Chapter4: Quantum Control

We now consider control systems of the form

$$\dot{x} = (\Omega_0 + \sum_i u_i \Omega_i)x,$$

Where $\Omega_0, \Omega_i \in \mathfrak{g}$ the Lie algebra of a Lie Group G. In quantum control $G = SU(n)$, special unitary matrices and $\mathfrak{g} = su(n)$ skew Hermitian matrices. Then we can write above equation as

$$\dot{U} = -i(H_0 + \sum_i u_i H_i)U,$$

where H_0, H_i are traceless Hermitian matrices. They are also called Hamiltonians. When we turn on our controls, we say we switch on our Hamiltonians.

Let us take the simplest example $G = SU(2)$ and $\mathfrak{g} = su(2)$. Recall dimension of $su(n)$ is $n^2 - 1$ and $su(2)$ is 3. The generators of $su(2)$ are $\{-i\sigma_x, -i\sigma_y, -i\sigma_z\}$ where

$$\sigma_x = \frac{1}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix};\ \sigma_y = \frac{1}{2}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix};\ \sigma_z = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{1}$$

$\sigma_x, \sigma_y, \sigma_z$ are called Pauli matrices. They are traceless Hermitian.

Consider the control system

$$\dot{U} = -i(\omega_0 \sigma_z + u(t)\sigma_x + v(t)\sigma_y)U,$$

U evolves on $SU(2)$. This system arises when we study dynamics of a spin in magnetic field. It appears in subject of NMR spectroscopy. Lets try to understand the physics of it. You are familiar with earth spinning on its axis. This gives earth a angular momentum. Now imagine our earth was charged. Then spinning will give earth a magnetic moment. Imagine a loop of wire carrying current (circulating charge), then it has a magnetic moment $M = I.A$, where I is the current and A area of the loop, from your basic physics. Now imagine a charge q going around in a loop of radius r, with angular velocity ω. Then it makes $\frac{\omega}{2\pi}$ rotations per sec. The current is then $\frac{q\omega}{2\pi}$ and its magnetic moment is $M = \frac{q\omega\pi r^2}{2\pi} = \frac{q}{2m}(mvr)$ where $l = mvr$ is the angular momentum. Then $M = \frac{q}{2m}L$, the ratio $\gamma = \frac{q}{2m}$ is called the gyromagnetic ratio, it relates angular momentum to magnetic moment.

Now suppose we have our charged spinning earth and we apply a magnetic field $B = (B_x, B_y, B_z)$, then $M = (m_x, m_y, m_z) = \gamma(l_x, l_y, l_z)$ will experience a torque. This torque is $M \times B$, and changes the angular momentum as

$$\dot{L} = M \times B.$$

Relating $M = \gamma L$, we have,

$$\dot{M} = \gamma M \times B = -\gamma(B_z\Omega_z + B_x\Omega_y + B_z\Omega_z)M, \tag{2}$$

where $\Omega_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$, $\Omega_y = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}$ and $\Omega_z = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ are generator of rotation. For reasons that will become clear as we go on denote $\omega_0 = -\gamma B_z$ and $u = -\gamma B_x$ and $v = -\gamma B_y$ and we get,

$$\dot{M} = (\omega_0\Omega_z + u\Omega_x + v\Omega_y)M,$$

Note $M(t) = \Theta M(0)$ where

$$\dot{\Theta} = (\omega_0\Omega_z + u\Omega_x + v\Omega_y)\Theta, \quad \Theta(0) = I$$

Then $\Theta \in SO(3)$. M rotates in B. It precesses around B.

What concerns us is spin of a atomic nuclie. Many atomic nuclie like hydrogen, carbon, nitrogen have a quantum mechanical property called spin which gives the nucleus a angular momentum and hence magnetic moment. However because of quantum mechanics this angular momentum is quantized. If we measure its value in say z direction, we will only find two values $\frac{\hbar}{2}$ and $-\frac{\hbar}{2}$, spinning up and spinning down. The state of the nucleus is then written as a two dimensional vector which is $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ when spinning up and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ when spinning down. In general the state is defined by a two dimensional complex vector $\psi = \begin{bmatrix} a \\ b \end{bmatrix}$. For the spinning earth, we saw that its magnetic moment precesses in a magnetic field given by Eq. (2). The two dimesnional vector will also precess in a magnetic field with equation given by

$$\dot{\psi} = i\gamma(B_z\sigma_z + B_x\sigma_x + B_y\sigma_y)\psi = -i(\omega_0\sigma_z + u\sigma_x + v\sigma_y)\psi, \tag{3}$$

where in Eq. (2), we have replaced the generator of rotations in real three dimensions $\Omega_x, \Omega_y, \Omega_z$ with $-i\sigma_x, -i\sigma_y, -i\sigma_z$ generator of rotations in complex two dimensions.

The evolution of ψ a two dimensional complex vector is given by $\psi(t) = U\psi(0)$, where

$$\dot{U} = -i(\omega_0\sigma_z + u\sigma_x + v\sigma_y)U, \tag{4}$$

where U is in $SU(2)$.

In practice, in a NMR experiment, we have very large number of atoms of order 10^{23} and each atom/nucleus has a spin state defined by a vector ψ_k, each ψ_k sees same magnetic field and hence evolves according to equation

$$\dot{\psi}_k = -i(\omega_0 \sigma_z + u\sigma_x + v\sigma_y)\psi_k, \tag{5}$$

We can form an average subspace spanned by these ψ_k as $\rho = \frac{1}{N}\sum \psi_k \psi_k^\dagger$, then ρ evolves as

$$\dot{\rho} = [-i(\omega_0 \sigma_z + u\sigma_x + v\sigma_y), \rho] \tag{6}$$

ρ is a two dimensional Hermitian matrix and can be written as

$\rho = \frac{1}{2}I + l_x\sigma_x + l_y\sigma_y + l_z\sigma_z$, where $L = (l_x, l_y, l_z)'$ represents average (x, y, z) angular momentum of the of the ensemble. This average or classical angular momentum evolves as

$$\dot{L} = (\omega_0 \Omega_z + u\Omega_x + v\Omega_y)L, \tag{7}$$

and denoting $M = \gamma L$ we have the same Eq. (2). These are called Bloch equations. Thus we see how evolution of spin state of individual nuclie evloves as two dimensional complex vector and how the average angular momentum and magnetic moment of the spin ensemble evolves as a three dimensional Bloch vector.

Lets think of an ensemble in which all spins are up. Then all $\psi_k = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\rho = \frac{1}{2}I + \sigma_z$. Thus $l_z = 1$ and we have an ensemble with net z angular momentum 1.

Lets think of an ensemble in which all spins are down. Then all $\psi_k = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $\rho = \frac{1}{2}I - \sigma_z$. This $l_z = -1$ and we have an ensemble with net z angular momentum -1.

Lets think of an ensemble in which all spins are $\psi_k = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\rho = \frac{1}{2}I + \sigma_x$. This $l_x = 1$ and we have an ensemble with net x angular momentum 1.

Now lets understand the basic NMR experiment. In an NMR experiment we have spins in a strong magnetic field along say z direction of order $10 - 20$ Tesla. Earths magnetic field is around 10^{-5} tesla. In this magnetic field, up spins have lower energy than down spins and so in thermal equilibrium, we have more spins up. The ratio of up to down spins is given by Boltzmann disprtibution and is $\exp(\frac{\Delta E}{kT})$ where $\Delta E = \mu \cdot B$ is energy difference between down and up spins, which is small, as magnetic moment μ of a nuclear spin is small. Thus at room temperature at such high fields, only 1 in 10^5 spins preferentiably points up. Thus

$\rho = \frac{1}{2}I + \alpha\sigma_z$, where $\alpha \sim 10^{-5}$, none the less the sample has a net angular momentum along z and hence has a net magnetic moment along z direction. Thus in Eq. (2) we start with $M = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. Now we turn on x and y magnetic fields and rotate this vector to $(1, 0, 0)'$.

How this is done will be discussed shortly. But imagine we have rotated M to $(1,0,0)'$ and we switch off u, v in Eq. (2). Then M just rotates around B_0 and we have an evolution $M(t) = (\cos\omega_0 t, \sin\omega_0 t, 0)$. This rotating magnetic moment will induce an emf in a nearby coil with a frequency ω_0 and hence we can measure ω_0. At fields of 14 tesla the ω_0 for hydrogen is 600 MHz, for carbon is 150 MHz, and for nitrogen is 60 MHz. Thus frequency of the induced emf tells us about chemical composition of the sample. This NMR can tell us about composition of the sample. Now we come to the question of how we use u, v to rotate M from $(0,0,1)$ to $(1,0,0)$.

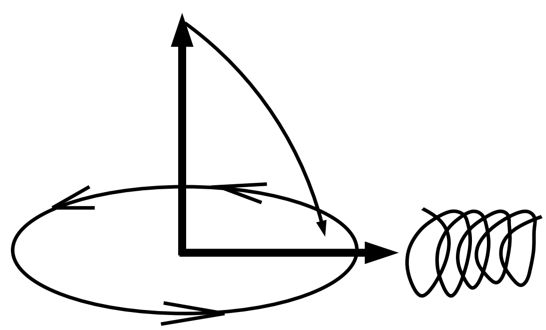


Figure 1: Figure shows how the magnetic moment $M(0)$ along z direction is rotated to transverse plain and it then rotates around z field and induces a EMF in the coil.

In Eq. (2), B_0 is much larger than B_x, B_y which are actually produced by rf-coil. To give an idea if ω_0 is 600 MHz, then u, v are only around 60 kHz. Around 10^5 times smaller. Then we ask how can such small u, v effect a change in $M(0)$. Beacuse suppose we choose $u = 1$ and $v = 0$. Then since ω_0 is 10^5 times u. The Eq. (2) essentially is rotating around z axis. The figure 2 below

shows how the magnetic moment $M(0)$ along z direction just rotates around an axis with a small tilt of z axis when we apply a constant control u. Then a constant control u will not rotates $M(0)$ to transverse plain as desired, because u is too small compared to ω_0. What works and is used is instead a oscillatory control input, $(u, v) = (A\cos\omega_0 t, A\sin\omega_0 t)$, with frequency same as ω_0. To understand how this control works, Eq. (2)

$$\dot{X} = (\omega_0\Omega_z + A\cos\omega_0 t\Omega_x + A\sin\omega_0 t\Omega_y)X, \quad A \ll \omega_0$$

we can write the above equation as

$$\dot{X} = (\omega_0\Omega_z + A\exp(\omega_0 t\Omega_z)\Omega_x\exp(-\omega_0 t\Omega_z))X,$$

Lets make a change of cordinates $Y = \exp(-\omega_0 t\Omega_z)X$, then

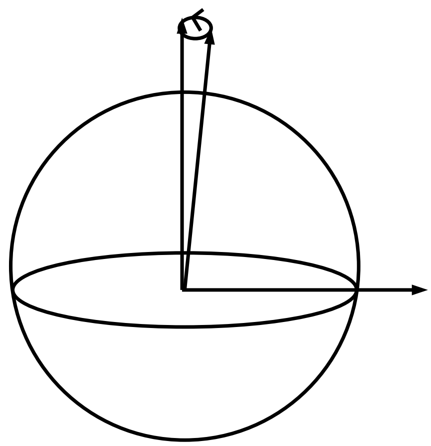


Figure 2: Figure shows how the magnetic moment $M(0)$ along z direction just rotates around an axis with a small tilt of z axis when we apply a constant control u.

$$\dot{Y} = A\Omega_x Y,$$

this is great as ω_0 has disappeared and Y starting from $Y(0) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ rotates to $Y(T) = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$ at $T = \frac{\pi}{2A}$.

Then $X(T) = \exp(\omega_0 T \Omega_z) Y(T)$, a vector on the equator. Thus we have been able to bring the Bloch vector in Eq. (2) to equator by use of an oscillatory controls. This is the first lesson in quantum control. The controls we apply are much weaker compared to drift in the system so constant control laws donot work. We need oscillatory controls. We need to excite the system on resonance.

We said at B_0 of 14 T we have for hydrogen $\omega_0 = 600$ MHz. This is not strictly true. Hydrogen nucleus has electrons around it. These moving/hovering electrons produce local magnetic fields and change the field from B_0 to $B_0(1-\sigma_0)$ and hence ω_0 changes from to $\omega_0(1-\sigma_0) = \omega_0 + \Delta\omega$. This σ_0 is of order few parts per million, i.e. 10^{-6} and hence when $\omega_0 = 600$ Mhz we have $\Delta\omega$ of order few kHz. This σ_0 also called chemical shift is characteristic of a electronic environment of nucleus. We can measure $\Delta\omega$, when we measure frequencies in our EMF. For example in Ethanol molecule we have three hydrogen, each with different chemical environment and hence three different $\Delta\omega$. When we find three different $\Delta\omega$ is our experiment at certain specific values, then we know we have a fingerprint spectrum of Ethanol. This way chemical shifts help us identify the molecules. Not only does NMR give information about the chemical composition but also the chemical shifts can identify compounds.

Now how do we rotate M to equator when we have many $\Delta\omega$.

To understand how control works now, consider Eq. (2)

$$\dot{X} = ((\omega_0 + \Delta\omega)\Omega_z + A\cos(\omega_0 t + \phi)\Omega_x + A\sin(\omega_0 t + \phi)\Omega_y)X, \quad A \ll \omega_0$$

Lets as before make a change of cordinates $Y = \exp(-\omega_0 t \Omega_z)X$, then

$$\dot{Y} = (\Delta\omega\Omega_z + A\cos\phi\Omega_x + A\sin\phi\Omega_y)Y,$$

this is great as before ω_0 has disappeared but $\Delta\omega$ stays and we have to now choose A and ϕ as functions to time so that $Y(0) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ rotates to $Y(T) = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$.

This is an important control problem because we want our control to work for all $\Delta\omega$ in a given range. Size of $\Delta\omega$ and A are comparable. We will study this problem in detail soon. It is called broadband control. At this point it is suffice to believe that we can do the desired maneuver by choice of $A(t)$ and $\phi(t)$.

In summary, we learnt about single spin $\frac{1}{2}$ whose state is a 2 dimensional complex vector evolving as

$$\dot{\psi} = -i(\underbrace{\omega_0\sigma_z}_{H_0} + u\underbrace{\sigma_x}_{H_1} + v\underbrace{\sigma_y}_{H_2})\psi, \tag{8}$$

As we saw this equation evolves as $\psi(t) = U(t)\psi(0)$, where $U(t) \in SU(2)$.

This is the simplest example of a quantum control system, where H_0 is the drift Hamiltonian and H_1, H_2 are control Hamiltonians.

Now as a general rule, we can have a quantum system A of dimension n_1, which means its state is a n_1 dimensional complex vector evolving as

$$\dot{\psi} = -iH\psi,$$

where H is a $n_1 \times n_1$ Hermitian matrix, such that $\psi(t) = U(t)\psi(0) = \exp(-iHt)\psi(0)$, where $-iH$ is skew Hermitian and $U(t) \in SU(n)$.

If we have a quantum system A of dim n_1 and a quantum system B of dim n_2, then when we bring the two systems together and make them interact, we get a a quantum system of dim $n_1 \times n_2$, whose state is a complex vector in a vector space of size $n_1 n_2$, spanned by a basis of the form $e_i \otimes f_j$ where e_i are basis for space A and f_j are basis for space B. A state ψ that can be written as $\psi_a \otimes \psi_b$ is called a separable state, else it has the form $\psi = \sum_{ij} \alpha_{ij} e_i \otimes f_j$ and is called an entangled state.

The hamiltonian for the joint system

$$H = \sum H_a^i \otimes H_b^j,$$

where H_a^i are Hamiltonians for system A and H_b^j are Hamiltonians for system B. Hamiltonians of the form $H_a \otimes I$ and $I \otimes H_b$ are called local Hamiltonians, beacuse if we have a seprable space $\psi_a \otimes \psi_b$ and we evolve it under $H_a \otimes I$, then

$$\exp(-iH_a \otimes I) = \exp(-iH_a) \otimes I$$

and

$$\exp(-iH_a \otimes I)\psi_a \otimes \psi_b = (\exp(-iH_a)\psi_a) \otimes \psi_b$$

The Hamiltonian only evoves A part of the subsystem. Similarly

$$\exp(-iI \otimes H_b)\psi_a \otimes \psi_b = \psi_a \otimes (\exp(-iH_b)\psi_b).$$

On the other hand if we have a Hamiltonian of the form $H_a \otimes H_b$, we call it an interaction Hamiltonian.

Then Hamiltonians for the joint system are of the general form

$$\{H_a \otimes I, I \otimes H_b, H_a \otimes H_b\}.$$

If we count dimensions there are $n_1^2 - 1$ (traceless Hermitian) Hamiltonians of the form $H_a \otimes I$ and $n_2^2 - 1$ of form $I \otimes H_b$ and $(n_1^2 - 1)(n_2^2 - 1)$ of form $H_a \otimes H_b$ and if we count them all we get total of $(n_1 n_2)^2 - 1$ which is indeed the dimension of Hamiltonians for a $n_1 n_2$ dimensional space.

To make all this concrete consider again spin $\frac{1}{2}$. It state space is 2 dimensional complex space with basis $|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $|1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. They are up-down states of spin. Like classical bits, a spin $\frac{1}{2}$ is called a quantum bit or qubit. However unlike a classical bit we can evolve our spin and prepare a state

$$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle).$$

This is called a superposition of 0 and 1. Now lets consider 2 spin $\frac{1}{2}$. Then the basis of our state space are

$$|00\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$|01\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

$$|10\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

$$|11\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Two spin $\frac{1}{2}$ are called coupled qubits. The Hamiltonians for the coupled qubit system are of the following kind

$$\{ \quad -i\sigma_x \otimes I, -i\sigma_y \otimes I, -i\sigma_z \otimes I, -iI \otimes \sigma_x, -iI \otimes \sigma_y, -iI \otimes \sigma_z,$$
$$-i\sigma_x \otimes \sigma_x, -i\sigma_x \otimes \sigma_y, -i\sigma_x \otimes \sigma_z, -i\sigma_y \otimes \sigma_x, -i\sigma_y \otimes \sigma_y, -i\sigma_y \otimes \sigma_z, -i\sigma_z \otimes \sigma_x, -i\sigma_z \otimes \sigma_y, -i\sigma_z \otimes \sigma_z\}$$

They ae 15 in all of these $\sigma_\alpha \otimes I$ and $I \otimes \sigma_\beta$ are local Hamiltonians and $\sigma_\alpha \otimes \sigma_\beta$ are interaction Hamiltonians where $\alpha, \beta \in \{x, y, z\}$.

Now suppose we start in the state $|00\rangle$ and evolve this state under Hamiltonian $-i\sigma_x \otimes I$ for time π then we get

$$\exp(-i\sigma_y \otimes I\pi)(\begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix}) = (\exp(-i\pi\sigma_y) \begin{bmatrix} 1 \\ 0 \end{bmatrix}) \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Now direct calculation shows that

$$\exp(-i\theta\sigma_y) = \cos\frac{\theta}{2}I - 2i\sin\frac{\theta}{2}\sigma_y = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Then

$$\exp(-i\sigma_y \otimes I\pi)(\begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix}) = (\begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix}) = |10\rangle$$

Thus by evolving the system under the given Hamiltonian we invert the state of the first spin. We have built an inverter. This is like a inverter in boolean/computer circuits but now done on a qubit. We say we have built an inverter gate.

Now in quantum mechanics we donot distinguish between state vector ψ and $\exp(i\alpha)\psi$, they differ by a so called global phase and are considered state. Therefore, we can also invert by

$$\exp(-i\sigma_x \otimes I\pi)(\begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix}) = (-i\begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix}) = -i|10\rangle$$

which is same as $|10\rangle$.

Now can we do something more interesting. Can we say evolve an Hamiltonian that will swap the state of two spins. Such that

$$|10\rangle \to |01\rangle, \quad |01\rangle \to |10\rangle$$

To do this we have to make the qubits interact using an interaction Hamiltonian.

Lets evolve under the hamiltonian

$$U = \exp(-i\pi(\sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y))$$

$$\sigma_x\sigma_x + \sigma_y\sigma_y = \frac{1}{4}(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \otimes \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \otimes \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}) = \frac{1}{2}\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Then

$$
\begin{aligned}
U|01\rangle &= U(\begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix}) = -i \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = |10\rangle \\
U|10\rangle &= U(\begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix}) = -i \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = |01\rangle \\
U|00\rangle &= |00\rangle \\
U|11\rangle &= |11\rangle
\end{aligned}
$$

We have built a SWAP gate. Another interesting gate is so called C-NOT gate. It inverts the state of the second qubit conditioned on the state of first qubit. If the state of first qubit is 0 we don't do anything else we invert.

$$
|00\rangle \to |00\rangle, \quad |01\rangle \to |01\rangle, \quad |10\rangle \to |11\rangle, \quad |11\rangle \to |10\rangle
$$

Let

$$
U = \exp(-i\pi(\frac{I}{2} - \sigma_z) \otimes \sigma_x) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \end{bmatrix}
$$

Then check we have built a CNOT gate.

Now we can generalize all this. We can have say n qubits. The state space is 2^n dimensional. The state

$$
|000\ldots 0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \cdots \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix}
$$

is all qubits in state zero. We can evolve a Hamiltonian $\sigma_x \otimes I \cdots \otimes I$, which is a local Hamiltonian that will only evolve the first qubit. Similarly a hamiltonian of the form $\sigma_x \otimes \sigma_x \otimes I \cdots \otimes I$, will make first two qubits interact and do a two qubit operation. So we can evolve Hamiltonians and do single qubit and two qubit operations. Now we can do any Boolean operation on n qubits and we have built a quantum computer. We can do operations as in classical computer but at the same time generate superpositions and do more powerful things we cannot do in classical computers. This allows us to do things we cannot do on a classical computer. Like we can factor very large integers in polynomial time. This is not possible on classical computers else we will break all existing crypto-systems which rely on the fact that it is hard to factor large integers.

Next we ask physically how do we get these Hamiltonians that we use to evolve our system. We saw for a single spin, when we but the spin in a magnetic field we get the evolution in Eq. (8), we then have our Hamiltonians, one drift H_0 and two control H_1 and H_2.

Interaction Hamiltonians arise because spins have magnetic moments and magnets interact. For example two magnetic moments in space μ_1 and μ_2 have hamiltonian (energy) as

$$H = \frac{\mu_0}{4\pi r^3}(\mu_1 \cdot \mu_2 - 3(\mu_1 \cdot \hat{r})(\mu_2 \cdot \hat{r}))$$

where r is the distance between moments and $\hat{r}$ is the unit vector connecting them.

Once again consider two qubits (spin $\frac{1}{2}$) and consider the evolution of the state vector ψ as

$$\dot{\psi} = -i\{u_1 \underbrace{\sigma_x \otimes I}_{H_1} + u_2 \underbrace{\sigma_y \otimes I}_{H_2} + u_3 \underbrace{I \otimes \sigma_x}_{H_3} + u_4 \underbrace{I \otimes \sigma_y}_{H_4} + J \underbrace{\sigma_z \otimes \sigma_z}_{H_0}\}\psi \tag{9}$$

Observe $\psi(t) = U(t)\psi(0)$, where $U(t) \in SU(4)$. Can we produce any unitary transformation on ψ. This is same as asking, is my system

$$\dot{U} = -i\{u_1\sigma_x \otimes I + u_2\sigma_y \otimes I + u_3 I \otimes \sigma_x + u_4 I \otimes \sigma_y + J \underbrace{\sigma_z \otimes \sigma_z}\}U \tag{10}$$

controllable. Observe we have four control Hamiltonians, which are local Hamiltonians. The first two rotate qubit 1 and last two rotate qubit 2. The drift hamiltonian is a interaction Hamiltonian and arises fron spin-spin interaction. The local Hamiltonians are produced by applying magnetic fields to the spins. Now to answer controllability question we have to use lie brackets.

By calculations like $[-i\sigma_x \otimes I, -i\sigma_y \otimes I] = -i\sigma_z \otimes I$ we can show that brackets of H_1, H_2, H_3, H_4 generate all local Hamiltonians

$$\{-i\sigma_x \otimes I, -i\sigma_y \otimes I, -i\sigma_z \otimes I, -iI \otimes \sigma_x, -iI \otimes \sigma_y, -iI \otimes \sigma_z\}$$

Now we can take brackets with drift and find we generate all the interaction generators,

$$\{-i\sigma_x\otimes\sigma_x, -i\sigma_x\otimes\sigma_y, -i\sigma_x\otimes\sigma_z, -i\sigma_y\otimes\sigma_x, -i\sigma_y\otimes\sigma_y, -i\sigma_y\otimes\sigma_z, -i\sigma_z\otimes\sigma_x, -i\sigma_z\otimes\sigma_y, -i\sigma_z\otimes\sigma_z\}$$

In taking Lie brackets we used the following identities

$$[A \otimes B, C \otimes D] = [A, C] \otimes BD + CA \otimes [B, D]$$

and

$$
\begin{aligned}
\sigma_x \sigma_y &= -\sigma_y \sigma_x = \frac{i}{2}\sigma_z \\
\sigma_y \sigma_z &= -\sigma_z \sigma_y = \frac{i}{2}\sigma_x \\
\sigma_z \sigma_x &= -\sigma_x \sigma_z = \frac{i}{2}\sigma_y
\end{aligned}
$$

For example,

$$[-i\sigma_z \otimes \sigma_z, -i\sigma_x \otimes I] = [-i\sigma_z, -i\sigma_x] \otimes \sigma_z + (-i\sigma_x - i\sigma_z) \otimes [\sigma_z, I] = -i\sigma_y \otimes \sigma_z$$

$$[-i\sigma_z \otimes \sigma_z, -i\sigma_x \otimes \sigma_x] = [-i\sigma_z, -i\sigma_x] \otimes \sigma_z\sigma_x + (\sigma_x\sigma_z) \otimes [-i\sigma_z, -i\sigma_x] = \frac{1}{2}(\sigma_y \otimes \sigma_y - \sigma_y \otimes \sigma_y)$$

The Lie algebra $\mathfrak{g} = su(4)$ is 15 dimensional and spanned by

$$\mathfrak{g} = \{-i\sigma_x \otimes I, -i\sigma_y \otimes I, -i\sigma_z \otimes I, -iI \otimes \sigma_x, -iI \otimes \sigma_y, -iI \otimes \sigma_z,$$
$$-i\sigma_x \otimes \sigma_x, -i\sigma_x \otimes \sigma_y, -i\sigma_x \otimes \sigma_z, -i\sigma_y \otimes \sigma_x, -i\sigma_y \otimes \sigma_y, -i\sigma_y \otimes \sigma_z, -i\sigma_z \otimes \sigma_x, -i\sigma_z \otimes \sigma_y, -i\sigma_z \otimes \sigma_z\}$$

This vector space $\mathfrak{g}$ has two orthogonal subspaces

$$\mathfrak{k} = \{-i\sigma_x \otimes I, -i\sigma_y \otimes I, -i\sigma_z \otimes I, -iI \otimes \sigma_x, -iI \otimes \sigma_y, -iI \otimes \sigma_z\}$$

the local generators and the interaction generators

$$\mathfrak{p} = \{-i\sigma_x \otimes \sigma_x, -i\sigma_x \otimes \sigma_y, -i\sigma_x \otimes \sigma_z, -i\sigma_y \otimes \sigma_x, -i\sigma_y \otimes \sigma_y, -i\sigma_y \otimes \sigma_z, -i\sigma_z \otimes \sigma_x, -i\sigma_z \otimes \sigma_y, -i\sigma_z \otimes \sigma_z\}$$

k is 6 dimensional and $\mathfrak{p}$ is 9 dimesnional and in total 15 dimensions. You should check that following commutations relations hold

$$[\mathfrak{k}, \mathfrak{k}] \subset \mathfrak{k}, \quad [\mathfrak{p}, \mathfrak{k}] \subset \mathfrak{p}, \quad [\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{k} \tag{11}$$

In general decomposition of a Lie algebras $\mathfrak{g}$, into a direct sum of two vector subspaces

$$\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k}$$

such that Eqs. 18 are true, is called a Cartan decomposition.

Now coming back to Eq. 9, we have shown controllability. We can generate any $U \in SU(4)$ but how to do it in minimum possible time. Now there is an important time scale sepration in such problems. u_1, u_2, u_3, u_4 are much larger compared to J. In practice they are in kHz range while J is of order of Hz. We say our control is fast while our drift is slow. Under such a time scale sepration we can say much about our time optimal control. Before we delve into it, we introduce a notaion used in NMR literature.

Given two spins or qubits, we call the first one I and second one S. Then the hamiltonian $\sigma_x \otimes I$ is written as I_x and $I \otimes \sigma_x$ as S_x and

$$\sigma_x \otimes \sigma_x = (\sigma_x \otimes I)(I \otimes \sigma_x) = I_x S_x$$

In this notation

$$\mathfrak{g} = \{-iI_x, -iI_y, -iI_z, -iS_x, -iS_y, -iS_z,$$
$$-iI_xS_x, -iI_xS_y, -iI_xS_z, -iI_yS_x, -iI_yS_y, -iI_yS_z, -iI_zS_x, -iI_zS_y, -iI_zS_z\}$$

Now we state our time optimal control probelm more generally, Let G be a compact Lie Group with Lie algebra $\mathfrak{g}$ and consider the control system on G

$$\dot{\Theta} = (X_d + \sum_i u_i X_i)\Theta$$

such that $\{X_d, X_i\}_{LA} = \mathfrak{g}$, system is controllable.

Further we have a cartan decomposition

$$\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k}$$

such that $X_d \in \mathfrak{p}$ and $X_i \in \mathfrak{k}$. We assume u_i can be made large compared to size of drift X_d. We study how to time optimally steer such a system. This will make our quantum computer fast.

Suppose we did not have drift then our system will be

$$\dot{\Theta} = \sum_i u_i X_i \Theta$$

Now this is not a controllable system, as $\{X_i\}_{LA} \neq \mathfrak{g}$, then without drift we cannot go everywhere.

However we assume that $\{X_i\}_{LA} = \mathfrak{k}$ a subalgebra of $\mathfrak{g}$.

Let $K = \exp(\mathfrak{k})$ be the subgroup of G generated by $\mathfrak{k}$.

Let us take an example. Let $G = SU(n)$ and consider the control system,

$$\dot{U} = (X_d + \sum_j u_j(t) X_{1j})U, \quad U(0) = I, \tag{12}$$

where $U \in SU(n)$. Where $X_{1j} \in \mathfrak{k} = so(n)$ is skew symmetric matrices with one in the $1j$ spot and

$$X_d = -i \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}, \quad \sum \lambda_i = 0.$$

Observe $\{X_{1j}\}_{LA}$, the Lie algebra (X_j and its matrix commutators) generated by generators X_{1j} is all of $so(n)$.

We have a cartan decomposition of $su(n)$, we have

$$su(n) = -iS \oplus so(n),$$

where S is traceless symmetric.

If we only had the system

$$\dot{U} = (\sum_j u_j(t) X_{1j})U, \quad U(0) = I, \tag{13}$$

we can steer the system to any point on $K = SO(n)$ and if we donot assume any bound on control, then we can steer it in as small time as possible. Because if $u_j(t)$ steer the system to U_F in time T then $N u_j(Nt)$ steers the system to U_F in time $\frac{T}{N}$.

So we can use fast controls to go anywhere on K and we do it so fast that X_d hardly evolves. Then we can imagine doing a control sequence like the following

$$K_{n+1} \exp(X_d t_n) K_n \dots K_2 \exp(X_d t_1) K_1$$

Using fast controls we generate K_1 , then evolve drift for t_1, then generate t_2 and so on. This way we can steer our system.

Coming back to Eq. (12), we can first understand how to generate any $U_F \in SU(N)$.

Any $U_F \in SU(N)$ can be written as

$$U_F = \exp(iS)K,$$

where S traceless symmetric and $K \in SO(n)$. We can diagonalize S with $SO(n)$ as

$$S = K_1 \begin{bmatrix} \mu_1 & 0 & \dots & 0 \\ 0 & \mu_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mu_n \end{bmatrix} K_1^{-1}, \quad \sum \mu_i = 0.$$

Then

$$U_F = K_1 \exp(-i \begin{bmatrix} \mu_1 & 0 & \dots & 0 \\ 0 & \mu_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mu_n \end{bmatrix}) K_2 \tag{14}$$

This decoposition of $U_F \in SU(N)$ is called KAK decompostion. In the above, using our controls we can generate K_1, K_2, how do we generate the middle part ,we need the drift. Observe using $SO(n)$ we can permute the drift. For X_d in Eq. (12), we can find permutation matrix in P in $SO(n)$ such that permute the diagonal of X_d

$$Ad_P(X_d) = P X_d P^{-1} = \begin{bmatrix} \lambda_{\sigma(1)} & 0 & \dots & 0 \\ 0 & \lambda_{\sigma(2)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_{\sigma(n)} \end{bmatrix}$$

Then we can find perumutations P_i such that

$$\sum_i \alpha_i Ad_{P_i}(X_d) = \begin{bmatrix} \mu_1 & 0 & \dots & 0 \\ 0 & \mu_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mu_n \end{bmatrix}$$

To see how it works let us say

$$X_d = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & -1 & \dots & 0 \\ \vdots & \vdots & 0 & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

writing the diagonal as a vector $\begin{bmatrix} 1 \\ -1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ we want to find α_i, such that

$$\alpha_1 \begin{bmatrix} 1 \\ -1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ 1 \\ -1 \\ \vdots \\ 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \vdots \\ \mu_n \end{bmatrix},$$

just solve sequentially, $\alpha_1 = \mu_1$ etc.

Then

$$\exp(-i \begin{bmatrix} \mu_1 & 0 & \dots & 0 \\ 0 & \mu_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mu_n \end{bmatrix}) = \exp(-i \sum_i \alpha_i Ad_{P_i}(X_d)) = \prod \exp(-i\alpha_i Ad_{P_i}(X_d)) = \prod P_i \exp(-i\alpha_i X_d) P_i^{-}$$

and in Eq. 14, we know now how to generate the central part.

Then we write our control strategy as

$$U_F = K_2 \prod \exp(-i\alpha_i Ad_{P_i}(X_d)) K_1, \quad \alpha_i > 0 \tag{15}$$

such that $Ad_{P_i}(X_d)$ all commute. We can further massage this as

$$\tilde{K}_2 \prod \exp(-i\alpha_i Ad_{Q_i}(X_d))$$

such that $Ad_{Q_i}(X_d)$ all commute. Then we go in commuting drift directions and in the end just jump on $SO(n)$. We will show that it is indeed the time optimal strategy. We have to move in commuting directions.

The fastest way to get to any U_F is to express it as in Eq. 15 and find the smallest $\sum \alpha_i$. We will now show this.

We consider Eq. (12) and let

$$\dot{K} = (\sum_j u_j(t) X_{1j}) K, \quad K(0) = I, \tag{16}$$

Let $V = K'U$, then

$$\dot{V} = \underbrace{K' X_d K}_{Ad_K(X_d)} V, \quad V(0) = I, \tag{17}$$

We want to steer U to U_F in minimum time. Since any K can be synthsized in Eq. (16) in no time, we can think of Eq. (17) as a control system with controls $Ad_K(X_d)$ where $K \in SO(n)$. The goal is to steer V to the coset KU_F in minimum time. Then once we reach the coset KU_F, we can reach U_F immediately as we can use our fast controls.

So let us understand how V evolves, at time t we can decompose V at $V = K_1 A K_2$, let us evolve for small time step Δt under $Ad_K(X_d)$. Then we have

$$V(t + \Delta T) = \exp(Ad_K(X_d)\Delta t) K_1 A K_2 = K_1 \exp(Ad_{\tilde{K}}(X_d)\Delta t) A K_2$$

Now $Ad_{\tilde{K}}(X_d) = -iQ$, for some traceless symmetric matrix Q. Then we can write

$$-iQ = \Lambda + R,$$

where Λ is the diagonal part and R offdiagonal part. Now A is off the form

$$A = \begin{bmatrix} \exp(i\phi_1) & 0 & \dots & 0 \\ 0 & \exp(i\phi_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \exp(i\phi_n) \end{bmatrix}$$

Let $\Omega \in so(n)$, then

$$(A \Omega A')_{ij} = \Omega_{ij}(\cos(\phi_i - \phi_j) + i \sin(\phi_i - \phi_j))$$

Lets assume $\phi_i - \phi_j \neq n\pi$, then $\sin(\phi_i - \phi_j) \neq 0$. We can choose Ω_{ij}, such that $i \sin(\phi_i - \phi_j)\Omega_{ij} = R_{ij}$, then

$$A \Omega A' = R + \Omega_1,$$

where $\Omega_1 \in so(n)$. Then

$$\begin{aligned} K_1 \exp(-iQ\Delta t) A K_2 &= K_1 \exp((\Lambda + R)\Delta t) A K_2 = K_1 \exp(-\Omega_1 \Delta t) \exp((\Lambda + R + \Omega_1)\ \Delta t) A K_2 \\ &= K_1 \exp(-\Omega_1 \Delta t) \exp(\Lambda \Delta t) \exp((R + \Omega_1)\Delta t) A K_2 \\ &= \underbrace{K_1 \exp(-\Omega_1 \Delta t)}_{K_1(t+\Delta t)} \underbrace{\exp(\Lambda \Delta t) A}_{A(t+\Delta t)} \underbrace{\exp(\Omega \Delta t) K_2}_{K_2(t+\Delta t)} \end{aligned}$$

From $V(t) = K_1(t) A(t) K_2(t)$, we evolve to $V(t + \Delta t) = K_1(t + \Delta t) A(t + \Delta t) K_2(t + \Delta t)$ where $A(t + \Delta t) = \exp(\Lambda \Delta t) A$. Now observe,

$Ad_{\tilde{K}}(X_d) = -iQ = \Lambda + R$, where Λ is diagonal of $-iQ$. Now we claim that

$$\Lambda = \sum_i \alpha_i P_i(X_d), \quad \alpha_i \geq 0$$

where $\sum_i \alpha_i = 1$ and P_i are permutation matrices.

Remark 1 Birkhoff convexity states, a real $n \times n$ matrix A is doubly stochastic ($\sum_i A_{ij} = \sum_j A_{ij} = 1$, for $A_{ij} \geq 0$) iff it can be written as convex hull of permutation matrices P_i (only one 1 and everything else zero in every row and column). Given $\Theta \in SO(n)$ and $X = -i \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$, we have $diag(\Theta X \Theta^T) = B\ diag(X)$ where $diag(X)$ is a column vector containing diagonal entries of X and $B_{ij} = (\Theta^{ij})^2$ and hence B is a doubly stochastic matrix which can be written as convex sum of permutations. Therefore $B\ diag(X) = \sum_i \alpha_i P_i\ diag(X)$, i.e. diagonal of a symmetric matrix $\Theta X \Theta^T$, lies in convex hull of its eigenvalues and its permutations. This is called Schur convexity.

Then we have

$$A(t + \Delta t) = \exp(\sum_i \alpha_i P_i(X_d)\Delta t)A(t); \quad A(T) = \exp(\sum_i \alpha_i P_i(X_d)T)A(0)$$

Then any U_F that can be reached in time T has the form

$$U_F = K_1 \exp(\sum_i \alpha_i P_i(X_d)T)K_2$$

This says, we can reach U_F in time T by going in commuting drift directions $P_i(X_d)$. Therefore our strategy of reaching U_F by going in commuting drift directions is optimal.

An important ingredient of the proof is the KAK decomposition. We first give a proof of it.

Theorem 1 Let $U \in SU(n)$, then $U = \Theta_1 \exp(\Omega)\Theta_2$ where $\Theta_1, \Theta_2 \in SO(n)$ and

$$\Omega = -i \begin{bmatrix} \lambda_1 & \dots & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix},$$

where $\sum_i \lambda_i = 0$.

Observe UU^T is in $SU(n)$. The eigenvalues of UU^T are of the form $\exp(j\theta)$.

$$UU^T z = \exp(j\theta) z.$$

$$\exp(-j\frac{\theta}{2})U^T z = \exp(j\frac{\theta}{2})(U^T)^* z.$$

$$(C + iD)z = (C - iD)z.$$

$$D(x + iy) = 0.$$

This implies $UU^T x = \exp(j\theta)x$ and $UU^T y = \exp(j\theta)y$. This implies $UU^T = \Theta\Sigma\Theta'$, where columns of Θ are real, perpendicular, and

$$\Sigma = \begin{bmatrix} \exp(-i\lambda_1) & \dots & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \exp(-i\lambda_n) \end{bmatrix}$$

where $\Sigma \in SU(n)$. Let $U = \Theta\Sigma^{\frac{1}{2}}V$. $UU^T = \Theta\Sigma\Theta' = \Theta\Sigma^{\frac{1}{2}}VV^T\Sigma^{\frac{1}{2}}\Theta'$.
Implying $VV^T = \mathbf{1}$. Then $U = \Theta\Sigma^{\frac{1}{2}}V$, where Θ, V can be chosen in $SO(n)$ and

$$\Sigma^{\frac{1}{2}} = \begin{bmatrix} \exp(-i\mu_1) & \dots & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \exp(-i\mu_n) \end{bmatrix},$$

where $\sum \mu_i = 2m\pi$. Choose $\mu_n \rightarrow \mu_n - 2m\pi$ so that $\sum \lambda_i = 0$ and result follows.**q.e.d**

We talked about $G = SU(n)$. All this can be generalized to general compact group G. Let Lie algebra $\mathfrak{g}$ has a cartan decompostion

$$\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k}, \tag{18}$$

when $\mathfrak{g} = su(n)$, we defined inner product of $su(n)$ as $tr(X'Y)$, where $X, Y \in su(n)$. In general we can choose basis of $\mathfrak{g}$ and in this basis $ad_X(\cdot) = [X, \cdot]$ is a matrix. The define

$$\langle X, Y \rangle = -tr(ad_X ad_Y).$$

This is an inner product called killing form. We only need Lie brackets to define it so it is intrinsic to g. If g has no abelian ideals (semisimple) $\langle X, Y \rangle$ is positive definite inner product.

As $\mathfrak{g} = su(n)$, choose as a basis of $\mathfrak{g}$, elements Ω_{kl}, $-i\Sigma_{kl}$ and $-iD_{l,l+1}$, where for $k < l$, Ω_{kl} is skew symetric with 1 in kl spot and Σ_{kl} is traceless symmetric with 1 in kl spot and

0 elsewhere abd D traceless diagonal with $1, -1$ in l and $l+1$ digonal spot and 0 elsewhere. Then in these basis lets compute what ad_X^2 looks like, where

$$X = -i \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}.$$

Then note $ad_X^2 \Omega_{kl} = -(\lambda_k - \lambda_l)^2 \Omega_{kl}$ and $ad_X^2 \Sigma_{kl} = -(\lambda_k - \lambda_l)^2 \Sigma_{kl}$ and $ad_X^2 D_{l,l+1} = 0$. Then in the chosen basis, ad_X^2 is diagonal and its trace is $\langle X, X \rangle = \sum_{kl} (\lambda_k - \lambda_l)^2$.

One can check in this killing form inner product $\langle p, k \rangle = 0$. Inside $\mathfrak{p}$ is maximally commuting subsalgebra $\mathfrak{a}$. When $\mathfrak{g} = su(n)$ and $\mathfrak{p} = iS$, for traceless symmetric S, we have $\mathfrak{a}$ as diagonal S. Clearly its all commuting and you cannot add any off diagonal matrix and still commute. Then let $K = \exp(\mathfrak{k})$ then any U_F in G can be written as

$$U_F = K_1 A K_2,$$

where $A \in \exp(\mathfrak{a})$. This is called KAK decomposition. we have seen it for the special case of $G = SU(N)$. We now sketch in general a proof when $\mathfrak{g}$ is compact semisimple. In this case with the cartan decomposition as in 18. we have for any $U_F \in G$, we can write it as

$$U_F = \exp(X) K_2,$$

where $K_2 \in exp(\mathfrak{k}) = K$ and $X \in \mathfrak{p}$. This is a fact that uses arguments about geodesics. We won's prove it, you can think of it a parametrization of G using directions in $\mathfrak{p}$ and in $\mathfrak{k}$. Now X can be diagonalized by K, i.e. there exists a $K_1 \in K$ such that $X = K_1 a K_1^{-1}$ for $a \in \mathfrak{a}$. To show this, we use the fact that $\mathfrak{a}$ has a regular element a_r such that if $Y \in \mathfrak{p}$ and $[Y, a_r] = 0$, then $Y \in \mathfrak{a}$.

For example when $\mathfrak{g} = su(n)$ and $\mathfrak{p} = iS$ for traceless symmetric S and $\mathfrak{a} = iD$ where D is traceless diagonal. Then a_r is D with all entries unequal, beacuse consider $Y \in iS$, then for $a_r = -i \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$ and $[a_r, Y]_{ij} = (\lambda_i - \lambda_j) Y_{ij}$. since $\lambda_i - \lambda_j \neq 0$ we have $Y_{ij} = 0$ and hence $Y \in \mathfrak{a}$

Now for general $\mathfrak{g}$, lets maximize

$$J = \langle K X K^{-1}, a_r \rangle$$

over choice of K. Suppose maximum is found at K_0. Then lets perturb as $K_0 \to \exp(ht) K_0$, where $h \in \mathfrak{k}$. Then

$$J(t) = \langle \exp(ht) K_0 X K_0^{-1} \exp(-ht), a_r \rangle$$

Then

$$\frac{dJ(t)}{dt}|_0 = \langle [h, K_0 X K_0^{-1}], a_r \rangle,$$

for maximum we have $\langle [h, K_0 X K_0^{-1}], a_r \rangle = 0$.

Now note for $X, Y, Z \in \mathfrak{g}$, we have $\langle [X,Y], Z \rangle = \langle X, [Y,Z] \rangle$. Note every Lie algebra satisfies Jacobi identity

$$ad_{[X,Y]}(Z) = [[X,Y],Z] = [[X,Z],Y] + [X,[Y,Z]] = ad_X ad_Y Z - ad_Y ad_X Z$$

Then

$$\langle [X,Y], Z \rangle = tr(ad_{[X,Y]} ad_Z) = tr(ad_X ad_Y ad_Z - ad_Y ad_X ad_Z) = tr(ad_X (ad_Y ad_Z - ad_Z ad_Y)) = \langle X, [Y,Z] \rangle$$

Then

$$\langle [h, K_0 X K_0^{-1}], a_r \rangle = \langle h, [K_0 X K_0^{-1}, a_r] \rangle = 0$$

Note $K_0 X K_0^{-1}, a_r \in \mathfrak{p}$ and hence $[K_0 X K_0^{-1}, a_r] \in \mathfrak{k}$. If $[K_0 X K_0^{-1}, a_r] \neq 0$, then I can choose $h = [K_0 X K_0^{-1}, a_r]$ and $\langle h, [K_0 X K_0^{-1}, a_r] \rangle \neq 0$. Therefore $[K_0 X K_0^{-1}, a_r] = 0$ and sice a_r is regular, we have $K_0 X K_0^{-1} \in \mathfrak{a}$. Therefore $K_0 X K_0^{-1} = a \in \mathfrak{a}$ and hence $X = K_0^{-1} a K_0$.

Then

$$U_F = \exp(K_0^{-1} a K_0) K_2 = K_0^{-1} \exp(a) K_0 K_2 = K_1 \exp(a) K_2,$$

the KAK decomposition.

$\mathfrak{a}$ is called the *cartan subalgebra* of $\mathfrak{g}$. If $\mathfrak{a}$ is cartan subalgebra, then so is $Ad_{K_1}(\mathfrak{a})$, for $K_1 \in K$. Infact $\mathfrak{p}$ is a union of such cartan subalgebras

$$\mathfrak{p} = \cup_K Ad_K(\mathfrak{a})$$

This is because for any $X \in \mathfrak{p}$, we have shown that $X = K_0^{-1} a K_0$ for some $a \in \mathfrak{a}$ and $K_0 \in K$. Then $X \in Ad_{K_0^{-1}}(\mathfrak{a})$.

In time optimal control problem for the SU(n) system in Eq. 12, we used an important convexity argument based on Birkhoff and Schur convexity. This result has an important generalization for a general compact semisimple $\mathfrak{g}$. It is called *Kostant Convexity*. Let $X \in \mathfrak{p}$. Lets look at $Ad_K(X)$ for all K. It is called orbit of X. It cuts $\mathfrak{a}$ is finite points X_i, called Weyl points. For example in $SU(n)$ case let $X = -i \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$ and $\mathfrak{a} = iD$ where

D is real tracelss diagonal matrices. Then $Ad_K(X)$ cuts $\mathfrak{a}$ in n points call X_i, the various permutations of X.

In SU(n) case we looked at $Ad_K(X)$, which for general K is not diagonal. We looked at its diagonal part, which is we project it on $\mathfrak{a}$. Then the projection (the diagonal part of $Ad_K(X)$) is in convex hull of X_i

This is true in general $\mathfrak{g}$. $Ad_K(X)$, for general K is not in $\mathfrak{a}$. The projection of $Ad_K(X)$ on $\mathfrak{a}$(w.r.t. to say killing form) is in convex hull of X_i.

The proof goes something like following.

Remark 2 Kostant Convexity Given the decomposition $\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k}$, let $\mathfrak{a} \subset \mathfrak{p}$ and $X \in \mathfrak{a}$,. Let $\mathcal{W}_i \in \exp(\mathfrak{k})$ such that $\mathcal{W}_i X \mathcal{W}_i \in \mathfrak{a}$ are distinct, Weyl points. Then projection (w.r.t killing form) of $Ad_K(X)$ on $\mathfrak{a}$ lies in convex hull of these Weyl points. The $\mathcal{C}$ be the convex hull and let projection $P(Ad_K(X))$ lie outside this Hull. Then there is a separating hyperplane a, such that $\langle Ad_K(X), a\rangle < \langle \mathcal{C}, a\rangle$. W.L.O.G we can take a to be a regular element. We minimize $\langle Ad_K(X), a\rangle$, with choice of K and find that minimum happens when $[Ad_K(X), a] = 0$, i.e. $Ad_K(X)$ is a Weyl point. Hence $P(Ad_K(X)) \in \sum_i \alpha_i \mathcal{W}_i X \mathcal{W}_i^{-1}$, for $\alpha_i > 0$ and $\sum_i \alpha_i = 1$. The result is true with a projection w.r.t inner product that satisfies $\langle x, [y, z]\rangle = \langle [x, y], z]\rangle$, like standard inner product on $\mathfrak{g} = su(n)$.

Remark 3 Stabilizer: Let $\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k}$ be cartan decomposition of real semisimple Lie algebra $\mathfrak{g}$ and $\mathfrak{a} \in \mathfrak{p}$ be its Cartan subalgebra. Let $a \in \mathfrak{a}$. Then $ad_a^2 : \mathfrak{p} \to \mathfrak{p}$ is symmetric in basis orthonormal wrt to the killing form. To see this let e_i be basis for $\mathfrak{p}$. Since they are orthonormal, we have

$$(ad_a^2)_{ij} = \langle e_i, [a[a, e_j]]\rangle = -\langle [a, e_i], [a, e_j]\rangle = \langle [a, [a, e_i]], e_j\rangle = (ad_a^2)_{ji}$$

We can diagonalize ad_a^2. Let Y_i be eigenvectors with nonzero (negative) eigenvalues $-\lambda_i^2$. Let $X_i = \frac{[a,Y_i]}{\lambda_i}$, $\lambda_i > 0$.

$$ad_a(Y_i) = \lambda_i X_i, \quad ad_a(X_i) = -\lambda_i Y_i.$$

X_i are independent, as $\sum \alpha_i X_i = 0$ implies $-\sum \alpha_i \lambda_i Y_i = 0$. Since Y_i are independent, X_i are independent. Given $X \perp X_i$, then $[a, X] = 0$, otherwise we can decompose it in eigenvectors of ad_a^2, i.e., $[a, X] = \sum_i \alpha_i a_i + \sum_j \beta_j Y_j$, where a_i are zero eigenvectors of ad_a^2. Since $0 = \langle X[a[a, X]\rangle = -\|[a, X]\|^2$, which means $[a, X] = 0$. This is a contradiction. Y_i are orthogonal, implies X_i are orthogonal, $\langle [a, Y_i][a, Y_j]\rangle = \langle [a, [a, Y_i]Y_j\rangle = \lambda_i^2 \langle Y_i Y_j\rangle = 0$. Let $\mathfrak{k}_0 \in \mathfrak{k}$ satisfy $[a, \mathfrak{k}_0] = 0$. Then $\mathfrak{k}_0 = \{X_i\}^{\perp}$.

$\tilde{Y}_i$ denote eigenvectors that have λ_i as non-zero integral multiples of π. $\tilde{X}_i$ are ad_a related to $\tilde{Y}_i$. We now reserve Y_i for non zero eigenvectors that are not integral multiples of π.

Let

$$\mathfrak{f} = \{a_i\} \oplus \tilde{Y}_i, \quad \mathfrak{h} = \mathfrak{k}_0 \oplus \tilde{X}_i,$$

$\tilde{X}_i, X_l, k_j$ where k_j forms a basis of $\mathfrak{k}_0$, forms a basis of $\mathfrak{k}$. Let $A = \exp(a)$.

$$AkA^- = A(\sum_i \alpha_i X_i + \sum_l \alpha_l \tilde{X}_l + \sum_j \alpha_j k_j)A^-,$$

where $k \in \mathfrak{k}$

$$AkA^- = \sum_i \alpha_i[\cos(\lambda_i)X_i - \sin(\lambda_i)Y_i] + \sum_l \pm\alpha_l \tilde{X}_l + \sum_j \alpha_j k_j. \tag{19}$$

The range of $A(\cdot)A^-$ in $\mathfrak{p}$, is perpendicular to $\mathfrak{f}$. Given $Y \in \mathfrak{p}$ such that $Y \in \mathfrak{f}^\perp$. The norm $\|X\|$ of $X \in \mathfrak{k}$, such that $\mathfrak{p}$ part of $AXA^{-1}|_{\mathfrak{p}} = Y$ satisfies

$$\|X\| \leq \frac{\|Y\|}{\sin \lambda_s}. \tag{20}$$

where λ_s^2 is the smallest nonzero eigenvalue of $-ad_a^2$ such that λ_s is not an integral multiple of π.

A^2kA^{-2} stabilizes $\mathfrak{h} \in \mathfrak{k}$ and $\mathfrak{f} \in \mathfrak{p}$. If $k \in \mathfrak{k}$, is stabilized by $A^2(\cdot)A^{-2}$, then in eq. 19 $\lambda_i = n\pi$, i.e., $k \in \mathfrak{h}$. This means $\mathfrak{h}$ is an sub-algebra, as the Lie bracket of $[y, z] \in \mathfrak{k}$ for $y, z \in \mathfrak{h}$ is stabilized by $A^2(\cdot)A^{-2}$.

Let $H = \exp(\mathfrak{h})$, be an integral manifold of $\mathfrak{h}$. It can be shown that $\exp(\mathfrak{h})$ is compact.

Let $y \in \mathfrak{f}$, then there exists a $h_0 \in \mathfrak{h}$ such that $\exp(h_0)y\exp(-h_0) \in \mathfrak{a}$. We maximize the function $\langle a_r, \exp(h)y\exp(h)\rangle$, over the compact group $\exp(\mathfrak{h})$, for regular element $a_r \in \mathfrak{a}$ and $\langle ., . \rangle$ is the killing form. At the maxima, we have at $t = 0$,

$$\frac{d}{dt}\langle a_r, \exp(h_1 t)(\exp(h_0)y\exp(-h_0))\exp(-h_1 t)\rangle = 0.$$

$$\langle a_r, [h_1 \exp(h_0)y\exp(-h_0)]\rangle = -\langle h_1, [a_r \exp(h_0)y\exp(-h_0)]\rangle,$$

if $\exp(h_0)y\exp(-h_0) \neq \mathfrak{a}$, then $[a_r,\ \exp(h_0)y\exp(-h_0)] \in \mathfrak{k}$. The bracket $[a_r,\ \exp(h_0)y\exp(-h_0)]$ is Ad_{A^2} invariant and hence belong to $\mathfrak{h}$. We can choose h_1 so that gradient is not zero. Hence $\exp(h_0)y\exp(-h_0) \in \mathfrak{a}$. For $z \in \mathfrak{p}$ such that $z \in \mathfrak{f}^\perp$, we have $\exp(h_0)z\exp(-h_0) \in \mathfrak{a}^\perp$.

$$\langle \mathfrak{a}, \exp(h_0)z\exp(-h_0)\rangle = \langle \exp(-h_0)\mathfrak{a}\exp(h_0), z\rangle = 0,$$

as $\exp(-h_0)\mathfrak{a}\exp(h_0)$ is Ad_{A^2} invariant, hence $\exp(-h_0)\mathfrak{a}\exp(h_0) \in \mathfrak{f}$. In above, we worked with killing form. For $\mathfrak{g} = su(n)$, we may use standard inner product.

We come back to

$$\dot{U} = (X_d + \sum_j u_j(t)X_j)U, \quad U(0) = I, \tag{21}$$

where $U \in G$, a compact group. We assume $\{X_j\}_{LA} = \mathfrak{k}$ and a Cartan decomposition $\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k}$ where $X_d \in \mathfrak{p}$. Let $\mathfrak{a}$ be a Cartan subalgebra containing X_d. As before $K = \exp(\mathfrak{k})$ can be generated fast. We llok at where all can we steer U in time T. Once again we define

$$\dot{K} = (\sum_j u_j(t) X_j) K, \quad \Theta(0) = I, \tag{22}$$

Let $V = K'U$, then

$$\dot{V} = \underbrace{K' X_d K}_{Ad_K(X_d)} V, \quad V(0) = I, \tag{23}$$

So let us understand how V evolves, at time t we can decompose V at $V = K_1 A K_2$, let us evolve for small time step Δt under $Ad_K(X_d)$. Then we have

$$V(t + \Delta T) = \exp(Ad_K(X_d)\Delta t) K_1 A K_2 = K_1 \exp(Ad_{\tilde{K}}(X_d)\Delta t) A K_2$$

Now $Q = Ad_{\tilde{K}}(X_d) = a + \sum_i \alpha_i Y_i$, for $a \in \mathfrak{a}$.

Now let A be such that corresponding $\mathfrak{f} = \mathfrak{a}$.

Then let $\Omega \in \mathfrak{k}$ by chosen as $\Omega = \sum_{\beta_i} X_i$, then

$$A \Omega A^{-1} = \sum_i \beta_i [\cos(\lambda_i) X_i - \sin(\lambda_i) Y_i] \tag{24}$$

Choose $\beta_i = -\frac{\alpha_i}{\sin(\lambda_i)}$ so that

$$A \Omega A^{-1} = \sum_i \alpha_i Y_i + \Omega_1 \tag{25}$$

where $\Omega_1 \in \mathfrak{k}$.

$$\begin{aligned} K_1 \exp(Q\Delta t) A K_2 &= K_1 \exp((a + \sum_{\alpha_i} Y_i)\Delta t) A K_2 = K_1 \exp(-\Omega_1 \Delta t) \exp(((a + \sum_{\alpha_i} Y_i + \Omega_1)\ \Delta t) A K_2 \\ &= K_1 \exp(-\Omega_1 \Delta t) \exp(a\Delta t) \exp((\sum_i \alpha_i Y_i + \Omega_1)\Delta t) A K_2 \\ &= \underbrace{K_1 \exp(-\Omega_1 \Delta t)}_{K_1(t+\Delta t)}\ \underbrace{\exp(a\Delta t) A}_{A(t+\Delta t)}\ \underbrace{\exp(\Omega \Delta t) K_2}_{K_2(t+\Delta t)} \end{aligned}$$

From $V(t) = K_1(t) A(t) K_2(t)$, we evolve to $V(t + \Delta t) = K_1(t + \Delta t) A(t + \Delta t) K_2(t + \Delta t)$ where $A(t + \Delta t) = \exp(a\Delta t) A$. Now observe,

Now by we claim that

$$a = \sum_i \alpha_i \mathcal{W}_i(X_d), \quad \alpha_i \geq 0$$

where $\sum_i \alpha_i = 1$ and $\mathcal{W}_i(X_d)$ are the Weyl points of X_d. They all belong to $\mathfrak{a}$ and hence commute.

$$A(t+\Delta t) = \exp(\sum_i \alpha_i \mathcal{W}_i(X_d)\Delta t)A(t); \quad A(T) = \exp(\sum_i \alpha_i \mathcal{W}_i(X_d)T)A(0)$$

Then any U_F that can be reached in time T has the form

$$U_F = K_1 \exp(\sum_i \alpha_i \mathcal{W}_i(X_d)T)K_2$$

This says, we can reach U_F in time T by going in commuting drift directions $P_i(X_d)$. Therefore our strategy of reaching U_F by going in commuting drift directions is optimal.

We assumed $\mathfrak{a} = \mathfrak{f}$. In general $\mathfrak{f}$ may be larger. Then $a \in \mathfrak{f}$ and we can find $h_0 \in \mathfrak{h}$ such that $\exp(h_0)a\exp(-h_0) = a_1 \in \mathfrak{a}$. Then substituting for a

$$\begin{aligned} K_1\exp(Q\Delta t)AK_2 &= K_1\exp(-\Omega_1\Delta t)\exp(a\Delta t)A\exp(\Omega\Delta t)K_2 \\ &= K_1\exp(-\Omega_1\Delta t)\exp(-h_0)\exp(a_1\Delta t)\exp(h_0)A\exp(\Omega\Delta t)K_2 \\ &= K_1\exp(-\Omega_1\Delta t)\exp(-h_0)\exp(a_1\Delta t)AA^{-1}\exp(h_0)A\exp(\Omega\Delta t)K_2 \\ &= \underbrace{K_1\exp(-\Omega_1\Delta t)\exp(-h_0)}_{K_1(t+\Delta t)}\underbrace{\exp(a_1\Delta t)A}_{A(t+\Delta t)}\underbrace{\exp(h_1)\exp(\Omega\Delta t)K_2}_{K_2(t+\Delta t)} \end{aligned}$$

where $h_1 \in \mathfrak{h}$. Note

$Q = a + z$, where $z \in \mathfrak{f}^\perp$. Then $Q = \exp(-h_0)a_1\exp(h_0) + z$ and $\exp(h_0)Q\exp(-h_0) = a_1 + \exp(h_0)z\exp(-h_0)$. Then since $\exp(h_0)z\exp(-h_0) \in a^\perp$ as shown before, we have a_1 as orthogonal projection of $\exp(h_0)Q\exp(-h_0) = Ad_K(X_d)$ on $\mathfrak{a}$. Hence by Kostant convexity

$$a_1 = \sum_i \alpha_i \mathcal{W}_i(X_d), \quad \alpha_i \geq 0$$

where $\sum_i \alpha_i = 1$ and $\mathcal{W}_i(X_d)$ are the Weyl points of X_d.

Now lets apply this general theory to the two qubit example we started our discussion with. If you recall the system is given in Eq. 9. The control generators $-iH_j$ generate the Lie algebra $\mathfrak{k}$, which is the local generators

$$\mathfrak{k} = \{-i\sigma_x \otimes I, -i\sigma_y \otimes I, -i\sigma_z \otimes I, -iI \otimes \sigma_x, -iI \otimes \sigma_y, -iI \otimes \sigma_z\}$$

and $\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k}$, where $\mathfrak{p}$ is the interaction generators

$$\mathfrak{p} = \{-i\sigma_x \otimes \sigma_x, -i\sigma_x \otimes \sigma_y, -i\sigma_x \otimes \sigma_z, -i\sigma_y \otimes \sigma_x, -i\sigma_y \otimes \sigma_y, -i\sigma_y \otimes \sigma_z, -i\sigma_z \otimes \sigma_x, -i\sigma_z \otimes \sigma_y, -i\sigma_z \otimes \sigma_z\}$$

Inside $\mathfrak{p}$ is Cartan subalgebra

$$\mathfrak{a} = \{-i\sigma_x \otimes \sigma_x, -i\sigma_y \otimes \sigma_y, -i\sigma_z \otimes \sigma_z\}$$

The genral $k \in \mathfrak{k}$ has the form

$$k = \sum_j \alpha_j - i\sigma_j \otimes I + \sum_j \alpha_j - iI \otimes \sigma_j$$

$$\begin{aligned} \exp k &= \exp(-i \sum_j \alpha_j \sigma_j \otimes I + -i \sum_j \beta_j I \otimes \sigma_j) = \exp(-i \sum_j \alpha_j \sigma_j \otimes I) \exp(-i \sum_j \beta_j I \otimes \sigma_j) \\ &= (\exp(-i \sum_j \alpha_j \sigma_j) \otimes I)(I \otimes \exp(-i \sum_j \beta_j \sigma_j)) = \exp(-i \sum_j \alpha_j \sigma_j) \otimes \exp(-i \sum_j \beta_j \sigma_j) \end{aligned}$$

Hence $K = SU(2) \otimes SU(2)$. K is the subgroup of local operation. Now we have drift $X_d = -i\sigma_x \otimes \sigma_x = -iI_z S_z \in \mathfrak{a} \subset \mathfrak{p}$ and $\mathcal{W}(X_d) = \{\pm I_z S_z, \pm I_x S_x, \pm I_y S_y\}$. Then from our general theory all unitary transformations that can be produced at time T are

$$U_F = K_1 \exp(T(\alpha_x I_x S_x + \alpha_2 I_y S_y + \alpha_3 I_z S_z)) K_2, \quad |\alpha_x| + |\alpha_y| + |\alpha_z| \leq 1$$

To generate U_F we produce K_1, K_2 fast and go in commuting directions $\{-i\sigma_x \otimes \sigma_x, -i\sigma_y \otimes \sigma_y, -i\sigma_z \otimes \sigma_z\}$.

1 Excercise

1. For $\alpha_x^2 + \alpha_y^2 + \alpha_z^2 = 1$ show that

$$\exp(-i\theta(\alpha_x \sigma_x + \alpha_y \sigma_y + \alpha_z \sigma_z)) = \cos\frac{\theta}{2} I - 2i \sin\frac{\theta}{2}(\alpha_x \sigma_x + \alpha_y \sigma_y + \alpha_z \sigma_z).$$

2. Let $U \in SU(2)$, then show we can write U as

$$U = \begin{bmatrix} \cos\alpha e^{i\delta} & \sin\alpha e^{i\phi} \\ -\sin\alpha e^{-i\phi} & \cos\alpha e^{-i\delta} \end{bmatrix}$$

Further show U can be written as

$$U = \cos\frac{\theta}{2} I - 2i \sin\frac{\theta}{2}(\alpha_x \sigma_x + \alpha_y \sigma_y + \alpha_z \sigma_z),$$

where $\alpha_x^2 + \alpha_y^2 + \alpha_z^2 = 1$.

3. For $\alpha, \beta, \gamma, \delta \in \{x, y, z\}$ show that $[\sigma_\alpha \otimes \sigma_\beta, \sigma_\gamma \otimes \sigma_\delta] = 0$ if $\alpha \neq \gamma$ and $\beta \neq \delta$.

4. Let $\mathfrak{g}$ be Lie algebra of Lie group G. Show that for $X, Y \in \mathfrak{g}$ the killing form

$$\langle X, Y \rangle = \langle \Theta X \Theta^{-1}, \Theta Y \Theta^{-1} \rangle$$

for $\Theta \in G$.

5. Let $\mathfrak{g}$ have a abelian ideal $\mathfrak{a}$. Show that for $X \in \mathfrak{a}$, the killing form $\langle X, X \rangle = 0$.

6. Let $\mathfrak{g}$ be Lie algebra of compact Lie group G. Show that the killing form

$$\langle X, X \rangle \geq 0.$$

7. Let $\mathfrak{g}$ be semi-simple Lie algebra of a compact Lie group G. Show that the killing form

$$\langle X, X \rangle > 0.$$

8. Let $\mathfrak{g}$ be a Lie algebra. Show that if the killing form $\langle X, X \rangle > 0$, then $\mathfrak{g}$ is semisimple.

9. Show that for $\mathfrak{g} = su(n)$, $\langle X, X \rangle > 0$. Therefore show it is semisimple.

10. Let $\mathfrak{g} = su(2n)$. Let $\mathfrak{k} = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$, space of block diagonal traceless skew Hermitian matrices. let $\mathfrak{p} = \begin{bmatrix} 0 & Z \\ -Z' & 0 \end{bmatrix}$. Show that $\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k}$ is a Cartan decomposition.

11. In above problem show that $\mathfrak{a} = \begin{bmatrix} 0 & \Lambda \\ -\Lambda & 0 \end{bmatrix}$, where Λ is real diagonal is a cartan subalgebra.

12. In above let $a_r = \begin{bmatrix} 0 & \Lambda \\ -\Lambda & 0 \end{bmatrix} \in \mathfrak{a}$ where Λ is real diagonal. Derive a condition on diagonal entries of Λ so that a_r is a regular element of $\mathfrak{a}$.

13. Let $U \in SU(2n)$, then show

$$U = \begin{bmatrix} K_1 & 0 \\ 0 & K_2 \end{bmatrix} \exp(\begin{bmatrix} 0 & \lambda \\ -\lambda & 0 \end{bmatrix}) \begin{bmatrix} K_3 & 0 \\ 0 & K_4 \end{bmatrix},$$

where $\begin{bmatrix} K_1 & 0 \\ 0 & K_2 \end{bmatrix}, \begin{bmatrix} K_3 & 0 \\ 0 & K_4 \end{bmatrix} \in SU(n) \times SU(n) \times U(1)$ (Block diagonal special unitary matrices) and

$$\begin{bmatrix} 0 & \lambda \\ -\lambda & 0 \end{bmatrix} = \begin{bmatrix} 0 & \dots & 0 & \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \lambda_n \\ -\lambda_1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & -\lambda_n & 0 & \dots & 0 \end{bmatrix}$$

14. For $\mathfrak{g} = su(4)$ and decomposition of $\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k}$ into local and interaction generators The cartanalgebra

$$\mathfrak{a} = \{-i\sigma_x \otimes \sigma_x, -i\sigma_y \otimes \sigma_y, -i\sigma_z \otimes \sigma_z\}.$$

For $a_r = -i(a_x\sigma_x \otimes \sigma_x + a_y\sigma_y \otimes \sigma_y + a_z\sigma_z \otimes \sigma_z) \in \mathfrak{a}$, when is a_r a regular element.

2 Control of spin dynamics under dissipation

We come back to equation of collection of spins ψ_k in magnetic field

$$\dot{\psi}_k = -i(\omega_0\sigma_z + u\sigma_x + v\sigma_y)\psi_k, \tag{26}$$

Recalled we formed an average subspace spanned by these ψ_k as $\rho = \frac{1}{N}\sum \psi_k\psi_k^\dagger$, then ρ evolves as

$$\dot{\rho} = [-i(\omega_0\sigma_z + u\sigma_x + v\sigma_y), \rho] \tag{27}$$

ρ is a two dimensional Hermitian matrix and can be written as

$\rho = \frac{1}{2}I + l_x\sigma_x + l_y\sigma_y + l_z\sigma_z$, where $L = (l_x, l_y, l_z)'$ represents average (x, y, z) angular momentum of the of the ensemble. This average or classical angular momentum evolves as Bloch equation

$$\dot{L} = (\omega_0\Omega_z + u\Omega_x + v\Omega_y)L, \tag{28}$$

Now we start with say $\psi_k = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\rho = \frac{1}{2}I + \sigma_z$. Now we rotate ψ_k to $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ which is done by evolving the system as $\exp(-i\frac{\pi}{2}\sigma_y)\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Now our ensemble has $\rho = \frac{1}{2}I + \sigma_z$. Now we let our ensemble evolve under the drift hamiltonian $-i\omega_0\sigma_z$. Then the system evolves as

$$\exp(-i\omega_0 t\sigma_z)\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} e^{-\frac{i}{2}\omega_0 t} \\ e^{\frac{i}{2}\omega_0 t} \end{bmatrix}$$

Then $\rho = \frac{1}{2}I + \cos\omega_0 t\sigma_x + \sin\omega_0 t$. In practice every ψ_k sees a slightly diffrent ω due to local fluctuations of the magnetic field. Hence we really have

$$\psi_k(t) = \frac{1}{\sqrt{2}}\begin{bmatrix} e^{-\frac{i}{2}\omega_k t} \\ e^{\frac{i}{2}\omega_k t} \end{bmatrix}$$

The correct model is that all ω_k start as ω_0 and then begin to diffuse away. As a result we find

$\rho = \frac{1}{2}I + \frac{1}{N}\sum(\cos\omega_k t\sigma_x + \sin\omega_k t\sigma_y) \to \frac{1}{2}$. With time the magnetic moments which started together on the transverse plain begin to move away leading to a zero net magnetization or magnetic moment. This is called the phenomenon of decoherence.

We capture it by introducing a decay R in the Bloch equations. The equations look

$$\dot{M} = \begin{bmatrix} -R & -\omega_0 & v \\ \omega_0 & -R & -u \\ -v & u & 0 \end{bmatrix} M,$$

where R is called the transverse relation rate.

It is put in the density matrix equation as

$$\dot{\rho} = [-i(\omega_0\sigma_z + u\sigma_x + v\sigma_y), \rho] - R[\sigma_z[\sigma_z, \rho]] \tag{29}$$

We come back to the system in Eq. 9,

$$\dot{\psi} = -i\{u_1 I_x + u_2 I_y + u_3 S_x + u_4 S_y + \pi J I_z S_z\}\psi \tag{30}$$

Observe $\psi(t) = U(t)\psi(0)$, where $U(t) \in SU(4)$ and

$$\dot{U} = -i\{u_1 I_x + u_2 I_y + u_3 S_x + u_4 S_y + \pi J I_z S_z\}U \tag{31}$$

We now study an important experiment in NMR called INEPT. We know protons have larger gyromagnetic ratio γ than carbon. It is about 4 times larger. This means that in a ensemble of 10^5 protons if we have 4 excess spins pointing in along B_0, for carbon we will only have 1 excess spin. The net magnetic moment of proton will be much larger that carbon. Therefore in experiments we get stronger signal when we observe proton compared to when we detect carbon. Let us denote more sensitive nuclie by I and other one by S. Now imagine that I have 18 spin pairs of IS. Of then half of them have spin S up and half down This is to say there is no preferential alignment of S as it is insensitive. Of these first half 5 have spin I up and 4 spin down and same for the other half. So how does the density matrix look like.

$$\rho = \frac{1}{18}(5|00\rangle\langle 00| + 4|10\rangle\langle 10| + 5|01\rangle\langle 01| + 5|11\rangle\langle 11|)$$

$$\rho = \frac{1}{18}(5|0\rangle\langle 0| \otimes I + 4|0\rangle\langle 0| \otimes I) = \frac{I}{4} + \frac{1}{18}\sigma_z \otimes I$$

Thus our ensemble has a net polarization on spin I. There is no net polarization on spin S. if we try to detect spin S we won't get a good signal. Now we show how using Control, we can transfer this polarization from I to S spin. Then we will be able to detect and observe S spin.

This is how control works. We go through the following set of operations. Recall under unitary propagator U density matrix evolves as $\rho \to U\rho U'$

Then

$$\sigma_z \otimes I \;\to\; \exp(i\frac{\pi}{2}I_x)\ (\sigma_z \otimes I)\ \exp(-i\frac{\pi}{2}I_x) = \sigma_x \otimes I \tag{32}$$

$$\sigma_x \otimes I \;\to\; \exp(-i\frac{\pi}{2}2I_zS_z)\ (\sigma_x \otimes I)\ \exp(i\frac{\pi}{2}2I_zS_z) = 2\sigma_y \otimes \sigma_z \tag{33}$$

$$2\sigma_y \otimes \sigma_z \;\to\; \exp(-i\frac{\pi}{2}(I_x + S_x))\ (2\sigma_y \otimes \sigma_z)\ \exp(-i\frac{\pi}{2}(I_x + S_x)) = -2\sigma_z \otimes \sigma_y \tag{34}$$

$$-2\sigma_z \otimes \sigma_y \;\to\; \exp(-i\frac{\pi}{2}2I_zS_z)\ (2\sigma_z \otimes \sigma_y)\ \exp(i\frac{\pi}{2}2I_zS_z) = I \otimes \sigma_x \tag{35}$$

$$I \otimes \sigma_x \;\to\; \exp(i\frac{\pi}{2}S_y)\ (I \otimes \sigma_x)\ \exp(-i\frac{\pi}{2}S_y) = I \otimes \sigma_z \tag{36}$$

Now we have an ensemble in which spin S is polarized. Now we can Observe spin S as there is more polarization on it.

After steps 32 and 33, we can convert

$$2\sigma_y \otimes \sigma_z \to \exp(-i\frac{\pi}{2}I_x)\ (2\sigma_y \otimes \sigma_z)\ \exp(-i\frac{\pi}{2}I_x) = 2\sigma_z \otimes \sigma_z \tag{37}$$

and we say we have created a two spin order, by transferring $\sigma_z \otimes I \to 2\sigma_z \otimes \sigma_z$. Observe steps 1 (Eq. 32) and 3 (Eq. 37) are fast steps because they involve evolving control Hamiltonians and hence take no time. Step 2 (Eq. 33) is a slow step. It involves evolving interaction, coupling hamiltonian in Eq. 30 for time $\frac{1}{2J}$. In writing these steps we have neglected all decoherence, and assume step 2 evolves as

$$\sigma_x \otimes I \rightarrow (\sigma_x \otimes I)\cos(\pi J t) + 2\sigma_y \otimes \sigma_z \sin(\pi J t) \tag{38}$$

which after time $\frac{1}{2J}$ reaches the target state.

However in presence of decoherence, this evolution is

$$\sigma_x \otimes I \rightarrow \exp(-Rt)\{(\sigma_x \otimes I)\cos(\pi J t) + 2\sigma_y \otimes \sigma_z \sin(\pi J t)\} \tag{39}$$

where R is the relaxation rate. Now we shouldn't wait for full $\frac{1}{2J}$ rather we should find the time when $\exp(-Rt)\sin(\pi J t)$ is maximum. This way we maximize the transfer to target state. Now we ask is this the best we can do. Important is to note that while the states $\sigma_x \otimes I$ and $\sigma_y \otimes \sigma_z$ decay under decoherence, the starting and end state $\sigma_x \otimes I$ and $\sigma_z \otimes \sigma_z$ donot. Then we can write a state space description for this transfer. Let the magnitudes of these states be $x_1 = \langle \sigma_z \otimes I \rangle$, $x_2 = \langle \sigma_x \otimes I \rangle$, $x_3 = \langle 2\sigma_y \otimes \sigma_z \rangle$ and $x_4 = \langle 2\sigma_z \otimes \sigma_z \rangle$. Then consider the control system

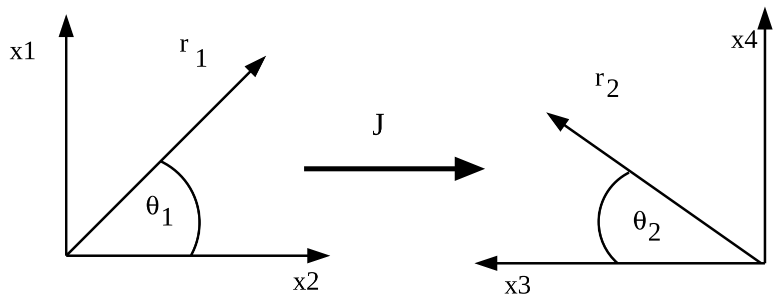
$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 & -u(t) & 0 & 0 \\ u(t) & -R & -\pi J & 0 \\ 0 & \pi J & -R & -v(t) \\ 0 & 0 & v(t) & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \tag{40}$$

Here u, v represent our fast operations that can immdiately transfer between $x_1 - x_2$ and $x_3 - x_4$. R is relaxation rate and describes how x_2, x_3 decay and πJ is due to coupling that evolves x_2 to x_3. We want to find starting from $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$, what is the largest value of x_4 possible, i.e., what is the larget η so that we can get to $\begin{bmatrix} 0 \\ 0 \\ 0 \\ \eta \end{bmatrix}$.

The states are depicted in the following vector diagram

where $r_1 = \sqrt{x_1^2 + x_2^2}$ and $r_2 = \sqrt{x_3^2 + x_4^2}$. Note with fast u, v we can control θ_1 and θ_2 fast. Let $u_1 = \cos\theta_1$ and $u_2 = \cos\theta_2$. Then we can write an equation for r_1 and r_2 and it takes the form

$$\frac{d}{dt}\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} -R\cos\theta_1^2 & -\pi J\cos\theta_1\cos\theta_2 \\ \pi J\cos\theta_1\cos\theta_2 & -R\cos\theta_2^2 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \pi J \begin{bmatrix} -\xi u_1^2 & -u_1 u_2 \\ \pi u_1 u_2 & -\xi u_2^2 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}, \tag{41}$$



where $\xi = \frac{R}{\pi J}$. The goal is to find $-1 \leq u_1(t), u_2(t) \leq 1$, that maximize transfer to the final state r_2, starting from the initial state $(r_1, r_2) = (1, 0)$.

We maximize the gain,

$$\frac{\Delta r_2^2}{-\Delta r_1^2} = \frac{(-\xi p^2 + p)\Delta t}{(\xi + p)\Delta t},$$

where $p = \frac{u_2 r_2}{u_1 r_1}$, positive, as r_2 decreases for negative p. Differentiating with p, we get

$$\frac{(1 - 2\xi p)(p + \xi) - p(1 - \xi p)}{(p + \xi)^2} = -\frac{(p - \eta_1)(p - \eta_2)}{(p + \xi)^2}$$

where $\eta_1 = -(\sqrt{1 + \xi^2} + \xi)$ and $\eta_2 = \sqrt{1 + \xi^2} - \xi$. Slope is increasing between roots η_1 and η_2 and decreasing outside. Maximum is at $\eta = \eta_2$. Substituting this value of p gives

$$\frac{\Delta r_2^2}{-\Delta r_1^2} = \eta^2,$$

with $-\Delta r_1^2 > 0$. Thus

$$\Delta(\eta^2 r_1^2 + r_2^2) = 0,$$

along the optimal trajectory. Using optimal return function for the problem as

$$V(r_1, r_2) = \eta^2 r_1^2 + r_2^2$$

we obtain for $m_1 = u_1 r_1$ and $m_2 = u_2 r_2$,

$$\frac{dV}{dt} = J \begin{bmatrix} \eta^2 & 1 \end{bmatrix} \begin{bmatrix} -\xi m_1^2 - m_1 m_2 \\ m_1 m_2 - \xi m_2^2 \end{bmatrix} = J \begin{bmatrix} m_1 & m_2 \end{bmatrix} \underbrace{\begin{bmatrix} -\xi \eta^2 & \frac{(1-\eta^2)}{2} \\ \frac{(1-\eta^2)}{2} & -\xi \end{bmatrix}}_{A} \begin{bmatrix} m_1 \\ m_2 \end{bmatrix}.$$

A is negative semi-definite for $\eta = \sqrt{1 + \xi^2} - \xi$, with null vector at $(1, \eta)$. Therefore, $\frac{dV}{dt} \leq 0$ with $\frac{dV}{dt} = 0$, for $\frac{m_2}{m_1} = \frac{u_2 r_2}{u_1 r_1} = \eta$.

$$\frac{d}{dt}\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} -u_1^2 & au_1u_2 \\ bu_1u_2 & -u_2^2 \end{bmatrix}\begin{bmatrix} r_1 \\ r_2 \end{bmatrix}, \tag{42}$$

$0 \le u_i \le 1$, $b > 0$ and $-1 \le \frac{a+b}{2} \le 1$ (system is dissipative in norm).

We maximize the gain,

$$\frac{\Delta r_2^2}{-\Delta r_1^2} = \frac{(-p^2 + bp)\Delta t}{(1 - ap)\Delta t},$$

where $p = \frac{u_2 r_2}{u_1 r_1}$, positive. Differentiating with p, we get

$$\frac{a(p^2 - \frac{2p}{a} + \frac{b}{a})}{(1 - ap)^2} = a\frac{(p - \eta_1)(p - \eta_2)}{(1 - ap)^2}$$

where $\eta_1 = a^{-1} - \sqrt{a^{-1}(a^{-1} - b)}$, $\eta_2 = a^{-1} + \sqrt{a^{-1}(a^{-1} - b)}$. When a is negative, slope is increasing between roots η_1 and η_2 and decreasing outside. Maximum is at η_2 ($p > 0$). When a is positive, slope is decreasing between roots η_1 and η_2 with $p < a^{-1}$ for $-\Delta r_1^2$ to be positive. Hence maximum is at η_1. In both case we have

$$a\eta + b\eta^{-1} = 2, \tag{43}$$

for maximum argument $p = \eta$. When $a = 0$, $\eta = \frac{b}{2}$.

When we substitute the value of p, it gives

$$\frac{\Delta r_2^2}{-\Delta r_1^2} = \frac{(-p^2 + bp)\Delta t}{(1 - ap)\Delta t} = \eta^2.$$

Using $V = \eta^2 r_1^2 + r_2^2$ as the return function for the largest value of r_2^2 possible, we find on differentiation

$$\frac{dV}{dt} = \begin{bmatrix} m_1 & m_2 \end{bmatrix} \underbrace{\begin{bmatrix} -\eta^2 & \frac{(b+a\eta^2)}{2} \\ \frac{(b+a\eta^2)}{2} & -1 \end{bmatrix}}_{A} \begin{bmatrix} m_1 \\ m_2 \end{bmatrix},$$

where, $m_1 = u_1 r_1$ and $m_2 = u_2 r_2$. A is negative semi-definite for $\eta = \eta_1 = a^{-1} - \sqrt{a^{-1}(a^{-1} - b)}$, when a positive and $\eta = \eta_2 = a^{-1} + \sqrt{a^{-1}(a^{-1} - b)}$, when a negative and $\eta = \frac{b}{2}$, when $a = 0$. The null vector is at $(1, \eta)$. Therefore, $\frac{dV}{dt} \le 0$ with $\frac{dV}{dt} = 0$, for $\frac{m_2}{m_1} = \frac{u_2 r_2}{u_1 r_1} = \eta$.

Consider special choice of a and b which is motivated later. Let $a = \frac{\chi}{\xi}\cos(\theta + \gamma)$ and $b = \frac{\chi}{\xi}\cos(\theta - \gamma)$, where $\chi = \sqrt{1 + \frac{k_c^2}{J^2}}$ and $\xi = \frac{k_a}{J}$ and $\cos\theta = \frac{-k_c}{\sqrt{k_c^2 + J^2}}$, which depend of three parameters in the system dynamics k_a, k_c, J. See Eq. 45 subsequently. This gives rise to the following system.

$$\frac{d}{dt}\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} -u_1^2 & \frac{\chi}{\xi}\cos(\theta+\gamma)u_1u_2 \\ \frac{\chi}{\xi}\cos(\theta-\gamma)u_1u_2 & -u_2^2 \end{bmatrix}\begin{bmatrix} r_1 \\ r_2 \end{bmatrix}, \tag{44}$$

where we choose $0 \le u_i \le 1$ to maximum transfer to r_2 starting from $(r_1, r_2) = (1, 0)$.

Then using Eq. 43,

$$\eta\cos(\theta+\gamma) + \eta^{-1}\cos(\theta-\gamma) = 2\frac{\xi}{\chi}.$$

$$(\eta+\eta^{-1})\cos\theta\cos\gamma + (\eta^{-1}-\eta)\sin\theta\sin\gamma = 2\frac{\xi}{\chi}.$$

$$(\eta+\eta^{-1})^2\cos^2\theta + (\eta^{-1}-\eta)^2\sin^2\theta \ge 4\frac{\xi^2}{\chi^2}.$$

$$(\eta^{-1}-\eta) \ge 2\zeta,$$

where $\zeta^2 = \frac{k_a^2-k_c^2}{k_a^2+k_c^2}$. Maximum of $\eta < 1$ is obtained, when we choose $\tan\gamma = \frac{1-\eta^2}{1+\eta^2}\tan\theta$,

$$(\eta^{-1}-\eta) = 2\zeta, \quad \eta = \sqrt{1+\zeta^2} - \zeta.$$

We can use

$$V(r_1, r_2) = \eta^2 r_1^2 + r_2^2$$

as return function for the system in 44. We obtain for $m_1 = u_1 r_1$ and $m_2 = u_2 r_2$,

$$\frac{dV}{dt} = \begin{bmatrix} \eta^2 & 1 \end{bmatrix}\begin{bmatrix} -m_1^2 & m_1m_2\frac{\chi}{\xi}\cos(\theta+\gamma) \\ \frac{\chi}{\xi}\cos(\theta-\gamma)m_1m_2 & -m_2^2 \end{bmatrix} = \begin{bmatrix} m_1 & m_2 \end{bmatrix}\underbrace{\begin{bmatrix} -\eta^2 & c \\ c & -1 \end{bmatrix}}_{A}\begin{bmatrix} m_1 \\ m_2 \end{bmatrix},$$

where $c = \eta\frac{\chi}{2\xi}(\eta\cos(\theta+\gamma) + \eta^{-1}\cos(\theta-\gamma))$. Where for special choice of γ, we have

$c^2 = \eta^2$ and the matrix is negative semi-definite with null space $(m_1, m_2) = (1, \eta)$, else its negative definite with $\frac{dV}{dt} < 0$.

The system in Eq. 44, arises from following transfer problem, which is of fundamental and practical interest in NMR spectroscopy. Given the control system

$$\frac{d}{dt}\begin{bmatrix} z_1 \\ y_1 \\ x_1 \\ x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} 0 & u(t) & -v(t) & 0 & 0 & 0 \\ -u(t) & -k_a & 0 & J & -k_c & 0 \\ v(t) & 0 & -k_a & -k_c & -J & 0 \\ 0 & -J & -k_c & -k_a & 0 & v(t) \\ 0 & -k_c & J & 0 & -k_a & -u(t) \\ 0 & 0 & 0 & -v(t) & u(t) & 0 \end{bmatrix}\begin{bmatrix} z_1 \\ y_1 \\ x_1 \\ x_2 \\ y_2 \\ z_2 \end{bmatrix}, \tag{45}$$

find optimal $(u(t), v(t))$ such that starting from $(z_1, y_1, x_1, x_2, y_2, z_2) = (1, 0, 0, 0, 0, 0)$, we obtain the largest value of $(0, 0, 0, 0, 0, \eta)$?

The above optimal control problem can be solved in closed form. Consider the vectors (x_2, y_2) and (x_1, y_1) with length l_2 and l_1. Writing equation for $r_1 = \sqrt{l_1^2 + z_1^2}$ and $r_2 = \sqrt{l_2^2 + z_2^2}$, with $u_i = cos(\beta_i) = \frac{l_i}{r_i}$, where $u_i > 0$, we get the system

$$\frac{d}{dt}\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} -\xi u_1^2 & \chi\cos(\theta+\gamma)u_1 u_2 \\ \chi\cos(\theta-\gamma)u_1 u_2 & -\xi u_2^2 \end{bmatrix}\begin{bmatrix} r_1 \\ r_2 \end{bmatrix}, \tag{46}$$

where γ is angle between l_1 and l_2 which is taken as a control variable. This Eq. is the scaled model in Eq. 44.

The optimal solution is then given by the following two invariants of motion. The ratio

$$\frac{u_2 r_2}{u_1 r_1} = \frac{l_2}{l_1} = \sqrt{1+\zeta^2} - \zeta = \eta; \quad \zeta = \sqrt{\frac{k_a^2 - k_c^2}{k_a^2 + J^2}} \tag{47}$$

is maintained constant and the angle γ between vectors l_2 and l_1 is maintained constant at $\tan\gamma = \frac{1-\eta^2}{1+\eta^2}\tan\theta$. The maximum transfer of efficiency is then η.

It is worthwhile to point out that researchers in magnetic resonance have developed novel pulse sequences that have improved the transfer described in Eq. (45), however the fundamental limits of the transfer described here was not known. Fig. 3 shows plot of transfer efficiency of various state of the art pulse sequences as a function of the ratio $\frac{k_a}{J}$ for $k_c = .75$. The CROP pulse sequence obtained by solving the above transfer problem using methods of optimal control (Eq. 47) performs better than all state of the art methods and provide significant improvement is sensitivity. Furthermore methods of optimal control help to provide limits on how close can a quantum dynamical system be driven to a target state.

Generalizing, we can consider a general dissipative control system

$$\dot{r} = A(u_i, u_j)r; \quad A(u_i, u_j) = \begin{bmatrix} a_{11}u_1^2 & \dots & a_{1j}u_1u_j & \dots & a_{1n}u_1u_n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{j1}u_ju_1 & \dots & a_{jj}u_j^2 & \dots & a_{jn}u_ju_n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1}u_nu_1 & \dots & a_{nj}u_nu_j & \dots & a_{nn}u_n^2 \end{bmatrix} = Au \circ u \tag{48}$$

where $\circ$ is Hadamard product and $u = (u_1, \dots, u_n)'$, with $\{A\}_{ij} = a_{ij}$ ($A + A^T$ is negative definite). By making a change of time variable to τ, where
$\frac{d\tau}{dt} = \sum_i u_i^2 r_i^2$, we can define $m_i = \frac{u_i r_i}{\sqrt{\sum_i u_i^2 r_i^2}}$ and $p_i = \frac{r_i^2}{2}$, we then have the system

$$\frac{dp}{d\tau} = Am \circ m. \tag{49}$$

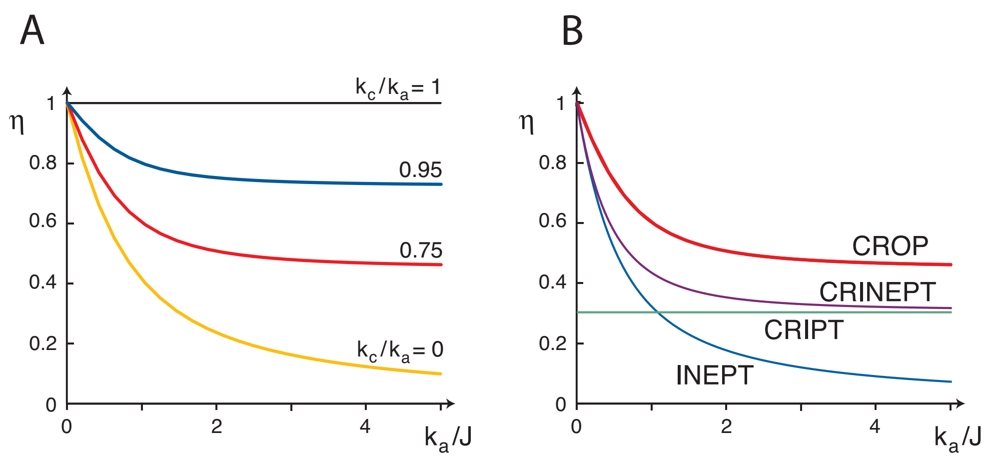


Figure 3: Figure A shows the transfer efficiency for CROP sequence as a function of $\frac{k_a}{J}$, for different values of $\frac{k_c}{k_a}$. Figure B shows efficiency of various state of the art pulse sequences as a function of $\frac{k_a}{J}$ for the transfer in Eq. (45) for $k_c = .75$. The CROP pulse sequences developed using optimal control of system in Eq. (45) provide the optimal transfer.

By changing u_i we change m_i and hence we treat m as control.

One possible transfer is to transfer from a given initial state $p = (1, 0, \ldots, 0)$ to maximum possible value p_n.

The reachable set takes the form ($diagA$ is diagonal of matrix A), for $\alpha_i > 0$

$$diag(\sum_i \alpha_i A m_i m_i^T) = diag(AM),$$

where M is a positive semidefinite (PSD) matrix. Thus the problem reduces to finding optimal PSD M_0 such that $diag(AM_0) = (-1, 0, \ldots, 0, \eta^2)$ for maximum possible η.

Remark 4 : Cone Separation Let $\mathcal{C} = diag(AM)$. Then $\mathcal{C}$ is a convex cone. It has non-empty interior as we can find m_i, $i = 1, \ldots, n$ such that $Am_i \circ m_i$ are independent, else they lie in a subspace annihilated by e. If $e_k \neq 0$ then $e^T Ay \circ y \neq 0$, where y is 1 in the k^{th} spot. Then $\sum_i \alpha_i Am_i \circ m_i$ for $\alpha_i > 0$ is an interior point. Note by Caratheodory's theorem, $\mathcal{B} = \{diag(\sum_i \alpha_i A m_i m_i^T), \|m_i\| = 1, \sum_i \alpha_i = 1, i = 1, \ldots n+1\}$ generates $\mathcal{C}$. Since A is negative definite $A + A' < 0$, we have $\mathcal{B}$ bounded away from zero and $\frac{\mathcal{B}}{\|\mathcal{B}\|}$ a compact set. For $y_i \in \mathcal{C}$ converging to y_i^*, $\frac{y_i}{\|y_i\|} \in \frac{\mathcal{B}}{\|\mathcal{B}\|}$ converges to $\frac{y_i^*}{\|y_i^*\|} \in \frac{\mathcal{B}}{\|\mathcal{B}\|}$. Hence $y_i^* \in \mathcal{C}$. $\mathcal{C}$ is a closed convex cone. Note $diag(AM_0)$ cannot be an interior point of $\mathcal{C}$ else we can proceed in direction $(0, \ldots, 0, 1)$ and improve the efficiency. Hence $x_0 = diag(AM_0)$ is a boundary point of $\mathcal{C}$ and there exists $\lambda = (\lambda_1, \ldots, \lambda_n)$ such that $\lambda^T x_0 = 0$ and $\lambda^T \mathcal{C} \leq 0$. Let $\lambda = diag(\Lambda)$.

Then we have

$$tr(\Lambda AM) \leq 0, \quad \Lambda A + A'\Lambda \leq 0 \tag{50}$$
$$tr((\Lambda A + A'\Lambda)M_0) = 0 \tag{51}$$

In the following, we consider few examples. Applying remark 4 to

$$A = \begin{bmatrix} -\xi & -1 \\ 1 & -\xi \end{bmatrix},$$

as in Eq. (**??**), we obtain $\lambda' x_0 = \lambda' \begin{bmatrix} -1 \\ \eta^2 \end{bmatrix} = 0$. For $\lambda = (\eta^2, 1)$,

$$\begin{bmatrix} \eta^2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -\xi & -1 \\ 1 & -\xi \end{bmatrix} + \begin{bmatrix} -\xi & 1 \\ -1 & -\xi \end{bmatrix} \begin{bmatrix} \eta^2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -2\xi\eta^2 & 1-\eta^2 \\ 1-\eta^2 & -2\xi \end{bmatrix}, \tag{52}$$

which is semidefinite, when

$$(\eta^{-1} - \eta) = 2\xi,$$

which gives $\eta = \sqrt{1+\xi^2} - \xi$.

As another example consider

$$A = \begin{bmatrix} -1 & a \\ b & -1 \end{bmatrix},$$

where $b > 0$ and $-2 < (a+b) < 2$.

$$\frac{1}{2}(\begin{bmatrix} \eta^2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & a \\ b & -1 \end{bmatrix} + \begin{bmatrix} -1 & a \\ b & -1 \end{bmatrix} \begin{bmatrix} \eta^2 & 0 \\ 0 & 1 \end{bmatrix}) = \begin{bmatrix} -\eta^2 & \frac{a\eta^2+b}{2} \\ \frac{a\eta^2+b}{2} & -1 \end{bmatrix}, \tag{53}$$

which is semidefinite, when

$$a\eta + b\eta^{-1} = 2, \tag{54}$$

and null vector is $m = (1, \eta)$. Eq. 54 implies that

$\eta = a^{-1} + \sqrt{\frac{1-ab}{a^2}}$, when a is negative. This choice of root ensures $\eta > 0$. When a is positive, we have $\eta = a^{-1} - \sqrt{\frac{1-ab}{a^2}}$ as $\dot{p}_1 < 0$ in Eq. 49 for this choice.

As another example consider Eq. (44),

$$A = \begin{bmatrix} -\xi & \chi\cos(\theta+\gamma) \\ \chi\cos(\theta-\gamma) & -\xi \end{bmatrix},$$

we have

$$\Lambda A + A'\Lambda = \begin{bmatrix} -2\xi\eta^2 & \chi(\eta^2\cos(\theta+\gamma)+\cos(\theta-\gamma)) \\ \chi(\eta^2\cos(\theta+\gamma)+\cos(\theta-\gamma)) & -2\xi \end{bmatrix}, \qquad (55)$$

which is semidefinite, when,

$$(\eta\cos(\theta+\gamma)+\eta^{-1}\cos(\theta-\gamma))^2 = \frac{4\xi^2}{\chi^2},$$

which gives maximum value of $\eta = \sqrt{1+\zeta^2} - \zeta$, when $\tan\gamma = \tan\theta\frac{1-\eta^2}{1+\eta^2}$, for $\zeta^2 = \frac{k_a^2-k_c^2}{k_a^2+k_c^2}$, as in Eq. (44).

As another example, consider

$$A = \begin{bmatrix} -\xi & -1 & 0 \\ 1 & -\xi & -1 \\ 0 & 1 & -\xi \end{bmatrix}.$$

for

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

we have

$$S = \Lambda A + A'\Lambda = \begin{bmatrix} -2\xi\lambda_1 & \lambda_2-\lambda_1 & 0 \\ \lambda_2-\lambda_1 & -2\xi\lambda_2 & 1-\lambda_2 \\ 0 & 1-\lambda_2 & -2\xi \end{bmatrix}.$$

Note $\begin{bmatrix} \lambda_1 & \lambda_2 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ \eta \end{bmatrix} = 0$ implies $\lambda_1 > 0$, we consider

$$uu^T = \Lambda A + A'\Lambda = S,$$

which is the case when S has a two-dimensional nullspace.

This is not possible as $S_{13} = 0$, while $u_1^2 = S_{11} \neq 0$ and $u_3^2 = S_{33} \neq 0$. This says that semidefinite S cannot be rank 1. Hence it only has a one-dimensional null space. The null vector m must satisfy

$$Am \circ m = \begin{bmatrix} -1 \\ 0 \\ \eta^2 \end{bmatrix},$$

which makes $m_2 = \frac{m_1-m_3}{2\xi}$ or $m_2 = 0$. On substituting we get

$$\frac{d}{dt}\begin{bmatrix} p_1 \\ p_3 \end{bmatrix} = \begin{bmatrix} -m_1^2 & (1+\xi^2)^{-1}m_1m_3 \\ (1+\xi^2)^{-1}m_1m_3 & -m_3^2 \end{bmatrix}. \tag{56}$$

From Eq. 42, we have $a = b = (1+\xi^2)^{-1}$. The maximum transfer is η^2, with $\eta = a^{-1} - \sqrt{a^{-1}(a^{-1}-b)} = 1+\xi^2 - \xi\sqrt{2+\xi^2}$.

Case $m_2 = 0$, produces no transfer.

3 Finite Time Optimal Control

We now consider the system,

$$\frac{d}{dt}\begin{bmatrix} r_1(t) \\ r_2(t) \end{bmatrix} = \begin{bmatrix} -\xi u_1^2 & -u_1u_2 \\ u_1u_2 & -\xi u_2^2 \end{bmatrix}\begin{bmatrix} r_1(t) \\ r_2(t) \end{bmatrix}. \tag{57}$$

We consider the problem of steering $(1,0)$ to maximum possible value of r_2 in finite time T. In the finite time case, the optimal return function $V(r_1, r_2, t)$ has explicit dependence on time and by definition

$$V(r_1, r_2, t) = \max_{u_1,u_2} V(r_1 + \delta t(-\xi u_1^2 r_1 - u_1u_2r_2), r_2 + \delta t(-\xi u_2^2 r_2 + u_1u_2r_1), t+\delta t).$$

Expanding in δt , we obtain the well known Hamilton Jacobi Bellman equation.

$$\frac{\partial V}{\partial t} + \max_{u_1,u_2}\begin{bmatrix} \frac{\partial V}{\partial r_1} & \frac{\partial V}{\partial r_2} \end{bmatrix}\begin{bmatrix} -\xi u_1^2 & -u_1u_2 \\ u_1u_2 & -\xi u_2^2 \end{bmatrix}\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = 0. \tag{58}$$

From equation (58), define the adjoint variables $(\lambda_1, \lambda_2) = (\frac{\partial V}{\partial r_1}, \frac{\partial V}{\partial r_2})$. Let $\mathbb{H} = -\lambda_1 r_1[\xi u_1^2 - (a-b)u_1u_2 + \xi ab u_2^2]$, where $a = \frac{\lambda_2}{\lambda_1}$ and $b = \frac{r_2}{r_1}$. Then equation (58) can be written as

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

For the finite time problem $\max_{u_1,u_2}\mathbb{H} > 0$. This implies $(a-b)^2 > 4\xi^2 ab$, i.e. $\frac{a-b}{2\xi ab}\frac{a-b}{2\xi} > 1$. We consider three separate cases for the problem.

1. **Case I:** If $(a-b) \leq 2\xi$, then the maximum of $\mathbb{H}$ is obtained for $u_2 = 1$ and $u_1 = \frac{a-b}{2\xi}$.
2. **Case II:** If $(a-b) > 2\xi$ and $\frac{a-b}{ab} > 2\xi$, then the maximum of $\mathbb{H}$ is obtained for $u_1 = 1$ and $u_2 = 1$.
3. **Case III:** If $b^{-1} - a^{-1} = \frac{a-b}{ab} \leq 2\xi$, then the maximum of $\mathbb{H}$ is obtained for $u_1 = 1$ and $u_2 = \frac{a-b}{2\xi ab}$.

From equation (58), the adjoint variables $(\lambda_1, \lambda_2) = (\frac{\partial V}{\partial r_1}, \frac{\partial V}{\partial r_2})$ satisfy the equations $\dot{\lambda_1} = -\frac{\partial \mathbb{H}}{\partial r_1}$ and $\dot{\lambda_2} = -\frac{\partial \mathbb{H}}{\partial r_2}$, i.e.

$$\frac{d}{dt}\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} \xi u_1^2 & -u_1u_2 \\ u_1u_2 & \xi u_2^2 \end{bmatrix}\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}, \tag{59}$$

where $(\lambda_1(T), \lambda_2(T)) = (0, 1)$. From equation (57, 59), we deduce that $V = \lambda_1 r_1 + \lambda_2 r_2$ is a constant for optimal trajectory and equals the optimal cost $r_2(T) = \lambda_1(0)$. Writing the equation for adjoint variables backward in time, let $\sigma = T - t$ then

$$\frac{d}{d\sigma}\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} -\xi u_1^2 & u_1u_2 \\ -u_1u_2 & -\xi u_2^2 \end{bmatrix}\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix},$$

where $(\lambda_1(\sigma), \lambda_2(\sigma))_{\sigma=0} = (0, 1)$. Now $u_1(\sigma)$ and $u_2(\sigma)$ should be chosen to maximize $\lambda_1(\sigma)|_{\sigma=T}$. Observe this is exactly the same optimization problem as (57), where the roles of u_1 and u_2 have been switched. From the symmetry of these two optimization problems, we then have

$$\begin{aligned} u_1^*(t) &= u_2^*(T-t) \\ r_1(t) = \lambda_2(T-t) \quad &; \quad r_2(t) = \lambda_1(T-t) \\ ab(\frac{T}{2}) = 1 \quad &; \quad V = 2r_1(\frac{T}{2})r_2(\frac{T}{2}) \\ (b^{-1} - a^{-1})(T-t) &= (a-b)(t) \end{aligned}$$

Observe from (57, 59), that $ab(t)$ is monotonically increasing and since $ab(0) = 0$ and $ab(\frac{T}{2}) = 1$, we have $ab(t) < 1$ for $t < \frac{T}{2}$ and therefore $\frac{(a-b)}{2\xi ab} > 1$ for $t < \frac{T}{2}$ (else $\frac{a-b}{2\xi} > 1$ and $ab < 1$ implies the stated). Therefore $u_2^*(t) = 1$ for $t < \frac{T}{2}$. Depending on $a(0)$, we have two cases. **Case A** In this case $\frac{a(0)}{2\xi} \geq 1$. Then we start in the case II discussed above and verify that in this case $a - b$ is increasing for $ab < 1$, implying $\frac{a-b}{2\xi ab} > 1$ Therefore we stay in this case for all $t \in [0, \frac{T}{2}]$ and therefore $u_1^* = u_2^*(t) = 1$ for all t. Since $b(0) = 0$, we have $b(\frac{T}{2}) = \tan\frac{T}{2}$. Similarly,

$$a(\frac{T}{2}) = \frac{a(0) + \tan(\frac{T}{2})}{1 - a(0)\tan(\frac{T}{2})}.$$

If $ab(\frac{T}{2}) = 1$, then above equation implies that $\tan(T) \leq \frac{1}{2\xi}$, as $a(0) > 2\xi$.

Case B If $\frac{a(0)}{2\xi} < 1$, then $u_1^*(0) = \frac{a(0)}{2\xi}$ and the system begins in case I. Let $\kappa(t)$ satisfy

$$\frac{d\kappa}{dt} = \frac{\kappa^2 - 2\kappa + 1}{2\xi} - 2\xi\kappa, \quad \kappa(0) = 0.$$

The solution to this equation is given by $\kappa(t) = 1 + 2\xi^2 - 2\xi\sqrt{1+\xi^2}\coth(\sqrt{1+\xi^2}t + 2\beta)$, where $\sinh(\beta) = \xi$. It can be verified that in case I, the optimal trajectory satisfies $\frac{b}{a}(t) =$

$\kappa(t)$. $a-b$ is increasing. After time τ, $\frac{a-b}{2\xi}$ becomes equal to 1 and the system switches to case II. Putting $\frac{a-b}{2\xi}=1$ and $\frac{b}{a}(t)=\kappa(t)$, we get $\frac{r_2(\tau)}{r_1(\tau)}=\frac{2\xi\kappa(\tau)}{1-\kappa(\tau)}$ (denote this ratio by $\tan\theta_1$, see Fig 4, Panel B). Then again by symmetry at time $T-\tau$ we have $\frac{1}{2\xi}(\frac{1}{b}-\frac{1}{a})=1$ and the system switches from case II to case III. In case III, verify $\frac{b}{a}(t)=\kappa(T-t)$ and the switching to this case occurs at $\tan\theta_2=\frac{r_2}{r_1}=\frac{1-\kappa(\tau)}{2\xi}$. Thus the system spends $T-2\tau$ in region II. Then we have

$$T-2\tau=\tan^{-1}\frac{1-\kappa(\tau)}{2\xi}-\tan^{-1}\frac{2\xi\kappa(\tau)}{1-\kappa(\tau)}.$$

We now derive an explicit expression for $r_2(T)$. For $t\geq T-\tau$,

$$V(t)=\sqrt{r_2^2(t)+\kappa(T-t)r_1^2(t)}, \tag{60}$$

is constant along the system trajectories and equals the optimal return function $r_2(T)$. At $t=T-\tau$, we have $\frac{r_2(T-\tau)}{r_1(T-\tau)}=\tan\theta_2=\frac{1-\kappa(\tau)}{2\xi}$ and therefore from (60), we have

$$V(T-\tau)=R_1\sqrt{\sin^2\theta_2+\cos^2\theta_2-2\xi\sin\theta_2\cos\theta_2}, \tag{61}$$

where $R_1=\sqrt{r_1^2(t)+r_2^2(t)}$ for $t=T-\tau$. Also note $V(\frac{T}{2})=2r_1(\frac{T}{2})r_2(\frac{T}{2})$. At time $t=\frac{T}{2}$, we then have $\frac{r_2}{r_1}=\tan(\frac{\theta_1+\theta_2}{2})$ and therefore

$$V(\frac{T}{2})=R_2^2\sin(\theta_1+\theta_2) \tag{62}$$

where $R_2=\sqrt{r_1^2(\frac{T}{2})+r_2^2(\frac{T}{2})}$. Note between $\frac{T}{2}$ and $T-\tau$, the system evolves under $u_1=u_2=1$. Therefore $R_1=R_2\exp(-(\frac{T}{2}-\tau))$. Since V is constant, equating (61) and (62), we get equation $V(T-\tau)=V(\frac{T}{2})=\eta_T$ (63).

At time τ, the optimal trajectory (r_1,r_2) passes from phase I to II and makes an angle θ_1 with the r_1 axis and at time $T-\tau$ the optimal trajectory passes from phase II to phase III and makes an angle θ_2 with the r_1 axis (see Fig. 4). The optimal efficiency η_T for the finite time T is expressed in terms of these angles as

$$\eta_T=\frac{\exp(\xi(\theta_1-\theta_2))(1-\xi\sin 2\theta_2)}{\sin(\theta_1+\theta_2)}. \tag{63}$$

In the limit, T goes to infinity $\tau=\frac{T}{2}$ and $\theta_1=\theta_2=\tan^{-1}\sqrt{1+\xi^2}-\xi$ and η_T approaches η in (**??**). This corresponds to the unconstrained time case we discussed initially.

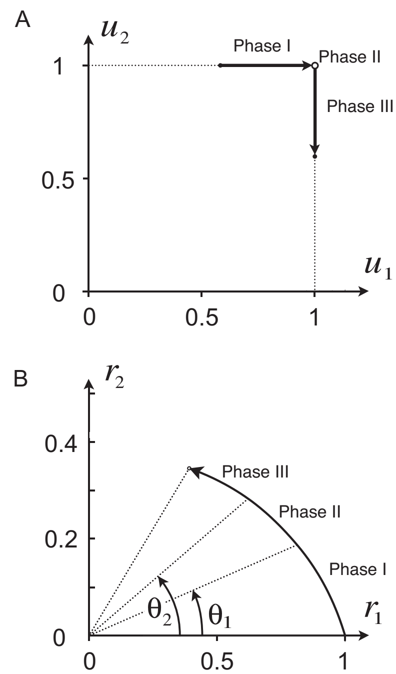


Figure 4: Phase trajectory of the controls u_1 and u_2 (panel A) and $\vec{r}(t)$ (panel B) for a finite-time ROPE sequence ($\xi = 1$).

We now study control of coupled spin dynamics in presence of longitudinal relaxation. For this we consider the system,

$$\frac{d}{dt}\begin{bmatrix} z_1 \\ x_1 \\ x_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} -k_1 & -u(t) & 0 & 0 \\ u(t) & -k & -J & 0 \\ 0 & J & -k & v(t) \\ 0 & 0 & -v(t) & -k_2 \end{bmatrix}\begin{bmatrix} z_1 \\ x_1 \\ x_2 \\ z_2 \end{bmatrix}, \quad k_1, k_2 \leq k \tag{64}$$

where goal is to drive the system from $(1, 0, 0, 0)'$ to maximum possible value of z_2. Since the controls can be made much larger than the natural parameters in the system, we define $r_1 = \sqrt{z_1^2 + x_1^2}$, $r_2 = \sqrt{z_1^2 + x_1^2}$, $\tan\theta_1 = \frac{z_1}{x_1}$ and $\tan\theta_2 = \frac{z_2}{x_2}$. Writing an equation for r_1 and r_2, gives us scaled equation

$$\frac{d}{dt}\begin{bmatrix} r_1(t) \\ r_2(t) \end{bmatrix} = \begin{bmatrix} -(\frac{k_1}{J} + \xi_1 u_1^2) & -u_1 u_2 \\ u_1 u_2 & -(\frac{k_2}{J} + \xi_2 u_2^2) \end{bmatrix}\begin{bmatrix} r_1(t) \\ r_2(t) \end{bmatrix}, \; \xi_i = \frac{k - k_i}{J} \tag{65}$$

In the finite time case, the optimal return function $V(r_1, r_2, t)$ has explicit dependence on time and by definition,

$$\frac{\partial V}{\partial t} + \max_{u_1,u_2} \begin{bmatrix} \frac{\partial V}{\partial r_1} & \frac{\partial V}{\partial r_2} \end{bmatrix} \begin{bmatrix} -(\frac{k_1}{J} + \xi_1 u_1^2) & -u_1 u_2 \\ u_1 u_2 & -(\frac{k_2}{J} + \xi_2 u_2^2) \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = 0. \tag{66}$$

Let $\mathbb{H} = -\lambda_1 r_1[(\frac{k_1}{J} + \xi_1 u_1^2) - (a-b)u_1 u_2 + (\frac{k_2}{J} + \xi_2 u_2^2)ab]$, where $a = \frac{\lambda_2}{\lambda_1}$ and $b = \frac{r_2}{r_1}$. Then equation (66) can be rewritten as

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

For the finite time problem $\frac{\partial V}{\partial t} - \lambda_1 r_1 \frac{k_1}{J} - \lambda_2 r_2 \frac{k_2}{J} < 0$. This implies $\frac{a-b}{2\xi_2 ab}\frac{a-b}{2\xi_1} > 1$. We consider three separate cases for the problem

1. **Case I:** If $(a-b) \leq 2\xi_1$, then the maximum of $\mathbb{H}$ is obtained for $u_2 = 1$ and $u_1 = \frac{a-b}{2\xi_1}$.
2. **Case II:** If $(a-b) > 2\xi_1$ and $\frac{a-b}{ab} > 2\xi_2$, then the maximum of $\mathbb{H}$ is obtained for $u_1 = 1$ and $u_2 = 1$.
3. **Case III:** If $b^{-1} - a^{-1} = \frac{a-b}{ab} \leq 2\xi_2$, then the maximum of $\mathbb{H}$ is obtained for $u_1 = 1$ and $u_2 = \frac{a-b}{2\xi_2 ab}$.

From equation (66), the adjoint variables $(\lambda_1, \lambda_2) = (\frac{\partial V}{\partial r_1}, \frac{\partial V}{\partial r_2})$ satisfy the equations $\dot{\lambda}_1 = -\frac{\partial \mathbb{H}}{\partial r_1}$ and $\dot{\lambda}_2 = -\frac{\partial \mathbb{H}}{\partial r_2}$, i.e.

$$\frac{d}{dt}\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} (\frac{k_1}{J} + \xi_1 u_1^2) & -u_1 u_2 \\ u_1 u_2 & (\frac{k_2}{J} + \xi_2 u_2^2) \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}, \tag{67}$$

where $(\lambda_1(T), \lambda_2(T)) = (0, 1)$. From equation (65, 67), we deduce that $V = \lambda_1 r_1 + \lambda_2 r_2$ is a constant for optimal trajectory and equals the optimal cost $r_2(T) = \lambda_1(0)$.

Observe from (65, 67), that $ab(t)$ is monotonically increasing. $(a-b)$ is increasing in case I and II. $(a-b)/ab$ is decreasing in case II and III. $\frac{(a-b)}{2\xi_2 ab} > 1$ to begin with. Therefore $u_2^*(t) = 1$ to begin with. Depending on $a(0)$, we have following cases. If we start in case I, we can increase $a-b$ and switch to case II. If $(a-b)/ab$ decreases below a threshold, we switch to case III and stay there. We may start in case II and switch to case III and stay there or we may always stay in II. In more detail,
If $\frac{a(0)}{2\xi} < 1$, then $u_1^*(0) = \frac{a(0)}{2\xi}$ and the system begins in case I. Let $\kappa_1(t) = \frac{b}{a}$. It satisfies

$$\frac{d\kappa_1}{dt} = \frac{\kappa_1^2 - 2\kappa_1 + 1}{2\xi_1} - 2\xi_1\kappa_1, \quad \kappa_1(0) = 0.$$

The solution to this equation is given by $\kappa_1(t) = 1 + 2\xi_1^2 - 2\xi_1\sqrt{1+\xi_1^2}\coth(\sqrt{1+\xi_1^2}t + 2\beta)$, where $\sinh(\beta) = \xi_1$. After time τ_1, $\frac{a-b}{2\xi_1}$ becomes equal to 1 and the system switches to case II. Putting $\frac{a-b}{2\xi_1} = 1$ and $\frac{b}{a}(t) = \kappa_1(t)$, we get $\frac{r_2(\tau_1)}{r_1(\tau_1)} = \frac{2\xi\kappa_1(\tau_1)}{1-\kappa_1(\tau_1)}$ (denote this ratio by $\tan\theta_1$).

In the case $\frac{a(0)}{2\xi} > 1$. Then we start in the case II discussed above and verify that in this case $a - b$ is increasing and $a - b/ab$, decreasing. Therefore we stay in this case for where $u_1^* = u_2^*(t) = 1$, before we may switch to case III.

In phase III, Let $\kappa_2(t) = \frac{b(T-t)}{a(T-t)}$. It satisfies

$$\frac{d\kappa_2}{dt} = \frac{\kappa_2^2 - 2\kappa_2 + 1}{2\xi_2} - 2\xi_2\kappa_2, \quad \kappa_2(0) = 0.$$

The solution to this equation is given by $\kappa_2(t) = 1 + 2\xi_2^2 - 2\xi_2\sqrt{1+\xi_2^2}\coth(\sqrt{1+\xi_2^2}t + 2\beta)$, where $\sinh(\beta) = \xi_2$.

At time $T-\tau_2$, when we switch to case III, we have $\frac{a-b}{2\xi_2 ab} = 1$ with $\frac{b}{a}(T-\tau_2) = \kappa_2(\tau_2)$. The switching to this case occurs at $\tan\theta_2 = \frac{r_2}{r_1} = \frac{1-\kappa_2(\tau_2)}{2\xi_2}$. Thus the system spends $T - \tau_1 - \tau_2$ in region II. Then we have

$$\begin{aligned} T - \tau_1 - \tau_2 &= \tan^{-1}\frac{1-\kappa_2(\tau_2)}{2\xi_2} - \tan^{-1}\frac{2\xi_1\kappa_1(\tau_1)}{1-\kappa_1(\tau_1)} &(68)\\ &= \tan^{-1}\frac{1-\kappa_1(\tau_1)}{2\xi_1} - \tan^{-1}\frac{2\xi_2\kappa_2(\tau_2)}{1-\kappa_2(\tau_2)}. &(69)\end{aligned}$$

where the last equation follows from duality of r and λ, where we drive the λ equation backwards in time from initial value $\lambda(T) = (0, 1)$ to maximum possible value of $\lambda_1(0)$. Given $\xi_1 < \xi_2$, we stay all the time in case II, when $T \le \tan^{-1}\frac{1}{2\xi_2}$. For $T > \tan^{-1}\frac{1}{2\xi_2}$, we have II for time $T - \tau_2$ followed by III for time τ_2, for a total time $\tau_2 + \tan^{-1}\frac{1-\kappa_2(\tau_2)}{2\xi_2}$, as long as

$$\tan^{-1}\frac{1-\kappa_2(\tau_2)}{2\xi_2} + \tan^{-1}\frac{2\xi_2\kappa_2}{1-\kappa_2(\tau_2)} \le \tan^{-1}\frac{1}{2\xi_1}$$

If T is made even larger, we see all three phases. When $\xi_2 < \xi_1$, we stay all the time in case II, when $T \le \tan^{-1}\frac{1}{2\xi_1}$. For larger $T \ge \tan^{-1}\frac{1}{2\xi_1}$, we have I for time τ_1 followed by II for time $T - \tau_1$, as long as

$$\tan^{-1}\frac{1-\kappa_1(\tau_1)}{2\xi_1} + \tan^{-1}\frac{2\xi_1\kappa_1}{1-\kappa_1(\tau_1)} \le \tan^{-1}\frac{1}{2\xi_2}.$$

For even larger T, we have all three stages.