ky}$.

The result of the algorithm for $n = 1$, $n = 2$ are shown in figures.

## 3.4 Conclusions

Our main contribution in this chapter has been to put the design of pulse sequences in coherent spectroscopy on solid geometrical foundations. Problems of design of control sequences in high resolution NMR has been formulated and solved as optimal control problems on compact Lie groups. We have computed fundamental bounds for coherence transfer in 2 spin system. Our equivalence theorem gives a analytical characterization of the minimum time required to implement any unitary transformation in a quantum system. We believe this work will give a new viewpoint and tools to treat other problems in NMR and control of quantum systems.

We here enumerate some of the directions for future work in this area which we believe can benefit a lot from the system theory point of view.

1. *Relaxation and Dissipative Effects:* In all our optimal control formulations, we neglected the relaxation effects which lead to decoherence in the system. We assumed that if control is implemented fast enough then we can make appropriate transfers in times short enough for decoherence to be significant. However this is not a valid assumption in spectroscopy of macromolecules like proteins, as these molecules tumble slowly in the solution and hence relaxation effects are very important. Therefore we need to include these effects into the evolution which gives us the *Master Equation* of evolution of density matrix

$$\dot{\rho} = -i[H, \rho] + \hat{\Gamma}(\sigma(t) - \sigma_0) \tag{3.10}$$

   in which the relaxation superoperator $\hat{\Gamma}$ can be expressed as a double commutator

$$\hat{\Gamma}(\sigma(t) - \sigma_0) = \sum_p \sum_q j^q(w_p)[A_p^q[A_p^{-q}, ]].$$

   This is where the power of system theoretic approach to these problems really comes in. We can invoke maximum principle again to find optimal control laws for such equations. We plan to incorporate such dissipative terms in our future study.

2. *Robust control:* Till now we assumed that the Hamiltonian was known exactly and is same for all the molecules. This is far from true. The two most important deviations are

   - *Deviation in Larmor Frequencies:* The Zeeman or the drift Hamiltonian for all spins is not the same, and there is a spread in the Larmor frequency, even if we have the same nuclei type. Therefore pulse sequences that will perform optimally for one set of nuclei will not be best for the others. Thus we need to design pulse sequences that optimize the expected transfers. The other source of uncertainty is

   - *RF inhomogeneity:* We assumed throughout that all spins see the same control, that is control field is homogeneous throughout the sample, but that is not true at all. In fact there may be deviation in amplitude ranging from 5 to 10 percent.

   System theory provides a framework for dealing with such problems.

3. *Selective Transfers:* Till now we worked on problems where we wanted to optimize certain transfers. In some cases, not only are we interested in maximizing certain transfers, but at the same time we are interested in minimizing or suppressing other transfers, as this increases the resolution of the spectra. In a control theory framework, we can associate appropriate costs to the transfer and design pulse sequences to optimize these costs.

# Chapter 4

# Feedback Stabilization of nonholonomic Systems

In this chapter, we introduce the problem of feedback stabilization of nonholonomic control systems. Our main result in this chapter is the construction of feedback control laws which asymptotically stabilize first brackett controllable systems. Feedback stabilization problems are concerned with obtaining feedback laws which guarantee that an equilibrium of the closed-loop system is asymptotically stable. The formal problem statement is as follows.

**Problem:** Let the control system in local coordinates be given by

$$\dot{x} = f(x,u) \ \ ; \ f(x_0,0) = 0, x \in \mathbb{R}^n, \ u \in \mathbb{R}^m.$$

Find a smooth function $u(x)$ such that the equilibrium point $x = x_0$ is asymptotically stable.

For a linear time-invariant system, if all unstable eigenvalues of the system are controllable, then the system can be stabilized by a linear time-invariant static state feedback. For nonholonomic systems the situation is more complex. The linearization of nonholonomic system about any equilibrium point is not asymptotically stabilizable. There are fundamental topological problems associated with the existence of continuous time-invariant feedback laws for nonholonomic systems. Before we delve into these issues any deeper, we will introduce basic definitions and results from classical stability theory.

**Definition 5** Given the vector differential equation

$$\dot{x} = f(x,t), \tag{4.1}$$

we say that $f \in E$, if $f$ is continuous and satisfies the Lipschitz condition such that existence and uniqueness of solution to the differential Equation (4.1) is guaranteed. We use $x(t)$ to denote that well-defined solution, which takes on value $x(0) = x_0$ at $t = 0$. A constant solution, $x(t) = x_0$, is said to be an *equilibrium solution*, or a *equilibrium point*, of the differential equation.

**Definition 6** The equilibrium solution of the differential Equation (4.1) is said to be *stable* if there exists, for each $\epsilon > 0$, a number $\delta > 0$ such that the inequality $\|x_0\| < \delta$ implies $\|x(t)\| < \epsilon$.
The equilibrium is said to be *quasi-asymptotically stable* if there is a number $\delta_0 > 0$ such that, from $\|x_0\| < \delta_0$, the relation $lim_{t\rightarrow\infty} x(t) = 0$ follows.
The equilibrium is said to be *asymptotically stable* if it is both stable and quasi-asymptotically stable.

We can extend this concept of stability to a point to stability to a submanifold.

**Definition 7** Let $x \in \mathbb{R}^n$ and $\theta \in M$ be a compact differentiable manifold. Following Zubov (1957), we define the submanifold $N = 0 \times M$ to be a asymptotically stable submanifold of the differential equations

$$\begin{aligned} \dot{x} &= f(x, \theta, t) \quad (f \in E) \\ \dot{\theta} &= g(x, \theta, t) \quad (g \in E) \end{aligned}$$

if there exists, for each $\epsilon > 0$, a scalar $\delta > 0$ such that the inequality $\|x(0)\| < \delta$ implies that the solution $(x(t), \theta(t))$, corresponding to the initial condition $(x(0), \theta(0))$, satisfies $\|x(t)\| < \delta, \ \forall t > 0$ and if there is a scalar $\delta_0 > 0$ such that, for $\|x(0)\| < \delta_0$,

$$\lim_{t\rightarrow\infty} x(t) = 0. \tag{4.2}$$

If equation (4.2) is satisfied for all $x(0)$, we say that submanifold $N$ is asymptotically stable in the large.

Having reviewed the basic definitions from stability theory, let us first recapitulate what we know about the stabilization problem for a linear system.

**Linear System**

**Theorem 13** Consider the standard linear system

$$\dot{x} = Ax + Bu, \quad x \in \mathbb{R}^n \tag{4.3}$$

where $A$ and $B$ are constant matrices. The null solution of (4.3) is stabilizable if and only if all the modes associated with eigenvalues with non-negative real parts are controllable.

**Proof:** By change of basis, we can write (4.3), as

$$\frac{d}{dt}\begin{bmatrix} x_u \\ x_l \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}\begin{bmatrix} x_u \\ x_l \end{bmatrix} + \begin{bmatrix} B_u \\ 0 \end{bmatrix} u$$

with range $(B_u, A_{11}B_u, \ldots, A_{11}^{n-1}B_u) = dim\ x_u$. Observe that the eigenvalues of $A_{22}$ must have negative real parts if $x_l$ is to go to zero as $t$ goes to $\infty$. It is well known that if $(B, AB, \ldots A^{n-1})$ is of rank $n$ then there exists a $m \times n$ matrix $K$ such that $(A + BK)$ has its eigenvalues in the open left-half plane. Thus the assertion follows.

**Remark 5** A systematic approach to the design of feedback controller can be based on the minimization of the integral of some positive function of the error and the control effort. For the linear systems, this might take the following well known form. Let the linear system $\dot{x} = Ax + Bu$ be controllable, then the control law $u(x)$ which stabilizes the null solution and among all stabilizing feedback control laws minimizes the cost function

$$\int_0^\infty (x^T Qx + u^T u)dt$$

is given by $u(x) = -B^T Kx$, where $K$ is the solution to the quadratic matrix equation

$$A^T K + KA - KBB^T K + Q = 0.$$

In light of the above discussion, we might expect that for a general affine control system

$$\dot{x} = f(x) + \sum_{i=1}^{n} u_i g_i(x),$$

we might accept similar behavior, in particular we can ask that if every initial state in the neighborhood of $x_0$ can be steered to $x_0$, then does there exist a feedback control law which makes $x_0$ asymptotically stable. It was shown by BROCKETT (1983) that the answer to this question is no provided we want a continuous feedback control law. Before we state the theorem we would like to take one more detail into account.

If we have

$$\dot{x} = f(x, u)\ ; f(x_0, 0) = 0,$$

with $f(.,.)$ continuously differentiable with respect to both arguments, and if we define $A = (\frac{\partial f}{\partial x})_{x_0}$ and $B = (\frac{\partial f}{\partial u})_0$, then the control system $\dot{x} = Ax + Bu$ represents the linearized system at $(x_0, 0)$. From the Theorem 13, we can infer that there exists a stabilizing control law for $\dot{x} = f(x, u)$ with $f(0,0) = 0$ provided the unstable modes of the linearized system are controllable and there exists no stabilizing control law if the linearized system has an unstable mode which is uncontrollable. The interesting phenomenon begins if the linearized system has uncontrollable purely imaginary eigenvalues. The following theorem gives a necessary condition for existence of stabilizing control law, which is decisive for a large class of problems of interest in nonlinear control.

**Theorem 14** *(***Brockett:***)* Let $\dot{x} = f(x, u)$ be given with $f(x_0, 0) = 0$ and $f(.,.)$ continuously differentiable in a neighborhood of $(x_0, 0)$. A necessary condition for the existence of a continuously differentiable control law $u(x)$ such that $u(x_0) = 0$, which makes $x_0$ asymptotically stable is that

- The linearized system should have no uncontrollable modes associated with eigenvalues whose real part is positive.

- There exists a neighborhood of $(x_0, 0)$ such that, for each $y \in N$, there exists a control $u_y(.)$ defined on $[0, \infty)$ such that the control steers the solution of $\dot{x} = f(x, u_y)$ from $x = y$ at $t = 0$ to $x = 0$ at $t = \infty$.

- The mapping $f$ maps every neighborhood of $x_0$ onto some neighborhood of zero.

**Proof:** The necessity of the first two conditions follow from the previous discussion. To prove the last condition, observe that, if $x_0$ is an equilibrium point of $\dot{x} = f(x)$, which is asymptotically stable, then there exists a Lyapunov function $V(x)$ such that $V(x)$ is positive for $x \neq 0$ and vanishes at $x_0$, is continuously differentiable, and has level sets $V^{-1}(\alpha)$ which are homotopy spheres. Observe that the vector field $f(x)$ is normal to the homotopy sphere $V^{-1}(\alpha)$, is nonzero on $V^{-1}(\alpha)$, and always points inwards. Compactness of $V^{-1}(\alpha)$ implies that, if $\|\xi\|$ is sufficiently small, the vector field associated with $\dot{x} = f(x) + \xi$ also points inwards on $V^{-1}(\alpha)$. By evaluating at time $t = 1$ the solution of $\dot{x} = f(x) + \xi$ which passes through $x$ at $t = 0$, we get a continuous map of $\{x | V(x) \leq \alpha\}$ into itself. Applying the Lefschetz fixed-point theorem, we see that this map has a fixed point which must be the equilibrium point of $\dot{x} = f(x)+\xi$. This implies we can solve for $f(x) = \xi$ for all $\xi$ sufficiently small. Hence the proof.

**Remark** If we have the control system

$$\dot{x} = f(x) + \sum_{i=1}^{n} u_i g_i(x) \quad ; \; x \in \mathbb{R}^n,$$

then the above theorem implies that the stabilization cannot have a solution if there is a smooth distribution $D$ which contains $f(.)$ and $g_1(.), \ldots,\ g_m(.)$ with $dim D < n$. In particular, there exist no continuous stabilizing control law for

$$\begin{aligned} \dot{x_1} &= u_1 \\ \dot{x_2} &= u_2 \\ \dot{x_3} &= x_1 u_2 - x_2 u_1. \end{aligned}$$

We will now like to rephrase the necessary onto condition of Theorem 14 in more degree theoretic language, which in fact gives us stronger necessary condition for stability. For this, we state the following theorem due to KRASNOSELSKII AND ZABREIKO (1983). We first define the notion of of a singular point.

**Definition 8** Let $x_0$ be a isolated zero of a continuous vector field $f(x) \in \mathbb{R}^n,\ x \in \mathbb{R}^n$, and let $S^{n-1}(x_0, r)$ denote a sphere of radius $r$ around $x_0$. The degree of the Gauss map $H : S^{n-1}(x_0, r) \to S^{n-1}(0, 1),\ H(x) = \frac{f(x)}{\|f(x)\|}$ for sufficiently small radius $r$ is called the *index of the singular point* $x_o$.

**Theorem 15** If the equilibrium point 0 of

$$\dot{x} = f(x)$$

is asymptotically stable then the index of the singular point zero of the vector field $f(x)$ is $(-1)^n$.

This theorem implies that near the equilibrium point $x_0$, the Gauss map $h(x) = \frac{f(x)}{\|f(x)\|}$ is homotopic to the antipodal map. In particular this implies that the Gauss map $h(x) = \frac{f(x)}{\|f(x)\|}$ is onto $S^{n-1}$ (as it has nonzero degree). This immediately implies that, for the system

$$\begin{aligned} \dot{x_1} &= u_1 \\ \dot{x_2} &= u_2 \\ \dot{x_3} &= x_1 u_2 - x_2 u_1, \end{aligned}$$

there exists no continuous feedback law because the map $h : S^{n-1} \to S^{n-1}$ is not onto as we cannot reach the point $(0, 0, \pm 1)$. Thus we see that feedback stabilization of nonholonomic systems can be seen as a topological problem.

A number of approaches have been proposed for the stabilization of nonholonomic systems. These approaches can be broadly classified into three categories (KOLMANOVSKY AND MCCLAMROCH (1995)).

**Discontinuous time-invariant control laws**

Discontinuous time-invariant control laws for stabilization of nonholonomic systems have been used by many researchers, ( ASTOLFI (1994); SUSSMANN (1979); LAFFERRIERE AND SONTAG (1991); BLOCH AND S.DRAKUNOV (1996); GULDNER AND UTKIN (1994); SAMSON AND AIT-ABDERRAHIM (1991); BLOCH ET AL. (1998)). These control laws are either piecewise continuous functions of state or sliding mode controllers. An example for stabilization of nonholonomic integrator using the sliding mode approach is given by the following control law (BLOCH AND S.DRAKUNOV (1996))

$$\begin{aligned} u_1 &= -x_1 + 2x_2 \; sign(x_3 - \frac{x_1 x_2}{2}) \\ u_2 &= -x_2 + 2x_1 \; sign(x_3 - \frac{x_1 x_2}{2}), \end{aligned}$$

where sign(.) denotes the signum function. The disadvantage of these controllers is that they may cause chattering.

**Time-Varying Stabilization**

The use of time-varying feedbacks originated in the mobile robot work by SAMSON AND AIT-ABDERRAHIM (1991). CORON (1991) showed any driftless controllable system can be asymptotically stabilized to a equilibrium point using smooth time-periodic static-state feedback. Various methods for designing time-varying controllers have been proposed in literature ( CORON AND POMET (1992); POMET (1992); MURRAY AND SASTRY (1993); BROCKETT (1996); MORGANSEN AND BROCKETT (1999)). As an illustration consider the following time-varying smooth feedback law for the nonholonomic integrator

$$\begin{aligned} u_1 &= -x_1 - x_3 \cos(t) \\ u_2 &= -x_2 - x_3^2 \sin(t). \end{aligned}$$

**Hybrid Feedback Laws**

Typically, hybrid controllers combine continuous time features with discrete event features. These have been used by BLOCH AND S.DRAKUNOV (1996); I. KOLMANOVSKY AND MCCLAMROCH (1994). The essential feature of these controllers is that there is a low level time invariant feedback controller supervised by discrete event system. The supervisor configures the low level feedback controllers and switches between them to provide stability.

We instead take a approach motivated by the following question.

**Problem Statement 5** Consider a control system $\dot{x} = f(x,u)$, $x \in \mathbb{R}^n, u \in \mathbb{R}^m$, which cannot be stabilized by a continuous static feedback control $u : \mathbb{R}^n \to \mathbb{R}^m$. Is it possible to embed this system in a higher dimensional manifold $\mathbb{R}^n \times N$ and find smooth static feedback laws $u(x,\theta)$ *and* $g(x,\theta)$, $(x,\theta) \in \mathbb{R}^n \times N$, such that, for any initial condition $(x(0),\theta(0))$, the solution of

$$\dot{x} = f(x, u(x,\theta)) \tag{4.4}$$

$$\dot{\theta} = g(x,\theta) \tag{4.5}$$

satisfies $\lim_{t\to\infty}(x(t),\theta(t)) \in 0 \times N$

Observe that, by finding such a system, we would have achieved the goal of stabilizing the original system $\dot{x} = f(x,u)$. Thus, the problem of stabilization to a point in the original space has been transformed to the problem of stabilization to a submanifold in the enlarged state space.

**Remark 6** Stabilization of nonholonomic systems by time-varying feedback is a special case of the above described situation, where the dynamical system $\dot{x} = f(x,u(x,t))$, $x \in M$, where $M$ is some manifold, can be seen as embedded in the larger space $M \times \mathbb{R}$ with

$$\dot{x} = f(x, u(x,\theta))$$

$$\dot{\theta} = 1,$$

where $\theta \in \mathbb{R}$. Also, stabilization of $x$ to some point $x_0$ in the manifold $M$ can be thought of as stabilization of $(x,\theta)$ to the submanifold $x_0 \times \mathbb{R}$ in the enlarged state space $M \times \mathbb{R}$.

As shown by Coron (1991), it is always possible to stabilize a driftless, controllable system by a smooth periodic time-varying control law. However, time-varying control laws make two special choices in the above raised Problem 5 by choosing $N = \mathbb{R}$ and $g(x, \theta) = 1$. We will show that by embedding the system in a higher dimensional space and finding control laws in the enlarged space, we can find simpler feedback control laws.

Before we illustrate some simple dynamic feedback laws for nonholonomic systems, we present a generalization of Brockett's theorem on asymptotic stability (Brockett (1983)) by smooth periodic time varying control laws.

**Theorem 16** *Let $\dot{x} = f(x,t)$ be given with $f(x_0, t) = 0$, where $x \in \mathbb{R}^n$ and $t \in \mathbb{R}$. In addition, let $f(x,t)$ be periodic with period $T$*

$$f(x,t) = f(x, t+T).$$

*Let $\phi_a(x,t)$ denote the unique solution to $\dot{x} = f(x,t) + a$. If $x = x_0$ is an asymptotically stable point for the above system, then it must hold true that, for every nghd $\Omega$ of $x_0$, there exist a neighborhood $\Omega_1$ of $0$ such that, for all $a \in \Omega_1$, the equation*

$$x = \phi_a(x, T) \quad x \in \Omega$$

*has a solution.*

**Proof:** We first observe that the above set of differential equations can be written as an autonomous system

$$\dot{x} = f(x, \theta) \tag{4.6}$$

$$\dot{\theta} = 1, \tag{4.7}$$

where $(x, \theta) \in R^n \times S^1$. If $x_0$ is a asymptotically stable point of original system, then $N = x_0 \times S^1$ is asymptotically stable submanifold of the autonomous system (4.6). Hence, there exists a Lyapunov function $v$, such that $v(p) \geq 0$ for all $p \in R^n \times S^1$ and vanishes only at $p \in N$, is continuously differentiable and has level sets $M_c = v^{-1}(c)$, $c > 0$, which are homotopy tubes such that the vector field $f(x, \theta)$ points inward at all points on $M_c$ (for more details see Wilson (1967)).

Compactness of $N$ implies the compactness of $M_c$ which further entails that, if $a$ is sufficiently small, say $\|a\| < \epsilon$ for some $\epsilon > 0$, then the vector fields associated with

$$\dot{x} = f(x, \theta) + a \tag{4.8}$$

$$\dot{\theta} = 1 \tag{4.9}$$

points inward on $M_c$. Thus, the flow $\phi_a(x, \theta, t)$ induces a continuous map of $\{p|v(p) \leq c\}$ into itself. In particular, $\phi_a(., \theta_0, T)$ is a continuous map of the homotopy ball $\{(x, \theta_0)|v(x, \theta_0) \leq c\}$ into itself. Applying the Lefschetz fixed-point formula, we see that the map $\phi_a(., \theta_0, T)$ has a fixed point. Thus there exits a $\xi \in \Omega = v(x, \theta_0) \leq c$ such that

$$(\xi, \theta_0) = \phi_T(\xi, \theta_0)$$

Hence the proof. **Q.E.D.**

In the limit $T \to 0$ we recover the well known onto condition.

In the remaining part of this chapter, we will illustrate our approach by constructing smooth dynamic feedback laws for nonholonomic systems. The main result in this chapter is the constructive solution to the problem of feedback stabilization of nonholonomic systems which are first bracket controllable. We will approach this problem by first constructing feedback laws for the generalization of nonholonomic integrator, called the *general position area system* or *the so(n) system*. The system is described by the equations BROCKETT (1981)

$$\dot{x} = u \tag{4.10}$$
$$\dot{y} = xu^T - ux^T, \tag{4.11}$$

where $x,\ u$ are column vectors in $\mathbb{R}^n$ and $y \in so(n),\ n \geq 2$. Here, $so(n)$ is the Lie algebra of the $n \times n$ skew-symmetric matrices: $y^T = -y$.

The importance of the $so(n)$ system is that it is the canonical form of a class of driftless controllable systems of the form $\dot{x} = B(x)u,\ u \in \mathbb{R}^n,\ x \in \mathbb{R}^{\frac{n(n+1)}{2}}$, whose first derived algebra spans the tangent space $T\mathbb{R}^{\frac{n(n+1)}{2}}$ at any point (Recall that if $E^0$ is a sub-bundle of the tangent bundle spanned by the control fields, then the first derived algebra is given by $E^1 = E^0 + [E^0 E^0]$). BROCKETT (1981) showed that such a systems can be transformed to the form of (4.10-4.11) up to a suitable order in the neighborhood of a given point such as the origin.

To fix ideas, we first analyze in detail the dynamic stabilization for the nonholonomic integrator.

## 4.1 Stabilization of the nonholonomic integrator

We will motivate our stabilization approach by the following discussion. Consider the nonholonomic integrator

$$\begin{aligned} \dot{x}_1 &= u_1 \\ \dot{x}_2 &= u_2 \\ \dot{x}_3 &= x_1 u_2 - x_2 u_1. \end{aligned} \tag{4.12}$$

Observe that by choosing $u_1 = -x_1$ and $u_2 = -x_2$, one can exponentially stabilize $x_1$ and $x_2$ to zero. Motion in the $x_3$ direction is produced by generating areas in the $x_1 - x_2$ plane. Is there a natural way to stabilize $x_3$ to zero? Let us add another dimension to the above system by introducing the variable $\theta \in \mathbb{R}$. In the $(x_1, x_2, x_3, \theta)$ space consider a one-parameter family of embedded submanifolds $\{S_r\}$, $r \geq 0$, defined by $S_r = \{(x_1, x_2, x_3, \theta) \in \mathbb{R}^4 : x_1 = r\cos\theta,\ x_2 = r\sin\theta\}$. Observe that if $(x_1, x_2, x_3, \theta) \in S_R$, $R > 0$ , then $\dot{x}_3(t) = x_1\dot{x}_2 - x_2\dot{x}_1 = R^2\ \dot{\theta}(t)$. If we let $\dot{\theta}(t) = -x_3$, then $\dot{x}_3(t) = -R^2\ x_3$ and hence $x_3$ is exponentially stabilized to zero. Thus, in the enhanced space parametrized by $(x_1, x_2, x_3, \theta)$, we have identified a family of submanifolds $\{S_r\}$ on which there is a natural way to stabilize $x_3$.

Our control strategy will be to design $u_1$ and $u_2$ to stabilize $(x_1, x_2, x_3, \theta)$ to the submanifold $S_{q(t)}$ and let $\dot{\theta} = -x_3$ to stabilize $x_3$ as explained above. $q(t)$ is such that it goes to zero as $x_3$ goes to zero, thereby stabilizing all $x_1, x_2, x_3$ to zero. This can be achieved in different ways, and we will present two simple ways of accomplishing the above in the following theorems.

**Theorem 17** *Let* $\theta \in \mathbb{R}$ *and* $(x_1, x_2, x_3) \in \mathbb{R}^3$. *If* $u_1 = -x_1 + x_3\cos\theta$ *and* $u_2 = -x_2 - x_3\sin\theta$, *then there exists* $\epsilon > 0$ *such that, for any initial condition* $\|(x_1(0), x_2(0), x_3(0))\| < \epsilon$, *the solution of*

$$\begin{aligned} \dot{x}_1 &= u_1 \\ \dot{x}_2 &= u_2 \\ \dot{x}_3 &= x_1 u_2 - x_2 u_1 \\ \dot{\theta} &= x_3 \end{aligned}$$

*satisfies* $\lim_{t\to\infty} (x_1(t), x_2(t), x_3(t)) = (0, 0, 0)$.

**Proof:** In terms of the variables

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \tag{4.13}$$

the above equations take the form

$$\begin{bmatrix} \dot y_1 \\ \dot y_2 \\ \dot y_3 \end{bmatrix} = \begin{bmatrix} -1 & -y_3 & 1 \\ y_3 & -1 & 0 \\ 0 & 0 & -y_2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \tag{4.14}$$

$$\dot\theta = y_3. \tag{4.15}$$

Observe that the equations for $(y_1, y_2, y_3)$ do not depend on $\theta$, so they can be treated as an autonomous system of equations in $\mathbb{R}^3$. Define the Lyapunov function

$$V(y_1, y_2, y_3) = (y_1 - y_3)^2 + y_2^2 + y_3^2.$$

From the equations (4.14)

$$\dot V(y_1, y_2, y_3) = \begin{bmatrix} (y_1 - y_3) & y_2 \end{bmatrix} \begin{bmatrix} -2 & -y_3 \\ -y_3 & -2 \end{bmatrix} \begin{bmatrix} (y_1 - y_3) \\ y_2 \end{bmatrix}. \tag{4.16}$$

Notice that if $\|y_3\| < 1$, then $\dot V(y_1, y_2, y_3) \le 0$. Let

$$B = \{(y_1, y_2, y_3) \in \mathbb{R}^3 : V(y_1, y_2, y_3) < 1\}.$$

Observe that $B \in \{\|y_3\| < 1\}$ is therefore a positively invariant set. Let

$$S = \{(y_1, y_2, y_3) \in B : \dot V(y_1, y_2, y_3) = 0\}.$$

From equation (4.16), if $(y_1, y_2, y_3) \in S$, then $y_1 = y_3 and\ y_2 = 0$. Substituting this in equation (4.14) we conclude that if $(y_1, y_2, y_3) \in S$, then $\dot y_1 = 0$, $\dot y_2 = y_3^2$, and $\dot y_3 = 0$. It follows that the largest invariant set contained in $S$ is $\{0\}$. Hence, by LaSalle's stability theorem, if $(y_1(0), y_2(0), y_3(0)) \in B$, then $\lim_{t \to \infty}(y_1(t), y_2(t), y_3(t)) = 0$.

Let $\Omega(\epsilon) = \{(x_1, x_2, x_3) \in \mathbb{R}^3 | x_1^2 + x_2^2 + x_3^2 < \epsilon\}$, $\epsilon > 0$. By equation (4.13), we have

$$y_1^2(t) + y_2^2(t) = x_1^2(t) + x_2^2(t) \quad , \quad y_3(t) = x_3(t),$$

from which we can deduce that, if $(x_1, x_2, x_3) \in \Omega(\frac{1}{3})$, then $(y_1, y_2, y_3) \in B$. This shows that if $(x_1(0), x_2(0), x_3(0)) \in \Omega(\frac{1}{3})$, then $\lim_{t \to \infty} (y_1(t), y_2(t), y_3(t)) = (0, 0, 0)$ and therefore $\lim_{t \to \infty} (x_1(t), x_2(t), x_3(t)) = (0, 0, 0)$. **Q.E.D.**

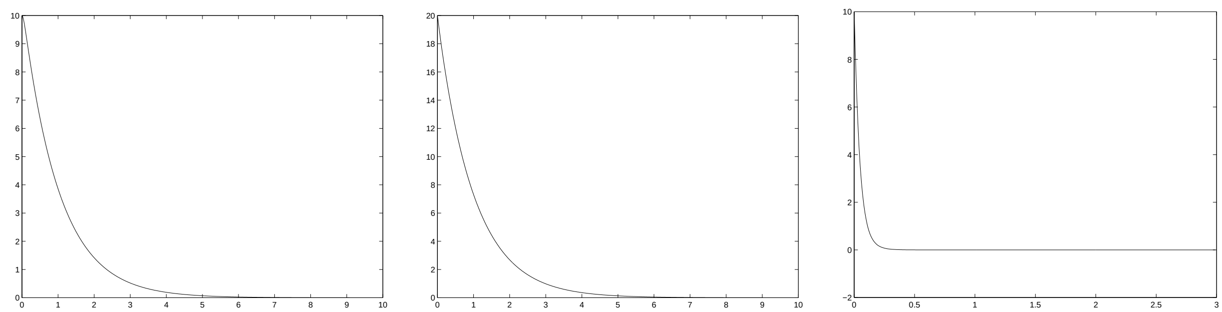

Figure 4.1: The above panels show the result of stabilization of nonholonomic system, with the graphs showing the evolution of $x_1, x_2, x_3$ from left to right for the set of initial conditions $x_1 = 10,\ x_2 = 20,\ x_3 = 10$.

**Remark 7** We only proved asymptotic stability here, With little more work it can be shown that the above system of equations also globally stabilize $x_1, x_2, x_3$ to the origin. Let $e^T = [1, 0] \in \mathbb{R}^2$, $x^T = [x_1, x_2] \in \mathbb{R}^2$, and $u^T = [u_1, u_2] \in \mathbb{R}^2$. Let

$$y = \begin{bmatrix} 0 & x_3 \\ -x_3 & 0 \end{bmatrix} \ ; \ \Theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

then observe

$$xu^T - ux^T = \begin{bmatrix} 0 & (x_1u_2 - x_2u_1) \\ -(x_1u_2 - x_2u_1) & 0 \end{bmatrix}.$$

Hence the control law in Theorem 17 can be rewritten as $u = -x + y\Theta e$ , and

$$\begin{aligned} \dot{x} &= u \ ; \ x \in \mathbb{R}^2 \\ \dot{y} &= xu^T - ux^T \ ; \ y \in so(2) \\ \dot{\Theta} &= y\Theta. \end{aligned}$$

The results of simulations for stabilizing the above system is shown in Figure 4.1.

**Remark 8** The main point we want to emphasize here is that the topological restriction imposed by Brockett's necessary condition for feedback stabilization can be overcome by embedding the system in some higher-dimensional space, which provides a lot of flexibility in the choice of control.

We can now extend this viewpoint to find stabilizing control laws for the general position-area system.

## 4.2 General position-area system

Recall that the system is described by the following set of equations,

$$\dot{x} = u \tag{4.17}$$

$$\dot{y} = xu^T - ux^T, \tag{4.18}$$

where $x$ and $u$ are column vectors in $\mathbb{R}^n$ and $y \in so(n)$, $n \geq 2$. We will find smooth stabilizing control laws for the above system by embedding it in $\mathbb{R}^{n+1} \times so(n) \times O(n)$. We motivate the choice of $O(n)$ by the following discussion. To understand how the general problem has an additional level of complexity with respect to the special case of $n = 2$, we start by looking at the qualitative nature of the trajectories that need to be generated in order to stabilize $y$. For $n = 2$, we saw that motion in $x_3$ direction was produced by generating areas in the $x_1 - x_2$ plane. Now, we need to generate $n(n-1)/2$ areas ($dy_{ij} = x_i dx_j - x_j dx_i$) for stabilizing $y \in so(n)$. This can be achieved as follows. Let $e \in \mathbb{R}^n$ be a unit vector which evolves as $\dot{e}(t) = ye$. Suppose we can make $x(t) = qe(t)$, where q is a positive constant, then the norm of $y(t)$ from equation (4.18) evolves as

$$\frac{d\ tr(yy^T)}{dt} = -2\ q^2\ \|ye\|^2,$$

which suggests that norm of $y$ will decrease until $e$ begins to lie in the null space of $y$. Essentially, what is happening is that the vector $e$ is generating areas in $\mathbb{R}^n$ to stabilize a subspace of $yy^T$. To stabilize all $n$ orthogonal subspaces of $yy^T$, we will evolve $n$ orthonormal vectors in $\mathbb{R}^n$, each of which will generate areas to stabilizes a subspace of $yy^T$. This is naturally achieved by introducing $\Theta \in O(n)$ such that $\dot{\Theta} = y\Theta$. The columns of $\Theta$ then form the desired orthonormal frame. We arrange matters so that we can switch between these orthonormal vectors in a smooth way. This is done using a *selector function* introduced in the following definition. Now by letting $q$ gradually go to zero as $y$ goes to zero, we can stabilize both $x$ and $y$.

**Definition 9** *(***Selector Function:***)*Let $e(t) = (e_1(t),\ e_2(t),\ \cdots,\ e_n(t)) \in \mathbb{R}^n$ be a $C^1$ function of time. We will call $e(t)$ a selector function of period $T$ and strength $\epsilon > 0$ if it satisfies the following properties

- $e(t) = e(t+T)$ such that $\|e(t)\| \leq 1$ and $\|\dot{e}(t)\|$ is bounded,

- $\int_t^{t+T} \|e_i(\tau)\|\ d\tau \geq \epsilon$, $\forall\ t$, and $e_i(t) \cdot e_j(t) = 0$ if $i \neq j,\ i, j \in 1, \ldots, n$.

We then say that $e \in \mathcal{SF}(n, T, \epsilon)$.

**Lemma 4** Suppose $e(t) \in \mathbb{R}^n$ is a selector function and $y \in so(n)$, then

$$e^T(t)\ y\ \dot{e}(t) = 0.$$

**Proof:** If $e(t) = 0$, then the proposition is trivial. Suppose $e_i(t) \neq 0$ for some $i \in 1, \ldots, n$, then, by the definition of selector function, $e_j(t) = 0$ if $i \neq j$. Differentiating $e_i(t) \cdot e_j(t) = 0$, we conclude that $e_i(t) \cdot \dot{e}_j(t) = 0$ if $i \neq j$. This shows that $e^T(t)\ y\ \dot{e}(t) = y_{ii}\ e_i(t)\dot{e}_i(t)$, but $y_{ii} = 0$ because $y$ is skew-symmetric. **Q.E.D.**

**Lemma 5** Let $e \in \mathcal{SF}(n, T, \epsilon)$, $y(t) \in so(n)$, and $\Theta \in O(n)$ such that $\dot{\Theta} = w(t)\Theta$, where $w(t) \in so(n)$. Suppose $[y(t), w(t)] = 0\ \forall t$, then

$$\int_t^{t+T} \|y\Theta e(\tau)\|\ d\tau > \epsilon \int_t^{t+T} \frac{\|y(\tau)\|}{nT} - 2\|\dot{y}(\tau)\|\ d\tau.$$

**Proof:** Let $z = \Theta^T y \Theta$, then it follows that $\|z\| = \|y\|$ and $\|y\Theta e\| = \|ze\|$. Because $[y(t), w(t)] = 0$, it follows that $\|\dot{z}(t)\| = \|\dot{y}(t)\|$. Let $u_i^T = (0, \ldots, 1, \ldots 0)$ with 1 in the $i^{th}$ position. Let $\Lambda(t) = \int_t^{t+T} \|\dot{z}(\tau)\|\ d\tau$. Observe that $\|z(t)\ u_k\| \geq \frac{\|z(t)\|}{n}$ for some $\bar{k} \in 1, \ldots, n$. Then,

$$\begin{aligned}
\int_t^{t+T} \|ze(\tau)\|\ d\tau &\geq \int_t^{t+T} \|e_{\bar{k}}(\tau)\| \|zu_{\bar{k}}\|\ d\tau \\
&\geq \int_t^{t+T} \|e_{\bar{k}}(\tau)\| (\|zu_{\bar{k}}(t)\| - \Lambda) \\
&\geq \epsilon(\ \|zu_{\bar{k}}(t)\| - \Lambda\ ) \qquad (4.19) \\
&\geq \epsilon(\ \frac{\|z(t)\|}{n} - \Lambda\ ), \qquad (4.20)
\end{aligned}$$

where (4.19) follows from the definition of selector function. Also notice that

$$\begin{aligned}
\int_t^{t+T} \|z(\tau)\|\ d\tau &\leq \int_t^{t+T} (\ \|z(t)\| + \Lambda\ )\ d\tau \\
&\leq T\ (\ \|z(t)\| + \Lambda\ ). \qquad (4.21)
\end{aligned}$$

Combining inequalities (4.20)-(4.21) we get the desired result. Notice in particular that if $y(\tau) = y_0$ a constant, then

$$\int_t^{t+T} \|y_0 \Theta e(\tau)\|\ d\tau > \frac{\epsilon}{n} \|y_0\|$$

**Q.E.D.**

We now present a feedback stabilization law for the position-area system.

**Theorem 18** Let $\mathcal{S}$ be a subspace of $so(n)$ and $\mathcal{P} : so(n) \to so(n)$ be a projection operator onto this subspace. Let $x \in \mathbb{R}^n$, $q \in \mathbb{R}$, $\Theta \in O(n)$, $y \in \mathcal{S}$, and $e \in \mathcal{SF}(n, T, \epsilon)$. If

$$u = -x + \|y\|\Theta e + q(y\Theta e + \Theta \dot{e}),$$

then for

$$\begin{aligned} \dot{x} &= u && (4.22)\\ \dot{y} &= \mathcal{P}[xu^T - ux^T] && (4.23)\\ \dot{q} &= -(q - \|y\|) && (4.24)\\ \dot{\Theta} &= y\Theta, && (4.25) \end{aligned}$$

the submanifold $N = \{\ (x, y, q, \Theta) \in \mathbb{R}^n \times so(n) \times \mathbb{R} \times O(n) : x = 0, y = 0, q = 0\ \}$ is asymptotically stable in the large.

**Proof:** First notice that, for $\mathcal{S} = so(n)$ and $\mathcal{P}$ the identity operator, equations (4.22)-(4.23) reduce to the position-area system. Let $(x(t), y(t), \Theta(t), q(t))$ be the solution of equations (4.22)-(4.24) for a given initial condition $(x(0), y(0), \Theta(0), q(0))$. To simplify notation, we will often drop the time index $t$ and just write the solution as $(x, y, \Theta, q)$. Let $p = \Theta e$, $\bar{p} = \Theta \dot{e}$, and $r = x - qp$. Then $u = -x + \|y\|p + qyp + q\bar{p}$. From equations (4.22)-(4.24), we get $\dot{r} = -r$. Notice that

$$\begin{aligned} \frac{d\ tr(y^T y)}{dt} &= 2\ tr(y^T \mathcal{P}[xu^T - ux^T]) && (4.26)\\ &= 2\ tr(y^T[xu^T - ux^T]), && (4.27) \end{aligned}$$

where equation (4.27) follows from the fact $y \in \mathcal{S}$ and $\mathcal{P}$ is a projection on $\mathcal{S}$. Substituting for $x(t)$ and $u(t)$ we get

$$\begin{aligned} \frac{d\|y\|^2}{dt} &= -4\ \{\ q^2 tr(p^T y y^T p) + q^2 tr(\bar{p}^T y p)\\ &\quad + \|y\| tr(p^T y r) +\ q tr(\bar{p}^T y r)\\ &\quad + q tr(p^T y y^T r)\}. && (4.28) \end{aligned}$$

First, observe that $tr(\bar{p}^T yp) = 0$ from Lemma (4). Now using $tr(AB) < \|A\|\|B\|$, from equation (4.28), we get

$$\frac{d\|y\|^2}{dt} \leq -4q^2\|yp\|^2 + 4\|q\|\|y\|\|yp\|\|r\| + 4\|y\|\|yp\|\|r\| + 4\|q\|\|\bar{p}\|\|y\|\|r\| \tag{4.29}$$

$$\frac{d\|y\|^2}{dt} \leq -2(\|q\|\|yp\| - \|y\|\|r\|)^2 - 2q^2\|yp\|^2 + 2\|y\|^2\|r\|^2 + 4\|y\|\|yp\|\|r\| + 4\|q\|\|\bar{p}\|\|y\|\|r\| \tag{4.30}$$

$$\frac{d\|y\|}{dt} \leq \|y\|\|r\|^2 + 2\|yp\|\|r\| + 2\|q\|\|\bar{p}\|\|r\| \tag{4.31}$$

Because $\|\bar{p}(t)\|$ and $\|p(t)\|$ is bounded by the definition of selector function and $r(t) = r(0)e^{-t}$, equation (4.31) can be written as

$$\frac{d\|y\|}{dt} \leq A\|q\| e^{-t} + B\|y\|e^{-t}. \tag{4.32}$$

for positive constants $A$ and $B$. From equations (4.32) and (4.24), we can deduce that $\|y(t)\|$ is bounded and $q(t)$ is bounded. Therefore, for the given initial condition $(y(0), q(0))$, there exists $M < \infty$ such that $\|y(t)\| < M$ and $\|q(t)\| < M, \forall t$. Hence, we can rewrite equation (4.29) as

$$\frac{d\|y\|^2}{dt} \leq -4q^2\|yp\|^2 + M_1 e^{-t}, \tag{4.33}$$

for some positive constant $M_1$ which depends on the initial condition $(x(0), y(0), q(0))$. Defining $V(t) = \|y(t)\|^2 + M_1 e^{-t}$, observe that we have $\dot{V}(t) \leq -4q^2\|yp\|^2 \leq 0$. As $V(t) \geq 0$ and non-increasing, it follows that $\lim_{t\to\infty} \dot{V}(t) = 0$, implying that $\lim_{t\to\infty} \frac{d\|y(t)\|}{dt} = 0$, i.e. $\lim_{t\to\infty} y(t) = y_0$ for some $y_0 \in so(n)$. Therefore, by equation (4.24), $\lim_{t\to\infty} q(t) = \|y_0\|$. We now argue that $\|y_0\| = 0$.
As $\lim_{t\to\infty} \frac{d\|y(t)\|}{dt} = 0$, from equation (4.33), we obtain $lim_{t\to\infty}\|y_0\|^2\|y_0 p(t)\|^2 = 0$. Since $p(t) = \Theta(t)e(t)$, where $e(t)$ is a selector function, we conclude, from Lemma, 5 that $\|y_0\| = 0$. Therefore, $\lim_{t\to\infty} q(t) = 0$ and, from $x(t) = qp + ae^{-t}$, it follows $\lim_{t\to\infty} x(t) = 0$. **Q.E.D**.

A simulation of the stabilization of the general position-area system, using the feedback control law given in Theorem 18 is shown in Figure 4.2, where the following selector

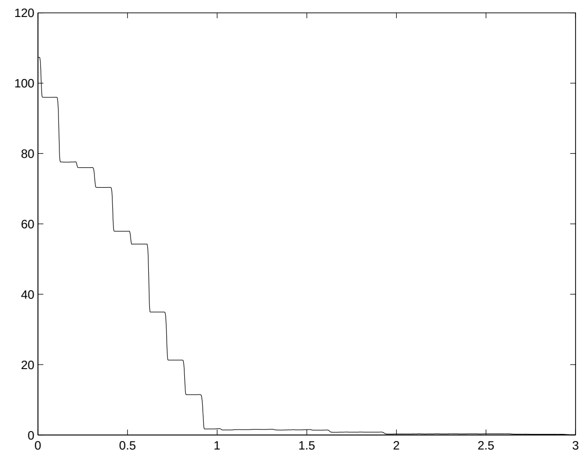


Figure 4.2: The panel shows the result of stabilization of the position-area system for $n = 10$, with the graphs showing the evolution of $\|y\|$ plotted against time.

function $e(t) \in \mathbb{R}^n$ with period $T = 1$ was used in the simulations

$$\begin{aligned} e_1(t) &= sin^2(\frac{\pi n t}{T}) \ ; \ 0 \le t \le \frac{T}{n} \\ &= 0 \ , \ \frac{T}{n} < t \le T \\ e_{k+1}(t) &= e_k(t - \frac{T}{n}) \ , \ k\{\in 1 \dots n\}. \end{aligned}$$

We now extend the result of Theorem 18 to the class of drift-free systems $\dot{z} = B(z)u$, which are first bracket controllable. As shown in BROCKETT (1981), for such systems, we can choose coordinates in a neighborhood of a point, say $z = 0$, so that equations of motion take the form (4.39)-(4.40). In the following theorem, we present feedback laws that asymptotically stabilize such systems. First, we state a modification of the result due to Krasovskii. For details see HAHN (1950a).

**Theorem 19** Let $M$ be a compact differentiable manifold, $x \in \mathbb{R}^n$, and $\theta \in M$. Let

$$\begin{aligned} \dot{x} &= f(x, \theta, t) \quad (f \in E) \\ \dot{\theta} &= g(x, \theta, t) \quad (g \in E). \end{aligned}$$

The existence of a Lyapunov function $v(x, \theta, t)$ satisfying inequalities of the form

$$v < a_1 \|x\|^\gamma, \quad \dot{v} < -a_2 \|x\|^{\gamma+\eta} \tag{4.34}$$

for sufficiently small $\|x\|$ is necessary and sufficient for the solution $(x(t), \theta(t))$ of the differential equation (4.34) to satisfy an estimate of the form

$$\|x(t)\|^{-\eta} - \beta\|x(0)\|^{-\eta} > \alpha t, \quad \forall t \geq 0, \tag{4.35}$$

for small initial values $\|x(0)\|$. Here, $a_1$, $a_2$, $\eta$, $\gamma$, $\alpha$ and $\beta$ are positive constants.
In case an estimate of the type (4.35) is satisfied, then $v \in C_1$ can be determined such that, in addition to (4.34), the inequality

$$\|\frac{\partial v}{\partial x_i}\| < a_3\|x\|^{\gamma-1} \quad (i = 1, \ldots, n; a_3 > 0, \gamma > 1) \tag{4.36}$$

is valid. As a result, the solutions of the modified differential equation

$$\begin{aligned} \dot{x} &= f(x, \theta, t) + h(x, \theta, t) \quad (f + h \in E) && (4.37) \\ \dot{\theta} &= g(x, \theta, t) && (4.38) \end{aligned}$$

with sufficiently small initial values $\|x(0)\|$ also satisfy an estimate of the form (4.35) if $\lim_{\|x\| \to 0} \frac{\|h(x,\theta,t)\|}{\|x\|^{\eta+1}} = 0$.

**Remark 9** We now present a constructive solution to feedback stabilization of a system of the form

$$\dot{z} = B(z)u \quad , z \in R^n,$$

which is first bracket controllable. It was shown by Brockett (1981) that such a system can be approximated near origin by a system

$$\begin{aligned} \dot{x} &= u + R(x, y, u) \\ \dot{y} &= \mathcal{P}[xu^T - ux^T] + R^1(x, y, u), \end{aligned}$$

where $R(x, y, u) \in \mathbb{R}^n$ and $R^1(x, y, u) \in so(n)$ have vanishing first partials with respect to $x$ and $y$ at the origin and be linear in $u$, such that $R(x, y, 0) = 0$, and $R^1(x, y, 0) = 0$. Therefore, we will find stabilizing feedback law for this system and, if the rate of convergence is large enough, we can claim asymptotic stability of the original system by invoking Theorem 19.

**Theorem 20** Let $\mathcal{S}$ be a subspace of $so(n)$ and $\mathcal{P} : so(n) \to so(n)$ be a projection operator onto this subspace. Let $x \in \mathbb{R}^n$, $u \in \mathbb{R}^n$, $q \in \mathbb{R}$, $\Theta \in O(n)$, $y \in \mathcal{S}$, and $e \in \mathcal{SF}(n, T, \epsilon)$. Let

$R(x,y,u) \in \mathbb{R}^n$ and $R^1(x,y,u) \in \mathcal{S}$ have vanishing first partials with respect to $x$ and $y$ at the origin and be linear in $u$, such that $R(x,y,0) = 0$, and $R^1(x,y,0) = 0$. If

$$u = -x + \|y\|\Theta e + q(\|y\|^{-\nu} y\Theta e + \Theta \dot{e}) \quad , \; 0 < \nu < 1$$

then, for

$$\begin{aligned}
\dot{x} &= u + R(x,y,u) && (4.39)\\
\dot{y} &= \mathcal{P}[xu^T - ux^T] + R^1(x,y,u) && (4.40)\\
\dot{q} &= -(q - \|y\|) && (4.41)\\
\dot{\Theta} &= \frac{y}{\|y\|^\nu}\,\Theta, && (4.42)
\end{aligned}$$

the submanifold $N = \{\ (x,y,q,\Theta) \in \mathbb{R}^n \times so(n) \times \mathbb{R} \times O(n) : x = 0, y = 0, q = 0\ \}$ is asymptotically stable.

**Proof:** First notice that, by definition, $R$ and $R^1$ satisfy

$$\lim_{\|x,y,q\| \to 0} \frac{R}{\|x,y,q\|^{3-\nu}} = 0 \quad , \quad \lim_{\|x,y,q\| \to 0} \frac{R^1}{\|x,y,q\|^{3-\nu}} = 0. \tag{4.43}$$

We will show that, for small initial values $\|x(0), y(0), q(0)\| < \delta_0$, the solutions to equations

$$\begin{aligned}
\dot{x} &= u && (4.44)\\
\dot{y} &= \mathcal{P}[xu^T - ux^T] && (4.45)\\
\dot{q} &= -(q - \|y\|) && (4.46)\\
\dot{\Theta} &= \frac{y}{\|y\|^\nu}\,\Theta \quad , \; 0 < \nu < 1 && (4.47)
\end{aligned}$$

satisfy, for $\alpha > 0$, $\beta > 0$ and $t > 0$, an estimate of the form

$$\|x(t), y(t), q(t)\|^{-2+\nu} - \beta \|x(0), y(0), q(0)\|^{-2+\nu} > \alpha t \tag{4.48}$$

which, using the result (19) and (4.43), proves the theorem.

Let $p = \Theta e$, and $r = x - qp$. From equations (4.44), (4.46), and (4.47), we get

$$\dot{r} = -r. \tag{4.49}$$

From equations (4.49) and (4.45), we obtain

$$\begin{aligned}
\frac{d\ \|y\|^2}{dt} &= -4\{q^2\ \|y\|^{-\nu}\ \|yp\|^2 + \|y\| tr(p^T yr) + \\
&\quad q\ \|y\|^{-\nu} tr(p^T y y^T r) + q\ tr(\bar{p}^T yr)\} && (4.50)
\end{aligned}$$

$$\frac{d\ \|y\|}{dt} \quad = \quad -2q^2\ \|y\|^{-(1+\nu)}\ \|yp\|^2 + o(r,y,q) \tag{4.51}$$

where $o(r,y,q) < b\|r,y,q\|$ for some positive $b$. Observe that for $r = 0$, equation (4.50) reduces to $\frac{d\ \|y\|^2}{dt} = -4q^2\ \|y\|^{-\nu}\|yp\|^2$. We will first show that the solutions to the system of reduced equations

$$\frac{d\ \|y\|^2}{dt} \quad = \quad -4q^2\ \|y\|^{-\nu}\|yp\|^2 \tag{4.52}$$
$$\dot{q} \quad = \quad -(q - \|y\|) \tag{4.53}$$
$$\dot{\Theta} \quad = \quad \frac{y}{\|y\|^\nu}\ \Theta \tag{4.54}$$

satisfy, for $\alpha_1 > 0$, $\beta_1 > 0$ and $t > 0$, an estimate of the form

$$\|y(t), q(t)\|^{-2+\nu} - \beta_1\|y(0), q(0)\|^{-2+\nu} > \alpha_1 \tag{4.55}$$

for sufficiently small $\|y(0), q(0)\|$. Observe that, from equation (4.45), if $r = 0$ then,

$$\|\dot{y}\| \quad \leq \quad 2q^2\ \|y\|^{-\nu}\|p\|\ \|yp\|. \tag{4.56}$$

From equation (4.53), we have

$$\begin{aligned} q(t) \quad &= \quad q(0)e^{-t} + \int_0^t e^{-(t-\tau)}\|y(\tau)\|d\tau \\ &\geq \quad (q(0) - \|y(0)\|)\ e^{-t} + \|y(t)\| \end{aligned} \tag{4.57}$$

where the last inequality follows from the fact that $\|y(t)\|$ is non-increasing function of time (equation (4.52)). Let $M = q(0) - \|y(0)\|$, from (4.52) and (4.57), we have

$$\begin{aligned} \frac{d\ \|y\|^2}{dt} \quad &\leq \quad -4\ \|y\|^{-\nu}\{Me^{-t} + \|y(t)\|\ \}^2\ \|yp\|^2 \\ \frac{d\ \|y\|}{dt} \quad &\leq \quad -2\|yp\|^2\|y\|^{1-\nu} + 4\ \|M\|\ \|y(0)\|^{2-\nu}\ e^{-t} \\ &\quad\quad +2\ M^2\ \|y(0)\|^{1-\nu}\ e^{-2t}. \end{aligned} \tag{4.58}$$

We first show that solutions to

$$\frac{d\ \|y\|}{dt} = -\|yp\|^2\|y\|^{1-\nu} \tag{4.59}$$

satisfy for $\alpha_2 > 0$, $\beta_2 > 0$ and $t > 0$,

$$\|y(t)\|^{-2+\nu} - \beta_2\|y(0)\|^{-2+\nu} > \alpha_2 t. \tag{4.60}$$

From equation (4.56) and Lemma 5, it follows that

$$\int_t^{t+T} \|yp(\tau)\| \; d\tau \quad \geq \quad \gamma \int_t^{t+T} \|y(\tau)\| \; d\tau$$

for some positive constant $\gamma$, implying that

$$\int_t^{t+T} \|yp(\tau)\|^2 \; \|y\|^{1-\nu} d\tau \geq \beta \int_t^{t+T} \|y(\tau)\|^{3-\nu} \; d\tau \tag{4.61}$$

for some positive constant $\beta$. Therefore (4.60) follows, and we can deduce (4.55) from (4.58) and (4.53). Using Result 19, there exists a Lyapunov function $v(y, q, \Theta)$ satisfying (4.34) and (4.36) for $\gamma = 2$ and $\eta = 2 - \nu$, where $\gamma$, $\eta$ as defined in Result 19. Consider the Lyapunov function

$$v_1(r, y, q, \Theta) = v(y, q, \Theta) + r^2.$$

Observe from equations (4.49), (4.51), (4.46) and (4.47), $v_1$ satisfies estimates of the form (4.34), for sufficiently small $\|r, y, q\|$, hence by Result 19, we conclude that the assertion (4.48) is valid and thus the proof of the theorem follows. **Q.E.D.**

## 4.3 Conclusions

Our main purpose of this chapter was to motivate the idea of dynamic feedback stabilization for nonholonomic systems in order to circumvent topological problems associated with smooth feedback stabilization of nonholonomic systems. We then illustrated our approach by giving a constructive solution to the unsolved problem of stabilizing a general first bracket controllable system. Our approach is natural as it addresses the topological issues directly. In this chapter, the choice of the auxiliary manifold and design of feedback laws were motivated by qualitative arguments. In the following chapter, we derive (in analogy to the LQG controller for linear systems) the dynamic feedback control laws as solutions to a variational problem, we call the *Riemannian Regulator problem.* This will give us a systematic way for designing feedback control laws.

# Chapter 5

# Dynamic Stabilization based on Gauge Extentions

## 5.1 Introduction

In this chapter, we introduce the concept of a control system, by which additional control degrees of freedom are introduced into the system dynamics by embedding the state space of the system in a higher dimensional manifold. The choice of higher dimensional space is dictated by the symmetries of the system. Our approach is analogous to, and inspired by, gauge transformations in physics, where interactions are brought into the system dynamics by lifting the global symmetries to local symmetries via introduction of a gauge group. We will discuss at some length the analogies between our approach and gauge theories in physics. We then study the problem of optimal feedback stabilization of nonholonomic systems, and we show that after we have make a suitable gauge extension, smooth feedback control for stabilizing nonholonomic systems can be obtained as solutions to variational problems, which we call the Riemannian regulator problem BROCKETT (2000). The chapter is organized as follows, we first introduce the Riemannian regulator problem and derive optimal control laws for stabilizing system trajectories. We then show how problems of stabilizing nonholonomic systems can be reduced to the above case by the process of gauge extension. To fix ideas, we analyze the stabilization of the nonholonomic integrator in this setting. Finally we extend the results to two important generalizations of the nonholonomic integrator. We present a variational solution to the problem of feedback stabilization of the *general position area system or the $so(n)$ system.* We also present another important generalization of the nonholonomic integrator, the $sl(n)$ system described by

$$\dot{x} = u \tag{5.1}$$
$$\dot{y} = xu - ux, \tag{5.2}$$

where $x,\ u \in sym(n)$, the space of $n \times n$ symmetric matrices and $y \in so(n),\ n \geq 2$.

Finally we will present our results in more generality, by considering a Lie algebra generalization of these two cases.

**Notation:** We will use $O(n)$ to denote the group of $n$-dimensional real orthogonal matrices, $so(n)$ the Lie algebra of $n \times n$ skew symmetric matrices, and $sym(n)$ the space of $n \times n$ symmetric matrices. We use the symbols $[.\ ,\ .]$ to denote the commutator and $\{.\ ,\ .\}$ the anticommutator between the matrices. Given the vector space $V$ with a Riemannian metric $G$ on it, we denote by $< \cdot\ ,\ \cdot >_G$ the scalar product between the elements. If $x = [x_1, \dots, x_n] \in \mathbb{R}^n$ and $\phi : \mathbb{R}^n \to \mathbb{R}$ is a differentiable function, we use $\phi_x$ to denote the gradient with respect to the standard Euclidean metric, $\phi_x = [\frac{\partial \phi}{\partial x_1}, \dots, \frac{\partial \phi}{\partial x_n}]^T \in \mathbb{R}$

## 5.2 Riemannian Regulator Problem

Let the control system $\dot{x} = B(x)u$ be given, where $x \in X$, a Riemannian manifold with metric $G$. Suppose for now, $B(x)$, is full rank $\forall x$, and the metric $G$ on $M$ is defined by $G^{-1}(x) = B(x)B^T(x)$. Let $\phi : M \to R$, be a differentiable function. Consider the infinite time optimization problem of finding $u(x,t)$ which minimizes

$$\eta = \int_0^\infty u^T u + \ < \nabla\phi, \nabla\phi >_G \ dt,$$

and stabilizes the system to $x_0$. Please note $\nabla\phi$ is the gradient with respect to the metric $G$, i.e $\nabla\phi = G^{-1}\phi_x$. We call this the Riemannian regulator problem. The solution is given by the following theorem Brockett (2000)

**Theorem 21** *(***Brockett:***)* Let $(X, G)$ be a Riemannian manifold and $\phi : X \to \mathbb{R}$ a positive definite differentiable function. Let the control system $\dot{x} = B(x)u$, $x \in X$ be given and let $G^{-1}(x) = B(x)B^T(x)$. Suppose $\phi_x = 0$ if and only if $x = x_0$, then the feedback control law

$$u = -B^T(x)\phi_x$$

stabilizes the system to $x_0$ and among all stabilizing feedback laws it minimizes the cost function

$$\eta = \int_0^\infty u^T u + < \nabla\phi, \nabla\phi >_G dt$$

**Proof:** We first prove stability of the control law. Let $\phi$ be the Lyapunov function, then

$$\frac{d\phi}{dt} = -(B(x)\phi_x)^T(B(x)\phi_x).$$

Therefore $\frac{d\phi}{dt} \leq 0$ and hence by La Salle's invariance principle all trajectories of the above control system go to the set $S = \{\frac{d\phi}{dt} = 0\}$. Observe $B(x)$ is full rank as the Riemannian metric $G = (BB^T)^{-1}$ is nonsingular. Therefore $\frac{d\phi}{dt} = 0$ if and only if $\phi_x = 0$, hence all trajectories are stabilized to $x_0$.

To prove optimality observe the cost function

$$\begin{aligned}
\eta &= \int_0^\infty u^T u + < \nabla\phi, \nabla\phi >_G dt \\
&= \int_0^\infty < \dot{x}, \dot{x} >_G + < \nabla\phi, \nabla\phi >_G dt \\
&= \int_0^\infty < \dot{x} + \nabla\phi, \dot{x} + \nabla\phi >_G -2 < \dot{x}, \nabla\phi >_G dt \\
&= \int_0^\infty < \dot{x} + \nabla\phi, \dot{x} + \nabla\phi >_G dt \ + 2(\phi(x(0)) - \phi(x_0))
\end{aligned}$$

where in deriving the last equality we have used the fact that $< \dot{x}, \nabla\phi >_G dt = d\phi$ and being a perfect differential just integrates out. As the system trajectories approach $x_0$, we have $\phi(x(\infty)) = \phi(x_0)$. Now observe the integral $\int_0^\infty < \dot{x} + \nabla\phi, \dot{x} + \nabla\phi >_G dt$ is minimized for $\dot{x} = -\nabla\phi$ and therefore the optimal feedback control law is $u = -B^T\phi_x$ and the return function for the feedback control law is $[\ 2\phi(x) - 2\phi(x_0)\ ]$. **Q.E.D**

We now state a slight modification of the above theorem, which addresses the case when the metric $G$ is singular.

**Theorem 22** Let $X$ be a differentiable manifold and $\phi : X \to \mathbb{R}$ a positive definite differentiable function. Let the control system $\dot{x} = B(x)u$, $x \in X$ be given. Let $G(x) = B(x)B^T(x)$ be a quadratic form on $X$. Let $\mathcal{N}(B^T(x))$ denote the null space of $B^T(x)$. Suppose $\phi_x \notin \mathcal{N}(B^T(x))$, $\forall x \neq x_0$, and $\phi_x = 0$ if $x = x_0$, then the feedback control law

$$u = -B^T(x)\phi_x$$

stabilizes the system to $x_0$ and among all stabilizing feedback laws it minimizes the cost function

$$\eta = \int_0^\infty u^T u + \phi_x^T G(x) \phi_x dt$$

**Proof:** The proof of stability is exactly along the same lines as theorem 21, by use of LaSalle's invariance and with $\phi$ as the the Lyapunov function, we get the result. To prove optimality, let $K = (B^T B)^{-1} B^T$ and let $\nabla\phi = B(x)B^T(x)\phi_x$ and express $u = K\dot{x}$. Observe that

$$\begin{aligned}
\eta &= \int_0^\infty u^T u + \phi_x^T G(x)\phi_x dt \\
&= \int_0^\infty < K\dot{x}, K\dot{x} > + \phi_x^T G(x)\phi_x dt \\
&= \int_0^\infty < K\dot{x} + K\nabla\phi, K\dot{x} + K\nabla\phi > -2 < K\dot{x}, K\nabla\phi > dt \\
&= \int_0^\infty < K\dot{x} + K\nabla\phi, K\dot{x} + K\nabla\phi > dt \ + 2(\phi(x(0)) - \phi(x_0))
\end{aligned}$$

where in deriving the last equality we have used the fact that $< K\dot{x}, K\nabla\phi > dt =< u, \ B^T d\phi > dt =< \dot{x}, d\phi > dt$, and being a perfect differential just integrates out and because of stability $\phi(x(\infty)) = \phi(x_0)$. Now observe the integral $\int_0^\infty < K\dot{x} + K\nabla\phi, K\dot{x} + K\nabla\phi > dt$ is minimized for $\dot{x} = -\nabla\phi$ and therefore the optimal feedback control law is $u = -B^T\phi_x$ and the return function for the feedback control law is $[\ 2\phi(x) - 2\phi(x_0)\ ]$. **Q.E.D**

**Remark 10** In keeping with the notation of theorem 21, we will use the notation

$$< \nabla\phi, \nabla\phi >_G = \phi_x^T G(x)\phi_x$$

for the remainder of the chapter.

We now show how the above theorems can be used to find optimal stabilizing laws for the nonholonomic systems.

## 5.3 Gauge Extension

In this section we introduce the concept of gauge extension of nonholonomic systems. The biggest challenge in smooth feedback stabilization of nonholonomic systems comes from the problem of fewer control degrees of freedom than the dimension of the state space of the system. We will show that by embedding the state space in a higher dimensional manifold, additional control degrees of freedom can be introduced in the system dynamics, so that we have as many controls as the state space dimension. We will call this approach gauge completion of nonholonomic systems. We will show that by using gauge completion, we can find optimal stabilization laws as solutions to Riemannian regulator type problems. This approach has striking parallelism to gauge theories in physics, where interactions are introduced into the free system dynamics by lifting local symmetries to global symmetries, thereby introducing a gauge potential, which transforms according to the gauge transformation laws. General relativity, Electromagnetism, and Yang Mills theories are some familiar examples. We will elaborate on this line of thought more, once the basic mathematical paradigm is established. We begin with some definitions

**Definition 10 Symmetry Group:** Let $X = R^n$ and $U = R^m$ represent the state and control space of the control system $\dot{x} = f(x, u)$, where $x \in X$ and $u \in U$. Let $H$ be a matrix lie group that acts on $X \times U$ effectively via group actions $\psi = (\psi_1, \psi_2) : H \times X \times U \to X \times U$. We call $H$ to be a symmetry group of the control system $\dot{x} = f(x, u)$, if the action of the group leaves the control system invariant, i.e if $h \in H$, and $(p, v) = \psi(h, x, u)$ then $\dot{p} = f(p, v)$.

With the control system $\dot{x} = f(x, u)$ and the action of the symmetry group $H$ on the space $X \times U$, we can associate a flow on $X$, we call the symmetry flow. This is defined as follows

**Definition 11 Gauge Extension and Gauge Controls:** Let $X = R^n$ and $U = R^m$ represent the state and control space of the control system $\dot{x} = f(x, u)$, where $x \in X$ and $u \in U$. Let $H$ be a symmetry group for this control system and let $\mathcal{H}$ denote the Lie algebra of left invariant vector fields on the lie group $H$. Each $\xi \in \mathcal{H}$ corresponds to a $w \in \mathfrak{h}$ where $\mathfrak{h}$ represents the tangent space $T_e$ at the identity element $e$ of the group. Corresponding to each $w \in \mathfrak{h}$, there is an induced vector field $g(\tilde{x}, w)$ on $X$ defined as follows. If $\tilde{x} = \psi_1(h, x, u)$, then $g(\tilde{x}, w)$ is obtained by differentiating $\psi_1(h\ exp(tw), x, u)$ with respect to $t$ at $t = 0$. Therefore the flow $\dot{x} = f(x, u)$ and $\dot{h} = \xi(h)$, under the mapping $\tilde{x} = \psi_1(h, x, u)$ and $v = \psi_2(h, x, u)$, induces a flow on $X$ of the form $\dot{\tilde{x}} = f(\tilde{x}, v) + g(\tilde{x}, w)$. This defines a new control system $\dot{\tilde{x}} = \tilde{f}(\tilde{x}, v, w)$, and we call this the gauge extension of the system $\dot{x} = f(x, u)$. The additional controls $w$, we get in the system dynamics will be referred to as gauge controls.

Thus observe that we have introduced additional control degrees of freedom $w$ in the control system, using the flow $\dot{h} = \xi(h)$.

**Example 3** Consider the $so(n)$ system defined in Section 5.1. For this system, the state space is $X = R^n \times so(n)$ and the control space is $U = R^n$. Let $\Theta \in O(n)$, then observe that the orthogonal group acts on $X \times U$ via the group action $\psi(\Theta, x, y, u) = (\Theta x, \Theta y \Theta^T, \Theta u)$. If $(p, z, v) = (\Theta x, \Theta y \Theta^T, \Theta u)$ then observe that

$$\begin{aligned} \dot{p} &= v \\ \dot{z} &= pv^T - vp^T \end{aligned}$$

Therefore $H = O(n)$ is a symmetry group of the above system. If $\dot{\Theta} = w\Theta$ then the map $(p, z, v) = (\Theta x, \Theta y \Theta^T, \Theta u)$ gives the gauge extended system

$$\begin{aligned} \dot{p} &= v + wp \\ \dot{z} &= pv^T - vp^T + [w, z] \end{aligned}$$

We now define gauge completion.

**Definition 12 Gauge Completion:** Let $X = R^n$ and $U = R^m$ represent the state and control space of the control system $\dot{x} = f(x, u)$, where $x \in X$ and $u \in U$ and $m < n$. Let $H$ be a symmetry group for the control system. The gauge extension $\dot{\tilde{x}} = \tilde{f}(\tilde{x}, v, w)$, is said to be a gauge completion of $\dot{x} = f(x, u)$, if the mapping $\tilde{f} : R^n \times R^m \times T_e H \to R^n$ maps some nghd of $(0, 0, 0) \in R^n \times R^m \times T_e H$ onto some nghd of 0 in $R^n$.

To illustrate this concept we look at some examples

**Example 4** Recall the nonholonomic integrator is defined by the system of equations

$$\dot{x_1} = u_1 \tag{5.3}$$

$$\dot{x_2} = u_2 \tag{5.4}$$

$$\dot{x_3} = x_1 u_2 - x_2 u_1 \tag{5.5}$$

Let $x = [x_1, x_2, x_3]^T$ and $u = [u_1, u_2]^T$, then the above equations 5.3-5.5 can also be written as $\dot{x} = B(x)u$ where

$$B(x) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -x_2 & x_1 \end{bmatrix}$$

Observe that the orthogonal group $SO(2)$ is a symmetry group for the system under the action defined as follows. Let $\Theta \in SO(2)$, writing $\Theta$ as

$$\Theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Consider the group action $\psi : SO(2) \times \mathbb{R}^3 \to \mathbb{R}^3$ defined by $f : (\Theta, x_1, x_2, x_3) \to (y_1, y_2, y_3)$ where

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{5.6}$$

$$y_3 = x_3 \tag{5.7}$$

and let $v = [v_1, v_2]^T$

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{5.8}$$

The flow defined by the following differential equation on $\mathbb{R}^3 \times SO(2)$

$$\dot{x} = B(x)u \tag{5.9}$$

$$\dot{\theta} = v_3 \tag{5.10}$$

induces the flow on $\mathbb{R}^3$ defined by

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -y_2 \\ 0 & 1 & y_1 \\ -y_2 & y_1 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \tag{5.11}$$

By $\tilde{v}$, we denote the augmented control vector $\tilde{v} = [v_1, v_2, v_3]^T$. Equation 5.11 can then be written as $\dot{y} = C(y)\tilde{v}$. Observe that the matrix $C(y)$ is full rank except when $y_1 = y_2 = 0$. Thus the flow $\dot{y} = C(y)\tilde{v}$ is a gauge completion of the flow $\dot{x} = B(x)u$.

We will now show how the gauge extension and results of the Theorem 22 can be used to find optimal feedback laws for stabilization of nonholonomic systems. Before doing that we first discuss some connections between our approach and gauge theories in physics.

### 5.3.1 Gauge Theories and Gauge Extension

We now try to elaborate on the analogies between gauge theories in physics and gauge completion. To elaborate our point we will treat in some detail the example from electromagnetism. The main principle that forms the basis of these theories is that by making global symmetries of the system, local symmetries, one incorporates an interacting potential also known as the gauge potential in the system dynamics, which transforms in a prescribed way to insure local invariance of the theory. This transformation law is also known as the gauge transformation. We elaborate on this using the following example from electromagnetism.

**Example 5 (U(1) gauge of Electromagnetism:)**Let $x \in R^3$ and $t \in R$, and let $\psi : R^4 \rightarrow C^1$. We use $\psi$ to denote the wavefunction of a particle in quantum mechanics which evolves according to the well known Schroedinger equation

$$i\frac{\partial\psi(x,t)}{\partial t} = \frac{1}{2m}(-i\nabla)^2\psi(x,t). \tag{5.12}$$

Observe that equation 5.12 has a $U(1)$ symmetry group that is if we let $\psi \rightarrow \psi' = e^{i\alpha}\psi$, where $\alpha$ is a constant, then the equation for $\psi'$ remains unchanged

$$\frac{\partial\psi'(x,t)}{\partial t} = \frac{i}{2m}(\nabla)^2\psi'(x,t).$$

Now lets see what happens when we want this global phase invariance to be a local invariance, that is we now want the above equation to remain unchanged under a space-time dependent phase transformation $\psi \rightarrow \psi' = e^{i\alpha(x,t)}\psi$. It is easy to see that this local phase change is not an invariance of the free particle wave equation 5.12. If we wish to satisfy the demands of local phase invariance then we need-alter this free equation 5.12. It can be easily shown that the modified equation takes the form

$$(\frac{i\partial}{\partial t} - qV)\psi(x,t) = \frac{1}{2m}(-i\nabla - qA)^2\psi(x,t) \tag{5.13}$$

where $(A, V)$ are the gauge potentials which transform in the following way

$$\begin{aligned} A \rightarrow A' &= A + q^{-1}\nabla\alpha(x,t) && (5.14) \\ V \rightarrow V' &= V - q^{-1}\frac{\partial\alpha(x,t)}{\partial t} && (5.15) \end{aligned}$$

when we modify $\psi \rightarrow \psi' = e^{i\alpha(x,t)}\psi$. This modified equation has the right local invariance properties and this transformation law is called the gauge transformation.

Let us look at the concept of gauge completion from this perspective. We again investigate the general position area example and interpret it from the gauge theory point of view.

**Example 6 (O(n) gauge of position-area system:)** Recall that the $so(n)$ system is defined by

$$\begin{aligned}\dot{x} &= u \ ; \ x \in R^n \quad u \in R^n \\ \dot{y} &= xu^T - ux^T \ ; \ y \in so(n)\end{aligned}$$

We have seen that $O(n)$ is a symmetry group of the above system. Lets see how does this equation needs to be changed if we relax this global rotation, and let it be a local symmetry of the system, i.e if we want the system to be invariant under a coordinate change $x \to p = \Theta(t)x$, $u \to v = \Theta(t)u$, and $y \to z = \Theta(t)x\Theta^T(t)$. For this to be true one can verify that the modified equation takes the form

$$\begin{aligned}\dot{x} &= u + wx \\ \dot{y} &= xu^T - ux^T + [w, y]\end{aligned}$$

where $w \in so(n)$ is analogous to the gauge potential and transforms as $w \to w' = \dot{\Theta}\Theta^T(t) + \Theta(t)w\Theta^T(t)$. Under this gauge transformation law $O(n)$ is a local symmetry group of the above system. Observe that this modified equation is the gauge extension of the $so(n)$ system under the symmetry group $O(n)$.

We will now present smooth optimal feedback laws for stabilization of a class of nonholonomic systems, where the first derived algebra of vector fields span the tangent space at all points of the state space. We begin with the stabilization of nonholonomic integrator.

## 5.4 Stabilization of the Nonholonomic Integrator

Recall from Example 4 that the nonholonomic integrator defined by

$$\begin{aligned}\dot{x_1} &= u_1 \\ \dot{x_2} &= u_2 \\ \dot{x_3} &= x_1u_2 - x_2u_1\end{aligned}$$

has $SO(2)$ as a symmetry group and the gauge extension was shown to be

$$\frac{d}{dt}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -y_2 \\ 0 & 1 & y_1 \\ -y_2 & y_1 & 0 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

Our strategy will be the following. We will find the stabilizing control $v = [v_1, v_2, v_3]$ for the gauge extended system as a solution to the Riemannian regulator problem. Since the symmetry group $SO(2)$ preserves norms under the group action, this flow lifts to a unique flow on $SO(2) \times \mathbb{R}^3$ and stabilizes $[x_1, x_2, x_3]$.

**Theorem 23** Let $y = [y_1, y_2, y_3]^T \in \mathbb{R}^3$, and $v = [v_1, v_2, v_3]^T \in \mathbb{R}^3$. Let the control system $\dot{y} = C(y)v$ be given, where

$$C(y) = \begin{bmatrix} 1 & 0 & -y_2 \\ 0 & 1 & y_1 \\ -y_2 & y_1 & 0 \end{bmatrix}.$$

Consider the quadratic $G$ on $\mathbb{R}^3$ defined by $G(y) = C(y)C^T(y)$. Let $\phi : \mathbb{R}^3 \to \mathbb{R}$, be a positive definite differentiable function. Let $S = \{(y_1, y_2, y_3) \in \mathbb{R}^3 : y_1 = 0, y_2 = 0, y_3 \neq 0\}$. If $\phi_{y_1}^2 + \phi_{y_2}^2 > 0$ for $y \in S$, then the feedback control law $v(y) = -C^T(y)\phi_y$, stabilizes the system, and among all stabilizing control laws for the system $\dot{y} = C(y)v$, it minimizes the cost functional defined by

$$\eta = \int_0^\infty v^T v + \ < \nabla\phi, \nabla\phi >_G \ dt.$$

**Proof:** Optimality of the control follows from Theorem 22. To prove stability observe that $C(y)$ is full rank except on the set $S$, and the null space is spanned by the vector $[0, 0, 1]^T$. Therefore if $\phi_{y_1}^2 + \phi_{y_2}^2 > 0$ for $y \in S$, then $\phi_y \notin \mathcal{N}(C^T(y))$ for $y \neq 0$ and therefore by Theorem 22, we get stability. **Q.E.D.**

**Corollary 4** Let $y = [y_1, y_2, y_3]^T \in \mathbb{R}^3$, and $v = [v_1, v_2, v_3]^T \in \mathbb{R}^3$. Let the control system $\dot{y} = C(y)v$, where

$$C(y) = \begin{bmatrix} 1 & 0 & -y_2 \\ 0 & 1 & y_1 \\ -y_2 & y_1 & 0 \end{bmatrix}.$$

Consider the quadratic form $G$ on $\mathbb{R}^3$ be defined by $G(y) = C(y)C^T(y)$. Let $\phi(y_1, y_2, y_3) = \frac{1}{2}(y_1^2 + y_2^2 + y_3^2) + y_1 y_3 + y_2 y_3$. The feedback control law

$$v(y) = -\begin{bmatrix} 1 & 0 & -y_2 \\ 0 & 1 & y_1 \\ -y_2 & y_1 & 0 \end{bmatrix} \begin{bmatrix} y_1 + y_3 \\ y_2 + y_3 \\ y_1 + y_2 + y_3 \end{bmatrix}$$

stabilizes the system trajectories, and among all stabilizing control laws for the system $\dot{y} = C(y)v$, it minimizes the cost functional defined by

$$\eta = \int_0^\infty v^T v + < \nabla\phi, \nabla\phi >_G .$$

**Proof:** $\phi$ satisfies all the conditions in theorem 23. **Q.E.D.**

We now extend the stabilization results to the general position area system.

## 5.5 Stabilization of the General Position Area System

Recall, as shown in Example 4, that the general position area system defined by

$$\begin{aligned} \dot{x} = u \ \ ; \ x \in \mathbb{R}^n \\ \dot{y} = xu^T - ux^T \ \ ; \ y \in so^n \end{aligned}$$

has $O(n)$ as symmetry group. Recall that the action of $O(n)$ on the above system is as follows. Let $\Theta \in O(n)$, then $x \to p = \Theta x$, $y \to z = \Theta y \Theta^T$, and $u \to v = \Theta u$. Recall that under this symmetry group action, the flow $\dot{\Theta} = w\Theta$, induces the gauge extension

$$\dot{p} = v + wp \ ; \ p \in \mathbb{R}^n \tag{5.16}$$

$$\dot{z} = pv^T - vp^T + [w, z] \ ; \ z \in so^n \tag{5.17}$$

Thus we observe that by making the symmetries of the system time varying, we have introduced additional $\frac{n(n-1)}{2}$ controls into the system dynamics, through the skew symmetric matrix $w$. Thus we have transformed the problem of controlling $\frac{n(n+1)}{2}$ state variable with just $n$ controls, to a problem with as many controls $(v, w)$ as the number of state variables $(p, z)$. We now show that analogous to the case of the nonholonomic integrator, we can carry out the gauge completion program for the general position area system.

**Lemma 6** The gauge extension

$$\begin{aligned} \dot{p} &= v + wp \ ; \ p \in \mathbb{R}^n \ \ ; \ w \in so^n \\ \dot{z} &= pv^T - vp^T + [w, z] \ ; \ z \in so^n \end{aligned}$$

is a gauge completion of

$$\begin{aligned} \dot{p} &= v \ ; \ p \in \mathbb{R}^n \\ \dot{z} &= pv^T - vp^T \ ; \ z \in so^n \end{aligned}$$

**Proof :** For the case $n = 2$, the system is reduced to the nonholonomic integrator case, and we saw there that the above assertion was true. For $n > 2$, it suffices to observe that $\mathcal{L} = so(n)$ for $n > 2$ is a semi-simple Lie algebra. Therefore the first derived algebra $[\mathcal{L}, \mathcal{L}] = \mathcal{L}$, and hence we can find $z$ and $w$ such that $[z, w] = w'$ for some given $w'$. Let $p = 0$ and then we are free to choose $v$. Hence equations 5.16-5.17 satisfy the onto condition for stability, which concludes the proof

We will now present gradient flows for the stabilization of the above system (5.16)-(5.17). Let $L = (p, z) \in \mathbb{R}^n \times so(n)$ and $r = (v, w) \in \mathbb{R}^n \times so(n)$, then the system (5.16)-(5.17) can be compactly written as $\dot{L} = C(L)r$. In the spirit of theorem 22, we now want to find a positive definite function $\phi : \mathbb{R}^n \times so(n) \to \mathbb{R}$, such that the control law $r(L) = -C^T(L)\phi_l$ stabilizes the system. This is done in the following theorem.

**Theorem 24** Let $L = (p, z) \in \mathbb{R}^n \times so(n)$ and $r = (v, w) \in \mathbb{R}^n \times so(n)$. Let the control system $\dot{L} = C(L)r$ be defined by

$$\begin{aligned} \dot{p} &= v + wp \qquad (5.18) \\ \dot{z} &= pv^T - vp^T + [w, z] \qquad (5.19) \end{aligned}$$

Let $\phi : \mathbb{R}^n \times so(n) \to \mathbb{R}$ be $\phi(p, z) = \|p + ze\|^2 + tr(zz^T N)$, where $e = \frac{1}{\sqrt{n}}[1,\ 1,\ \dots,\ 1]^T$ is a unit vector and $N$ is a diagonal matrix, $N = diag(1, 2, \dots, n)$. Consider the quadratic form $G(L) = C(L)C^T(L)$ on $\mathbb{R}^n \times so(n)$. Then the feedback control law $r(L) = -C^T(L)\phi_L$ stabilizes the system to $(p = 0, z = 0)$ and among all feedback laws that stabilize system trajectories to $(p = 0, z = 0)$, it minimizes the cost function

$$\eta = \int_0^\infty \|v\|^2 + \|w\|^2 + < \nabla\phi, \nabla\phi >_G .$$

**Proof:** We first prove stability. First observe that

$$C^T(v, w) = (v - 2wp, -\frac{pv^T - vp^T}{2} - [w, z]).$$

Let $\phi$ be the lyapunov function, then

$$\frac{d\phi}{dt} = - < \phi_L, C(L)C^T(L)\phi_L >,$$

which allows us to rewrite the above equation as

$$\frac{d\phi}{dt} = -\|C^T(L)\phi_L\|^2.$$

All we need to show is that $\phi_L \in \mathcal{N}(C^T(L))$ only if $L = 0$. Observe that

$$C^T(L)\phi_L = (\phi_p - 2\phi_z p, -(\frac{p\phi_p^T - \phi_p p^T}{2} + [\phi_z, z])) \tag{5.20}$$

Note, if $\phi_L \in \mathcal{N}(C^T(L))$, then

$$\begin{aligned} \phi_p &= 2\phi_z p && (5.21) \\ \frac{p\phi_p^T - \phi_p p^T}{2} + [\phi_z, z] &= 0 && (5.22) \end{aligned}$$

Observe that from Equations (5.21)-(5.22), we obtain that if $\phi_L \in \mathcal{N}(C^T(L))$ then

$$pp^T\phi_z^T - \phi_z pp^T + [\phi_z, z] = 0 \tag{5.23}$$

from which follows that

$$\begin{aligned} \phi_z p &= 0 && (5.24) \\ [\phi_z, z] &= 0 && (5.25) \end{aligned}$$

To see this multiply both sides of Equation 5.23 by $\phi_z$ and take the trace. Therefore from Equation (5.21), $\phi_p = 0$. Notice that $\phi_p = 2(p + ze)$, therefore if $\phi_p = 0$ then $p = -ze$. Notice that

$$\phi_z = pe^T - ep^T + \{N + ee^T, z\}.$$

Substituting for $p = -ze$ we obtain $\phi_z = \{N, z\}$. Substituting in 5.23, $[\phi_z, z] = 0$, we obtain $[\{N, z\}, z] = 0$ implying $[N, z^2] = 0$, and therefore $z^2$ is a diagonal matrix. Therefore if $\phi_z p = 0$ then $\{N, z\}ze = 0$. Since $z^2$ is diagonal, we obtain $\{N, z\}ze = 0$ if and only if $z = 0$ and hence $p = 0$. Hence the proof. The optimality of the control law for the given cost function follows from theorem (22). **Q.E.D.**

We can now lift this flow to $R^n \times so(n) \times O(n)$ and since $\|p\| = \|\Theta x\| = \|x\|$ and $\|z\| = \|\Theta y \Theta^T\| = \|y\|$, the control law $u(x, y) = \Theta^T v(\Theta x, \Theta y \Theta^T)$ stabilizes the trajectories of the system

$$\begin{aligned} \dot{x} &= u && (5.26) \\ \dot{y} &= xu^T - ux^T && (5.27) \\ \dot{\Theta} &= w(\Theta x, \Theta y \Theta^T)\Theta && (5.28) \end{aligned}$$

to $x = 0, y = 0$. We now extend the stabilization results to the general $sl(n)$ system.

## 5.6 General sl(n) System

Recall that the $sl(n)$ system is defined by

$$\dot{x} = u \;\; ; \;\; x \in sym(n)$$
$$\dot{y} = xu - ux \;\; ; \;\; y \in so^n$$

Consider the map $F$ mapping $M = sym(n) \times so(n) \times O(n)$ to $N = sym(n) \times so(n)$ , defined as $(p, r) = F(x, y, \Theta)$, where

$$\begin{aligned} p &= \Theta\, x\Theta^T \;\; ; \;\; p \in sym(n) \\ z &= \Theta\, y\, \Theta^T \;\; ; \;\; z \in so^n \end{aligned}$$

Let $v = \Theta\, u\Theta^T$, and $\dot{\Theta} = w\Theta$, then we have

$$\dot{p} = v + [w, p] \;\; ; \;\; p \in sym(n) \tag{5.29}$$
$$\dot{z} = [p, v] + [w, z] \; ; \;\; z \in so^n \tag{5.30}$$

We now proceed as in the previous section to find optimal feedback laws that stabilize the system of Equations 5.29-5.30.

**Theorem 25** Let $L = (x, z) \;\in\; sym(n) \times so(n)$ and $v = (u, w) \;\in sym(n) \times so(n)$. Let the control system $\dot{L} = C(L)v$ be defined by

$$\dot{x} = u + [w, x] \tag{5.31}$$
$$\dot{z} = [x, u] + [w, z] \tag{5.32}$$
$$\tag{5.33}$$

Let $R$ be a $n \times n$ real matrix defined as

$$R = \frac{1}{\sqrt{n}} \begin{bmatrix} 0 & 1 & & & \\ 1 & 0 & & & \\ & & 0 & 1 & \\ & & 1 & 0 & \\ & & & & \ddots \end{bmatrix}$$

and $N$ a diagonal matrix, $N = diag(1, 2, \dots, n)$. Let $\phi : sym(n) \times so(n) \to \mathbb{R}$ be $\phi(x, z) = \|x + [z, R]\|^2 + tr(zz^T N)$. Let $G(L) = C(L)C^T(L)$ be a quadratic form on $sym(n) \times so(n)$.

Then the feedback control law $v(L) = -C^T(L)\phi_L$ stabilizes the system to $(x = 0, z = 0)$ and minimizes the cost function

$$\eta = \int_0^\infty \|u\|^2 + \|w\|^2 + < \nabla\phi, \nabla\phi >_G .$$

**Proof:** We first prove stability. Let $\phi$ be the Lyapunov function, then

$$\frac{d\phi}{dt} = - < \phi_L, C(L)C^T(L)\phi_L >$$

which allows us to rewrite the above equation as

$$\frac{d\phi}{dt} = -\|C^T(L)\phi_L\|^2.$$

All we need to show is that $\phi_L \in \mathcal{N}(C^T(L))$ only if $L = 0$. Observe

$$C^T(L)\phi_L = (\phi_x - [\phi_z, x], -([x, \phi_x] + [\phi_z, z])) \tag{5.34}$$

Note if $\phi_L \in \mathcal{N}(C^T(L))$ then

$$\begin{aligned} \phi_x &= [\phi_z, x] && (5.35)\\ [x, \phi_x] + [\phi_z, z] &= 0 && (5.36)\\ & && (5.37) \end{aligned}$$

Observe that from Equation (5.55)-(5.56), it follows that

$$[x, [\phi_z, x]] + [\phi_z, z] = 0$$

implying that $[\phi_z, x] = 0$ (multiply both sides of equation by $\phi_z$ and take the trace and therefore $[\phi_z, z] = 0$.) Therefore from Equation (5.55), we obtain $\phi_x = 0$. Notice that $\phi_x = 2(x + [z, R])$, therefore if $\phi_x = 0$ then $x = -[z, R]$. Also it holds that

$$\phi_z = [x, R] + [[z, R], R] + \{N, z\}.$$

Substituting for $x = -[z, R]$, we get $\phi_z = \{N, z\}$. Since $[\phi_z, z] = 0$ implies that $[\{N, z\}, z] = 0$, which is $[N, z^2] = 0$, that is $z^2$ is diagonal. $z^T z$ is diagonal implies that

$$z = \begin{bmatrix} 0 & \lambda_1 & & & \\ -\lambda_1 & 0 & & & \\ & & 0 & \lambda_2 & \\ & & -\lambda_2 & 0 & \\ & & & & \ddots \end{bmatrix}$$

from which follows that

$$[z,R] = \begin{bmatrix} \lambda_1 & 0 & & & \\ 0 & -\lambda_1 & & & \\ & & \lambda_2 & 0 & \\ & & 0 & -\lambda_2 & \\ & & & & \ddots \end{bmatrix}.$$

Therefore $\|[z,R]\| = \frac{\|z\|}{\sqrt{n}}$. Note $[\phi_z, x] = [\{N, z\}, [z, R]]$, therefore if $[\phi_z, x] = 0$, implies $\|z\| = 0$ and hence $\|x\| = 0$. Thus the proof. The optimality of the control law for the given cost function follows from Theorem (22). **Q.E.D.**

We now extend the results for the above system to systems which have higher degree of nonholonomy. We first introduce the following notation.

**Notation:** Let $x,\ u \in sym(n)$. By $[x^k, u]$, we will denote the $k^{th}$ order lie bracket $[x,[x,\ldots[x,k]]$.

Consider the following generalization of the $sl(n)$ system. Let

$$\begin{aligned} \dot{x} &= u \ ;\ x \in sym(n) \\ \dot{y} &= [x^{2k-1}, u] \ ;\ y \in so^n,\ k \in 1,\ldots,n \end{aligned}$$

Consider the map $F$ mapping $M = sym(n) \times so(n) \times O(n)$ to $N = sym(n) \times so(n)$ , $(p,r) = F(x,y,\Theta)$

$$\begin{aligned} p &= \Theta\ x\Theta^T \ ;\ p \in sym(n) \\ z &= \Theta\ y\ \Theta^T \ ;\ z \in so^n \end{aligned}$$

Let $v = \Theta\ u\Theta^T$, and $\dot{\Theta} = w\Theta$, then we have

$$\dot{p} = v + [w,p] \ ;\ p \in sym(n) \tag{5.38}$$

$$\dot{z} = [p^{2k-1}, v] + [w,z] \ ;\ z \in so^n \ , \quad k \in 1,\ldots,n \tag{5.39}$$

We now proceed as in previous section to find optimal feedback laws that stabilize the system of Equations 5.38-5.39.

**Theorem 26** Let $L = (x, z) \in sym(n) \times so(n)$ and $v = (u, w) \in sym(n) \times so(n)$. Let the control system $\dot{L} = C(L)v$ be defined by

$$\dot{x} = u + [w, x] \tag{5.40}$$
$$\dot{z} = [x^{2k-1}, u] + [w, z] \; ; \; k \in 1, \dots, n \tag{5.41}$$
$$\tag{5.42}$$

Let $R$ be a $n \times n$ real matrix defined as

$$R = \frac{1}{\sqrt{n}} \begin{bmatrix} 0 & 1 & & & \\ 1 & 0 & & & \\ & & 0 & 1 & \\ & & 1 & 0 & \\ & & & & \ddots \end{bmatrix}$$

and $N$ a diagonal matrix, $N = \ diag(1, 2, \dots, n)$. Let $\phi : sym(n) \times so(n) \to \mathbb{R}$ be $\phi(x, z) = \|x + [z, R]\|^2 + tr(zz^T N)$. Let $G(L) = C(L)C^T(L)$ be a quadratic form on $sym(n) \times so(n)$. Then the feedback control law $v(L) = -C^T(L)\phi_L$ stabilizes the system to $(x = 0, z = 0)$ and among all feedback laws that stabilize system trajectories to $(x = 0, z = 0)$, it minimizes the cost function

$$\eta = \int_0^\infty \|u\|^2 + \|w\|^2 + < \nabla\phi, \nabla\phi >_G .$$

The proof is on exactly same lines as theorem (25).

We now present the above results in much more generality. We will look at a Lie algebraic generalization of $so(n)$ and the $sl(n, \mathbb{R})$ systems.

## 5.7 Lie Algebra Generalization

We now present a Lie algebra generalization of the previous results for the $so(n)$ and the $sl(n, \mathbb{R})$ systems (Bloch et al. (1998)).

Let $\mathfrak{g}$ be a real semisimple Lie algebra, and let $B : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ be the killing form on $\mathfrak{g}$. Furthermore assume that $\mathfrak{g}$ has a direct decomposition $\mathfrak{g} = \mathfrak{h} + \mathfrak{m}$, where $\mathfrak{h}$ is a compactly embedded sub-algebra and the subspace $\mathfrak{m}$ is the orthogonal complement of $\mathfrak{h}$ relative to $B$. Under these assumptions, the commutation relations $[\mathfrak{h}, \mathfrak{m}] \subseteq \mathfrak{m}$ and $[\mathfrak{m}, \mathfrak{m}] \subseteq \mathfrak{h}$, and the restriction $B|_{\mathfrak{h} \times \mathfrak{h}}$, of the killing form $B$ to $\mathfrak{h} \times \mathfrak{h}$, is negative definite. In addition we assume that no ideal of $\mathfrak{g}$, is contained in $\mathfrak{h}$. This assumption, together with the commutation relations for $\mathfrak{h}$ and $\mathfrak{m}$, imply that $[\mathfrak{m},\mathfrak{m}]=\mathfrak{h}$. In particular the representation of $\mathfrak{h}$ on $\mathfrak{m}$ is faithful as the kernel of this representation is $B$ orthogonal in $\mathfrak{h}$ to $[\mathfrak{m},\mathfrak{m}]$. We now consider the stabilization of the following system

$$\dot{x} = u \tag{5.43}$$
$$\dot{y} = [u,x] \tag{5.44}$$

where $x,u\in\mathfrak{m}$, $y\in\mathfrak{h}$. It is possible to analyze this system without any loss of generality by considering more specialized types of Lie algebra. Using the hypotheses on $\mathfrak{g}$, there exist $B$ orthogonal ideals $\mathfrak{g}_+$ and $\mathfrak{g}_-$, with the following properties.

- $\mathfrak{g}=\mathfrak{g}_+ + \mathfrak{g}_-$ (direct sum)

- $\mathfrak{h}_\pm = \mathfrak{g}_\pm \bigcap \mathfrak{h}$ is compactly embedded in $\mathfrak{g}_\pm$ and contains no ideal of $\mathfrak{g}_\pm$.

- $\mathfrak{g}_+ = \mathfrak{h}_+ + \mathfrak{m}_+$ is of non-compact type and $\mathfrak{g}_- = \mathfrak{h}_- + \mathfrak{m}_-$ is of compact type where $\mathfrak{m}_\pm = \mathfrak{g}_\pm \bigcap \mathfrak{m}$.

If we let $x=x_+ + x_-$, $u=u_+ + u_-$ and $y=y_+ + y_-$ denote the corresponding decompositions, then the system 5.43 decomposes into the systems

$$\dot{x_\pm} = u_\pm \tag{5.45}$$
$$\dot{y_\pm} = [u_\pm, x_\pm] \tag{5.46}$$

It follows to stabilize the system 5.43 in $\mathfrak{g}$ it is enough to stabilize simultaneously the systems 5.45. It will thus be assumed from now that $\mathfrak{g}$ is either of non-compact type or of compact type. This implies that the restriction $B|_{\mathfrak{m}\times\mathfrak{m}}$ of $B$ to $\mathfrak{m}$ is positive definite if $\mathfrak{g}$ is of non-compact type and negative definite if $\mathfrak{g}$ is of compact type.

In order to discuss the compact and non-compact cases simultaneously, let $\epsilon=1$ if $\mathfrak{g}$ is of non-compact type and $\epsilon=-1$ if $\mathfrak{g}$ is of compact type. We will use the inner product on $\mathfrak{g}$ defined by the Killing form:

$$< x_1+y_1, x_2+y_2 >= \epsilon B(x_1,x_2) - B(y_1,y_2) \tag{5.47}$$
$$\tag{5.48}$$

for $x_1,x_2\in\mathfrak{m}$, $y_1,y_2\in\mathfrak{h}$. The corresponding norm will be denoted by $\|.\|$.

Let $\Theta \in Aut(\mathfrak{g})$, the group of automorphisms of the Lie algebra $\mathfrak{g}$. Let $\Theta(0) = I$ the identity map and let $\dot{\Theta} = adw\ \Theta$, where $w \in \mathfrak{h}$. Observe that under this flow, $\Theta$ is constrained to live in $Aut(\mathfrak{g})$. Note the norm $(p, z) = \Theta(x, y)$ preserves the norm of $(x, y)$ and under the map

$$(p, z) = L(x, y),$$

we obtain from equation 5.43, that

$$\dot{p} = v + [w, p] \tag{5.49}$$

$$\dot{z} = [p, v] + [w, z] \tag{5.50}$$

We will now present feedback control laws $v(p, z)$ and $w(p, z)$, which asymptotically stabilize the above system.
For $Y \in \mathfrak{h}$, let

$$D(Y) = -(ad_Y)^2|_{\mathfrak{m}}$$

denote a nonnegative symmetric operator, acting on $m$. It can be easily shown that there exists a constant $0 < \eta < 1$, such that

$$tr(C(Y)) > \eta\|Y\|^2,$$

for all $Y \in \mathfrak{h}$ (Bloch et al. (1998)). If we choose a basis for $\mathfrak{m}$, then let $n \in \mathfrak{h}$ denote the unique element, such that the operator $D(n)$ corresponds to $N = diag(1, 1, 2, 2, \dots, m/2)$. Also Let $r \in \mathfrak{m}$ represent the unique vector that has the representation $\frac{1}{\sqrt{n}}[1,\ 1,\ \dots,\ 1]^T$.

**Remark 11** As in theorem 24, it follows, that for $Y \in \mathfrak{h}$, if $ad_n^2 ad_Y^2|_{\mathfrak{m}} = ad_Y^2 ad_n^2|_{\mathfrak{m}}$, then $ad_{[n,Y]}|_m = 0$. Also, if $D(Y)D(n)r = 0$ then it follows that $Y = 0$.

**Theorem 27** Let $L = (x, z)\ \in\ \mathfrak{m} + \mathfrak{h}$ and $v = (u, w)\ \in \mathfrak{m} + \mathfrak{h}$. Let the control system $\dot{L} = C(L)v$ be defined by

$$\dot{x} = u + [w, x] \tag{5.51}$$

$$\dot{z} = [x, u] + [w, z]\ ;\ k \in 1, \dots, n \tag{5.52}$$

$$\tag{5.53}$$

Let $n \in \mathfrak{h}$ and $r \in \mathfrak{m}$. Let $\phi :\to \mathbb{R}$ be $\phi(x, z) = \|x + [z, r]\|^2 + tr(ad_z^2 ad_n^2)$. Let $G(L) = C(L)C^T(L), X, Y\ \in\ \mathfrak{g}$ be a quadratic form on $\mathfrak{g}$. Then the feedback control law $v(L) =$

$-C^T(L)\phi_L$ stabilizes the system to $(x = 0, z = 0)$ and among all feedback laws that stabilize system trajectories to $(x = 0, z = 0)$, it minimizes the cost

$$\eta = \int_0^\infty \|u\|^2 + \|w\|^2 + < \nabla\phi, \nabla\phi >_G .$$

**Proof:** We first prove stability. Let $\phi$ be the Lyapunov function, then

$$\frac{d\phi}{dt} = - < \phi_L, C(L)C^T(L)\phi_L >$$

rewriting the above equation as

$$\frac{d\phi}{dt} = -\|C^T(L)\phi_L\|^2.$$

All we need to show is that $\phi_L \in \mathcal{N}(C^T(L))$ only if $L = 0$. Observe

$$C^T(L)\phi_L = (\phi_x - [\phi_z, x], -([x, \phi_x] + [\phi_z, z])) \tag{5.54}$$

Note if $\phi_L \in \mathcal{N}(C^T(L))$ then

$$\begin{aligned} \phi_x &= [\phi_z, x] && (5.55) \\ [x, \phi_x] + [\phi_z, z] &= 0 && (5.56) \\ & && (5.57) \end{aligned}$$

Observe from equation (5.55)-(5.56) it follows that

$$[x, [\phi_z, x]] + [\phi_z, z] = 0$$

implying that $[\phi_z, x] = 0$ (multiply both sides of equation by $\phi_z$ and take the trace and therefore $[\phi_z, z] = 0$. Therefore from equation (5.55), we obtain $\phi_x = 0$. Notice that $\phi_x = 2(x + [z, r])$, therefore if $\phi_x = 0$ then $x = -[z, r]$. Also

$$\phi_z = [x, r] + [[z, r], r] + Q.$$

where $ad_Q = ad_z ad_n^2 + ad_n^2 ad_z$. Substituting for $x = -[z, r]$, we get $\phi_z = Q$. Since $[\phi_z, z] = 0$ implies $[Q, z] = 0$. Thus $ad_{[Q,z]} = [ad_Q, ad_z] = 0$. Substituting for $Q$ we get

$$[ad_z^2, ad_n^2] = 0.$$

which implies $[z, n] = 0$. Note $[\phi_z, x] = [Q, [z, r]]$, which implies that

$$\begin{aligned} [z[n[n[z, r]]]] + [n[n[z[z, r]]]] &= 0 && (5.58) \\ [z[z[n[n, r]]]] &= 0. && (5.59) \end{aligned}$$

where (5.59) follows from equation (5.58)by repeated use of Jacobi identity. From the remark 11 and equation (5.59) it follows that $z = 0$, implying $p$ is zero. The proof for optimality directly follows from theorem (25) **Q.E.D.**

## 5.8 Conclusions

Our major contribution in this chapter was to derive feedback stabilization laws for a class of nonholonomic systems as solutions to variational problems. It was shown that by introducing dynamics in the symmetry group of the system, we can introduce additional controls in the system dynamics. We presented smooth feedback stabilization laws for systems more general than considered before. Similarities between our approach and gauge theories in physics, were pointed out. The future work in this area involves next extending these variational methods to systems which require higher brackets for controllability.

# Appendix A

# Riemannian Geometry of Lie Groups and Homogeneous Spaces

In this chapter, we recapitulate the basics of differential geometry and geometric control. We follow SMITH (1993) and BROCKETT (1979) in our exposition. We assume that the reader knows, or will find out elsewhere the definition of a manifold. vector field etc BOOTHBY (1976); KOBAYASHI AND NOMIZU (1969). Let $M$ be a $C^\infty$ differentiable manifold. Denote the set of $C^\infty$ functions on $M$ by $C^\infty(M)$, the tangent plane at $p$ in $M$ by $T_p$ or $T_pM$, and the set of $C^\infty$ vector fields on $M$ by $\mathfrak{X}(M)$.

**Riemannian structures**

**Definition 13** Let $M$ be a differentiable manifold. A Riemannian structure on $M$, is a tensor field $g$ of type $(0,2)$, which for all $X, Y \in \mathfrak{X}(M)$ and $p \in M$ satisfies $g(X,Y) = g(Y,X)$ and $g_p : T_p \times T_p \to R$ is positive definite. We use the notation $< X, Y >= g_p(X,Y)$ and $\|X\| = g_p(X,X)^{\frac{1}{2}}$, where $X, Y \in T_p$ is often used. Let $t \to \gamma(t)$, $t \in [a,b]$ be a curve segment in $M$. The length of $\gamma$ is defined by

$$L(\gamma) = \int_a^b g_{\gamma(t)}(\dot\gamma(t), \dot\gamma(t))^{\frac{1}{2}} dt.$$

A Riemannian manifold is a connected differentiable manifold with a Riemannian structure. Because $M$ is connected, any two points can be joined by a curve. The infimum of length of all curve segments joining $p$ and $q$ yields a metric on $M$, called the Riemannian metric denoted by $d(p,q)$.

**Affine connections**

Let $M$ be a differentiable manifold. An affine connection on $M$ is a function $\nabla$ which assigns to each vector $X \in \mathfrak{X}(M)$ an R-linear map $\nabla : \mathfrak{X}(M) \to \mathfrak{X}(M)$, which satisfies

$$\begin{aligned} \nabla_{fX+gY} &= f\nabla_X + g\nabla_Y \\ \nabla_X(fY) &= f\nabla_X(Y) + (Xf)Y, \end{aligned}$$

for all $f, g \in C^\infty(M)$, $X, Y \in \mathfrak{X}(M)$. The expression of these ideas using coordinates is very useful. Let $M$ be an $n$-dimensional differentiable manifold with affine connection $\nabla$, and $(U, x_1, \ldots, x_n)$ a coordinate chart on $M$. These coordinates induce a basis $\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_1}$ on $\chi(U)$. Then there exists $n^3$ functions $\gamma_{ij}^k$, $1 \leq i, j, k \leq n$, on $U$ such that $\partial_{\frac{\partial}{\partial x_i}} \frac{\partial}{\partial x_j} = \sum_{i,j,k} \gamma_{ij}^k \frac{\partial}{\partial x_k}$. The $\gamma_{ij}^k$ are called the Christoffel symbol of the connection.

**Geodesics and parallelism**

Let $M$ be a differentiable manifold with affine connection $\nabla$. Let $\gamma : I \to M$ be a smooth curve with tangent vectors $X(t) = \dot{\gamma}(t)$, where $I \in R$ is an open interval. The curve $\gamma$ is called a geodesic if $\nabla_X X = 0$, for all $t \in I$. Let $Y(t) \in T_{\gamma(t)}$, $(t \in I)$ be a smooth family of tangent vectors defined along $\gamma$. The family $Y(t)$ is said to be parallel along $\gamma$ if $\nabla_X Y = 0$ for all $t \in I$.

For every $p \in M$ and $X \neq 0$ in $T_p$, there exists a unique geodesic $t \to \gamma_X(t)$ such that $\gamma_X(0) = p$ and $\dot{\gamma}_X(0) = X$. We define the exponential map $exp_p : T_p \to M$ by $exp_p(X)\gamma_X(1)$ for all $X \in T_p$. Oftentimes the map $exp_p$ will be denoted by $exp(tX)$. A neighborhood $N_p$ of $p$ is a normal neighborhood if $N_p = expN_0$, where $N_0$ is a star-shaped neighborhood of the origin in $T_p$ and exp maps $N_0$ diffeomorphically onto $N_p$.

Let $(U, x_1, \ldots, x_n)$ be a coordinate chart on an n-dimensional differentiable manifold with affine connection $\nabla$. Geodesics in $U$ satisfy the $n$ second order nonlinear differential equations

$$\frac{d^2 x_k}{dt^2} + \sum_{i,j} \frac{dx_i}{dt}\frac{dx_j}{dt}\gamma_{ij}^k = 0.$$

**Riemannian connections**

Given a Riemannian structure $g$ on a differentiable manifold $M$, there exists a unique connection $\nabla$ on $M$, called the *Riemannian* or $Levi-Civita$ connection, which for $X, Y \in \mathfrak{X}(M)$ satisfies

$$\begin{aligned} \nabla_X Y - \nabla_Y X &= [X, Y] \\ \nabla g &= 0 \end{aligned}$$

Length minimizing curves on $M$ are geodesics of the Levi-Civita connection. We shall use this connection throughout this thesis. For every $p \in M$, there exists a normal neighborhood $N_p = expN_0$ of $p$ such that $d(p, exp_p X) = \|X\|$ for all $X \in N_0$, where $d$ is the Riemannian metric corresponding to $g$.

### A.0.1 Lie Groups and Homogeneous Spaces

We review the basic structure of Lie Groups and homogeneous spaces in this section.
Lie Groups

**Definition 14** A Lie Group $G$ is a differentiable manifold and a group such that the map $G \times G \to G$ defined by $(g,\ k) \to gk^{-1}$ is $C^{\infty}$.

**Definition 15** A Lie Algebra $\mathfrak{g}$ over $R$ is a vector space over $R$ with a bilinear operation $[.,.] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$(called the bracket)such that for all $x, y, z \in \mathfrak{g}$,

$$\begin{aligned} [x, x] &= 0 \\ [x,\ [y,\ z]] + [y,\ [z,\ x]] + [z,\ [x,\ y]] &= 0 \end{aligned}$$

Let $G$ be a Lie Group and $g \in G$. Left multiplication is denoted by the map $l_g : G \to G, k \to gk$, and similarly for right multiplication $r_g : k \to kg$. Let $X$ be a vector field on $G$. $X$ is said to be left invariant if for each $g \in G$,

$$l_g * (X) = X \circ l_g.$$

The notation $f_*$, is used here and elsewhere to denote $df$, the differential of a map $f$. Specifically, note that if $X$ is a left invariant vector field on $G$ then $X_g = l_{g*} X_e$, i.e, the value of $X$ at any point $g \in G$ is determined by its value at the identity $e$. Thus there is a one to one correspondence between the left invariant fields on $G$ and tangent vectors $T_e G$. We shall define $\mathfrak{g}$to be the vector space $T_e G$, and for $X \in \mathfrak{g}$, denote the corresponding left invariant vector field by $\tilde{X}$.

For every $X$ in $\mathfrak{g}$, there is a unique homomorphism $\phi : R \to G$, called the one parameter subgroup of G generated by $X$, such that $\dot{\phi}(0) = X$. Define the exponential map $exp : \mathfrak{g} \to G$, by setting $expX = \phi(1)$. The one-parameter subgroup $t \to \phi(t)$ generated by $X$ is denoted by $t \to exptX$. For matrix groups, this exponential map corresponds to matrix exponentials, i.e., $exptX = e^{Xt} = I + tX + (t^2/2!)X^2 + \ldots$.

Let $G$ be a Lie Group with Lie algebra $\mathfrak{g}$. Consider the action of $G$ on itself by conjugation, i.e., $a : (g, k) \to gkg^{-1}$. Denote the automorphism $k \to gkg^{-1}$ of $G$ by $a_g$. Define the adjoint representation $Ad : G \rightarrow Aut(\mathfrak{g})$ by the map $g \rightarrow (da_g)_e$, where $Aut(\mathfrak{g})$ is the group of automorphisms of a Lie algebra $\mathfrak{g}$. If $G$ is a matrix group with $g \in G$ and $\omega \in \mathfrak{g}$, we have $Ad_g(\omega) = g\omega g^{-1}$. Denote the differential of $Ad$ at the identity by $ad$, i.e.,

$$ad = dAd_e,$$

so that $ad : \mathfrak{g} \rightarrow End(\mathfrak{g})$, is a map from the Lie algebra $\mathfrak{g}$ to its vector space of endomorphisms $End(\mathfrak{g})$. It may be verified that $ad_X Y = [X, Y]$, for $X$ and $Y$ in $\mathfrak{g}$. The functions $Ad : G \rightarrow Aut(\mathfrak{g})$ and $ad : \mathfrak{g} \rightarrow End(\mathfrak{g})$ are related by

$$Ad \circ exp = exp \circ ad,$$

i.e., for $X \in \mathfrak{g}$, $Ad_{expX} = e^{ad_X}$.

**Definition 16** Let $\mathfrak{g}$ be a Lie algebra. The $killingform$ of $\mathfrak{g}$ is the bilinear form $\phi$ on $\mathfrak{g} \times \mathfrak{g}$ defined by

$$B(X, Y) = tr(ad_X \circ ad_Y).$$

**Definition 17** A lie algebra $\mathfrak{g}$ is called semisimple if $\mathfrak{g} \neq 0$ and has no Abelian ideals $\neq 0$.

**Theorem 28** *(Cartan)*A lie algebra $\mathfrak{g}$ is semisimple if and only if its $killing\ form$, $B(.,.)$ is non-degenerate.

**Homogeneous Spaces** Let $G$ be a Lie group and $H$ a closed subgroup of $G$. Then the (right) coset space $G/H = \{Hg : g \in G\}$, admits the structure of a differentiable manifold such that the natural projection $\pi : G \rightarrow G/H$, $g \rightarrow Hg$, and the action of $G$ on $G/H$ defined by $(Hk, g) \rightarrow Hkg$, are $C^\infty$. The dimension of $G/H$ is given by dim $G$ - dim $H$. Define the origin of $G/H$, by $o = \pi(e)$.

**Definition 18** Let $G$ be a Lie Group and $H$ a closed subgroup of $G$. The differentiable manifold $G/H$ is called a homogeneous space.

Let $\mathfrak{g}$ and $\mathfrak{h}$ be the Lie algebras of $G$ and $H$, respectively, and let $\mathfrak{m}$ be a vector subspace of $\mathfrak{g}$ such that $\mathfrak{g} = \mathfrak{m} + \mathfrak{h}$ is a direct sum. Then there exists a neighborhood of $0 \in \mathfrak{m}$, which is mapped homeomorhically onto a neighborhood of the origin $o \in G/H$ by the mapping $\pi \circ exp|_{\mathfrak{m}}$. The tangent plane $T_o(G/H)$ at the origin can be identified with the vector subspace $\mathfrak{m}$.

**Definition 19** Let $G$ be a connected Lie Group, $H$ a closed subgroup of $G$, and $\mathfrak{g}$ and $\mathfrak{h}$ be the Lie algebras of $G$ and $H$, respectively. The homogeneous space $G/H$ is said to be reductive if there exists a vector subspace $\mathfrak{m}$ of $\mathfrak{g}$ such that $\mathfrak{g} = \mathfrak{m} + \mathfrak{h}$ (direct sum), and $\mathfrak{m}$ is $Ad_H$-invariant, i.e., $Ad_H(\mathfrak{m}) \in \mathfrak{m}$.

For example the homogeneous space $G/H$, is reductive if $H$ is compact. We will solely be interested in examples where $H$ is compact and so all the homogeneous spaces, we will be working with will be reductive.

**Invariant affine connections**

**Definition 20** Let $G$ be a Lie transformation group acting on a differentiable manifold $M$. An affine connection $\nabla$ on $M$ is said to be $G-invariant$ if for all $g \in G$ , $X,\ Y, \in \mathfrak{X}(M)$,

$$l_{g^*}(\nabla_X Y) = \nabla_{(l_{g^*}X)}(l_{g^*}Y).$$

If $M = G$ is a Lie group we have the following classification. Let $\tilde{X}$ and $\tilde{Y}$ be left invariant vector fields on $G$ corresponding to $X$ and $Y \in \mathfrak{g}$, respectively. There is a one to one correspondence between invariant affine connections on $G$ and the set of bilinear function $\alpha : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ given by the formula

$$\alpha(X, Y) = (\nabla_{\tilde{X}}\tilde{Y})_e.$$

Geodesics on $G$ coincide with one-parameter subgroups if and only if $\alpha(X, X) = 0$ for all $X \in \mathfrak{g}$. The classical Cartan-Schouten invariant affine connections on $G$ correspond to $\alpha(X, Y) = 0$, $\alpha(X, Y) = \frac{1}{2}[X, Y]$ and $\alpha(X, Y) = [X, Y]$.

Let $G/H$ be a reductive homogeneous space with a fixed decomposition of the Lie algebra, $\mathfrak{g} = \mathfrak{m} + \mathfrak{h}$, $Ad_H(\mathfrak{m}) \subset \mathfrak{m}$, and $\pi : G \to G/H$ the natural projection. Any element $X \in \mathfrak{g}$ can then be uniquely decomposed into the sum of the elements in $\mathfrak{m}$ and $\mathfrak{h}$ respectively. There is one to one correspondence between affine connections on $G/H$ and set of bilinear functions $\alpha : \mathfrak{m} \times \mathfrak{m} \to \mathfrak{m}$ which are $Ad_H$ invariant.

Now we state a theorem, which characterizes the form of geodesics and parallel translation in homogeneous space $G/H$.

**Theorem 29** *(Nomizu)*On the reductive homogeneous space $G/H$, there exists a unique invariant connection which is torsion free and satisfies that the curve $t \to \gamma_X(t) \in G/H$, defined by $\gamma_X(t) = \pi(exptX)$, is a geodesic in $G/H$. It is defined by the function $\alpha(X, Y) = \frac{1}{2}[X, Y]_m$ on $\mathfrak{m} \times \mathfrak{m}$.

This connection is called the canonical torsion free connection on $G/H$. In case $G/H$ is a symmetric homogeneous space, that is $[m,m] \in h$, then we have $\alpha(X,Y) = 0$

**Invariant Riemannian metrics**

**Definition 21** A Riemannian structure $g$ on $G/H$ is said to be (right) invariant if

$$g_{p.k}(l_{k^*}X, l_{k^*}Y) = g_p(X,Y)$$

for all $p \in G/H$, $k \in G$ and $X, Y \in T_p$.

Let $G$ be a Lie group which admits a bi-invariant metric $<,>$. Then there is a corresponding left invariant metric, called the normal metric, on the homogeneous space $G/H$ with fixed decomposition $\mathfrak{g} = \mathfrak{m} + \mathfrak{h}$, arising from the restriction of $<,>$ to $\mathfrak{m}$. The Levi-Civita connection of this metric is the canonical torsion free connection on $G/H$. Hence geodesics under this connection are locally length minimizing under the normal metric.
We have given a brief overview of the geometrical structure we would need in this thesis. We now look at the formalism of geometric control and recapitulate some of the the examples from robotics where one makes use of such a formalism to make assertions about the controllability of the systems.

## A.1 Geometric Control

By a control system we will mean a dynamical system whose laws unlike the laws of classical physics are not completely specified but depend on choice of certain parameters, called controls, that can vary and by which one can control the behavior of the system . In this manuscript we will assume the space of all configurations of the system to be an $n$ dimensional manifold $M$ and the dynamics of the system described by vector fields on this manifold. The control parameters are assumed to take values in an arbitrary subset $U \in \mathbb{R}^m$ and the dynamics is described by a mapping

$$F : M \times U \to TM,$$

such that for each $u \in U$, $F_u : M \to TM$ defined by $F_u(x) = F(x,u)$ for $x$ in $M$ is a smooth vector field. The system then evolves according to the differential equation

$$\dot{x} = F(x,u).$$

The control functions can be of several types. A control $u$ is called a feedback control, or a closed loop control, if $u : M \to U$. When $u$ is a smooth map, and when $F$ is smooth, the corresponding vector field $x \to F(x, u(x))$ is a smooth vector field. Any integral curve of this vector field is called a closed-loop trajectory.
A control $u$ is called an open loop control if $u : \mathbb{R}^1 \to U$. Also a control can be a combination of both types, that is a mapping $u : M \times \mathbb{R} \to U$. The trajectories corresponding to this choice of control are the solutions of the time-varying differential system

$$\frac{dx}{dt} = F(x, u(x,t)).$$

If we choose a co-ordinate system for the manifold $M$, then the vector field $F$ can be written as

$$F = \sum_{i=1}^{n} f_i(x) \frac{\partial}{\partial x_i}.$$

and then in local coordinates the system evolution takes the form

$$\dot{x}_i = f_i(x); \quad i = 1, 2, \cdots n.$$

In this manuscript we will only be interested in a special class of control systems called the "Affine Control System". If $\{X_0, X_1, \cdots, X_n\}$ denote a family of smooth vector fields on the manifold $M$, then the control system we are interested in takes the form

$$\dot{x} = X_0(x) + \sum_{i=1}^{n} u_i X_i(x). \tag{A.1}$$

Given a system of the above form, the questions we will be interested in addressing in various specific scenarios take the following form.

- The existence of a measurable control function $u = (u_1, \ \ldots u_m) : \mathbb{R} \to U$, which can transfer the state from some given initial configuration $x(0)$ to some terminal configuration $x(T)$.
- The minimum time required to reach the terminal state and the existence and uniqueness of control laws which accomplish transfer in minimum time (time optimal control).
- The existence and uniqueness of control laws which minimize a given cost function.
- Existence and properties of feedback control $u : M \to \mathbb{R}$, which stabilizes the trajectories of the closed loop system $\dot{x} = F(x, u)$, to some given point $x_0 \in M$.

**Accessibility**

We begin by addressing the first question, the problem of controllability and accessibility for a nonlinear control system. The mathematical concept that lies at the heart of the matter is the concept of a lie bracket. Smooth vector fields act as derivations on the space of smooth functions. If $X$ denotes a vector field and $f$ a smooth function on $M$ , then $X(f)$ will denote the function $x \to X(x)(f)$. There is a very important algebraic construction associated with vector fields on a manifold, namely the construction of a lie bracket. If $F$ and $G$ represent smooth vector fields on the manifold $M$, then their lie bracket $[F,G]$ is defined by its action $[F,G](f) = G(F(f)) - F(G(f))$. In local coordinates, if $F = \sum_{i=1}^{n} f_i(x)\frac{\partial}{\partial x_i}$, and $G = \sum_{i=1}^{n} g_i(x)\frac{\partial}{\partial x_i}$ then the lie bracket $[F,G]$ takes the form

$$[F,G] = \sum_{i=1}^{n}\sum_{j=1}^{n}(g_j\frac{\partial f_i}{\partial x_j} - f_i\frac{\partial g_i}{\partial x_j})\frac{\partial}{\partial x_i}.$$

The lie bracket of vector fields satisfy the following properties

$$\begin{aligned} [F,\ G] &= -[G,\ F] \\ [F,\ [G,\ H]] + [G,\ [H,\ F]] + [H,\ [G,\ F]] &= 0 \end{aligned} \tag{A.2}$$

Let $\{f_i\}$, be a collection of smooth vector fields on the manifold $M$. Then the distribution $\Delta = span\{f_i\}$ assigns a subspace of the tangent space to each point in $M$ in a smooth way. At any point the distribution is a linear subspace of the tangent space

$$\Delta_x = span\{f_1(x), \ldots, f_m(x)\} \subset T_x(M).$$

A distribution is involutive if it is closed under the Lie bracket, that is iff $f,\ g \in \Delta,\ [f,g] \in \Delta$.

Consider a class of Affine control system, called the drift free or driftless control systems, which take the form

$$\dot{x} = \sum_{i=1}^{m} u_i f_i(x) \quad ,\ x \in M;\ u = (u_1, \ldots, u_m) \in U \in \mathbb{R}^m. \tag{A.3}$$

We assume that that $f_i$ are smooth, linearly independent vector fields on $M$. The accessibility properties of this system can be inferred by looking at the distribution $\Delta = span\{f_i\}$.

**Theorem 30** (Versions of Frobenius Theorem) Let $\Delta = span\{f_i\}$, be an involutive collection of vector fields which are

- Analytic on a analytic manifold $M$. Then given any point $x_0 \in M$, there exists a maximal submanifold $N$ containing $x_0$ such that $\Delta_x$, spans the tangent space of $N$ at each point $x \in N$.

- $C^\infty$ on a $C^\infty$ manifold $M$ with dim $\Delta_x$, being constant. Then given any point $x_0 \in M$, there exists a maximal submanifold $N$ containing $x_0$, such that $\Delta_x$, spans the tangent space of $N$ at each point $x \in N$.

We denote the manifold $N$ by $exp\{f_i\}x_0$. Thus for the control system (A.3), the distribution $\Delta = span\{f_i\}$ is involutive, then $N$ precisely characterizes the reachable set. However if $\Delta = span\{f_i\}$ is not involutive then the reachable set is bigger and is characterized by the following theorem of Chow.

Given, a collection of vector fields, $\{f_i\}$, we denote the smallest Lie algebra of vector fields which contain them by $\{f_i\}_{LA}$.

**Theorem 31** (Versions of Chow's Theorem) Let $\Delta = span\{f_i\}$, be a collection of vector fields on the manifold $M$, such that $\{f_i\}_{LA}$ is

- analytic on a analytic manifold $M$. Then given any point $x_0 \in M$, there exists a maximal submanifold $N \in M$ containing $x_0$ such that $\{exp\{f_i\}\}x_0 = \{exp\{f_i\}_{LA}\}x_0 = N$.

- $C^\infty$ on a $C^\infty$ manifold $M$ with $dim(span\{f_i\}_{LA})$ constant on $M$. Then given any point $x_0 \in M$, there exists a maximal submanifold $N$ containing $x_0$, such that $\{exp\{f_i\}\}x_0 = \{exp\{f_i\}_{LA}\}x_0 = N$

Chow's Theorem goes a long way, in answering the controllability question for driftless control systems,

$$\dot{x} = \sum_{i=1}^{m} u_i f_i(x) \quad , \; x \in M; \; u = (u_1, \ldots, u_m) \in U \in \mathbb{R}^m.$$

We assume that that the vector fields $f_i$ are smooth, linearly independent vector fields on $M$. We say that the system A.3 is controllable if for any $x_0$, $x_1 \in M$ there exists a $T > 0$ and a measurable $u : [0, T] \to U$, such that for the system A.3, $x(0) = x_0$ and $x(T) = x_1$. If $dim(\{f_i\}_{LA}(x)) = dim\ T_x(M)$, for all $x \in M$, then by chow's theorem the system is controllable and the points on the manifold $M$ can be reached by letting $u_i's$ take on the values zero and one.

The situation is more complex, if we have drift in the system. If we now consider the general affine control system, $\dot{x} = f(x) + \sum_{i=1}^{m} u_i g_i(x)$. Then we have

**Theorem 32** *(Krener, Lobry, Sussman-Jurdjevic)* Suppose $f$ and $\{g_i\}_{i=1}^m$ are vector fields on the manifold $M$, and suppose $\{f, g_i\}$ meets either of the conditions of the $Chow's$ theorem. Then the reachable set for $\dot{x} = f(x) + \sum_{i=1}^m u_i g_i(x)$, contains an open subset of the manifold $N = exp\{f, g_i\}_{LA} x_0$.

We consider a special case of this situation, when the drift generates periodic orbits. In this case more can be said about the reachable set.

**Theorem 33** Suppose $f$ and $\{g_i\}_{i=1}^m$ are vector fields on the manifold $M$, and suppose $\{f, g_i\}$ meets either of the conditions of the $Chow's$ theorem. Suppose for each $x_0$, the solution of $\dot{x} = f(x)$, is periodic with a least period $T(x_0) < M$, Then the reachable set for $\dot{x} = f(x) + \sum_{i=1}^m u_i g_i(x)$, is $\{exp\{f, g\}_{LA}\}_G x_0$.

We now consider a special case which will help us prove controllability in quantum systems. In case the state space of the control system $\dot{x} = f(x) + \sum_{i=1}^m u_i g_i(x)$ is a compact Lie Group $G$ and $\{f, g_i\}$ are right invariant vector fields on $G$ then we can make a strong assertion about controllabilityJURDJEVIC AND SUSSMANN (313-329).

**Theorem 34** *(Jurdjevic-Sussmann)* Given the right invariant control system

$$\dot{x} = f(x) + \sum_{i=1}^{m} u_i g_i(x).$$

on a compact connected Lie group $G$. If $\{f, g_i\}_{LA} = L(G)$, the lie algebra of right invariant vector fields on $G$. The reachable set $exp\{f, g_i\}_{LA} x_0 = G$.

**Nonholonomic Control Systems**

We now look up at some nonholonomic systems and the equations that model there dynamics. These examples which will reappear during the thesis. We present them here to illustrate the concepts of controllability we described above.

**Example 7** A simple example of a nonholonomic control systems is provided by a wheeled mobile robot called unicycle as shown in the figure A.1. The system consists of a platform on a wheel that can only roll on the ground. There are two control inputs to the system. Let $u$ denote the driving velocity and $v$ the steering velocity, and let $x, y, \phi$, denote the position of the center of mass of the robot and the angle the wheeled robot makes with the

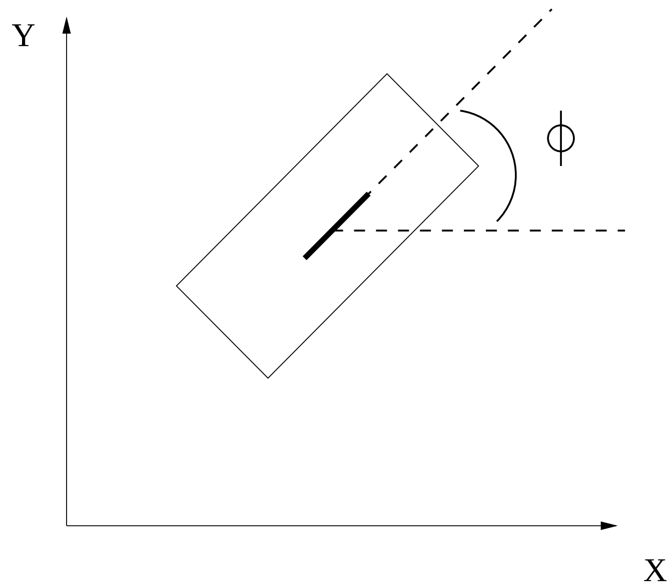


Figure A.1: The panel shows a unicycle the simplest form of a mobile robot.

$x$ axis. Then the kinematics of the system takes the form

$$\begin{aligned}\dot{x} &= u\cos(\phi)\\ \dot{y} &= u\sin(\phi)\\ \dot{\phi} &= v\end{aligned}$$

By suitable transformation of state and control variables, the equations can be transformed to the familiar nonholonomic integrator,

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = u_1\begin{bmatrix} 1 \\ 0 \\ -x_2 \end{bmatrix} + u_2\begin{bmatrix} 0 \\ 1 \\ x_1 \end{bmatrix} = u_1 g_1 + u_2 g_2$$

Observe $g_1 = \frac{\partial}{\partial x_1} - x_2\frac{\partial}{\partial x_3}$ and $g_2 = \frac{\partial}{\partial x_2} - x_1\frac{\partial}{\partial x_3}$, and $[g_1, g_2] = 2\frac{\partial}{\partial x_3}$, spans the tangent space at each point in $\mathbb{R}^3$. Hence the nonholonomic system is controllable.

Now consider a generalization of this system, called the **general position area system**. The system takes the following form

**Example 8** Let $x \in \mathbb{R}^n$, $u \in \mathbb{R}^n$ and $y \in so(n)$ the space of $n \times n$ skew symmetric matrices.

Then

$$\begin{aligned} \dot{x} &= u \\ \dot{y} &= xu^T - ux^T \end{aligned}$$

Observe the system can be written in a standard canonical form $\dot{X} = \sum_{i=1}^{n} u_i g_i(X)$, where $u = [u_1, \ldots, u_n]^T$. and $X \in \mathbb{R}^{\frac{n(n+1)}{2}}$. It is easy to see that span $\{g_i\}_{LA} = \mathbb{R}^{\frac{n(n+1)}{2}}$ and therefore the general position area system is controllable. The importance of this system is that it is the canonical form of first brackett controllable systems.

Till now we considered examples where the first brackets of the vector fields are enough to span the whole space, now we look at a nonholonomic systems system that needs higher order of brackets of the vector fields to generate the lie algebra, that spans the tangent space of the state space.

**Example 9** Consider the system shown in figure A.2. The system is useful for illustrating manipulation of spherical objects between two fingers. The system consists of a sphere between two parallel plates. The lower plate is fixed, and the upper plate is allowed to move in $X - Y$ plane with velocities $u_1$ in $X$ direction and $v_1$ in the $Y$ direction. However by just using two controls we can move the sphere anywhere in the plane with any given orientation, that is the system is controllable. By suitable change of coordinates BROCKETT AND L.DAI (1992), the state of the system can be written in terms of $(x, y, z, m, n) \in \mathbb{R}^5$

$$\begin{aligned} \dot{x} &= u \\ \dot{y} &= v \\ \dot{z} &= xv - yu \\ \dot{m} &= x^2 v \\ \dot{n} &= y^2 u \end{aligned}$$

Observe the system can be written in a standard canonical form

$$\frac{d}{dt}\begin{bmatrix} x \\ y \\ z \\ m \\ n \end{bmatrix} = u \begin{bmatrix} 1 \\ 0 \\ -y \\ 0 \\ y^2 \end{bmatrix} + v \begin{bmatrix} 0 \\ 1 \\ x \\ x^2 \\ 0 \end{bmatrix} = ug_1 + vg_2.$$

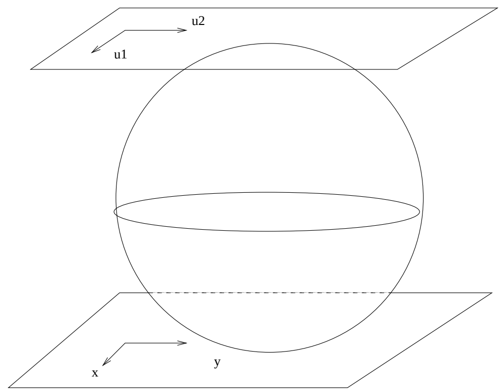


Figure A.2: The panel above, illustrates the ball and plate mechanism.

Observe $g_1 = \frac{\partial}{\partial x} - y\frac{\partial}{\partial z} + y^2\frac{\partial}{\partial n}$, $g_2 = \frac{\partial}{\partial y} + x\frac{\partial}{\partial z} + x^2\frac{\partial}{\partial m}$, $[g_1, g_2] = 2\frac{\partial}{\partial z} + 2x\frac{\partial}{\partial m} - 2y\frac{\partial}{\partial n}$, $[g_1, [g_1, g_2]] = 2\frac{\partial}{\partial m}$ and $[g_2, [g_1, g_2]] = -2\frac{\partial}{\partial n}$ spans the tangent space at each point in $\mathbb{R}^5$. Hence the above system is controllable.

Till now we only looked at examples of drift free system, lets take the example of systems with drift to see the consequence of theorems on controllability of systems with drift.

**Example 10** Let $\Theta \in SO(3)$, and let $\Omega_z = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $\Omega_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$ represent the generators of rotation around $x$ and $z$ axis. Let the control system

$$\dot{\Theta} = [\Omega_z + u\Omega_x]\Theta$$

be given. Since the group $SO(3)$ is compact, and $\{\Omega_x,\ \Omega_z\}_{LA} = \mathfrak{so}(3)$, the Lie algebra of the group $SO(3)$. The above system is controllable by theorem (34).

The above examples illustrate the usefulness of certain controllability theorems for nonlinear system stated above and provide motivation for adopting a differential geometric framework for studying such problems.

# Bibliography


A. Bloch, N. H. M., and M.Reyhanoglu (1992). Control and stabilization of nonholonomic dynamic systems. *IEEE Transactions on Automatic Control*, 37(11):1746–1757.

Astolfi, A. (1994). On the stabilization of nonholonomic systems. In *Proceedings of the 33rd IEEE Conference on Decision and Control*, pages 3481–3486.

Bloch, A., Drakunov, S., and Kinyon, M. (1998). Stabilization of nonholonomic systems using isospectral flows. *SIAM Journal of Control and Optimization.*

Bloch, A., and S.Drakunov (1996). Stabilization and tracking in the nonholonomic integrator via sliding modes. *Systems and Control Letters*, 29:91–99.

Bloch, A. M., and Rojo, A. G. (2000). Control of squeezed states. *Preprint.*

Bloch, F., Hansen, W., and M.Packard (1946). *Phys. Rev*, 69:127.

Boothby, W. (1976). *Introduction to Differential Geometry and Lie Groups.* Academic Press.

Brockett, R. (1979). *Nonlinear control theory.* ES-209 Class Notes.

Brockett, R., and Khaneja, N. (1999). *On the stochastic control of quantum ensembles.*

Brockett, R. W. (1981). *New Directions in Applied Mathematics*, chapter Control Theory and Riemannian Geometry, pages 11–27. Springer-Verlag New York Inc.

Brockett, R. W. (1983). *Differential Geometric Control Theory*, chapter Asymptotic Stability and Feedback Stabilization, pages 181–191. Burkhauser, Boston.

Brockett, R. W. (1991). Dynamical systems that sort lists, diagonalize matrics, and solve linear programming problems. *Linear algebra and its applications*, 146:79–91.

Brockett, R. W. (1996). Characteristic phenomena and model problems in nonlinear control. In *Proceedings of the 1996 IFAC Congress*, volume G, pages 135–140.

Brockett, R. W. (2000). Sub-riemannian gradients. In *Preprint.*

Brockett, R. W., and L.Dai (1992). *Nonholonomic Motion Planning*, chapter Nonholonomic kinematics and the role of elliptic functions in constructive controllability, pages 1–22. Kluwer.

Butkovskiy, A. G., and Samoilenko, Y. (1990). *Control of Quantum Mechanical Processes and Systems.* Kluwer Academic , Dordrecht.

Carr, H., and Purcell, E. (1954). *Phys. Rev.*, 94:630.

CAVANAGH, J., FAIRBROTHER, W. J., PALMER, A. G., and SKELTON, N. J. (1996). *Protein NMR Spectroscopy, Principles and Practice.* Academic Press.

C.COHEN-TANNOUDJI (1992). *Laser cooling and trapping of neutral atoms: theory*, 219:153–164.

CORON, J.-M. (1991). Global asymptotic stabilization for controllable systems without drift. *Mathematics of Control,Signals and Systems*, 5:295–312.

CORON, J.-M., and POMET, J.-B. (1992). A remark on the design of time-varying stabilizing feedback laws for controllable systems without drift. In *Proceedings of IFAC Nonlinear Control Systems Design Symposium*, pages 413–417.

CORY, D., FAHMY, A., and HAVEL, T. (1997). Ensemble quantum computing by nmr spectroscopy. *Proc. Natl. Acad,. Sci.*, 94:1634–1639.

ENOS, M. E. (1993). *Dynamics and Control of Mechanical Systems: The Falling Cat and Related Problems.* Americal Mathematical Society.

ERNST, R., BODENHAUSEN, G., and WOKAUN, A. (1987). *Principles of Nuclear Magnetic Resonance in One and Two Dimensions.* Oxford University Press.

GLASER, S., and QUANT, J. (1996). *Advances in Magnetic and Optical Resonance*, volume 19, pages 59–252. Academic Press.

GLASER, S., T.SCHULTE-HERBRÜGGEN, SIEVEKING, M., SCHEDLETZKY, O., NIELSEN, N., SØRENSEN, O., and GRIESINGER, C. (1998). Unitary control in quantum ensembles, maximizing signal intensity in coherent spectroscopy. *Science*, (280):421–424.

GOLDMAN, M. (1988). *Quantum Description of High-Resolution NMR in Liquids.* Oxford University Press.

GULDNER, J., and UTKIN, V. (1994). Stabilization of nonholonomic mobile robots using lyapunov functions for navigation and sliding mode control. In *Proceedings of the 33rd IEEE Conference on Decision and Control*, pages 2967–2972.

HAHN, E. (1950a). *Phys. Rev.*, 80:580.

HAHN, H. (1950b). *Phys. Rev*, 80:580.

HERBRÜGGEN, T. S. (1998). *Aspects and Prospects of High Resolution Nuclear Magnetic Resonance.* PhD thesis, ETH Zurich.

HUANG, G., TARN, T., and J.CLARK (1983). *J. Math. Phys.*, 24:2608.

I. KOLMANOVSKY, M. R., and MCCLAMROCH, N. (1994). Discontinuous feedback stabilization of nonholonomic systems in extended power form. In *Proceedings of 33rd IEEE Conference on Decision and Control.*

JURDJEVIC, V., and SUSSMANN, H. J. (313-329). Control systems on lie groups. *Journal of Differential Equations*, 12:663–670.

KANE, T., and SHUR, M. (1969). A dynamical explanation of the falling cat phenomenon. *International Journal of Solid Structures*, 5:663–670.

KOBAYASHI, S., and NOMIZU, K. (1969). *Foundations of Differential Geometry Vol.1 and 2.* Interscience Publishers.

KOLMANOVSKY, I., and MCCLAMROCH, N. H. (1995). Development in nonholonomic control problems. *IEEE Control Systems*, pages 20–36.

KRASNOSELSKII, M., and ZABREIKO, P. (1983). *Geometric Methods in Nonlinear Analysis.* Springer-Verlag, Berlin.

LAFFERRIERE, G. A., and SONTAG, E. D. (1991). Remarks on control lyapunov functions for discontinuous stabilizing feedback. In *Proceedings of the 32nd Conference on Decision and Control*, pages 2398–2403.

LLOYD, S. (1996). Universal quantum simulators. *Science*, 273(1073-1078).

LLOYD, S., and SLOTINE, J. J. E. (1998). Quantum feedback with weak measurements. *Preprint.*

MONTGOMERY, R. (1990). The isoholonomic problem and some applications. *Communications in Mathematical Physics*, 128(1):565–592.

MORGANSEN, K. A., and BROCKETT, R. W. (1999). Nonholonomic control based on approximate inversion. In *Proceedings of American Control Conference (To be published).*

MURRAY, R. M., and SASTRY, S. (1993). Nonholonomic motion planning: steering using sinusoids. *IEEE Trans. Aut. Control*, 38(5):700–716.

POMET, J.-B. (1992). Explicit design of time-varying stabilizing control laws for a class of controllable systems without drift. *Systems and Control Letters*, 18:147–158.

PURCELL, E., TORREY, H., and POUND, R. (1946). *Phys. Rev.*, 69:37.

RABITZ, H. (1993). *Nature*, 366:304.

RAMAKRISHNA, V., SALAPAKA, M., M.DAHLEH, H.RABITZ, and A.PIERCE (1995). *Physical Review A*, 51:960.

S. CHU, J.E. BJORKHOLM, A. C., and ASHKIN, A. (1985). *Phys. Rev. Lett.*, 55:48–51.

SAMSON, C., and AIT-ABDERRAHIM, K. (1991). Feedback stabilization of a nonholonomic wheeled mobile robots. In *Proceedings of International Conference on Intelligent Robots and Systems.*

SCHWEIGER, A. (1990). *Modern Pulsed and Coninuous Wave Electron Spin Resonance*, pages 43–118. Wiley, London.

SETO, D., and BROCKETT, R. (1999). Open-loop control using periodic excitations. *Preprint.*

SHAPERE, A., and WILCZEK, F. (1987). Self-propulsion at low reynolds number. *Physical Review Letters*, 58:2051–2054.

SMITH, S. T. (1993). *Geometric Optimization Methods for Adaptive Filtering.* PhD thesis, Harvard University.

SØRENSEN, O., EICH, G., LEVITT, M., BODENHAUSEN, G., and ERNST, R. (1983). *Prog. NMR Spectroscopy*, volume 16, page 163.

STOUSTRUP, J., O.SCHEDLETSKY, GLASER, S., GRIESINGER, C., NIELSEN, N., and SØRENSEN, O. (1995). Generalized bound on quantum dynamics: Efficiency of unitary transformations between non hermitian states. *Physical Review Letters*, (74):2921–2924.

SUSSMANN, H. (1979). Subanalytic sets and feedback control. *Journal of Differential Equations*, 31:31–52.

TAUBES, G. (1997). *Science*, 275:307.

VON NEUMANN, J. (1937). Some matrix-inequalities and metrization of metric spaces. *Tomsk Univ. Rev.*, 1:286–300.

WARREN, W. S., RABITZ, H., and DAHLEH, M. (1993). Coherent control of quantum dynamics: The dream is alive. *Science*, 259:1581–1589.

W.D. PHILLIPS, J. P., and METCALF, H. (1985). *Laser cooling and trapping of neutral atoms: theory*, J. Opt. Soc. Am:1751–1767.

WEITEKAMP, D., GARBOW, J., and A.PINES (1982). *J. Chem. Phys.*, 77:2870.

WILSON, F. (1967). The structure of the level surfaces of a lyapunov function. *J. of Diff. Eqs*, 4: 323–329.

WISEMAN, H., and MILBURN, G. (1993). *Phys. Rev. Lett.*, 70:548.

WOLF, J. (1984). *Spaces of Constant Curvature*. Publish or Perish.

ZUBOV, V. (1957). *The methods of Liapunov and their Applicatons*. Leningrad.