Optimal Control of Coupled Spin Dynamics in the Presence of Relaxation

A thesis presented

by

Dionisis Stefanatos

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 Physics

Harvard University
Cambridge, Massachusetts

June 2005
Abstract

In this thesis, we study methods for optimal manipulation of coupled spin dynamics in the presence of relaxation. We use these methods to compute analytical upper bounds for the efficiency of coherence and polarization transfer between coupled nuclear spins in multidimensional nuclear magnetic resonance (NMR) experiments, under the presence of relaxation. We derive relaxation optimized pulse sequences which achieve or approach these bounds and maximize the sensitivity of the experiments. The two systems examined in detail in this thesis, arising in the context of protein NMR spectroscopy, are a coupled heteronuclear spin pair and an Ising spin chain. New schemes for coherence and polarization transfer are found, which yield substantial gains in transfer efficiency, compared to the traditionally used pulse sequences. From the standpoint of mathematical control theory, the problems arising in connection with optimal manipulation of dissipative quantum dynamics, motivate the study of a class of dissipative bilinear control systems and the computation of their reachable sets. For these systems we show that the optimal solution, corresponding to the maximization of transfer efficiency, and the reachable set can be found by solving a semidefinite program.
To my family.
Acknowledgements

First, I would like to express my sincere thanks to my thesis advisor, Prof. Navin Khaneja. His passion for research, broad scientific interests and insistence for rigorous approaches, have guided and inspired me through my four-year graduate studies at Harvard. I am also grateful to Prof. Steffen Glaser from Technical University of Munich, for his valuable collaboration during these years, and to Profs. Kyriakos Hizanidis and Petros Maragos from National Technical University of Athens, because they helped me ‘open’ the gates of Harvard University. I would also like to express my appreciation towards the members of my doctoral committee, Prof. Roger Brockett, Prof. Lene Hau, and Prof. Federico Capasso, for their evaluation of this work.

I must also acknowledge the help and support of fellow students and friends. My officemates, Jr-Shin Li, Haidong Yuan, Andrew Johnson, Brent Pryor, Dan Iancu, and the rest members of Harvard Robotics Lab, Mohamed Ali Belabbas, Benjamin Peirce, Jason Adaska and Michael McElroy.

Finally, I would like to thank my family, whose love and support have accompanied me throughout my life. My father Gerasimos, my mother Argyro and my sister Elpida. This thesis is dedicated to you.
## Contents

1 Introduction 1

2 Fundamentals of Relaxation 9
   2.1 The Master Equation 10
   2.2 Spectral Density Functions 17
   2.3 Basic Relaxation Mechanisms in Liquid State NMR 25
      2.3.1 Dipolar Relaxation 25
      2.3.2 Chemical Shift Anisotropy 28
      2.3.3 Interference Effects 30

3 Optimal Control of a Coupled Spin Pair in the Presence of Relaxation 32
   3.1 Description of the System and Statement of the Problem 32
   3.2 Formulation of the Problem in Terms of Optimal Control 36
   3.3 Solution of the Optimal Control Problem 40
   3.4 Calculation of the Maximum Transfer Efficiency 46
   3.5 Calculation of Optimal Controls and of the Corresponding RF Field 48
   3.6 Summary of the Results and Discussion 50

4 Semidefinite Programming and Reachable Sets of Dissipative Bilinear Control Systems 59
   4.1 Statement of the Problem 59
   4.2 Reduction to a Semidefinite Program 62
   4.3 Remarks on the Rank of the Semidefinite Program Solutions 67
   4.4 Examples 70

5 Relaxation Optimized Polarization and Coherence Transfer in Ising Spin Chains 79
   5.1 Description of the System and Statement of the Problem 79
   5.2 Formulation of the Problem in Terms of Optimal Control 83
   5.3 Conventional Approach-Comments-Upper Bound for the Efficiency 85
   5.4 Numerical Calculation of the Optimal RF Field 87
   5.5 The General Case with Arbitrary Coupling Constants 92

6 Conclusions and Future Work 97
List of Figures

1.1 Basic features of an NMR experiment. (a) The static magnetic field $B_0$ is used to polarize the sample. The rf field $B_{rf}(t)$ is used to steer the bulk magnetization and generate the free induction decay (FID) signal. (b) FID profile. (c) The Fourier transform of FID gives a peak at the frequency $\omega_0$, characteristic of the nucleus. Relaxation determines the linewidth of the spectrum. ................................................................. 3

1.2 (a) Proton spectrum of a small peptide. The proton resonances are clearly resolved. (b) Proton spectrum of a protein. The spectrum has overlapping resonance lines. ................................................................. 5

1.3 (a) Basic steps of a 2D NMR experiment. Spins $I$ evolve for time $t_1$, while spins $S$ for time $t_2$. The NMR signal is recorded during the second time interval. (b) The Fourier transform of the recorded signal gives a peak at the frequencies $(\omega_I, \omega_S)$. ................................................................. 6

2.1 Decomposition of the total motion to overall tumbling and internal motions. The orientation of the interaction vector $OZ$ can be determined either by the angle $\Omega_{LF}$ in the laboratory frame (LF), or by the angle $\Omega_{MF}$ in the molecular fixed frame (MF) and the Euler angles, denoted collectively by $\omega$, that specify the orientation of MF with respect to LF. ......................... 19

2.2 Spectral density function for an isotropic rotor. Calculation was performed using (2.76) with $\tau_c = 10$ ns. ................................................................. 23

2.3 Polar angles of the interaction direction $OZ$ in the laboratory frame $Oxyz$. ................................................................. 26

3.1 Initial state $I_z$ (A) and final state $2I_zS_z$ (B), written as sums of single-transition operators, defined in the text. The transfer $I_z \rightarrow 2I_zS_z$ can be accomplished by selectively inverting $I_zS^\beta$ to $-I_zS^\beta$. The polarizations of the spins in the ellipses show schematically which energy level is excess populated for each single transition operator, placed under the corresponding ellipse. For example, for $I_zS^\alpha$ is $\alpha\alpha$ (both $I$ and $S$ are up), while for $-I_zS^\beta$ is $\beta\beta$, see also the figure bellow. ................................................................. 35

3.2 Energy level diagram for a weakly coupled spin 1/2 pair. The selective population inversion of the levels $\alpha\beta$ and $\beta\beta$ corresponds to the transfer of polarization $I_z$ (A) to longitudinal two-spin order $2I_zS_z$ (B). The dark circles represent excess population in the corresponding levels. ................................................................. 35
3.3 Representation of the system variables $r_1, r_2$ and the angles $\beta_1, \beta_2$ in terms of the expectation values $\langle I_x \rangle, \langle I_z \rangle, \langle 2I_p S_z \rangle, \langle 2I_z S_z \rangle$. The new control parameters are $u_1 = \cos \beta_1, u_2 = \cos \beta_2$. ............................................... 39

4.2 Construction of the reachable set of point I. .............................................. 66

4.1 (a) Trajectory in $p$-space following the control law presented in the text, for $r = 2$ and $N = 4$. It approximates the straight line from the initial point $I$ to the final point $F$. Note that the restrictions $p_i(t') \geq 0$ may not be satisfied when $N$ is small. (b) For large $N$ the trajectory coincides with the line $IF$, so the restrictions $p_i(t') \geq 0$ are satisfied. ............................................... 64

4.2 Construction of the reachable set of point I. .............................................. 66

3.4 Figure depicts the angles $\vartheta_1, \vartheta_2$ and $\varphi_1, \varphi_2$, the optimal trajectory $(r_1(t), r_2(t))$ and the trajectory for the adjoint variables $(\lambda_1(t), \lambda_2(t))$. ............................................... 45

3.5 Optimal pulse sequence when $T_A < T \leq T_B$ for (a) $k_1 < k_2$ and (b) $k_1 > k_2$. For case (b) we just interchanged the values of $k_1, k_2$ from case (a), keeping the same $k$. Observe the symmetry in the optimal controls. .......................... 52

3.6 Optimal pulse sequence when $T > T_B$ for (a) $k_1 < k_2$ and (b) $k_1 > k_2$. Again, for case (b) we just interchanged the values of $k_1, k_2$. The symmetry in the controls appears again. Note that the duration $T$ has been set equal to the optimal duration $T_{opt}$, which maximizes the optimal transfer efficiency $\eta_T$. For the values $k = J, k_1 = 0.05 J, k_2 = 0.25 J$ that we used in (a), it is $T_{opt} \approx 0.468 J^{-1}$. For case (b), $T_{opt}$ is the same. .......................... 52

3.7 Optimal pulse sequence for $k_1 = k_2$. Observe the symmetry in the controls. Specifically, the relation $u_1(t) = u_2(T - t)$ holds here. For the values $k = J, k_1 = k_2 = 0.20 J$ that we use, the optimal duration is $T_{opt} \approx 0.353 J^{-1}$. ... 54

3.8 Optimal transfer efficiency $\eta_T$ as a function of the total transfer time $T$ for $k = J$ and various values of $k_1, k_2$ (normalized with respect to $J$). Observe that for $k_1, k_2 \neq 0$ there is an optimal transfer time $T_{opt}$. .......................... 56

3.9 (b) Maximum transfer efficiency $\eta_T$ evaluated at $T_{opt}$ for each point of the line in $(k_1, k_2, k)$ space shown in (a). The line is parametrized by the parameter $0 \leq s \leq 1$ and has been chosen such that the increase of $s$ from $0$ to $1$ simulates the transition from slow molecular motion, where $k \gg k_1 \simeq k_2$, to rapid molecular motion, where $k \simeq k_1 \sim k_2$. Note the superiority of the relaxation optimized pulse element (ROPE) compared to the INEPT pulse sequence for the case $k \gg k_1 \simeq k_2$. .......................... 56

3.10 (a) The rf pulse sequence corresponding to the controls $u_1, u_2$ of Fig. 3a. The pulse sequence starts with a hard 13.30° pulse around $y$ axis, which establishes $u_1(0) = 0.229$, followed by phase I as shown above. During phase II, no rf pulses are applied. Approaching phase II the rf amplitude becomes large for a very short time period. This can experimentally be very well approximated by a hard pulse of small flip angle. Following phase III as shown above, we apply a final hard 35.20° pulse around the $x$ axis, which completes the transfer. (b) Time evolution of the expectation values of operators participating in the transfer, under the optimal pulse sequence shown in (a). .......................... 57

ix
5.1 The system that we study in this chapter is a linear chain of \( n \) weakly interacting spins \( 1/2 \) with Ising coupling between next neighbors. This system originates from protein NMR spectroscopy. Note that besides the coherent evolution, described by the Ising couplings, there is also dissipative evolution caused by the interaction of the system with its environment.

5.2 Steps for the polarization transfer \( I_{1z} \rightarrow I_{nz} \). From our previous work in chapter 3 we know the optimal way to do steps 1 and \( n \), in presence of the relaxation mechanisms mentioned in the text. In this chapter we investigate the intermediate steps. If the spins have well separated resonance frequencies, then we can selectively excite them. In this case we can examine separately each of the intermediate transfers. But the equations that describe these transfers are similar, thus we just need to study only one of them, for example the transfer \( 2I_{1z}I_{2z} \rightarrow 2I_{2z}I_{3z} \).

5.3 Energy level diagram for a weakly coupled three-spin \( 1/2 \) chain. The dark circles indicate excess population in the corresponding levels. The selective population inversion of the levels \( \alpha\alpha\beta, \alpha\beta\beta \) and \( \beta\beta\alpha, \beta\alpha\alpha \) corresponds to the transfer \( 2I_{1z}I_{2z} \rightarrow 2I_{2z}I_{3z} \) of longitudinal two-spin order.

5.4 Optimal pulse (dashed line) calculated using a numerical optimization method based on a steepest descent algorithm, for various values of the normalized relaxation parameter \( \xi \). The Gaussian pulse (solid line) approximates very well the optimal pulse shape and gives a similar efficiency. This suggests that instead of using the initial numerical optimization method, we can use Gaussian pulses of the form (5.27), optimized with respect to \( A \) and \( \sigma \) for each value of \( \xi \).

5.5 Efficiency for the conventional method (INEPT) and for our method (SPORTS ROPE), for the values of \( \xi \) shown in Table 5.1. We also plot the upper bound (5.26) for the efficiency.
5.6 (a) Time evolution of the transfer functions (operators) that participate in the transfer $z_1 \rightarrow z_3$, for $\xi = 1$ and the optimal Gaussian pulse shown in Fig. 5.4(c). (b) The angle $\theta_3 = \tan^{-1} (z_3/x_3)$ of the vector $r_3 = x_3\hat{x} + z_3\hat{z}$ with the $x$ axis, as a function of time. Observe that initially $r_3$ is parallel to $x$ axis ($\theta = 0$), but under the action of the Gaussian pulse is rotated gradually to $z$ axis ($\theta = \pi/2$).

5.7 (a) The Gaussian pulse that approximates the optimal pulse shape for $J_1 = k$, $J_2 = 0.5k$ ($\zeta_1 = 1$, $\zeta_2 = 0.5$). (b) The corresponding time evolution of the transfer functions. Note that $y_3$ and $x_4$ have a similar time evolution.

5.8 (a) The Mexican hat pulse that approximates the optimal pulse shape for $J_1 = 2k$, $J_2 = 0.5k$ ($\zeta_1 = 2$, $\zeta_2 = 0.5$). (b) Time evolution of the corresponding transfer functions. Note that the dashed line corresponds to $y_3$ while the dotted line to $x_4$. 

91

95
## List of Tables

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Modified Second-Order Spherical Harmonics</td>
<td>27</td>
</tr>
<tr>
<td>2.2</td>
<td>Relaxation Operators for Dipolar Interaction</td>
<td>28</td>
</tr>
<tr>
<td>2.3</td>
<td>Relaxation Operators for Chemical Shift Anisotropy</td>
<td>30</td>
</tr>
<tr>
<td>5.1</td>
<td>For various values of $\xi \in [0, 1]$, the optimal values of $A, \sigma$ and the corresponding efficiency are shown. We present also for comparison the efficiency achieved by the steepest descent method.</td>
<td>90</td>
</tr>
<tr>
<td>5.2</td>
<td>For $\zeta_2 = 0.5$ fixed we have calculated the upper bound $v$ for increasing values of $\zeta_1$. Observe that after $\zeta_1 = 1.5$, $v$ becomes insensitive in the further increase of $\zeta_1$.</td>
<td>94</td>
</tr>
<tr>
<td>5.3</td>
<td>For the two cases with $J_1 \neq J_2$ examined in the text, we show the optimal values of $A, \sigma, b$, the efficiency achieved by the conventional method (IN-EPT), our method (SPORTS ROPE), and the upper bound calculated using semidefinite programming.</td>
<td>95</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

The idea of controlling quantum systems can be traced back to the early days of quantum mechanics. For example, the possibility of using electromagnetic radiation to control the motion of neutral atoms was realized soon and even observed experimentally, Frisch (1933). In the subsequent years, the advances in experimental means and the emergence of important applications led to the growth of the field among physics-chemistry communities, but in a rather qualitative and intuitive way. The theoretical foundations of quantum control, using principles and ideas from control theory and optimization, were put in the 1980’s, Belavkin (1983), Huang et al. (1983), Pierce et al. (1988). Nowadays, quantum control is a rapidly expanding field with many applications, ranging from coherent spectroscopy and quantum information processing to coherent control of molecular dynamics and chemical reactions.

In most applications involving control and manipulation of quantum phenomena, the system of interest is not isolated but interacts with its environment. This leads to the phenomenon of relaxation, which in practice results in signal loss and ultimately limits the range of applications. Manipulating quantum systems in a manner that minimizes relaxation losses poses a fundamental challenge of utmost practical importance. A premier example is the control of coupled spin dynamics in nuclear magnetic resonance (NMR) spectroscopy, Ernst et al. (1987), Slichter (1990). In structural biology, high-resolution NMR spectroscopy is an important technique that allows one to study the structure and dynamics of biological macromolecules, such as proteins, in aqueous solution, Wuthrich (1986), Cavanagh et al. (1995). With increasing size of molecules or molecular complexes, the rotational tumbling of the molecules becomes slower and leads to increased relaxation losses. When these relaxation rates become comparable to spin-spin couplings, the sensitiv-
The goal of this thesis is to develop methods for relaxation optimized manipulation of spin systems that are frequently encountered in NMR experiments with large biomolecules. For the sake of completeness and with a view of understanding optimal control problems arising in protein NMR spectroscopy because of the phenomenon of relaxation, we describe briefly the basic principles of nuclear magnetic resonance.

It is a known fact that, among the various atomic nuclei, about a hundred isotopes possess an intrinsic angular momentum, called spin and written as \( hI \) (\( h \) is Planck’s constant divided by \( 2\pi \)). They also possess a magnetic moment \( \mu \) which is proportional to their angular momentum

\[
\mu = \gamma hI .
\] (1.1)

The coefficient \( \gamma \), characteristic of each nuclear species, is called its gyromagnetic ratio. Modern NMR experiments use a large static magnetic field \( B_0 \) (pointing, say, in the \( z \) direction) of the order of 5-20 Tesla to align the magnetic moments of atomic nuclei in a sample. The energy of a magnetic dipole \( \mu \) in a magnetic field \( B \) is, Jackson (1998),

\[
E = -\mu \cdot B .
\] (1.2)

Thus, the magnetic moments of the spins which were randomly oriented initially have the tendency to align themselves along the magnetic field \( B_0 \), as this is the lowest energy configuration. The resulting bulk magnetization \( M \) in the direction of \( B_0 \) is then controlled by an oscillating radiofrequency (rf) magnetic field \( B_{rf}(t) = (B_x(t), B_y(t)) \) in the \( xy \)-plane, whose magnitude is smaller than \( B_0 \) by 4 to 5 orders. This field exerts a torque \( N \) on \( M \), Jackson (1998),

\[
N = M \times B .
\] (1.3)

Because of Eq. (1.1), the bulk magnetization is proportional to the bulk angular momentum \( J \) with proportionality constant the gyromagnetic ratio \( \gamma \)

\[
M = \gamma J .
\] (1.4)

From Newton’s second law it is

\[
\frac{dJ}{dt} = N ,
\] (1.5)

thus, using additionally Eqs. (1.3) and (1.4), we find that \( M \) evolves as

\[
\frac{dM}{dt} = \gamma M \times B .
\] (1.6)
In component form the above vector equation is

\[
\frac{d}{dt} \begin{bmatrix} M_x \\ M_y \\ M_z \end{bmatrix} = \gamma \begin{bmatrix} 0 & -B_0 & B_y(t) \\ B_0 & 0 & -B_z(t) \\ -B_y(t) & B_x(t) & 0 \end{bmatrix} \begin{bmatrix} M_x \\ M_y \\ M_z \end{bmatrix}.
\]  

(1.7)

An oscillating \( \mathbf{B}_{rf}(t) = (B_x(t), B_y(t)) = (A \cos \omega_0 t, A \sin \omega_0 t) \) at the Larmor frequency \( \omega_0 = \gamma B_0 \), transfers the magnetization vector \( \mathbf{M}(0) = (0, 0, 1) \) to the \( xy \)-plane, see Fig. 1.1(a), in \( \pi/(2\gamma A) \) units of time. This can be easily seen by transforming the above equations in a frame rotating with angular velocity \( \omega_0 \), Cavanagh et al. (1995). Another way to see why the oscillating magnetic field at frequency \( \omega_0 \) rotates \( \mathbf{M} \) towards the \( xy \)-plane is by using elementary quantum mechanics. Consider the case where the atomic nuclei in the sample have spin \( I = 1/2 \). The energy difference between the states spin (magnetic mo-
ment) up and spin (magnetic moment) down is \(2 \cdot \gamma \hbar / 2 \cdot B_0 = h \gamma B_0 = h \omega_0\), see Eqs. (1.1), (1.2). By shining electromagnetic radiation at the resonant frequency \(\omega_0\) of the transition, some of the spins in the sample flip and this shows up as a rotation of the bulk magnetization to the \(xy\)-plane. Alternatively, the system absorbs energy and from the formula \(E = -M \cdot B\) we see that the angle between \(M\) and \(B\) should increase from its initial value 0. When \(M\) reaches the \(xy\)-plane, the oscillating field is switched off and, according to Eq. (1.7), the magnetization precesses around the static magnetic field \(B_0\) with a frequency \(\omega_0\). This precessing magnetization induces a current, also termed as free induction decay (FID), in a nearby coil. Overtime, the magnetic moments making the magnetization vector \(M\) experience local fluctuations in the ambient field \(B_0\), causing them to dephase out. The resulting phenomenon is termed relaxation, is a characteristic property of open quantum systems, as we mentioned above, and is discussed in detail in the next chapter. This gives to the FID signal a decaying envelope, see Fig. 1.1(b). When this signal is Fourier transformed, shows a peak at \(\omega_0\), a characteristic of the nuclei since \(\omega_0 = \gamma B_0\). The decaying envelope gives the peak a linewidth which is determined by the decay rate, see Fig. 1.1(c). For the commonly used values of \(B_0\), the Larmor frequencies are in the radiofrequency (rf) range. At a field of 14 Tesla, the Larmor frequencies of proton (\(^1H\)), nitrogen (\(^{15}N\)) and carbon (\(^{13}C\)) are approximately 600 Mhz, 60 Mhz and 150 Mhz, respectively. NMR is therefore an important analytical tool in chemistry, as the peaks of the spectra reveal the chemical composition of molecules.

The experiments in structure determination of biological macromolecules using NMR, Wuthrich (1986), Cavanagh et al. (1995), begin with collecting a proton spectra as depicted in Fig. 1.2(a). Protons in the molecule have different electronic environments. These electronic currents generate a local magnetic field which alters the static field \(B_0\) to \(B_0(1 - \sigma)\), hence altering the Larmor frequency of these protons. The resulting proton spectrum has many peaks, as depicted in Fig. 1.2(a). These shifts in the Larmor frequency are characteristic of the chemical environment of the spins and are called chemical shifts. Based on these shifts, it is possible to assign the various peaks in the spectrum of a molecule to protons in different amino-acids, a process called frequency labelling, Wuthrich (1986). In NMR spectroscopy of proteins, one starts with a knowledge of composition of the protein i.e. the sequence of amino-acids making the protein. The different chemical environment of amino-acids gives its protons their chemical shifts. In theory then, using these shifts, the peaks in the spectrum can be uniquely assigned to protons in various amino-acids. Following this, a series of experiments are used to selectively excite protons and study the
Figure 1.2: (a) Proton spectrum of a small peptide. The proton resonances are clearly resolved. (b) Proton spectrum of a protein. The spectrum has overlapping resonance lines.

effect on other protons. This effect, nuclear Overhauser effect (NOE), is proportional to $r^{-6}$, where $r$ is the distance between the nuclei, Solomon (1955). These experiments then provide distance information between protons of various amino-acids. Once the distance information is known, one can solve for a folded configuration of the protein, which satisfies these distance constraints. In practice, the proton spectrum of a large protein molecule is poorly resolved, see Fig. 1.2(b). Larger proteins have larger relaxation rates hence the peaks in the spectrum have bigger linewidths. Also sheer larger number of protons leads to crowding of the spectra. The frequency labelling task becomes impossible.

To circumvent this problem, methods of two-dimensional NMR were invented, Ernst et al. (1987). The 2D NMR experiments generate a two dimensional spectrum, where each peak is labelled by the Larmor frequencies of a coupled spin pair. For example, the first label is the Larmor frequency of a proton (spin $I$) and the second label is the Larmor frequency of another spin $S$, coupled to proton, say $^{15}\text{N}$. As a result, two protons with overlapping Larmor frequencies, which cannot be resolved in their one-dimensional spectra, now can be distinguished by their nitrogen frequencies. This way, the one-dimensional spectrum of protons can be resolved and peaks can be assigned a unique frequency label. The basic steps of a 2D NMR experiment are shown in Fig. 1.3(a). The goal is to generate a map of
§1.0 Introduction

Figure 1.3: (a) Basic steps of a 2D NMR experiment. Spins $I$ evolve for time $t_1$, while spins $S$ for time $t_2$. The NMR signal is recorded during the second time interval. (b) The Fourier transform of the recorded signal gives a peak at the frequencies $(\omega_I, \omega_S)$.

frequencies $(\omega_I, \omega_S)$ for a coupled spin pair ensemble.

The 2D experiments begin with putting a sample containing coupled spin pairs $IS$ in a static magnetic field $B_0$ which then defines the Larmor frequencies $\omega_I, \omega_S$ for the spins. Initially, spins $I$ are excited by an oscillating magnetic field with frequency corresponding approximately to their Larmor frequency. The net magnetization of spins $I$ is driven to the transverse plane, where it precesses for some time $t_1$. Then, by applying the appropriate rf magnetic field and using the evolution of $IS$ coupling between the spins, the magnetization is transferred from spins $I$ to spins $S$. Spins $S$ are then excited and the precession of their magnetization on the transverse plane is recorded for some time $t_2$. The experiment is performed again by incrementing the value of $t_1$. At the end, the recorded signal $s(t_1, t_2)$ is indexed by the time $t_1$ of the precession of spins $I$ and $t_2$, the time of the precession of spins $S$. Signal $s(t_1, t_2)$ has the form

$$s(t_1, t_2) = \eta e^{-R_I t_1} e^{-R_S t_2} \cos \omega_I t_1 \cos \omega_S t_2,$$

(1.8)

where $R_I$ and $R_S$ are the relaxation rates for spins $I$ and $S$ respectively and $\eta$ is the efficiency of the intermediate step involving transfer of magnetization. This transfer efficiency is limited by the phenomenon of relaxation. If we perform a Fourier transform of the signal, this gives a peak at frequencies $(\omega_I, \omega_S)$, see Fig. 1.3(b). The decay rates $R_I$ and $R_S$ determine the linewidths and the efficiency $\eta$ governs the sensitivity of the experiments.

The basic idea of a 2D NMR experiment can be extended to 3 or 4 dimensional experiments which generate frequency information of spins in a coupled spin network. In NMR spectroscopy of proteins, many such elaborate experiments have been developed to improve
resolution of the data, Cavanagh et al. (1995). The goal of this thesis is to find optimal 
rf excitations that maximize the magnetization transfer for various topologies of coupled 
spins, encountered in common protein NMR experiments, and also to derive upper limits 
for the maximum achievable efficiency. In other words, we want to calculate the optimal 
magnetic field that maximizes the transfer efficiency $\eta$ in Eq. (1.8) above and, additionally, 
the corresponding maximum value of $\eta$. From mathematical control theory perspective, 
these problems involve computing the reachable sets of bilinear control systems with drift, 
as we shall see in subsequent chapters. It should be emphasized that the methods developed 
in this thesis are not limited to NMR applications, but are also useful for answering impor-
tant questions in quantum information theory. It is a fundamental problem to understand 
the extent to which an open quantum system can be controlled. For example, how much 
entanglement can be produced in the presence of relaxation? Questions like this are related 
to optimal control of open quantum systems.

The rest of the thesis is organized as follows

- In chapter 2 we recapitulate the fundamentals of relaxation in liquid state NMR. We 
derive the master equation which describes the density matrix evolution for the open 
quantum system of interest. We also present the basic relaxation mechanisms in liquid 
state NMR: The dipole-dipole (DD) interaction and the chemical shift anisotropy 
(CSA), as well as their interference effects.

- In chapter 3 we use optimal control methods to compute analytical bounds for the 
efficiency of polarization-coherence transfer between a pair of coupled nuclear spins, 
in the presence of longitudinal and transverse relaxation. We also derive relaxation 
optimized pulse sequences which achieve these bounds and maximize the sensitivity 
of the experiments.

- In chapter 4, motivated by optimal control problems arising in the previous chapter, 
we study a class of dissipative bilinear control systems. For these systems it is shown 
that the optimal solution, corresponding to the maximum transfer efficiency, and the 
reachable set can be found by solving a semidefinite program. As a practical result, 
we obtain an upper bound for the polarization-coherence transfer efficiency along an 
Ising spin chain, a system studied in the next chapter.

- In chapter 5 we examine the problem of optimal polarization-coherence transfer 
through an Ising spin chain in the presence of transverse relaxation. We study in
detail the basic step involved in this transfer, the so-called spin order transfer. We find an upper bound for the efficiency of this step, using the theory of the preceding chapter, and calculate numerically optimal pulses that approach this bound. Finally, we show that in many cases these optimal pulses can be very well approximated by a Gaussian profile.

- In chapter 6, which concludes the thesis, we give some guidelines for future work regarding protein NMR experiments.
Chapter 2

Fundamentals of Relaxation

Relaxation is the name given to the phenomenon of evolution towards statistical equilibrium of macroscopic systems. The case in which we are interested is that when the whole system can be modeled as being composed of two weakly interacting parts: the spin system, consisting of all degrees of freedom dependent on the spin operators of the nuclei, and the so-called lattice, consisting of all other degrees of freedom of the liquid sample, associated with the molecular rotations and translations. Nuclear magnetic relaxation is in that case the evolution of the spin system towards thermal equilibrium with the lattice, and is called spin-lattice relaxation. The lattice is an ‘infinite’ thermostat, in the sense that its own thermal equilibrium state is not affected by the exchange of energy with the spin system taking place during relaxation.

The purpose of this chapter is not to give a complete theory of relaxation, but to show the basic principle of how relaxation phenomena can be taken into account in a quantitative manner. This can be done only by a quantum calculation that we describe together with the physical approximations on which it is based. Having derived the quantitative description of relaxation, we show how the parameters that appear there are connected to well-defined characteristics of the dynamics of molecular motions. Finally, we describe the most important relaxation mechanisms in liquids. The theory presented here is used in subsequent chapters to examine specific examples of spin systems encountered in liquid state NMR spectroscopy.
2.1 The Master Equation

The total system (spin system+lattice) is a closed quantum system. From quantum mechanics we know that the density matrix $\rho$ of such a system evolves unitarily in time according to the Liouville-von Neumann equation, see for example Merzbacher (1998),

$$\frac{d\rho}{dt} = -i[H(t), \rho], \quad (2.1)$$

where $H(t)$ is the Hamiltonian of the system. Note that in the above equation we take for simplicity $\hbar = 1$, so the Hamiltonian corresponds to units $E/\hbar$. The goal of this section is to derive a differential equation that describes the evolution of the density matrix of the spin system alone (the so-called master equation). Note that the spin system alone is an open system, since it interacts with the surrounding medium (the lattice). So, we expect that the evolution described by the master equation is not unitary.

Our presentation is based on Goldman (2001). We follow the semiclassical approach, where the spin system is treated quantum-mechanically and the lattice classically. The Hamiltonian of the total system is written as the sum of a deterministic Hamiltonian, which describes evolution under coherent processes and acts only on the spin system, $H_d(t)$, and a stochastic Hamiltonian, $H_1(t)$, which couples the spin system to the lattice

$$H(t) = H_d(t) + H_1(t). \quad (2.2)$$

The deterministic Hamiltonian is also the sum of two terms. A main, static Hamiltonian, $H_0$, and a time-dependent one, $H_{rf}(t)$, that describes the effect of the applied radio-frequency magnetic field. We derive the master equation for $H_{rf} = 0$. The result remains valid for $H_{rf} \ll H_0$, Abragam (1983), in other words when the $rf$ magnetic field is much smaller than the static field, $B_{rf} \ll B_0$. Since we take $H_d = H_0$, the total Hamiltonian $H$ becomes

$$H = H_0 + H_1(t). \quad (2.3)$$

The Hamiltonian $H_1(t)$ is a random function of time. The random character of $H_1(t)$ is due to the fact that molecules in solutions are constantly being bombarded with solvent molecules and undergo random ‘Brownian’ motion as a result. This stochastic Brownian motion modulates randomly the various magnetic interactions involving the nuclear spins, thus it is the principle mechanism of relaxation in liquid state NMR spectroscopy. We assume that $H_1(t)$ has a vanishing average value. If it is not, we can incorporate the average value to $H_0$, so the resulting $H_1(t)$ has a zero average.
We begin by removing the static Hamiltonian $H_0$, so as to single out the effect of the random perturbation $H_1(t)$. This is done by the use of an interaction representation. We replace all operators $Q$ in the laboratory frame by the operators

$$Q \rightarrow \tilde{Q}(t) = \exp(iH_0 t) Q \exp(-iH_0 t).$$

(2.4)

The evolution of the density matrix in this representation is given by

$$\frac{d\tilde{\rho}}{dt} = -i[H_1(t), \tilde{\rho}].$$

(2.5)

The Hamiltonian $\tilde{H}_1(t)$ has a double time dependence: that due to its random character and that due to the passage to the interaction representation.

By formal integration of Eq. (2.5), we obtain

$$\tilde{\rho}(t) = \tilde{\rho}(0) - i \int_0^t [\tilde{H}_1(t'), \tilde{\rho}(t')] dt',$n

(2.6)

a form which is inserted into the right-hand side of Eq. (2.5). We obtain

$$\frac{d\tilde{\rho}}{dt} = -i[\tilde{H}_1(t), \tilde{\rho}(0)] - \int_0^t [\tilde{H}_1(t), [\tilde{H}_1(t'), \tilde{\rho}(t')]] dt'.$n

(2.7)

This expression is rigorous. For physical reasons, we subject it to two modifications:

1. We take an ensemble average of all terms. The reason is that remote parts of a large system relax independently of each other. Each of them experiences a random coupling with a different evolution history, leading to a different local density matrix $\tilde{\rho}$. These parts are simulated by different members of a Gibbs ensemble. All have identical initial $\tilde{\rho}(0)$ and, since $H_1(t)$ has a vanishing average, the first term on the right hand side of Eq. (2.7) vanishes.

2. In the remaining term on the right hand side of Eq. (2.7), we replace $\tilde{\rho}(t)$ by

$$\tilde{\rho}(t) \rightarrow \tilde{\rho}(t) - \tilde{\rho}_{eq},$$

(2.8)

where $\tilde{\rho}_{eq}$ is the thermal equilibrium form of the density matrix for the Hamiltonian $H_0$. This is the phenomenological account for the finite lattice temperature and we can justify it only if we consider the full quantum-mechanical approach.

We obtain in place of Eq. (2.7)

$$\frac{d\tilde{\rho}}{dt} = - \int_0^t [\tilde{H}_1(t), [\tilde{H}_1(t'), \tilde{\rho}(t') - \tilde{\rho}_{eq}]] dt',$n

(2.9)
where the ensemble average is denoted by an overbar.

In the next step we expand the spin-lattice Hamiltonian in the form

\[ H_1(t) = \sum_\alpha V_\alpha F_\alpha(t) = \sum_\alpha V_\alpha^\dagger F_\alpha^*(t), \tag{2.10} \]

where the \( V_\alpha \) are spin operators, which we call relaxation operators, and the \( F_\alpha(t) \) are random functions of time. The equality of the two forms is a consequence of the fact that \( H_1 \), being a Hamiltonian, is Hermitian. The decomposition (2.10) is made in such a way that

\[ [H_0, V_\alpha] = \omega_\alpha V_\alpha, \tag{2.11} \]

whence

\[ \tilde{V}_\alpha(t) = e^{iH_0 t} V_\alpha e^{-iH_0 t} = e^{i\omega_\alpha t} V_\alpha. \tag{2.12} \]

Operators \( V_\alpha \) satisfying (2.11) are eigenoperators of \( H_0 \), with corresponding eigenfrequencies \( \omega_\alpha \). Such a decomposition of \( H_1 \) in eigenoperators of \( H_0 \) is always possible. As an example, let us choose

\[ V_\alpha = |i\rangle \langle i| H_j |j\rangle \langle j|, \tag{2.13} \]

where \( |i\rangle \) and \( |j\rangle \) are eigenkets of \( H_0 \). Then

\[ \omega_\alpha = \langle i| H_0 |i\rangle - \langle j| H_0 |j\rangle. \tag{2.14} \]

The operator \( V_\alpha \) defined by Eq. (2.13) has only one nonvanishing matrix element. In usual problems, operators \( V_\alpha \) have several nonvanishing matrix elements, and there are several possible choices for the \( V_\alpha \). We can write (2.9) under the form

\[ \frac{d\tilde{\rho}}{dt} = -\sum_{\alpha, \beta} \int_0^t \{ [\tilde{V}_\alpha(t), [V_\beta^\dagger(t'), \tilde{\rho}(t') - \tilde{\rho}_{eq}]] F_\alpha(t) \overline{F_\beta^*(t')} dt'. \tag{2.15} \]

Note that, since the different density matrices \( \tilde{\rho}(t') \) in different members of the Gibbs ensemble experiencing different time evolution of the random functions \( F_\alpha(t), F_\beta^*(t') \), it is essential to use a joint average over the spin part, i.e. \( \tilde{\rho}(t') \), and the lattice part, i.e. \( F_\alpha(t), F_\beta^*(t') \), under the integral of Eq. (2.15).

In order to proceed further, it is necessary to make some approximations. First we employ the so-called Markovian approximation. We suppose that the fluctuation of the random perturbation is fast compared with the evolution through relaxation of the physical variables under study. More specifically, let \( \tau_c \) be the time scale for the fluctuation of the random functions \( F(t) \); it is the time scale \( t - t' = \tau_c \) over which a typical product \( F_\alpha(t) \overline{F_\beta^*(t')} \)
decays by a substantial amount (a better definition of $\tau_c$ will be given in the next section). We assume that the quantity $\langle Q \rangle$ (expectation value of the operator $Q$, it depends on the density matrix) we are interested in has a slow evolution on the time scale $\tau_c$. Justification of the validity of this assumption can be done \textit{a posteriori}: we first derive the evolution equation for $\langle Q \rangle(t)$ under the assumption that it is slow, compared with $\tau_c$, and then we determine which are the conditions for the result to be consistent with this assumption.

In liquids, $\tau_c$ is on the order of several times an inter-collision time ($10^{-14} - 10^{-12}$ s), so about $10^{-12} - 10^{-9}$ s, while the time scale for a significant change in $\langle Q \rangle(t)$ is on the order of ms or longer, as we can assure beforehand, Cavanagh et al. (1995). Thus the Markovian approximation is valid. Note that what we really want to guarantee with this approximation is that the density matrix in (2.15) changes slightly during a time interval $\tau_c$, but this statement is termed in a more physical language by using expectation values of (possibly observable) operators. The second approximation is that we choose $t \gg \tau_c$ in (2.15).

The shortness of $\tau_c$ compared to the time scale of evolution of $\tilde{\rho}$ and the choice of $t \gg \tau_c$ have some important consequences, as we show next. A first consequence is that we may replace $\tilde{\rho}(t')$ by $\tilde{\rho}(t)$ on the right-hand side of (2.15), since only values of $t'$ differing from $t$ by only a few times $\tau_c$ contribute to the integral. Another consequence is that, since each member of the Gibbs ensemble has experienced its random perturbation for many times $\tau_c$, the effect of the difference in their random evolutions averages out, so that the various $\tilde{\rho}(t')$ are equal. If we use the symbol $\tilde{\sigma}$ to denote the Gibbs ensemble average $\tilde{\sigma} = \bar{\tilde{\rho}}$, then it is also $\tilde{\sigma} = \tilde{\rho}$, and we may replace $\tilde{\rho}(t'), \tilde{\rho}(t)$ by $\tilde{\sigma}(t)$. As a result, there is a decoupling between the average over the spin part and that over the lattice part. Equation (2.15) is then replaced by

$$
d\tilde{\sigma}/dt = - \sum_{\alpha, \beta} \int_0^t [V_\alpha(t), [\bar{V}_\beta(t'), \tilde{\sigma}(t) - \tilde{\sigma}_{eq}]] F_{\alpha}(t) F_{\beta}^*(t') dt' . \quad (2.16)$$

The matrix $\tilde{\sigma}$ is the average of $\tilde{\rho}$ over the lattice degrees of freedom, i.e. the reduced density matrix for the spin system. We limit ourselves to stationary random functions, a realistic assumption in most practical cases, that is such that

$$F_{\alpha}(t) F_{\beta}^*(t') = C_{\alpha\beta}(|t - t'|) \quad (2.17)$$

From the expression (2.16), together with Eq. (2.12) and the definition (2.17), we obtain

$$
d\tilde{\sigma}/dt = - \sum_{\alpha, \beta} [V_\alpha, [\bar{V}_\beta(t), \tilde{\sigma}(t) - \tilde{\sigma}_{eq}]] \int_0^t C_{\alpha\beta}(t - t') e^{i(\omega_\alpha t - \omega_\beta t')} dt' . \quad (2.18)$$
If we make the change of variables $t - t' = \tau$, the integral term in the above relation becomes

$$\int_0^t C_{\alpha\beta}(t-t') e^{i(\omega_\alpha t - \omega_\beta t')} dt' = e^{i(\omega_\alpha - \omega_\beta)t} \int_0^t C_{\alpha\beta}(\tau) e^{i\omega_\beta \tau} d\tau .$$  \hspace{1cm} (2.19)

From the definition of $\tau_c$ we know that the correlation function $C_{\alpha\beta}$ is nonzero for a time interval equal to a few times $\tau_c$ around $\tau = 0$. Using this and the assumption $t \gg \tau_c$, we can extend the upper limit of the integral in the right hand side of Eq. (2.19) to infinity.

Thus

$$\int_0^t C_{\alpha\beta}(t-t') e^{i(\omega_\alpha t - \omega_\beta t')} dt' \approx e^{i(\omega_\alpha - \omega_\beta)t} \int_0^\infty C_{\alpha\beta}(\tau) e^{i\omega_\beta \tau} d\tau = e^{i(\omega_\alpha - \omega_\beta)t} J_{\alpha\beta}(\omega_\beta) ,$$  \hspace{1cm} (2.20)

where $J_{\alpha\beta}$, called a spectral density, is defined by

$$J_{\alpha\beta}(\omega) = \int_0^\infty C_{\alpha\beta}(\tau) e^{i\omega \tau} d\tau .$$  \hspace{1cm} (2.21)

Eq. (2.18) becomes

$$\frac{d\tilde{\sigma}}{dt} = -\sum_{\alpha, \beta} [V_\alpha, [V_\beta^+, \tilde{\sigma} - \tilde{\sigma}_{eq}]] e^{i(\omega_\alpha - \omega_\beta)t} J_{\alpha\beta}(\omega_\beta) .$$  \hspace{1cm} (2.22)

We next show that when the frequency difference $|\omega_\alpha - \omega_\beta|$ is large enough (in a sense to be explained) then the contribution of the corresponding term to the variation of $\tilde{\sigma}$ is negligible compared to the contribution of terms with $\omega_\alpha = \omega_\beta$. According to Eq. (2.22), a general measure of the contribution of the term $\alpha\beta$ to $\tilde{\sigma}$ is

$$\int_0^t e^{i(\omega_\alpha - \omega_\beta)\tau'} \times J_{\alpha\beta}(\omega_\beta) .$$  \hspace{1cm} (2.23)

When $\omega_\alpha \neq \omega_\beta$ this gives

$$\frac{e^{i(\omega_\alpha - \omega_\beta)t} - 1}{i(\omega_\alpha - \omega_\beta)} \times J_{\alpha\beta}(\omega_\beta) ,$$  \hspace{1cm} (2.24)

while for $\omega_\alpha = \omega_\beta$ gives

$$t J_{\alpha\beta}(\omega_\beta) .$$  \hspace{1cm} (2.25)

Comparing the last two expressions we see that for times $t$ that satisfy

$$t \gg \frac{1}{|\omega_\alpha - \omega_\beta|} ,$$  \hspace{1cm} (2.26)

we can ignore the contribution of the nonsecular term. Now if the characteristic relaxation time for our system (the time for which a significant change in $\tilde{\sigma}$ happens) satisfies the above inequality for every pair of different frequencies, then the oscillating terms $e^{i(\omega_\alpha - \omega_\beta)t}$ average to zero much more rapidly than relaxation occurs and we can eliminate these terms.
from Eq. (2.22). This is called the secular approximation, since we keep only terms with 
\( \omega_\alpha = \omega_\beta \). Furthermore, if none of the frequencies are degenerate, the only secular and
non-zero terms, i.e. terms with \( \omega_\alpha = \omega_\beta \), are those with \( \alpha = \beta \). This is the case that we
consider in this thesis. Thus, we need to take into account only the correlation functions
\( C_{\alpha\alpha} \) and the corresponding spectral densities \( J_\alpha \). For simplicity we denote them as \( C_\alpha \)
and \( J_\alpha \) respectively. Then, Eq. (2.22) takes the form
\[
\frac{d\tilde{\sigma}}{dt} = -\sum_\alpha [V_\alpha, [V_\alpha^\dagger, \tilde{\sigma} - \tilde{\sigma}_{eq}]]J_\alpha(\omega_\alpha) .
\] (2.27)

This is the equation for the evolution of the density matrix \( \tilde{\sigma} \) in the interaction
representation. Using the above equation, the relation
\[
\sigma = e^{-iH_0 t} \tilde{\sigma} e^{iH_0 t}
\] (2.28)
and the property
\[
U[A, B]U^\dagger = [UAU^\dagger, UBU^\dagger],
\] (2.29)
for unitary operators \( U \), we can find the corresponding evolution equation for \( \sigma \) in the
initial Schrödinger representation
\[
\frac{d\sigma}{dt} = -i[H_0, \sigma] - \sum_\alpha [V_\alpha, [V_\alpha^\dagger, \sigma - \sigma_{eq}]]J_\alpha(\omega_\alpha) .
\] (2.30)
The spectral density function \( J_\alpha \) is given by
\[
J_\alpha(\omega) = \int_0^\infty C_\alpha(\tau)e^{i\omega\tau} d\tau .
\] (2.31)
Being Fourier transforms over positive time only, the spectral densities are complex
\[
J_\alpha(\omega) = j_\alpha(\omega) + ik_\alpha(\omega) .
\] (2.32)
It can be shown quite general that their real part corresponds to relaxation proper, whereas
their imaginary part produces a small shift of the resonance frequencies, known as the
dynamical shift, which can be accounted by redefining the unperturbed Hamiltonian \( H_0 \).
Thus, the evolution of the reduced density matrix \( \sigma \) of the spin system is described, in the
laboratory frame, by the master equation
\[
\frac{d\sigma}{dt} = -i[H_0, \sigma] - \sum_\alpha j_\alpha(\omega_\alpha)[V_\alpha, [V_\alpha^\dagger, \sigma - \sigma_{eq}]] .
\] (2.33)
Observe that this evolution is not unitary, something expected since the spin system alone is an open quantum system. If we define the relaxation superoperator $L(\cdot)$ as follows

$$L(\cdot) = -\sum_\alpha j_\alpha(\omega_\alpha)[V_\alpha, [V_\alpha^\dagger, \cdot]] ,$$

then Eq. (2.33) takes the form

$$\frac{d\sigma}{dt} = -i[H_0, \sigma] + L(\sigma - \sigma_{eq}) ,$$

which emphasizes the non unitary character of evolution.

In practice, more than one stochastic Hamiltonian capable of causing relaxation (more than one relaxation mechanisms) may be operative. In this circumstance, Eq. (2.10) is generalized to

$$H_1(t) = \sum_m \sum_\alpha V_m^\alpha F_\alpha^m(t) ,$$

where the summation over the index $m$ refers to different relaxation mechanisms or stochastic Hamiltonians. Using (2.36) rather than (2.10) in the derivation given above, we find a generalization of (2.33)

$$\frac{d\sigma}{dt} = -i[H_0, \sigma] - \sum_{m,n} \sum_\alpha j_{mn}^\alpha(\omega_\alpha)[V_m^\alpha, [V_n^\dagger, \sigma - \sigma_{eq}]] ,$$

where

$$j_{mn}^\alpha(\omega) = \text{Re} \left\{ \int_0^\infty C_{mn}^\alpha(\tau)e^{i\omega\tau}d\tau \right\}$$

and

$$C_{mn}^\alpha(\tau) = F_m^\alpha(t)F_n^{\alpha^*}(t + \tau) .$$

The occurrence of cross terms ($m \neq n$) is referred to as interference or cross-correlation effect in relaxation.

Usually, we are not much interested in the evolution of the density matrix itself as in that of measurable spin variables. Let $Q$ be the operator corresponding to such a variable. The expectation value of $Q$ is

$$\langle Q \rangle = \text{Tr}\{Q\sigma\} ,$$

and

$$\frac{d}{dt} \langle Q \rangle = \text{Tr}\{Q\frac{d\sigma}{dt}\} .$$

Using Eq. (2.37) and the property

$$\text{Tr}\{A[B, C]\} = \text{Tr}\{[A, B]C\} ,$$

it is not difficult to show that
twice in succession, we find
\[
\frac{d}{dt} \langle Q \rangle = \langle -i [Q, H_0] \rangle - \sum_{m,n} \sum_{\alpha} j^{mn}_{\alpha}(\omega_\alpha) \{\langle [Q, V^m_{\alpha}], V^{n\dagger}_{\alpha} \rangle \} - \langle [Q, V^m_{\alpha}], V^{n\dagger}_{\alpha} \rangle \} \quad \text{eq}.
\]
Note that the first term in the above equation is the familiar term from Ehrenfest’s theorem in quantum mechanics. The double commutators \([Q, V^m_{\alpha}], V^{n\dagger}_{\alpha} \) will often yield operators differing from \(Q\). In this case one will speak of cross-relaxation processes. It is necessary to calculate the evolution of these operators as well. The procedure must be repeated until reaching a closed system of operators with coupled relaxation processes. We emphasize the difference between cross-relaxation and cross-correlation. Cross-relaxation denotes the inter-conversion between different operators, regardless of the nature of the relaxation mechanism. On the other hand, cross-correlation effects indicate the interference between relaxation mechanisms of different nature.

### 2.2 Spectral Density Functions

In the preceding section we saw that spectral density functions play a central role in the relaxation part of the master equation. A general expression for the spectral density function \(j^{mn}_{\alpha}(\omega)\) is given by Eq. (2.38). It is the real part of the Fourier transform (over positive times only) of the correlation function \(C^{mn}_{\alpha}(\tau)\) given in Eq. (2.39). Correlation functions account for the extent of correlation between the fluctuating spatial functions. These time fluctuations arise from molecular motions. Thus spectral density functions measure the contribution to relaxation rates from motions with frequency components in the range \(\omega\) to \(\omega + d\omega\). To relate relaxation and dynamics, one must first calculate the correlation function and then the corresponding spectral density.

The relaxation mechanisms of interest in the present context arise from tensorial operators of rank two. As a result, the corresponding stochastic functions \(F^m_{\alpha}\) can be written in the form
\[
F^m_{\alpha}(t) = c^m(t) \tilde{Y}^{(2)}_q(\Omega^m_{LF}(t)) \quad \text{eq},
\]
where \(c^m(t)\) is a function of physical constants and spatial variables, \(\tilde{Y}^{(2)}_q\) is a modified second-order spherical harmonic function, related to the corresponding ordinary spherical harmonic through
\[
\tilde{Y}^{(2)}_q = \sqrt{\frac{4\pi}{5}} Y_q^{(2)},
\]
and \(\Omega^m_{LF} = (\theta^m, \phi^m)\) represents polar angles in the laboratory frame. These angles define the orientation of a unit vector that points in the principal direction for the interaction
corresponding to the relaxation mechanism labeled by index $m$. An expression similar to Eq. (2.44) holds for $F_m^n$, a random function corresponding to relaxation mechanism $n$. Then, Eq. (2.39) for the correlation function gives

$$C_{qm}^{mn}(\tau) = c^m(0)c^n(\tau)\tilde{Y}_q^{(2)}(\Omega_{LF}(0))\tilde{Y}^{(2)*}_{q'}(\Omega_{LF}(\tau)),$$  

(2.46)

where we set $t = 0$ without loss of generality, since we examine only stationary correlation functions.

Note that in the above formulas the index $\alpha$ of the previous section has been replaced by the coherence index $q$. This replacement works because we assume that stochastic functions $F_m^q, F_n^{q'}$, with $q \neq q'$, are uncorrelated (this is indeed the case when these functions are proportional to spherical harmonics with constant coefficients). Thus, even if there are relaxation operators with the same eigenfrequency $\omega_\alpha$ (so that the secular condition is satisfied) but different coherence, the corresponding correlation is zero. The only obstacle in the full conformation with the formalism of section 2.1 is that, as we will see, operators $V^m_q$ corresponding to $F^m_q$ may not satisfy the condition (2.11). In this case, we have to decompose further $V^m_q$ to a sum of operators each of which satisfies (2.11). These operators are the actual relaxation operators $V^m_\alpha$ of the preceding section. From Eq. (2.46) we observe that correlation functions depend only on the coherence index $q$, so relaxation operators of the same $q$ ($V_\alpha$ coming from the decomposition of $V_q$) correspond to the same correlation function and thus to the same spectral density. Nevertheless, there exists a difference. For each of these operators the spectral density function is evaluated at the corresponding eigenfrequency $\omega_\alpha$, determined from (2.11), and this way enters the master equation.

The fluctuations of the angles $\Omega_{LF}^n, \Omega_{MF}^n$ are due to both internal motions and overall tumbling of the molecule. These two contributions can be separated by using the matrix elements of finite rotations $D_{qk}^{(2)}$, Edmonds (1960),

$$\tilde{Y}_q^{(2)}(\Omega_{LF}(0)) = \sum_k D_{qk}^{(2)}(\omega(0))\tilde{Y}_k^{(2)}(\Omega_{MF}(0)),$$  

(2.47)

$$\tilde{Y}_q^{(2)*}(\Omega_{LF}(\tau)) = \sum_{k'} D_{qk'}^{(2)*}(\omega(\tau))\tilde{Y}_{k'}^{(2)*}(\Omega_{MF}(\tau)),$$  

(2.48)

where $\omega$ represent the three Euler angles $\alpha, \beta$ and $\gamma$ relating the molecular fixed reference frame to the laboratory frame. Thus $\omega$ accounts for the rotational diffusion of the molecule. On the other hand $\Omega_{MF}$ denotes the two polar angles of the interaction vector in the molecular frame, and its time dependence therefore reflects internal motions. The angles $\Omega_{LF}, \omega$ and $\Omega_{MF}$ are shown schematically in Fig. 2.1. Using Eq. (2.47) and (2.48), Eq.
2.2 Spectral density functions

Figure 2.1: Decomposition of the total motion to overall tumbling and internal motions. The orientation of the interaction vector $OZ$ can be determined either by the angle $\Omega_{LF}$ in the laboratory frame (LF), or by the angle $\Omega_{MF}$ in the molecular fixed frame (MF) and the Euler angles, denoted collectively by $\omega$, that specify the orientation of MF with respect to LF.

(2.46) takes the form

$$C_{mn}^{q} (\tau) = \sum_{k, k'} c_{m}^{n}(0)c_{n}^{m}(\tau)Y_{k}^{(2)}(\omega(0))Y^{(2)*}_{k'}(\omega(\tau)) \tilde{Y}^{(2)}_{k}(\Omega_{MF}(0)) \tilde{Y}^{(2)*}_{k'}(\Omega_{MF}(\tau)) .$$  \hspace{1cm} (2.49)

If overall and internal motions are independent, Eq. (2.49) can be rewritten as

$$C_{q}^{mn} (\tau) = \sum_{k, k'} D_{qk}^{(2)}(\omega(0)) D^{(2)*}_{qk'}(\omega(\tau)) \times c_{m}^{n}(0)c_{n}^{m}(\tau) Y^{(2)}_{k}(\Omega_{MF}(0)) Y^{(2)*}_{k'}(\Omega_{MF}(\tau)) ,$$  \hspace{1cm} (2.50)

where one can now estimate the ensemble average of each contribution independently.

Note that the functions $c_{m}^{n}, c_{n}^{m}$ are incorporated in the second average, which gives the contribution of internal motions. The reason is that these functions depend on spatial variables, like for example the internuclear distance in a dipole interaction, that fluctuate because of the internal motions only.

The first average denotes motional averaging due to rotational diffusion of the whole molecule. We can calculate it using the formula

$$\overline{D_{qk}^{(2)}(\omega(0))D^{(2)*}_{qk'}(\omega(\tau))} = \int \int p(\omega_0, 0)p(\omega, \tau|\omega_0, 0)D_{qk}^{(2)}(\omega_0)D^{(2)*}_{qk'}(\omega)d\omega_0d\omega ,$$  \hspace{1cm} (2.51)

where $p(\omega_0, 0)$ is the probability that the Euler angles ($\alpha, \beta$ and $\gamma$) are $\omega_0$ at time $t = 0$ and $p(\omega, \tau|\omega_0, 0)$ is the conditional probability that the Euler angles are $\omega$ at time $t = \tau$ if they were $\omega_0$ at time $t = 0$. 
It is assumed that all initial orientations are equally probable, so that
\[ p(\omega_0, 0) = \left( \int d\omega_0 \right)^{-1} = \left( \int_0^{2\pi} \int_0^{\pi} \int_0^{2\pi} d\alpha_0 \sin \beta_0 d\beta_0 d\gamma_0 \right)^{-1} = \frac{1}{8\pi^2}. \] (2.52)

The crucial step is to calculate the conditional probability. Let \( \psi(\omega, t) \) be the (unconditioned) probability that the Euler angles are \( \omega \) at time \( t \). It has been shown, Favro (1960), that for a spherically symmetric molecule (treated as a rigid body) subjected to rotational Brownian motion, \( \psi \) satisfies the diffusion equation
\[ \frac{\partial \psi}{\partial t} = -D L^2 \psi, \] (2.53)
where \( D \) is the diffusion coefficient and \( L^2 \) is the square magnitude of the angular momentum operator for a rigid rotor. The form of this operator, expressed in terms of the Euler angles, is Edmonds (1960)
\[ L^2 = -\frac{\partial^2}{\partial \beta^2} - \cot \beta \frac{\partial}{\partial \beta} - \frac{1}{\sin^2 \beta} \left( \frac{\partial^2}{\partial \alpha^2} + \frac{\partial^2}{\partial \gamma^2} - 2 \cos \beta \frac{\partial^2}{\partial \alpha \partial \gamma} \right). \] (2.54)

The conditional probability \( p(\omega, \tau | \omega_0, 0) \) is the solution of (2.53) which satisfies the initial condition
\[ \psi(\omega, 0) = \delta(\omega - \omega_0). \] (2.55)

We solve this problem. Suppose that we know a complete orthonormal set \( \{ \psi_n \} \) of eigenfunctions of the operator \( L^2 \), with corresponding eigenvalues \( E_n \). Then, we can seek the solution of (2.53) in the expansion form
\[ \psi(\omega, t) = \sum_n a_n(t) \psi_n(\omega). \] (2.56)

By substituting (2.56) in (2.53) and taking into account the relation \( L^2 \psi_n = E_n \psi_n \) as well as the orthonormality of the eigenfunctions, it follows that
\[ \frac{da_n}{dt} = -D E_n a_n, \] (2.57)
thus
\[ a_n(t) = a_n(0) e^{-D E_n t}. \] (2.58)

The completeness of the eigenfunctions implies that the function \( \delta(\omega - \omega_0) \) can also be expanded in terms of \( \psi_n \) as, Landau and Lifshitz (1977)
\[ \delta(\omega - \omega_0) = \sum_n \psi_n^*(\omega_0) \psi_n(\omega), \] (2.59)
§ 2.2 SPECTRAL DENSITY FUNCTIONS

whence

\[ a_n(0) = \psi_n^*(\omega_0) \]  

(2.60)

and

\[ p(\omega, t|\omega_0, 0) = \psi(\omega, t) = \sum_n \psi_n^*(\omega_0) \psi_n(\omega) e^{-DE_n t} . \]  

(2.61)

A complete set of eigenfunctions of the operator \( L^2 \) are the wavefunctions of the symmetric top, Landau and Lifshitz (1977), which are also appropriate for the spherical top, and are given by

\[ \psi_n(\omega) = \sqrt{\frac{2L + 1}{8\pi^2}} D_{QK}^{(L)}(\omega) . \]  

(2.62)

The corresponding eigenvalues are

\[ E_n = L(L + 1) . \]  

(2.63)

Substituting the above equations in (2.61) we find that the conditional probability at time \( t = \tau \) is given by

\[ p(\omega, \tau|\omega_0, 0) = \sum_{L,Q,K} \frac{2L + 1}{8\pi^2} D_{QK}^{(L)}(\omega_0) D_{QK}^{(L)}(\omega) e^{-DL(L+1)\tau} . \]  

(2.64)

Eq. (2.64) in turn is substituted into Eq. (2.51). The integrals can easily be done using the orthogonality relation for the \( D \) matrices, Edmonds (1960),

\[ \frac{1}{8\pi^2} \int D_{qk}^{(l)}(\omega) D_{qk'}^{(l')*}(\omega) d\omega = \frac{1}{2l + 1} \delta_{ll'}\delta_{kk'} . \]  

(2.65)

We finally obtain

\[ \frac{D_{qk}^{(2)}(\omega(0)) D_{qk'}^{(2)*}(\omega(\tau))}{5} = e^{-6D\tau} \delta_{kk'} . \]  

(2.66)

If we substitute this result in (2.50) we get

\[ C_{mn}^{(q)}(\tau) = \frac{1}{5} e^{-\tau/\tau_c} \sum_k c_m(0)c_n(\tau) \tilde{Y}_{k}^{(2)}(\Omega_{MF}^{(0)}) \tilde{Y}_{k}^{(2)*}(\Omega_{MF}^{(\tau)}) , \]  

(2.67)

where \( \tau_c = 1/6D \) is the rotational correlation time of the spherically symmetric molecule. It is approximately the average time for the molecule to rotate by one radian. The correlation time varies with molecular size, solvent viscosity and temperature, but generally \( \tau_c \) is of the order of picoseconds for small molecules and of the order of nanoseconds for biological macromolecules in aqueous solution, Cavanagh et al. (1995). Note that the result is independent of the coherence number \( q \). Thus we can use the symbol \( C_{mn}^{(q)} \) instead of \( C_{mn}^{(q)} \) and rewrite (2.67) as

\[ C_{mn}^{(q)}(\tau) = \frac{1}{5} e^{-\tau/\tau_c} \sum_k c_m(0)c_n(\tau) \tilde{Y}_{k}^{(2)}(\Omega_{MF}^{(0)}) \tilde{Y}_{k}^{(2)*}(\Omega_{MF}^{(\tau)}) . \]  

(2.68)
This result can be written in the product form

\[ C_{mn}(\tau) = C_O(\tau)C_I^{mn}(\tau), \quad (2.69) \]

where

\[ C_O(\tau) = \frac{1}{5} e^{-\tau/\tau_c} \quad (2.70) \]

is the correlation function for the overall molecular tumbling, and

\[ C_I^{mn}(\tau) = \sum_k c^m(0)c^n(\tau)Y^{(2)}_k(\Omega_{MF}(0))Y^{(2)\ast}_k(\Omega_{MF}(\tau)) \quad (2.71) \]

is the correlation function for the internal motions.

In the absence of internal motions (rigid molecule) \( c^m, c^n, \Omega_{MF}^m \) and \( \Omega_{MF}^n \) are constant. Then, Eq. (2.71) becomes

\[ C_I^{mn} = c^m c^n \sum_k Y^{(2)}_k(\Omega_{MF})Y^{(2)\ast}_k(\Omega_{MF}) \quad (2.72) \]

Using the well known addition theorem for spherical harmonics, we find

\[ C_I^{mn} = c^m c^n P_2(\cos \theta_{mn}) \quad (2.73) \]

where \( \theta_{mn} \) denotes the (constant) angle between the two (rigid) vectors that correspond to the relaxation mechanisms \( m \) and \( n \), and \( P_2 \) is the second order Legendre polynomial.

For autocorrelated relaxation (\( m=n \)), \( \theta_{mn} = 0 \) and \( P_2(\cos \theta_{mn}) = 1 \). The total correlation function \( C^{mn} \) is

\[ C^{mn}(\tau) = \frac{1}{5} e^{-\tau/\tau_c} \times c^m c^n P_2(\cos \theta_{mn}) \quad (2.74) \]

and the corresponding spectral density function is

\[ j^{mn}(\omega) = c^m c^n P_2(\cos \theta_{mn})J(\omega) \quad (2.75) \]

where

\[ J(\omega) = \text{Re} \left( \int_0^\infty C_O(\tau) \exp(i\omega \tau) d\tau \right) = \frac{1}{5} \frac{\tau_c}{1 + \omega^2 \tau_c^2} \quad (2.76) \]

We observe that the functional form of the spectral density function for a rigid molecule is Lorentzian. A graph of \( J(\omega) \) versus \( \omega \) is shown in Fig. 2.2. The plot of \( J(\omega) \) is relatively constant for \( \omega^2 \tau_c^2 \ll 1 \) and then begins to decrease rapidly at \( \omega^2 \tau_c^2 \approx 1 \). If molecular motion is sufficiently rapid to satisfy \( \omega^2 \tau_c^2 \ll 1 \), the extreme narrowing condition obtains and \( J(\omega) \approx J(0) \). For sufficiently slow molecular motion, \( \omega^2 \tau_c^2 \gg 1 \), \( J(\omega) \propto \omega^{-2} \), and the slowly tumbling regime or spin diffusion limit is reached. One other important observation
Figure 2.2: Spectral density function for an isotropic rotor. Calculation was performed using (2.76) with \( \tau_c = 10 \) ns.

is that when we measure relaxation rates for rigid molecules, we obtain information about the constants \( e^m, e^n \) and \( \theta_{mn} \), i.e. about the molecular structure.

The internal motions can be taken into account in a way similar to that we followed for the case of the overall tumbling. In other words, one can calculate \( C_I^{mn}(\tau) \) by using an equation similar to Eq. (2.51), where the conditional probability is derived by an appropriate model describing the internal motions. Here we follow a different approach. We present the Lipari-Szabo ‘model-free’ method, Lipari and Szabo (1982a,b). The significant advantage of this method is that its application does not require the specification of the microscopic motional model. It has been extensively used to analyze spin relaxation in proteins Palmer (1993), Wagner (1993).

We start by examining the quantities \( C_I^{mn}(0) \) and \( C_I^{mn}(\infty) \) from Eq. (2.71). Using the addition theorem of spherical harmonics, we find that

\[
C_I^{mn}(0) = e^m e^n P_2(\cos \theta_{mn}).
\]  

(2.77)

For \( C_I^{mn}(\infty) \) we use the following property of correlation functions

\[
\lim_{t \to \infty} A(0)B(t) = \overline{A} \times \overline{B},
\]

(2.78)

and we get

\[
C_I^{mn}(\infty) = \sum_k e^m \tilde{Y}_k^{(2)}(\Omega_M^m) \times e^n \tilde{Y}_k^{(2)\ast}(\Omega_M^n).
\]

(2.79)
Further calculation of $C_{I}^{mn}(0), C_{I}^{mn}(\infty)$ requires specification of the microscopic motional model.

In the Lipari-Szabo method, the correlation function is approximated by a function with exponential time dependence that is exact for $\tau = 0$ and $\tau = \infty$:

$$C_{I}^{mn}(\tau) = C_{I}^{mn}(\infty) + [C_{I}^{mn}(0) - C_{I}^{mn}(\infty)]e^{-\tau/\tau_{e}},$$  \hspace{1cm} (2.80)

where $\tau_{e}$ is the effective correlation time characterizing the internal motions. The above relation can be rewritten in the form

$$C_{I}^{mn}(\tau) = C_{I}^{mn}(0)[S_{mn} + (1 - S_{mn})e^{-\tau/\tau_{e}}],$$  \hspace{1cm} (2.81)

where

$$S_{mn} = \frac{C_{I}^{mn}(\infty)}{C_{I}^{mn}(0)}$$  \hspace{1cm} (2.82)

is the so-called order parameter. We observe that in this approach the correlation function is described by the two parameters $S_{mn}$ and $\tau_{e}$. The total correlation function (2.69) becomes

$$C^{mn}(\tau) = \frac{1}{5} e^{-\tau/\tau_{e}} \times C_{I}^{mn}(0)[S_{mn} + (1 - S_{mn})e^{-\tau/\tau_{e}}]$$

$$= C_{I}^{mn}(0)\left[\frac{1}{5} S_{mn}e^{-\tau/\tau_{e}} + \frac{1}{5} (1 - S_{mn})e^{-\tau/\tau_{i}}\right],$$  \hspace{1cm} (2.83)

where $\tau_{i}$ is defined such that

$$\frac{1}{\tau_{i}} = \frac{1}{\tau_{c}} + \frac{1}{\tau_{e}}.$$  \hspace{1cm} (2.84)

The corresponding spectral density is

$$j_{mn}(\omega) = \frac{C_{I}^{mn}(0)}{5} \left[ \frac{S_{mn} \tau_{e}}{1 + \omega^{2} \tau_{c}^{2}} + \frac{(1 - S_{mn}) \tau_{i}}{1 + \omega^{2} \tau_{i}^{2}} \right].$$  \hspace{1cm} (2.85)

The model-free approach, so-called because $S_{mn}$ and $\tau_{e}$ have model-independent significance, consists of using the above spectral density to least-squares fit relaxation data by treating $S_{mn}$ and $\tau_{e}$ as adjustable parameters.

We conclude this section by examining some important limiting cases of the above formula. In the absence of internal motions (rigid molecule), $c^{m}, c^{n}, \Omega_{MF}^{m} \text{ and } \Omega_{MF}^{n}$ are constant. Then $C_{I}^{mn}(0) = C_{I}^{mn}(\infty) = c^{m}c^{n}P_{2}(\cos \theta_{mn})$, so $S_{mn} = 1$ and Eq. (2.85) reduces to Eq. (2.75). For slow internal motions $\tau_{e} \gg \tau_{c}$ and hence $\tau_{i} \simeq \tau_{c}$. The contribution of the order parameter is eliminated from Eq. (2.85) and we recover again Eq. (2.75). In other words, relaxation measurements are insensitive to motions on a time scale slower than the overall tumbling of the molecule. On the other hand, when $\tau_{e} \ll \tau_{c}$ (fast internal motion limit), Eq. (2.85) simplifies to

$$j_{mn}(\omega) = \frac{C_{I}^{mn}(0)}{5} \left[ \frac{S_{mn} \tau_{c}}{1 + \omega^{2} \tau_{c}^{2}} \right].$$  \hspace{1cm} (2.86)
2.3 Basic Relaxation Mechanisms in Liquid State NMR

A very large number of physical interactions give rise to stochastic Hamiltonians capable of mediating spin relaxation. These various relaxation mechanisms in NMR spectroscopy have been well studied, Redfield (1965), Ernst et al. (1987). In liquid solutions, the most important relaxation mechanisms are the intramolecular magnetic dipole-dipole interaction (DD) and the chemical shift anisotropy (CSA) interaction, as well as their interference effects (DD-CSA cross-correlated relaxation and DD-DD interference between dipolar interactions of different pairs of spins). In this section we describe these basic relaxation mechanisms.

2.3.1 Dipolar Relaxation

We examine the case of a system with two nuclear spins, \( I \) and \( S \). We know that the (classical) interaction energy of two magnetic dipoles \( m_I, m_S \), is given by the bilinear expression

\[
E_1 = \frac{1}{r^3} [m_I \cdot m_S - 3(m_I \cdot n)(m_S \cdot n)] ,
\]

where \( r \) is the distance between the two particles carrying the magnetic moments, assumed to be point particles, and \( n = r/r \) is the unit vector in the direction joining them. Owing to the quantization of the magnetic moments it is

\[
m_I = \gamma_I \hbar I , \quad m_S = \gamma_S \hbar S ,
\]

where \( \gamma_I, \gamma_S \) are the gyromagnetic ratios and \( I, S \) are the spin angular momentum operators for spins \( I \) and \( S \), respectively. For spins \( 1/2 \) these vector operators are proportional to the vector operator \( \sigma \) with elements the Pauli spin matrices, i.e.

\[
I = \frac{1}{2} \sigma ,
\]

where

\[
\sigma = \begin{bmatrix}
\sigma_x \\
\sigma_y \\
\sigma_z 
\end{bmatrix}
\]

and

\[
\sigma_x = \begin{bmatrix}
0 & 1 \\
1 & 0
\end{bmatrix} , \quad \sigma_y = \begin{bmatrix}
0 & -i \\
i & 0
\end{bmatrix} , \quad \sigma_z = \begin{bmatrix}
1 & 0 \\
0 & -1
\end{bmatrix} ,
\]
in a representation in which $\sigma_z$ is diagonal. A similar relation holds for $S$. The dipolar Hamiltonian corresponding to the classical energy (2.87) is

$$H_1 = \gamma I_\gamma \gamma h \frac{\mathbf{I} \cdot \mathbf{S} - 3(\mathbf{I} \cdot \mathbf{n})(\mathbf{S} \cdot \mathbf{n})}{r^3}. \quad (2.92)$$

Note that in the above equation we write in the pre-factor $\frac{1}{\sqrt{2}}$ instead of $\frac{1}{\sqrt{h}}$, since our analysis is based on the Liouville-von Neumann equation (2.1) with $\hbar = 1$, so the Hamiltonian actually corresponds to $E/\hbar$. Hamiltonian $H_1$ is randomly modulated, either by the variation of the distance $r$ or by the variation of the interaction direction $\mathbf{n}$. $H_1(t)$ is the stochastic Hamiltonian corresponding to the DD relaxation mechanism.

In a system of axes where $\mathbf{OZ}$ is along $\mathbf{n}$, $H_1$ takes the form

$$H_1 = \gamma I_\gamma \gamma h \frac{\mathbf{I} \cdot \mathbf{S} - 3I_ZS_Z}{r^3}. \quad (2.93)$$

For relaxation calculations, one must write this interaction in a form analogous to the expansion (2.10). To do so, we use the laboratory frame $\mathbf{Oxyz}$, where Oz is along the external static magnetic field. Let $\theta$ and $\phi$ be the polar angles of $\mathbf{OZ}$ in this frame, see Fig. 2.3. We can express the operators $I_A$, $S_A$, $A = X, Y, Z$ as functions of the operators $I_\alpha$, $S_\alpha$, $\alpha = x, y, z$ and the angles $\theta$ and $\phi$. For example

$$I_Z = I_z \cos \theta + I_x \sin \theta \cos \phi + I_y \sin \theta \sin \phi \quad (2.94)$$

Through an elementary calculation, we obtain the following form for Eq. (2.93)

$$H_1 = \sum_{q=-2}^{2} F_q^*(\theta, \phi)V_q(I, S). \quad (2.95)$$
where

\[
\begin{align*}
F_0 &= -c_{DD}^{(2)}(r)\tilde{Y}_0^{(2)}(\theta, \phi), \\
F_{\pm 1} &= \pm c_{DD}^{(2)}(r)\tilde{Y}_{\pm 1}^{(2)}(\theta, \phi), \\
F_{\pm 2} &= -c_{DD}^{(2)}(r)\tilde{Y}_{\pm 2}^{(2)}(\theta, \phi), \\
c_{DD}^{(2)}(r) &= \sqrt{6} \gamma I \gamma S \tilde{r}^3, \\
V_0 &= \frac{1}{\sqrt{6}} (3I_zS_z - I \cdot S), \\
V_{\pm 1} &= \mp \frac{1}{2} (I_{\pm}S_z + I_zS_{\pm}), \\
V_{\pm 2} &= \frac{1}{2} I_{\pm}S_{\pm},
\end{align*}
\]  

(2.96)

(2.97)

(2.98)

and \(I_{\pm} = I_x \pm iI_y, S_{\pm} = S_x \pm iS_y\) are the raising and lowering operators for spins \(I\) and \(S\), respectively. The modified second-order spherical harmonics are given in Table 5.7. Note that the functions \(F_q\) are in the form (2.44), with \(c^{m}(t) = c_{DD}^{(2)}(r(t))\) and \(\Omega_{m_{F}}^{m}(t) = \{\theta(t), \phi(t)\}\). A sign difference does not play any role in the calculation of correlation functions. Observe that \(c_{DD}^{(2)}(t)\) is a function of the physical constants \(\gamma_I, \gamma_S, \hbar\) and of the spatial variable \(r\), the internuclear distance, that changes only because of internal motions, as we mentioned in section 2.2. Also note that for constant \(r\) and isotropic distribution of the interaction direction \(n\) in space, i.e. of the angles \(\theta\) and \(\phi\), the average of the stochastic Hamiltonian is \(\overline{H}_1 = 0\), since \(\overline{F}_q = c_{DD}^{(2)}(r)\tilde{Y}_q^{(2)} = 0\).

In section 2.2 we pointed out that operators \(V_q\) do not necessarily satisfy the eigenfrequency condition (2.11), which is crucial for the derivation of the master equation. Thus, we must further decompose \(V_q\) into operators \(V_\alpha\) that satisfy (2.11). First, we need to specify the static Hamiltonian \(H_0\). For a pair of uncoupled spins \(I\) and \(S\), subjected only to the external static magnetic field along the \(z\) direction, this Hamiltonian is

\[
H_0 = \omega_I I_z + \omega_S S_z,
\]

(2.99)

### Table 2.1: Modified Second-Order Spherical Harmonics

<table>
<thead>
<tr>
<th>(q)</th>
<th>0</th>
<th>(\pm 1)</th>
<th>(\pm 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Y_q^{(2)})</td>
<td>(\frac{1}{2} (3 \cos^2 \theta - 1))</td>
<td>(\sqrt{\frac{3}{2}} \sin \theta \cos \theta e^{\pm i\phi})</td>
<td>(\sqrt{\frac{3}{8}} \sin^2 \theta e^{\pm 2i\phi})</td>
</tr>
</tbody>
</table>
Table 2.2: Relaxation Operators for Dipolar Interaction

<table>
<thead>
<tr>
<th>$V_\alpha$</th>
<th>$\frac{2}{\sqrt{6}}I_zS_z$</th>
<th>$-\frac{1}{2\sqrt{6}}I_\pm S_\mp$</th>
<th>$\mp\frac{1}{2}I_\pm S_\pm$</th>
<th>$\pm\frac{1}{2}I_\pm S_\pm$</th>
<th>$\frac{1}{2}I_\pm S_\pm$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\omega_\alpha$</td>
<td>0</td>
<td>$\pm(\omega_I - \omega_S)$</td>
<td>$\pm\omega_I$</td>
<td>$\pm\omega_S$</td>
<td>$\pm(\omega_I + \omega_S)$</td>
</tr>
<tr>
<td>$q$</td>
<td>0</td>
<td>0</td>
<td>$\pm 1$</td>
<td>$\pm 1$</td>
<td>$\pm 2$</td>
</tr>
</tbody>
</table>

where $\omega_I, \omega_S$ are the Larmor frequencies of the two spins. For spins with very different Larmor frequencies (to avoid degeneracy), the decomposition of $V_q$ is done as follows, Goldman (1988),

\[
V_0 = \frac{2}{\sqrt{6}}I_zS_z - \frac{1}{2\sqrt{6}}I_\pm S_\mp - \frac{1}{2\sqrt{6}}I_\pm S_\pm ,
\]

\[
V_{\pm 1} = \mp\frac{1}{2}I_\pm S_\pm + \frac{1}{2}I_zS_\pm ,
\]

\[
V_{\pm 2} = \frac{1}{2}I_\pm S_\pm .
\]

In Table 2.2 we show the various operators $V_\alpha$ and the corresponding eigenfrequencies $\omega_\alpha$ and coherence numbers $q$. These are the operators that should be used in the master equation. The corresponding spectral density functions are found using the functions $F_q$ presented above and the theory of section 2.2.

2.3.2 Chemical Shift Anisotropy

As we mentioned in the introduction, the electronic environment of the nucleus generates a weak local magnetic field which alters slightly the static field $B_0$, hence shifting the Larmor frequency of the nucleus (the so-called chemical shift). The electronic environment of a nucleus is not isotropic in general, nor is the chemical shift produced by this environment. The chemical shift Hamiltonian of a nuclear spin $I$ is a bilinear combination with respect to the components of the external field and of the spin, of the form

\[
H_{cs} = \gamma_I I \cdot \sigma \cdot B = \gamma_I \sum_{\alpha, \beta} B_\alpha \sigma_{\alpha\beta} I_\beta ,
\]

where $\alpha, \beta = x', y', z'$ are axes linked to the molecule and $\sigma_{\alpha\beta}$ are the components of the chemical shift tensor $\sigma$ (don’t confuse it with the vector with elements the Pauli matrices, presented above).
There is no physical reason why the tensor $\sigma$ should be symmetrical. However, despite much effort, it has not been possible to date to get an experimental proof of the hypothetical asymmetry of this tensor. We will therefore assume for simplicity that it is symmetrical. Since the matrix of this tensor in a given set of axes is symmetrical with real elements, it can be diagonalized by an appropriate orthogonal matrix. This transformation corresponds to a rotation or, what amount to the same, to a change of axes. In this new set of axes OXYZ, called the principal axes of the tensor, Eq. (2.103) takes the form

$$H_{cs} = \gamma I (\sigma_X B_X I_X + \sigma_Y B_Y I_Y + \sigma_Z B_Z I_Z),$$

where $\sigma_X, \sigma_Y, \sigma_Z$ are called the principal values of the chemical shift tensor. When this tensor is axially symmetric along the Z axis, then

$$\sigma_X = \sigma_Y = \sigma_\perp, \quad \sigma_Z = \sigma_\parallel.$$  

(2.105)

It has been shown, Goldman (1984), that an anisotropic CSA tensor can be rewritten as the sum of two axially symmetric tensors. Thus, we just need to examine the relaxation caused by an axially symmetric CSA tensor. When (2.105) is satisfied, the chemical shift Hamiltonian can be written as

$$H_{cs} = \gamma I \left[ \frac{\sigma_\parallel + 2\sigma_\perp}{3} B \cdot I + \frac{\sigma_\parallel - \sigma_\perp}{3} (2B_Z I_Z - B_X I_X - B_Y I_Y) \right].$$

(2.106)

The first term on the right-hand side, which is a scalar product, is invariant under rotations. It is what we have called the chemical shift $\sigma$ in the introduction, with

$$\sigma = \frac{1}{3} (\sigma_\parallel + 2\sigma_\perp).$$

(2.107)

This part of $H_{cs}$ is incorporated in the static Hamiltonian $H_0 = -\gamma I B \cdot I = \omega I I_z$ as a shift in the Larmor frequency of spin $I$

$$\omega' = \omega I (1 - \sigma).$$

(2.108)

As for the second term, called the anisotropic chemical shift interaction, its structure

$$H_1 = \gamma I \frac{\sigma_\parallel - \sigma_\perp}{3} (B \cdot I - 3B_Z I_Z).$$

(2.109)

is comparable to that of the dipolar interaction (2.93). It is modulated as a function of time by the motion, because of the variation of the orientation of the axis OZ.
In complete analogy with the passage from Eq. (2.93) to Eq. (2.95), the form of $H_1$ in the laboratory frame of axes $Oxyz$ (the external static magnetic field is along $z$ axis, $B = B_0 \hat{z}$, see Fig. 2.3) is

$$H_1 = \sum_{q=-1}^{1} F_q^*(\theta, \phi)V_q(I) ,$$

(2.110)

where

$$F_q = e^{CSA} \tilde{Y}_q^{(2)} ,$$

(2.111)

$$c^{CSA} = \sqrt{\frac{2}{3}} \gamma I B_0 (\sigma_{||} - \sigma_{\perp}) ,$$

(2.112)

$$V_0 = \frac{2}{\sqrt{6}} I_z ,$$

(2.113)

$$V_{\pm 1} = \frac{1}{2} I_{\pm} .$$

(2.114)

In this specific case, operators $V_q$ satisfy Eq. (2.11), so they are eigenoperators $V_\alpha$ of the static Hamiltonian $H_0$. In Table 2.3 we show these operators along with the corresponding eigenfrequencies and coherence numbers. Note that although in the table we write $\omega_I$ for the frequency, we actually mean the frequency $\omega'_I$ that includes the chemical shift. The same observation holds for Table 2.2.

### 2.3.3 Interference Effects

Interference phenomena between different relaxation mechanisms arise because of the existent correlation between these mechanisms. Consider for example Eqs. (2.110), (2.95), that describe the stochastic Hamiltonians for the dipolar and the anisotropic chemical shift interactions, respectively. Assume for simplicity that the chemical shift tensor is axially symmetric along the $I - S$ direction. This means that the angle $\Omega = \{\theta, \phi\}$ is the same in

<table>
<thead>
<tr>
<th>$V_\alpha$</th>
<th>$\frac{2}{\sqrt{6}} I_z$</th>
<th>$\frac{1}{2} I_{\pm}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\omega_\alpha$</td>
<td>0</td>
<td>$\pm \omega_I$</td>
</tr>
<tr>
<td>$q$</td>
<td>0</td>
<td>$\pm 1$</td>
</tr>
</tbody>
</table>
both equations. Interference is a consequence of the fact that both stochastic Hamiltonians depend on the same spherical harmonics of the same angle $\Omega$.

As we have already mentioned, interference terms that appear in the master equation have the form $[V_{\alpha}^{DD}, [V_{\alpha}^{ICSA}, \cdot]]$. Not all the pairs of operators in tables 2.2 and 2.3 can give rise to such terms. Only those operators that correspond to the same coherence $q$ and, additionally, to the same frequency $\omega_{\alpha}$, are capable to produce interference effects. For example, we mention the pairs $(V_{\alpha}^{DD}, V_{\alpha}^{CSA}) = (\frac{2}{\sqrt{6}} I_z S_z, \frac{2}{\sqrt{6}} I_z)$, with $q = 0, \omega_{\alpha} = 0$, and $(V_{\alpha}^{DD}, V_{\alpha}^{CSA}) = (\pm \frac{1}{2} I_{\pm} S_z, \frac{1}{2} I_{\pm})$, with $q = \pm 1, \omega_{\alpha} = \pm \omega_I$.

Cross-correlation effects between DD and CSA interactions are not the only interference phenomena encountered in biological macromolecules. Interference between the dipolar interactions of different pairs of spins is also frequently encountered. The prototypical example for DD-CSA cross-correlation is that appeared in a $^1$H–$^{15}$N or $^1$H–$^{13}$C pair. The prototypical example of the DD-DD cross-correlation is the interference between the dipolar interactions in an $I_2 S$ or $I_3 S$ spin system such as a methylene ($I_2$ represents the two methylene protons, $S$ represents either a remote proton or the methylene $^{13}$C) or methyl group ($I_3$ represents the three methyl protons, $S$ represents either a remote proton or the methyl $^{13}$C).

Interference effects can result in cross-relaxation between pairs of operators for which cross-relaxation would not be observed otherwise. Thus, the observation of otherwise ‘forbidden’ cross-relaxation pathways is one of the hallmarks of interference effects.
Chapter 3

Optimal Control of a Coupled Spin Pair in the Presence of Relaxation

In this chapter, we develop methods for optimal manipulation of coupled spin dynamics in the presence of relaxation. These methods are used to compute analytical bounds for the optimal efficiency of polarization-coherence transfer between a pair of coupled nuclear spins in presence of longitudinal and transverse relaxation. We also derive relaxation optimized pulse sequences which achieve these bounds and maximize the sensitivity of the experiments in spectroscopic applications. The main core of the text that follows has been published in the 2003 International Conference “Physics and Control” Proceedings, Stefanatos et al. (2003), and in Physical Review A, Stefanatos et al. (2004).

3.1 Description of the System and Statement of the Problem

The system that we study in this chapter is an isolated pair of heteronuclear spins 1/2 (spins that belong to different nuclear species), denoted $I$ (for example $^1\text{H}$) and $S$ (for example $^{13}\text{C}$ or $^{15}\text{N}$), with indirect interaction between them (mediated by the surrounding electrons). The relaxation mechanisms taken into account here are dipole-dipole interaction and chemical shift anisotropy. We want to apply the theory presented in the preceding chapter and write down a master equation for this system. We start from the deterministic Hamiltonian $H_d(t)$. It is

$$H_d(t) = H_0 + H_{rf}(t),$$

where $H_0$ is the static Hamiltonian and $H_{rf}$ describes the effect of the applied radio-frequency (rf) magnetic field. More details about $H_{rf}$ will be given later. The static
Hamiltonian is also a sum of two terms

$$H_0 = H_Z + H_{ind} ,$$

(3.2)

where $H_Z$ is the Zeeman Hamiltonian for the spins $I$ and $S$

$$H_Z = \omega_I I_z + \omega_S S_z$$

(3.3)

and $H_{ind}$ is the Hamiltonian for the indirect interaction between them. The Larmor frequencies $\omega_I, \omega_S$ are such that to incorporate the effect of the isotropic part of chemical shift, see (2.108). We say now a few words about the origin of indirect interactions. It was mentioned above that the indirect interactions between spins are mediated through the electrons. Note that the great majority of NMR experiments is performed on nuclear spins of atoms belonging to diamagnetic molecules, in which the orbital and spin angular momenta of the individual electrons are coupled in such a way that their bulk angular momentum and magnetic moment both vanish. But the magnetic field produced by the magnetic moment of one spin, modifies the electronic ground state in such a way that the electronic system achieves a small magnetization proportional to that field. This electronic magnetization produces a small extra magnetic field that changes the magnetic field experienced by the other spin. The general form of $H_{ind}$ for two spins is

$$H_{ind} = \mathbf{I} \cdot \mathbf{J} \cdot \mathbf{S} ,$$

(3.4)

where $\mathbf{J}$ is a tensor. The only effective part of this interaction in liquids is its average over all relative orientations of the spins in space. In isotropic liquids it is of the form

$$H_{ind} = 2\pi J \mathbf{I} \cdot \mathbf{S} ,$$

(3.5)

where $J$ is the scalar coupling constant. In the weak coupling limit ($J \ll |\omega_I - \omega_S|$), the indirect interaction Hamiltonian is simplified to the form, Goldman (1988),

$$H_{ind} = 2\pi J I_z S_z .$$

(3.6)

For heteronuclear spins $I$ and $S$ the weak coupling condition is always satisfied. Thus, the static Hamiltonian for our system is

$$H_0 = \omega_I I_z + \omega_S S_z + 2\pi J I_z S_z .$$

(3.7)

This differs from $H_0$ in (2.99) by the indirect interaction term. So, instead of using the full interaction representation, we go to an intermediate one defined by the operator

$$U = \exp(iH_Zt) = \exp[i(\omega_I I_z + \omega_S S_z)t] .$$

(3.8)
Observe that this representation corresponds to a doubly rotating frame (a frame rotating with different frequency for each spin). If we follow the same steps as in the preceding chapter then, as long as $J^{-1} \gg \tau_c$ (this is indeed the case since the correlation time $\tau_c$ is of the order of ns while the inverse coupling constant $J^{-1}$ is of the order of ms, Cavanagh et al. (1995)), we end up with the master equation

$$\frac{d\tilde{\sigma}}{dt} = -i\pi J[2I_z S_z, \tilde{\sigma}] + L(\tilde{\sigma} - \tilde{\sigma}_{eq}),$$

(3.9)

where $\tilde{\sigma}$ is the density matrix in the doubly rotating frame, related to that in the laboratory frame by

$$\tilde{\sigma} = U\sigma U^\dagger.$$

(3.10)

$L$ is the relaxation superoperator corresponding to DD and CSA relaxation mechanisms. If we want to show explicitly the control Hamiltonian, then we should add in the coherent evolution of (3.9) the term

$$\tilde{H}_{rf} = \omega_x(t)I_x + \omega_y(t)I_y,$$

(3.11)

where $\omega_x, \omega_y$ are the components of the rf magnetic field (the available controls) in the rotating frame. But let us forget for a moment this Hamiltonian. The controls will be reintroduced by an indirect way in the following section.

In this chapter, we address the problem of finding the maximum efficiency for the transfers $\sigma_{\text{initial}} \rightarrow \sigma_{\text{target}}$

$$I_\alpha \rightarrow 2I_\beta S_\gamma,$$

(3.12)

and

$$I_\alpha \rightarrow S_\beta,$$

(3.13)

where $\alpha, \beta$ and $\gamma$ can be $x, y$ or $z$. These transfers are of central importance for two-dimensional NMR spectroscopy and are conventionally accomplished by the INEPT, Morris and Freeman (1979), and refocused INEPT, Burum and Ernst (1980), pulse sequence elements, respectively. We describe INEPT (Insensitive Nuclei Enhanced by Polarization Transfer) in the next section, where we also formulate our problem in terms of optimal control theory. Observe that we can concentrate on the transfers

$$I_z \rightarrow 2I_z S_z,$$

(3.14)

and

$$I_z \rightarrow S_z.$$

(3.15)
3.1 Description of the system and statement of the problem

\[ I_z = I_z S^\alpha + I_z S^\beta \]
\[ 2I_z S_z = I_z S^\alpha - I_z S^\beta \]

Figure 3.1: Initial state \( I_z \) (A) and final state \( 2I_z S_z \) (B), written as sums of single-transition operators, defined in the text. The transfer \( I_z \rightarrow 2I_z S_z \) can be accomplished by selectively inverting \( I_z S^\beta \) to \(-I_z S^\beta\). The polarizations of the spins in the ellipses show schematically which energy level is excess populated for each single transition operator, placed under the corresponding ellipse. For example, for \( I_z S^\alpha \) is \( \alpha\alpha \) (both \( I \) and \( S \) are up), while for \(-I_z S^\beta\) is \( \beta\beta \), see also the figure below.

Figure 3.2: Energy level diagram for a weakly coupled spin 1/2 pair. The selective population inversion of the levels \( \alpha\beta \) and \( \beta\beta \) corresponds to the transfer of polarization \( I_z \) (A) to longitudinal two-spin order \( 2I_z S_z \) (B). The dark circles represent excess population in the corresponding levels.
3.2 Formulation of the Problem in Terms of Optimal Control

Having found how to accomplish these optimally, we can use the same scheme for the transfers (3.12), (3.13). We just need to add the appropriate initial and final $90^\circ$ hard pulses to make the rotations $I_\alpha \rightarrow I_z$, $2I_zS_z \rightarrow 2I_\beta S_\gamma$, $S_z \rightarrow S_\beta$. These rotations can be done, in the ideal case, instantaneously, with no loss.

We give now a more physical picture for transfer (3.14). In the following, we use $\alpha$ and $\beta$ to denote the Zeeman eigenstates of a single spin 1/2 (spin up and spin down, respectively). The initial state $I_z$, representing polarization on spin $I$ and shown schematically in Figs. 3.1(A), 3.2(A), can be decomposed as a sum of two single-transition operators

$$I_zS^\alpha = \frac{I_z}{2} + I_zS_z, \quad I_zS^\beta = \frac{I_z}{2} - I_zS_z.$$  \hspace{1cm} (3.16)

$I_zS^\alpha$ represents a coherence between the levels $\alpha\alpha$ and $\beta\alpha$, with the first level excess populated and the second depleted, see Fig. 3.2(A). The frequency of the corresponding single transition (flip of spin $I$ only) is $\omega/I/2\pi + J/2$, or just $J/2$ in the rotating frame. $I_zS^\beta$ represents a coherence between the levels $\alpha\beta$ and $\beta\beta$, with the first excess populated and the second depleted, see Fig. 3.2(A). The frequency of the corresponding single transition is $-J/2$ in the rotating frame. Obviously, transfer (3.14) can be accomplished by selectively inverting $I_zS^\beta$ to $-I_zS^\beta$, see Figs. 3.1, 3.2. In the absence of relaxation this can be done by weak rf irradiation ($\omega_{rf} = \gamma I B_{rf} \ll 2\pi J$, so that the other transition is not affected) at the frequency $-J/2$ of the corresponding transition. In the presence of relaxation this procedure is not optimal, since during the slow rotation from $I_zS^\beta$ to $-I_zS^\beta$, part of the population is lost. Our goal is to find the optimal rf excitation that minimizes these relaxation losses.

A similar physical picture can be drawn for polarization transfer (3.15) from spin $I$ to spin $S$. This transfer is implemented in two steps. First we make the transfer $I_z \rightarrow 2I_zS_z$ and then a similar transfer $2I_zS_z \rightarrow S_z$. The later transfer can be accomplished by selectively inverting the populations of the levels $\beta\beta$ and $\beta\alpha$, see Fig. 3.2(B), so finally the excess populated levels are $\alpha\alpha$ and $\beta\alpha$, a configuration corresponding to polarization $S_z$ (note the similarity with the configuration shown in Fig. 3.2(A) which corresponds to $I_z$).

3.2 Formulation of the Problem in Terms of Optimal Control

Let the initial density operator be $\sigma(0)$ and let $\sigma(t)$ denote the density operator at time $t$. The maximum efficiency of transfer between $\sigma(0)$ and a target operator $Q$ is defined as the largest possible value of $\langle Q \rangle = \text{Tr}\{Q\sigma(t)\}$. For transfers (3.14) and (3.15) it is $\sigma(0) = \tilde{\sigma}(0)$ and $\text{Tr}\{Q\sigma(t)\} = \text{Tr}\{Q\tilde{\sigma}(t)\}$, so we can use $\tilde{\sigma}(t)$ instead of $\sigma(t)$ and thus the master equation
3.2 FORMULATION OF THE PROBLEM IN TERMS OF OPTIMAL CONTROL

(3.9). Here we concentrate on transfer (3.14). As we mentioned above, transfer (3.15) can be implemented in two steps, each of them corresponding to a transfer like (3.14). Using Eq. (3.9) and the results of the preceding chapter, we can derive evolution equations for the ensemble averages of the operators that participate in the building of \( \langle Q \rangle = \langle 2I_z S_z \rangle \). These equations can be found in Goldman (1984). For simplicity, we ignore terms that arise because of cross-correlation effects between the DD and CSA relaxation mechanisms. We find

\[
\begin{align*}
\frac{d}{dt} \langle I_z \rangle &= -\pi k_1 (\langle I_z \rangle - I_0) - \pi k_{cr} (\langle S_z \rangle - S_0), \\
\frac{d}{dt} \langle I_x \rangle &= -\pi k_{t1} (\langle I_x \rangle - I_0) - \pi J \langle 2I_y S_z \rangle, \\
\frac{d}{dt} \langle 2I_y S_z \rangle &= \pi J (\langle I_z \rangle - \pi k_{t2} (2I_y S_z), \\
\frac{d}{dt} \langle 2I_z S_z \rangle &= -\pi k_2 \langle 2I_z S_z \rangle,
\end{align*}
\]

where

\[
\begin{align*}
k_1 &= D \left\{ 6(1 + \alpha^2) J(\omega_I) + 2J(\omega_I - \omega_S) + 12J(\omega_I + \omega_S) \right\}, \\
k_{cr} &= D \left\{ 12J(\omega_I + \omega_S) - 2J(\omega_I - \omega_S) \right\}, \\
k_{t1} &= D \left\{ 4(1 + \alpha^2) J(0) + 3(1 + \alpha^2) J(\omega_I) + J(\omega_I - \omega_S) + 6J(\omega_S) + 6J(\omega_I + \omega_S) \right\}, \\
k_{t2} &= D \left\{ 4(1 + \alpha^2) J(0) + 3(1 + \alpha^2) J(\omega_I) + J(\omega_I - \omega_S) + 6J(\omega_I + \omega_S) \right\}, \\
k_2 &= D \left\{ 6(1 + \alpha^2) J(\omega_I) + 6J(\omega_S) \right\},
\end{align*}
\]

and

\[
\begin{align*}
D &= \frac{1}{4\pi} \frac{2\gamma_S^2 \gamma r^6}{\gamma_S^2}, \\
\alpha &= \frac{2B_0 (\sigma_\parallel - \sigma_\perp) \rho^3}{3\gamma_S \hbar}.
\end{align*}
\]

The spectral function \( J(\omega) \) is given in (2.76). It should not be confused with the coupling constant \( J \). \( I_0, S_0 \) are the thermal equilibrium values for \( \langle I_z \rangle, \langle S_z \rangle \), respectively.

In order to simplify things, we ignore the small cross-relaxation between spins \( I \) and \( S \), i.e. we set \( k_{cr} = 0 \) (this approximation is very good for large molecules where \( \tau_{c}^{-1} \ll |\omega_I - \omega_S| \)). We also take the two transverse relaxation rates \( k_{t1}, k_{t2} \) to be equal, i.e. we set \( k_{t1} = k_{t2} = k \). This is a very good approximation for systems where \( \tau_{c}^{-1} \ll \omega_S \), see Eqs. (3.23), (3.24) and (2.76). Note that the methods presented here can be extended to
§3.2 FORMULATION OF THE PROBLEM IN TERMS OF OPTIMAL CONTROL

account for the case \( k_{t1} \neq k_{t2} \). Finally, we set \( I_0 = 0 \). Let us justify this step. In general, polarization \( \langle I_z \rangle \) relaxes towards its thermal equilibrium polarization \( I_0 \). For example, this thermal correction must be taken into account in transfer steps that are applied in the preparation period of an experiment, Ernst et al. (1987). However, if coherences of some non-zero order are selected at a given point in the pulse sequence using phase cycles or magnetic field gradients, the thermal correction can be omitted from all subsequent pulse sequence elements, Levitt and Bari (1994). Here, we focus on this case, corresponding to all mixing steps in multidimensional NMR experiments. These steps are always applied after an evolution period in which a non-zero coherence order is selected. Hence, we can set \( I_0 = 0 \). With these simplifications, the evolution equations for the ensemble averages take the form

\[
\begin{align*}
\frac{d}{dt} \langle I_z \rangle &= -\pi k_1 \langle I_z \rangle, \\
\frac{d}{dt} \langle I_x \rangle &= \pi k \langle I_x \rangle - \pi J \langle 2I_y S_z \rangle, \\
\frac{d}{dt} \langle 2I_y S_z \rangle &= \pi J \langle I_x \rangle - \pi k \langle 2I_y S_z \rangle, \\
\frac{d}{dt} \langle 2I_z S_z \rangle &= -\pi k_2 \langle 2I_z S_z \rangle.
\end{align*}
\] (3.28) (3.29) (3.30) (3.31)

Now we can describe how the transfer (3.14) takes place. We start from operator \( I_z \), which relaxes with rate \( k_1 \) (longitudinal relaxation). The two heteronuclear spins have well separated frequencies, allowing for fast selective manipulation of each spin on a time-scale determined by the coupling \( J \) and the relaxation rates \( k \), and \( k_1 \) or \( k_2 \). Thus, using the magnetic field in \( y \)-direction, we can rotate \( I_z \) to \( I_x \), see Fig. 3.3. Operator \( I_x \) evolves under the \( J \) coupling to \( 2I_y S_z \) and also relaxes with rate \( k \) (transverse relaxation). As the operator \( 2I_y S_z \) is produced, it also relaxes with rate \( k \). Finally, using the magnetic field in \( x \)-direction, we can rotate \( 2I_y S_z \) to \( 2I_z S_z \), see Fig. 3.3. Operator \( 2I_z S_z \) relaxes with rate \( k_2 \) (longitudinal relaxation).

Define

\[
\begin{align*}
\alpha_1(t) &= \sqrt{\langle I_x \rangle^2 + \langle I_z \rangle^2}, \\
\alpha_2(t) &= \sqrt{\langle 2I_y S_z \rangle^2 + \langle 2I_z S_z \rangle^2}.
\end{align*}
\] (3.32)

Using (3.28)-(3.31), evolution equations for \( \alpha_1 \), \( \alpha_2 \) can be found. We get the system

\[
\frac{d}{dt} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \pi J \begin{bmatrix} -k-k_1 J \cos^2 \beta_1 - k_1 J \cos \beta_1 \cos \beta_2 \\ k_2 J \cos \beta_1 \cos \beta_2 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix},
\] (3.33)

where \( \cos \beta_1 = \langle I_x \rangle / \langle I_z \rangle \), \( \cos \beta_2 = \langle 2I_y S_z \rangle / \langle 2I_z S_z \rangle \), see Fig. 3.3. Using the magnetic field in \( y \)-direction, we can control the angle \( \beta_1 \). Analogously, using the magnetic field in
§3.2 FORMULATION OF THE PROBLEM IN TERMS OF OPTIMAL CONTROL

3.2 formulation of the problem in terms of optimal control

\[ \langle I_x \rangle, \langle I_z \rangle, \langle 2I_y S_z \rangle, \langle 2I_z S_z \rangle \]

\[ J, r_1, r_2, \beta_1, \beta_2 \]

\[ 0.1, 0.2, 0.3, 0.4 \]

Figure 3.3: Representation of the system variables \( r_1, r_2 \) and the angles \( \beta_1, \beta_2 \) in terms of the expectation values \( \langle I_x \rangle, \langle I_z \rangle, \langle 2I_y S_z \rangle, \langle 2I_z S_z \rangle \). The new control parameters are \( u_1 = \cos \beta_1, u_2 = \cos \beta_2 \).

\( x \)-direction, we can control the angle \( \beta_2 \). Let

\[ \xi_1 = \frac{k - k_1}{J}, \quad \bar{\xi}_1 = \frac{k_1}{J}, \quad \xi_2 = \frac{k - k_2}{J}, \quad \bar{\xi}_2 = \frac{k_2}{J}. \quad (3.34) \]

Denoting \( u_1 = \cos \beta_1, u_2 = \cos \beta_2 \) and rescaling time according to \( t_{new} = \pi J t_{old} \), we find the following equation for \( r_1, r_2 \) (by abuse of notation we use the same \( t \) for the new time)

\[ \frac{d}{dt} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} -\xi_1 u_1^2 - \bar{\xi}_1 & -u_1 u_2 \\ u_1 u_2 & -\xi_2 u_2^2 - \bar{\xi}_2 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}. \quad (3.35) \]

The initial conditions are

\[ r_1(0) = 1, \quad r_2(0) = 0. \quad (3.36) \]

Note that starting from the ensemble average equations, we brought the system equations in a form where the principles of optimal control can be applied. The problem that we face is: Given the dynamical system above, how should \( u_1(t), u_2(t) \), with \( 0 \leq u_1, u_2 \leq 1 \), be chosen so that starting from \( r_1(0) = 1 \) we achieve the largest value \( r_2(T) \) for a specified finite time \( T \). In spectroscopic applications this would correspond to the maximum efficiency for the transfer of \( I_z \) to \( 2I_z S_z \), because once \( r_2 \) is maximized we can apply a hard pulse in the \( x \)-direction and transfer it to \( 2I_z S_z \) with no loss. Transfer (3.14) is conventionally accomplished by the INEPT pulse sequence: At time \( t = 0^- \) we apply a 90° hard pulse which rotates \( I_z \) to \( I_x \). Then, we let this operator evolve under the \( J \) coupling towards \( 2I_y S_z \) for the whole time interval \([0, T]\), i.e. we keep \( \beta_1 = \beta_2 = 0 \) (\( u_1(t) = u_2(t) = 1 \))
during this interval. Finally, by applying a $90^\circ$ hard pulse at $t = T^+$, we rotate the operator $2I_yS_z$ to $2I_zS_z$. Schematically, the INEPT pulse sequence is

$$I_z \rightarrow I_x \rightarrow 2I_yS_z \rightarrow 2I_zS_z,$$

where the solid arrow represents hard pulses, while the dashed arrow represents the evolution under the $J$ coupling. In the subsequent sections we show that in the presence of relaxation this is not the optimal procedure. Having found the optimal controls $u_1(t)$, $u_2(t)$, we can calculate the corresponding magnetic field components $\omega_x(t)$, $\omega_y(t)$, which achieve the maximum efficiency.

### 3.3 Solution of the Optimal Control Problem

The solution of the optimal control problem depends on the relative magnitude of the parameters $\xi_1, \xi_2$. In the following, we solve the problem for the nonsymmetric case $\xi_1 > \xi_2$, i.e $k_1 < k_2$. The results for the case $\xi_1 < \xi_2$ ($k_1 > k_2$) are analogous. At the end of this section we describe also what happens for the symmetric case $\xi_1 = \xi_2$.

To find the optimal controls $(u_1^*(t), u_2^*(t))$ in Eq. (3.35), we use the principle of dynamic programming, Bryson and Ho (1975), and solve for the maximum achievable value of $r_2$ for all initial points $(r_1, r_2)$. Starting from $(r_1, r_2)$, we denote the maximum achievable value of $r_2$ by $V(r_1, r_2, t)$, also called the optimal return function for the point $(r_1, r_2)$ at time $t$. Note here that for the finite time problem ($T$ finite), the optimal return function has explicit dependence on time, Bryson and Ho (1975). If we start at $(r_1, r_2)$, then by making a choice of controls in (3.35) and letting the dynamical system evolve, after small time $\delta t$ we can make a transition to all points $(\tilde{r}_1, \tilde{r}_2)$ which are related to $(r_1, r_2)$ by

$$\begin{bmatrix} \tilde{r}_1 \\ \tilde{r}_2 \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} + \delta t \begin{bmatrix} -\xi_1 u_1^2 - \xi_1 & -u_1 u_2 \\ u_1 u_2 & -\xi_2 u_2^2 - \xi_2 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}.$$  

(3.38)

From all points $(\tilde{r}_1, \tilde{r}_2)$ that can be reached by appropriate choice of $(u_1, u_2)$ in small time $\delta t$, we should choose to go to that $(\tilde{r}_1, \tilde{r}_2)$ for which $V(\tilde{r}_1, \tilde{r}_2, t + \delta t)$ is the largest. But now note by definition of $V$ that $V(r_1, r_2, t) = \max_{\tilde{r}_1, \tilde{r}_2} V(\tilde{r}_1, \tilde{r}_2, t + \delta t)$. This can be rewritten as

$$V(r_1, r_2, t) = \max_{u_1, u_2} V\{r_1 + \delta t[(-\xi_1 u_1^2 - \xi_1) r_1 - u_1 u_2 r_2], r_2 + \delta t[(-\xi_2 u_2^2 - \xi_2) r_2 + u_1 u_2 r_1], t + \delta t\}.$$  

(3.39)
for infinitesimal $\delta t$. The right-hand side of the above expression can be expanded (Taylor series expansion) in powers of $\delta t$ and retaining only the terms linear in $\delta t$ (for $\delta t$ approaching zero), we obtain the well known Hamilton Jacobi Bellman equation

$$
\frac{\partial V}{\partial t} + \max_{u_1, u_2} H(u_1, u_2) = 0 ,
$$

where

$$
H(u_1, u_2) = \left[ \frac{\partial V}{\partial r_1} \frac{\partial V}{\partial r_2} \right] \begin{bmatrix} -\xi_1 u_1^2 - \tilde{\xi}_1 - u_1 u_2 \\ u_1 u_2 - \xi_2 u_2^2 - \tilde{\xi}_2 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}
$$

is the Hamiltonian for the optimal control problem. Let

$$
\lambda_1 = \frac{\partial V}{\partial r_1}, \quad \lambda_2 = \frac{\partial V}{\partial r_2}.
$$

If, additionally, we set

$$
a = \frac{\lambda_2}{\lambda_1}, \quad b = \frac{r_2}{r_1},
$$

then the Hamiltonian can be expressed as

$$
H(u_1, u_2) = -\lambda_1 r_1 [\xi_1 u_1^2 - (a - b) u_1 u_2 + ab \xi_2 u_2^2 + \tilde{\xi}_1 + ab \tilde{\xi}_2].
$$

The optimal return function is a non-decreasing function of $r_1, r_2$ (starting from a larger $r_1(0)$ or $r_2(0)$ you can achieve a larger $r_2(T)$) so $\lambda_1, \lambda_2, a \geq 0$. Since $\lambda_1, r_1 \geq 0$, in order to maximize $H$ in the square $0 \leq u_1, u_2 \leq 1$ it is equivalent to minimize the function

$$
F(u_1, u_2) = \xi_1 u_1^2 - (a - b) u_1 u_2 + ab \xi_2 u_2^2 + \tilde{\xi}_1 + ab \tilde{\xi}_2
$$

or the quadratic form

$$
Q(u_1, u_2) = \xi_1 u_1^2 - (a - b) u_1 u_2 + ab \xi_2 u_2^2,
$$

which is the part of $F(u_1, u_2)$ that contains explicitly the controls $u_1, u_2$. Note that $a, b, \xi_1, \xi_2, u_1, u_2 \geq 0$. If $(a - b) \leq 0$ then the solution to the minimization of $Q(u_1, u_2)$ is the trivial one $u_1 = u_2 = 0$. Therefore $(a - b) > 0$. Now suppose that $(a - b)^2 \leq 4 \xi_1 \xi_2 ab$, i.e. $(a - b)^2 / 4 \xi_1 \xi_2 = ab - \epsilon$ for $\epsilon \geq 0$. Then

$$
Q = \left[ \frac{(a - b)}{2 \sqrt{\xi_1}} u_2 - \xi_1 u_1 \right]^2 + \epsilon \xi_2 u_2^2
$$

and the solution to the minimization of $Q$ is again the trivial one $u_1 = u_2 = 0$. Therefore the acceptable case is $(a - b)^2 > 4 \xi_1 \xi_2 ab$. Using the conditions

$$
a - b > 0, \quad (a - b)^2 > 4 \xi_1 \xi_2 ab,
$$

\section*{Solution of the Optimal Control Problem}
that we just derived, we minimize \( Q \) in the square \( 0 \leq u_1, u_2 \leq 1 \). We find three separate cases (actually, there is one more case but, since it is not important for the rest of the problem, we do not present it)

1. **Case I:** If \( a - b < 2\xi_1 \) and \( (a-b)/ab \geq 2\xi_2 \), then the minimum of \( Q \) (maximum of \( H \)) is obtained for \( u_2 = 1 \) and \( u_1 = (a-b)/2\xi_1 \).

2. **Case II:** If \( a - b \geq 2\xi_1 \) and \( (a-b)/ab \geq 2\xi_2 \), then the minimum of \( Q \) is obtained for \( u_1 = 1 \) and \( u_2 = 1 \).

3. **Case III:** If \( a - b \geq 2\xi_1 \) and \( (a-b)/ab < 2\xi_2 \), then the minimum of \( Q \) is obtained for \( u_1 = 1 \) and \( u_2 = (a-b)/2\xi_2ab \).

It is a standard result, Bryson and Ho (1975) that, along the optimal trajectory \((r_1^*(t), r_2^*(t))\), the adjoint variables

\[
(\lambda_1(t), \lambda_2(t)) = \left( \frac{\partial V}{\partial r_1}, \frac{\partial V}{\partial r_2} \right) \bigg|_{(r_1^*(t), r_2^*(t))}
\]

satisfy the equations \( \dot{\lambda}_1 = -\partial H/\partial r_1 \) and \( \dot{\lambda}_2 = -\partial H/\partial r_2 \), i.e.

\[
\frac{d}{dt} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} \xi_1 u_1^2 + \xi_1 & -u_1 u_2 \\ u_1 u_2 & \xi_2 u_2^2 + \xi_2 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix},
\] (3.49)

with the terminal conditions

\[
\lambda_1(T) = 0 \quad \lambda_2(T) = 1.
\] (3.50)

For the optimal trajectory starting at \((r_1, r_2) = (1, 0)\), we have \( b(0) = 0 \), so depending on \( a(0) \) we have the following cases

**Case A:** If \( a(0) \geq 2\xi_1 \) (first condition) then we start in case II discussed above. We stay there for the whole interval \( T \) if (note that \( a-b \) is increasing) \([a(T) - b(T)]/a(T)b(T) \geq 2\xi_2 \) (second condition). In case II, \( u_1 = u_2 = 1 \) and using (3.35), (3.49) we can find \( a(T), b(T) \) in terms of \( a(0) \) and \( b(0) = 0 \). It is \( a(T) = [a(0) + \tan T]/[1 - a(0) \tan T], b(T) = \tan T \) so

\[
\frac{a(T) - b(T)}{a(T)b(T)} = \frac{a(0)(1 + \tan^2 T)}{\tan T[a(0) + \tan T]}.
\] (3.51)

Now note that \( a(T) = \lambda_2(T)/\lambda_1(T) = \infty \) since \( \lambda_1(T) = 0 \), so must be \( a(0) = 1/\tan T \) if we spend the whole interval in case II. For \( \tan T \leq 1/2\xi_1 \) it is \( a(0) \geq 2\xi_1 \), so we are consistent with the first condition, and \([a(T) - b(T)]/a(T)b(T) = 1/\tan T \geq 2\xi_1 > 2\xi_2 \), so we are also consistent with the second condition. The conclusion is that for \( \tan T \leq 1/2\xi_1 \) we start in
case II and stay there for the whole time interval $T$. This corresponds to the INEPT pulse sequence.

**Case B:** If $a(0) < 2\xi_1$ we start in case I. In this case it is $u_2 = 1$, $u_1 = (a - b)/2\xi_1$ and the equation for the state variables becomes

$$
\frac{d}{dt} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} -\xi_1 u_1^2 - \bar{\xi}_1 & -u_1 \\ u_1 & -\xi_1 - \bar{\xi}_1 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}.
$$

(3.52)

But $\xi_1 + \bar{\xi}_1 = \xi_2 + \bar{\xi}_2 = k/J$ so the above equation is the same as

$$
\frac{d}{dt} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} -\xi_1 u_1^2 - \bar{\xi}_1 & -u_1 \\ u_1 & -\xi_1 - \bar{\xi}_1 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}.
$$

(3.53)

If we make the transformation $\bar{r}_i = e^{\xi_1 t} r_i$ we get

$$
\frac{d}{dt} \begin{bmatrix} \bar{r}_1 \\ \bar{r}_2 \end{bmatrix} = \begin{bmatrix} -\xi_1 u_1^2 - \bar{\xi}_1 & -u_1 \\ u_1 & -\xi_1 - \bar{\xi}_1 \end{bmatrix} \begin{bmatrix} \bar{r}_1 \\ \bar{r}_2 \end{bmatrix}.
$$

(3.54)

Analogously, for the adjoint variables we find the equation

$$
\frac{d}{dt} \begin{bmatrix} \bar{\lambda}_1 \\ \bar{\lambda}_2 \end{bmatrix} = \begin{bmatrix} \xi_1 u_1^2 - u_1 \\ u_1 & \xi_1 - \bar{\xi}_1 \end{bmatrix} \begin{bmatrix} \bar{\lambda}_1 \\ \bar{\lambda}_2 \end{bmatrix},
$$

(3.55)

where $\bar{\lambda}_i = e^{-\xi_1 t} \lambda_i$. Now observe that $\bar{b}/\bar{a} = \bar{r}_2 \bar{\lambda}_1 / \bar{r}_1 \bar{\lambda}_2 = b/a$. If we set $\bar{b}/\bar{a} = b/a = \kappa_1(t)$ then $\kappa_1(0) = 0$ and $\kappa_1(t)$ satisfies the differential equation

$$
\frac{d\kappa_1}{dt} = \frac{(\kappa_1 - 1)^2}{2\xi_1} - 2\xi_1 \kappa_1.
$$

(3.56)

This can be proved using (3.54), (3.55). So $(b/a)(t)$ satisfies (3.56) for case I. After time $\tau_1, (a - b)/2\xi_1$ becomes equal to 1 and the system switches to case II. Using $[a(\tau_1) - b(\tau_1)]/2\xi_1 = 1$ and that $(b/a)(\tau_1) = \kappa_1(\tau_1)$ we find

$$
a(\tau_1) = \frac{\lambda_2(\tau_1)}{\lambda_1(\tau_1)} = \frac{2\xi_1}{1 - \kappa_1(\tau_1)} \equiv \tan \varphi_1,
$$

(3.57)

$$
b(\tau_1) = \frac{r_2(\tau_1)}{r_1(\tau_1)} = \frac{2\xi_1 \kappa_1(\tau_1)}{1 - \kappa_1(\tau_1)} \equiv \tan \varphi_1.
$$

(3.58)

As time goes by $(a - b)/ab$ decreases. If this fraction does not reach the value $2\xi_2$ in the remaining time interval $[\tau_1, T]$ then the system remains in case II and we call this **Case B1**. Else, the system switches to case III and we call this **Case B2**. We examine first the full scenario (Case B2). Suppose that $(a - b)/ab = 2\xi_2$ at time $T - \tau_2$. This is the switching time
from case II to case III. If we work as before we find that in case III, \((b/a)(t) = \kappa_2(T - t)\) where \(\kappa_2(t)\) satisfies the differential equation

\[
\frac{d\kappa_2}{dt} = \frac{(\kappa_2 - 1)^2}{2\xi_2} - 2\xi_2\kappa_2, \quad \kappa_2(0) = 0. \tag{3.59}
\]

For \(t = T - \tau_2\) it is \(b(T - \tau_2)/a(T - \tau_2) = \kappa_2(\tau_2)\) and \([a(T - \tau_2) - b(T - \tau_2)]/a(T - \tau_2)b(T - \tau_2) = 2\xi_2\). Using these we find

\[
a(T - \tau_2) = \frac{\lambda_2(T - \tau_2)}{\lambda_1(T - \tau_2)} = \frac{1 - \kappa_2(\tau_2)}{2\xi_2\kappa_2(\tau_2)} \equiv \tan \varphi, \tag{3.60}
\]

\[
b(T - \tau_2) = \frac{\nu_2(T - \tau_2)}{\nu_1(T - \tau_2)} = \frac{1 - \kappa_2(\tau_2)}{2\xi_2} \equiv \tan \psi. \tag{3.61}
\]

In order to find the switching times \(\tau_1, T - \tau_2\) we have to connect (3.57), (3.58) with (3.60), (3.61). We can do so by examining the evolution of the system while it is in case II, i.e. in the time interval \([\tau_1, T - \tau_2]\). In case II it is \(u_1 = u_2 = 1\) so the system equation (3.35) becomes

\[
\begin{align*}
\frac{d}{dt} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} &= \begin{bmatrix} -\xi & -1 \\ 1 & -\xi \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}, \tag{3.62}
\end{align*}
\]

where \(\xi = \xi_1 + \xi_2 = \xi_1 + \xi_2\). This is the same as the system equation for a damped harmonic oscillator with angular frequency \(\omega = 1\). From \(t = \tau_1\) to \(t = T - \tau_2\) the vector \((r_1, r_2)\) is rotated through the angle

\[
\psi_2 - \psi_1 = T - \tau_1 - \tau_2 = \tan^{-1} \frac{1 - \kappa_2(\tau_2)}{2\xi_2} - \tan^{-1} \frac{2\xi_1\kappa_1(\tau_1)}{1 - \kappa_1(\tau_1)} \tag{3.63}
\]

The evolution equation for the adjoint variables in case II is

\[
\frac{d}{dt} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} \xi & -1 \\ 1 & \xi \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}. \tag{3.64}
\]

The vector \((\lambda_1, \lambda_2)\) is rotated through the angle

\[
\varphi_2 - \varphi_1 = T - \tau_1 - \tau_2 = \tan^{-1} \frac{1 - \kappa_2(\tau_2)}{2\xi_2\kappa_2(\tau_2)} - \tan^{-1} \frac{2\xi_1}{1 - \kappa_1(\tau_1)} \tag{3.65}
\]

Equations (3.63), (3.65) constitute a system of two equations for the two unknowns \(\tau_1, \tau_2\) which can be solved numerically. Fig. 3.4 depicts the angles \(\theta_1, \theta_2\) and \(\varphi_1, \varphi_2\), the optimal trajectory \((r_1(t), r_2(t))\) and the trajectory for the adjoint variables \((\lambda_1(t), \lambda_2(t))\). For case B1 there is only one switching time \((\tau_1)\) and working analogously we can easily find that it satisfies the equation

\[
T - \tau_1 = \varphi_2 - \varphi_1 = \frac{\pi}{2} - \tan^{-1} \frac{2\xi_1}{1 - \kappa_1(\tau_1)}. \tag{3.66}
\]
As we showed above, case B happens when the duration of the experiment satisfies $T > T_A = \tan^{-1}(1/2\xi_1)$. The question is when we are in case B1 and when in case B2. Intuitively, we expect that up to some duration $T_B$ we have to apply the controls of case B1 and for $T > T_B$, where the phenomenon of relaxation degrades more the performance, we need to apply the more fancy controls of case B2. $T_B$ can be obtained from Eqs. (3.63) and (3.65) by substituting $\tau_2 = 0$. So, we have a system of two equations with unknowns the duration $T_B$ and the switching time $\tau_1$ and thus we can find $T_B$. For $T < T_B$ the system of Eqs. (3.63) and (3.65) gives $\tau_2 < 0$, while for $T > T_B$ gives $\tau_2 > 0$. Note that all the times that appear in this section are normalized according to $t_{\text{new}} = \pi J t_{\text{old}}$.

The results for $\xi_1 < \xi_2$ ($k_1 > k_2$) are analogous. The basic difference is for the case B1 (that with only one switching time). Here, we start in phase II ($u_1 = u_2 = 1$) and at time $t = T - \tau_2$ we switch to case III. The time $\tau_2$ is calculated by

$$T - \tau_2 = \vartheta_2 - \vartheta_1 = \tan^{-1} \frac{1 - \kappa_2(\tau_2)}{2\xi_2}.$$  

For case B2 (that with the two switching times) we have three phases, as before, but now we spend more time in phase III than in phase I. The times $\tau_1, \tau_2$ are still calculated using (3.63), (3.65). The duration $T_A$ is given now by $T_A = \tan^{-1}(1/2\xi_2)$.

Finally, we examine what happens for $\xi_1 = \xi_2 \equiv \zeta$. In this case it is $\kappa_1(t) = \kappa_2(t) \equiv \kappa(t)$, and the equations (3.63), (3.65) take the form

$$T - \tau_1 - \tau_2 = \tan^{-1} \frac{1 - \kappa(\tau_2)}{2\zeta} - \tan^{-1} \frac{2\zeta \kappa(\tau_1)}{1 - \kappa(\tau_1)}.$$  

(3.68)
\[ T - \tau_1 - \tau_2 = \tan^{-1} \frac{1 - \kappa(\tau_2)}{2\zeta(\tau_2)} - \tan^{-1} \frac{2\zeta}{1 - \kappa(\tau_1)} \]  

(3.69)

If we set
\[ f(\tau_1, \tau_2) = \tan^{-1} \frac{1 - \kappa(\tau_2)}{2\zeta} - \tan^{-1} \frac{2\zeta\kappa(\tau_2)}{1 - \kappa(\tau_1)}, \]  

(3.70)

it is not hard to see that
\[ f(\tau_2, \tau_1) = \tan^{-1} \frac{1 - \kappa(\tau_2)}{2\zeta(\tau_2)} - \tan^{-1} \frac{2\zeta}{1 - \kappa(\tau_1)}, \]  

(3.71)

thus we end up with the system
\[ T - \tau_1 - \tau_2 = f(\tau_1, \tau_2), \]  

(3.72)

\[ T - \tau_1 - \tau_2 = f(\tau_2, \tau_1). \]  

(3.73)

These equations imply
\[ f(\tau_1, \tau_2) = f(\tau_2, \tau_1). \]  

(3.74)

This gives
\[ \tau_1 = \tau_2 \equiv \tau, \]  

(3.75)

so for the symmetric case $\xi_1 = \xi_2$ the optimal pulse sequence spends equal time $\tau$ in phases I and III. Note that for $\xi_1 = \xi_2$ there is no intermediate case (B1, with only one switching time). Here, the optimal pulse sequence is either the INEPT or that with the three phases (and the two switching times). The time $\tau$ can be calculated by using either of (3.68), (3.69) with $\tau_1 = \tau_2 = \tau$.

### 3.4 Calculation of the Maximum Transfer Efficiency

Here we derive the value of $\eta_T = r_2(T)$ for each of the cases presented in section 3.3.

**Case A (INEPT):** For $u_1 = u_2 = 1$ the system equation is (3.62). With initial conditions $r_1(0) = 1, r_2(0) = 0$ we find
\[ \eta_T = r_2(T) = e^{-\xi T} \sin T. \]  

(3.76)

**Case B2:** Using equations (3.35) and (3.49) we can show that
\[ V(t) = \lambda_1(t)r_1(t) + \lambda_2(t)r_2(t), \quad 0 \leq t \leq T \]  

(3.77)

is a constant along the system trajectories and equals the optimal return function $V(t) = \lambda_1(0) = r_2(T)$. Analogously we can show that
\[ V_1(t) = e^{-\xi_1 t} \sqrt{\lambda_1^2(t) + \kappa_1(t)\lambda_2^2(t)}, \quad 0 \leq t \leq \tau_1 \]  

(3.78)
§3.4 Calculation of the Maximum Transfer Efficiency

and

\[ V_2(t) = e^{-\xi_2(T-t)} \sqrt{r_2^2(t) + \kappa_2(T-t)r_1^2(t)} , \quad T - \tau_2 \leq t \leq T , \]  

(3.79)

are constants along the system trajectories with \( V_1(t) = \lambda_1(0) \), \( V_2(t) = r_2(T) \). So they are also equal to the optimal return function. Define

\[ R(t) = \sqrt{r_1^2(t) + r_2^2(t)} , \quad \Lambda(t) = \sqrt{\lambda_1^2(t) + \lambda_2^2(t)} . \]  

(3.80)

Since \( r_2/r_1 = \tan \vartheta \) and \( \lambda_2/\lambda_1 = \tan \varphi \) it is also \( r_2 = R \sin \vartheta, r_1 = R \cos \vartheta \) and \( \lambda_2 = \Lambda \sin \varphi, \lambda_1 = \Lambda \cos \varphi \). Using these relations and Eq. (3.78) we find

\[ V_1(\tau_1) = e^{-\xi_1 \tau_1} \Lambda_1 \sqrt{1 - \xi_1 \sin 2 \varphi_1} , \]  

(3.81)

where \( \Lambda_1 = \Lambda_1(\tau_1) \). Using (3.77) we find

\[ V(\tau_1) = \Lambda_1 R_1 \cos(\varphi_1 - \vartheta_1) , \]  

(3.82)

where \( R_1 = R(\tau_1) \). Since \( V(\tau_1) = V_1(\tau_1) \) from (3.81), (3.82) we easily deduce that

\[ R_1 = e^{-\xi_1 \tau_1} \sqrt{1 - \xi_1 \sin 2 \varphi_1} \cos(\varphi_1 - \vartheta_1) . \]  

(3.83)

From (3.79) we find

\[ V_2(T - \tau_2) = e^{-\xi_2 \tau_2} R_2 \sqrt{1 - \xi_2 \sin 2 \vartheta_2} . \]  

(3.84)

But \( R_2 = R_1 e^{-\xi(T - \tau_1 - \tau_2)} \) (case II, damped harmonic oscillator). Using the above relations we find finally

\[ \eta_T = e^{-[\xi_1 \tau_1 + \xi(T - \tau_1 - \tau_2) + \xi_2 \tau_2]} \times \frac{\sqrt{1 - \xi_1 \sin 2 \varphi_1 \sqrt{1 - \xi_2 \sin 2 \vartheta_2}}}{\cos(\varphi_1 - \vartheta_1)} . \]  

(3.85)

**Case B1:** Following a procedure analogous to the above we can find that the maximum transfer efficiency in this case is given by

\[ \eta_T = e^{-[\xi_1 \tau_1 + \xi(T - \tau_1)]} \sqrt{1 - \xi_1 \sin 2 \varphi_1} . \]  

(3.86)

This is the result for \( k_1 < k_2 \). For \( k_1 > k_2 \) the result is analogous

\[ \eta_T = e^{-[k_2 \tau_2 + k(T - \tau_2)]} \sqrt{1 - \xi_2 \sin 2 \vartheta_2} . \]  

(3.87)
3.5 Calculation of Optimal Controls and of the Corresponding RF Field

Case I ($0 \leq t \leq \tau_1$): The optimal controls are

$$u_1 = \frac{a - b}{2\xi_1} = \frac{1}{2\xi_1} \lambda_2(t) [1 - \kappa_1(t)], \quad u_2 = 1. \quad (3.88)$$

Function $\kappa_1(t)$ is the solution of Eq. (3.56) with initial condition $\kappa_1(0) = 0$. We can easily find that

$$\kappa_1(t) = 1 + 2\xi_1^2 - 2\xi_1 \sqrt{1 + \xi_1^2} \coth \bar{\phi}_1(t), \quad (3.89)$$

where

$$\bar{\phi}_1(t) = t \sqrt{1 + \xi_1^2} + 2 \sinh^{-1} \xi_1, \quad (3.90)$$

for $t$ normalized or

$$\bar{\phi}_1(t) = \pi J t \sqrt{1 + \xi_1^2} + 2 \sinh^{-1} \xi_1, \quad (3.91)$$

for $t$ not normalized. We still need to calculate $\lambda_1(t), \lambda_2(t)$. Using Eq. (3.49) for the adjoint variables, we can see that $\lambda_2(t)$ satisfies the equation

$$\dot{\lambda}_2(t) = \left[\xi_1 + \sqrt{1 + \xi_1^2} \coth \bar{\phi}_1(t)\right] \lambda_2 \quad (3.92)$$

or

$$\dot{\bar{\lambda}}_2(t) = \sqrt{1 + \xi_1^2} \coth \bar{\phi}_1(t) \bar{\lambda}_2, \quad (3.93)$$

where as usual $\bar{\lambda}_2 = e^{-\xi_1 t} \lambda_2$. We can solve Eq. (3.93) in the interval $0 \leq t \leq \tau_1$ with final condition $\bar{\lambda}_2(\tau_1) = \bar{\Lambda}_2$ (we calculate $\bar{\lambda}_2$ and $\bar{\lambda}_1 = \bar{\lambda}_1(\tau_1)$ later). The solution is

$$\lambda_2(t) = e^{\xi_1 t} [\xi_1 + \sqrt{1 + \xi_1^2} \coth \bar{\phi}_1(t)] \lambda_2, \quad (3.94)$$

Now using the expression $\eta_T = e^{-\xi_1 t} \sqrt{\Lambda_1^2(t) + \kappa_1(t) \lambda_2^2(t)}$ for the optimal return function we have

$$\lambda_1(t) = \sqrt{e^{2\xi_1 t} \eta_T^2 - \kappa_1(t) \lambda_2^2(t)}, \quad (3.95)$$

Using again the expression for $\eta_T$ but with $t = \tau_1$ we find

$$\eta_T = e^{-\xi_1 \tau_1} \sqrt{\Lambda_1^2 + \kappa_1(\tau_1) \lambda_2^2} = \sqrt{\Lambda_1^2 + \kappa_1(\tau_1) \lambda_2^2}, \quad (3.96)$$

where $\Lambda_1 = \lambda_1(\tau_1), \Lambda_2 = \lambda_2(\tau_1)$ and $\bar{\lambda}_1, \bar{\lambda}_2$ as above. Combining Eqs. (3.94)-(3.96) we get

$$\lambda_1(t) = e^{\xi_1 t} \sqrt{\Lambda_1^2 + \kappa_1(\tau_1) \lambda_2^2} \frac{\sinh^2 \bar{\phi}_1(t)}{\sinh^2 \bar{\phi}_1(\tau_1)}, \quad (3.97)$$
Using Eqs. (3.94) and (3.97) we find
\[ u_1(t) = \sqrt{\frac{\bar{\Lambda}_1^2 [1 + \cosh \phi_1(t)]}{2\Lambda_1^2 \sinh^2 \phi_1(\frac{t}{2}) + \bar{\Lambda}_2^2 \cosh \phi_1(\tau_1) - \bar{\Lambda}_2^2 \cosh \phi_1(t)}} , \quad (3.98) \]
where
\[ \phi_1(t) = 2\pi J t \sqrt{1 + \xi_1^2} + 2 \sinh^{-1}(\xi_1) . \quad (3.99) \]
In order to find \( \bar{\Lambda}_1, \bar{\Lambda}_2 \) we use Eq. (3.96) and that \( \bar{\Lambda}_2/\bar{\Lambda}_1 = \Lambda_2/\Lambda_1 = \tan \phi_1 \). We get
\[ \bar{\Lambda}_1 = \eta_T \sqrt{1 + \kappa_1(\tau_1)} \tan^2 \varphi_1 , \quad \bar{\Lambda}_2 = \frac{\eta_T}{\sqrt{\kappa_1(\tau_1) + \tan^2 \varphi_1}} , \quad (3.100) \]
where \( \tan \varphi_1 \) is given by Eq. (3.57).

Case II \( (\tau_1 \leq t \leq \tau_2) \): The optimal controls are \( u_1(t) = u_2(t) = 1 \).

Case III \( (\tau_2 \leq t \leq T) \): The optimal controls are
\[ u_1 = 1 , \quad u_2 = \frac{a - b}{2\xi_2 ab} = \frac{1}{2\xi_2 r_2(t)} [1 - \kappa_2(T - t)] . \quad (3.101) \]
Following a procedure analogous to that for case I we finally find
\[ u_2(t) = \sqrt{\frac{R_1^2 [1 + \cosh \phi_2(T - t)]}{2R_2^2 \sinh^2 \phi_2(\frac{t}{2}) + R_1^2 \cosh \phi_2(\tau_2) - R_2^2 \cosh \phi_2(T - t)}} , \quad (3.102) \]
where
\[ \phi_2(t) = 2\pi J t \sqrt{1 + \xi_2^2} + 2 \sinh^{-1}(\xi_2) , \quad (3.103) \]
\[ R_1 = \frac{\eta_T}{\sqrt{\kappa_2(\tau_2) + \tan^2 \vartheta_2}} , \quad R_2 = \frac{\eta_T}{\sqrt{1 + \kappa_2(\tau_2) \tan^2 \vartheta_2}} , \quad (3.104) \]
\[ \kappa_2(t) = 1 + 2\xi_2^2 - 2\xi_2 \sqrt{1 + \xi_2^2} \coth(\pi J t) \sqrt{1 + \xi_2^2} + 2 \sinh^{-1}(\xi_2) \quad (3.105) \]
and \( \tan \vartheta_2 \) is given by Eq. (3.61).

We finally calculate the rf field that produces the optimal controls \( u_1, u_2 \). Since \( u_1 = \cos \beta_1 \), it is
\[ \dot{\beta}_1 = -\frac{\dot{u}_1}{\sqrt{1 - u_1^2}} . \quad (3.106) \]
Using Eq. (3.98) for \( u_1 \) we find, after some lengthy calculation, that
\[ \dot{\beta}_1 = -\pi J \frac{u_1(1 + u_1^2)}{\sqrt{1 - u_1^2}} \tanh \left( \frac{\phi_1}{2} \right) \sqrt{1 + \xi_1^2} . \quad (3.107) \]
But it is also \( \tan \beta_1 = \langle I_x \rangle / \langle I_x \rangle \), so
\[ \dot{\beta}_1 = \cos^2 \beta_1 \frac{d}{dt} \left( \frac{\langle I_x \rangle}{\langle I_x \rangle} \right) = u_1^2 \frac{d}{dt} \left( \frac{\langle I_z \rangle}{\langle I_x \rangle} \right) . \quad (3.108) \]
3.6 Summary of the Results and Discussion

In this section we summarize the results of the previous sections and attempt to explain them physically. We describe first the characteristics of the optimal pulse sequence for the case \( \xi_1 > \xi_2 \), i.e. for \( k_1 < k_2 \). Note that the presence of finite longitudinal relaxation rates results in an optimal transfer duration \( T_{opt} \) in which the maximum transfer efficiency is achieved. We compute this \( T_{opt} \) by finding the optimal pulse sequence for every choice of transfer duration \( T \) and then locating the \( T \) that gives the best transfer efficiency. Depending on the values of the problem parameters, we find three important cases in the optimal solution:

- **\( T \leq T_A \) (case A)** \( T_A = \cot^{-1}(2\xi_1)/\pi J \), for \( \xi_1 > \xi_2 \): In this case \( u_1(t) = u_2(t) = 1 \) throughout, i.e. \( \beta_1 \) and \( \beta_2 \) in Fig. 3.3 are always kept zero and this solution corresponds to the INEPT pulse sequence.
**3.6 SUMMARY OF THE RESULTS AND DISCUSSION**

- **$T_A < T \leq T_B$ (case B1)** (we describe how we calculate $T_B$ below): In this case the optimal pulse sequence has two distinct phases, see Fig. 3.5(a). There is a switching time $\tau_1$ such that for $0 \leq t \leq \tau_1$ (phase I), $u_2(t) = 1$ and $u_1(t)$ is increased gradually from a value $u_1(0) < 1$ to $u_1(\tau_1) = 1$. Then, for time $\tau_1 \leq t \leq T$ (phase II), the optimal controls are $u_1(t) = u_2(t) = 1$.

- **$T > T_B$ (case B2)**: Here the optimal pulse sequence has three distinct phases, see Fig. 3.6(a). There are two switching times, $\tau_1$ and $T - \tau_2$. Phases I and II are the same as above: For $0 \leq t \leq \tau_1$ (phase I), $u_2(t) = 1$ and $u_1(t)$ is increased gradually from a value $u_1(0) < 1$ to $u_1(\tau_1) = 1$. For time $\tau_1 \leq t \leq T - \tau_2$ (phase II), the optimal controls are $u_1(t) = u_2(t) = 1$. Finally, for $T - \tau_2 \leq t \leq T$ (phase III), we have $u_1(t) = 1$ and $u_2(t)$ is decreased from $u_2(T - \tau_2) = 1$ to $u_2(T) < 1$.

We now give physical explanation for the existence of these three cases. For small enough $T$, the major limitation for the transfer $r_1(0) \rightarrow r_2(T)$ is not relaxation, but the limited available time. The optimal choice $u_1 = u_2 = 1$ maximizes (absolute value) the off-diagonal elements $\pm u_1 u_2$, which accomplish the transfer $r_1(t) \rightarrow r_2(t)$, as can be seen from the system equation (3.35). It also maximizes the diagonal elements, i.e. the relaxation rates of $r_1(t), r_2(t)$. But for small available time $T$, the gain that we get by maximizing the desired transfer at each moment $t$ is more important than the (small) relaxation losses. As time $T$ increases, the relaxation degrades more the performance and the choice $u_1 = u_2 = 1$ ceases to be optimal. With $u_1 < 1$ or $u_2 < 1$ we may reduce the transfer rate of $r_1(t) \rightarrow r_2(t)$, but at the same time we decrease also the instantaneous relaxation rates $\xi_i u_i^2 + \xi_i, i = 1, 2$. Since for large enough $T$ the relaxation dominates, we conclude that by an appropriate choice of $u_1 \leq 1$ or $u_2 \leq 1$ we may get a better efficiency for the transfer $r_1(0) \rightarrow r_2(T)$. This appropriate choice corresponds to the cases B1 and B2. Note that for $k_1 < k_2$, the system in case B2 spends more time in phase I ($u_1 < 1, u_2 = 1$) than in phase III ($u_1 = 1, u_2 < 1$), see Fig. 3.5(a). This happens because for $k_1 < k_2$ and $u_1 = u_2 = u < 1$, the relaxation rate $\xi_1 u_1^2 + \xi_1$ is lower than the rate $\xi_2 u_2^2 + \xi_2$ (note $\xi_1 u_1^2 + \xi_1 - \xi_2 u_2^2 - \xi_2 = (\xi_1 - \xi_2) u^2 + \xi_1 - \xi_2 < \xi_1 - \xi_2 + \xi_1 - \xi_2 = 0$, since $\xi_1 + \xi_1 = \xi_2 + \xi_2 = k/J$). Based on the above observation about the duration of phases I and III, we expect that as we increase $T$, from values where case A holds to values where case B2 is the optimal, there must be an intermediate range of values of $T$ where the optimal pulse sequence has no phase III at all. This is case B1.

The duration $T_B$ above which the optimal pulse sequence is different than INEPT is $T_A = \cot^{-1}(2\xi_1)/\pi J$, for $\xi_1 > \xi_2$. We can explain the dependence of this quantity on the
Figure 3.5: Optimal pulse sequence when $T_A < T \leq T_B$ for (a) $k_1 < k_2$ and (b) $k_1 > k_2$. For case (b) we just interchanged the values of $k_1, k_2$ from case (a), keeping the same $k$. Observe the symmetry in the optimal controls.

Figure 3.6: Optimal pulse sequence when $T > T_B$ for (a) $k_1 < k_2$ and (b) $k_1 > k_2$. Again, for case (b) we just interchanged the values of $k_1, k_2$. The symmetry in the controls appears again. Note that the duration $T$ has been set equal to the optimal duration $T_{opt}$, which maximizes the optimal transfer efficiency $\eta_T$. For the values $k = J, k_1 = 0.05J, k_2 = 0.25J$ that we used in (a), it is $T_{opt} \simeq 0.468J^{-1}$. For case (b), $T_{opt}$ is the same.
parameters $k, k_1$. Note that $\xi_1 = (k - k_1)/J$, so $T_A$ is a decreasing function of $k$ and an increasing function of $k_1$. For larger $k$ (larger transverse relaxation) it is more costly to have the vectors $r_1, r_2$ parallel to the $xy$ plane, i.e. it is more costly to have $u_1 = u_2 = 1$, see Fig. 3.3. This explains why $T_A$, which determines the range of values of $T$ where the INEPT pulse sequence is optimal, is decreased. Now for larger $k_1$ (larger longitudinal relaxation) it is more costly to have the vector $r_1$ parallel to the $z$ axis, i.e. to have $u_1 < 1$. This explains why $T_A$, and with it the range of optimality of INEPT, is increased.

The switching time $\tau_1$ for case B1 is calculated using the following equation

$$T - \tau_1 = \frac{1}{\pi J} \left( \frac{\pi}{2} - \varphi_1 \right), \quad (3.114)$$

where

$$\varphi_1 = \tan^{-1} \frac{2 \xi_1}{1 - \kappa_1(\tau_1)} \quad (3.115)$$

and

$$\kappa_1(t) = 1 + 2 \xi_1^2 - 2 \xi_1 \sqrt{1 + \xi_1^2 \coth(\pi J t) \left(\frac{1}{1 + \xi_1^2} + 2 \sinh^{-1} \xi_1\right)} \quad (3.116)$$

The switching times for case B2 are calculated by solving the following system of equations

$$T - \tau_1 - \tau_2 = \frac{\vartheta_2 - \vartheta_1}{\pi J}, \quad (3.117)$$

$$T - \tau_1 - \tau_2 = \frac{\varphi_2 - \varphi_1}{\pi J}, \quad (3.118)$$

where $\varphi_1$ as above and

$$\vartheta_1 = \tan^{-1} \frac{2 \xi_1 \kappa_1(\tau_1)}{1 - \kappa_1(\tau_1)}, \quad \vartheta_2 = \tan^{-1} \frac{1 - \kappa_2(\tau_2)}{2 \xi_2}, \quad \varphi_2 = \tan^{-1} \frac{1 - \kappa_2(\tau_2)}{2 \xi_2 \kappa_2(\tau_2)}. \quad (3.119)$$

Function $\kappa_2(t)$ is given by a similar formula as $\kappa_1(t)$, with $\xi_1$ replaced by $\xi_2$. The relation of these angles to the optimal control problem is explained in section 3.3. The time $T_B$ mentioned above can be found by solving Eqs. (3.117), (3.118) for $\tau_2 = 0$, i.e. with unknowns $T_B$ and $\tau_1$. In other words, we find the time $T_B$ for which the optimal pulse sequence develops the additional phase III by setting the duration of that phase equal to zero.

The results for $\xi_1 < \xi_2$ ($k_1 > k_2$) are analogous. The basic difference is for the case B1 (that with only one switching time). Here, we start in phase II ($u_1 = u_2 = 1$) and at time $t = T - \tau_2$ we switch to phase III, as it is shown in Fig. 3.5(b). The time $\tau_2$ is calculated by

$$T - \tau_2 = \frac{\vartheta_2}{\pi J} = \frac{1}{\pi J} \tan^{-1} \frac{1 - \kappa_2(\tau_2)}{2 \xi_2}. \quad (3.120)$$
For case B2 (that with the two switching times) we have three phases, as before, but now we spend more time in phase III than in phase I, see Fig. 3.6(b). The times $\tau_1, \tau_2$ are still calculated using (3.117), (3.118). The duration $T_A$ is given now by $T_A = \cot^{-1}(2\xi_2)/\pi J$.

We describe also what happens for the symmetric case $\xi_1 = \xi_2 = \zeta$. In this case, the optimal pulse sequence spends equal time $\tau$ in phases I and III, see Fig. 3.7. Note that when $\xi_1 = \xi_2$ there is no intermediate case (B1, with only one switching time). Here, the optimal pulse sequence is either the INEPT or that with the three phases (and the two switching times). The time $\tau$ can be calculated by solving the equation

$$T - 2\tau = \frac{1}{\pi J} \left[ \tan^{-1} \frac{1 - \kappa(\tau)}{2\zeta} - \tan^{-1} \frac{2\zeta\kappa(\tau)}{1 - \kappa(\tau)} \right],$$

where $\kappa(\tau)$ is given by the formula for $\kappa_1(t)$, with $\xi_1$ replaced by $\zeta$.

For each of the cases presented above, the maximum transfer efficiency $\eta_T = r_2(T)$ is calculated in section 3.4. The results are

- **case A (INEPT)**
  \[ \eta_T = e^{-\pi kT} \sin(\pi JT) \]  

- **case B1 (one switch)**
  \[ \eta_T = e^{-\pi[k_1 \tau_1 + k(T - \tau_1)]} \sqrt{1 - \xi_1 \sin 2\varphi_1}, \]  

for $k_1 < k_2$ and
  \[ \eta_T = e^{-[k_2 \tau_2 + k(T - \tau_2)]} \sqrt{1 - \xi_2 \sin 2\varphi_2}, \]
for $k_1 > k_2$

- **case B2 (two switches)**

\[
\eta_T = e^{-\pi[k_1 \tau_1 + k(T-\tau_1-\tau_2)+k_2 \tau_2]} \times \frac{\sqrt{1-\xi_1 \sin^2 2\varphi_1} \sqrt{1-\xi_2 \sin^2 2\varphi_2}}{\cos(\varphi_1 - \vartheta_1)}. \tag{3.125}
\]

We can easily prove that the efficiency remains the same if the values of $k_1$ and $k_2$ are swapped. In Fig. 3.8 we plot the efficiency $\eta_T$ as a function of $T$ for various values of the parameters $k, k_1, k_2$. We observe that for $k, k_1, k_2 \neq 0$ none of the operators $I_z, 2I_z S_z$ is protected against relaxation, while for $k_1 = 0$ or $k_2 = 0$ at least one of them is. We can explain the existence of this optimal $T$ intuitively. For small $T$ the available time for the application of the controls is not enough (we cannot ‘put’ much control to the system) and thus the maximum efficiency that we get is small. For large $T$ the phenomenon of relaxation dominates, since there is no operator protected against it, and thus the maximum efficiency that we achieve is poor. So, there must be an intermediate time $T$ such that $\eta_T$ becomes maximum. This time is $T_{\text{opt}}$. In Fig. 3.9(b) we plot the maximum efficiency, calculated at $T_{\text{opt}}$ for each choice of the parameters, along a specific line in $(k_1, k_2, k)$ space. This line is shown in Fig. 3.9(a) and has been chosen to simulate the transition from the slow molecular motion (slowly tumbling regime), where $k \gg k_1 \simeq k_2$, to the rapid molecular motion, where $k \simeq k_1 \sim k_2$, see (3.21)-(3.25).

For $0 \leq t \leq \tau_1$ (phase I), the optimal control is given by

\[
u_1(t) = \Phi_1(t) e^{-\pi \frac{\tilde{\Lambda}_1^2}{2} \frac{1 + \cosh \varphi_1(t)}{2} \sinh^2 \frac{\varphi_1(t)}{2}} + \tilde{\Lambda}_1 \cosh \varphi_1(t), \tag{3.126}
\]

where

\[
\Phi_1(t) = 2\pi J t \sqrt{1 + \xi_1^2} + 2 \sinh^{-1} \xi_1 \tag{3.127}
\]

and

\[
\tilde{\Lambda}_1 = \frac{\eta_T}{\sqrt{1 + \kappa_1(\tau_1) \tan^2 \varphi_1}}, \quad \tilde{\Lambda}_2 = \frac{\eta_T}{\sqrt{\kappa_1(\tau_1) + \frac{1}{\tan^2 \varphi_1}}} \tag{3.128}
\]

For $T - \tau_2 \leq t \leq T$ (phase III), the optimal control is

\[
u_2(t) = \Phi_2(t) e^{-\pi \frac{\tilde{R}_2^2}{2} \frac{1 + \cosh \varphi_2(T-t)}{2} \sinh^2 \frac{\varphi_2(T-t)}{2}} + \tilde{R}_2 \cosh \varphi_2(T-t), \tag{3.129}
\]
§3.6 SUMMARY OF THE RESULTS AND DISCUSSION

Figure 3.8: Optimal transfer efficiency $\eta_T$ as a function of the total transfer time $T$ for $k = J$ and various values of $k_1, k_2$ (normalized with respect to $J$). Observe that for $k_1, k_2 \neq 0$ there is an optimal transfer time $T_{opt}$.

Figure 3.9: (b) Maximum transfer efficiency $\eta_T$ evaluated at $T_{opt}$ for each point of the line in $(k_1, k_2, k)$ space shown in (a). The line is parametrized by the parameter $0 \leq s \leq 1$ and has been chosen such that the increase of $s$ from 0 to 1 simulates the transition from slow molecular motion, where $k \gg k_1 \simeq k_2$, to rapid molecular motion, where $k \simeq k_1 \sim k_2$. Note the superiority of the relaxation optimized pulse element (ROPE) compared to the INEPT pulse sequence for the case $k \gg k_1 \simeq k_2$.
Figure 3.10: (a) The rf pulse sequence corresponding to the controls \(u_1, u_2\) of Fig. 3a. The pulse sequence starts with a hard 13.30° pulse around \(y\) axis, which establishes \(u_1(0) = 0.229\), followed by phase I as shown above. During phase II, no rf pulses are applied. Approaching phase II the rf amplitude becomes large for a very short time period. This can experimentally be very well approximated by a hard pulse of small flip angle. Following phase III as shown above, we apply a final hard 35.20° pulse around the \(x\) axis, which completes the transfer. (b) Time evolution of the expectation values of operators participating in the transfer, under the optimal pulse sequence shown in (a).

where

\[
\phi_2(t) = 2\pi J t \sqrt{1 + \xi_2^2} + 2 \sinh^{-1} \xi_2
\]  

\[
(3.130)
\]

and

\[
\bar{R}_1 = \frac{\eta T}{\sqrt{\kappa_2(\tau_2) + \tan^2 \vartheta_2}}, \quad \bar{R}_2 = \frac{\eta T}{\sqrt{1 + \kappa_2(\tau_2) \tan^2 \vartheta_2}}.
\]

The corresponding rf amplitude for phase I is given by

\[
\omega_y(t) = 2\pi J \frac{u_1(t)}{\sqrt{1 - u_1^2(t)}} \tanh \left[ \frac{\phi_1(t)}{2} \right] \sqrt{1 + \xi_1^2}
\]

\[
(3.132)
\]

and for phase III by

\[
\omega_x(t) = 2\pi J \frac{u_2(t)}{\sqrt{1 - u_2^2(t)}} \tanh \left[ \frac{\phi_2(T - t)}{2} \right] \sqrt{1 + \xi_2^2}.
\]

\[
(3.133)
\]

The details of these calculations are described in section 3.5.

The optimal rf amplitudes for a specific choice of the parameters are shown in Fig. 3.10(a). The optimal transfer strategy from \(I_z \rightarrow 2I_zS_z\) is then as follows. We start with an initial hard pulse that precedes phase I and rotates the vector \(r_1\), Fig. 3.3, from the angle \(\beta_1(0^-) = 90^\circ\) (parallel to \(z\)-axis) to the initial angle \(\beta_1(0^+) = \cos^{-1} u_1(0)\). During phase I \((0^+ \leq t \leq \tau_1)\), we rotate \(r_1\) slowly towards the \(x\)-axis using the field \(\omega_y(t)\). During
§3.6 SUMMARY OF THE RESULTS AND DISCUSSION

phase II \((\tau_1 \leq t \leq T - \tau_2)\), no rf pulses are applied. We just let \(r_1\) evolve towards \(r_2\) on the \(xy\) plane. In phase III \((T - \tau_2 \leq t \leq T^-)\), we rotate \(r_2\) slowly from the \(y\)-axis towards the \(z\)-axis using the field \(\omega_z(t)\), up to the angle \(\beta_2(T^-) = \cos^{-1} u_2(T)\). Following phase III, we apply a final hard pulse that rotates \(r_2\) from \(\beta_2(T^-)\) to \(\beta_2(T^+) = 90^\circ\), in order to complete the transfer to \(2I_zS_z\). This optimal pulse sequence takes the place of INEPT. Note that approaching phase II the rf amplitude becomes large for a very short time period. This can experimentally be very well approximated by a hard pulse of small flip angle. In Fig. 3.10(b) we plot the time evolution of the expectation values of operators participating in the transfer, under the optimal pulse sequence shown in Fig. 3.10(a).

Up to this point we have considered in detail the optimal way for the transfer \(I_z \to 2I_zS_z\). The optimal transfer \(I_z \to S_z\) is achieved by first doing the transfer \(I_z \to 2I_zS_z\) optimally, followed by optimal transfer of \(2I_zS_z \to S_z\). The optimal pulse sequence for the later transfer is analogous to the pulse sequence discussed above, with spin \(S\) playing the role of spin \(I\) and \(k_1, k_2\) now representing the relaxation rates of longitudinal operators \(2I_zS_z\) and \(S_z\) respectively. Now \(k\) represents the transverse relaxation rates for operators \(2I_zS_x, S_x\) etc.
Chapter 4

Semidefinite Programming and Reachable Sets of Dissipative Bilinear Control Systems

The optimal control problem examined in the preceding chapter motivates the study of a specific class of dissipative bilinear control systems and the calculation of their reachable sets. In this chapter we show that the optimal solution and the reachable set of these systems can be found by solving a semidefinite program (semidefinite programming is a branch of convex optimization with many applications in engineering, see Vandenberghe and Boyd (1996)). As a practical result, we obtain an upper bound for the polarization-coherence transfer efficiency along an Ising spin chain, a system studied in the next chapter. The core of the text that follows was submitted in the form of a paper to the 44th IEEE Conference on Decision and Control, 2005, see Stefanatos and Khaneja (2005).

4.1 Statement of the Problem

Recall Eqs. (3.28)-(3.31) of the preceding chapter with $k_1 = k_2 = 0$, i.e. when the longitudinal relaxation is negligible compared to the transverse relaxation $k$, Khaneja et al. (2003) (this is a very good approximation for slowly tumbling molecules in the spin diffusion limit, where $|\omega_I - \omega_S|\tau_c \gg 1$, as we can see from Eqs. (3.21)-(3.25)).

\[
\frac{d}{dt} \langle I_z \rangle = -\omega_y \langle I_x \rangle , \tag{4.1}
\]

\[
\frac{d}{dt} \langle I_x \rangle = \pi k \langle I_x \rangle - \pi J \langle 2I_y S_z \rangle + \omega_y \langle I_z \rangle , \tag{4.2}
\]
\[ \frac{d}{dt} \langle 2I_y S_z \rangle = \pi J \langle I_x \rangle - \pi k \langle 2I_y S_z \rangle - \omega_x \langle 2I_z S_z \rangle, \quad (4.3) \]

\[ \frac{d}{dt} \langle 2I_z S_z \rangle = \omega_x \langle 2I_y S_z \rangle. \quad (4.4) \]

Here, it is shown explicitly the effect of the rf magnetic field. If we set

\[ x_1 = \langle I_z \rangle, \quad x_2 = \langle 2I_z S_z \rangle, \quad y_1 = \langle 2I_y S_z \rangle, \quad y_2 = \langle I_x \rangle, \quad v_1 = \omega_y, \quad v_2 = \omega_x, \quad (4.5) \]

and use \( k \) and \( J \) instead of \( \pi k, \pi J \), the above equations take the following system form

\[
\begin{pmatrix}
\dot{x}_1 \\
\dot{x}_2 \\
\dot{y}_1 \\
\dot{y}_2
\end{pmatrix} =
\begin{pmatrix}
0 & -v_1 & 0 & 0 \\
0 & 0 & 0 & -v_2 \\
v_1 & 0 & -k & -J \\
0 & v_2 & J & -k
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2 \\
y_1 \\
y_2
\end{pmatrix}.
\quad (4.6)
\]

The optimal control problem that we faced before can be expressed as follows: Given the dynamical system (4.6) and starting from the initial state \( \mathbf{e}_1 = (1, 0, 0, 0)^T \), what is the maximum achievable value of \( x_2 \) and what are the optimal controls \( v_1(t) \in \mathbb{R} \) and \( v_2(t) \in \mathbb{R} \) that achieve this value?

We showed in the preceding chapter that if we define

\[ r_i = \sqrt{x_i^2 + y_i^2}, \quad (4.7) \]

then the above system can be transformed to

\[
\begin{pmatrix}
\dot{r}_1 \\
\dot{r}_2
\end{pmatrix} =
\begin{pmatrix}
-\xi u_1^2 & -u_1 u_2 \\
u_2 u_1 & -\xi u_2^2
\end{pmatrix}
\begin{pmatrix}
r_1 \\
r_2
\end{pmatrix}.
\quad (4.8)
\]

Here \( u_1 = y_1/r_1 \) and \( u_2 = y_2/r_2 \) are the new control parameters which take their values in the interval \([-1, 1]\) and \( \xi = k/J \). The initial problem of maximum transfer from \( x_1 \) to \( x_2 \) has been transformed to the following equivalent question: Given the dynamical system (4.8) and the initial state \( (r_1(0), r_2(0)) = (1, 0) \), find the optimal control \( (u_1(t), u_2(t)) \), such that \( r_2 \) is maximized.

Motivated by this example, which originates from a real physical system, let us consider the following \( n \)-dimensional generalization of system (4.6):

\[
\begin{pmatrix}
\dot{x} \\
\dot{y}
\end{pmatrix} =
\begin{pmatrix}
0 & -V \\
V & A
\end{pmatrix}
\begin{pmatrix}
x \\
y
\end{pmatrix},
\quad (4.9)
\]

where \( x = (x_1, x_2, \ldots, x_n)^T \), \( y = (y_1, y_2, \ldots, y_n)^T \), \( V = \text{diag}(v_1, v_2, \ldots, v_n) \) and \( A = \{a_{ij}\} \) is such that its symmetric part \( A + A^T \) is negative definite. This condition insures that
the norm of the vector \((x, y)\) can only decrease. This models the physics in open quantum systems, where dissipation can only reduce coherence in the system dynamics. Furthermore, \(A\) is such that any two states \(y_i\) and \(y_j\) are coupled by its off-diagonal elements, not necessarily directly (we say \(A\) is irreducible).

**Problem 1** Given the dynamical system (4.9) and the starting state \((x(0), y(0))\), find the optimal control \((v_1(t), v_2(t), \ldots, v_n(t))\) which maximizes \(x_n\).

It is shown in the following section that the negative definiteness condition on \(A\) is a sufficient condition for the existence of an optimal solution.

If we define \(r_i = \sqrt{x_i^2 + y_i^2}\) and work as in the 2-dimensional case, we find that \(r_i\) satisfies the equation

\[
\frac{dr_i}{dt} = \sum_{j=1}^{n} a_{ij} u_i u_j r_j ,
\]  

(4.10)

where \(u_i = y_i / r_i\). Problem 1 has been transformed to the following.

**Problem 2** Given the dynamical system defined by (4.10) for \(i = 1, 2, \ldots, n\), with \(A = \{a_{ij}\}\) irreducible and such that \(A + A^T\) negative definite, and the starting state \((r_1(0), r_2(0), \ldots, r_n(0))\), with \(r_i(0) \geq 0\), find the optimal control \((u_1(t), u_2(t), \ldots, u_n(t))\), \(|u_i| \leq 1\), which maximizes \(r_n\), while it preserves \(r_i(t) \geq 0\).

Observe that if \(T_1 < T_2\), then the maximum achievable value in time \(T_1\) cannot exceed the corresponding value in time \(T_2\), since by putting \(u_i = 0\) the evolution in the interval \((T_1, T_2]\) can be stopped. Therefore, Problem 2 is considered as an infinite horizon problem.

Multiplying the \(i^{th}\) equation of system (4.10) with \(2r_i\), we get

\[
\frac{d}{dt} (r_i^2) = \sum_{j=1}^{n} 2a_{ij} u_i u_j r_i r_j
\]  

(4.11)

and from this

\[
\frac{d}{dt} (r_i^2) = U^2 \sum_{j=1}^{n} 2a_{ij} u_i r_i u_j r_j ,
\]  

(4.12)

where

\[
U = \sqrt{\sum_{i=1}^{n} (u_i r_i)^2} .
\]  

(4.13)

By setting

\[
p_i = r_i^2 , \quad m_i = \frac{u_i r_i}{U} ,
\]  

(4.14)
and rescaling time according to \( dt' = U^2 dt \), equation (4.12) becomes

\[
\frac{dp_i}{dt'} = \sum_{j=1}^{n} 2a_{ij}m_im_j. \tag{4.15}
\]

The initial optimal control problem has been transformed to the following one.

\textbf{Problem 3} Given the dynamical system defined by (4.15) for \( i = 1, 2, \ldots, n \) and the starting point \( p(0) = (p_1(0), p_2(0), \ldots, p_n(0))^T, \) \( p_i(0) \geq 0, \) find the unit vector \( m(t') = (m_1(t'), m_2(t'), \ldots, m_n(t'))^T \) that maximizes \( p_n \), while it preserves \( p_i(t') \geq 0. \) Matrix \( A = \{a_{ij}\} \) is irreducible and such that \( A + A^T \) is negative definite.

Note that, although Problem 2 is an infinite horizon problem, Problem 3 defined above may achieve its maximum for a finite final time \( T_f \). There is no inconsistency here, since the times for the two systems are related through \( dt' = U^2 dt \), so \( T_f = \int_0^{T_f} dt' = \int_0^\infty U^2 dt. \) If \( U(t) \to 0 \) sufficiently fast as \( t \to \infty \), then \( T_f \) is finite. As we will see, this is indeed the case.

In the next sections, we study problems 2 and 3 in detail. Having found an optimal solution for the latter, we can easily find a corresponding optimal control law for the former.

### 4.2 Reduction to a Semidefinite Program

In the following, the inner product \( \langle \cdot, \cdot \rangle \) in the space of symmetric \( n \times n \) matrices \( \text{Sym}_n \) is defined in the usual way as the trace of the matrix product, i.e. \( \langle A, B \rangle = \text{tr}(AB) \) for \( A, B \in \text{Sym}_n \). Note also that \( A \preceq 0 \) denotes that matrix \( A \in \text{Sym}_n \) is positive semidefinite, \( A \prec 0 \) that is negative definite etc.

\textbf{Theorem 1} Let us define matrices \( A_i \in \text{Sym}_n, i = 1, 2, \ldots, n, \) by the relation

\[
A_i = \begin{bmatrix}
  a_{i1} & O_{i1} & \vdots & O_{i2} \\
  O_{i1} & \ddots & \ddots & \vdots \\
  \vdots & \ddots & \ddots & O_{i3} \\
  a_{i(i-1)} & a_{i(i+1)} & \cdots & a_{in} \\
  O_{i3} & \vdots & \ddots & O_{i4} \\
  a_{i(i-1)} & a_{i(i+1)} & \cdots & a_{in} \\
  \vdots & \ddots & \ddots & \vdots \\
  a_{i1} & a_{i2} & \cdots & a_{in}
\end{bmatrix},
\]

where \( a_{ij} \) are the elements of matrix \( A \) given in Problem 3 and \( O_{il}, l = 1, 2, 3, 4, \) are zero matrices with appropriate sizes. The solution of Problem 3 can be reduced to the solution
of the following semidefinite program:

\[
\text{Find } \mathcal{E} = \max_M \langle A_n, M \rangle \\
\text{subject to } \langle A_i, M \rangle = -p_i(0), \quad i = 1, 2, \ldots, n - 1 \\
\text{and } M \succeq 0.
\]

The maximum achievable value of \( p_n \) is \( p_n(0) + \mathcal{E} \).

**Proof:** Let \( T_f \) be the time when \( p_n \) achieves its maximum, i.e. the final time. From equation (4.15) it is

\[
p_i(T_f) = p_i(0) + \sum_{j=1}^{n} 2a_{ij} \int_0^{T_f} m_i(t') m_j(t') dt'.
\] (4.16)

Observe that if we define the positive semidefinite matrix \( M \) through the relation

\[
M = \int_0^{T_f} m(t') m^T(t') dt',
\] (4.17)

then (4.16) becomes

\[
p_i(T_f) = p_i(0) + \langle A_i, M \rangle.
\] (4.18)

One other important observation is that the end point of the optimal trajectory should lie on the line \((0, 0, \ldots, 0, p_n)\) in \( p \)-space. Suppose that the end point has a component \( p_k > 0 \) for some \( k \neq n \). If \( p_k \) is directly coupled to \( p_n \) then choose \( m = (0, 0, \ldots, 0, m_k, 0, \ldots, 0, m_n)^T \) such that \( m_n(a_{nk}m_k + a_{nn}m_n) > 0 \) and evolve the system until \( p_k = 0 \). Thereby we get a greater value of \( p_n \). If \( p_k \) is not directly coupled to \( p_n \), we can still transfer from \( p_k \) to \( p_n \) using intermediate states (because matrix \( A \) is irreducible). We conclude that at the final time \( T_f \) the end point of the optimal trajectory should lie on the line \((0, 0, \ldots, 0, p_n)\). Thus, we have to maximize \( p_n(T_f) = p_n(0) + \langle A_n, M \rangle \) under the conditions \( p_i(T_f) = p_i(0) + \langle A_i, M \rangle = 0, \, i = 1, 2, \ldots, n - 1 \). Equivalently, we have to solve the following semidefinite program: Find \( \mathcal{E} = \max_M \langle A_n, M \rangle \) subject to \( \langle A_i, M \rangle = -p_i(0) \) for \( i = 1, 2, \ldots, n - 1 \) and \( M \succeq 0 \).

Having found an optimal \( M \), we can always find an appropriate unit vector \( m(t') \) such that \( M = \int_0^{T_f} m(t') m^T(t') dt' \) and \( p_i(t') \geq 0 \). Since \( M \succeq 0 \), it can always be decomposed in the form

\[
M = \sum_{k=1}^{r} \lambda_k m_k m_k^T.
\]
where $\lambda_k$ are the positive eigenvalues of $M$, $m_k$ are the corresponding (real) normalized eigenvectors and $r$ is the rank of $M$. Now let $N$ be a positive integer. Rewrite the above relation in the form

$$M = N \sum_{k=1}^{r} \Delta \lambda_k m_k m_k^T,$$

where $\Delta \lambda_k = \lambda_k / N$, and define the times $t_k'$ through

$$t_0' = 0, \quad t_k' = \sum_{l=1}^{k} \Delta \lambda_l \quad \text{for} \quad k = 1, 2, \ldots, r.$$

Let us forget for a moment the restrictions $p_i(t') \geq 0$. If we apply the control

$$m(t') = m_k \quad \text{for} \quad t_{k-1}' \leq t' < t_k', \quad k = 1, 2, \ldots, r$$

and repeat for $N$ times, then on the one hand the requirement $\int_0^{T_f} m(t')m^T(t')dt' = M$ is satisfied and on the other hand the trajectory in $p$-space approximates the line joining the initial point $I(p_1(0), p_2(0), \ldots, p_n(0))$ to the final point $F(0, 0, \ldots, p_n(T_f))$, see Fig 4.1(a). If $N$ is large enough then the trajectory actually follows this line, see Fig. 4.1(b), thus the restrictions $p_i(t') \geq 0$ are satisfied. Note that $T_f = \sum_{k=1}^{r} \lambda_k = \text{tr}(M)$ is finite, if $\text{tr}(M) < +\infty$. In the special case where $r = 1$, it is $M = \lambda mm^T$ and thus $m(t') = m$ for $t' \in [0, T_f]$, $T_f = \lambda$.

The conclusion is that we just need to solve the semidefinite program defined above. The maximum achievable value of $p_n$ is $p_n(T_f) = p_n(0) + \mathcal{E}$. ■
We show next how this control law can be applied to system (4.10) in Problem 2. For $0 \leq t' \leq t'_1$, $m(t') = \mathbf{m}_1 = \text{constant}$. Since, additionally, $\mathbf{m}_1$ is a unit vector, we can assume without loss of generality that its first component $m_1 \neq 0$. Consider the ratios
\[
    \frac{u_ir_i}{u_1r_1} = \frac{m_i(t')}{m_1(t')} = s_i, \quad i = 1, 2, \ldots, n.
\]
For $0 \leq t' \leq t'_1$, $s_i$ are constant. Define
\[
    \mathcal{M} = \max_i \left( \frac{s_ir_1}{r_i} \right), \quad i = 1, 2, \ldots, n.
\]
The optimal policy can be realized as
\[
    u_1 = \frac{1}{\mathcal{M}}
\]
and
\[
    u_i = \frac{s_ir_1}{r_i} u_1, \quad i = 2, 3, \ldots, n.
\]
where $i = 2, 3, \ldots, n$. With the above choice we insure that $|u_i| \leq 1$. Using this feedback law we can evolve system (4.10) in time $t$ and calculate the function $U(t) = \sum_{i=1}^n (u_ir_i)^2$. Then, we can find $t' = \int_0^t U^2 dt$. When $t' = t'_1$, we switch to $m(t') = \mathbf{m}_2$ and repeat the above procedure. If the rank of $M$ is $r = 1$ then the ratios $s_i$ keep the same value for all times. Note that the maximum achievable value of $r_n$ is
\[
    r_n(\infty) = \sqrt{p_n(T_f)} = \sqrt{p_n(0) + \mathcal{E}} = \sqrt{r_n^2(0) + \mathcal{E}}. \tag{4.19}
\]

In the above discussion we implicitly assumed that an optimal solution exists, and we used for $\mathcal{E}$ the characterization “maximum” instead of the more formal “supremum”. We show below that the negative definiteness of $A + A^T$ is a sufficient condition for the existence of an optimal solution. The following lemma is used.

**Lemma 1** If $B > 0$ and $M \succeq 0$, $B, M \in \text{Sym}_n$, then $\langle B, M \rangle \geq 0$.

**Proof:** Since $B \in \text{Sym}_n$ it can be diagonalized by an orthogonal matrix $O$, $B = O\Delta O^T$, where $\Delta = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ and $\lambda_i > 0$ are the eigenvalues of the positive definite matrix $B$. It is
\[
    \langle B, M \rangle = \text{tr}(BM) = \text{tr}(O\Delta O^T M) = \text{tr}(\Delta O^T MO) = \text{tr}(\tilde{M}) = \sum_{i=1}^n \lambda_i \tilde{m}_{ii},
\]
where $\tilde{M} = O^T MO$ and $\tilde{m}_{ii}$ are its diagonal elements. But $\tilde{M}^T = \tilde{M}$ and $x^T \tilde{M}x = (Ox)^T MOx \geq 0$ for every $x \in \mathbb{R}^n$, since $M \succeq 0$. So, it is also $\tilde{M} \succeq 0$ and thus $\tilde{m}_{ii} \geq 0$. Since, additionally, $\lambda_i > 0$, we conclude that $\langle B, M \rangle = \sum_{i=1}^n \lambda_i \tilde{m}_{ii} \geq 0$. \qed
§4.2 REDUCTION TO A SEMIDEFINITE PROGRAM

Theorem 2 If $A + A^T \prec 0$ then the semidefinite program defined in Theorem 1 has an optimal solution.

Proof: First we show that the set $S$ of all matrices $M \succeq 0$ satisfying the equality constraints $\langle A_i, M \rangle = -p_i(0), i = 1, 2, \ldots, n - 1,$ is non-empty. Indeed, the matrix

$$M = \text{diag}(-p_1(0)/2a_{11}, -p_2(0)/2a_{22}, \ldots, -p_n(0)/2a_{nn})$$

satisfies these conditions and, additionally, it is $M \succeq 0$, since $p_i(0) \geq 0$ and $a_{ii} < 0$ ($A + A^T \prec 0$). Note that $S$ is closed and convex. Now consider the function $f : S \to \mathbb{R}$ defined by $f(M) = \langle A_n, M \rangle$ and the matrix $B = -\sum_{i=1}^{n} A_i = -(A + A^T) \succ 0$. From Lemma 1 and for $M \in S$, we have

$$\langle B, M \rangle \geq 0 \Rightarrow \langle A_n, M \rangle \leq \sum_{i=1}^{n-1} \langle A_i, M \rangle = \sum_{i=1}^{n-1} p_i(0) < +\infty \Rightarrow f(M) < +\infty.$$ 

Thus $\sup_{M \in S} f(M) < +\infty$ and since $S$ is closed the supremum is achieved for a $M_0 \in S$, so it is actually a maximum. The existence of an optimal solution is established.

We finally show how the semidefinite programming formalism can be used for calculating the reachable set of point $I(p_1(0), p_2(0), \ldots, p_n(0))$. Consider the line $\varepsilon$ parallel to $p_n$-axis, with $p_i = \text{constant} \geq 0$, $i = 1, 2, \ldots, n - 1$. The maximum achievable value of $p_n$ on $\varepsilon$, starting from $I$, can be found by solving the following semidefinite program: Find $\max_M \langle A_n, M \rangle$ subject to $\langle A_i, M \rangle = p_i - p_i(0)$ for $i = 1, 2, \ldots, n - 1$ and $M \succeq 0$. If this program has a solution $M_0$ such that $p_n = p_n(0) + \langle A_n, M_0 \rangle \geq 0$, then let $P$ be the point $(p_1, p_2, \ldots, p_n)$ of $\varepsilon$, see Fig. 4.2. This point belongs to the reachable set of $I$. Additionally, every point $N(p_1, p_2, \ldots, p_{n-1}, p'_n)$ of $\varepsilon$ with $0 \leq p'_n \leq p_n$, see Fig. 4.2, belongs also to the reachable set (first arrive at $P$ and then use $m = (0, 0, \ldots, 1)^T$ to go down, since (4.15) gives $p_n = a_{nn} < 0, \dot{p}_i = 0$ for $i \neq n$). Thus, the segment $PQ$, where $Q(p_1, p_2, \ldots, p_{n-1}, 0)$, is in the reachable set. By repeating the above procedure for all the allowed $\varepsilon \parallel p_n$, the reachable set of $I$ can be constructed.
4.3 Remarks on the Rank of the Semidefinite Program Solutions

In the preceding section we saw that the bigger the rank of the optimal $M$ for the semidefinite program, the more complicated is the optimal control law. Thus, it would be useful to know if there exist low rank optimal solutions and, additionally, rank upper bounds for them. Even more, we would like to know for what matrices $A$ the corresponding semidefinite program has solutions of the lowest possible rank $r = 1$. In this section we present a series of results in these directions.

**Lemma 2** Let us fix $A_1, A_2, \ldots, A_k \in \text{Sym}_n$ and $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{R}$. If there is a matrix $M \succeq 0$ such that

$$\langle A_i, M \rangle = \alpha_i, \quad i = 1, 2, \ldots, k,$$

then there is a matrix $M_0 \succeq 0$ such that

$$\langle A_i, M_0 \rangle = \alpha_i, \quad i = 1, 2, \ldots, k$$

and, additionally,

$$\text{rank}M_0 \leq \left\lfloor \frac{\sqrt{8k + 1} - 1}{2} \right\rfloor,$$

where $\lfloor \cdot \rfloor$ denotes the integer part of the embraced number.

**Proof:** See Barvinok (2002), chapter II, proposition 13.1.

**Proposition 1** If $A + A^T \prec 0$ then there is an optimal solution $M_0$ to the semidefinite program defined in Theorem 1, with

$$\text{rank}M_0 \leq \left\lfloor \frac{\sqrt{8n + 1} - 1}{2} \right\rfloor.$$

**Proof:** From Theorem 2 we have that, since $A + A^T \prec 0$, the semidefinite program has an optimal solution $M \succeq 0$, which satisfies

$$\langle A_i, M \rangle = -p_i(0), \quad i = 1, 2, \ldots, n - 1, \quad \langle A_n, M \rangle = \mathcal{E}.$$  

According to Lemma 2, there exists a $M_0 \succeq 0$ such that

$$\langle A_i, M_0 \rangle = -p_i(0), \quad i = 1, 2, \ldots, n - 1, \quad \langle A_n, M_0 \rangle = \mathcal{E}$$

and

$$\text{rank}M_0 \leq \left\lfloor \frac{\sqrt{8n + 1} - 1}{2} \right\rfloor.$$

Obviously $M_0$ is also an optimal solution.
4.3 Remarks on the rank of the semidefinite program solutions

**Corollary 1** If \( A + A^T \prec 0 \) and \( A \) is \( 2 \times 2 \), then the semidefinite program has an optimal solution of rank \( r \leq 1 \).

**Proof:** Apply Proposition 1 for \( n = 2 \). ■

In general, the bound imposed by Lemma 2 is the best possible. However, there is one special case where it can be sharpened.

**Lemma 3** For some positive integer \( r \), let us fix \( k = (r+2)(r+1)/2 \) matrices \( A_1, A_2, ..., A_k \in \text{Sym}_n \), where \( n \geq r + 2 \), and \( k \) numbers \( \alpha_1, \alpha_2, ..., \alpha_k \in \mathbb{R} \). If there is a matrix \( M \succeq 0 \) such that
\[
\langle A_i, M \rangle = \alpha_i, \quad i = 1, 2, ..., k
\]
and the set of all such matrices is bounded, then there is a matrix \( M_0 \succeq 0 \) such that
\[
\langle A_i, M_0 \rangle = \alpha_i, \quad i = 1, 2, ..., k
\]
and, additionally,
\[
\text{rank} M_0 \leq r.
\]

**Proof:** See Barvinok (2002), chapter II, proposition 13.4. ■

**Proposition 2** If \( A + A^T \prec 0 \) and \( A \) is \( 3 \times 3 \), then the semidefinite program has an optimal solution of rank \( r \leq 1 \).

**Proof:** Since \( A + A^T \prec 0 \), the semidefinite program has an optimal solution \( M \succeq 0 \), which satisfies
\[
\langle A_i, M \rangle = -p_i(0), \quad i = 1, 2, \quad \langle A_n, M \rangle = \mathcal{E}.
\]
From Lemma 3 we see that the choice \( r = 1 \) gives \( k = 3 = n \), since \( n = 3 \) according to the hypothesis of the proposition. In order to apply Lemma 3, we just need to show that the set of optimal matrices, i.e. all the matrices \( M \succeq 0 \) satisfying the above relations, is bounded. Consider the matrix \( B = -\sum_{i=1}^{3} A_i = -(A + A^T) \succ 0 \). For a matrix \( M \succeq 0 \) in the set of optimal solutions, we have
\[
\langle B, M \rangle = p_1(0) + p_2(0) - \mathcal{E} < +\infty.
\]
But from Lemma 1, we have also \( \langle B, M \rangle = \sum_{i=1}^{3} \lambda_i \tilde{m}_{ii} \), where \( \lambda_i \) are the eigenvalues of \( B \) and \( \tilde{m}_{ii} \) the diagonal elements of the matrix \( \tilde{M} = O^T MO \), \( O \) the orthogonal matrix diagonalizing \( B \). Combining these we find that
\[
\sum_{i=1}^{3} \lambda_i \tilde{m}_{ii} < +\infty.
\]
Since $\lambda_i > 0$ and $\tilde{m}_{ii} \geq 0$, the above relation implies that

$$\tilde{m}_{ii} < +\infty,$$

thus

$$\text{tr}(M) = \text{tr}(\tilde{M}) < +\infty.$$

But

$$\langle M, M \rangle = \text{tr}(M^2) \leq (\text{tr}(M))^2 < +\infty,$$

since $M \succeq 0$. So indeed the set of optimal $M$ is bounded and we can apply Lemma 3 with $r = 1$. This means that there is an optimal $M_0 \succeq 0$ with $\text{rank} M_0 \leq 1$. Note that the bound that Proposition 1 gives in this case is only $\lfloor (\sqrt{8 \cdot 3 + 1} - 1)/2 \rfloor = 2$. 

**Lemma 4** Let us call an $n \times n$ matrix $A = \{a_{ij}\}$ $r$-diagonal if $a_{ij} = 0$ unless $|i - j| < r$. Suppose that the matrices $A_1, A_2, \ldots, A_k \in \text{Sym}_n$ are $r$-diagonal and there exists a matrix $M \succeq 0$ such that

$$\langle A_i, M \rangle = \alpha_i \in \mathbb{R}, \quad i = 1, 2, \ldots, k.$$

Then there exists a matrix $M_0 \succeq 0$ such that

$$\langle A_i, M_0 \rangle = \alpha_i, \quad i = 1, 2, \ldots, k$$

and, additionally,

$$\text{rank} M_0 \leq r.$$

**Proof:** See Barvinok (2002), chapter IV, corollary 10.3, problem 3. 

**Proposition 3** If $A + A^T \prec 0$ and $A$ is $r$-diagonal, then the semidefinite program has an optimal solution of rank $\leq r$.

**Proof:** Since $A + A^T \prec 0$, there exists an optimal solution of the semidefinite program. Since $A$ is $r$-diagonal, the corresponding $A_i$ are also $r$-diagonal. Thus, we can apply Lemma 4, which assures the existence of an optimal solution of rank $\leq r$. 

We conclude this section by noting that there is strong numerical evidence that the following conjecture is true.

**Conjecture 1** If $A + A^T \prec 0$ then the semidefinite program has an optimal solution of rank $r = 1$. 

4.4 Examples

In this section we solve problems 2 and 3 for some specific systems. We start from the system with
\[
A = \begin{bmatrix}
-\xi & -1 \\
1 & -\xi
\end{bmatrix}, \quad \xi > 0,
\]
which corresponds to system (4.8) appeared in the introduction. It is not necessary to solve numerically the corresponding semidefinite program, because we can attack this particular case analytically. Since
\[
A + A^T = \text{diag}(-2\xi, -2\xi) \prec 0
\]
and \(A\) is \(2 \times 2\), from Corollary 1 we have that there is an optimal solution to the semidefinite program with rank \(r \leq 1\). Thus, there is an optimal constant vector \(m = (m_1, m_2)^T\), solution to Problem 3. The system equation (4.15) with \(A\) given above and \(m\) constant gives
\[
p_1(T_f) = p_1(0) - (\xi m_1^2 + m_1 m_2)T_f, \quad p_2(T_f) = p_2(0) + (m_2 m_1 - \xi m_2^2)T_f.
\]
Optimality requires
\[
p_1(T_f) = 0 \Rightarrow T_f = \frac{p_1(0)}{\xi m_1^2 + m_1 m_2},
\]
so
\[
p_2(T_f) = p_2(0) + \frac{m_2 m_1 - \xi m_2^2}{\xi m_1^2 + m_1 m_2}p_1(0).
\]
In order to maximize \(p_2(T_f)\), we just need to maximize the coefficient of \(p_1(0)\). If we set \(m_2/m_1 = x\), then this coefficient takes the form
\[
f(x) = \frac{x - \xi x^2}{x + \xi}.
\]
Before maximizing \(f\), we find the allowed values of variable \(x\). It should be \(p_2(T_f) \geq p_2(0) \Rightarrow x - \xi x^2 \geq 0\) and \(p_1(T_f) \leq p_1(0) \Rightarrow x + \xi \geq 0\). These are both satisfied when \(x \in [0, 1/\xi]\). We calculate the maximum of \(f\) in this interval. It is not difficult to verify that
\[
f'(x) = -\frac{\xi(x^2 + 2\xi x - 1)}{(x + \xi)^2}
\]
becomes zero at the point
\[
x_0 = \sqrt{1 + \xi^2} - \xi
\]
of the interval \([0, 1/\xi]\). Also verify that \(f'(x) > 0\) for \(x \in [0, x_0]\) and \(f'(x) < 0\) for \(x \in (x_0, 1/\xi]\). So \(f(x_0)\) is a maximum in the interval \([0, 1/\xi]\). After some manipulation we find that
\[
f(x_0) = x_0^2.
\]
The maximum achievable value of $p_2$ is

$$p_2(T_f) = p_2(0) + x_0^2 p_1(0)$$

(4.27)

and the optimal unit vector is

$$m = \left( \frac{1}{\sqrt{1 + x_0^2}}, \frac{x_0}{\sqrt{1 + x_0^2}} \right).$$

(4.28)

The optimal trajectory in $p$-space is a straight line joining the points $(p_1(0), p_2(0))$ and $(0, p_2(T_f))$.

The maximum achievable value of $r_2$ is

$$r_2(\infty) = \sqrt{r_2^2(0) + x_0^2 r_1^2(0)}. $$

(4.29)

If the starting state is the point $(r_1(0), r_2(0)) = (1, 0)$, the maximum transfer efficiency takes the value

$$r_2(\infty) = x_0 = \sqrt{1 + \xi^2} - \xi. $$

(4.30)

For $\xi = 1$ this efficiency is $\sqrt{2} - 1$. The optimal controls $u_1, u_2$ for system (4.8), can be found by using the method described in section 4.2. If we define

$$\mathcal{M} = \max\left(1, \frac{x_0 r_1}{r_2}\right),$$

the optimal policy can be realized as

$$u_1 = \frac{1}{\mathcal{M}}, \quad u_2 = \frac{x_0 r_1}{r_2} u_1.$$  

Observe that the initial point $(1, 0)$ is a stationary point of the optimal control policy $(r_2(0) = 0 \Rightarrow \mathcal{M} = \infty \Rightarrow u_1 = 0 \Rightarrow u_2 = 0)$. This optimal policy in the infinite horizon case should then be interpreted as the limit of optimal control policy for the corresponding finite time problem (Finite time for the $r_i$ system, don’t confuse it with the finite time problem for $p_i$ which corresponds to the infinite horizon problem for $r_i$. It is the special case $k_1 = k_2 = 0$ of the finite time problem solved in the preceding chapter. The solution for this particular case can be found in Khaneja et al. (2003)). In practice, we give a small but finite value in $r_2(0)$ (an initial ‘kick’ from zero) which makes the optimal control law applicable. In Fig. 4.3(a) we plot the optimal controls $u_1$ and $u_2$. In Fig. 4.3(b) we depict $r_1(t), r_2(t)$ and in Fig. 4.3(c) the corresponding optimal trajectory in $r$-space. For all these figures it is $\xi = 1$ and $(r_1(0), r_2(0)) = (1, 0)$. 


Figure 4.3: (a) Optimal controls $u_1(t)$ and $u_2(t)$ for system (4.8) when $\xi = 1$ and $(r_1(0), r_2(0)) = (1, 0)$. (b) The corresponding state variables $r_1(t)$ and $r_2(t)$. (c) The optimal trajectory in $r$-space.
Remark 1 The closure of the reachable set of point \((1, 0)\) is

\[
\mathcal{R}(1, 0) = \{ r_1, r_2 \geq 0 \mid \sqrt{r_2^2 + x_0^2} \leq x_0 \},
\]

where \(x_0 = \sqrt{1 + \xi^2} - \xi\). This set is depicted in Fig. 4.3(c) for \(\xi = 1\). The closure of the reachable set \(\mathcal{R}(1, 0, 0, 0)\) for the corresponding bilinear system (4.6) is

\[
\{(x_1, x_2, y_1, y_2) \in \mathbb{R}^4 \mid \sqrt{(x_2^2 + y_2^2) + x_0^2(x_1^2 + y_1^2)} \leq x_0 \}.
\]

The next case that we examine is the system with

\[
A = \begin{bmatrix}
-\xi & -1 & 0 \\
1 & -\xi & -1 \\
0 & 1 & -\xi
\end{bmatrix}, \quad \xi > 0.
\]

Since \(A + A^T = \text{diag}(-2\xi, -2\xi, -2\xi) < 0\) and \(A\) is \(3 \times 3\), from Proposition 2 we have that the semidefinite program has a solution of rank \(r \leq 1\). Now let us become more specific, so set \(\xi = 1\) and consider the starting point \((p_1(0), p_2(0), p_3(0)) = (1, 1, 0)\). The corresponding matrices \(A_i\) are

\[
A_1 = \begin{bmatrix}
-2 & -1 & 0 \\
-1 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}, \quad A_2 = \begin{bmatrix}
0 & 1 & 0 \\
1 & -2 & -1 \\
0 & -1 & 0
\end{bmatrix}, \quad A_3 = \begin{bmatrix}
0 & 0 & 0 \\
0 & 0 & 1 \\
0 & 1 & -2
\end{bmatrix}.
\]

If we solve numerically the corresponding semidefinite program using some appropriate software package, for example SDPT3, Toh et al. (1999), we find that the optimal matrix \(M \succeq 0\) is

\[
M = \begin{bmatrix}
0.1775 & 0.3225 & 0.1304 \\
0.3225 & 0.5856 & 0.2368 \\
0.1304 & 0.2368 & 0.0958
\end{bmatrix},
\]

and the maximum achievable value of \(p_3\) is

\[
p_3(T_f) = p_3(0) + \langle A_3, M \rangle = \langle A_3, M \rangle = 0.2821.
\]

It is easy to verify that this matrix has two zero eigenvalues and one nonzero, so its rank is indeed \(r = 1\). It can be written in the form \(M = \lambda mm^T\), where \(\lambda = 0.8589 \approx T_f\) is the nonzero eigenvalue and

\[
m = (m_1, m_2, m_3)^T = (0.4546, 0.8257, 0.3339)^T.
\]
the corresponding eigenvector. This unit vector is the optimal solution for Problem 3.

The maximum achievable value of \( r_3(\infty) = \sqrt{p_3(T_f)} \) is 0.5311. We find the optimal \( u_1, u_2, u_3 \). Let us set

\[
x_0 = \frac{m_2}{m_1} = 1.8163, \quad y_0 = \frac{m_3}{m_1} = 0.7345.
\]

If we define

\[
M = \max(1, \frac{x_0 r_1}{r_2}, \frac{y_0 r_1}{r_3}),
\]

then the optimal policy can be realized as

\[
u_1 = \frac{1}{M}, \quad u_2 = \frac{x_0 r_1}{r_2} u_1, \quad u_3 = \frac{y_0 r_1}{r_3} u_1.
\]

Observe that the initial point \((1, 1, 0)\) is a stationary point of the optimal control policy \((r_3(0) = 0 \Rightarrow M = \infty \Rightarrow u_1 = 0 \Rightarrow u_2 = u_3 = 0)\). The situation is similar with that in the previous example. Again, the optimal policy in the infinite horizon case should be interpreted as the limit of the optimal policy for the corresponding finite time problem.

In practice, a small but finite value is given to \( r_3(0) \). In Fig. 4.4(a) we plot the optimal controls \( u_1, u_2 \) and \( u_3 \). In Fig. 4.4(b) we depict \( r_1(t), r_2(t), r_3(t) \) and in Fig. 4.4(c) the corresponding optimal trajectory in \( r \)-space.

Another interesting case to examine is the same system with \( \xi > 0 \) unspecified and starting point \((p_1(0), 0, p_3(0))\). This problem can be solved analytically and has the practical application that it gives an analytic upper bound for our ability to coherently control a specific dissipative quantum (spin) system, examined in the next chapter. As before, we know that there is an optimal constant vector \( m = (m_1, m_2, m_3)^T \). From equations (4.15) we find

\[
\begin{align*}
p_1(T_f) &= p_1(0) - (\xi m_1^2 + m_1 m_2) T_f, \\
p_2(T_f) &= p_2(0) + (m_2 m_1 - \xi m_2^2 - m_2 m_3) T_f, \\
p_3(T_f) &= p_3(0) + (m_3 m_2 - \xi m_3^2) T_f.
\end{align*}
\]

Optimality requires

\[
p_1(T_f) = 0 \Rightarrow T_f = \frac{p_1(0)}{\xi m_1^2 + m_1 m_2} \tag{4.32}
\]

and

\[
p_2(T_f) = 0 \Rightarrow m_2 m_1 - \xi m_2^2 - m_2 m_3 = 0. \tag{4.33}
\]

So, we have to maximize

\[
p_3(T_f) = p_3(0) + \frac{m_3 m_2 - \xi m_3^2}{\xi m_1^2 + m_1 m_2} p_1(0) \tag{4.34}
\]
Figure 4.4: (a) Optimal controls $u_1(t), u_2(t)$ and $u_3(t)$ for system (4.10), with $A$ the $3 \times 3$ matrix given in the text, when $\xi = 1$ and $(r_1(0), r_2(0), r_3(0)) = (1, 1, 0)$.  (b) The corresponding state variables $r_1(t), r_2(t)$ and $r_3(t)$.  Observe that $r_2(t)/r_1(t) = 1$ throughout. Remember that the optimal trajectory in $p$-space is a straight line ending at the point $(0, 0, p_3(T_f))$, so $p_2(t)/p_1(t) = p_2(0)/p_1(0) = 1$ for the starting point $(1, 1, 0)$.  (c) The optimal trajectory in $r$-space.
subject to the constraint (4.33). We just need to maximize the coefficient of $p_1(0)$ under the same condition. If we set

$$\frac{m_2}{m_1} = x, \quad \frac{m_3}{m_1} = y,$$

then this coefficient takes the form

$$g(x, y) = \frac{xy - \xi y^2}{x + \xi},$$

while the condition becomes

$$x(1 - \xi x - y) = 0 \Rightarrow y = 1 - \xi x.$$  \hspace{1cm} (4.36)

Note that $x=0$ gives $g \leq 0$ so it is rejected. Using (4.36), $g$ becomes a function of $x$ only

$$f(x) = g(x, y(x)) = \frac{-\xi (1 + \xi^2) x^2 + (1 + 2\xi^2)x - \xi}{x + \xi}. \hspace{1cm} (4.37)$$

We find the allowed values of $x$. A natural requirement is $p_3(T_f) \geq p_3(0) \Rightarrow y(x - \xi y) \geq 0 \Rightarrow (y \geq 0$ and $x - \xi y \geq 0)$ or $(y \leq 0$ and $x - \xi y \leq 0)$. Using (4.36) we find that the first option implies $x_1 \leq x \leq x_2$ and the second $x_2 \leq x \leq x_1$, where $x_1 = \xi/(1 + \xi^2)$, $x_2 = 1/\xi$. Since $x_1 < x_2$, only the first option is acceptable, so it must be $x \in [x_1, x_2]$. For such $x$, the similar requirement $p_1(T_f) \leq p_1(0) \Rightarrow x + \xi \geq 0$ is satisfied. So we maximize $f$ in the interval $[x_1, x_2] = [\xi/(1 + \xi^2), 1/\xi]$. Solving the equation $f'(x_0) = 0$, we find

$$x_0 = \sqrt{\xi^2 + 2} - \xi.$$ \hspace{1cm} (4.38)

Indeed, $x_0 \in [x_1, x_2]$. The corresponding maximum value of $f$ is

$$f_{\text{max}} = \frac{x_0^4}{4}.$$ \hspace{1cm} (4.39)

The maximum achievable value of $p_3$ is

$$p_3(T_f) = p_3(0) + f_{\text{max}} p_1(0).$$ \hspace{1cm} (4.40)

Condition (4.33) implies that in the optimal case it is $\dot{p}_2 = 0$, so it is also $\dot{r}_2 = 0$. If $r_2(0) = 0$ then $r_2(t) = 0$ and, as we can see from (4.10), there is no transfer from $r_1$ to $r_3$. What we actually examine here is the limiting case $r_2(0) = \epsilon \to 0^+$, where $\epsilon$ is an arbitrarily small positive number. We can still use condition (4.33), i.e. $\dot{r}_2 = 0$. The transfer $r_1 \to r_3$ takes place through $r_2$ which is held to the small constant value $r_2 = \epsilon$. The maximum achievable value of $r_3$, which corresponds to the limit $\epsilon \to 0^+$, is

$$r_3(\infty) = \sqrt{r_3^2(0) + f_{\text{max}} r_1^2(0)}. \hspace{1cm} (4.41)$$
Figure 4.5: (a) Optimal controls $u_1(t)$, $u_2(t)$ and $u_3(t)$ for system (4.10), with $A$ the $3 \times 3$ matrix given in the text, when $\xi = 1$ and $(r_1(0), r_2(0), r_3(0)) = (1, \epsilon, \epsilon)$, $0 < \epsilon \ll 1$. Here we take $\epsilon = 0.01$ for convenience. (b) The corresponding state variables $r_1(t), r_2(t)$ and $r_3(t)$. Note that transfer $r_1 \to r_3$ takes place through $r_2$ which is held to the small constant value $r_2 = \epsilon$. Thus, this transfer requires more time compared to the preceding examples. (c) The optimal trajectory in $r$-space.
If the starting state is the point \((1, \epsilon, 0)\), where \(\epsilon \to 0^+\), the maximum efficiency is

\[
r_3(\infty) = \sqrt{f_{\text{max}}} = \frac{x_0^2}{2} = \frac{(\sqrt{\xi^2 + 2} - \xi)^2}{2}.
\] (4.42)

For \(\xi = 1\) we find that this efficiency is \(2 - \sqrt{3}\). In Fig. 4.5 we plot the optimal controls \(u_1(t), u_2(t), u_3(t)\), the state variables \(r_1(t), r_2(t), r_3(t)\) and the optimal trajectory in \(r\)-space.

Observe that the starting point is actually \((1, \epsilon, \epsilon)\). It is necessary to give a small positive initial value to \(r_3\), since the point \((1, \epsilon, 0)\) is still a stationary point of the optimal policy.

If the starting point is \((1, \epsilon, \epsilon)\), then by solving the corresponding semidefinite program we find numerically the same efficiency as in (4.42), in the limit \(\epsilon \to 0^+\).
Chapter 5

Relaxation Optimized Polarization and Coherence Transfer in Ising Spin Chains

In this chapter we generalize the problem of optimal polarization-coherence transfer in the presence of relaxation, introduced in chapter 3 for a coupled spin pair, to a more complex coupled spin system. Specifically, we address this problem for an Ising spin chain in the presence of transverse relaxation. This system is motivated again by experiments in multidimensional NMR spectroscopy of proteins, where polarization or coherence transfer along spin chains, in backbone and side chain assignments, is frequently encountered. Using the theory presented in the preceding chapter, we obtain upper bounds for the efficiency of polarization or coherence transfer along a spin chain. We then calculate numerically optimal pulse sequences that approach these bounds. The core of the text that follows was submitted in the form of a paper to Physical Review A, see Stefanatos et al. (2005).

5.1 Description of the System and Statement of the Problem

The system that we study in this chapter is a linear chain of \( n \) weakly interacting spins \( 1/2 \), placed in a static external magnetic field in the \( z \) direction (NMR experimental setup) and with Ising type coupling between next neighbors, see Fig. 5.1. In a suitably chosen (multiple) rotating frame, which rotates with each spin at its resonance (Larmor) frequency,
§5.1 DESCRIPTION OF THE SYSTEM AND STATEMENT OF THE PROBLEM

The system that we study in this chapter is a linear chain of \( n \) weakly interacting spins 1/2 with Ising coupling between next neighbors. This system originates from protein NMR spectroscopy. Note that besides the coherent evolution, described by the Ising couplings, there is also dissipative evolution caused by the interaction of the system with its environment.

Figure 5.1: The system that we study in this chapter is a linear chain of \( n \) weakly interacting spins 1/2 with Ising coupling between next neighbors. This system originates from protein NMR spectroscopy. Note that besides the coherent evolution, described by the Ising couplings, there is also dissipative evolution caused by the interaction of the system with its environment.

The Hamiltonian \( H_f \) that governs the free evolution of the system is given by

\[
H_f = \sum_{i=1}^{n-1} 2J_i I_i z I_{i+1} z .
\]  

Hamiltonian (5.1) is valid in the so-called weak coupling limit, where the resonance frequencies of the spins \( i \) and \( i+1 \) satisfy \( |\omega_i - \omega_{i+1}| \gg J_i \) and thus the Heisenberg coupling \((I_i \cdot I_{i+1})\), which is the characteristic indirect coupling between spins in isotropic liquids, can be approximated by the Ising coupling \((I_i z I_{i+1} z)\), Levitt (2001). Here, we examine in detail the case where the coupling constants \( J_i \) between the spins are equal. This choice simplifies the problem and is also an appropriate model for realistic situations. At the end of the chapter we show how the theory can be extended to the general case of arbitrary couplings. The common coupling constant is written in the form \( \sqrt{2} J \) for normalization reasons, so (5.1) takes the form

\[
H_f = \sqrt{2} J \sum_{i=1}^{n-1} I_i z I_{i+1} z .
\]  

Besides the coherent evolution, described by the Ising couplings, there is also dissipative evolution caused by the interaction of the system with its environment and expressed by the relaxation superoperator \( L \) in the master equation (2.35). The relaxation mechanisms that contribute to the super-operator \( L \) for our liquid state NMR system are, as we have
explained in chapter 2, dipole-dipole (DD) interaction and chemical shift anisotropy (CSA). We neglect interference effects. The corresponding relaxation operators can be found in chapter 2. Here we focus on slowly tumbling molecules in the spin diffusion limit, where the correlation time of the molecular tumbling is much longer than the inverse of the resonance frequencies of the spins, see section 2.2. In this limit and for DD and CSA relaxation mechanisms, as in our system, longitudinal relaxation rates (rates for operators with components only in the \( z \) direction, like \( I_{iz}, 2I_{iz}I_{(i+1)z} \)) are negligible compared with transverse relaxation rates (rates for operators with components in \( x \) and \( y \) directions, like \( 2I_{iz}I_{(i+1)x}, 2I_{iz}I_{(i+1)y} \)), see Eqs. (3.21)-(3.25).

The problem that we address in this chapter is to find the maximum efficiency for the polarization-coherence transfer along the spin chain

\[
I_{1\alpha} \rightarrow I_{n\beta},
\]

(5.3)

where \( \alpha, \beta \) can be \( x, y \) or \( z \), under the presence of the relaxation mechanisms mentioned above. The available controls are the components \( \omega_x \) and \( \omega_y \) (in the rotating frame) of the transverse radiofrequency (rf) magnetic field. This comes from the NMR experimental setup. If the resonance frequencies of the spins are well separated, each spin can be selectively excited (addressed) by an appropriate choice of the components of the rf field at its resonance frequency. We want also to calculate the rf fields that maximize transfer (5.3).

Note that we can concentrate on the polarization transfer

\[
I_{1z} \rightarrow I_{nz}.
\]

(5.4)

Having found how to accomplish this optimally, we can use the same scheme for transfer (5.3). We just need to add an initial and a final 90° hard pulse to make the rotations \( I_{1\alpha} \rightarrow I_{1z}, I_{nz} \rightarrow I_{n\beta} \). The way that polarization transfer (5.4) is done is shown in Fig. 5.2. From our previous work in chapter 3 we know the optimal way to do the first and the last step (the problem examined there is more general since we considered both longitudinal and transverse relaxation). In this chapter we investigate how can we make the intermediate steps optimally, i.e. the transfers

\[
2I_{(i-1)z}I_{iz} \rightarrow 2I_{iz}I_{(i+1)z},
\]

(5.5)

where \( i = 2, \ldots, n - 1 \). Each of these transfers is done by irradiating the intermediate spin \( i \). If the spins have well separated frequencies, then we can selectively excite them, thus we can examine each intermediate transfer separately from the others. Additionally, the
§5.1 DESCRIPTION OF THE SYSTEM AND STATEMENT OF THE PROBLEM

Figure 5.2: Steps for the polarization transfer $I_{1z} \rightarrow I_{nz}$. From our previous work in chapter 3 we know the optimal way to do steps 1 and n, in presence of the relaxation mechanisms mentioned in the text. In this chapter we investigate the intermediate steps. If the spins have well separated resonance frequencies, then we can selectively excite them. In this case we can examine separately each of the intermediate transfers. But the equations that describe these transfers are similar, thus we just need to study only one of them, for example the transfer $2I_{1z}I_{2z} \rightarrow 2I_{2z}I_{3z}$.

Figure 5.3: Energy level diagram for a weakly coupled three-spin 1/2 chain. The dark circles indicate excess population in the corresponding levels. The selective population inversion of the levels $\alpha\alpha\beta, \alpha\beta\beta$ and $\beta\beta\alpha, \beta\alpha\alpha$ corresponds to the transfer $2I_{1z}I_{2z} \rightarrow 2I_{2z}I_{3z}$ of longitudinal two-spin order.
§5.2 Formulation of the Problem in Terms of Optimal Control

Equations that describe these transfers are similar, so we actually need to study only one of them. In the following we focus on the transfer with \( i = 2 \)

\[
2I_{1z}I_{2z} \rightarrow 2I_{2z}I_{3z}.
\] (5.6)

Using the energy level diagram shown in Fig. 5.3, we can give a more physical picture for the full polarization transfer \( I_{1z} \rightarrow I_{3z} \) along a three-spin chain. The configuration of the initial state, corresponding to polarization \( I_{1z} \), has excess populations in the levels \( \alpha\alpha\alpha, \alpha\alpha\beta, \alpha\beta\alpha, \alpha\beta\beta \) (this configuration is not shown in Fig. 5.3). By selectively inverting the populations of \( \alpha\beta\beta, \beta\beta\beta \) and \( \alpha\beta\alpha, \beta\beta\alpha \), the spin system acquires the configuration shown in Fig. 5.3, with excess populations in the levels \( \alpha\alpha\alpha, \alpha\alpha\beta, \beta\beta\beta, \beta\beta\alpha \). This corresponds to state \( 2I_{1z}I_{2z} \). The next step is the selective population inversion shown in Fig. 5.3 (\( \alpha\alpha\beta \) with \( \alpha\beta\beta \) and \( \beta\beta\alpha \) with \( \beta\alpha\alpha \)). The new configuration (excess populations in \( \alpha\alpha\alpha, \alpha\beta\beta, \beta\beta\beta, \beta\alpha\alpha \)) corresponds to \( 2I_{2z}I_{3z} \). Thus, the intermediate transfer \( 2I_{1z}I_{2z} \rightarrow 2I_{2z}I_{3z} \) of longitudinal two-spin order has been accomplished. The final step is to selectively invert \( \alpha\beta\beta, \alpha\beta\alpha \) and \( \beta\beta\beta, \beta\beta\alpha \). The final configuration (excess populations in \( \alpha\alpha\alpha, \alpha\beta\alpha, \beta\alpha\alpha, \beta\beta\alpha \), not shown in Fig. 5.3, corresponds to polarization \( I_{3z} \). The transfer \( I_{1z} \rightarrow I_{3z} \) is completed.

5.2 Formulation of the Problem in Terms of Optimal Control

In this section we formulate the problem of transfer (5.6) in the context of optimal control. The first step is to write down evolution equations for the ensemble averages of the operators that participate in this transfer (with ensemble average of an operator \( O \) we mean \( \langle O \rangle = \text{trace}\{\rho O\} \)). This can be done by using the master equation (2.35) with the appropriate for our system Hamiltonian \( H \) and relaxation superoperator \( L \). The Hamiltonian \( H \) is the sum of the free evolution Hamiltonian \( H_f \) given in (5.2) and the Hamiltonian \( H_{rf} \) that describes the effect of the transverse rf magnetic field. For transfer (5.6) only the first two terms of (5.2) participate, so we can set \( H_f = \sqrt{2J}(I_{1z}I_{2z} + I_{2z}I_{3z}) \). For the same transfer, we selectively excite only the second spin with the rf field. So \( H_{rf} = \omega_x(t)I_{2x} + \omega_y(t)I_{2y} \), where \( \omega_x, \omega_y \) are the components of the rf field (the available controls) in the rotating frame. Using Ehrenfest’s theorem we can derive the coherent evolution of the various ensemble averages under \( H = H_f + H_{rf} \). Next, we need to add the dissipative evolution caused by the relaxation term \( L \). \( L \) has no effect on the longitudinal operators \( (2I_{1z}I_{2z}, 2I_{2z}I_{3z}) \) since we consider the spin diffusion limit, as we mentioned above. It affects the transverse operators
(2I_{1z}I_{2x}, 2I_{2x}I_{3z}), which relax at the same rate \( k \). Putting all this together, we find the following system of equations (the operators appear in the order that they are produced by the coherent evolution)

\[
\frac{d}{dt} \langle 2I_{1z}I_{2x} \rangle = \omega_x \langle 2I_{1z}I_{2y} \rangle - \omega_y \langle 2I_{1z}I_{2x} \rangle ,
\]

\[
\frac{d}{dt} \langle 2I_{1z}I_{2x} \rangle = \omega_y \langle 2I_{1z}I_{2z} \rangle - k \langle 2I_{1z}I_{2x} \rangle - J \langle \sqrt{2}(2I_{1z}I_{2y}I_{3z} + \frac{I_{2y}}{2}) \rangle ,
\]

\[
\frac{d}{dt} \langle 2I_{1z}I_{2y} \rangle = -\omega_x \langle 2I_{1z}I_{2z} \rangle - k \langle 2I_{1z}I_{2y} \rangle + J \langle \sqrt{2}(2I_{1z}I_{2x}I_{3z} + \frac{I_{2x}}{2}) \rangle ,
\]

\[
\frac{d}{dt} \langle \sqrt{2}(2I_{1z}I_{2y}I_{3z} + \frac{I_{2y}}{2}) \rangle = -\omega_x \langle \sqrt{2}(2I_{1z}I_{2x}I_{3z} + \frac{I_{2x}}{2}) \rangle + J \langle 2I_{1z}I_{2x} \rangle
\]

\[-k \langle \sqrt{2}(2I_{1z}I_{2y}I_{3z} + \frac{I_{2y}}{2}) \rangle + J \langle 2I_{2x}I_{3z} \rangle ,
\]

\[
\frac{d}{dt} \langle \sqrt{2}(2I_{1z}I_{2x}I_{3z} + \frac{I_{2x}}{2}) \rangle = \omega_y \langle \sqrt{2}(2I_{1z}I_{2x}I_{3z} + \frac{I_{2x}}{2}) \rangle - J \langle 2I_{1z}I_{2y} \rangle
\]

\[-k \langle \sqrt{2}(2I_{1z}I_{2x}I_{3z} + \frac{I_{2x}}{2}) \rangle - J \langle 2I_{2x}I_{3z} \rangle ,
\]

\[
\frac{d}{dt} \langle \sqrt{2}(2I_{1z}I_{2x}I_{3z} + \frac{I_{2x}}{2}) \rangle = \omega_x \langle \sqrt{2}(2I_{1z}I_{2y}I_{3z} + \frac{I_{2y}}{2}) \rangle - \omega_y \langle 2I_{1z}I_{2x}I_{3z} \rangle
\]

We write these equations in a more clear form, to facilitate inspection. If we rescale time according to \( t_{new} = Jt_{old} \) and we set

\[
x_1 = \langle 2I_{1z}I_{2x} \rangle, \quad y_1 = -\langle 2I_{1z}I_{2y} \rangle, \quad z_1 = \langle 2I_{1z}I_{2z} \rangle \quad (5.7)
\]

\[
x_2 = \langle \sqrt{2}(2I_{1z}I_{2x}I_{3z} + \frac{I_{2x}}{2}) \rangle, \quad y_2 = \langle \sqrt{2}(2I_{1z}I_{2y}I_{3z} + \frac{I_{2y}}{2}) \rangle, \quad z_2 = \langle \sqrt{2}(2I_{1z}I_{2x}I_{3z} + \frac{I_{2x}}{2}) \rangle \quad (5.8)
\]

\[
x_3 = -\langle 2I_{2x}I_{3z} \rangle, \quad y_3 = \langle 2I_{2y}I_{3z} \rangle, \quad z_3 = \langle 2I_{2z}I_{3z} \rangle \quad (5.9)
\]

\[
\xi = \frac{k}{J}, \quad \Omega_x = \frac{\omega_x}{J}, \quad \Omega_y = \frac{\omega_y}{J} \quad (5.10)
\]

then the above evolution equations take the form (by abuse of notation we use the same \( t \) for the new time)

\[
\dot{x}_1 = -\Omega_x y_1 - \Omega_y x_1 \label{eq:5.11}
\]

\[
\dot{x}_1 = \Omega_y z_1 - \xi x_1 - y_2 \label{eq:5.12}
\]

\[
\dot{y}_1 = \Omega_x z_1 - \xi y_1 - x_2 \label{eq:5.13}
\]
5.3 CONVENTIONAL APPROACH-COMMENTS-UPPER BOUND FOR THE EFFICIENCY

\[
\begin{align*}
\dot{y}_2 &= -\Omega_x x_2 + x_1 - \xi y_2 - x_3, \\
\dot{x}_2 &= \Omega_y z_2 + y_1 - \xi x_2 - y_3, \\
\dot{z}_2 &= \Omega_x y_2 - \Omega_y x_2, \\
\dot{x}_3 &= -\Omega_y z_3 - \xi x_3 + y_2, \\
\dot{y}_3 &= -\Omega_x z_3 - \xi y_3 + x_2, \\
\dot{z}_3 &= \Omega_x y_3 + \Omega_y x_3.
\end{align*}
\] (5.14) (5.15) (5.16) (5.17) (5.18) (5.19)

Careful inspection of equations (5.11)-(5.19) shows that there is a characteristic pathway for transfer $z_1 \rightarrow z_3$ ($2I_1 z_2 \rightarrow 2I_2 z_3$). This path is $z_1 \rightarrow x_1 \rightarrow y_2 \rightarrow x_3 \rightarrow z_3$ and it is encountered when $\Omega_x = 0, \Omega_y \neq 0$. The corresponding equations are (in matrix form)

\[
\begin{bmatrix}
\dot{z}_1 \\
\dot{x}_1 \\
\dot{y}_2 \\
\dot{x}_3 \\
\dot{z}_3
\end{bmatrix} =
\begin{bmatrix}
0 & -\Omega_y & 0 & 0 & 0 \\
\Omega_y & -\xi & -1 & 0 & 0 \\
0 & 1 & -\xi & -1 & 0 \\
0 & 0 & 1 & -\xi & -\Omega_y \\
0 & 0 & 0 & \Omega_y & 0
\end{bmatrix}
\begin{bmatrix}
z_1 \\
x_1 \\
y_2 \\
x_3 \\
z_3
\end{bmatrix}
\] (5.20)

We can shortly describe this evolution. Using the rf field $\Omega_y$ we rotate $z_1$ to $x_1$. As we can see from the above equations, $x_1$ is rotated to $y_2$ and this to $x_3$. Finally, using again $\Omega_y$, we rotate $x_3$ to $z_3$. Our goal is to find that $\Omega_y(t)$ which maximizes the transfer $z_1 \rightarrow z_3$ for the evolution equation (5.20), starting from the point $(z_1, x_1, y_2, x_3, z_3) = (1, 0, 0, 0, 0)$. We would also like to find the maximum achievable value of $z_3$.

5.3 Conventional Approach-Comments-Upper Bound for the Efficiency

The conventional method used for doing this transfer is the INEPT (Insensitive Nuclei Enhanced by Polarization Transfer) pulse sequence. We describe it briefly. At time $t = 0^-$ we apply a $90^\circ$ hard pulse in the $y$ direction, which rotates instantaneously $z_1$ to $x_1$. The initial condition for $x_1$ changes from $x_1(0^-) = 0$ to $x_1(0^+) = 1$. Next, we let the system evolve freely. As we can see from (5.20), this evolution is governed by the equations

\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{y}_2 \\
\dot{x}_3
\end{bmatrix} =
\begin{bmatrix}
-\xi & -1 & 0 \\
1 & -\xi & -1 \\
0 & 1 & -\xi
\end{bmatrix}
\begin{bmatrix}
x_1 \\
y_2 \\
x_3
\end{bmatrix}
\] (5.21)
Using the initial condition \((x_1, y_2, x_3) = (1, 0, 0)\), we find that
\[
x_3(t) = \frac{1}{2} e^{-\xi t}[1 - \cos(\sqrt{2}t)].
\]
(5.22)

The value of \(x_3\) is maximized at the time \(t = t_m\) which satisfies
\[
\frac{dx_3}{dt} \bigg|_{t = t_m} = 0 \quad \Rightarrow \quad \sin(\sqrt{2}t_m) = \xi[1 - \cos(\sqrt{2}t_m)]
\]
(5.23)

At time \(t = t_m^-\) we apply a second 90° hard pulse in the \(y\) direction, which rotates instantaneously \(x_3\) to \(z_3\). The INEPT pulse sequence is completed and the transfer efficiency is
\[
\eta_l = z_3(t_m^+) = x_3(t_m^-) = \frac{1}{2} e^{-\xi t_m}[1 - \cos(\sqrt{2}t_m)],
\]
(5.24)

where \(t_m\) is calculated from (5.23).

The basic idea of our approach to the problem is to use the rf field \(\Omega_y(t)\) to rotate gradually \(z_1\) to \(x_1\) and \(x_3\) to \(z_3\), instead of using hard pulses, pretty much as we did for the two spin system in chapter 3. This method is better than the conventional one presented above. For the particular case that we study here, we can explain this using the system equation (5.20). Suppose that the available time \(T\) for the application of the rf field is specified. If we choose \(\Omega_y(t)\) to make the rotations \(z_1 \to x_1\), \(x_3 \to z_3\) fast enough, like in the conventional method, the intermediate transverse variables \(x_1, y_2\) and \(x_3\) are built quickly. As we can observe from (5.20), these variables are exposed to the phenomenon of relaxation which limits the performance. But if we choose \(\Omega_y(t)\) to make these rotations in the appropriate rate, then we can accomplish the desired transfer by building the intermediate transverse variables gradually, thus reducing the relaxation losses.

The maximum achievable value of \(z_3\) corresponds to the limiting case where the duration \(T\) of the experiment approaches infinity. This can be easily proved. Suppose that we have two available times \(T_1\) and \(T_2\), with \(T_1 < T_2\). Suppose also that we have found an optimal policy \(P_1\) (the optimal control \(\Omega_y(t)\)) for the time interval \([0, T_1]\), which gives us a transfer efficiency \(z_3(T_1) = \eta_1\). For the available time \(T_2\), if we apply the policy \(P_1\) for \(t \in [0, T_1]\) and then set \(\Omega_y(t) = 0\) for \(t \in (T_1, T_2]\), we get the same efficiency \(\eta_1\). Of course, we can find an optimal policy \(P_2\) for the whole time interval \([0, T_2]\). Let \(\eta_2\) be the transfer efficiency under this policy. Obviously it is \(\eta_2 \geq \eta_1\). Thus the longer the available time the better efficiency we may get. The physical explanation is that the longitudinal (\(z\)) direction is protected against relaxation. The maximum transfer efficiency is achieved in the limiting case where the available time approaches infinity. In control language, we need to consider our problem as an infinite horizon problem.
Finding an analytic solution for this optimal control problem is not an easy task. The difficulty comes from the fact that the control Hamiltonian for system (5.20) is linear in the control variable $\Omega_y$ and, additionally, this variable is unbounded. Nevertheless, we can compute an analytic upper bound for the maximum achievable value of $z_3$, using the theory presented in the preceding chapter. Consider the system

$$
\begin{bmatrix}
\dot{z}_1 \\
\dot{z}_2 \\
\dot{z}_3 \\
\dot{x}_1 \\
\dot{y}_2 \\
\dot{x}_3
\end{bmatrix}
= 
\begin{bmatrix}
0 & 0 & 0 & -v_1 & 0 & 0 \\
0 & 0 & 0 & 0 & -v_2 & 0 \\
0 & 0 & 0 & 0 & 0 & -v_3 \\
v_1 & 0 & 0 & -\xi & -1 & 0 \\
0 & v_2 & 0 & 1 & -\xi & -1 \\
0 & 0 & v_3 & 0 & -1 & -\xi
\end{bmatrix}
\begin{bmatrix}
z_1 \\
z_2 \\
z_3 \\
x_1 \\
y_2 \\
x_3
\end{bmatrix}.
$$

(5.25)

Observe that system (5.25) reduces to system (5.20) for $v_1 = -v_3 = \Omega_y$ and $v_2 = 0$. Thus, if we know the maximum achievable value of $z_3$ starting from $(1, 0, 0, 0, 0, 0)$ and evolving under system (5.25), then this is an upper bound for the maximum achievable value of $z_3$ with evolution described by the original system (5.20). In chapter 4 we calculated this bound $\upsilon$ analytically, Eq. (4.42), which we rewrite here for convenience

$$
\upsilon = \left(\frac{\sqrt{\xi^2 + 2} - \xi}{2}\right)^2.
$$

(5.26)

We emphasize that this is an upper bound for the efficiency; we cannot achieve $\upsilon$ using the real system (5.20). What we can do is to calculate numerically, for each value of $\xi$, a rf field $\Omega_y$ that, when applied to the real system (5.20), gives a value of $z_3$ approaching this upper bound. This is the subject of the next section.

## 5.4 Numerical Calculation of the Optimal RF Field

Having established an analytical upper bound (5.26) for the efficiency, we now try to find numerically a rf field $\Omega_y(t)$ that approaches this bound, for each value of the parameter $\xi$. The methods that we employ in this section use the equations (5.20) of the original system.

At first, we use a numerical optimization method based on a steepest descent algorithm. For the application of the method we use a finite time window $T$, although the maximum efficiency is achieved for $T \to \infty$, as mentioned above. Using a big enough $T$ is sufficient, since the major improvement in the efficiency by the use of gradual rotations for systems with only transverse relaxation present is achieved for finite times, see for example Khaneja...
et al. (2003). For the values of normalized relaxation $\xi$ that we are interested in ($0 - 1$), a time interval $T = 10$ (normalized time units) is enough. For larger values of $\xi$ we can use even shorter $T$. The optimal rf field $\Omega_y(t)$ that we find with this method, for various $\xi$, is shown in Fig. 5.4. Note that as $\xi$ increases, the optimal pulse, calculated for a fixed time window, becomes shorter in time and acquires a larger peak value. The reason for this is that when the available time is finite, then the transfer $z_1 \rightarrow z_3$ should be done faster for larger $\xi$, in order to reduce the increased relaxation losses. Now observe that the optimal pulse shape can be very well approximated by a Gaussian profile of the form

$$\Omega_y(t) = A \exp \left[ - \left( \frac{t - T/2}{\sqrt{2} \sigma} \right)^2 \right],$$  \hspace{1cm} (5.27)$$

with $A, \sigma$ appropriately chosen. As a result, the efficiency that we find using the appropriate Gaussian pulse is very close to that we find using the original pulse. This suggests that instead of using the initial numerical optimization method, we can use Gaussian pulses of the form (5.27), optimized with respect to $A$ and $\sigma$ for each value of $\xi$. The optimal $A, \sigma$ are found by numerical simulations. For each $\xi$ we simulate the equations of system (5.20) with $\Omega_y(t)$ given by (5.27), for many values of $A$ and $\sigma$. We choose those values that give the maximum $z_3(T)$. In Table 5.1 we show the optimal $A, \sigma$, for various values $\xi \in [0, 1]$. We also show the corresponding efficiency, as well as the efficiency achieved by the initial numerical optimization method. Observe how close lie these two groups of values, especially for large relaxation rates ($\xi > 0.5$). The choice of the Gaussian shape is indeed successful.

Fig. 5.5 shows the efficiency of the conventional method (INEPT), the efficiency of our method (SPORTS ROPE, SPin ORder TranSfer with Relaxation Optimized Pulse Element) and the upper bound $\nu$ from (5.26), for the values of the relaxation parameter $\xi$ shown in Table 5.1. Note that for large $\xi$ (large relaxation rates), SPORTS ROPE gives a significant improvement over INEPT.

In Fig. 5.6(a) we plot the time evolution of the various transfer functions (operators) that participate in the transfer $z_1 \rightarrow z_3$, when the optimal Gaussian pulse for $\xi = 1$, shown in Fig. 5.4(c), is applied to system (5.20). Observe the gradual building of the intermediate variables $x_1, y_2$ and $x_3$. In Fig. 5.6(b) we plot the angle $\theta_3 = \tan^{-1} (z_3/x_3)$ of the vector $r_3 = x_3 \hat{x} + z_3 \hat{z}$ with the $x$ axis, as a function of time. Observe that initially $r_3$ is parallel to $x$ axis ($\theta = 0$), but under the action of the Gaussian pulse is rotated gradually to $z$ axis ($\theta = \pi/2$). The gradual rotation of $x_3$ to $z_3$ is the spirit of SPORTS ROPE.

We remark that for the general transfer $I_{i_z} \rightarrow I_{nz}$, more than one intermediate steps $2I_{i-1_z}I_{iz} \rightarrow 2I_{iz}I_{i+1_z}$ are necessary. Since the equations that describe the $i^{th}$ transfer
Figure 5.4: Optimal pulse (dashed line) calculated using a numerical optimization method based on a steepest descent algorithm, for various values of the normalized relaxation parameter $\xi$. The Gaussian pulse (solid line) approximates very well the optimal pulse shape and gives a similar efficiency. This suggests that instead of using the initial numerical optimization method, we can use Gaussian pulses of the form (5.27), optimized with respect to $A$ and $\sigma$ for each value of $\xi$. 

\begin{align*}
\xi &= 0 \\
\xi &= 0.5 \\
\xi &= 1 \\
\xi &= 2 \\
\xi &= 5 \\
\xi &= 10
\end{align*}
Table 5.1: For various values of $\xi \in [0, 1]$, the optimal values of $A, \sigma$ and the corresponding efficiency are shown. We present also for comparison the efficiency achieved by the steepest descent method.

<table>
<thead>
<tr>
<th>$\xi$</th>
<th>$A$</th>
<th>$\sigma$</th>
<th>Gaussian Pulse</th>
<th>Steepest Descent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>1.11</td>
<td>1.30</td>
<td>0.2510</td>
<td>0.2512</td>
</tr>
<tr>
<td>0.95</td>
<td>1.09</td>
<td>1.32</td>
<td>0.2661</td>
<td>0.2662</td>
</tr>
<tr>
<td>0.90</td>
<td>1.07</td>
<td>1.34</td>
<td>0.2824</td>
<td>0.2825</td>
</tr>
<tr>
<td>0.85</td>
<td>1.05</td>
<td>1.36</td>
<td>0.3000</td>
<td>0.3001</td>
</tr>
<tr>
<td>0.80</td>
<td>1.03</td>
<td>1.38</td>
<td>0.3190</td>
<td>0.3191</td>
</tr>
<tr>
<td>0.75</td>
<td>1.02</td>
<td>1.39</td>
<td>0.3396</td>
<td>0.3397</td>
</tr>
<tr>
<td>0.70</td>
<td>1.00</td>
<td>1.41</td>
<td>0.3619</td>
<td>0.3620</td>
</tr>
<tr>
<td>0.65</td>
<td>0.98</td>
<td>1.43</td>
<td>0.3861</td>
<td>0.3863</td>
</tr>
<tr>
<td>0.60</td>
<td>0.97</td>
<td>1.44</td>
<td>0.4124</td>
<td>0.4126</td>
</tr>
<tr>
<td>0.55</td>
<td>0.96</td>
<td>1.44</td>
<td>0.4410</td>
<td>0.4413</td>
</tr>
<tr>
<td>0.50</td>
<td>0.95</td>
<td>1.44</td>
<td>0.4721</td>
<td>0.4726</td>
</tr>
<tr>
<td>0.45</td>
<td>0.94</td>
<td>1.45</td>
<td>0.5060</td>
<td>0.5067</td>
</tr>
<tr>
<td>0.40</td>
<td>0.93</td>
<td>1.46</td>
<td>0.5428</td>
<td>0.5439</td>
</tr>
<tr>
<td>0.35</td>
<td>0.92</td>
<td>1.46</td>
<td>0.5830</td>
<td>0.5846</td>
</tr>
<tr>
<td>0.30</td>
<td>0.91</td>
<td>1.46</td>
<td>0.6270</td>
<td>0.6292</td>
</tr>
<tr>
<td>0.25</td>
<td>0.90</td>
<td>1.47</td>
<td>0.6750</td>
<td>0.6780</td>
</tr>
<tr>
<td>0.20</td>
<td>0.89</td>
<td>1.48</td>
<td>0.7277</td>
<td>0.7315</td>
</tr>
<tr>
<td>0.15</td>
<td>0.88</td>
<td>1.48</td>
<td>0.7855</td>
<td>0.7900</td>
</tr>
<tr>
<td>0.10</td>
<td>0.85</td>
<td>1.52</td>
<td>0.8494</td>
<td>0.8536</td>
</tr>
<tr>
<td>0.05</td>
<td>0.79</td>
<td>1.60</td>
<td>0.9203</td>
<td>0.9232</td>
</tr>
<tr>
<td>0.00</td>
<td>0.73</td>
<td>1.71</td>
<td>0.9999</td>
<td>1.0000</td>
</tr>
</tbody>
</table>
§5.4 NUMERICAL CALCULATION OF THE OPTIMAL RF FIELD

Figure 5.5: Efficiency for the conventional method (INEPT) and for our method (SPORTS ROPE), for the values of $\xi$ shown in Table 5.1. We also plot the upper bound (5.26) for the efficiency.

Figure 5.6: (a) Time evolution of the transfer functions (operators) that participate in the transfer $z_1 \rightarrow z_3$, for $\xi = 1$ and the optimal Gaussian pulse shown in Fig. 5.4(c). (b) The angle $\theta_3 = \tan^{-1}(z_3/x_3)$ of the vector $r_3 = x_3\hat{x} + z_3\hat{z}$ with the $x$ axis, as a function of time. Observe that initially $r_3$ is parallel to $x$ axis ($\theta = 0$), but under the action of the Gaussian pulse is rotated gradually to $z$ axis ($\theta = \pi/2$).
are the same as (5.20), we just need to apply the same Gaussian pulse but centered, in the
frequency domain, at the resonance frequency of spin $i$. In this sequence of Gaussian pulses
we should add at the beginning and at the end the optimal pulses for the first and the final
step, respectively, see Fig. 5.2. These pulses can be calculated using the theory presented
in chapter 3.

5.5 The General Case with Arbitrary Coupling Constants

In this section we study transfer (5.6) but for the general form of the free evolution Hamil-
tonian $H_f$

$$H_f = 2J_1I_{1z}I_{2z} + 2J_2I_{2z}I_{3z}.$$  \hfill (5.28)

If we work as in section 5.2 with $H = H_f + H_{rf}$, $H_{rf} = \omega_xI_{2x} + \omega_yI_{2y}$, we can derive
evolution equations for the average values of the various operators participating in the
transfer. Again, a characteristic pathway can be identified in the resulting set of equations.
As before, this path is encountered when $\Omega_x = 0, \Omega_y \neq 0$. The corresponding equations in
matrix form are

$$
\begin{bmatrix}
\dot{z}_1 \\
\dot{x}_1 \\
\dot{y}_2 \\
\dot{y}_3 \\
\dot{x}_4 \\
\dot{z}_4
\end{bmatrix} =
\begin{bmatrix}
0 & -\Omega_y & 0 & 0 & 0 & 0 \\
\Omega_y & -1 & -\zeta_1 & -\zeta_2 & 0 & 0 \\
0 & \zeta_1 & -1 & 0 & -\zeta_2 & 0 \\
0 & \zeta_2 & 0 & -1 & -\zeta_1 & 0 \\
0 & 0 & \zeta_2 & \zeta_1 & -1 & -\Omega_y \\
0 & 0 & 0 & 0 & \Omega_y & 0
\end{bmatrix}
\begin{bmatrix}
z_1 \\
x_1 \\
y_2 \\
y_3 \\
x_4 \\
z_4
\end{bmatrix},
$$

(5.29)

where

$$x_1 = \langle 2I_{1z}I_{2z} \rangle, \quad z_1 = \langle 2I_{1z}I_{2z} \rangle, \quad (5.30)$$

$$y_2 = \langle I_{2y} \rangle, \quad y_3 = \langle 4I_{1z}I_{2y}I_{3z} \rangle, \quad (5.31)$$

$$x_4 = -\langle 2I_{2x}I_{3z} \rangle, \quad z_4 = \langle 2I_{2x}I_{3z} \rangle, \quad (5.32)$$

$$\zeta_1 = \frac{J_1}{k}, \quad \zeta_2 = \frac{J_2}{k}, \quad \Omega_y = \frac{\omega_y}{k}. \quad (5.33)$$

Note that in system (5.29) there is one more equation than in system (5.20). This happens
because the term $y_2 = \langle \sqrt{2}(2I_{1z}I_{2y}I_{3z} + I_{2y}/2) \rangle$ in (5.20) splits into two terms, $y_2 = \langle I_{2y} \rangle$
and $y_3 = \langle 4I_{1z}I_{2y}I_{3z} \rangle$ in (5.29), which evolve differently for $\zeta_1 \neq \zeta_2$. 


An upper bound for the maximum achievable value of $z_4$, starting from $(z_1, x_1, y_2, y_3, x_4, z_4) = (1, 0, 0, 0, 0, 0)$, can be found by using the theory presented in chapter 4. Consider the system

$$\begin{bmatrix}
\dot{z}_1 \\
\dot{z}_2 \\
\dot{z}_3 \\
\dot{z}_4 \\
\dot{x}_1 \\
\dot{y}_2 \\
\dot{y}_3 \\
\dot{x}_4
\end{bmatrix} = \begin{bmatrix}
0 & 0 & 0 & 0 & -v_1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -v_2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -v_3 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -v_4 \\
v_1 & 0 & 0 & 0 & -1 & -\zeta_1 & -\zeta_2 & 0 \\
v_2 & 0 & v_3 & 0 & \zeta_2 & 0 & -1 & -\zeta_1 \\
v_3 & 0 & 0 & 0 & \zeta_2 & 0 & -1 & -\zeta_1 \\
v_4 & 0 & 0 & 0 & \zeta_2 & \zeta_1 & -1 & 0
\end{bmatrix} \begin{bmatrix}
z_1 \\
z_2 \\
z_3 \\
z_4 \\
x_1 \\
y_2 \\
y_3 \\
x_4
\end{bmatrix},$$

(5.34)

Observe that system (5.34) reduces to system (5.29) for $v_1 = -v_4 = \Omega_y, v_2 = v_3 = 0$. Thus, if we know the maximum achievable value of $z_4$ starting from $(1, 0, 0, 0, 0, 0, 0, 0)$ and evolving under system (5.34), then this is an upper bound for the maximum achievable value of $z_4$ with evolution described by the original system (5.29). In chapter 4 we showed that the maximum efficiency for systems like (5.34) can be found by solving an appropriate semidefinite program. We describe the semidefinite program that we need to solve for the particular case that we examine. Let us define the following matrices

$$A_1 = \begin{bmatrix}
-2 & -\zeta_1 & -\zeta_2 & 0 \\
-\zeta_1 & 0 & 0 & 0 \\
-\zeta_2 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}, \quad A_2 = \begin{bmatrix}
0 & \zeta_1 & 0 & 0 \\
\zeta_1 & -2 & 0 & -\zeta_2 \\
0 & 0 & 0 & 0 \\
0 & -\zeta_2 & 0 & 0
\end{bmatrix},$$

$$A_3 = \begin{bmatrix}
0 & 0 & \zeta_2 & 0 \\
0 & 0 & 0 & 0 \\
\zeta_2 & 0 & -2 & -\zeta_1 \\
0 & 0 & -\zeta_1 & 0
\end{bmatrix}, \quad A_4 = \begin{bmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & \zeta_2 & 0 \\
0 & 0 & 0 & \zeta_1 \\
0 & \zeta_2 & \zeta_1 & -2
\end{bmatrix}.$$

Using Theorem 1 of the preceding chapter, we easily see that the semidefinite program of interest is the following: Find

$$\mathcal{E} = \max_M \langle A_n, M \rangle$$

subject to

$$\langle A_1, M \rangle = -1, \quad \langle A_2, M \rangle = 0, \quad \langle A_3, M \rangle = 0$$

and $M$ positive semidefinite symmetric $4 \times 4$ matrix. The maximum achievable value of $z_4$ for system (5.34) is $\nu = \sqrt{\mathcal{E}}$ and this is an upper bound for the corresponding value for system (5.29).
§5.5 THE GENERAL CASE WITH ARBITRARY COUPLING CONSTANTS

Table 5.2: For $\zeta_2 = 0.5$ fixed we have calculated the upper bound $\nu$ for increasing values of $\zeta_1$. Observe that after $\zeta_1 = 1.5$, $\nu$ becomes insensitive in the further increase of $\zeta_1$.

<table>
<thead>
<tr>
<th>$\zeta_1$</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1</th>
<th>1.5</th>
<th>2</th>
<th>2.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\zeta_2$</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>$\nu$</td>
<td>0.1715</td>
<td>0.1934</td>
<td>0.2100</td>
<td>0.2220</td>
<td>0.2300</td>
<td>0.2345</td>
<td>0.2360</td>
<td>0.2360</td>
<td>0.2360</td>
</tr>
</tbody>
</table>

This semidefinite program can be solved by using some appropriate software package, for example SDPT-3, Toh et al. (1999). As an illustration, we consider the case $J_1 = J_2 = k/\sqrt{2}$ ($\zeta_1 = \zeta_2 = 1/\sqrt{2}$). This corresponds to $J = k$ ($\xi = 1$) for system (5.20). We find $\mathcal{E} = 0.0718 \Rightarrow \nu = \sqrt{\mathcal{E}} = 0.2679$. On the other hand, using formula (5.26) with $\xi = 1$ we find $\nu = 0.2679$, i.e. the same value. Calculating numerically the upper bound for various values of $\zeta_1$ and $\zeta_2$, leads us to the following very important observation. If we keep $\zeta_2$ fixed and calculate the upper bound $\nu$ for increasing values of $\zeta_1$, we find that after some point the further increase of $\zeta_1$ has no effect to the value of $\nu$. This is shown in Table 5.2.

For the general case $J_1 \neq J_2$ and working analogously with the case $J_1 = J_2 = J$, we can calculate the efficiency accomplished by the conventional method, INEPT. We can also use a steepest descent algorithm to calculate numerically the optimal pulse $\Omega_y(t)$ and compare the efficiency achieved with the upper bound obtained by solving the corresponding semidefinite program. About the optimal pulse shape, we find after numerical investigation that if the values of $\zeta_1$ and $\zeta_2$ are relatively close to each other then again it can be very well approximated by a Gaussian profile of the form (5.27). Otherwise, if $\zeta_1$ is much different than $\zeta_2$, the optimal pulse shape is more complicated. It develops wings and we have to use more complicated profiles to approximate it, for example the Mexican hat profile

$$\Omega_y(t) = A \left[ 1 - \left( \frac{t - T/2}{b} \right)^2 \right] \exp \left[ - \left( \frac{t - T/2}{\sqrt{2}\sigma} \right)^2 \right].$$

(5.35)

For $J_1 = k$, $J_2 = 0.5k$ ($\zeta_1 = 1, \zeta_2 = 0.5$) the Gaussian profile is a good approximation to the optimal pulse shape. In Fig. 5.7(a) we plot the optimal Gaussian pulse and in Fig. 5.7(b) the time evolution of the various transfer functions under this pulse. For $J_1 = 2k$, $J_2 = 0.5k$ ($\zeta_1 = 2, \zeta_2 = 0.5$) the optimal pulse shape is more complicated and a Mexican hat profile gives a better approximation then a Gaussian. The optimal Mexican hat pulse is shown in Fig. 5.8(a) and the corresponding evolution of the transfer functions in Fig. 5.8(b). Note that for both cases function $y_3$ is developed less than function $y_2$. This happens because
Figure 5.7: (a) The Gaussian pulse that approximates the optimal pulse shape for $J_1 = k, J_2 = 0.5k$ ($\zeta_1 = 1, \zeta_2 = 0.5$). (b) The corresponding time evolution of the transfer functions. Note that $y_3$ and $x_4$ have a similar time evolution.

Figure 5.8: (a) The Mexican hat pulse that approximates the optimal pulse shape for $J_1 = 2k, J_2 = 0.5k$ ($\zeta_1 = 2, \zeta_2 = 0.5$). (b) Time evolution of the corresponding transfer functions. Note that the dashed line corresponds to $y_3$ while the dotted line to $x_4$.

Table 5.3: For the two cases with $J_1 \neq J_2$ examined in the text, we show the optimal values of $A, \sigma, b$, the efficiency achieved by the conventional method (INEPT), our method (SPORTS ROPE), and the upper bound calculated using semidefinite programming.

<table>
<thead>
<tr>
<th>$\zeta_1$</th>
<th>$\zeta_2$</th>
<th>$A$</th>
<th>$\sigma$</th>
<th>$b$</th>
<th>INEPT</th>
<th>SPORTS ROPE</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>0.5</td>
<td>1.32</td>
<td>1.07</td>
<td>-</td>
<td>0.1593</td>
<td>0.2214</td>
<td>0.2345</td>
</tr>
<tr>
<td>2.0</td>
<td>0.5</td>
<td>2.40</td>
<td>0.85</td>
<td>1.12</td>
<td>0.1755</td>
<td>0.2230</td>
<td>0.2360</td>
</tr>
</tbody>
</table>
$y_3$ is connected to $x_1$ by the smallest coupling $\zeta_2$ while is connected to $x_4$ by the largest coupling $\zeta_1$. As a result, $y_3$ is built less and depleted faster, that’s why is developed less than $y_2$ in the course of time. If the couplings were equal, then $y_2$ and $y_3$ would have the same time evolution and if $\zeta_1 < \zeta_2$ then $y_3$ would be developed more than $y_2$. In Table 5.3 we show the values of the parameters ($A$, $\sigma$ and $b$) used for the two cases, the efficiency achieved by the conventional method (INEPT), our method (SPORTS ROPE), and the upper bound calculated using semidefinite programming. Note that although the efficiency of SPORTS ROPE and the upper bound are pretty much the same for the two cases (see the observation in the preceding paragraph), the efficiency of INEPT is different. It is larger for the second case, because INEPT exploits better the larger $\zeta_1$.

We conclude this chapter by noting that the general method for calculating efficiency upper bounds using semidefinite programming, can be extended to the more general case where interference phenomena between the various relaxation mechanisms are taken into account.
Chapter 6

Conclusions and Future Work

In this thesis we applied the principles of optimal control theory to calculate upper bounds for the polarization-coherence transfer in coupled spin dynamics, under the presence of general relaxation mechanisms. To demonstrate the basic ideas we used two specific quantum systems, originating from liquid state NMR spectroscopy: A coupled spin pair and an Ising spin chain. For these systems, the dominant relaxation mechanisms were dipole-dipole interaction and chemical shift anisotropy. We also calculated optimal pulse sequences (radiofrequency magnetic field pulses) that achieve or approach the derived upper bounds. New transfer schemes were found, which yield substantial gains in transfer efficiency, compared to the traditionally used INEPT pulse sequence. From the perspective of mathematical control theory, the problems related to calculating the maximum transfer efficiency motivated the study of a class of dissipative bilinear control systems. For these systems it was shown that the optimal solution and the reachable set can be found by solving a semidefinite program.

The above methods can also be applied to other important liquid state NMR systems. These include the $I_2S$ and $I_3S$ spin systems ($I$ can be for example $^1H$, $S$ can be $^{13}C$). Combining the results for these systems and for the systems presented in the thesis, we can design relaxation optimized pulse sequences for larger spin systems, as those commonly encountered in backbone and side chain assignments in protein NMR spectroscopy. A substantial increase in sensitivity of the experiments is expected, leading to better signal to noise ratio in protein NMR spectroscopy and better conclusions about the protein structure. Note that before applying these pulse elements to real proteins, it is necessary to make them broadband (this requirement comes from the fact that NMR experiments deal with an ensemble of systems, with variation in their internal parameters). Some recent results
in this direction can be found in Khaneja et al. (2004) and Fruh et al. (2005).

The methods developed here are not limited to NMR applications but is expected to find applications in the coherent control of other quantum mechanical phenomena in the presence of dissipation and decoherence. They are useful for answering important questions in quantum information theory. It is a fundamental problem to understand the extent to which an open quantum system can be controlled, i.e. where all the state of a quantum-mechanical systems can be steered in presence of relaxation? How much entanglement can be produced in presence of decoherence and dissipation and what is the optimal way to synthesize unitary gates in open quantum systems so as to maximize their fidelity? All these problems are related to optimal control of quantum-mechanical systems in presence of relaxation.
Bibliography


