TIME OPTIMAL CONTROL OF A RIGHT INVARIANT SYSTEM ON A COMPACT LIE GROUP

JODY-LYNN STORM

ABSTRACT. In this paper we will study the pulse sequences in NMR spectroscopy and quantum computing as a time control problem. Radio frequency pulses are used to execute a unitary transfer of state. Sequences of pulses should be as short as possible to minimize decoherence. We model the problem as a controllable right invariant system on a compact Lie group. We investigate the minimum time required to steer the system from an initial point to a specified final point.

1. Introduction

Richard Feynman was one of the first individuals to recognize that there is an exponential slowdown when classical computers are used to simulate quantum systems. He went on to suggest that the use of quantum computers to simulate quantum systems should be exponentially faster than their classical counterparts.

In 1994, Peter Shor devised an algorithm for quantum computers that could factor integers exponentially faster than the best-known classical factoring algorithm. Shor's algorithm means, that if quantum computers could be built, then cryptographic systems based on factoring, like those commonly used in banking, could be broken. Expanding on Shor's algorithm, in 1996, Lov Grover created a quantum algorithm that could search databases faster than anything possible on classical computers.

A technique that has been used to build elementary quantum computers is nuclear magnetic resonance (NMR), a physical phenomenon based upon the quantum mechanical magnetic properties of an atom's nucleus. The elementary particles, neutrons and protons, composing an atomic nucleus have the intrinsic quantum mechanical property of angular momentum, called spin. In NMR, unitary transformations are used to manipulate an ensemble of nuclear spins. However, the sequence of pulses that generate the desired unitary operator should be as short as possible in order to minimize the effects of decoherence; a major obstacle faced in building quantum computers. Quantum systems want to wander from their computational path and entangle with the rest of the environment. Transferring quantum states

Date: October 28, 2008.

I would like to thank Dr. George Wilkens for his patience, knowledge and support.

as quickly as possible helps minimize the impact of decoherence, and thus motivates the minimum time problem.

In this paper we will study the design of pulse sequences in NMR spectroscopy (measurement of a quantity as function of either wavelength or frequency) as a time optimal control problem on a compact Lie group. In quantum computation, allowed operations are unitary matrices, which are effectively rotations, so our primary interest will be in the unitary group $U(n)$, although the discussion is general enough to include other compact Lie groups such as $SO(n)$. The question we will address is the problem of time control. If we are given a controllable right invariant system on a compact Lie group, what is the minimum time required to steer the system from an initial point to a specified final point?

We will present a mathematical formulation for the problem of finding the shortest pulse sequences in NMR. We will also show that the problem of computing minimum time to produce a unitary propagator can be reduced to finding the shortest length paths on certain coset spaces.

Control Systems. The time evolution of a quantum system is described by the time dependent Schroedinger equation

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

where $H(t)$ is the Hamiltonian and $U(t)$ is the unitary displacement. We can write the Hamiltonian as

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

where H_d is the part of the Hamiltonian internal to the system, called the *drift*, and $\sum_{i=1}^{m} v_i(t)H_i(t)$ is the part of the Hamiltonian that can be externally changed, which we call the *control*.

We are interested in finding the minimum time required to transfer the system

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

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

2. Preliminaries

Definition. A *smooth (or differentiable) manifold* of dimension n is a second-countable, Hausdorff topological space M with a collection of pairs (U_α, ϕ_α), called charts, where U_α is an open subset of M and $\phi_\alpha : U_\alpha \to \mathbb{R}^n$ so that:

(a) Each ϕ_α is a homeomorphism of U_α onto an open subset V_α of $\mathbb{R}^n$.
(b) $\bigcup_\alpha U_\alpha = M$.

(c) For every α, β the transition functions $\phi_{\alpha\beta} = \phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha(U_\alpha \cap U_\beta) \to \phi_\beta(U_\alpha \cap U_\beta)$ are smooth, in the sense of smooth functions between subsets of $\mathbb{R}^n$, *i.e.* $\phi_{\alpha\beta}$ is $\mathcal{C}^\infty$. In this case the charts (U_α, ϕ_α) and (U_β, ϕ_β) are called *compatible*.

(d) The family $\{(U_\alpha, \phi_\alpha)\}$ is maximal relative to the conditions (b) and (c).

Such a family of sets and maps satisfying (b),(c), and (d) constitutes a *differentiable structure* on M.

Example. The Euclidean space $\mathbb{R}^n$ is an n-dimensional manifold whose differentiable structure consists of all charts that are compatible with the chart (U, ϕ) where $U = \mathbb{R}^n$ and $\phi : U \to \mathbb{R}^n$ is the identity map.

Definition. Let p be a point of a manifold M. Let $\mathcal{F}(M)$ be the set of all smooth real-valued functions on M. A *tangent vector* to M at p is a real-valued function $v : \mathcal{F}(M) \to \mathbb{R}$ that satisfies:

(1) $v(af + bg) = av(f) + bv(g)$,

(2) $v(fg) = v(f)g(p) + f(p)v(g)$, for $a, b \in \mathbb{R}$, $f, g \in \mathcal{F}(M)$.

For each $p \in M$, let $T_p(M)$ be the set of all tangent vectors to M at p. Then under the operations

$$(v + w)(f) = v(f) + w(f),$$
$$(av)(f) = av(f),$$

the set $T_p(M)$ is a real vector space, called the *tangent space* of M at p.

By a standard proposition in differential geometry, the dimension of $T_p(M)$ equals that of M. The *tangent bundle* of M is the union, over all p in M, of the tangent spaces $T_p(M)$. The tangent bundle is denoted by TM.

Definition. Let $F\colon M \to N$ be a function, where M and N are smooth manifolds. Then F is *smooth* (*differentiable*) if for every smooth, real valued function $f \in \mathcal{F}(N)$ the composite function $f \circ F \in \mathcal{F}(M)$. Let $p \in M$. The *derivative of* F *at* p is the function $dF_p\colon T_p(M) \to T_{F(p)}(N)$ such that for every $v_p \in T_p(M)$ and every $f \in \mathcal{F}(N)$, $dF_p(v_p)f = v_p\,(f \circ F)$.

A *vector field* X on a manifold M is a function that assigns to each point $p \in M$ a tangent vector X_p to M at p. Thus $X : M \to TM$ with $X_p \in T_pM$. If X is a vector field on M and $f \in \mathcal{F}(M)$, then Xf denotes the real-valued function on M given by

$$Xf(p) = X_p(f), \quad \forall p \in M.$$

The vector field X is called *smooth* if the function Xf is smooth for all $f \in \mathcal{F}(M)$.

Definition. Let G be a smooth manifold. Then G is called a *Lie Group* if:

(a) G is a group

(b) The group operations $G \times G \to G$ defined by $(x, y) \mapsto xy$ and $G \to G$ defined by $x \mapsto x^{-1}$ are smooth (differentiable) functions.

Example. The set $\mathbb{R}^n$ is a Lie group under vector addition.

Definition. A *Lie algebra* $\mathfrak{g}$ is a vector space over a field F with a binary operation $[.\,,.] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ called the *Lie bracket*, which satisfies the following axioms:

(a) Bilinearity

$$[aX + bY, Z] = a[X, Z] + b[Y, Z], \quad [Z, aX + bY] = a[Z, X] + b[Z, Y]$$

for all scalars $a, b \in F$ and all $X, Y, Z \in \mathfrak{g}$.

(b) Skew-symmetry:

$$[X, Y] = -[Y, X]$$

for all $X, Y \in \mathfrak{g}$. When F is a field of characteristic two, then:

$$[X, X] = 0$$

for all $X \in \mathfrak{g}$

(c) The Jacobi identity:

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$$

for all $X, Y, Z \in \mathfrak{g}$.

Example. The set of smooth vector fields on a manifold M is a vector space over $\mathbb{R}$. Given a pair of smooth vector fields X and Y, we define their commutator to be the vector field $[X, Y]$ whose action on the smooth function f is given by $[X, Y]\, f = Y(X\, f) - X(Y\, f)$. Direct calculations confirm that the commutator satisfies the axioms of a Lie bracket and thus provides a Lie algebra structure to the smooth vector fields on M.

Proposition 1. *Let X be a smooth vector field on a smooth manifold M, and let $p \in M$. Then there exists an open neighborhood U of p, an open interval I around 0, and a smooth mapping $\phi : I \times U \to M$ such that the curve $\gamma_q : I \to M$ defined by $\gamma_q(t) = \phi(t, q)$ for $q \in U$ is the unique curve that satisfies $\frac{\partial \phi}{\partial t} = X_{\gamma_q(t)}$ and $\gamma_q(0) = q$. In this case, we call γ_q an integral curve of the vector field X.*

If t is kept constant, the above proposition shows that the assignment $q \to \gamma_q(t)$ defines a function $\phi_t : U \to M$ on a neighborhood U of p. We call ϕ_t the *flow* of X. The flow has the following properties:

(a) ϕ_0 is the identity map of U
(b) $\phi_s \circ \phi_t = \phi_{s+t}$ for all $s, t, s + t \in I$
(c) each flow is a diffeomorphism with $\phi_t^{-1} = \phi_{-t}$.

Definition. A vector field X on a Lie group G is *right-invariant* if $X \circ R_a = dR_a(X)$ for all $a \in G$, or more explicitly $X_{ga} = (dR_a)_g(X_g)$ for all $a, g \in G$, where $R_a : G \to G$ defined by $R_a(g) = ga$ is called *right translation* on G. Left-invariant vector fields and left translation are defined similarly. A vector field that is both left-invariant and right-invariant is called *bi-invariant*.

We note the following:

- A right-invariant vector field is determined by its value at the identity e of the Lie Group G since $X_a = dR_a(X_e)$ for all $a \in G$. Also, since multiplication in G is smooth, so is a right-invariant vector field.
- Let $L(G)$ denote the set of right-invariant vector fields on G. This vector subspace is closed under the operation of commutator, and thus is a subalgebra of the Lie algebra of vector fields on G.

Throughout this paper we will let G be a compact and connected semi-simple Lie group (see appendix) with a bi-invariant metric $<,>_G$, *i.e.* left and right translation preserves $<,>_G$. Let e be the identity element of G. Let K be a compact subgroup of G. Let $L(G)$ and $L(K)$ be the Lie algebras of the right-invariant vector fields on G and K, respectively. A direct argument shows that $L(K)$ is a subalgebra of $L(G)$. The first of the above bullet-points shows that the vector spaces $L(G)$ and $L(K)$ are canonically isomorphic to their respective tangent spaces $T_e(G)$ and $T_e(K)$. The canonical isomorphism induces Lie algebra structures on $T_e(G)$ and $T_e(K)$, and we denote these Lie algebras by $\mathfrak{g}$ and $\mathfrak{k}$. Note that $\mathfrak{k}$ is a subalgebra of $\mathfrak{g}$.

Definition. A *right action* of a Lie group G on a manifold M is a smooth map $\lambda : M \times G \to M$ such that $\lambda(m, e) = m$ and $\lambda(m, ab) = \lambda(\lambda((m, b), a))$ for all $a, b \in G$ and $m \in M$. We often denote $\lambda(m, g)$ by mg.

Definition.

(a) An action is called *transitive* if for any $m, n \in M$ there exists $g \in G$ such that $mg = n$.
(b) Let $m \in M$. The set $G_m = \{g \in G : mg = m\}$ is called the *isotropy group* at m.
(c) The *orbit* of a point $m \in M$ is the set $mG = \{mg : g \in G\}$.

Given a Lie group G and a closed Lie subgroup $K \subset G$, there is a natural differentiable structure on the set $G/K = \{Kg : g \in G\}$ of all right cosets of $K \in G$. The resulting manifold is called the *coset manifold*. The map $G/K \times G \to G/K$ defined by $(Kg_1, g_2) = Kg_1g_2$ for $g_1, g_2 \in G$ is called the *natural action* of G on G/K, and this action is transitive. We call a coset manifold with this transitive action a *homogeneous space*.

Consider the direct sum decomposition $\mathfrak{g} = \mathfrak{m} \oplus \mathfrak{k}$ where $\mathfrak{m} = \mathfrak{k}^{\perp}$. Let $o = \pi(e)$ be the origin of the homogeneous space G/K. Recall that G/K admits the structure of a differentiable manifold.

There exists a neighborhood of $0 \in \mathfrak{m}$ which is mapped diffeomorphically onto a neighborhood of o by the map $\pi \circ \exp|_{\mathfrak{m}}$ (for our purposes it suffices to consider matrix Lie groups where exp denotes the usual exponential map). Thus we have an identification between $\mathfrak{m}$ and the tangent space $T_o(G/K)$.

The group G acts on $\mathfrak{g}$ by the *adjoint action* $Ad_G : \mathfrak{g} \times G \to \mathfrak{g}$ defined as follows. Given $U \in G$ and $X \in \mathfrak{g}$,

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

where we write $Ad_U(X)$ to denote $Ad_G(X, U)$.

Example. Let $G = SU(n) = \{A \in GL(n, \mathbb{C}) : AA^{\dagger} = I, \det A = 1\}$ and $\mathfrak{g} = \mathfrak{su}(n) = \{A \in M(n, \mathbb{C}) : A^{\dagger} = -A, \operatorname{tr} A = 0\}$. Then for $A, B \in \mathfrak{su}(n)$, the bi-invariant metric on $SU(n)$ may be represented by $\langle A, B \rangle_G = \operatorname{tr}(A^{\dagger} B)$. Now fix $U \in G$. If $A \in \mathfrak{su}(n)$, then $Ad_U(A) = U^{\dagger} A U$.

The decomposition $\mathfrak{g} = \mathfrak{m} \oplus \mathfrak{k}$ satisfies the following relations,

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

If, in addition,

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

then we call G/K a *Riemannian Symmetric space.*

Example. Let $G = SO(3)$ and let $K = SO(2)$. Fix $m = e_3 = (0, 0, 1) \in \mathbb{R}^3$. Then the isotropy group $G_m = K = SO(2)$ and the orbit $G_m = G/K = S^2$ is the Riemannian symmetric space of the 2-sphere.

Suppose $\mathfrak{h} \subset \mathfrak{m}$ is a subalgebra of $\mathfrak{g}$, *i.e.* $[\mathfrak{h}, \mathfrak{h}] \subset \mathfrak{h}$. Since $[\mathfrak{m}, \mathfrak{m}] \subset \mathfrak{k}$, we have that $\mathfrak{h}$ is abelian, *i.e.* $[\mathfrak{h}, \mathfrak{h}] = 0$.

Theorem 1. *If $\mathfrak{h}$ and $\mathfrak{h}'$ are two maximal abelian subalgebras of $\mathfrak{m}$. Then*

(a) *There exists an element $X \in \mathfrak{h}$ whose centralizer in $\mathfrak{m}$ is equal to $\mathfrak{h}$.*
(b) *There is an element $k \in K$ such that $Ad_k(\mathfrak{h}) = \mathfrak{h}'$.*
(c) $\mathfrak{m} = \bigcup_{k \in K} Ad_k(\mathfrak{h})$.

By the above theorem, the maximal abelian subalgebras of $\mathfrak{m}$ are all conjugate by Ad_K. Hence they all have the same dimension. We call this dimension the *rank* of the Riemannian Symmetric space G/K. The maximal abelian subalgebras of $\mathfrak{m}$ are called the *Cartan subalgebras* of the pair $(\mathfrak{g}, \mathfrak{k})$.

Corollary 1. *Let G/K be a Riemannian symmetric space. Let $\mathfrak{h}$ be a Cartan subalgebra of the pair $(\mathfrak{g}, \mathfrak{k})$ and define $A = \exp(\mathfrak{h}) \subset G$. Then $G = KAK$.*

Proof. $G = KP$, where $P = \exp(\mathfrak{m}) = \exp(\cup_{k\in K} Ad_k(\mathfrak{h})) = \cup_{k\in K} Inn_k(\exp(\mathfrak{h})) = \cup_{k\in K} Inn_k(A) \subset KAK$. Thus $G = KKAK = KAK$. □

The space G/K is a union of maximal abelian subgroups $Inn_k(A)$, $k \in K$. We call these maximal subgroups *maximal tori* (see appendix).

Consider the following control system on G (here we assume that G is a matrix Lie group).

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

The Lie algebra generated by $\{H_d, H_1, ..., H_m\}$ is denoted $\{H_d, H_1, ..., H_m\}_{LA}$. Assume that $\mathfrak{g} = \{H_d, H_1, ..., H_m\}_{LA}$. Since G is compact and connected, G is controllable, *i.e.* we can join e with any other point of G by concatenation of solutions $U(t)$ to (1). Let $\mathfrak{k} = \{H_1, ..., H_m\}_{LA}$ and let $K = \exp\{H_1, ..., H_m\}_{LA}$ be the Lie subgroup generated by $\mathfrak{k}$. We assume that K is closed and that $H_d \in \mathfrak{m}$. Recall that $\mathfrak{g} = \mathfrak{m} \oplus \mathfrak{k}$ where $\mathfrak{m} = \mathfrak{k}^{\perp}$. We will also assume that $Ad_K(\mathfrak{m}) \subset \mathfrak{m}$, in which case the homogeneous space G/K is called *reductive*. All of the examples will fall into this category.

Notation: Let $\mathcal{C}$ be the set of all locally bounded measurable functions defined on the interval $[0, \infty)$ with values in $\mathbb{R}^m$. Let $\mathcal{C}(T)$ denote the restriction of $\mathcal{C}$ to the interval $[0, T]$. Assume throughout that $v = (v_1, ..., v_m) \in \mathcal{C}$ in equation (1). For $v \in \mathcal{C}$, let $U(t)$ denote the solution to equation (1) such that $U(0) = U_0$. If there exists a $t \in [0, T]$ such that $U(t) = U'$ we say that the control v steers U_0 into U' in t units of time and U' is *reachable* from U_0 at time t.

Definition. The set of all $U' \in G$ reachable from U_0 at time t will be denoted by $R(U_0, t)$. Let

$$\mathbf{R}(U_0, T) = \bigcup_{0\le t\le T} R(U_0, t)$$

$$\mathbf{R}(U_0) = \bigcup_{0\le t\le \infty} R(U_0, t).$$

We call $\mathbf{R}(U_0)$ the *reachable set* of U_0.

Note. By right-invariance we have the following relations

$$R(U_0, T) = R(e, T)U_0, \qquad \mathbf{R}(U_0, T) = \mathbf{R}(e, T)U_0, \qquad \mathbf{R}(U_0) = \mathbf{R}(e)U_0.$$

It may not be the case that $\mathbf{R}(U_0, T)$ is closed. Denote the closure of $\mathbf{R}(U_0, T)$ by $\overline{\mathbf{R}(U_0, T)}$.

Definition. Let $U_F \in G$ and define

$$t^*(U_F) = \inf\{t \ge 0 : U_F \in \overline{\mathbf{R}(e, t)}\}$$

$$t^*(KU_F) = \inf\{t \ge 0 : \text{for some } k \in K, kU_F \in \overline{\mathbf{R}(e, t)}\}.$$

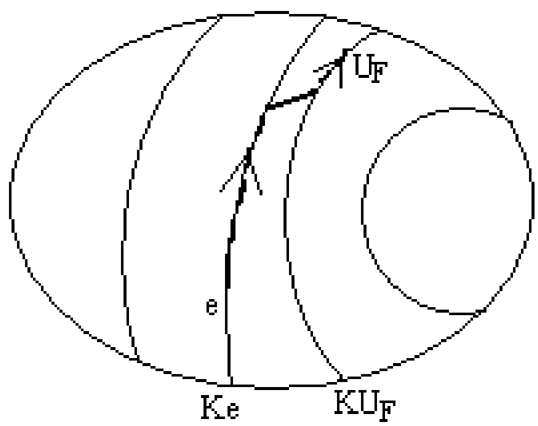


FIGURE 1

We call $t^*(U_F)$ and $t^*(KU_F)$ the *infimizing times* of U_F and KU_F respectively.

3. Time Optimal control

We will now show that if $U_F \in K$, then $t^*(U_F) = 0$. To illustrate this we will let v in equation (1) be arbitrarily large, allowing us to move on K as quickly as we wish. As $\|v\| \to \infty$, we can come arbitrarily close to any point in K, in an arbitrarily small amount of time. This can be accomplished with very little influence from H_d. Similarly, we will have $t^*(U_F) = t^*(kU_F)$ for $k \in K$. Thus computing $t^*(U_F)$ is reduced to finding the minimum time to steer the system (1) from Ke to KU_F.

The above figure shows the time optimal path between e and U_F belonging to G. The dashed lines represent the fast portion of the curve corresponding to movement in the cosets Ke and KU_F. The solid line represents the slow portion of the path connecting the two cosets. In NMR the dashed line in a single coset represents a *pulse* and the solid line between cosets represents the *drift*, or slow portion of the curve (also known as the evolution of couplings in NMR). The minimum time problem corresponds to finding the shortest path between cosets, hence, the shortest path in the space G/K.

Lemma 1. *Let $\mathcal{C}(T)$ be given the topology of weak convergence. For each $v \in \mathcal{C}(T)$, let $U : [0, T] \to G$ be the solution to the right-invariant control system* (1) *that is determined v. Then the mapping $(v, t) \mapsto U(t)$ from $\mathcal{C}(T) \times [0, T]$ into G is continuous.*

We now use Lemma 1 to prove the following

Lemma 2. *For the right-invariant control system in equation* (1), $t^*(U_F) = t^*(KU_F)$.

Proof. Suppose $U_F \in K$. Thus $KU_F = Ke$. We have $t^*(U_F) \geq t^*(KU_F)$ since KU_F contains U_F. Also, since $K = \exp\{H_1, ..., H_m\}_{LA}$, given any $T \geq 0$, there exists a $\hat{v} \in \mathcal{C}(T)$ such that the solution $U(t)$ to equation

$$\dot{U} = [\sum_{i=1}^{m} \hat{v}(t)_i H_i]U, \qquad U(0) = e$$

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

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

Set $\tau = \alpha t$. Thus

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

By Lemma 1, as $\alpha \to \infty$, $U(\tau) \mid_{\tau=T} = U_F$ or $\lim_{\alpha\to\infty} U(t) \mid_{t=\frac{\tau}{\alpha}} = U_F$. Hence $U_F \in \overline{\mathbf{R}(e,T)}$, for all $T > 0$. It follows that $t^*(U_F) = 0$. Since $0 = t^*(U_F) \geq t^*(Ke) \geq 0$, we have that $t^*(U_F) = t^*(Ke)$. Now let $U_F \in G$. By right invariance $(Ke)U_F = (KU_F)$. Since our choice of $U_F \in G$ was arbitrary, $t^*(U_F) = t^*(KU_F)$ for any $U_F \in G$. □

Finding $t^*(U_F)$ is now reduced to finding the minimum time to steer the system (1) from Ke to KU_F. Next we will use Lemma 2 to show the relationship between computing $t^*(U_F)$ and computing the minimum length paths for a related problem to follow.

Let $P \in G$. We define the right-invariant *adjoint* control system associated with (1) to be

$$\dot{P} = H(t)P, \qquad P : [0,T] \to G \tag{2}$$

where $H(t) \in Ad_K(H_d) = \{k^{-1}H_d k \mid k \in K\}$, *i.e* H is a function with values in $Ad_K(H_d)$.

For the control system (2), we say that $KU_F \in B(U_0, t')$ if there exists a control $H[0, t']$ which steers $P(0) = U_0$ to $P(t') \in KU_F$ in t' units of time. Let

$$\mathbf{B}(U_0, T) = \bigcup_{0\leq t\leq T} B(U_0, t).$$

From Lemma 1 we have that $\mathbf{B}(U_0, T)$ is closed. The following defines the minimum time required to steer the system (2) from a fixed U_0 to the coset KU_F.

Definition. Let U_0, U_F be fixed. We define the *minimum coset time* (from U_0 to KU_F) to be

$$L^*(KU_F) = \inf\{t \geq 0 : KU_F \in \mathbf{B}(U_0, t)\}.$$

Theorem 2. ***(Equivalence Theorem)*** *The infimizing time $t^*(U_F)$ of steering the system*

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

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

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

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

Proof. Let $Q : [0, T] \to K$ be a solution to the control system

$$\dot{Q} = [\sum_{i=1}^{m} v_i H_i]Q, \qquad Q(0) = e. \tag{3}$$

Let $P \in G$ such that P evolves according to the following equation

$$\dot{P} = (Q^{-1} H_d Q)P, \qquad P(0) = e. \tag{4}$$

Consequently, the product $Q(t)P(t)$ satisfies

$$\frac{dQP}{dt} = [H_d + \sum_{i=1}^{m} v_i H_i]QP, \qquad Q(0)P(0) = e,$$

which is the same evolution as (1).

Note. The motivation for the above is to write $U = QP$, where Q is motion in K. In this form we may represent U in terms of Q and P. By substitution in (1) we have

$$\dot{U} = (QP)^{\dot{}} = Q\dot{P} + \dot{Q}P = H_d QP + (\sum_{i=1}^{m} v_i H_i)QP$$

which is implied by the following system

$$\dot{Q}P = (\sum_{i=1}^{m} v_i H_i)QP \Leftrightarrow \dot{Q} = (\sum_{i=1}^{m} v_i H_i)Q$$
$$Q\dot{P} = H_d QP \Leftrightarrow \dot{P} = Q^{-1} H_d QP$$

Since $U(0) = Q(0)P(0) = e$, by the uniqueness theorem for differential equations, $U(t) = Q(t)P(t)$. Thus, given a solution $\hat{U}(t)$ of equation (1) with initial condition $\hat{U}(0) = e$ there exist unique curves $\hat{P}(t)$ and $\hat{Q}(t)$ defined through equations (3) and (4) satisfying $\hat{U}(t) = \hat{Q}(t)\hat{P}(t)$.

If $U_F \in \overline{\mathbf{R}(e,T)}$, then there exists a sequence of controls $v^r[0,T]$ such that the corresponding solutions $U^r(t)$ of (1) satisfy $U^r(T) \to U_F$. Thus, the solutions $P^r(t)$ of (4) satisfy $\lim_{r\to\infty} P^r(T) \in KU_F$. Since $\mathbf{B}(e,T)$ is closed, it follows that $KU_F \in \mathbf{B}(e,T)$. Therefore, $t^*(U_F) \geq L^*(KU_F)$.

Now if $KU_F \in \mathbf{B}(e,T)$, then there exists a control $\bar{H}[0,T]$ such that the corresponding solution $\bar{P}(t)$ of (4) satisfies $\bar{P}(T) \in KU_F$. Since $\bar{H}(t) \in Ad_K(H_d)$ we can write $\bar{H}(t)$ as $\bar{Q}(t)^{-1}H_d\bar{Q}(t)$ for $\bar{Q}(t) \in K$. It may not be the case that $\bar{Q}(t)$ is a solution to (3). However, there exists a family of controls $v^s(t)$ such that for the corresponding solution $Q^s(t)$ of

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

we have

$$\lim_{s\to\infty} \int_0^T \|\bar{Q}(t) - Q^s(t)\|dt = 0.$$

Thus,

$$\lim_{s\to\infty} \int_0^T \|\bar{H}(t) - (Q^s(t))^{-1}H_dQ^s(t)\|dt = 0.$$

By using Lemma 1 we claim that the family of differential equations

$$\dot{P}^s = [(Q^s)^{-1}(t)H_dQ^s(t)]P^s, \qquad P^s(0) = e$$

satisfies $\lim_{s\to\infty} P^s(T) = \bar{P}(T) \in KU_F$.

Hence for all s we have

$$\left.\begin{aligned} Q^s(T) &\to \bar{Q}(T) \\ P^s(T) &\to \bar{P}(T) \end{aligned}\right\} \quad \Rightarrow U^s(T) \to \bar{Q}(T)\bar{P}(T)$$

Since $\bar{Q}(T) \in K$ and $\bar{P}(T) \in KU_F$ we have that $\overline{U} = \lim_{s\to\infty} U^s(T) \in KU_F$. Thus $\overline{U} \in \overline{\mathbf{R}(e,T)}$ which implies $t^*(KU_F) \leq T$. Since the choice of T was arbitrary, $t^*(KU_F) \leq L^*(KU_F)$ □

We will now compute $t^*(U_F)$ based on the properties of $Ad_K(H_d)$. Consider the following classifications:

(1) *Riemannian Symmetric Case* which may be split into two cases;
   - The *Pulse-drift-pulse sequence*: Characteristic of a single spin system. Here the rank of G/K is 1.
   - The *Chained Pulse-drift-pulse sequence*: Characteristic of a two spin system. Here the rank of G/K is greater than 1 and we have a finite chain of pulse-drift-pulse sequences.

(2) *Chatter Sequence*: Characteristic of a system with spin greater than 2. Here G/K is no longer Riemannian Symmetric *i.e* $[\mathfrak{m},\mathfrak{m}] \not\subseteq \mathfrak{k}$.

For this paper, we will only concern ourselves with the Riemannian symmetric case.

Pulse-drift-pulse sequence. The control system (1) evolves on G and induces a control system on the coset space G/K by the projection map π. The adjoint control system (2) is a representation of this control system. Moreover, relative to the bi-invariant metric, solutions to the adjoint control system have constant speed $\|\dot{P}\| = \|H(t)P\| = \|H(t)\| = \|H_d\|$. We observe that time is therefore proportional to the length of a solution curve, with proportionality factor $\|H_d\|$. If necessary, we may multiply the bi-invariant metric $\langle\, ,\rangle_G$ by a positive real number so that $\|H_d\| = 1$. We will always assume that this normalization has been done, so that along solutions to the adjoint equation time equals length.

From Theorem 1 and the rank-1 assumption we have that $Ad_K(H_d)$ is the unit sphere in $\mathfrak{m}$. We conclude that the projected control system on G/K fixes the speed of curves to unit speed curves, but allows the velocity vectors to point in any direction. Thus every unit speed curve emanating from $o = \pi(e)$ is the projection of some solution $P(t)$ to the adjoint equation. Finally, since the projection map $\pi : G \to G/K$ is a Riemannian isometry, where G/K is given the normal metric, we see that $L^*(KU_F)$ equals the Riemannian distance, in G/K, from o to $\pi(U_F)$. The next theorem shows how we may use this observation to derive a method for computing $t^*(U_F)$.

Theorem 3. *Let G be a compact Lie group with a bi-invariant metric $<,>$, and K be a closed subgroup of G. Let $\mathfrak{g}$ and $\mathfrak{k}$ be their Lie algebras with decomposition $\mathfrak{g} = \mathfrak{m} \oplus \mathfrak{k}$ where $\mathfrak{m} = \mathfrak{k}^{\perp}$. Consider the right-invariant control system*

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

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

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

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

from $P(0) = e$ to KU_F. From above recall that $L^*(KU_F)$ is the Riemannian distance between $o = \pi(e)$ and $\pi(U_F)$ under the metric $<,>_n$. By a standard theorem of Riemannian symmetric spaces the geodesics under the normal metric that emanate from o take the form $\pi(\exp(\tau H))$ for $H \in \mathfrak{m}$ (Kobayashi and Nomizu [6]). If $U_F = Q_1 \exp(tH_d) Q_2 = (Q_1 Q_2) Q_2^{-1} \exp(tH_d) Q_2 = (Q_1 Q_2) \exp(t Q_2^{-1} H_d Q_2)$ for $Q_1, Q_2 \in K$, then $\pi(U_F) = \pi(\exp(t Q_2^{-1} H_d Q_2))$. Thus this curve is a geodesic that connects o to $\pi(U_F)$. It takes the form $\pi(\exp(t Q_2^{-1} H_d Q_2))$ and has length $L = t$. Therefore minimizing t over all ways of writing $U_F = Q_1 \exp(tH_d) Q_2$ yields the Riemannian distance between o and $\pi(U_F)$. □

Remark. Roughly speaking, the time optimal trajectory for system (1), which steers the system from $U(0) = e$ to $U_F = Q_1 \exp(\alpha H_d) Q_2$, takes the form

$$e \to Q_2 \to exp(\alpha H_d) Q_2 \to Q_1 \exp(\alpha H_d) Q_2$$

where the first and last steps of the chain take no time and the second step (drift) takes the required time t.

We will illustrate these ideas through the following examples.

Corollary 2. *Let $U \in G = SU(2)$ and let*

$$I_x = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, I_y = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, I_z = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

represent the Pauli spin matrices. Consider the unitary evolution of the single spin system

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

where the control $v \in \mathbb{R}$. Thus $H_d = -iI_z$ and $H_1 = -iI_x$. By direct computation we also have $\|-iI_z\| = 1$. Let $K = \{\exp(-iI_x t) | t \in \mathbb{R}\}$ be the one parameter subgroup generated by I_x. Given $U_F \in SU(2)$, there exists a unique $\alpha \in [0, 2\pi]$ such that $U_F = U_1 \exp[-i\alpha I_z] U_2$, where $U_1, U_2 \in K$. The infimizing time $t^*(U_F) = |\alpha \mod [-2\pi, 2\pi]|$.

Proof. The Lie algebra $\mathfrak{g} = \mathfrak{su}(2)$ has the decomposition $\mathfrak{m} = span\{iI_y, iI_z\}, \mathfrak{k} = span\{iI_x\}$ with $Ad_K(-\alpha i I_z) = \mathfrak{m}$. Suppose $U_F = U_1 exp(-\alpha i I_z) U_2$, for $U_1, U_2 \in K$. Since $\exp(-\alpha i I_z)$ has a period of 4π, there exists a unique $\alpha \in [-2\pi, 2\pi]$ for which $U_F = U_1 \exp(-\alpha i I_z) U_2$ holds. Thus $U_F = U_1 \exp(-\beta i I_z) U_2$, where $\beta = |\alpha \mod [-2\pi, 2\pi]|$. By Theorem 3 we are done. □

Corollary 3. *Let $\Theta \in G = SO(3)$ and let*

$$\Omega_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \Omega_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \Omega_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

represent the generators of rotation around the x-axis, y-axis and z-axis. Consider the control system

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

where the control $u \in \mathbb{R}$. Let $K = \{\exp(\Omega_x t) | t \in \mathbb{R}\}$ represent the one parameter subgroup generated by Ω_x and let $\Theta_F = \Theta_1 \exp(\alpha \Omega_z) \Theta_2$, where $\Theta_1, \Theta_2 \in K$. Then the infimizing time $t^*(\Theta_F) = |\alpha \mod [-\pi, \pi]|$.

Proof. The Lie algebra $\mathfrak{g} = \mathfrak{so}(3)$ has the decomposition $\mathfrak{m} = span\{\Omega_y, \Omega_z\}$, $\mathfrak{k} = span\{\Omega_x\}$ with $Ad_K(\Omega_z) = \mathfrak{m}$. Suppose $\Theta_F = \Theta_1 \exp(\alpha \Omega_z) \Theta_2$, where $\Theta_1, \Theta_2 \in K$. Notice that $\exp(\alpha \Omega_z)$ is periodic with a period of 2π. As in the proof of Corollary 2, we have $t^*(\Theta_F) = |\alpha \mod [-\pi, \pi]|$. □

Corollary 4. *Let $\Theta \in G = SO(n)$ and let the control system*

$$\dot{\Theta} = [\Omega_d + \sum_{i=1}^{m} v_i \Omega_i]\Theta, \qquad \Theta(0) = I$$

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

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

Proof. Notice that $Ad_K(\Omega_d) = \mathfrak{m}$. The rest of the proof is similar to that of Corollary 2. □

Chained Pulse-drift-pulse sequence. Now we consider the case where the Riemannian symmetric space G/K has rank > 1. To better understand this particular case we will need some background information.

Definition. Given the decomposition $\mathfrak{g} = \mathfrak{m} + \mathfrak{k}$, let $\mathfrak{h} \subset \mathfrak{m}$ represent the maximal abelian subalgebra containing H_d. Let $\Delta_{H_d} = \mathfrak{h} \cap Ad_K(H_d)$ denote the maximal commuting set contained in the adjoint orbit of H_d. The set Δ_{H_d} is called the *Weyl orbit* of H_d. Let $\mathfrak{c}(H_d) = \{\sum_{i=1}^n \beta_i H_i : \beta_i \geq 0, \sum_{i=1}^n \beta_i = 1, H_i \in \Delta_{H_d}\}$ denote the convex hull of the Weyl orbit of H_d, with vertices given by the elements of the Weyl orbit of H_d.

We will compute the infimizing time for the system (1) in the following theorems, which generalize the rank one case.

Remark. Recall from corollary 1 that if $A = \exp(\mathfrak{h})$, where $\mathfrak{h}$ is a maximal abelian subalgebra contained in $\mathfrak{m}$, then $G = KAK$. Thus, given any $U_F \in G$, we can express $U_F = Q_1 \exp(Z) Q_2 = Q_1 Q_2 \exp(Ad_{Q_2}(Z))$, where $Q_1, Q_2 \in K$ and $Z \in \mathfrak{h}$. Suppose $Z = \sum_{i=1}^n \beta_i H_i$, $\beta_i \geq 0$, $\sum_{i=1}^n \beta_i = 1$, $H_i \in \Delta_{H_d}$. By choosing $H(t)$ to be $Ad_{Q_2}(H_i)$ for β_i units of time we can steer the adjoint control system $\dot{P} = H(t)P$ from the identity to the coset $KU_F = K\exp(Ad_{Q_2}(Z))$. The following theorem states that the fastest way to get to the coset KU_F is to flow on the *maximal torus* (see appendix), $Ad_{Q_2}(A)$, $Q_2 \in K$ containing the coset KU_F.

Theorem 4. ***(Kostant's Convexity Theorem)*** *Given the direct sum decomposition $\mathfrak{g} = \mathfrak{m} + \mathfrak{k}$, let $\mathfrak{h} \subset \mathfrak{m}$ represent a maximal abelian subalgebra containing $H_d \in \mathfrak{m}$. Let $\Gamma : \mathfrak{m} \to \mathfrak{h}$ be the orthogonal projection of $\mathfrak{m}$ onto $\mathfrak{h}$. Then the image of $Ad_K(H_d)$ under Γ is $\mathfrak{c}(H_d)$, where $\mathfrak{c}(H_d)$ is the convex hull of the Weyl orbit of H_d.*

Theorem 5. ***(Time Optimal Tori Theorem)*** *Let G be a compact matrix Lie group and K be a closed subgroup with Lie algebras $\mathfrak{g}$ and $\mathfrak{k}$, respectively, such that*

G/K is a Riemannian symmetric space. Let the direct sum decomposition $\mathfrak{g} = \mathfrak{m} + \mathfrak{k}$, where $\mathfrak{m} = \mathfrak{k}^{\perp}$, be given. Consider the right invariant control system

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

where $v_i \in \mathbb{R}$, $H_d \in \mathfrak{m}$, $\{H_1, \ldots, H_m\}_{LA} = \mathfrak{k}$. Then any $U_F = Q_1 \exp(\alpha Y) Q_2$, where $\alpha > 0$, $Q_1, Q_2 \in K$, and $Y \in \mathfrak{c}(H_d)$, belongs to the closure of the reachable set. The infimizing time $t^*(U_F)$ is the smallest value of $\alpha > 0$ such that we can solve

$$U_F = Q_1 \exp(\alpha Y) Q_2$$

where $Q_1, Q_2 \in K$ and Y belongs to the convex hull $\mathfrak{c}(H_d)$.

Proof. (sketch) We will compute $t^*(U_F)$ by finding the minimum time T to steer the system $\dot{P} = HP$ from $P(0) = e$ to the coset KU_F. Let $A = \exp(\mathfrak{h})$ be the maximal torus in G generated by $\mathfrak{h}$. By corollary 1, $G/K = Inn_K(A)$. Hence, we only need to compute T for $U_F \in A$. Let $H = H_i \in \Delta_{H_d}$ for α_i units of time. So we have solutions $P_i(t) = \exp(tH_i)$, $0 \le t \le \alpha_i$, of the adjoint control system. Since A is abelian and $U_F \in A$, there exist solutions $P_i(t)$ as we have just described such that

$$U_F = \exp(\sum_{i=1}^{p} \alpha_i H_i) = \Pi \exp(\alpha_i H_i) = \Pi P_i(\alpha_i).$$

It is evident that the time to reach U_F via the above solution to the adjoint equation is given by $\sum_{i=1}^{p} \alpha_i$. Let T be the infimum of these times taken over all trajectories of this type that join the identity to U_F. Note that $T \ge t^*(U_F)$.

Next we show that $T \le t^*(U_F)$ (we can reach the coset KU_F no sooner than time T). Let $\bar{P}(t)$ be the shortest trajectory of $\dot{P} = HP$ which steers $P(0) = e$ to KU_F. Let $\pi_A : G/K \to A$ be the projection such that $\pi_A(k^{-1}A_1 k) = A_1$, for $A_1 \in A$. This map is unique modulo a Weyl group action. So we will define an explicit projection which makes the mapping unique. Let $a(t) = (\pi_A \circ \bar{P})(t) \in A$ be a continuous path onto A, obtained from the projection of $\bar{P}(t)$. Thus, by the chain rule $\dot{a}(t) = d\pi_A \dot{\bar{P}}(t)$, where $\dot{\bar{P}} = H\bar{P}$ for $H \in Ad_K(H_d)$. Let $a(t_F) = \pi_A(KU_F) = U_F$ where $t_F = t^*(U_F)$. We can write $\dot{a} = \Omega\, a = d\pi_A(H\bar{P})$ where $\Omega = \Gamma(Ad_{\tilde{k}}(H_d))$ for some $\tilde{k} \in K$. By Kostant's convexity theorem $\Omega \in \mathfrak{c}(H_d)$. Hence, $\Omega = \sum_{i=1}^{p} \beta_i H_i$, $\sum_{i=1}^{p} \beta_i = 1$, $\beta_i \ge 0$, $H_i \in \Delta_{H_d}$. Thus, $\dot{a} = \Omega a = (\sum_{i=1}^{p} \beta_i H_i)a$, $a(0) = e$. So we have

$$a(t) = \exp(t \sum_{i=1}^{p} \beta_i H_i) = \exp(\sum_{i=1}^{p} t\beta_i H_i).$$

$$\Rightarrow a(t_F) = \exp(t_F \sum_{i=1}^{p} \beta_i H_i) = \exp(\sum_{i=1}^{p} t_F \beta_i H_i).$$

Note that for each H_i we have $\|H_i\| = \|H\| = 1$. Therefore, since $\sum_{i=1}^{p} \beta_i = 1$, we have

$$t^*(U_F) = t_F = \sum_{i=1}^{p} t_F \beta_i \geq \sum_{i=1}^{p} \alpha_i = T.$$

□

4. Conclusion

In this paper, we presented a mathematical formulation of the problem of finding the shortest pulse sequences in NMR. We showed how the problem of computing minimum time to produce a unitary propagator can be reduced to finding the shortest length paths on certain coset spaces. A remarkable feature of time optimal control, and key to our results, is that the control is zero most of the time with pulses in between. We are now left with the open problem of computing minimum time to produce a unitary propagator in the Chatter Sequence case. In this case the coset space is no longer Riemannian Symmetric.

Note. As of May 8, 2006, the largest quantum information processor to be constructed was based on a 12-qubit system (Science Daily), which is quite small. The following is a list of candidates for quantum computing that are currently being explored.

(1) Superconductor-based quantum computers (including SQUID-based quantum computers)
(2) Trapped ion quantum computer
(3) Optical lattices
(4) Topological quantum computer
(5) Quantum dot on surface (e.g. the Loss-DiVincenzo quantum computer)
(6) Nuclear magnetic resonance on molecules in solution (liquid NMR)
(7) Solid state NMR Kane quantum computers
(8) Electrons on helium quantum computers
(9) Cavity quantum electrodynamics (CQED)
(10) Molecular magnet
(11) Fullerene-based ESR quantum computer
(12) Optic-based quantum computers (Quantum optics)
(13) Diamond-based quantum computer
(14) BoseEinstein condensate-based quantum computer
(15) Transistor-based quantum computer - string quantum computers with entrainment of positive holes using a electrostatic trap
(16) Spin-based quantum computer
(17) Adiabatic quantum computation

The large number of candidates clearly shows that the topic is still in its infancy.

5. Appendix

The background needed to properly understand the control theory problem took some interesting turns, one of which was in the classification of compact Lie groups; a subject that deserves to be explored in greater detail. Thus, an in-depth study of the classification of compact Lie groups has been added as an appendix to this paper.

Representation.

Definition. A (finite dimensional) *representation* of a Lie group G is a continuous homomorphism $\phi : G \to Aut(V)$, where V is a (finite dimensional) vector space. The dimension of the representation is the dimension of the vector space V. Denote the representation of G in V by (G, V).

Definition. Let (G, V) be a representation. A subspace U of V is called invariant if $gU \subset U$ for all $g \in G$.

Definition. A representation is called *irreducible* if the only invariant subspaces are $\{0\}$ and V.

Definition. Two representations $\phi_1 : G \to Aut(V_1)$ and $\phi_2 : G \to Aut(V_2)$ are said to be *equivalent* (denoted by $\phi_1 \cong \phi_2$) if V_1 and V_2 are isomorphic, *i.e.*, there exists a linear isomorphism $A : V_1 \to V_2$ such that $A(\phi_1(g)(v)) = \phi_2(g)(A(v))$, for all $g \in G$ and $v \in V_1$.

Hence the following diagram commutes:

$$\begin{array}{ccc} V_1 & \xrightarrow{\phi_1(g)} & V_1 \\ {\scriptstyle A}\downarrow & & \downarrow{\scriptstyle A} \\ V_2 & \xrightarrow{\phi_2(g)} & V_2 \end{array}$$

An *automorphism* of a Lie group G is a map $\phi : G \to G$ that is a diffeomorphism and a group homomorphism. Let G be a Lie group and let $x \in G$. Then the map $I_x : G \to G$ sending each g to xgx^{-1} is a automorphism. Note that $I_x = R_{x^{-1}} \circ L_x$ is a diffeomorphism called an *inner automorphism* of G.

Definition. The adjoint representation of G is the homomorphism $Ad : G \to Aut(\mathfrak{g})$ given by $Ad(g) = (dI_g)_e$.

Definition. The adjoint representation of $\mathfrak{g}$ is the homomorphism $ad : \mathfrak{g} \to End(\mathfrak{g})$ given by $ad(X) = (dAd)_e(X)$.

Theorem 6. *The adjoint representation of* $\mathfrak{g}$ *satisfies* $ad(X)Y = [X,Y]$ *for all* $X, Y \in \mathfrak{g}$.

Proposition 2. *If* G *is a matrix group, then* $Ad(g)X = gXg^{-1}$ *for all* $g \in G, X \in \mathfrak{g}$.

Remark. If V is a vector space over $\mathbb{R}$, then we can define the vector space

$$V^{\mathbb{C}} = V \otimes_{\mathbb{R}} \mathbb{C} = \{X + iY : X, Y \in V, i = \sqrt{-1}\}$$

(or simply $V \otimes C$), whose dimension over $\mathbb{C}$ equals the dimension of V over $\mathbb{R}$. If $\mathfrak{g}$ is a Lie bracket over $\mathbb{R}$, then the complexification of $\mathfrak{g}$ is the Lie algebra $\mathfrak{g} \otimes \mathbb{C}$ (sometimes written as $\mathfrak{g} + i\mathfrak{g}$), with the bracket operation given by

$$[U + iV, X + iY] = [U, X] - [V, Y] + i([V, X] + [U, Y]).$$

Example. Let $G = SU(2)$ with Lie algebra $\mathfrak{g} = \mathfrak{su}(\mathfrak{2})$ consisting of matrices of the form $\begin{pmatrix} i\alpha & z \\ -\bar{z} & -i\alpha \end{pmatrix}$.

Let $A = \begin{pmatrix} x+iy & u+iv \\ -u+iv & x-iy \end{pmatrix} \in SU(2)$. We need to compute $Ad : SU(2) \to Aut(\mathfrak{su}(2))$, the linear transformation given by $Ad(A)B = ABA^{-1}$. Let

$$E_1 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, E_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, E_3 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

be a basis for $\mathfrak{su}(2)$. Calculating $Ad(A)E_1$ gives

$$\begin{aligned} Ad\begin{pmatrix} x+iy & u+iv \\ -u+iv & x-iy \end{pmatrix}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} &= \begin{pmatrix} x+iy & u+iv \\ -u+iv & x-iy \end{pmatrix}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}\begin{pmatrix} x-iy & -u-iv \\ u-iv & x+iy \end{pmatrix} \\ &= \begin{pmatrix} ix^2+iy^2-iu^2-iv^2 & -2ixu+2uy+2xv+2iyv \\ 2iux+2xv-2uy+2ivy & iu^2+iv^2-ix^2-iy^2 \end{pmatrix} \end{aligned}$$

By computations on the other basis elements we get

$$Ad\begin{pmatrix} x+iy & u+iv \\ -u+iv & x-iy \end{pmatrix} = \begin{pmatrix} x^2+y^2-u^2-v^2 & -2xv+2uy & 2xu+2yv \\ 2uy+2xv & x^2-y^2+u^2-v^2 & -2xy+2uv \\ -2xu+2yv & 2xy+2uv & x^2-y^2-u^2+v^2 \end{pmatrix}$$

Notice that this is a 3×3 matrix, which agrees with the dimension of the representation which is $\dim(\mathfrak{su}(2)) = 3$

Standard representation. We will define the standard representations of $GL(n,\mathbb{R})$, $O(n)$ and $SO(n)$ on $M_{n\times 1}\mathbb{R} \cong \mathbb{R}^n$ with the operation being matrix multiplication on $\mathbb{R}^n$, *i.e.*, $\phi(g)v = gv$. Similarly we can define the standard representations of $GL(n,\mathbb{C})$, $SU(n)$ and $U(n)$ on $M_{n\times 1}\mathbb{C} \cong \mathbb{C}^n$ with the operation being matrix multiplication on $\mathbb{C}^n$. Let $\tilde{\lambda}_n$ denote the standard representation of $GL(n,\mathbb{R})$ and let λ_n denote the standard representation of $SO(n)$ (or $O(n)$). Let $\tilde{\mu}_n$ denote the standard representation of $GL(n,\mathbb{C})$ and let μ_n denote the standard representation of $SU(n)$(or $U(n)$). Let v_n denote the standard representation of $Sp(n)$. A representation is called trivial, denoted by 1, if each group element acts as the identity transformation. Then the adjoint representations of these groups are equivalent to the following representations:

$$
\begin{aligned}
&Ad^{GL(n,\mathbb{R})} = \tilde{\lambda}_n \otimes_{\mathbb{R}} \tilde{\lambda}_n^*,\\
&Ad^{GL(n,\mathbb{C})} = \tilde{\mu}_n \otimes_{\mathbb{C}} \tilde{\mu}_n^*,\\
&Ad^{SO(n)} = \wedge^2 \lambda_n,\\
&Ad^{U(n)} \otimes \mathbb{C} = \mu_n \otimes_{\mathbb{C}} \mu_n^* = \mu_n \otimes_{\mathbb{C}} \bar{\mu}_n,\\
&Ad^{SU(n)} \otimes \mathbb{C} = \mu_n \otimes_{\mathbb{C}} \bar{\mu}_n - 1,\\
&Ad^{Sp(n)} \otimes \mathbb{C} = S^2 v_n
\end{aligned}
$$

For a better understanding of the above equivalences, we will show that $Ad^{GL(n,\mathbb{R})} = \tilde{\lambda}_n \otimes_{\mathbb{R}} \tilde{\lambda}_n^*$. Note that $GL(n,\mathbb{R}) \cong Aut(\mathbb{R}^n)$ and for $A \in GL(n,\mathbb{R})$, $x \in \mathbb{R}^n$ we have $\lambda_n(A)x = Ax$. Since $\mathbb{R}^n$ and $(\mathbb{R}^n)^*$ are finite vector spaces, $\mathbb{R}^n \otimes (\mathbb{R}^n)^* \cong Hom(\mathbb{R}^n,\mathbb{R}^n) = \mathfrak{gl}(n)$ by the mapping $\mathbb{R}^n \otimes (\mathbb{R}^n)^* \to Hom(\mathbb{R}^n,\mathbb{R}^n)$ defined by $(w \otimes \theta)v = \theta(v)w$. Thus we need to find an isomorphism ϕ which makes the following diagram commute

$$
\begin{array}{ccc}
\mathbb{R}^n \otimes (\mathbb{R}^n)^* & \xrightarrow{\tilde{\lambda}_n \otimes \tilde{\lambda}_n^*(g)} & \mathbb{R}^n \otimes (\mathbb{R}^n)^* \\
\downarrow \phi & & \downarrow \phi \\
\mathfrak{gl}(n) & \xrightarrow{Ad(g)} & \mathfrak{gl}(n)
\end{array}
$$

First consider the case when $n = 2$. Then we have the following maps:

$$
\begin{aligned}
\tilde{\lambda}_2 &: GL(2,\mathbb{R}) \to Aut(\mathbb{R}^2),\\
\tilde{\lambda}_2^* &: GL(2,\mathbb{R}) \to Aut(\mathbb{R}^2),\\
\tilde{\lambda}_2 \otimes \tilde{\lambda}_2^* &: GL(2,\mathbb{R}) \to Aut(\mathbb{R}^2 \otimes (\mathbb{R}^2)^*)
\end{aligned}
$$

where $v \in \mathbb{R}^2$ is a column vector and $u \in (\mathbb{R}^2)^*$ a row vector. Choose the standard basis

$e_1 = (1,0)^T, e_2 = (0,1)^T \in \mathbb{R}^2$ and $e^1 = (1,0), e^2 = (0,1) \in (\mathbb{R}^2)^*$. Thus

$$e_1 \otimes e^1 = (1,0)^T \otimes (1,0) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \in \mathfrak{gl}(2)$$

$$e_1 \otimes e^2 = (1,0)^T \otimes (0,1) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathfrak{gl}(2)$$

$$e_2 \otimes e^1 = (0,1)^T \otimes (1,0) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \in \mathfrak{gl}(2)$$

$$e_2 \otimes e^2 = (0,1)^T \otimes (0,1) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \in \mathfrak{gl}(2)$$

Hence for any n we have

$$\begin{aligned} \tilde{\lambda}_n \otimes \tilde{\lambda}_n^*(g) e_i \otimes e^j &= \tilde{\lambda}_n(g) e_i \otimes (\tilde{\lambda}_n^*(g) e^j \\ &= (ge_i) \otimes (e^j g^{-1}) \in \mathbb{R}^n \otimes (\mathbb{R}^n)^*, \end{aligned}$$

for each column vector e_i and row vector e^j. Define $\phi(a \otimes b) = ab$. Then

$$\begin{aligned} \phi(\tilde{\lambda}_n \otimes \tilde{\lambda}_n^*(g)) a \otimes b &= \phi(\tilde{\lambda}_n(g)a) \otimes (\tilde{\lambda}_n^*(g)b) \\ &= \phi((ga) \otimes (bg^{-1})) \\ &= (ga)(bg^{-1}) \\ &= g(ab)g^{-1} \\ &= Ad(g)ab \\ &= Ad(g)\phi(a \otimes b). \end{aligned}$$

And we are done.

Killing form.

Definition. The *Killing form* of a Lie algebra $\mathfrak{g}$ is the function $B : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ given by $B(X,Y) \oplus tr(adX \circ adY)$.

Proposition 3. *The Killing form has the following properties:*

(a) *It is a symmetric bilinear form on $\mathfrak{g}$.*
(b) *If $\mathfrak{g}$ is the Lie algebra of G, then B is the Ad-invariant, that is, $B(X,Y) = B(Ad(g)X, Ad(g)Y)$ for all $g \in G$ and $X, Y \in \mathfrak{g}$. in other words, each $Ad(g), g \in G$ is B-orthogonal.*
(c) *Each $ad(Z)$ is skew-symmetric with respect to B, that is, $B(ad(Z)X,Y) = -B(X, ad(Z)Y)$ or $B([X,Z],Y) = B(X,[Z,Y])$.*

Note. The Killing form of a Lie group G is understood to be the Killing form of its Lie algebra $\mathfrak{g}$.

Definition. (**semisimple**) A Lie group G is called *semisimple* if its Killing form is non-degenerate.

We can think of a semisimple Lie algebra $\mathfrak{g}$ as one that has no proper subspace $\mathfrak{h}$ with $[X,Y]=0$ if $X\in\mathfrak{h}$ and $Y\in\mathfrak{g}$.

Theorem 7. *If G is a compact semisimple Lie group, then its Killing form is negative definite, i.e., $B(X,X)<0$ for all $X\neq 0$ in G.*

Example. To compute the Killing form of $SU(2)$ it is sufficient to compute the Killing form on its Lie algebra $\mathfrak{su}(2)$. Let $X=\begin{pmatrix} i\theta & 0\\ 0 & -i\theta\end{pmatrix}$, $Y=\begin{pmatrix} i\phi & 0\\ 0 & -i\phi\end{pmatrix}\in$ $\mathfrak{su}(2)$ and consider the basis $\begin{pmatrix} i & 0\\ 0 & -i\end{pmatrix},\begin{pmatrix} 0 & 1\\ -1 & 0\end{pmatrix},\begin{pmatrix} 0 & i\\ i & 0\end{pmatrix}$ of $\mathfrak{su}(2)$. Then $adX=\begin{pmatrix} 0&0&0\\ 0&0&-2\theta\\ 0&2\theta&0\end{pmatrix}$ and $adY=\begin{pmatrix} 0&0&0\\ 0&0&-2\phi\\ 0&2\phi&0\end{pmatrix}$. So

$$B(X,Y)=tr(ad(X)ad(Y))=tr\begin{pmatrix} 0&0&0\\ 0&-4\theta\phi&0\\ 0&0&-4\theta\phi\end{pmatrix}=-8\theta\phi=4trXY.$$

Example. As with $SU(2)$, to compute the Killing form of $U(2)$ we compute the Killing form of $\mathfrak{u}(2)$ with the following basis:

$$\begin{pmatrix} i&0\\ 0&0\end{pmatrix},\begin{pmatrix} 0&0\\ 0&i\end{pmatrix},\begin{pmatrix} 0&1\\ -1&0\end{pmatrix},\begin{pmatrix} 0&i\\ i&0\end{pmatrix}.$$

Let $X=\begin{pmatrix} i\theta_1 & 0\\ 0 & -i\theta_2\end{pmatrix}$ and $Y=\begin{pmatrix} i\phi_1 & 0\\ 0 & -i\phi_2\end{pmatrix}\in\mathfrak{u}(2)$. Then

$$adX=\begin{pmatrix} 0&0&0&0\\ 0&0&0&0\\ 0&0&0&\theta_1-\theta_2\\ 0&0&-\theta_1-\theta_2&0\end{pmatrix}, adY=\begin{pmatrix} 0&0&0&0\\ 0&0&0&0\\ 0&0&0&\phi_1-\phi_2\\ 0&0&-\phi_1-\phi_2&0\end{pmatrix}.$$

Thus $B(X,Y)=tr(adX,adY)=tr\begin{pmatrix} 0&0&0&0\\ 0&0&0&0\\ 0&0&-(\theta_1-\theta_2)(\phi_1+\phi_2)&0\\ 0&0&0&-(\theta_1+\theta_2)(\phi_1-\phi_2)\end{pmatrix}=$
$4(\theta_1\phi_1+\theta_2\phi_2)-2(\theta_1+\theta_2)(\phi_1+\phi_2)=4trXY-2trXtrY.$
Notice that if $\theta_1=\theta_2=\phi_1=\phi_2=1$, then $B(X,Y)=0$. So $U(2)$ is not semisimple.

Example. As in the previous examples, to compute the Killing form of $SO(3)$ we only need to compute the Killing form of $\mathfrak{so}(3)$. Consider the following basis of

$\mathfrak{so}(3)$:

$$E_1 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, E_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, E_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

Since $tr(adE_i adE_j) = 0$ for $i \neq j$, it suffices to compute the Killing form for the matrices

$$X = \begin{pmatrix} 0 & \theta & 0 \\ -\theta & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, Y = \begin{pmatrix} 0 & \phi & 0 \\ -\phi & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

in $\mathfrak{so}(3)$. So we have

$$adX = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \theta \\ 0 & -\theta & 0 \end{pmatrix}, adY = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \phi \\ 0 & -\phi & 0 \end{pmatrix}$$

Thus $B(X,Y) = tr\left(\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \theta \\ 0 & -\theta & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \phi \\ 0 & -\phi & 0 \end{pmatrix}\right) = tr\begin{pmatrix} 0 & 0 & 0 \\ 0 & -\theta\phi & 0 \\ 0 & 0 & -\theta\phi \end{pmatrix} = -2\theta\phi =$ $trXY$.

By generalizing the above examples we obtain the following formulas:

$$\begin{aligned} U(n) &: B(X,Y) = 2ntrXY - 2trXtrY, \\ SU(n) &: B(X,Y) = 2ntrXY, \\ SO(n) &: B(X,Y) = (n-2)trXY, \\ Sp(n) &: B(X,Y) = 2(n+1)trXY. \end{aligned}$$

Maximal tori.

Definition. (**torus**) A torus in a Lie group G is a Lie subgroup that is isomorphic to a product of $S^1 \times ... \times S^1$. A torus T is a maximal torus in G if for any torus S in G with $T \subset S \subset G$, then $T = S$.

Example. The set $T = \left\{ \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \right\}$ is a maximal torus in $SU(2)$.

Remark.

- Any torus is contained in a maximal torus.
- If G is compact then any maximal torus T is a maximal connected abelian subgroup of G.

- If T is a connected subgroup of a compact Lie group G whose Lie algebra is a maximal abelian subalgebra of $\mathfrak{g}$, then T is a maximal torus in G.

Theorem 8. *Let G be a compact and connected Lie group. Then:*

(a) *Any element in G is contained in some maximal torus.*
(b) *Any two maximal tori are conjugate. That is, if T_1, T_2 are maximal tori in G, then there exists an element $g \in G$ such that $gT_1g^{-1} = T_2$.*

From the above theorem we can conclude that all maximal tori have the same dimension, so is an invariant for a compact and connected Lie group.

Definition. (**rank**) The *rank* of a compact and connected Lie group is the dimension of a maximal torus.

Proposition 4. *Let G be a compact and connected Lie group with Lie algebra $\mathfrak{g}$. Then:*

(a) *The exponential map is onto.*
(b) *There is a one-to-one correspondence between maximal tori T in G and maximal abelian subspaces $\mathfrak{h}$ in $\mathfrak{g}$ given by $T \leftrightarrow \mathfrak{h} = \exp \mathfrak{t}$, where $\mathfrak{t}$ is the Lie algebra of T.*
(c) *If T is a maximal torus in G with Lie algebra $\mathfrak{t}$, then $G = \cup_{g \in G} gTg^{-1}$ and $\mathfrak{g} = \cup_{g \in G} Ad(g)\mathfrak{t}$.*
(d) *The center of G is equal to $\cap_{maximaltori} T$.*
(e) *If S is a subset of G, we define the centralizer of S to be the set $C(S) = \{g \in G : gx = xgforallx \in S\}$. Then, if T is a maximal torus in G, $C(T) = T$.*
(f) *Maximal tori are also maximal abelian subgroups.*
(g) *For any $X \in \mathfrak{g}$, the closure of $\{\exp(tX)\}$ is a compact abelian subgroup of G, and hence is a torus.*

Example. Consider the set $T = \{diag(e^{i\theta_1}, \ldots, e^{i\theta_n})\}$. Clearly T is a torus in $U(n)$. Let $A \in U(n)$ such that $AT = TA$. Let T_J be the subgroup of matrices t_j with 1 on the j^{th} entry and 0 elsewhere. Then, for all $t_j \in T_j$ we have $t_jAe_j = At_je_j = Ae_j$, where e_j is the column vector with 1 in the j^{th} place and 0 elsewhere. Since A is fixed by T, there exists some $\lambda_j \in \mathbb{C}$ such that $Ae_j = \lambda_j e_j$. Also, $A \in U(n)$, thus λ_j has modulus 1. Therefore $\lambda_j = e^{i\phi_j}$ for some ϕ. Since this is true for each j, $A = diag(e^{i\phi_1}, \ldots, e^{i\phi_n})$. Thus $A \in T$ and T is a maximal torus in $U(n)$ and the rank of $U(n)$ is n.

Classification of compact and connected Lie groups. All the groups in this section are assumed to be compact and connected.

Definition. A Lie group is called *simple* if it is non-abelian and it does not contain any proper normal Lie subgroups. Equivalently, a Lie group is simple if its Lie algebra is simple, *i.e.* it is non-abelian and has no proper ideals.

Note. The above definition gives a correspondence between normal subgroups and ideals.

Theorem 9.

(a) *Let G be a compact and connected Lie group. Then there exists a covering space of G that is isomorphic to the direct product of a torus and a compact, connected and simply connected Lie group.*

(b) *Every compact, connected and simply connected Lie group is isomorphic o the direct product of simple, compact, connected and simply connected Lie groups.*

(c) *The simple, compact, connected and simply connected Lie groups are the following:*

$$SU(n)(\geq 2), \widetilde{SO}(2n+1)(n \geq 3), Sp(n)(n \geq 2), \widetilde{SO}(2n)(n \geq 4), G_2, F_4, E_6, E_7, E_8.$$

Remark. The Lie algebras of the first four groups are denoted by A_{n-1}, B_n, C_n, D_n respectively. The following isomorphisms hold: $A_1 \cong B_1 \cong C_1, B_2 \cong C_2, A_3 \cong D_3, D_2 \cong A_1 \oplus A_1$. The last five Lie groups are called the exceptional Lie groups. Their indices indicate the rank of the group and their dimensions are 14, 52, 78, 133 and 248 respectively.

Complex semisimple Lie algebras. A simply connected Lie group is determined by its Lie algebra, so there is a one-to-one correspondence (up to isomorphism) between compact semisimple Lie algebras and compact Lie groups. By complexifying these Lie algebras we obtain a one-to-one correspondence between these complexified Lie algebras and the complex semisimple Lie algebras. The complex semisimple Lie algebras are classified by their root systems, and the root systems are classified by their bases. The bases are described by the Dynkin diagrams. Hence, in this section, we end up with a one-to-one correspondence between compact simply connected Lie groups and Dynkin diagrams.

Definition. Let $\mathfrak{g}$ be a complex Lie algebra.

(1) The *adjoint representation* of $\mathfrak{g}$ is the homomorphism $ad : \mathfrak{g} \to End(g)$ given by $ad(X)(Y) = [X, Y]$ for all $X, Y \in \mathfrak{g}$.
(2) The *Killing form* of $\mathfrak{g}$ is the symmetric bilinear form given by $B(X, Y) = tr(adX \circ adY), (X, Y \in \mathfrak{g})$.
(3) $\mathfrak{g}$ is called *semisimple* if its Killing form is non-degenerate.
(4) $\mathfrak{g}$ is called *simple* if it is non-abelian and its only ideals are $\{0\}$ and $\mathfrak{g}$.
(5) A *Cartan subalgebra* $\mathfrak{h}$ of $\mathfrak{g}$ is a maximal abelian subalgebra of $\mathfrak{g}$ such that, for all $H \in \mathfrak{h}$ the endomorphism $ad(h)$ is diagonalizable.

Proposition 5. *A Lie algebra is semisimple if and only if it is isomorphic to a product of simple Lie algebras.*

Proposition 6.

(a) *Any complex Lie algebra contains a Cartan subalgebra.*

(b) *Let G be the group of automorphisms of $\mathfrak{g}$ generated by the following elements:*

$$\exp(adX) = \sum_{n=0}^{\infty} \frac{1}{n!}(adX)^n, (X \in \mathfrak{g})$$

Then any two Cartan subalgebras are conjugate under G. This group is called the adjoint group.

Note. If G is a compact Lie group with Lie algebra $\mathfrak{g}$, and $\mathfrak{h}_1, \mathfrak{h}_2$ are two Cartan subalgebras of $\mathfrak{g}$, then there exists a $g \in G$ such that $Ad(g)\mathfrak{h}_1 = \mathfrak{h}_2$.

Definition. The *rank* of a Lie algebra is the dimension of a Cartan subalgebra.

Let $\mathfrak{g}$ be a complex semisimple Lie algebra and $\mathfrak{h}$ a fixed Cartan subalgebra of $\mathfrak{g}$. Let $\mathfrak{h}^*$ be the dual space of $\mathfrak{h}$. Then for all $\alpha \in \mathfrak{h}^*$, let $\mathfrak{g}^\alpha = \{X \in \mathfrak{g} : ad(H)X = \alpha(H)X \text{ for all } H \in \mathfrak{h}\}$ denote the corresponding eigenspace of $\mathfrak{g}$. If $\alpha \in \mathfrak{h}^*$ such that $\alpha \neq 0$ and $\mathfrak{g}^\alpha \neq \{0\}$, then we call α a *root* of $\mathfrak{g}$. We call $\mathfrak{g}^\alpha$ the *root space* corresponding to α. Denote the set of all the roots of $\mathfrak{g}$ by R, called the *root system* of $\mathfrak{g}$ (relative to $\mathfrak{h}$). In particular, let $\mathfrak{g}^0$ be the set of all elements of $\mathfrak{g}$ that commute with $\mathfrak{h}$. Since $\mathfrak{h}$ is maximal abelian, $\mathfrak{g}^0 = \mathfrak{h}$. Also. for all $H \in \mathfrak{h}$, the endomorphisms $ad(H)$ are diagonalizable and commute, so they are simultaneously diagonalizable. Thus, for a fixed $\mathfrak{h}$ we obtain the *root space decomposition* of $\mathfrak{g}$

$$\mathfrak{g} = \mathfrak{h} \oplus \sum_{\alpha \in R} \mathfrak{g}^\alpha$$

For each $\alpha \in R$, let H_α denote the unique element in $\mathfrak{h}$ such that $B(H_\alpha, H) = \alpha(H)$ for all $H \in \mathfrak{h}$. This called the root vector for α.

Proposition 7.

(a) *If α is a root, then so is $-\alpha$.*

(b) *The roots span $\mathfrak{h}^*$ and the root vectors span $\mathfrak{h}$.*

(c) $[\mathfrak{g}^\alpha, \mathfrak{g}^\beta] \subset \mathfrak{g}^{\alpha+\beta}$. *If $\alpha + \beta \notin R$ the bracket is interpreted as 0.*

(d) *The killing form is non-degenerate on $\mathfrak{h}$.*

(e) *The subspace $[\mathfrak{g}^\alpha, \mathfrak{g}^{-\alpha}]$ of $\mathfrak{h}$ has dimension 1. Let $E_\alpha \in \mathfrak{g}^\alpha$ and $E_{-\alpha} \in \mathfrak{g}^{-\alpha}$. Then $[E_\alpha, E_{-\alpha}] = B(E_\alpha, E_{-\alpha})H_\alpha$.*

(f) *For each $\alpha \in R$ the dimension of each $\mathfrak{g}^\alpha$ is 1.*

(g) *If $\alpha \in R$ and $k\alpha \in R$ for some integer k, the $k = \pm 1$.*

Elements E_α of $\mathfrak{g}^\alpha$ with $[E_\alpha, E_{-\alpha}] = H_\alpha$ are called *root elements.* Let $\mathfrak{h}_{\mathbb{R}} = \sum_\alpha \mathbb{R} H_\alpha$.

Example. Consider the Lie algebra $\mathfrak{sl}(3,\mathbb{C}) = \{A \in M_3\mathbb{C} : trA = 0\}$ with the following basis:

$$H_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, H_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

$$X_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, X_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, X_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

$$X_4 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, X_5 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, X_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

The diagonal matrices H_1, H_2 form a basis for the Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{sl}(3,\mathbb{C})$. The dimension of $\mathfrak{sl}(3,\mathbb{C})$ is 8. The rank of $\mathfrak{sl}(3,\mathbb{C})$ is the dimension of $\mathfrak{h}$, which is 2. Hence the roots of $\mathfrak{sl}(3,\mathbb{C})$ are ordered pairs $\alpha = (\alpha_1, \alpha_2) \in \mathbb{C}^2$. Moreover, if X is a root vector, then X is a simultaneous eigenvector for $ad(H_1)$ and $ad(H_2)$ *i.e.*, $[H_1, X] = \alpha_1 X$ and $[H_2, X] = \alpha_2 X$. Below we have the bracket table of the basis elements.

$[\,,]$	H_1	H_2	X_1	X_2	X_3	X_4	X_5	X_6
H_1	0	0	X_1	$2X_2$	X_3	$-X_4$	$-2X_5$	$-X_6$
H_2	0	0	$-X_1$	X_2	$2X_3$	X_4	$-X_5$	$-2X_6$
X_1	$-X_1$	X_1	0	0	X_2	$H_1 - H_2$	$-X_6$	0
X_2	$-2X_2$	$-X_2$	0	0	0	$-X_3$	H_1	X_1
X_3	$-X_3$	$-2X_3$	$-X_2$	0	0	0	X_4	H_2
X_4	X_4	$-X_4$	$-H_1 + H_2$	X_3	0	0	$-X_5$	0
X_5	$2X_5$	X_5	X_6	$-H_1$	$-X_4$	0	0	0
X_6	X_6	$2X_6$	0	$-X_1$	$-H_2$	X_5	0	0

By the above table we have the following roots of $\mathfrak{sl}(3,\mathbb{C})$ with corresponding root vectors, where each root vector X_i is the span of the corresponding root spaces $\mathfrak{g}^{\alpha_i}$.

roots	*vectors*
$\alpha_1 = (1, -1)$	X_1
$\alpha_2 = (2, 1)$	X_2
$\alpha_3 = (1, 2)$	X_3
$\alpha_4 = (-1, 1)$	X_4
$\alpha_5 = (-2, -1)$	X_5
$\alpha_6 = (-1, -2)$	X_6

Notice that the roots are precisely the pairs made up of the non-zero eigenvalues of $ad(H_1)$ and $ad(H_2)$.

$$ad(H_1) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}, ad(H_2) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 \end{pmatrix}$$

Proposition 8.

(a) *The Killing form restricted to* $\mathfrak{h}_{\mathbb{R}}$ *is a real positive-definite bilinear form.*
(b) *Every root* α *takes real values when restricted to* $\mathfrak{h}_{\mathbb{R}}$.
(c) $\mathfrak{h}_{\mathbb{R}}$ *is a real form of* $\mathfrak{h}$, *that is* $\mathfrak{h} = \mathfrak{h}_{\mathbb{R}} \oplus i\mathfrak{h}_{\mathbb{R}}$.

Note. The $\{H_\alpha\}$ are not necessarily linearly independent.

Proposition 9.

(a) *The numbers* $N(\alpha, \beta) = \frac{2(\alpha,\beta)}{(\beta,\beta)}$ *are integers whose only possible values are* $0, \pm1, \pm2, \pm3$. *They are called the Cartan integers.*
(b) *For each* $\alpha \in R$ *we consider the reflection map* $S_\alpha : \mathfrak{h}_{\mathbb{R}}^* \to \mathfrak{h}_{\mathbb{R}}^*$ *with respect to the hyperplane orthogonal to* α, *given by* $S_\alpha(\lambda) = \lambda \frac{2(\alpha,\lambda)}{(\alpha,\alpha)}\alpha$. *Notice that* $S_\alpha(\alpha) = -\alpha$. *Then* $S_\alpha(R) = R$ *(the set of roots is invariant under all* S_α*).*

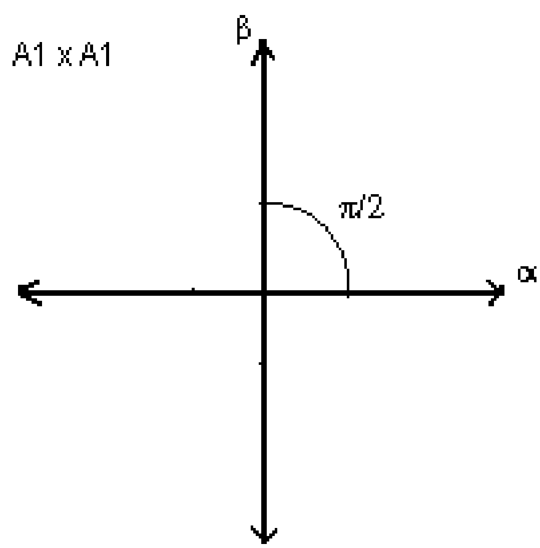
- The set $\{S_\alpha : \alpha \in R\}$ generates a group of isometries of $\mathfrak{h}_{\mathbb{R}}^*$ called the *Weyl group* of R (or of $\mathfrak{g}$) with respect to $\mathfrak{h}$.

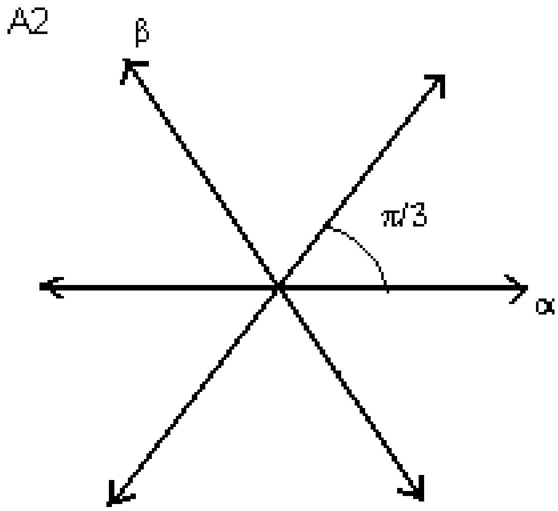
- For any $\alpha, \beta \in R$ with $\beta \neq \pm\alpha$, we have that $[E_\alpha, E_\beta] = N_{\alpha,\beta}E_{\alpha+\beta}$ for some complex number $N_{\alpha,\beta}$. These numbers are the the *structure constants* of $\mathfrak{g}$.

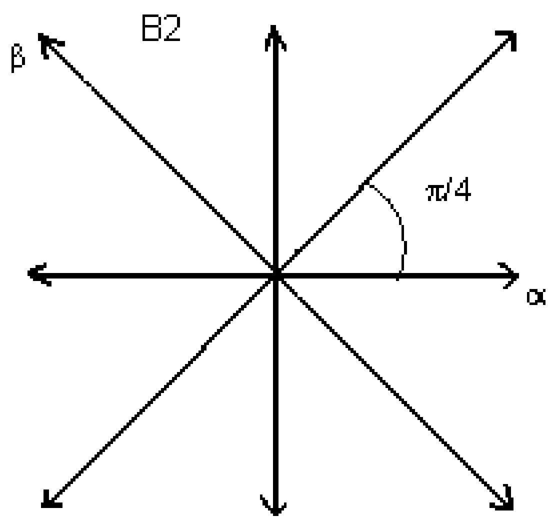
Recall that the cosine of the angle θ between α and β in Euclidean space is given by the formula $\|\alpha\|\|\beta\|\cos\theta = (\alpha, \beta)$. Thus, $N(\beta, \alpha) = 2\frac{\|\beta\|}{\|\alpha\|}$ and $N(\alpha, \beta)N(\beta, \alpha) = 4\cos^2\theta \geq 0$. Also, notice that $N(\alpha, \beta)$ and $N(\beta, \alpha)$ have the same sign, so the following are the only possibilities for angles and relative lengths when $\alpha \neq \pm\beta$ and $\|\beta\| \geq \|\alpha\|$.

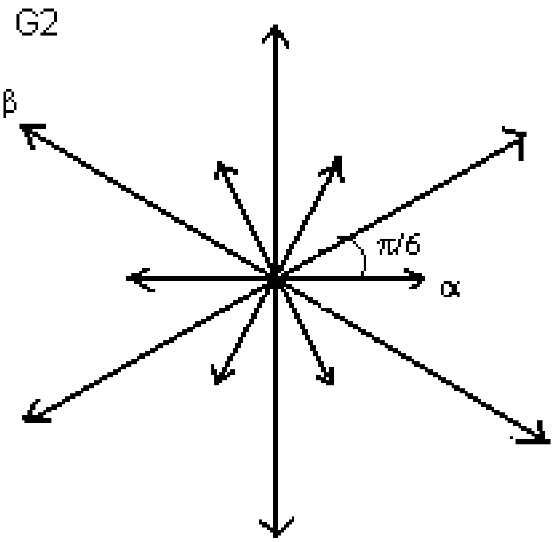
$N(\alpha,\beta)$	$N(\beta,\alpha)$	θ	$\|\beta\|^2/\|\alpha\|^2$
0	0	$\pi/2$	*undetermined*
1	1	$\pi/3$	1
-1	-1	$2\pi/3$	1
1	2	$\pi/4$	2
-1	-2	$3\pi/4$	2
1	3	$\pi/6$	3
-1	-3	$5\pi/6$	3

A reduced root system of rank 1 is called A_1. It consists of a pair of vectors, $\pm\alpha$. There are four non-isomorphic reduced root systems of rank 2; $A_1 \times A_1$, A_2, B_2 and G_2 (see theorem 9 and the following remark), represented by the following graphs. Notice that the root system of $\mathfrak{sl}(3, \mathbb{C})$ in the above example corresponds to the graph of A_2.




A2
β
π/3
α


B2
β
π/4
α



Proposition 10. *For $\alpha, \beta, \gamma\delta \in R$ we have the following:*

(a) $N_{\alpha,\beta} = -N_{\beta,\alpha}$.
(b) $N_{\alpha,\beta} = N_{\beta,\gamma} = N_{\gamma,\alpha}$ *for* $\alpha + \beta + \gamma = 0$
(c) $N_{\alpha,\beta}N_{\gamma,\delta} + N_{\alpha,\gamma}N_{\delta,\beta} + N_{\alpha,\delta}N_{\beta,\gamma} = 0$ *for* $\alpha + \beta + \gamma + \delta = 0$.
(d) $\alpha + \beta$ *is a root if and only if* $N_{\alpha,\beta} \neq 0$.
(e) *It is possible to choose the root elements* $\{E_\alpha\}$ *in such a way so that the structure constants are real numbers satisfying* $N_{\alpha,\beta} = -N_{-\alpha,-\beta}$.
(f) *(Chevalley) The structure constants can be chosen to be integers.*

Definition. Let $\mathfrak{g}$ be a complex semisimple Lie algebra with $\mathfrak{h}$ a Cartan subalgebra, and root system R. Let $H_1, \ldots, H_l$ (l =rank of $\mathfrak{g}$) be a basis for $\mathfrak{h}$. For each $\alpha \in R$ let E_α be root elements satisfying $E_\alpha, E_{-\alpha}] = H_\alpha$, and such that the structure constants are integers with $N_{\alpha,\beta} = -N_{-\alpha,-\beta}$. Then the set $\{H_1, \ldots, H_l; E_\alpha : \alpha \in R\}$ is said to be a *Weyl-Chevalley basis* for $\mathfrak{g}$.

Proposition 11. *Let R be the root system of a complex semisimple Lie algebra $\mathfrak{g}$(with respect to a fixed Cartan subalgebra). Then there exists a subset $\Pi = \{\alpha_1, \ldots, \alpha_l\}$ (l =rank of $\mathfrak{g}$) such that every root $\alpha \in R$ can be expressed uniquely as $\alpha_1\alpha_1 + \ldots + n_l\alpha_l$, where n_i are integers, either all positive or all negative.*

We call Π a *set of simple roots* for R. Π is called irreducible if there is no nontrivial disjoint union $\Pi = \Pi_1 \cup \Pi_2$ with $(\alpha, \beta) = 0$ for all $\alpha \in \Pi_1$ and $\beta \in \Pi_2$.

A root is called positive ($\alpha > 0$) if $\alpha = \sum_{i=1}^{l} n_i\alpha_i$ with all $n_i \geq 0$. Let R^+ denote the set of positive roots and let $R^- = \{-\alpha : \alpha \in R^+\}$.

Proposition 12. *R^+ is called an ordering in R and satisfies the following:*

(a) $R^+ \cap (-R^+) = \emptyset$, $R^+ \cup (-R^+) = R$.
(b) *For each* $\alpha, \beta \in R^+$ *with* $\alpha + \beta \in R^+$, *i.e, for each* $\alpha, \beta \in R$, *then* $\alpha > \beta$ *if and only if* $\alpha - \beta \in R^+$.

Definition. The *Dynkin diagram* of a root system R with a set of simple roots Π consists of a planar graph with l vertices labeled with $\alpha_1, \ldots, \alpha_l$ and $N(\alpha_i, \alpha_j)N(\alpha_j, \alpha_i)$ line segments joining the vertex α_i to the vertex α_j. If $N(\alpha, \beta) > 0$ and $(\beta, \beta) > (\alpha, \alpha)$, draw an arrow on the line segment from the vertex of β (long root) to the vertex from the vertex of α (short root).

Theorem 10. *Classification Assigning to each complex semisimple Lie algebra the Dynkin diagram of the root system of a Cartan subalgebra, sets up a one-to-one correspondence between the set of such Lie algebras (up to isomorphism) and fundamental root systems (up to equivalence). In particular, the simple Lie algebras correspond to irreducible fundamental systems.*

The following is a list of the complex Lie algebras along with their Dynkin diagrams

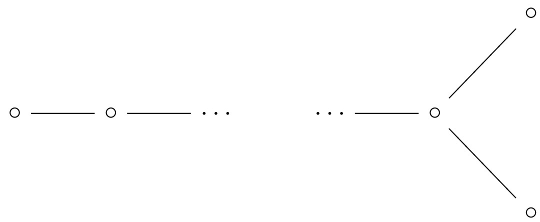
Name	*Description*	*Rank*	*Dimension*
A_l	$\mathfrak{sl}_{l+1}\mathbb{C}$	$l \geq 1$	$l(l+2)$
B_l	$\mathfrak{so}_{2l+1}\mathbb{C}$	$l \geq 2$	$l(2l+1)$
C_l	$\mathfrak{sp}_l\mathbb{C}$	$l \geq 3$	$l(2l+1)$
D_l	$\mathfrak{so}_{2l}\mathbb{C}$	$l \geq 4$	$l(2l+1)$
G_2	$--$	2	14
F_4	$--$	4	52
E_6	$--$	6	78
E_7	$--$	7	133
E_8	$--$	8	248

A_l $\quad \circ - \circ - \cdots \quad \cdots - \circ - \circ$

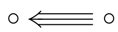
B_l $\quad \circ - \circ - \cdots \quad \cdots - \circ \Longrightarrow \circ$

C_l $\quad \circ - \circ - \cdots \quad \cdots - \circ \Longleftarrow \circ$

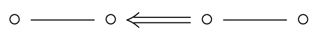
D_l



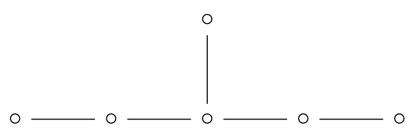
G_2



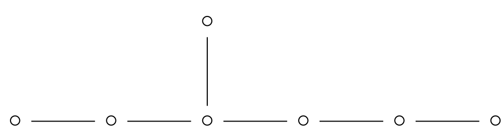
F_4



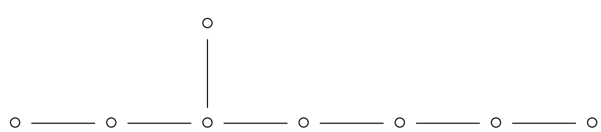
E_6



E_7



E_8



Definition. A real Lie algebra $\mathfrak{g}_0$ is called a *real form* of a complex Lie algebra $\mathfrak{g}$, if $\mathfrak{g}$ is isomorphic to the complexification of $\mathfrak{g}_0$, that is, $\mathfrak{g} = \mathfrak{g}_0 + i\mathfrak{g}_0$.

Remark. (Weyl) Every complex semisimple Lie algebra has a compact real form. Compact means that its Killing form is negative definite. All compact real forms of $\mathfrak{g}$ are conjugate via an inner automorphism.

Any real form $\mathfrak{g}_{\mathbb{R}}$ can be characterized as a fixed point set of a conjugate linear involution $\tau : \mathfrak{g} \to \mathfrak{g}$ ($\tau^2 = I, \tau$ is linear over the reals), which is an automorphism of $\mathfrak{g}$ considered as a real Lie algebra. If $\mathfrak{g} = \mathfrak{h} \oplus \sum_{\alpha \in R} \mathfrak{g}^{\alpha}$ is a root space decomposition of $\mathfrak{g}$, then we can construct a compact real form $\mathfrak{g}_0$ of $\mathfrak{g}$ by the conjugate linear map $\tau_0 : \mathfrak{g} \to \mathfrak{g}$ defined by

$$\tau|_{\mathfrak{h}_{\mathbb{R}}} = -Id, \qquad \tau(E_\alpha) = -E_{-\alpha}.$$

This is called the standard involution associated with the root space decomposition. The set of fixed points of τ is the compact real form given by

$$\mathfrak{g}_0 = i\mathfrak{h}_{\mathbb{R}} \oplus \bigoplus_{\alpha \in R^+} \mathbb{R}(E_\alpha - E_{-\alpha} \oplus \bigoplus_{\alpha \in R^+} \mathbb{R}i(E_\alpha + E_{-\alpha})$$

The elements of $iH_\alpha, E_\alpha - E_{-\alpha}, iE_\alpha + E_{-\alpha}$ generate a subalgebra of $\mathfrak{g}$ isomorphic to $\mathfrak{su}(2)$ given by

$$iH_\alpha \mapsto \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, E_\alpha - E_{-\alpha} \mapsto \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, E_\alpha + E_{-\alpha} \mapsto \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},$$

We can also use the root spaces to obtain homomorphisms into $\mathfrak{sl}(\mathbb{C}$ given by

$$iH_\alpha \mapsto \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, E_\alpha \mapsto \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, E_{-\alpha} \mapsto \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

References

[1] Andreas Arvanitoyeorgos, *An Introduction to Lie Groups and the Geometry of Homogeneous Spaces*. American Mathematical Society, 2003.

[2] Brian C. Hall, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction.* Springer Science+Business Media, 2003.

[3] James E. Humphreys, *Introduction to Lie Algebras and Representation Theory*. Springer-Verlag, 1972.

[4] N. Khaneja, R. Brocktett, and S. Glaser, "'Time Optimal Control of Spin Systems"', Phys. Rev. A, 63, 032308, 2001.

[5] Navin Khaneja, *Geometric Control in Classical and Quantum Systems* (Ph.d Thesis, Harvard University, 2000).

[6] Kobayashi, S., and Nomizu, K., *Foundations of Differential Geometry Vol.1 and 2*. Interscience Publishers, 1969.

[7] Richard W. Sharpe, *Differential Geometry: Cartan's Generalization of Klein's Erlangen Program*. Springer-Verlag, 1997.

Mathematics Department, University of Hawai'i, Mānoa, 2565 The Mall, Honolulu, HI 96822

E-mail address: jody@math.hawaii.edu