Communications in Analysis and Mechanics

https://www.aimspress.com/journal/cam

Communications in Analysis and Mechanics, 16(2): 345–387.
DOI: 10.3934/cam.2024017
Received: 15 November 2023
Revised: 21 February 2024
Accepted: 26 March 2024
Published: 08 May 2024

Research article

Time optimal problems on Lie groups and applications to quantum control

Velimir Jurdjevic*

Department of Mathematics, University of Toronto, Toronto,Ontario M5S 3G3, Canada

* **Correspondence:** Email: jurdj@math.utoronto.ca.

Abstract: In this paper we introduce a natural compactification of a left (right) invariant affine control system on a semi-simple Lie group G in which the control functions belong to the Lie algebra of a compact Lie subgroup K of G and we investigate conditions under which the time optimal solutions of this compactified system are "approximately" time optimal for the original system. The basic ideas go back to the papers of R.W. Brockett and his collaborators in their studies of time optimal transfer in quantum control ([1], [2]). We showed that every affine system can be decomposed into two natural systems that we call horizontal and vertical. The horizontal system admits a convex extension whose reachable sets are compact and hence posess time-optimal solutions. We then obtained an explicit formula for the time-optimal solutions of this convexified system defined by the symmetric Riemannian pair (G, K) under the assumption that the Lie algebra generated by the control vector fields is equal to the Lie algebra of K.

In the second part of the paper we applied our results to the quantum systems known as Icing n-chains (introduced in [2]). We showed that the two-spin chains conform to the theory in the first part of the paper but that the three-spin chains show new phenomena that take it outside of the above theory. In particular, we showed that the solutions for the symmetric three-spin chains studied by ([3], [4]) are solvable in terms of elliptic functions with the solutions completely different from the ones encountered in the two-spin chains.

Keywords: symplectic; manifolds; Lie-Poisson bracket; Lie algebras; co-adjoint orbits; extremal curves; integrable systems
Mathematics Subject Classification: Primary 53C17, 53C22, 53B21, 53C25, 30C80, 26D05; Secondary 49J15, 58E40.

1. Introduction

This study will address time-optimal solutions of affine systems defined by the pairs (G, K) where G is a semi-simple Lie group and K is a compact subgroup of G with a finite centre. Such pairs of Lie groups are reductive in the sense that the Lie algebra $\mathfrak{g}$ of G admits a decomposition $\mathfrak{g} = \mathfrak{p} + \mathfrak{k}$ with $\mathfrak{p}$ the orthogonal complement of the Lie algebra $\mathfrak{k}$ of K relative to the Killing form in $\mathfrak{g}$ that satisfies Lie algebra condition $[\mathfrak{p}, \mathfrak{k}] \subseteq \mathfrak{p}$. We will then consider time-optimal solutions of affine control systems of the form

$$\frac{dg}{dt} = X_0(g(t)) + \sum_{i=1}^{m} u_i(t)X_i(g(t))) \tag{1.1}$$

where $X_o, \dots, X_em$ are all left-invariant vector fields on G under the assumption that the drift element X_0 belongs to $\mathfrak{p}$ at the group identity and that the controlling vector fields $X_i, i = 1, \dots, m$ belong to $\mathfrak{k}$ at the group identity. We will write such systems as

$$\frac{dg}{dt} = g(t)(A + \sum_{i=1}^{m} u_i(t)B_i), \tag{1.2}$$

where $A = X_0(e)$ and $B_i = X_i(e)$,$i = 1, \dots, m$.

We will be particularly interested in the pairs (G, K) in which K is the set of fixed points by an involutive automorphism σ on G. Recall that $\sigma \neq I$ is an involutive automorphism on G that satisfies $\sigma^2 = I$ where I is the identity map in G. Then, the tangent map σ_* at e of σ is a Lie algebra isomorphism that satisfies $\sigma_*^2 = I$, where now I is the identity map on the Lie algebra $\mathfrak{g}$. Therefore $(\sigma_* + I)(\sigma_* - I) = 0$, and $\mathfrak{g} = ker(\sigma_* + I) \oplus ker(\sigma_* - I)$, i.e.,

$$\mathfrak{g} = \{X \in \mathfrak{g} : \sigma_* X = -X\} \oplus \{X \in \mathfrak{g} : \sigma_* X = X\}. \tag{1.3}$$

It follows that $\mathfrak{k} = \{X \in \mathfrak{g} : \sigma_*(X) = X\}$ is the Lie algebra of K and that $\mathfrak{p} = \{X \in \mathfrak{g} : \sigma_*(X) = -X\}$ is a vector space in $\mathfrak{g}$ that coincides with the orthogonal complement of $\mathfrak{k}$ and satisfies $[\mathfrak{p}, \mathfrak{p}] \subseteq \mathfrak{k}$. In the literature of symmetric Riemannian spaces the decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ subject to

$$[\mathfrak{k}, \mathfrak{k}] \subseteq \mathfrak{k}, [\mathfrak{p}, \mathfrak{k}] \subseteq \mathfrak{p}, [\mathfrak{p}, \mathfrak{p}] \subseteq \mathfrak{k} \tag{1.4}$$

is called a *Cartan decomposition* ([5], [6]). A symmetric pair is said to be of *compact type* if the Killing form is negative definite on $\mathfrak{p}$. Compact type implies that G is a compact Lie group (prototypical example $G = SU(n), K = SO(n, \mathbb{R})$). The pair (G, K) is said to be of *non-compact type* if the Killing form is positive definite on $\mathfrak{p}$ (prototypical example $G = SL(n, \mathbb{R}), K = SO(n, \mathbb{R})$) ([5]). We will assume that the pair (G, K) is one of these two types. In either case $Kl(X, Y)$ will denote the Killing form on $\mathfrak{g}$. Recall that Kl is non-degenerate on $\mathfrak{g}$.

This background information shows that in each affine system (1.1) there is a natural energy function

$$E = \frac{1}{2} \int_0^T \langle U(t), U(t) \rangle \, dt, U(t) = \sum_{i=1}^{m} u_i(t)B_i$$

where the scalar product $\langle \, , \, \rangle$ is the negative of the Killing form.This energy function induces a natural variational problem, called *affine-quadratic problem*, defined as follows: given two boundary conditions in G and a time interval $[0, T]$, find a solution $g(t)$ of (1.1) that satisfies $g(0) = g_0$, $g(T) = g_1$ whose energy of transfer $\int_0^T \langle U(t), U(t) \rangle \, dt$ is minimal. Remarkably, every affine system (1.1) is controllable on G whenever A is regular and the Lie algebra $\mathfrak{k}_v$ generated by $B_1, \dots, B_m$ is equal to $\mathfrak{k}$ and the corresponding extremal Hamiltonian system obtained by the maximum Principle is completely integrable ([7]).

In contrast to the above energy problem, time-optimal problems are more elusive due to the fact that the reachable sets need not be closed because the control functions are not bounded (it may happen that certain points in G that can be reached in an arbitrarily short time, but are not reachable in zero time, as will be shown later). More generally, it is known that any point of the group K_v generated by the exponentials in the Lie algebra $\mathfrak{k}_v$ generated by $B_1, \ldots, B_m$ belongs to the topological closure of the set of reachable points $\mathcal{A}(e, \leq T)$ in any positive time T, and yet it is not known (although it is generally believed) that each point in K_v can be reached in an arbitrarily short time from the group identity e. This lack of information about the boundary of the reachable sets in the presence of a drift vector still remains an impediment in the literature dealing with time optimality ([1, 8–10]).

In this paper we will adopt the definition of R. W. Brockett et al. ([1], [2]) according to which the optimal time T that g_1 can be approximately reached from g_0 is defined as $T = inf\{t : g_1 \in \bar{\mathcal{A}}(g_0, \leq t)\}$, where $\bar{\mathcal{A}}(g_0, \leq t)$ denotes the topological closure of the set of points reachable from g_0 in t or less units of time by the trajectories of (1.2). Then $\mathcal{T}(g)$ will denote the minimal time that g is approximately reachable from the group identity e.

It is evident that Brockett's definition of time optimality is invariant under any enlargement of the system that keeps the closure of the reachable set $\mathcal{A}(e, \leq t)$ the same. In particular, the optimal time is unchanged if the original system is replaced by

$$\frac{dg}{dt} = g(t)(A + U(t)), \tag{1.5}$$

where now $U(t)$ is an arbitrary curve in $\mathfrak{k}_v$. Let now K_v denote the Lie subgroup generated by the exponentials in $\mathfrak{k}_v$. We shall assume that K_v is a closed subgroup of K, which then implies that K_v is compact, since K is compact. Recall that every point in K_v belongs to the closure of $\mathcal{A}(e, \leq t)$ for any $t > 0$. Therefore $\mathcal{T}(h) = 0$ for any $h \in K_v$.

Each affine system (1.5) defines a distinctive *horizontal system*

$$\frac{dg}{dt} = g(t)Ad_h(t)A, h(t) \in K_v. \tag{1.6}$$

These two systems are related as follows: every solution $g(t)$ of (1.5) generated by a control $U(t) \in \mathfrak{k}_v$ defines a solution $\hat{g}(t) = g(t)h^{-1}(t)$ of the horizontal system whenever $\frac{dh}{dt} = h(t)U(t)$. Conversely, every solution $\hat{g}(t)$ of the horizontal system gives rise to a solution $g(t) = \hat{g}(t)h(t)$ of the affine system for $h(t)$ a solution of $\frac{dh}{dt} = h(t)U(t)$. It follows that $\mathcal{T}(\hat{g}) = \mathcal{T}(gh^{-1}) = \mathcal{T}(g)$, and that $\bar{\mathcal{A}}_h(e, \leq t) \subseteq \bar{\mathcal{A}}(e, \leq t)$, where $\mathcal{A}_h(e, \leq t)$ denotes the reachable set of the horizontal system.

The above horizontal system can be extended to the convexified system without altering the closure of the reachable sets $\mathcal{A}(g_0, \leq t)$. The convexified system is given by

$$\frac{dg}{dt} = g(t)\sum_{i=1}^{k} \lambda_i(t)Ad_{h_i(t)}(A), \lambda_i(t) \geq 0, \sum_{i=1}^{k} \lambda_i(t) = 1. \tag{1.7}$$

We will think of this system as a control system with $h_1(t), \ldots h_k(t)$ in K_v and $\lambda_1(t), \ldots \lambda_k(t)$ as the control functions, and we will use $\mathcal{A}_{conv}(e, \leq t)$ to denote the points in G reachable from e in t or less units of time by the solutions of (1.7).

The following proposition summarizes the relations between (1.5), (1.6) and (1.7).

Proposition 1. *$\mathcal{A}_{conv}(e, \leq T)$ is a compact set equal to $\bar{\mathcal{A}}_h(e, \leq T)$ for each $T > 0$. Therefore, $\mathcal{A}_{conv}(e, \leq t) = \bar{\mathcal{A}}_h(e, \leq t) \subseteq \bar{\mathcal{A}}(e, \leq t)$.*

This proposition is a paraphrase of the well known results in geometric control theory: Theorem 11 in [11], p. 88 implies that

$$\mathcal{A}_{conv}(e, \leq t) = \bar{\mathcal{A}}_h(e, \leq t) \subseteq \bar{\mathcal{A}}(e, \leq t)$$

and Theorem 11 in [11] on p.119 states that $\mathcal{A}_{conv}(e, \leq t)$ is compact.

Equation (1.7) may be regarded as the compactification of (1.6). The following proposition captures its essential properties.

Proposition 2. *The optimum time $\mathcal{T}(\mathbf{g})$ is equal to the minimum time required for a trajectory of the convexified system to reach the coset $\mathbf{g}K_v$ from the group identity.*

Proof. If $\mathbf{g} \in \bar{\mathcal{A}}(e, \leq T)$ then there is a sequence of trajectories $g_n(t)$ of (1.5) and a sequence of times $\{t_n\}$ such that $\lim g_n(t_n) = \mathbf{g}$. There is no loss in generality in assuming that $\{t_n\}$ converges to a time $t, t \leq T$. Let $\tilde{g}_n(t) = g_n(t)h_n(t), h_n(t) \in K_v$ denote the corresponding sequence of trajectories in (1.6). Since K_v is compact there is no loss in generality in assuming that $h_n(t_n)$ converges to an element h in K_v. Then $\lim \tilde{g}_n(t_n) = \mathbf{g}h$ and $\mathbf{g}h$ belongs to $\bar{\mathcal{A}}_h(e, \leq t)$. But then $\mathbf{g}h$ is reachable by the convexified system (1.7) since $\mathcal{A}_{conv}(e, \leq T) = \bar{\mathcal{A}}_h(e, \leq T)$.

Conversely if $\mathbf{g}h \in \mathcal{A}_{conv}(e, \leq T)$, then the same argument followed in reverse order shows that $\mathbf{g} \in \bar{\mathcal{A}}(e, \leq T)$. Therefore, $\mathcal{T}(\mathbf{g}) = T_{conv}$, where T_{conv} is the first time that a point of $\mathbf{g}K_v$ is reachable from e by a trajectory of the convexified system (1.7).

The paper is organized as follows. We begin with the algebraic preliminaries needed to show that the convex hull of $\{Ad_h(A), h \in K\}$ contains an open neighbourhood of the origin in $\mathfrak{p}$ whenever A is regular and $K_v = K$ (an element X in $\mathfrak{p}$ is regular if the set $\{P \in \mathfrak{p} : [P, X] = 0\}$ is an abelian subalgebra in $\mathfrak{g}$ contained in $\mathfrak{p}$). This result implies two important properties of the system. First, it shows that the stationary curve $g(t) = g(0)$ is a solution of the convexified system, which it turn implies that any coset gK can be reached in an arbitrarily short time by a trajectory of the convexified system. Second, it shows that the positive convex cone spanned by $\{Ad_h(A), h \in K\}$ is equal to $\mathfrak{p}$. Therefore, the convexified system is controllable whenever $[\mathfrak{p}, \mathfrak{p}] = \mathfrak{k}$. These facts then imply that any two points in G can be connected by a time-optimal trajectory of the convexified system, and they also imply that any point g_0 in G can be connected to any coset g_1K by a time optimal trajectory of the convexified system. We then follow these findings with the extremal equations obtained by the maximum principle. We show that the time- optimal solutions on G are either stationary, or are of the form

$$g(t) = g(0)e^{t(P+Q)}e^{-tQ}, \tag{1.8}$$

for some elements $P \in \mathfrak{p}$ and $Q \in \mathfrak{k}$.

The non-stationary solutions on G/K are of the form

$$\pi(g(0)e^{tP}), P \in \mathfrak{p}, \tag{1.9}$$

where π denotes the natural projection $\pi(g) = gK$. Since $\pi(g(0)e^{tP}), P \in \mathfrak{p}, ||P|| = 1$, coincide with the geodesics on G/K emanating from $\pi(g(0))$ (relative to its natural G-invariant metric) it follows that t is the length of the geodesic that connects $\pi(g(0))$ to $\pi(g(0)e^{tP})$. Evidently minimal time corresponds to the length of the shortest geodesic that connects these points.

Remark 1. *The papers of Brockett et al ([1] and [2]) claim that the time optimal solutions in* (1.1) *can be obtained solely from the horizontal system* (1.6)*, but that cannot be true for the following reasons: every trajectory $g(t)$ of the horizontal system $\frac{dg}{dt} = g(t)Ad_{h(t)}(A)$ is generated by a control $U(t) = Ad_{h(t)}(A)$ that satisfies $\|U(t)\|^2 = \|Ad_{h(t)}(A)\|^2 = \|A\|^2$. Hence $U(t)$ cannot be equal to zero, and $g(t)$ cannot be stationary.*

In the second part of the paper we apply our results to quantum systems known as Icing n-chains (introduced in [2]). We will show that the two-spin chains conform to the above theory and that their time-optimal solutions are given by equations (1.8). The three-spin systems, however, do not fit the above formalism due to the fact that the Lie algebra generated by the controlling vector fields does not meet Cartan's conditions (1.4). We provide specific details suggesting why the solutions fall outside the above theory. We end the paper by showing that the symmetric three-spin chain studied by ([3], [4]) is solvable in terms of elliptic functions. The solution of the symmetric three-spin system is both new and instructive, in the sense that it foreshadows the challenges in the more general cases.

2. Convexified horizontal systems

2.1. Algebraic background

We will continue with the symmetric pairs (G, K), with G semisimple and K a compact subgroup of G subject to Cartan's conditions (1.4). We recall that the Killing form is positive on $\mathfrak{p}$ in the non-compact cases, and is negative on $\mathfrak{p}$ in the compact cases. In either case $\mathfrak{g}$ admits a fundamental decomposition

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \cdots \oplus \mathfrak{g}_m, \mathfrak{g}_i = \mathfrak{p}_i \oplus [\mathfrak{p}_i, \mathfrak{p}_i], \mathfrak{p} = \mathfrak{p}_1 \oplus \cdots \oplus \mathfrak{p}_m \tag{2.1}$$

where each $\mathfrak{g}_i$ is a simple ideal in $\mathfrak{g}$ and $[\mathfrak{g}_i, \mathfrak{g}_j] = 0, i \neq j$ ([11], p.123). It then follows that $\mathfrak{p} \oplus [\mathfrak{p}, \mathfrak{p}] = \mathfrak{g}$, a fact that is important for controllability, as we shall see later on. As before, $\langle\ ,\ \rangle$ will denote a suitable scalar multiple of the Killing form.

We recall that an element X in $\mathfrak{p}$ is regular if the set $\mathfrak{h} = \{P \in \mathfrak{p} : [P, X] = 0\}$ is an abelian subalgebra in $\mathfrak{g}$ contained in $\mathfrak{p}$. It follows that $\mathfrak{h}$ is a maximal abelian algebra that contains X. It is easy to verify that the projection of a regular element on each factor $\mathfrak{p}_i$ is non-zero. The following proposition summarizes the essential relations between regular elements and maximal abelian sub-algebras in $\mathfrak{p}$.

Proposition 3. *i. Every maximal abelian algebra in $\mathfrak{p}$ contains a regular element.*
ii.. Any two maximal abelian algebras $\mathfrak{h}$ and $\mathfrak{h}^*$ in $\mathfrak{p}$ are K conjugate, i.e., $Ad_k(\mathfrak{h}) = \mathfrak{h}^*$ for some $k \in K$.
iii. $\mathfrak{p}$ is the union of maximal abelian algebras in $\mathfrak{p}$.

The above results, as well as the related theory of Weyl groups and Weyl chambers are well known in the theory of symmetric Riemannian spaces ([5], [6]), but their presentation is often directed to a narrow group of specialists and, as such, is not readily accessible to a wider mathematical community. For that reason, we will present all these theoretical ingredients in a self contained manner, and in the process we will show their relevance for the time-optimal problems defined above.

If $\mathfrak{h}$ is a maximal abelian algebra in $\mathfrak{p}$ then $\mathcal{F} = \{adX : X \in \mathfrak{h}\}$ is a collection of commuting linear transformations in $\mathfrak{g}$ because $[adX, adY] = ad[X, Y] = 0$ for any X and Y in $\mathfrak{h}$. In the non-compact case, $\mathfrak{g}$ is a Euclidean space relative to the scalar product $\langle X, Y\rangle_\sigma = -Kl(\sigma_* X, Y)$ induced by the automorphism σ. Relative to this scalar product each $adH, H \in \mathfrak{p}$ is a symmetric linear transformation in $gl(\mathfrak{g})$. Then, it

is well known that $\mathcal{F}$ can be simultaneously diagonalized over $\mathfrak{g}$. That is, there exist mutually orthogonal vector spaces $\mathfrak{g}_0$, $\mathfrak{g}_\alpha$, with α in some finite set Δ such that:
1. $\mathfrak{g}_0 = \cap_{H\in\mathfrak{h}} ker(adH)$.
2. $\mathfrak{g} = \mathfrak{g}_0 \oplus \sum_{\alpha\in\Delta} \mathfrak{g}_\alpha$,
3. $adH = \alpha(H)I$ on $\mathfrak{g}_\alpha$ for each $H \in \mathfrak{h}$, and $\alpha(H) \neq 0$ for some $H \in \mathfrak{h}$.
Additionally,

$$\alpha(H)\sigma_*\mathfrak{g}_\alpha = \sigma_*(adH(\mathfrak{g}_\alpha)) = (ad\sigma_*H)(\sigma_*\mathfrak{g}_\alpha) = -ad(H)(\sigma_*\mathfrak{g}_\alpha),$$

which implies that Δ is symmetric, that is $-\alpha \in \Delta$ for each $\alpha \in \Delta$. It is not hard to show that each $\alpha \in \Delta$ is a linear function on $\mathfrak{h}$, i.e., Δ is a subset of $\mathfrak{h}^*$. In the literature on symmetric spaces $\mathfrak{g}_\alpha$ are called root spaces, and elements $\alpha \in \Delta$ are called roots ([5]).

In the compact case, the Killing form is negative on $\mathfrak{g}$. Therefore $\mathfrak{g}$ is a Euclidean vector space relative to the scalar product $\langle\,,\,\rangle = -Kl$. Since $Kl(X, [Y, Z]) = Kl([X, Y], Z)$, $\langle ad(H)X, Y\rangle = -\langle X, ad(H)Y\rangle$. Hence each $ad(H)$ is a skew-symmetric linear operator on $\mathfrak{g}$. It follows that $\mathcal{F} = \{adH : H \in \mathfrak{h}\}$ is a family of commuting skew-symmetric operators on $\mathfrak{g}$ for each maximal abelian algebra $\mathfrak{h}$; as such, $\mathcal{F}$ can be simultaneously diagonalized, but this time over the complexified algebra $\mathfrak{g}^c$.

The complexified Lie algebra $\mathfrak{g}^c$ consists of elements $Z = X + iY, X, Y \in \mathfrak{g}$ with the obvious Lie algebra structure inherited from $\mathfrak{g}$. Then $\mathfrak{g}^c = \mathfrak{p}^c \oplus \mathfrak{k}^c$ with $\mathfrak{p}^c = \mathfrak{p} + i\mathfrak{p}$ and $\mathfrak{k}^c = \mathfrak{k} + i\mathfrak{k}$. It is evident that $\mathfrak{p}^c$ and $\mathfrak{k}^c$ satisfy Cartan's conditions

$$[\mathfrak{p}^c, \mathfrak{p}^c] \subseteq \mathfrak{k}^c, [\mathfrak{p}^c, \mathfrak{k}^c] \subseteq \mathfrak{p}^c, [\mathfrak{k}^c, \mathfrak{k}^c] \subseteq \mathfrak{k}^c$$

whenever $\mathfrak{p}$ and $\mathfrak{k}$ satisfy conditions (1.4).

In order to make advantage of the corresponding eigenspace decomposition we will regard $\mathfrak{g}^c$ as a Hermitian vector space with the Hermitian product

$$\langle\langle X + iY, Z + iW\rangle\rangle = \langle X, Z\rangle + \langle Y, W\rangle + i(\langle Y, Z\rangle - \langle X, W\rangle). \tag{2.2}$$

We recall that Hermitian means that $\langle\langle\,,\,\rangle\rangle$ is bilinear and satisfies

$$\langle\langle u, u\rangle\rangle \geq 0, \langle\langle v, u\rangle\rangle = \overline{\langle\langle u, v\rangle\rangle}, \tag{2.3}$$

for any u and v in $\mathfrak{g}^c$. One can easily show that for each $H \in \mathfrak{h}$

$$\langle\langle adH(X + iY), Z + iW)\rangle\rangle = -\langle\langle X + iY, adH(Z + iW)\rangle\rangle,$$

therefore each adH is a skew-Hermitian transformation on $\mathfrak{g}^c$.

It follows that $\mathcal{F} = \{adH, H \in \mathfrak{h}\}$ becomes a family of commuting skew-Hermitian operators on $\mathfrak{g}^c$, and consequently can be simultaneously diagonalized. If λ is an eigenvalue of a skew-symmetric transformation T, then λ is imaginary, because $Tx = \lambda x$ means that

$$\lambda||x||^2 = \langle Tx, x\rangle = -\langle x, Tx\rangle = -\bar{\lambda}||x||^2.$$

Hence $\lambda = -\bar{\lambda}$. We will write $\lambda = i\alpha$. So, if X_α is the eigenvector corresponding to $i\alpha \neq 0$ then $ad(H)(X_\alpha) = i\alpha(H)X_\alpha$, $H \in \mathfrak{h}$. It follows that $\alpha \in \mathfrak{h}^*$ because

$$i\alpha(\lambda H_1 + \mu H_2)X_\alpha = \lambda ad(H_1)(X_\alpha) + \mu ad(H_2)(X_\alpha) = i(\lambda\alpha(H_1) + \mu\alpha(H_2))X_\alpha,$$

hence $\alpha(\lambda H_1 + \mu H_2) = \lambda\alpha(H_1) + \mu\alpha(H_2)$. Then $\mathfrak{g}^c_\alpha$ will denote the eigenspace corresponding to $i\alpha$ for each non-zero eigenvalue $i\alpha$, that is,

$$\mathfrak{g}^c_\alpha = \{X \in \mathfrak{g}^c : ad(H)X = i\alpha(H)X, H \in \mathfrak{h}\}, \alpha(H) \neq 0, \text{ for some } H \in \mathfrak{h}.$$

Since

$$ad(H)\overline{X} = \overline{ad(H)X} = -i\alpha(H)\overline{X}, H \in \mathfrak{h},$$

$-i\alpha$ is a non-zero eigenvalue for each eigenvalue α. We will let $i\Delta$ denote the set of non-zero eigenvalues of $\{ad(H), H \in \mathfrak{h}\}$. As in non-compact case, Δ is a symmetric and a finite set in $\mathfrak{h}^*$. It then follows that the eigenspaces $\mathfrak{g}^c_\alpha$ corresponding to different eigenvalues are orthogonal with respect to $\langle\langle\, , \,\rangle\rangle$ and $\mathfrak{g}^c = \mathfrak{g}^c_0 + \sum_{\alpha\in\Delta} g^c_\alpha$, where $\mathfrak{g}^c_0$ is given by $\cap_{H\in\mathfrak{h}} ker(adH)$ and where the sum is direct.

Every $Z \in \mathfrak{g}^c$ can be written as $Z = Z_0 + \sum_{\alpha\in\Delta} Z_\alpha$ in which case

$$adH(Z) = \sum_{\alpha\in\Delta} i\alpha(H)Z_\alpha, Z_\alpha \in \mathfrak{g}^c_\alpha. \tag{2.4}$$

Then $Z \in \mathfrak{g}$ if and only if $Z_\alpha + \bar{Z}_\alpha = 0$ and $\bar{Z}_0 = Z_0$. If H is such that $\alpha(H) \neq 0$ for all α, then $adH(Z) = 0$ if and only if $Z_\alpha = 0$ for all α.

Suppose now that $Z \in \mathfrak{g} \cap \mathfrak{g}^c_0$ that is, suppose that $adH(Z) = 0$ for all $H \in \mathfrak{h}$. Then, $Z = X + Y$ for some $X \in \mathfrak{p}$, and $Y \in \mathfrak{k}$. Our assumption that $adH(X + Y) = 0$ yields $[H, X] = 0$ and $[H, Y] = 0$. Hence $X \in \mathfrak{h}$ and $Y \in \mathfrak{k}$ belongs to the Lie algebra $\mathfrak{m}$ in $\mathfrak{k}$ consisting of all elements Y such that $[H, Y] = 0$ for all $H \in \mathfrak{h}$.

Proposition 4. *For each $\alpha \in \Delta$ there exist non-zero elements $X_\alpha \in \mathfrak{p}$ and $Y_\alpha \in \mathfrak{k}$ such that*

$$adH(X_\alpha) = -\alpha(H)Y_\alpha, adH(Y_\alpha) = \alpha(H)X_\alpha, \textit{ compact case,} \tag{2.5}$$

and

$$adH(X_\alpha) = \alpha(H)Y_\alpha, adH(Y_\alpha) = \alpha(H)X_\alpha, \textit{non-compact case.} \tag{2.6}$$

In either case $[X_\alpha, Y_\alpha] \in \mathfrak{h}$.

Proof. Let us begin with the compact case with Z_α in $\mathfrak{g}^c_\alpha$ a non-zero element such that $adH(Z_\alpha) = i\alpha(H)Z_\alpha$ for some element $H \in \mathfrak{h}$ such that $\alpha(H) \neq 0$. If $Z_\alpha = U_\alpha + iV_\alpha$ with $U_\alpha \in \mathfrak{g}$ and $V_\alpha \in \mathfrak{g}$, then

$$adH(U_\alpha) = -\alpha(H)V_\alpha, adH(V_\alpha) = \alpha(H)U_\alpha.$$

These relations imply that neither $U_\alpha = 0$ nor $V_\alpha = 0$. Let now

$$U_\alpha = U^{\mathfrak{p}}_\alpha + U^{\mathfrak{k}}_\alpha, V_\alpha = V^{\mathfrak{p}}_\alpha + V^{\mathfrak{k}}_\alpha,$$

with $U^{\mathfrak{p}}_\alpha, V^{\mathfrak{p}}_\alpha$ in $\mathfrak{p}$ and $U^{\mathfrak{k}}_\alpha, V^{\mathfrak{k}}_\alpha$ in $\mathfrak{k}$. It follows that

$$adH(U^{\mathfrak{p}}_\alpha + U^{\mathfrak{k}}_\alpha) = -\alpha(H)(V^{\mathfrak{p}}_\alpha + V^{\mathfrak{k}}_\alpha), adH(V^{\mathfrak{p}}_\alpha + V^{\mathfrak{k}}_\alpha) = \alpha(H)(U^{\mathfrak{p}}_\alpha + U^{\mathfrak{k}}_\alpha).$$

Cartan relations (1.4) imply

$$adH(U^{\mathfrak{p}}_\alpha) = -\alpha(H)V^{\mathfrak{k}}_\alpha, adH(V^{\mathfrak{k}}_\alpha) = \alpha(H)U^{\mathfrak{p}}_\alpha,$$

$$adH(U_\alpha^{\mathfrak{k}}) = -\alpha(H)V_\alpha^{\mathfrak{p}}, adH(V_\alpha^{\mathfrak{p}}) = \alpha(H)U_\alpha^{\mathfrak{k}}$$

which, in turn, imply that both $U_\alpha^{\mathfrak{p}}$ and $V_\alpha^{\mathfrak{k}}$ are non-zero, and also imply that $U_\alpha^{\mathfrak{k}}$ and $V_\alpha^{\mathfrak{p}}$ are non-zero. Then $X_\alpha = U_\alpha^{\mathfrak{p}}$ and $Y_\alpha = V_\alpha^{\mathfrak{k}}$ satisfy

$$adH(X_\alpha) = -\alpha(H)Y_\alpha, adH(Y_\alpha) = \alpha(H)X_\alpha.$$

In the non-compact case, $Z_\alpha = X_\alpha + Y_\alpha$, $X_\alpha \in \mathfrak{p}$ and $Y_\alpha \in \mathfrak{k}$. Then $adH(Z_\alpha) = \alpha(H)Z_\alpha$, together with the Cartan conditions yield

$$adH(X_\alpha) = \alpha(H)Y_\alpha, adH(Y_\alpha) = \alpha(H)X_\alpha. \tag{2.7}$$

In either case,

$$\begin{aligned} adH([X_\alpha, Y_\alpha]) &= -[Y_\alpha, adH(X_\alpha)] + [X_\alpha, adH(Y_\alpha)] \\ &= \pm\alpha(H)[Y_\alpha, Y_\alpha] + \alpha(H)[X_\alpha, X_\alpha] = 0. \end{aligned}$$

Hence $[X_\alpha, Y_\alpha] \in \mathfrak{h}$.

There are many properties that both the compact and the non-compact spaces symmetric spaces share. In particular in both cases each root α defines a hyperplane $\{X \in \mathfrak{h} : \alpha(X) = 0\}$. The set $\cup_{\alpha\in\Delta}\{X \in \mathfrak{h} : \alpha(X) = 0\}$ is closed and nowhere dense in $\mathfrak{h}$. Therefore its complement $\mathcal{R}(\mathfrak{h})$, given by $\mathcal{R}(\mathfrak{h}) = \cap_{\alpha\in\Delta}\{X \in \mathfrak{h} : \alpha(X) \neq 0\}$, is open and dense in $\mathfrak{h}$. It is a union of finitely many connected components called *Weyl chambers*. Each Weyl chamber is defined as an equivalence class under the equivalence relation in $\mathcal{R}(\mathfrak{h})$ defined by $X \sim Y$ if and only if $\alpha(X)\alpha(Y) > 0$ for all roots $\alpha \in \Delta$. It is evident that each Weyl chamber is an open and convex subset in $\mathfrak{h}$.

Proposition 5. *An element $X \in \mathfrak{p}$ is regular in a maximal abelian algebra $\mathfrak{h}$ in $\mathfrak{p}$ if and only if $X \in \mathcal{R}(\mathfrak{h})$. That is, X is regular if and only if $\alpha(X) \neq 0$ for every root $\alpha \in \Delta$.*

Proof. The proof is almost identical in both the compact and the non-compact case. Suppose that X is regular in $\mathfrak{h}$ and suppose that $\alpha(X) = 0$ for some $\alpha \in \Delta$. Let $X_\alpha \in \mathfrak{p}$ and $Y_\alpha \in \mathfrak{k}$ be as in Proposition 4, that is

$$adH(X_\alpha) = -\alpha(H)Y_\alpha, adH(Y_\alpha) = \alpha(H)X_\alpha, H \in \mathfrak{h},$$

in the compact case, and

$$adH(X_\alpha) = \alpha(H)Y_\alpha, adH(Y_\alpha) = \alpha(H)X_\alpha, H \in \mathfrak{h},$$

in the non-compact case. If $\alpha(X) = 0$, then $adX(X_\alpha) = 0$ and therefore $X_\alpha \in \mathfrak{h}$. Hence $0 = adH(X_\alpha) = \pm\alpha(H)Y_\alpha$ which yields $Y_\alpha = 0$ since $\alpha \neq 0$, which contradicts our assumption that neither X_α nor Y_α are non-zero.

Conversely, assume that X is an element in $\mathfrak{h}$ such that $\alpha(X) \neq 0$ for any $\alpha \in \Delta$. Let $Y \in \mathfrak{p}$ be such that $[X, Y] = 0$. Then $0 = adX(Y) = \sum_{\alpha\in\Delta}\alpha(X)Y_\alpha$, where $Y = Y_0 + \sum Y_\alpha$. This relation implies that $Y_\alpha = 0$ for any $\alpha \neq 0$. Hence $Y = Y_0$, $Y_0 \in \mathfrak{g}_0 \cap \mathfrak{h}$. This shows that $Y \in \mathfrak{h}$, therefore X is regular.

Corollary 1. *The set of regular elements in $\mathfrak{p}$ is open and dense in $\mathfrak{p}$.*

The following proposition is of central importance.

Proposition 6. *Let X and X^* be regular elements in the maximal abelian algebras $\mathfrak{h}$ and $\mathfrak{h}^*$ in $\mathfrak{p}$. Consider now functions $F(h) = Kl(X^*, Ad_h(X)), h \in K$, in the non-compact case and $F(h) = -Kl(X^*, Ad_h(X)$ in the compact case. If $k \in K$ yields a critical point for the function $F(h)$, then $Ad_k(X) \in \mathfrak{h}^*$ and $Ad_k(\mathfrak{h}) = \mathfrak{h}^*$. When k yields the maximum for F then $Ad_k(X) \in C(X^*)$, and $Ad_k(C(X)) = C(X^*)$, where $C(X)$ and $C(X^*)$ denote the Weyl chambers that contain X and X^*.*

Proof. Let $\langle X, Y\rangle = \pm Kl(X, Y)$. If $U \in \mathfrak{k}$ then

$$F(ke^{tU}) = \langle X^*, Ad_k(X) + tadU(X) + \frac{t^2}{2}ad^2U(X) + \cdots\rangle.$$

When k is a critical point of F, then $\frac{d}{dt}F(ke^{tU})|_{t=0} = 0$, and when k is a maximal point then in addition $\frac{d^2}{dt^2}F(ke^{tU})|_{t=0} \leq 0$. In the first case,

$$0 = dF(k)(U) = \langle X^*, Ad_k[U, X]\rangle = -\langle [X^*, Ad_k(X)], Ad_kU\rangle = 0,$$

for any $U \in \mathfrak{k}$. It follows that $[X^*, Ad_k(X)] = 0$ because U is arbitrary and Ad_k is an isomorphism on $\mathfrak{k}$. Hence $Ad_k(X)$ belongs to the Cartan algebra that contains X^*, which is equal to $\mathfrak{h}^*$ since X^* is regular in $\mathfrak{h}^*$. If $Y \in \mathfrak{h}$ then $[Ad_k(Y), Ad_k(X)] = Ad_k([X, Y]) = 0$, therefore $Ad_k(Y) \in \mathfrak{h}^*$. Hence, $Ad_k(\mathfrak{h}) = \mathfrak{h}^*$.

Assume now that $F(k)$ is a maximal point for F. It follows that

$$\frac{d^2}{dt^2}F(ke^{tU})|_{t=0} = \langle X^*, Ad_k(ad^2U(X)\rangle \leq 0.$$

If we let $Ad_k(X) = X'$ and $Ad_k(U) = U'$ then the above can be written as

$$\langle adX^*adX'(U'), U'\rangle \leq 0, U' \in K.$$

If $T = adX^*adX'$ then T is negative semi-definite on $\mathfrak{k}$.

In the compact case T is a composition of two commuting skew-symmetric operators, hence is symmetric (relative to $\langle\, ,\, \rangle$ which is positive on $\mathfrak{k}$). In the non-compact case, T is a composition of two commuting symmetric operators, hence is symmetric again, but this time relative to a negative definite metric- since the Killing form is negative on $\mathfrak{k}$. Hence T is negative semi-definite on $\mathfrak{k}$ in the compact case, and positive semi-definite in the non-compact case. Therefore, the non-zero eigenvalues of T are positive in the non-compact case and negative in the compact case.

We will show now that $\alpha(X^*)\alpha(X') > 0$ for each $\alpha \in \Delta(\mathfrak{h}^*)$. In the compact case there are elements $X_\alpha \in \mathfrak{p}$ and $Y_\alpha \in \mathfrak{k}$ such that

$$ad(H)(X_\alpha) = -\alpha(H)Y_\alpha, adH(Y_\alpha) = \alpha(H)X_\alpha, H \in \mathfrak{h}^*,$$

for each $\alpha \in \Delta(\mathfrak{h}^*)$. Then,

$$adX^*(X_\alpha) = -\alpha(X^*)Y_\alpha, adX^*(Y_\alpha) = \alpha(X^*)X_\alpha,$$
$$adX'(X_\alpha) = -\alpha(X')Y_\alpha, adX'(Y_\alpha) = \alpha(X')X_\alpha.$$

Since X^* and X' are regular $\alpha(X^*)$ and $\alpha(X')$ are non-zero. We then have

$$T(Y_\alpha) = adX^*adX'(Y_\alpha) = adX^*\alpha(X')X_\alpha = -\alpha(X^*)\alpha(X')Y_\alpha.$$

It follows that Y_α is an eigenvector for T with $-\alpha(X^*)\alpha(X')$ the corresponding eigenvalue. Since the non-zero eigenvalues of T are negative we get $\alpha(X^*)\alpha(X') > 0$.

In the non-compact case

$$ad(H)(X_\alpha) = \alpha(H)Y_\alpha, adH(Y_\alpha) = \alpha(H)X_\alpha, H \in \mathfrak{h}^*,$$

for each $\alpha \in \Delta(\mathfrak{h}^*)$, therefore

$$T(Y_\alpha) = adX^*adX'(Y_\alpha) = adX^*\alpha(X')X_\alpha = \alpha(X^*)\alpha(X')Y_\alpha.$$

Thus $\alpha(X)\alpha(X'$ are the eigenvalues of T. Since T is positive semi-definite $\alpha(X)\alpha(X' > 0$ (neither $\alpha(X)$ nor $\alpha(X')$ can be zero because X and X' are regular.) Therefore $X' \in C(X^*)$ in both cases.

We now return to Proposition 3 with the proofs.

Proof. The first statement is obvious in view of Proposition 5, If $\mathfrak{h}$ is any Cartan algebra then take any $X \in \mathfrak{h}$ such that $\alpha(X) \neq 0$ for any $\alpha \in \Delta$.

Second statement follows from Proposition 6. To prove the last statement let P be an arbitrary element in $\mathfrak{p}$ and let X_0 be a regular element in $\mathfrak{h}$. There is an element $k \in K$ that attains the maximum for the function $F(k) = \langle P, Ad_k X_0\rangle$. Then $dF(k) = 0$ yields $[P, Ad_k X_0] = 0$. Therefore $P \in Ad_k(\mathfrak{h})$. This shows that every element $P \in \mathfrak{p}$ is contained in some maximal abelian algebra in $\mathfrak{p}$.

We are now ready to introduce another important theoretic ingredient, the Weyl group. If $\mathfrak{h}$ be any maximal abelian subalgebra in $\mathfrak{p}$ let

$$N(\mathfrak{h}) = \{h \in K : Ad_h(\mathfrak{h}) \subseteq \mathfrak{h}\}, C(\mathfrak{h}) = \{h \in K : Ad_h(X) = X, X \in \mathfrak{h}\}.$$

These groups are respectively called *the normalizer* of $\mathfrak{h}$ and *the centralizer* of $\mathfrak{h}$. Each group is a closed subgroup of K, and consequently, each group a Lie subgroup of K. Moreover, $C(\mathfrak{h})$ is normal in $N(\mathfrak{h})$. Any element U in the Lie algebra $n(\mathfrak{h})$ of $N(\mathfrak{h})$ satisfies $adU(X) \in \mathfrak{h}$ for any $X \in \mathfrak{h}$. But then $\langle [U, X], \mathfrak{h}\rangle = \langle U, [X, \mathfrak{h}]\rangle = 0$. Hence $[U, X] = 0$. Therefore, U belongs to the Lie algebra of the centralizer $C(\mathfrak{h})$. It follows that $N(\mathfrak{h})$ and $C(\mathfrak{h})$ have the same Lie algebra, which then implies that $N(\mathfrak{h})$ is an open cover of $C(\mathfrak{h})$, that is, the quotient group $N(\mathfrak{h})/C(\mathfrak{h})$ is finite. This quotient group is called *the Weyl group*.

We will follow S. Helgason and represent the elements of the Weyl group by the mappings $Ad_k|_{\mathfrak{h}}$ with $k \in N(\mathfrak{h})$ ([6]) in which case $\{Ad_k|_{\mathfrak{h}} : k \in N(\mathfrak{h})\}$ is denoted by $W(G, K)$. An interested reader can easily show that if $W_{\mathfrak{h}}(G, K)$ is the Weyl group associated with a Cartan algebra $\mathfrak{h}$ and $W_{\mathfrak{h}^*}(G, K)$ is the Weyl group associated with another Cartan algebra $\mathfrak{h}^*$ then

$$kW_{\mathfrak{h}}(G, K)k^{-1} = W_{\mathfrak{h}^*}(G, K), Ad_k(\mathfrak{h}) = \mathfrak{h}^*.$$

In that sense the Weyl group is determined by the pair (G, K) rather than a particular choice of a Cartan algebra.

Proposition 7. *If $Ad_k(C(\mathfrak{h})) = C(\mathfrak{h})$ for some $k \in K$, and some Weyl chamber $C(\mathfrak{h})$ in $\mathfrak{h}$, then $Ad_k|_{\mathfrak{h}} = Id$.*

The following lemma is useful for the proof of the proposition.

Lemma 1. *Let H be a regular element in a maximal abelian algebra $\mathfrak{h}$ in $\mathfrak{p}$. Then*

$$\{Z \in \mathfrak{g} : [Z, H] = 0\} = \mathfrak{h} + \{Q \in \mathfrak{k} : [Q, H] = 0\} = \mathfrak{h} + \{U \in \mathfrak{k} : [U, \mathfrak{h}] = 0\}.$$

Proof. If $Z = P + Q, P \in \mathfrak{p}, Q \in \mathfrak{k}$, then $[Z, H] = 0$ if and only if $[P, H] = 0$ and $[Q, H] = 0$. Therefore, $P \in \mathfrak{h}$ because H is regular. It follows that $\{Z \in \mathfrak{g} : [Z, H] = 0\} = \mathfrak{h} + \{Q \in \mathfrak{k} : [Q, H] = 0]\}$.

Now let V be an arbitrary point in $\mathfrak{h}$. Then for any $Q \in \mathfrak{k}$ such that $[Q, H] = 0$, $[[Q, V], H] = -[[H, Q], V] - [[V, H], Q] = 0$. Therefore $[Q, V] \in \mathfrak{h}$ since $[Q, V] \in \mathfrak{p}$ and H is regular. But then$\langle [Q, V], \mathfrak{h}\rangle = \langle Q, [V, \mathfrak{h}]\rangle = 0$ and hence $[Q, V] = 0$.

We now return to the proof of Proposition 7.

Proof. Since $C(\mathfrak{h})$ is open in $\mathfrak{h}$ and the set of regular elements is dense, there is a regular element X in $C(\mathfrak{h})$. Then $Ad_k(X) = X^*$ belongs to $C(\mathfrak{h})$. If $Z \in \mathfrak{h}$ then $[X^*, Ad_k Z] = [Ad_k X, Ad_k Z] = Ad_k[X, Z] = 0$ and therefore $Ad_k Z \in \mathfrak{h}$. This shows that $k \in N(\mathfrak{h})$ that is, $Ad_k|_{\mathfrak{h}} \in W(G, K)$. Since $W(G, K)$ is finite, the orbit $\{Ad_k^n(X^*), k = 0, 1, \dots\}$ is finite, and therefore there is a positive integer N such that $Ad_k^N(X^*) = X^*$. If N is the smallest such integer then let $H = \frac{1}{N-1}(X^* + Ad_k X^* + \cdots + Ad_k^{N-1} X^*)$. It follows that $Ad_k(H) = H$. Since $Ad_k(C(\mathfrak{h})) = C(\mathfrak{h})$, $Ad_k^n X^* \in C(\mathfrak{h})$, and since $C(\mathfrak{h})$ is convex, $H \in C(\mathfrak{h})$.

The above implies that k belongs to the centralizer of H. The Lie algebra of the centralizer in K is given by $\{U \in \mathfrak{k} : [U, H] = 0\}$. But this Lie algebra coincides with $\{U \in \mathfrak{k} : [U, V] = 0, V \in \mathfrak{h}\}$ as shown in the Lemma above. Since $Ad_k(H) = H$, $ke^{tH}k^{-1} = e^{tH}$. Therefore k belongs to the centralizer of the one parameter group $\{e^{tH}, t \in R\}$. Let T be the closure of $\{e^{tH}, t \in R\}$. Then, T is a connected abelian subgroup in G, i.e., T is a torus. Its centralizer in G is the maximal torus that contains T. Every maximal torus is connected, and consequently is generated by the exponentials in its Lie algebra. The Lie algebra of this centralizer is given by $\mathcal{L} = \{Z \in \mathfrak{g} : [Z, H] = 0\}$, which is equal to $\mathfrak{h} + \{U \in \mathfrak{k} : [U, \mathfrak{h}] = 0\}$ by the lemma above.

We now have $Ad_{e^{tU}} X = X$ for each $U \in \mathcal{L}$ and each $X \in \mathfrak{h}$. Since $k = \prod_{i=1}^m e^{U_i}$ for some choice of $U_1, \dots, U_m$ in $\mathcal{L}$, $Ad_k|_{\mathfrak{h}} = Id$, and therefore $X^* = X$.

Propositions 6 and 7 can be summarized as follows:

Proposition 8. *Let $C(\mathfrak{h})$ be a Weyl chamber in $\mathfrak{h}$. Then $\{Ad_k(C(\mathfrak{h})) : k \in W(G, K)\}$ acts simply and transitively on the set of Weyl chambers in $\mathfrak{h}$. Here acting simply means that if some $k \in W(G, K)$ takes a Weyl chamber $C(\mathfrak{h})$ onto itself, then $k = e$.*

Corollary 2. *If X_0 is any regular element in $\mathfrak{p}$ and if $C(\mathfrak{h})$ is a Weyl chamber associated with any maximal abelian subalgebra in $\mathfrak{p}$ then there is a unique $k \in K$ such that $Ad_k(X_0) \in C(\mathfrak{h})$.*

The Weyl group could be also defined in terms of the orthogonal reflections in $\mathfrak{h}$ around the hyperplane $\{X \in \mathfrak{p} : \alpha(X) = 0\}, \alpha \in \Delta$. The reader can readily verify that this reflection is given by $s_\alpha(H) = H - 2\frac{\alpha(H)}{\alpha(A)} A$ where $A \in \mathfrak{h}$ is the unit vector such that $\alpha(H) = \langle A, H\rangle, H \in \mathfrak{h}$. The following proposition is basic.

Proposition 9. *There exists $k \in N(\mathfrak{h})$ such that $Ad_k|_{\mathfrak{h}} = s_\alpha$.*

Proof. Let X_α and Y_α be non-zero vectors in $\mathfrak{g}$ as in Proposition 4 such that

$$adH(X_\alpha) = -\alpha(H)Y_\alpha, adH(Y_\alpha) = \alpha(H)(X_\alpha)$$

in the compact case, and

$$adH(X_\alpha) = \alpha(H)Y_\alpha, adH(Y_\alpha) = \alpha(H)(X_\alpha)$$

in the non-compact case. We have already shown $[X_\alpha, Y_\alpha] \in \mathfrak{h}$. Since

$$\langle H, [Y_\alpha, X_\alpha]\rangle = \langle [H, Y_\alpha], X_\alpha\rangle = \alpha(H)\langle X_\alpha, X_\alpha\rangle,$$

X_α could be rescaled so that $\langle H, [Y_\alpha, X_\alpha]\rangle = \alpha(H)$.

Let $A_\alpha \in \mathfrak{h}$ be such that $\alpha(H) = \langle A_\alpha, H\rangle, H \in \mathfrak{h}$. Then $[Y_\alpha, X_\alpha] = A_\alpha$. We now have

$$adA_\alpha(X_\alpha) = -\alpha(A_\alpha)Y_\alpha,\ adA_\alpha(Y_\alpha) = \alpha(A_\alpha)X_\alpha.$$

Therefore

$$adY_\alpha(A_\alpha) = -\alpha(A_\alpha)X_\alpha, \text{ and } ad^2Y_\alpha(A_\alpha) = -\alpha(A_\alpha)A_\alpha. \tag{2.8}$$

Hence,

$$\begin{gathered} Ad_{e^{tY_\alpha}}(A_\alpha) = e^{tadY_\alpha}A_\alpha = \\ \sum_{n=0}^{\infty} \tfrac{1}{2n!}t^{2n}ad^{2n}Y_\alpha(A_\alpha) + \sum_{n=0}^{\infty} \tfrac{1}{2n+1!}t^{2n+1}ad^{2n+1}Y_\alpha(A_\alpha) = \\ \sum_{n=0}^{\infty} \tfrac{t^{2n}}{2n!}(-\alpha(A_\alpha)^n A_\alpha + \sum_{n=0}^{\infty} \tfrac{t^{2n+1}}{2n+1}(-\alpha(A_\alpha)^{2n-1}X_\alpha = \\ \sum_{n=0}^{\infty}(-1)^n \tfrac{\theta^{2n}}{2n!}A_\alpha + \sum_{n=0}^{\infty}(-1)^n \tfrac{\theta^{2n+1}}{2n+1!}X_\alpha = \\ \cos t\theta A_\alpha + \sin t\theta X_\alpha, \end{gathered}$$

where $\theta = \sqrt{\alpha(A_\alpha)}$. When $t\theta = \pi$ then $Ad_{e^{tY}}(A_\alpha) = -A_\alpha$.

Moreover, if $H \in \mathfrak{h}$ is perpendicular to A_α then $\alpha(H) = 0$ and therefore, $adY(H) = \alpha(H)X = 0$. Hence $Ad_{e^{tY}}(H) = H$, and $Ad_{e^{tY}}|_{\mathfrak{h}} = s_\alpha$.

Proposition 10. *The Weyl group $W(G, K)$ is equal to the group generated by the reflections $Ad_k|_{\mathfrak{h}} = s_\alpha, \alpha \in \Delta$.*

Proof. Let W_s be the group generated by s_α, $\alpha \in \Delta$. Then W_s is a subgroup of $W(G, K)$. We will show that for any Ad_k in $W(G, K)$ there exists an element Ad_h in W_s such that $Ad_k(X) = Ad_h(X)$ for any $X \in \mathfrak{h}$. It suffices to show the equality on regular elements in $\mathfrak{h}$.

If X is a regular element in $\mathfrak{h}$, then let C^* be the Weyl chamber in $\mathfrak{h}^* = Ad_k(\mathfrak{h})$ that contains $X^* = Ad_k(X)$. Let Ad_{h^*} be the element of W_s that minimizes $\|X^* - Ad_h(X)\|$ over W_s. Then the line segment from $Ad_{h^*}(X)$ to X^* cannot cross any hypersurface $\alpha = 0$. Hence $\alpha(X^*)$ and $\alpha(Y)$ have the same signature at any point Y on the line segment from X^* to $Ad_{h^*}(X)$. It then follows that $Ad_{h^*}(X)$ and X^* belong to the same Weyl chamber. Then $Ad_k(X) = Ad_{h^*}(X)$ by the previous proposition.

2.2. Weyl group and controllability

Let $\mathfrak{h}$ be any maximal abelian algebra in $\mathfrak{g}$ contained in $\mathfrak{p}$, and let $\alpha_1, \ldots, \alpha_n$ be any basis in Δ. Then let $A_1, \ldots, A_n$ be the corresponding vectors in $\mathfrak{h}$ defined by $\langle X, A_i \rangle = \alpha_i(X), X \in \mathfrak{h}$. If X is a an element in $\mathfrak{h}$ that is orthogonal to each A_i, then $\alpha_i(X) = 0$ for each $\alpha_i \in \Delta$. That means that $ad(X) = 0$. Therefore $X = 0$, since the centre in $\mathfrak{g}$ consists of 0 alone. Hence $A_1, \ldots, A_n$ form a basis in $\mathfrak{h}$. With these observations at our disposal we now return to the convexified horizontal control system

$$\frac{dg}{dt} = \sum_{i=1}^{k} \lambda_i(t) g(t) Ad_{h_i(t)}(X_0), \lambda_i(t) \geq 0, \sum_{i=1}^{k} \lambda_i(t) = 1, \tag{2.9}$$

with $X_0 \in \mathfrak{p}$ regular, controlled by the coefficients $\lambda_1, \ldots, \lambda_k$ and the curves $h_1(t), i = 1, \ldots, k$ in K. There will be no loss in generality if the curves $h_i(t)$ are restricted to the solutions of $\frac{dh}{dt} = U(t)h(t)$ with $U(t)$ transversal to the Lie algebra $\{V \in \mathfrak{k} : [V, X_0] = 0\}$.

Proposition 11. *The convex hull of* $\{Ad_h(X_0), h \in N(\mathfrak{h})\}$ *contains an open neighbourhood of the origin in* $\mathfrak{h}$.

Proof. Let $O(X_0) = \{Ad_{h_i} X_0, i = 0, 1, \ldots, m\}$ denote the orbit of $W((G, K)$ through X_0. Assume that $Ad_{h_0}(X_0) = X_0$ and that $Ad_{h_i}(X_0) = s_{\alpha_i}, i = 1, \ldots, n$. We know that O acts simply and transitively on the Weyl chambers in $\mathfrak{h}$. Let

$$X = \frac{1}{m} \sum_{i=0}^{m} Ad_{h_i} X_0.$$

It follows that X is in the convex hull of the orbit $\{Ad_h X_0, h \in N(\mathfrak{h})\}$. Since $Ad_{h_j} Ad_{h_i} X_0 = Ad_{h_j h_i} X_o = Ad_{h_k} X_0$, where $k \in K$, each Ad_{h_j} permutes the elements in $O(X_0)$, which in turn implies that $Ad_{h_j} X = X$ for each $j = 1, \ldots, m$. Therefore, $X = 0$. Let now

$$\sigma(t) = \sum_{i=0}^{n} (\frac{1}{m} + t\varepsilon_i) s_{\alpha_i}(X_0) + \frac{1}{m} \sum_{i=n+1}^{m} Ad_{h_i}(X_0),$$

where $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_n$ are arbitrary numbers such that $\sum_{i=0}^{n} \varepsilon_i = 0$. Let

$$\lambda_i(t) = \frac{1}{m} + t\varepsilon_i, i = 0, \ldots, n, \lambda_i = \frac{1}{m}, i = n+1, \ldots, m.$$

Then, $\sum_{i=0}^{m} \lambda_i(t) = 1$, and for sufficiently small t, $\lambda_i > 0, i = 0, \ldots, m$. It follows that $\sigma(t)$ is contained in the convex hull of the Weyl orbit through X_0 for small t and satisfies $\sigma(0) = 0$. Then $\frac{d\sigma}{dt}(0) = -\sum_{i=1}^{n} \varepsilon_i \frac{\alpha_i(X_0)}{\alpha(A_i)} A_i$ and therefore the mapping $F(\lambda_0(t), \ldots, \lambda_m(t)) = \sum_{i=1}^{m} \lambda_i(t) Ad_{h_i} X_0$ is open at $\lambda_1 = \lambda_2 = \cdots = \lambda_m = \frac{1}{m}$.

Corollary 3. *The convexified horizontal system* (2.9) *admits a stationary solution* $g(t) = g_0$.

Proposition 12. *The convexified horizontal system is controllable.*

Proof. We will first show that the Lie algebra $\mathcal{L}$ generated by $\{Ad_h X_0 : h \in K\}$ is equal to $\mathfrak{g}$. Let V denote the vector space spanned by $\{Ad_h(X_0), h \in K\}$ and let $\mathcal{L}$ be the Lie algebra generated by V. If $U_1, \ldots, U_j$ are arbitrary elements in $\mathfrak{k}$ then $Ad_{h_1(t_1)\cdots h_j(t_j)}(X_0)$ is in V where $h_i(t_i) = e^{t_i U_i}$. Since V is a vector space, $\frac{\partial}{\partial t_i} Ad_{h_1(t_1)\cdots h_j(t_j)}(X_0)$ is in V. Therefore,

$$\frac{\partial}{\partial t_j} Ad_{h_1(t_1)\cdots h_j(t_j)}(X_0)_{t_j=0} = Ad_{h_1(t_1)\cdots h_{j-1}(t_{j-1})}(ad(U_j)(X_0)) \in V.$$

Further differentiations yield $ad(U_1 \circ ad(U_2) \circ \cdots \mathrm{ad}(U_j)(X_0) \in V$. This can be also written as $ad^j \mathfrak{k}(X_0) \subset V$.

Let now $\hat{V}$ be the vector space spanned by $\bigcup_{j=0}^{\infty} ad^j \mathfrak{k}(X_0)$. It follows that $\hat{V} \subseteq V$. Let now $\hat{V}^{\perp}$ denote its orthogonal complement in $\mathfrak{p}$. Both $\hat{V}$ and $\hat{V}^{\perp}$ are $ad(\mathfrak{k})$ invariant. If $Z \in \hat{V}$, $W \in \hat{V}^{\perp}$ and $Y \in \mathfrak{k}$, then

$$\langle Y, [Z, W] \rangle = \langle [Y, Z], W \rangle = 0.$$

Since Y is arbitrary $[Z, W] = 0$. Therefore $[\hat{V}, \hat{V}^{\perp}] = 0$, and hence $\hat{V} + [\hat{V}, \hat{V}]$ is an ideal in $\mathfrak{g}$. Let us now use the fundamental decomposition

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \cdots \mathfrak{g}_m, \mathfrak{g}_i = \mathfrak{p}_i + [\mathfrak{p}_i, \mathfrak{p}_i], \mathfrak{p} = \mathfrak{p}_1 \oplus \cdots \oplus \mathfrak{p}_m$$

defined in (2.1). It follows that the projection of $\hat{V} + [\hat{V}, \hat{V}]$ on each simple factor is equal to $\mathfrak{g}_i$ (since $X_0 \in \hat{V}$, and the projection of X_0 on each factor $\mathfrak{g}_i$ is non-zero). So $\hat{V} + [\hat{V}, \hat{V}] = \mathfrak{g}$. But then $\hat{V} + [\hat{V}, \hat{V}] \subseteq \mathcal{L}$ yields $\mathcal{L} = \mathfrak{g}$. Since $\hat{V} + [\hat{V}, \hat{V}] \subseteq V + [V, V] = \mathfrak{g}$, $V = \hat{V}$ and $V = \mathfrak{p}$.

To prove controllability it would suffice to show that the affine cone $\{\sum_{i=1}^{k} \lambda_i Ad_{h_i}(X_0), \lambda_i \geq 0, h_i \in K, i = 1, \ldots, k\}$ is equal to V which by the above is equal to $\mathfrak{p}$. Let P be an arbitrary point in $\mathfrak{p}$. Then, P belongs to some maximal abelian algebra $\mathfrak{h}$. By the preceding proposition the convex hull of $\{Ad_h X_0 : h \in K\}$ covers a neigborhood of the origin in $\mathfrak{h}$. If $\varepsilon > 0$ is any scalar such that εP is in this neighborhood, then $-\varepsilon P$ is also in this neghborhood, and hence is reachable by the convex hull of $\{Ad_h X_0 : h \in K\}$. But then $-P = \frac{1}{\varepsilon}(-\varepsilon P)$ is in the above affine cone.

The preceding results show that the convex cone spanned by $Ad_h(X_0)$ is a neighbourhood of the origin in $\mathfrak{p}$. It then follows that the positive cone $\sum \lambda_i Ad_{h_i}(X_0), \lambda_i \geq 0$, is equal to $\mathfrak{p}$. This implies that any time optimal trajectory of the compactified horizontal system is generated by a control on the boundary of the convex cone defined by $\{Ad_h(X_0), h \in K\}$. For if $g(t)$ is a trajectory generated by a control $U(t) = \sum_{i=1}^{k} \lambda_i(t) Ad_{h_i(t)}(X_0)$ in the interior of the convex set spanned by $Ad_h(X_0)$, then $\rho U(t)$ is in the same interior for some $\rho > 1$. But then $g(t)$ reparametrized by $s = \rho t$ steers e to $g(T)$ in $s = \frac{T}{\rho}$ units of time violating the time optimality of $g(t)$.

The time-optimal problem for the convexified system is related to the sub-Riemannian problem of finding the shortest length of a horizontal curve that connects two given points in G. In fact any horizontal curve $g(t)$ is a solution of $\frac{dg}{dt} = g(t)U(t)$ with $U(t) = \mathrm{Ad}_{h(t)} X_0$ and inherits the notion of length from G given by $\int_0^T \sqrt{\langle U(t), U(t) \rangle} dt$, where $\langle\, , \,\rangle$ denotes a suitable scalar multiple of the Killing form. Since $U(t) = Ad_{h(t)}(X_0)$ satisfies $\langle U(t), U(t) \rangle = ||X_0||^2 = 1$ when X_0 is a unit vector, the length of $g(t)$ in the interval $[0, T]$ is equal to the time it takes to reach $g(T)$ from $g(0)$. Therefore, the shortest time to reach a point g_1 from g_0 is equal to the minimum length of the horizontal curve to reach g_1 from g_0. As we showed above, the horizontal system is controllable, therefore any two points in G can be connected by a horizontal curve. But then any two points in G can be connected by a horizontal curve of minimal length by a suitable compactness argument.

3. Necessary conditions of optimality

3.1. Generalities-left-invariant Hamiltonians

We will now use the maximum principle to obtain the necessary conditions of optimality on the cotangent bundle T^*G. We recall that each optimal solution is the projection of an integral curve in T^*G of the Hamiltonian vector generated by a suitable Hamiltonian obtained from the maximum principle. To preserve the left-invariant symmetries, we will regard the cotangent bundle T^*G as the product $G \times \mathfrak{g}^*$ via the left-translations. In this formalism tangent vectors $v \in T_gG$ are identified with pairs $(g, X) \in G \times \mathfrak{g}$ via the relation $v = L_{g_*}X$, where L_{g_*} denotes the tangent map associated with the left translation $L_g(h) = gh$. Similarly, points $\xi \in T_g^*G$ are identified with pairs $(g, \ell) \in G \times \mathfrak{g}^*$ via $\xi = \ell \cdot L_{g_*}^{-1}$. If the optimal problem was defined over a right-invariant system, then the tangent bundle would be trivialized by the right translations, in which case the ensuing formalism would remain the same as in the left-invariant setting.

Then, $T(T^*G)$, the tangent bundle of the cotangent bundle T^*G, is naturally identified with $(G \times \mathfrak{g}^*) \times (\mathfrak{g} \times \mathfrak{g}^*)$, with the understanding that an element $((g, \ell), (A, a)) \in (G \times \mathfrak{g}^*) \times (\mathfrak{g} \times \mathfrak{g}^*)$ stands for the tangent vector (A, a) at the base point (g, ℓ).

We will make use of the fact that $G \times \mathfrak{g}^*$ is a Lie group in its own right since $\mathfrak{g}^*$, as a vector space, is an abelian Lie group. Then left-invariant vector fields in $G \times \mathfrak{g}^*$ are the left-translations of the pairs (A, a) by the elements (g, ℓ) in $G \times \mathfrak{g}^*$. The corresponding one-parameter groups of diffeomorphisms are given by $(g \exp(tA), \ell + ta), t \in R$. In terms of these vector fields the canonical symplectic form on T^*G is given by

$$\omega_{(g,\ell)}(V_1, V_2) = a_2(A_1) - a_1(A_2) - \ell([A_1, A_2]) \tag{3.1}$$

for any $V_1 = (gA_1, a_1)$ and $V_2 = (gA_2, a_2)$. ([7]).

The above differential form is invariant under the left-translations in $G \times \mathfrak{g}^*$, and is particularly revealing for Hamiltonian vector fields generated by the left-invariant functions on $G \times \mathfrak{g}^*$. A function H on $G \times \mathfrak{g}^*$ is said to be *left-invariant* if $H(gh, \ell) = H(g, \ell)$ for all $g, h \in G$ and all $\ell \in \mathfrak{g}^*$. It follows that the left-invariant functions are in exact correspondence with functions on $\mathfrak{g}^*$. Each left-invariant vector field $X(g) = (L_g)_*A$, $A \in \mathfrak{g}$, lifts to a linear function $\ell \to \ell(A)$ on $\mathfrak{g}^*$ because

$$h_X(\xi) = \xi(X(g)) = \ell \circ L_{g_*}^{-1} \circ (L_g)_*(A) = \ell(A), \xi \in T_g^*G.$$

Any function H on $\mathfrak{g}^*$ generates a Hamiltonian vector field $\vec{H}$ on $G \times \mathfrak{g}^*$ whose integral curves are the solutions of

$$\frac{dg}{dt}(t) = g(t)dH_{\ell(t)}, \quad \frac{d\ell}{dt}(t) = -\mathrm{ad}^* dH_{\ell(t)}(\ell(t)). \tag{3.2}$$

For when H is a function on $\mathfrak{g}^*$, then its differential at a point ℓ is a linear function on $\mathfrak{g}^*$, hence is an element of $\mathfrak{g}$ because $\mathfrak{g}^*$ is a finite dimensional vector space. If $\vec{H}_{(g,\ell)} = (A(g, \ell), a(g, \ell))$ for some vectors $A(g, \ell) \in \mathfrak{g}$ and $a(g, \ell) \in \mathfrak{g}^*$, then

$$b(dH_\ell) = b(A) - a(B) - \ell[A, B],$$

must hold for any tangent vector (B, b) at (g, ℓ). This implies that $A(g, \ell) = dH_\ell$, and $a = -\mathrm{ad}^* dH_\ell(\ell)$, where $(\mathrm{ad}^* A)(\ell)(B) = \ell[A, B]$ for all $B \in \mathfrak{g}$. This argument validates equations (3.2).

The dual space $\mathfrak{g}^*$ is a Poisson space with its Poisson structure $\{f, h\}(\ell) = \ell([dh, df])$ inherited from the symplectic form (3.1). Recall that a manifold M together with a bilinear, skew-symmetric form

$$\{\,,\,\} : C^\infty(M) \times C^\infty(M) \to C^\infty(M)$$

that satisfies

$$\{fg, h\} = f\{g, h\} + g\{f, h\}, \text{(Leibniz's rule), and}$$
$$\{f, \{g, h\}\} + \{h, \{f, g\}\} + \{g, \{h, f\}\} = 0, \text{ (Jacobi's identity),}$$

for all functions f, g, h on M, is called a *Poisson manifold*.

Every symplectic manifold is a Poisson manifold with the Poisson bracket defined by $\{f, g\}(p) = \omega_p(\vec{f}(p), \vec{g}(p))$, $p \in M$. However, a Poisson manifold need not be symplectic, because it may happen that the Poisson bracket is degenerate at some points of M. Nevertheless, each function f on M induces a Poisson vector field $\vec{f}$ through the formula $\vec{f}(g) = \{f, g\}$. It is known that every Poisson manifold is foliated by the orbits of its family of Poisson vector fields, and that each orbit is a symplectic submanifold of M with its symplectic form $\omega_p(\vec{f}, \vec{h}) = \{f, h\}(p)$ (this foliation is known as a *the symplectic foliation of M* ([7])).

It follows that each function H on $\mathfrak{g}^*$ defines a Poisson vector field $\vec{H}$ on $\mathfrak{g}^*$ through the formula $\vec{H}(f)(\ell) = \{H, f\}(\ell) = \ell([dH, df])$. The integral curves of $\vec{H}$ are the solutions of

$$\frac{d\ell}{dt}(t) = -\mathrm{ad}^* dH_{\ell(t)}(\ell(t)) \tag{3.3}$$

That is, each function H on $\mathfrak{g}^*$ may be considered both as a Hamiltonian on T^*G, as well as a function on the Poisson space $\mathfrak{g}^*$; the Poisson equations of the associated Poisson field are the projections of the Hamiltonian equations (3.2) on $\mathfrak{g}^*$.

Solutions of equation (3.3) are intimately linked with the coadjoint orbits of G. We recall that the coadjoint orbit of G through a point $\ell \in \mathfrak{g}^*$ is given by $\mathrm{Ad}^*_g(\ell) = \{\ell \circ \mathrm{Ad}_{g^{-1}}, \quad g \in G\}$.

The following proposition is a paraphrase of A.A. Kirillov's fundamental contributions to the Poisson structure of $\mathfrak{g}^*$ ([12]).

Proposition 13. *Let $\mathcal{F}$ denote the family of Poisson vector fields on $\mathfrak{g}^*$ and let $M = O_{\mathcal{F}}(\ell_0)$ denote the orbit of $\mathcal{F}$ through a point $\ell_0 \in \mathfrak{g}^*$. Then M is equal to the connected component of the coadjoint orbit of G that contains ℓ_0. Consequently, each coadjoint orbit is a symplectic submanifold of $\mathfrak{g}^*$.*

The fact that the Poisson equations evolve on coadjoint orbits implies useful reductions in the theory of Hamiltonian systems with symmetries. Our main results will make use of this fact.

On semi-simple Lie groups the Killing form, or any scalar multiple of it $\langle\,,\,\rangle$ is non-degenerate, and can be used to identify linear functions ℓ on $\mathfrak{g}$ with points $L \in \mathfrak{g}$ via the formula $\langle L, X\rangle = \ell(X)$, $X \in \mathfrak{g}$. Then, Poisson equation (3.3) can be expressed dually on $\mathfrak{g}$ as

$$\frac{dL}{dt} = [dH, L]. \tag{3.4}$$

The argument is simple:

$$\langle\frac{dL}{dt}, X\rangle = \frac{d\ell}{dt}(X) = -\ell([dH, X]) = \langle L, [X, dH]\rangle = \langle[dH, L], X\rangle.$$

Since X is arbitrary, equation (3.4) follows.

With the aid of Cartan's conditions (1.4) equation (3.4) can be written as

$$\frac{dL_{\mathfrak{k}}}{dt} = [dH_{\mathfrak{k}}, L_{\mathfrak{k}}] + [A, L_{\mathfrak{p}}], \frac{dL_{\mathfrak{p}}}{dt} = [dH_{\mathfrak{k}}, L_{\mathfrak{p}}] + [A, L_{\mathfrak{k}}] \tag{3.5}$$

where $dH_{\mathfrak{p}}$, $dH_{\mathfrak{k}}$, $L_{\mathfrak{p}}$ and $L_{\mathfrak{k}}$ denote the projections of dH and L on the factors $\mathfrak{p}$ and $\mathfrak{k}$.

Under the above identification coadjoint orbits are identified with the adjoint orbits $\mathcal{O}(L_0) = \{gL_0g^{-1} : g \in G\}$, and the Poisson vector fields $\vec{f_X}(\ell) = -\mathrm{ad}^*X(\ell)$ are identified with vector fields $\vec{X}(L) = [X, L]$. Each vector field $[X, L]$ is tangent to $\mathcal{O}(L_0)$ at L, and $\omega_L([X, L], [Y, L]) = \langle L, [Y, X]\rangle$, X, Y in $\mathfrak{g}$ is the symplectic form on each orbit $\mathcal{O}(L_0)$.

3.2. *Time-optimal extremals*

Let us now turn to the extremal equations associated with the time-optimal problem for the convexified horizontal system (1.7). The Hamiltonian lift is given by

$$H_0(\lambda_0, \ell) = -\lambda_0 + \sum_{i=1}^{m} \lambda_i(t)\ell(Ad_{h_i(t)}X_0), \ell \in \mathfrak{g}^*, \lambda_0 = 0, 1.$$

Suppose now that $\mathbf{g(t)}$ is a time-optimal curve generated by the controls $\lambda_i(t), \mathbf{h}_i(t), i = 1, \dots, k$. According to the maximum principle $\mathbf{g}(t)$ is the projection of an extremal curve $(\lambda_{\mathbf{0}}, \ell(t)) \in \mathbb{R} \times \mathfrak{g}^*$, $\ell(t) \neq 0$ when $\lambda_{\mathbf{0}} = 0$, that satisfies $H_0(\ell(t)) = 0$ and is further subject to:

$$-\lambda_0 + \sum_{i=1}^{m} \lambda_i(t)\ell(t)(Ad_{\mathbf{h}_i(t)}(X_0)) \geq -\lambda_0 + \sum_{i=1}^{m} \mu_i(t)\ell(t)(Ad_{h_i(t)}(X_0)) \tag{3.6}$$

for any $\mu_i(t) \geq 0$, $\sum_{i=1}^{k} \mu_i(t) = 1$, and any $h_i(t) \in K$.

The extremal curve $\ell(t)$ is called *abnormal* when $\lambda_0 = 0$. In such a case, $H(\ell(t)) = \sum_{i=1}^{m} \lambda_i(t)\ell(Ad_{\mathbf{h}_i(t)}X_0) = 0$. In the remaining case, $\lambda_0 = 1$, $H(\ell(t)) = 1$, and $\ell(t)$ is called a normal extremal. In either case,

$$\frac{d\ell}{dt} = -ad^*(\sum_{i=1}^{k} \lambda_i(t)Ad_{\mathbf{h}_i(t)}X_0)(\ell(t), \tag{3.7}$$

or, dually,

$$\frac{dL}{dt} = [\sum_{i=1}^{k} \lambda_i(t)Ad_{\mathbf{h}_i(t)}X_0, L(t)]. \tag{3.8}$$

When the terminal point is replaced by a terminal manifold S then a time-optimal trajectory must additionally satisfy the transversality condition $\ell(T)(V) = 0$ for all tangent vectors V in $T_{g(T)}S$. In particular, when $S = gK$, and when the tangent space T_gK is represented by $T_gK = g \times \mathfrak{k}$, then the transversality condition becomes $\ell(T)V = 0$ for all $V \in \mathfrak{k}$.

We will find it more convenient to work in $\mathfrak{g}$ rather than $\mathfrak{g}^*$. So, if L in $\mathfrak{g}$ corresponds to ℓ in $\mathfrak{g}^*$, then $L = L_{\mathfrak{p}} + L_{\mathfrak{k}}$ where $L_{\mathfrak{p}} \in \mathfrak{p}$ and $L_{\mathfrak{k}} \in \mathfrak{k}$.

Proposition 14. *Suppose that a time optimal control $X(t) = \sum_{i=1}^{k} \lambda_i(t) Ad_{\mathbf{h_i}(t)} X_0$ is the projection of an extremal curve $L(t)$. If $L(t)$ is abnormal, then $L_{\mathfrak{p}}(t) = 0$ and $L_{\mathfrak{k}}(t)$ is constant. In particular, the stationary solution $X(t) = 0$ is the projection of an abnormal extremal curve.*

If $L(t)$ is a normal extremal curve then $X(t) = Ad_{h(t)} X_0$ for some curve $h(t)$ in K.

Proof. If $L(t)$ is abnormal then

$$0 = \langle L_{\mathfrak{p}}(t), X(t) \rangle \geq \langle L_{\mathfrak{p}}(t), \sum_{i=1}^{k} \mu_i(t) Ad_{h_i(t)} X_0 \rangle$$

for arbitrary controls $\sum_{i=1}^{k} \mu_i(t) = 1$, and $h_1(t), \ldots, h_k(t)$ in K. This can hold only when $L_{\mathfrak{p}}(t) = 0$ (due to Proposition 11). But then equations (3.7) become

$$0 = [X(t), L_{\mathfrak{k}}], \frac{dL_{\mathfrak{k}}}{dt} = [X(t), L_{\mathfrak{p}}(t)] = 0.$$

Evidently these equations hold when $X(t) = 0$. So the stationary solution is the projection of an abnormal extremal.

In the normal case

$$H(L(t)) = \langle L_{\mathfrak{p}}(t), \sum_{i=1}^{k} \lambda_i(t) Ad_{h_i(t)}(X_0) \rangle = \sum_{i=1}^{k} \lambda_i(t) \langle L_{\mathfrak{p}}(t) Ad_{h_i(t)}(X_0) \rangle = 1.$$

So $L_{\mathfrak{p}}(t) \neq 0$. Let $\mathbf{h}(t) \in \{h_1(t), \ldots, h_k(t)\}$ corresponds to the maximal value of $\langle L_{\mathfrak{p}}(t), Ad_{h_i(t)}(X_0) \rangle$, $i = 1, \ldots, k$. Then,

$$\langle L_{\mathfrak{p}}(t), Ad_{\mathbf{h}(t)}(X_0) \rangle \geq \langle L_{\mathfrak{p}}(t), \sum_{i=1}^{k} \lambda_i(t) Ad_{h_i(t)(}(X_0) \rangle \geq \langle L_{\mathfrak{p}}(t), Ad_{\mathbf{h}(t)}(X_0) \rangle$$

can hold only if $X(t) = Ad_{\mathbf{h}(t)}(X_0)$.

It follows that the normal extremals are the solutions of the following system of equations:

$$\frac{dg}{dt} = g(t) Ad_{\mathbf{h(t)}}(X_0), \frac{dL_{\mathfrak{p}}}{dt} = [Ad_{\mathbf{h(t)}}(X_0) \rangle, L_{\mathfrak{k}}(t)], \frac{dL_{\mathfrak{k}}}{dt} = [Ad_{\mathbf{h(t)}}(X_0), L_{\mathfrak{p}}(t)]. \tag{3.9}$$

subject to the inequality

$$1 = \langle L_{\mathfrak{p}}(t), Ad_{\mathbf{h}(t)}(X_0) \rangle \geq \langle L_{\mathfrak{p}}(t), Ad_{h(t)}(X_0) \rangle, h(t) \in K.$$

3.3. Time-optimal solutions

Let us first note that there is no loss in generality in assuming that $||L_{\mathfrak{p}}(t)|| = 1$ for the following reasons: since $\mathbf{h(t)}$ is a critical point of H, $[Ad_{\mathbf{h}(t)}(X_0), L_{\mathfrak{p}}(t)] = 0$. Then,

$$2\tfrac{d}{dt}||L_{\mathfrak{p}}(t)|| = 2\langle L_{\mathfrak{p}}(t), \tfrac{dL_{\mathfrak{p}}}{dt} \rangle =$$
$$2\langle L_{\mathfrak{p}}(t), [Ad_{\mathbf{h}(t)}(X_0)], L_{\mathfrak{k}}] \rangle = -\langle [Ad_{\mathbf{h}(t)}(X_0), L_{\mathfrak{p}}(t)], L_{\mathfrak{k}} \rangle = 0.$$

Therefore $||L_{\mathfrak{p}}(t)||$ is constant. Hence the extremal equations are unaltered if $L_{\mathfrak{p}}$ is replaced by $\frac{1}{||L_{\mathfrak{p}}||} L_{\mathfrak{p}}$ and $L_{\mathfrak{k}}$ is replaced by $\frac{1}{||L_{\mathfrak{p}}||} L_{\mathfrak{k}}$.

Proposition 15. *Suppose that $(L_{\mathfrak{p}}(t), L_{\mathfrak{k}}(t))$ is a normal extremal curve generated by $\mathbf{h}(t)$ with $||L_{\mathfrak{p}}(t)|| = 1$. Then, $L_{\mathfrak{p}}(t) = Ad_{\mathbf{h}(t)}X_0$ and $L_{\mathfrak{k}}(t)$ is constant.*

Proof. According to the Cauchy-Schwarz inequality, $\langle X, Y\rangle \leq 1$ for any unit vectors X and Y in a finite dimensional Euclidean vector space, with $\langle X, Y\rangle = 1$ only when $X = Y$. In our case, $||Ad_{\mathbf{h}}X_0|| = 1$ and $||L_{\mathfrak{p}}|| = 1$, hence $\langle L_{\mathfrak{p}}, Ad_{\mathbf{h}}h(X_0)\rangle = 1$ occurs only when $L_{\mathfrak{p}} = Ad_{\mathbf{h}}(X_0)$. But then $\frac{dL_{\mathfrak{k}}}{dt} = [Ad_{h(t)}(X_0), L_{\mathfrak{p}}(t)] = 0$, and $L_{\mathfrak{k}}$ is constant.

Proposition 16. *The normal extremal curves project onto*

$$g(t) = g_0 e^{t(L_{\mathfrak{p}}(0)+L_{\mathfrak{k}})}e^{-tL_{\mathfrak{k}}}, ||L_{\mathfrak{p}}|| = 1. \tag{3.10}$$

The solutions that satisfy the transversality condition $L_{\mathfrak{k}} = 0$ are given by $g(t) = g_0e^{tP}$ for some $P \in \mathfrak{p}$ such that $||P|| = 1$.

Proof. Since $Ad_{\mathbf{h}(t)}X_0 = L_{\mathfrak{p}}(t)$, $L_{\mathfrak{p}}(t)$ is a solution of $\frac{dL_{\mathfrak{p}}}{dt} = [L_{\mathfrak{p}}(t), L_{\mathfrak{k}}]$. Since $L_{\mathfrak{k}}$ is constant, $L_{\mathfrak{p}}(t) = Ad_{e^{tL_{\mathfrak{k}}}}L_{\mathfrak{p}}(0)$. Then $\tilde{g}(t) = g(t)e^{tL_{\mathfrak{k}}}$ satisfies $\frac{d\tilde{g}}{dt} = \tilde{g}(t)(L_{\mathfrak{p}}(0) + L_{\mathfrak{k}})$, from which (3.10) easily follows. Since $L_{\mathfrak{k}}$ is constant, it is zero whenever it is zero at the terminal point. So the solution satisfies the transversality condition $L_{\mathfrak{k}}(T) = 0$ whenever $L_{\mathfrak{k}} = 0$ in the above formula.

Remark 2. *Formula* (3.10) *is not new. As far as I know, it appeared first in 1990 in ([13]) and it has also appeared in various contexts in my earlier writings ([11], [7]). But it has never before been obtained directly from the affine system* (1.1) *with controls in the affine hull $\sum_{i=1}^{k} \lambda_i Ad_{h_i}A, h_i \in K, \sum_{i=1}^{k} \lambda_i = 1$.*

Corollary 4. *Let π denote the natural projection from G onto G/K. Then $\pi(g_0e^{tP})$ is a geodesic in G/K that connects $\pi(g_0)$ to $\pi(g(t), g(t) = g_0e^{tP}$. Therefore $\mathcal{T}(g)$ is equal to the shortest length of a geodesic that connects $\pi(I)$ to $\pi(g)$.*

3.4. Fundamental example $(SU(2), SO(2))$

This example is not only typical of the general situation, but is also a natural starting point for problems in quantum control. Recall that $SU(2)$ consists of matrices $\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$ with a and b complex numbers such that $|a|^2 + |b|^2 = 1$. It follows that $g \in G$ whenever $g^{-1} = g^*$, where g^* is the matrix transpose of the complex conjugate of g. Hence the Lie algebra $\mathfrak{su}(2)$ of G consists of matrices $\frac{1}{2}\begin{pmatrix} ix_3 & x_1 + ix_2 \\ -x_1 + ix_2 & -ix_3 \end{pmatrix}$. We will assume that $\mathfrak{su}(2)$ is endowed with the trace metric $\langle X, Y\rangle = -2Tr(XY)$, in which case the skew-Hermitian matrices

$$A_x = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, A_y = \frac{1}{2}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, A_z = \frac{1}{2}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \tag{3.11}$$

form an orthonormal basis in $\mathfrak{su}(2)$. If $X = \frac{1}{2}\begin{pmatrix} iz & x + iy \\ -x + iy & -iz \end{pmatrix}$ is represented by the coordinates $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in R^3$ then the adjoint representation $X \to Ad_g(X)$ is identified with rotations in $\mathbb{R}^3$. If G_x, G_y, G_Z denote the rotations around the axes $\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}$, then A_x, A_y, A_z are the infinitesimal generators of G_x, G_y, G_z which explains the motivation behind the terminology. Relative to the Lie bracket $[A, B] = BA - AB$, A_x, A_y, A_z conform to the following Lie bracket table:

$$[A_x, A_y] = -A_z, [A_z, A_x] = -A_y, [A_y, A_z] = -A_x.$$

The automorphism $\sigma(g) = (g^T)^{-1}$ identifies $SO(2)$ as the group of fixed points by σ, and induces a Cartan decomposition $\mathfrak{g} = \mathfrak{p} + \mathfrak{k}$ with $\mathfrak{p}$ the linear span of A_y and A_z, and $\mathfrak{k}$ the linear span of A_x. Relative to the above decomposition,

$$\frac{dg}{dt} = g(t)(A_z + u(t)A_x) = g(t)\frac{1}{2}\begin{pmatrix} i & u(t) \\ -u(t) & -i \end{pmatrix}, g(t) \in SU_2, \tag{3.12}$$

is a prototypical affine system in G.

Since $[A_z, A_x]] = -A_y$, $G = \mathcal{A}(e, \leq T)$ for some $T > 0$, and since $SU(2)$ is simple, $\mathcal{A}(e, T) = G$ for some $T > 0$ ([14]). However, not all points of G can be reached from the identity in short time as noticed in [14]. For instance, points $g = \begin{pmatrix} x_0 + ix_1 & x_2 + ix_3 \\ -x_2 + ix_3 & x_0 - ix_1 \end{pmatrix}$ in $SU(2)$ with $x_1^2 + x_3^2 > 0$ cannot be reached from the identity in time less than $2(x_1^2 + x_3^2)$. The argument is simple:

$$\begin{aligned} \tfrac{dx_0}{dt} &= -\tfrac{1}{2}(x_1 + ux_2), \tfrac{dx_1}{dt} = \tfrac{1}{2}(x_0 - ux_3) \\ \tfrac{dx_2}{dt} &= \tfrac{1}{2}(ux_0 + x_3), \tfrac{dx_3}{dt} = \tfrac{1}{2}(ux_1 - x_2). \end{aligned}$$

Therefore,

$$x_1\frac{dx_1}{dt} + x_3\frac{dx_3}{dt} = \frac{1}{2}(x_0x_1 - x_2x_3),$$

and hence

$$x_1^2(t) + x_3^2(t) = \int_0^t (x_0(s)x_1(s) - x_2(s)x_3(s))\, ds \leq \frac{t}{2}$$

because $(x_0 - x_1)^2 + (x_2 + x_3)^2 = 1 - 2(x_0x_1 - x_2x_3) \geq 0$ implies that $2(x_0x_1 - x_2x_3) \leq 1$. So if a point g can be reached in time T, then $T \geq 2(x_1^2 + x_3^2)$.

However, not all points of $SU(2)$ can be reached in the shortest time. Below we will show that $-I$ can be reached in any positive time, but is not reachable at $T = 0$. To demonstrate, note that for any $X \in \mathfrak{su}(2)$, $X^2 = -\frac{1}{4}\|X\|^2 I$, and therefore,

$$e^{tX} = I\cos\frac{\|X\|}{2}t + \frac{2}{\|X\|}X\sin\frac{\|X\|}{2}t.$$

In particular when $X = A_z + uA_x, u \in R$, then $\|X\| = \sqrt{1 + u^2}$, and

$$e^{tX} = I\cos\frac{\sqrt{1+u^2}}{2}t + \frac{1}{\sqrt{1+u^2}}\begin{pmatrix} i & u \\ -u & -i \end{pmatrix}\sin\frac{\sqrt{1+u^2}}{2}t.$$

For any $t > 0$ there exists $u \in R$ such that $t\sqrt{1 + u^2} = 2\pi$, and therefore, $e^{tX} = -I$. Therefore, $-I$ can be reached in any positive time t but is not reachable at $T = 0$.

The preceding formula can be used to show that any element of $SO(2)$ lies in the closure of $\mathcal{A}(e, \leq t)$ for any $t > 0$. To do so, let θ be any number, and then let $u_n = 2n\theta$, Then, $e^{\frac{1}{n}X(u_n)} \in \mathcal{A}(e, \leq T)$ for any $T > 0$, provided that n is sufficiently large. An easy calculation shows that

$$\lim_{n\to\infty} e^{\frac{1}{n}X(u_n)} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

Hence $g = \begin{pmatrix} \cos\theta & \sin\theta \\ -sin\theta & \cos\theta \end{pmatrix}$ belongs to $\bar{\mathcal{A}}(e, \leq T)$. It seems likely that $g \in \mathcal{A}(e, \leq T)$, but that has not been verified, as far as I know.

Let us now return to the horizontal system given by

$$\frac{dg}{dt} = g(t)Ad_{h(t)}X_0, \frac{dh}{dt} = h(t)\begin{pmatrix} o & u(t) \\ -u(t) & 0 \end{pmatrix}. \tag{3.13}$$

It follows that

$$h(t) = \begin{pmatrix} \cos\theta(t) & \sin\theta(t) \\ -\sin\theta(t) & \cos\theta(t) \end{pmatrix}, \theta(t) = \int_0^t u(s)\, ds,$$

and therefore

$$\frac{dg}{dt} = g(t)(u_1(t)A_z + u_2(t)A_y), u_1(t) = \cos 2\theta(t), u_2(t) = -\sin 2\theta(t). \tag{3.14}$$

To pass to the convexified horizontal system we need to enlarge the controls to the sphere $u_1^2 + u_2^2 \leq 1$. It then follows that the time-optimal extremals are given by equation (3.10) except for the stationary extremal $g(t) = g_0$.

Let us interpret the above results in slightly different terms with an eye on the connections with quantum control. If $X = x_1A_x + x_2A_y + x_3A_z$ and $Y = y_1A_x + y_2A_y + y_3A_z$, then $Z = [X, Y] = z_1A_x + z_2A_y + z_3A_z$ is given by the vector product $z = y \times x$, where $x = (x_1, x_2, x_3)$, $y = (y_1, y_2, y_3)$, and $z = (z_1, z_2, z_3)$. Hence $[X, Y] = 0$ if and only if x and y are co-linear. Therefore, maximal abelian algebras in $\mathfrak{p}$ are one dimensional, and every non-zero element in $\mathfrak{p}$ is regular. It follows that the Weyl group consists of $\pm I$.

The equation $Ad_hX_0 = L_\mathfrak{p}$ is solvable for each $L_\mathfrak{p} \in \mathfrak{p}$ such that $||L_\mathfrak{p}|| = 1$. Then the line segment that connects $-L_\mathfrak{p}$ and $L_\mathfrak{p}$ is in the convex hull defined by Ad_hX_0. This shows that $\{Ad_hX_0 : h \in K\}$ is the unit circle in $\mathfrak{p}$ and the corresponding convex hull is the unit ball $\{L_\mathfrak{p} \in \mathfrak{p} : ||L_\mathfrak{p}|| \leq 1\}$. The coset extremals are given by

$$e^{tP} = I\cos||P||t + \frac{1}{||P||}P\sin||P||t, P \in \mathfrak{p}, ||P|| = 1. \tag{3.15}$$

These extremals reside on a two dimensional sphere S^2 because

$$e^{iP} = I\cos t\sqrt{a^2+b^2} + \tfrac{i}{\sqrt{a^2+b^2}}P\sin t\sqrt{a^2+b^2} =$$
$$\begin{pmatrix} \cos t\sqrt{a^2+b^2} + \frac{ia}{\sqrt{a^2+b^2}}\sin t\sqrt{a^2+b^2} & \frac{ib}{\sqrt{a^2+b^2}}\sin t\sqrt{a^2+b^2} \\ \frac{ib}{\sqrt{a^2+b^2}}\sin t\sqrt{a^2+b^2} & \cos t\sqrt{a^2+b^2} - \frac{ia}{\sqrt{a^2+b^2}}\sin t\sqrt{a^2+b^2} \end{pmatrix},$$

for any matrix $iP = \begin{pmatrix} ia & ib \\ ib & -ia \end{pmatrix}$ with a and b real. If $x = \cos t\sqrt{a^2+b^2}$, $y = \frac{a}{\sqrt{a^2+b^2}}\sin t\sqrt{a^2+b^2}$, *and* $z = \frac{b}{\sqrt{a^2+b^2}}\sin t\sqrt{a^2+b^2}$, then $x^2 + y^2 + z^2 = 1$. The decomposition $g = e^{iP}R$ corresponds to the Hopf fibration $S^3 \to S^2 \to S^1$.

Hopf fibration has remarkable applications in quantum technology due to the fact that a two level quantum system, called qubit, can be modelled by points in $SU(2)$, whereby all possible states of a particle are represented by complex linear combinations $\alpha(|0>)+\beta(|1>0)$, where $|0>$ and $|1>$ denote the basic levels (states) and where α and β are complex numbers such that $|\alpha|^2+|\beta|^2=1$. In this context, the particle can be either in state $|0>$ with probability $|\alpha|^2$, or in state $|1>$ with probability $|\beta|^2$. For this to make mathematical sense, the basic states are represented by two orthonormal vectors in some complex Hilbert space. Then, the states $\alpha|0>+\beta|1>0$ are identified with matrices $\begin{pmatrix}\alpha & \beta\\ -\bar{\beta} & \bar{\alpha}\end{pmatrix}$ in $SU(2)$.

In this setting, the quotient space G/K is called the Bloch sphere (see for instance [15]). In quantum mechanics points in G/K represent the observable states. It follows that each point g in a given coset is reached time-optimally according to the formula $g=e^{T(Q+P)}e^{-QT}, ||P||=1$ for some $T>0$, but the coset itself is reached time-optimally in the time equal to the length of a geodesic that connects $\pi(I)$ to $\pi(g)$ where π stands for the natural projection from G to G/K.

For instance, if $g_f=-I$, then $g_fK=K$. Therefore, $g(t)=I$, generated by $u(t)=0$, is the only trajectory of the convexified horizontal system that reaches the coset K in zero time. Any other optimal trajectory is of the form $g(t)=e^{t(Q+P)}e^{-Qt}$, and such trajectories cannot reach points in zero time.

4. Notable Riemannian symmetric pairs

4.1. $(SL(n),SO(n))$ *and* $(SU(n),SO(n))$

Each of these pairs of Lie groups is symmetric relative to the automorphism $\sigma(g)=(g^T)^{-1}$ where g^T denotes the matrix transpose. It follows that $K=SO(n)$ is the group of points in G fixed by σ. Then, $\mathfrak{g}$ is equal to $\mathfrak{sl}(n)$ when $G=SL(n)$ and is equal to $\mathfrak{su}(n)$ when $G=SU(n)$. In the first case the Lie algebra is equal to the space of $n\times n$ matrices with zero trace, while in the second case the Lie algebra consists of $n\times n$ complex skew-symmetric matrices with zero trace. Then, $\mathfrak{g}=\mathfrak{p}\oplus\mathfrak{k}$, where $\mathfrak{p}$ is equal to the space of symmetric matrices in $\mathfrak{sl}(n)$ and the space of symmetric matrices with imaginary entries in $\mathfrak{su}(n)$. These two Lie algebras are dual in the sense that the Cartan decomposition $\mathfrak{p}+\mathfrak{k}$ in $\mathfrak{sl}(n)$ corresponds to the Cartan decomposition $\mathfrak{k}+i\mathfrak{p}$ in $\mathfrak{su}(n)$ (see [6] for further details). In each case, the Killing form is equal to $2nTr(XY)$. It follows that it is positive on $\mathfrak{p}$ in $\mathfrak{sl}(n)$ and negative on $\mathfrak{p}$ in $\mathfrak{su}(n)$. Therefore, the pair $(SL(n),SO(n))$ is non-compact, while the pair $(SU(n),SO(n))$ is compact.

In $\mathfrak{sl}(n)$, each matrix X in $\mathfrak{p}$ can be diagonalized by some $Ad_h, h\in K$, and the set of all diagonal matrices $\mathcal{D}$ in $\mathfrak{p}$ forms an $n-1$ dimensional abelian algebra, which is also maximal since $[\mathcal{D},X]=0$ can only hold only if X is diagonal. It follows that $n-1$ is the rank of the underlying symmetric space. If X is a diagonal matrix with its diagonal entries $x_1,\dots,x_n$ then $ad(X)Y=\sum_{i,j}^n(x_i-x_j)Y_{ij}e_i\otimes e_j$ for any matrix $Y=\sum_{i,j}^n Y_{ij}e_i\otimes e_j$. Hence

$$adX(e_i\otimes e_j)=(x_i-x_j)e_i\otimes e_j, i\neq j, adX(\mathcal{D})=0, \tag{4.1}$$

that is, $\alpha(X)=x_i-x_j$ are the non-zero roots in $\mathcal{D}$. This implies that X is regular if and only if the diagonal entries of X are all distinct.

Weyl chambers in $\mathcal{D}$ are in one to one correspondence with the elements of the permutation group on n letters. For if $X=diag(x_1,\dots,x_n)$ and $Y=diag(y_1,\dots,y_n)$ are any regular elements in $\mathcal{D}$ then there exist unique permutations α and β on n letters such that $x_{\alpha(1)}>x_{\alpha(2)}>\cdots>x_{\alpha(n)}$ and $y_{\beta(1)}>y_{\beta(2)}>\cdots>y_{\beta(n)}$. If X and Y are in the same Weyl chamber, then $(x_i-x_j)(y_i-y_j)>0$ for all i

and j. It then follows by an easy argument that $\alpha = \beta$. The reasoning on $\mathfrak{su}(n)$ with diagonal matrices having imaginary entries is similar and will be omitted.

It follows that the Weyl orbit $Ad_h(X_0)$ in $\mathcal{D}$ consists of the diagonal matrices with diagonal entries a permutation of the diagonal entries of X_0. The convex hull spanned by these matrices coincides with the controls of the convexified system that reside in $\mathcal{D}$.

4.2. *Self-adjoint subgroups of $SL(n)$*

A subgroup G of $SL(n)$ is called *self-adjoint* if the matrix transpose g^T is in G for any g in G. Any self-adjoint group G admits an involutive automorphism $\sigma(g) = g^{T^{-1}}, g \in G$, with $K = SO(n) \cap G$ equal to the group of its fixed points.

It follows that the Lie algebra $\mathfrak{g}$ of G admits a Cartan decomposition $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$ where $\mathfrak{k} = \mathfrak{g} \cap \mathfrak{so}(n)$ and $\mathfrak{p} = Sym(n) \cap \mathfrak{g}$ with $Sym(n)$ the space of symmetric matrices in $\mathfrak{sl}(n)$. Since $\langle X, Y\rangle = 2nTr(XY)$ inherited from $\mathfrak{sl}(n)$ is positive on $\mathfrak{p}$ the pair (G, K) is a symmetric Riemannian pair of non-compact type.

One can show that $SO(p, q), p + q = n$, the group that preserves the scalar product $(x, y)_{p,q} = \sum_{i=1}^{p} x_i y_i - \sum_{i=p+1}^{n} x_i y_i$ is self-adjoint, as well as $Sp(n)$, the group that leaves the symplectic form $\sum_{i=1}^{n} x_i y_{n+i} - y_i x_{i+n}, x, y \in \mathbb{R}^n$ invariant.

When $G = SO(p, q)$ the Lie algebra $\mathfrak{g}$ consists of block matrices $M = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$ with A and C skew-symmetric $p \times p$ and $q \times q$ matrices and B an arbitrary $p \times q$ matrix.Then $M \in \mathfrak{p}$ if $A = C = 0$, and $M \in \mathfrak{k}$ if $B = 0$. The quotient space $SO(p, q)/K$ can be identified with an open subset of Grassmannians consisting of all q-dimensional subspaces in $\mathbb{R}^{(p+q)}$ on which $(x, x)_{p,q} > 0$, while the quotient spaces $Sp(n)/K$ can be identified with the generalized Poincaré plane $\mathcal{P}_n = \{X + iY, X^T = X, Y^T = Y, Y > 0\}$ ([7], pages 126, 127).

4.3. *Rank one symmetric spaces*

In rank-one symmetric spaces the Weyl group is minimal (it consists of two elements $\pm I$)), which accounts for an easier visualization of the general theory. We will use $(SO(1, n), K)$ together with its compact companion $(SO(n + 1), K)$, $K = \{1\} \times SO(n)$ to illustrate the relevance of the rank for the general theory. Both of the above cases can be treated simultaneously in terms of the parameter $\varepsilon = \pm 1$ and the scalar product $(x, y)_\varepsilon = x_0 y_0 + \varepsilon \sum_{i=1}^{n} x_i y_i$. In that spirit, $SO_\varepsilon(n + 1)$ will denote $SO(1, n)$ when $\varepsilon = -1$, and $SO(n + 1)$ when $\varepsilon = 1$.

Each group $SO_\varepsilon(n + 1)$ acts on points of $\mathbb{R}^{n+1}$ by the matrix multiplication and this action can be used to identify the quotient space $SO_\varepsilon(n + 1)/K$ with the orbit $O(e_0) = \{ge_0 : g \in SO_\varepsilon(n + 1)\}$ where $e_0 = (1, 0, \ldots, 0)^T$. Since $SO_\varepsilon(n + 1)$ preserves $(\, , \,)_\varepsilon$, $O(e_0)$ is the Euclidean sphere S^n when $\varepsilon = 1$ and the hyperboloid $\mathbb{H}^n$ when $\varepsilon = -1$.

Let now $\mathfrak{g}_\varepsilon = \mathfrak{so}_\varepsilon(n + 1)$ denote the Lie algebra of $SO_\varepsilon(n + 1)$ equipped with its natural scalar product $\langle X, Y\rangle = \frac{1}{2}Tr(XY)$, and let $\mathfrak{k}$ denote the Lie algebra of K. It is easy to check that the orthogonal complement $\mathfrak{p}_\varepsilon$ of $\mathfrak{k}$ is given by $= \{e_0 \wedge_\varepsilon u, u \in \mathbb{R}^{n+1}, (u, e_0)_\varepsilon = 0\}$, and that $\mathfrak{k}$ itself is given by $\mathfrak{k} = \{(u \wedge_\varepsilon v) : (u, e_0)_\varepsilon = (v, e_0)_\varepsilon = 0\}$, where

$$(u \wedge_\varepsilon v) = u \otimes_\varepsilon v - v \otimes_\varepsilon u \in \mathbb{R}^{n+1}, v \in \mathbb{R}^{n+1},$$

with $u \otimes_\varepsilon v$ the rank-one matrix defined by $(u \otimes_\varepsilon v)x = (v, x)_\varepsilon u, x \in \mathbb{R}^{n+1}$.

It follows that Cartan's relations

$$[\mathfrak{p}_\varepsilon, \mathfrak{p}_\varepsilon] = \mathfrak{k}_\varepsilon, [\mathfrak{p}_\varepsilon, \mathfrak{k}_\varepsilon] = \mathfrak{p}_\varepsilon, [\mathfrak{k}_\varepsilon, \mathfrak{k}_\varepsilon] \subseteq \mathfrak{k}_\varepsilon,$$

hold, as can be readily verified through the following general formula

$$[a \wedge_\varepsilon b, c \wedge_\varepsilon d] = (a,c)_\varepsilon(b \wedge_\varepsilon d) + (b,d)_\varepsilon(a \wedge_\varepsilon c) - (b,c)_\varepsilon(a \wedge_\varepsilon d) - (a,d)_\varepsilon(b \wedge_\varepsilon c).$$

Since $\langle e_0 \wedge_\varepsilon u, e_0 \wedge_\varepsilon v\rangle = -\varepsilon \sum_{i=1}^n u_i v_i$, the bilinear form $\langle\, ,\, \rangle$ is positive on $\mathfrak{p}_\varepsilon$ when $\varepsilon = -1$ and is negative when $\varepsilon = 1$. It follows that the pair (G_ε, K) is a compact type when $\varepsilon = 1$ and a non-compact type when $\varepsilon = -1$.

We now return to time optimality. The space $\mathfrak{p}_\varepsilon = \{u \wedge_\varepsilon e_0 : \langle u, e_0\rangle_\varepsilon = 0\}$ is n-dimensional. If $U = u \wedge_\varepsilon e_0$ and $V = v \wedge_\varepsilon e_0$ are arbitrary elements in $\mathfrak{p}_\varepsilon$ then $[U, V] = u \wedge_\varepsilon v$. Hence $[U, V] = 0$ if and only if u and v are parallel. Thus each maximal abelian algebra is one-dimensional and each non-zero element U in $\mathfrak{p}_\varepsilon$ is regular. The Weyl group consists of two elements I_1 and I_2 such that $Ad_{I_1}U = U$ and $Ad_{I_2}U = -U$.

If $h = \{1\} \times R$ for some $R \in SO(n)$, then $Ad_h X_0 = Rx_0 \wedge_\varepsilon e_0$. Since $SO(n)$ acts transitively on the spheres S^n, $Ad_K X_0 = S^n \wedge_\varepsilon e_0$. If $L_\mathfrak{p} = l \wedge_\varepsilon e_0$ then $Rx_0 = l$ yields $Ad_h X_0 = L_\mathfrak{p}$. The above shows that $\{Ad_h X_0, h \in K\} = \{x \wedge_\varepsilon e_0, ||x|| = ||x_0||\}$ and the convex hull is equal to $\{x \wedge_\varepsilon e_0 : ||x|| \leq ||x_0||\}$.

4.4. Compact Lie groups

Each semi-simple compact Lie group K is a symmetric space realized as the quotient $G/\tilde{K}$, with $G = K \times K$ and $\tilde{K} = \{(g, g) : g \in K\}$ under the automorphism $\sigma(g_1, g_2) = (g_2, g_1)$.

If $\mathfrak{k}$ denotes the Lie algebra of K then $\mathfrak{g} = \mathfrak{k} \times \mathfrak{k}$ is the Lie algebra of G, and $\tilde{\mathfrak{k}} = \{(X, X), X \in \mathfrak{k}\}$ is the Lie algebra of $\tilde{K}$. Then, $\mathfrak{p} = \{(X, -X) : X \in \mathfrak{k}\}$ is the orthogonal complement of $\tilde{\mathfrak{k}}$ in $\mathfrak{g}$ relative to the natural bi-invariant metric inherited from K. It then follows that $\tilde{\mathfrak{k}}$ and $\mathfrak{p}$ satisfy Cartan's decomposition (1.4). To pass to the quotient space $G/\tilde{K}$, note that G acts on K by the natural action

$$\tau((g_1, g_2), h) = g_1 h g_2^{-1}.$$

Since $h_2 h_1 h_1^{-1} = h_2$ the action is transitive. In particular the orbit through the group identity is identified with K.

Maximal abelian algebras in $\mathfrak{p}$ are in exact correspondence with maximal abelian algebras in $\mathfrak{k}$. Any $\tilde{X}_0 \in \mathfrak{p}$ is of the form $\tilde{X}_0 = (X_0, -X_0)$ for some $X_0 \in \mathfrak{k}$. If $h \in \tilde{K}$ is of the form $h = (g, g)$, then $Ad_h\tilde{X}_0 = (Ad_g(X_0), -Ad_g X_0)$. Therefore, time-optimal solutions associated with

$$\frac{d\tilde{g}}{dt} = \tilde{g}(t)(Ad_h(\tilde{X}_0)), \tilde{g} = (g_1(t), g_2(t)) \in G$$

are given by

$$g_1(t) = g_1(0)e^{t(P+Q)}e^{-tQ}, g_2(t) = g_2(0)e^{t(-P+Q)}e^{-tQ}, \tag{4.2}$$

for some elements $P \in \mathfrak{k}$ and $Q \in \mathfrak{k}$, with $h(t) = g_1(0)e^{t(P+Q)}e^{t(-P+Q)}g_2^{-1}(0)$ the projection on K in accordance with equation (3.10).

5. Applications to quantum control-n chains

5.1. Finite dimensional Schrödinger equation and the associated control systems

In non-relativistic quantum mechanics, time evolution of a finite dimensional quantum system is governed by a time dependent Schrödinger equation

$$\frac{dz}{dt} = -iH(t)z(t), \tag{5.1}$$

in an n-dimensional complex Hilbert space $\mathbb{H}^n$, where $H(t)$ is a fixed time varying Hermitian operator in $\mathbb{H}^n$ ([1]). Recall that $H(t)$ is Hermitian if $\langle H(t)z, w\rangle = \langle z, H(t)w\rangle$ for z, w in $\mathbb{H}^n$ where $\langle\ ,\ \rangle$ denotes the Hermitian quadratic form on $\mathbb{H}^n$.

In what follows, points in $\mathbb{H}^n$ will be represented by the coordinates $z_1, \ldots, z_n$ relative to an orthonormal basis in $\mathbb{H}^n$, and $\mathbb{H}^n$ will be identified with $\mathbb{C}^n$ with the Hermitian scalar product $\langle z, w\rangle = \sum z_i \bar{w}_i$ for any z and w in $\mathbb{C}^n$, with $\bar{w}_i$ the complex conjugate of w_i. Then, a matrix H is Hermitian if $H^* = H$, where H^* is equal to the complex conjugate of the matrix transpose of H.

Equation (5.1) is subordinate to the master equation

$$\frac{dg}{dt} = -iH(t)g(t), g(0) = I, \tag{5.2}$$

in the unitary group $U(n)$, in the sense that every solution $z(t)$ of (5.1) that satisfies $z(0) = z_0$ is given by $z(t) = g(t)z_0$. Recall that iH is skew-Hermitian for each Hermitian matrix H, hence every solution $g(t)$ of equation (5.2) that originates in $U(n)$ evolves in $U(n)$. It follows that $||z(t)|| = ||z_0||$, i.e., the reachable sets of (5.1) evolve on the spheres S^{2n-1}.

To be consistent with the first part of the paper, we will focus on the left-invariant form of the master equation

$$\frac{dg}{dt} = g(t)(iH(t)). \tag{5.3}$$

Of course, it is easy to go from one form to the other; if $g(t)$ is a solution of (5.2), then $g^{-1}(t)$ is a solution of (5.3) and vice versa.

As a way of bridging the language gap between quantum control literature and mainstream control theory, we will make a slight detour into the Kronecker products of matrices and the associated operations. For our purposes it suffices to work with square matrices. Then the Kronecker product $U \otimes V$ of any $n \times n$ matrix U and any $m \times m$ matrix V is equal to the $nm \times nm$ matrix with block entries $(u_{ij}V), i, j \leq n$. The Kronecker product enjoys the following properties:

$$\begin{gathered}(U \otimes V)(W \otimes Z) = UW \otimes VZ, (U \otimes V)^* = U^* \otimes V^* \\ Tr(U \otimes V) = Tr(U)Tr(V), Det(U \otimes V) = (DetU)^m(DetV)^n.\end{gathered} \tag{5.4}$$

It follows that $(U \otimes V) \in U(nm)$ for any $U \in U(n)$ and $V \in U(m)$: similarly, $U \otimes V$ is in $SU(mn)$ whenever $U \in SU(n)$ and $V \in SU(m)$ and n and m are of the same parity. It can be easily shown that

$$[U_1 \otimes V_1, U_2 \otimes V_2] = [U_1, U_2] \otimes V_2V_1 + U_1U_2 \otimes [V_1, V_2], \tag{5.5}$$

for any matrices U_1, U_2 of the same size, and any matrices V_1, V_2 also of the same size (recall our convention $[X, Y] = YX - XY$).

The following proposition assembles some facts that are relevant for the n-spin chains.

Proposition 17. *If $U \in \mathfrak{u}(n)$ (resp. $U \in \mathfrak{su}(n)$) and I_k is the k-dimensional identity matrix. then both $I_k \otimes U$ and $U \otimes I_k$ belong to $\mathfrak{u}(nk)$ (resp. $\mathfrak{su}(nk)$).*

However, if $U \in \mathfrak{u}(n)$ and $V \in \mathfrak{u}(m)$, then $i(U \otimes V) \in \mathfrak{u}(nm)$. Similarly, $i(U \otimes V)$ is in $\mathfrak{su}(nm)$ whenever $U \in \mathfrak{su}(n)$ and $V \in \mathfrak{su}(m)$.

Proof. $(I_k \otimes U)^* = I_k^* \otimes U^* = I_k \otimes (-U) = -(I_k \otimes U)$. Hence $I_k \otimes U \in \mathfrak{u}(n)$. If $Tr(U) = 0$ then $Tr(I_k \otimes U) = 0$. In addition,

$$(i(U \otimes V))^* = -i(U^* \otimes V^*) = -i(-U) \otimes (-V) = -i(U \otimes V).$$

We will now direct our attention to the n-spin chains introduced in [1] and [2]. These chains are defined in terms of the Kronecker products of Pauli matrices

$$I_x = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, I_y = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, I_z = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{5.6}$$

The n-spin chains oriented in the z-direction are defined by the Hamiltonians

$$H = \sum_{j=2}^{n} J_{(j-1)j} I_{(j-1)z} I_{jz} + \sum_{i=1}^{m} v_i(t) I_{ix} + u_i(t) I_{iy}), n \geq 2, m \leq n, \tag{5.7}$$

where J_{ij} are the coupling constants, and where I_{ix}, I_{iy}, I_{iz} denotes the matrix $X_1 \otimes X_2 \otimes \cdots \otimes X_n$ where $X_i = I_x$ (resp. $X_i = I_y, X_i = I_z$) in the i-th position and where all the remaining elements X_j are equal to the identity I_2. This kind of spin-chains are known as the Ising spin chains ([16], [17]). We will now address time optimality of the associated left-invariant master system (5.3). Each chain defines a pair of Lie algebras $(\mathcal{L}, \mathfrak{k}_v)$ where $\mathfrak{k}_v$, the vertical algebra, is the Lie algebra generated by the controlling vector fields I_{ix} and I_{iy}, $i = 1, \ldots m$, and where $\mathcal{L}$ is the controllability algebra generated by the drift element $\sum_{j=2}^{n} J_{(j-1)j} I_{(j-1)z} I_{jz}$ and $\mathfrak{k}_v$.

We will now consider two and three spin chains with a particular interest on the cases where $\mathcal{L} = \mathfrak{su}(n)$ for some integer n and where $\mathfrak{k}_v$ is a subalgebra of $\mathcal{L}$ such that the Cartan conditions (1.4) hold for the pair $(\mathfrak{p}, \mathfrak{k}_v)$ with $\mathfrak{p}$ equal to the orthogonal complement of $\mathfrak{k}_v$ in $\mathcal{L}$. For the sake of uniformity with the first part of the paper, we will work with the matrices A_x, A_y, A_z introduced in equations (3.11) rather than with the Pauli matrices I_x, I_y, I_z. Recall that

$$A_x = iI_y, A_y = iI_x, A_z = iI_z. \tag{5.8}$$

In this notation then

$$H = -(\sum_{j=2}^{n} J_{(j-1)j} A_{(j-1)z} A_{jz} + i \sum_{i=1}^{m} (v_i(t) A_{iy} + u_i(t) A_{ix})), n \geq 2, m \leq n, \tag{5.9}$$

As a preliminary first step, let us single out the symmetric (irreducible) Riemannian pairs (G, K) in which $G = SU(n)$ for some n. It is known that there are only three such Riemannian spaces

$$SU(n)/SO(n), SU(2n)/Sp(n) \text{ and } SU(p+q)/S(U(p) \times U(q)), \tag{5.10}$$

where $S(U(p)\times U(q)) = SU(p+q)\cap(U(p)\times U(q))$ ([6], p. 518).

The first symmetric space $(SU(n), SO(n))$, known as **Type AI**, has already been discussed in the preceding section. The second symmetric space, **Type AII**, occurs on $SU(2n)$ and is induced by the automorphism

$$\sigma(g) = J_n(g^{-1})^T J_n^{-1},\ J_n = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}.$$

Then $\sigma(g) = g$ if and only if ${g^{-1}}^T J_n = J_n g$, or $J_n = g^T J_n g$, which in turn means that $g\in Sp(n)$, where $Sp(n) = SU(2n)\cap Sp(2n,\mathbb{C})$. Then

$$\sigma_*(X) = \frac{d}{dt}J_n(e^{-tX})^T J_n^{-1}|_{t=0} = \frac{d}{dt}J_n e^{t\bar{X}} J_n^{-1}|_{t=0} = J_n\bar{X}J_n^{-1}.$$

It follows that $\mathfrak{k} = \{X\in\mathfrak{su}(2n) : J_n\bar{X}J_n^{-1} = X\}$ and $\mathfrak{p} = \{X\in\mathfrak{su}(2n) : J_n\bar{X}J_n^{-1} = -X\}$. If $X = \begin{pmatrix} X_{11} & X_{12} \\ -\bar{X}_{12}^T & X_{22} \end{pmatrix}$ is the decomposition of X into the $n\times n$ blocks, then

$$J_n\bar{X}J_n^{-1} = \begin{pmatrix} \bar{X}_{22} & X_{12}^T \\ -\bar{X}_{12} & \bar{X}_{11} \end{pmatrix}.$$

Therefore, $X\in\mathfrak{k}$ if and only if

$$X_{11} = \bar{X}_{22} \text{ and } X_{12} = X_{12}^T,$$

and $X\in\mathfrak{p}$ if and only if

$$X_{11} = -\bar{X}_{22}, Tr(X_{11}) = 0, \text{ and } X_{12}^T = -X_{12}.$$

The remaining symmetric space, **Type AIII**, is associated with the automorphism

$$\sigma(g) = I_{p,q}gI_{p,q}^{-1},\ g\in SU(p+q), \text{ where } I_{p,q} = \begin{pmatrix} -I_p & 0 \\ 0 & I_q \end{pmatrix}.$$

The induced automorphism on $\mathfrak{su}(p+q)$ is given by $\sigma_*(X) = I_{p,q}XI_{p,q}^{-1}$. Then

$$\mathfrak{k} = \{X\in\mathfrak{su}(p+q) : X = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\},\ \mathfrak{p} = \{X\in\mathfrak{su}(p+q) : X = \begin{pmatrix} 0 & C \\ -\bar{C}^T & 0 \end{pmatrix}$$

where A is a $p\times p$ matrix and B is a $q\times q$ matrix such that $Tr(A+B) = 0$, and where C is an arbitrary $p\times q$ matrix with complex entries. Then $S(U(p)\times U(q))$ denotes the subgroup of $SU(p+q)$ whose Lie algebra consists of matrices $X = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$, with $A\in\mathfrak{u}(p)$, $B\in\mathfrak{u}(q)$ such that $Tr(A+B) = 0$.

In all these cases the metric on $\mathfrak{p}$ coincides with the restriction of the canonical metric on $\mathfrak{su}(n)$ given by $\langle X, Y\rangle = -\frac{1}{2}Tr(XY) = \frac{1}{2}Tr(X\bar{Y}^T)$.

The relevance of these classical classifications for the problems of quantum control has already been noticed in the existing literature ([1] and [2] in regard to **Type AI**, and [18] in regard to **Type AIII**).

5.2. Two-spin chains

The two-spin chains given by

$$H = -(\sum_{j=2}^{2} J_{(j-1)j}A_{(j-1)z}A_{jz} + i\sum_{i=1}^{m}(u_i(t)A_{ix} + v_i(t)A_{iy})), m \leq 2,$$

give rise to the rescaled left-invariant master equation ($J_{(j-1)j} = 1$)

$$\frac{dg}{dt} = g(t)i(-A_z \otimes A_z) + \sum_{i=1}^{m} u_i(t)A_{ix} + v_i(t)A_{iy}, \tag{5.11}$$

where now A_{ix} and A_{iy} are the chains with A_x and A_y in the i-th position.

Let now $\mathfrak{k}_v$ denote the vertical subalgebra generated by the controlling vector fields $A_{iy}, A_{ix}, i = 1, \dots, m$. For $m = 1$ there are two controls u and v associated with the controlling matrices $A_x \otimes I_2$ and $A_y \otimes I_2$, and for $m = 2$ there are four controls u_1, u_2, v_1, v_2 associated with matrices $A_x \otimes I_2, I_2 \otimes A_x, A_y \otimes I_2, I_2 \otimes A_y$.

It is easy to verify that $\mathfrak{k}_v = \{X \otimes I_2 : X \in \mathfrak{su}(2)\}$ for $m = 1$, and $\mathfrak{k}_v = \{X \otimes I_2 + I_2 \otimes Y : X \in \mathfrak{su}(2), Y \in \mathfrak{su}(2)\}$ for $m = 2$. In the first case $\mathfrak{k}_v$ is a three-dimensional algebra isomorphic to $\mathfrak{su}(2)$, and in the second case it is a six dimensional Lie algebra isomorphic to $\mathfrak{su}(2) \times \mathfrak{su}(2)$.

Lemma 2. *If A and B are any matrices in $\mathfrak{su}(2)$, then*

$$AB = -\langle A, B\rangle I_2 + \frac{1}{2}[B, A], \text{ where } \langle A, B\rangle = -\frac{1}{2}Tr(AB). \tag{5.12}$$

The mapping ϕ defined by $\phi(iX \otimes Y) = iY \otimes X$,
$\phi(X \otimes I_2) = I_2 \otimes X, \phi(I_2 \otimes X) = X \otimes I_2$, X, Y in $\mathfrak{su}(2)$ is a Lie algebra isomorphism on $\mathfrak{su}(4)$.

Proof. If $A = \begin{pmatrix} ia_3 & a \\ -\bar{a} & -ia_3 \end{pmatrix}$ and $B = \begin{pmatrix} ib_3 & b \\ -\bar{b} & -ib_3 \end{pmatrix}$ then

$$AB + BA = -2(a_1b_1 + a_2b_2 + a_3b_3)I_2 = -2\langle A, B\rangle I_2.$$

Hence $2AB = -2\langle A, B\rangle I_2 + [B, A]$. This proves the first part of the lemma.

Then

$$\begin{aligned}
&[\phi(iA \otimes B), \phi(iC \otimes D)] = [iB \otimes A, iD \otimes C] = \\
&-[B, D] \otimes AC - DB \otimes [A, C] = \langle A, C\rangle[B, D] \otimes I_2 + \langle D, B\rangle I_2 \otimes [A, C] \\
&= \phi(\langle A, C\rangle I_2 \otimes [B, D] + \langle D, B\rangle \otimes [A, C] \otimes I_2) = \phi([iA \otimes B], iC \otimes D]),
\end{aligned}$$

and

$$[\phi(A \otimes I_2), \phi(i(B \otimes C))] = i(C \otimes [A, B]) = \phi([A \otimes I_2, i(B \otimes C)]).$$

Hence ϕ is an isomorphism.

Proposition 18. *Let $\mathcal{L}$ denote the Lie algebra generated by $i(A_z \otimes A_z)$ and $\mathfrak{k}_v$. When $m = 1$, $\mathcal{L} = \mathfrak{p} \oplus \mathfrak{k}$, $\mathfrak{p} = i(\mathfrak{su}(2) \otimes A_z)$ and $\mathfrak{k} = \mathfrak{su}(2) \otimes I_2$. If ϕ is the isomorphism from the previous lemma then $\phi(\mathcal{L}) = \begin{pmatrix} \mathfrak{su}(2) & 0 \\ 0 & \mathfrak{su}(2) \end{pmatrix}$ and*

$$\phi(\mathfrak{p}) = \{\begin{pmatrix} X & 0 \\ 0 & -X \end{pmatrix}, X \in \mathfrak{su}(2)\}, \phi(\mathfrak{k}_v) = \{\begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix}, X \in \mathfrak{su}(2)\}. \tag{5.13}$$

Proof. Evidently, $\mathfrak{k} = \mathfrak{k}_v$. Secondly, $[i(A_z \otimes A_z), X \otimes I_2] = i([A_z, X] \otimes A_z)$ for any X in $\mathfrak{su}(2)$. This implies that both $i(A_y \otimes A_z)$ and $i(A_x \otimes A_z)$ are in $\mathcal{L}$. Therefore $\mathfrak{p} \subset \mathcal{L}$. Since $\langle X \otimes I_2, Y \otimes A_z \rangle = -\frac{1}{2} Tr(XY) Tr(I_z) = 0$, $\mathfrak{k}_v$ and $\mathfrak{p}$ are orthogonal. Also, $[i(X \otimes A_z), i(Y \otimes A_z)] = -[X, Y] \otimes A_z^2 = \frac{1}{4}[X, Y] \otimes I_2$. Therefore $[\mathfrak{p}, \mathfrak{p}] \subseteq \mathfrak{k}_v$ $\mathcal{L} = \mathfrak{p} \oplus \mathfrak{k}$. Hence $\mathfrak{p}$ and $\mathfrak{k}$ satisfy Cartan's conditions (1.4) and consequently $\mathcal{L} = \mathfrak{p} \oplus \mathfrak{k}_v$.

If ϕ is the isomorphism from the preceding lemma, then $\phi(-2iX \otimes A_z)) = -2iA_z \otimes X = \begin{pmatrix} X & 0 \\ 0 & -X \end{pmatrix}$ for any $-2iX \otimes A_z$ in $\mathfrak{p}$, and $\phi(X \otimes I_2) = I_2 \otimes X = \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix}$ for $X \otimes I_2 \in \mathfrak{k}$. The linear span of these matrices is equal to $\begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix}$, X, Y in $\mathfrak{su}(2)$.

The above shows that the $m = 1$ chain can be represented on $G = SU(2) \times SU(2)$ as

$$\frac{dg_1}{dt} = g_1(t)(\frac{1}{2}A_z + u_1(t)A_x + v_1(t)A_y), \frac{dg_2}{dt} = g_2(t)(-\frac{1}{2}A_z + u_1(t)A_x + v_1(t)A_y).$$

The time-optimal solutions are of the form

$$g_1(t) = g_1(0)e^{t(P+Q)}e^{-tQ}, g_2(t) = g_2(0)e^{t(-P+Q)}e^{-tQ}, \tag{5.14}$$

$P \in \mathfrak{su}(2)$, $Q \in \mathfrak{su}(2)$, with $h(t) = g_1(0)e^{t(P+Q)}e^{t(-P+Q)}g_2^{-1}(0)$ the projection on $SU(2)$ (in accordance with (4.2)).

Proposition 19. *For $m = 2$, $\mathcal{L} = \mathfrak{su}(4)$. If*

$$\mathfrak{k}_v = \{X \otimes I_2 + I_2 \otimes Y, \{X, Y\} \subset \mathfrak{su}(2)\}, \mathfrak{p} = \{i(X \otimes Y) : \{X, Y\} \subset \mathfrak{su}(2)\},$$

then $\mathcal{L} = \mathfrak{p} + \mathfrak{k}_v$ and

$$[\mathfrak{p}, \mathfrak{k}_v] \subseteq \mathfrak{p}, [\mathfrak{p}, \mathfrak{p}] \subseteq \mathfrak{k}_v.$$

Proof. Let $\mathfrak{p} = \{i(X \otimes Y) : X \in \mathfrak{su}(2), Y \in \mathfrak{su}(2)\}$. It then follows that $\mathfrak{su}(4) = \mathfrak{p} \oplus \mathfrak{k}_v$ by an easy dimensionality argument. Straightforward calculations shows that $\mathfrak{p}$ and $\mathfrak{k}_v$ satisfy Cartan's conditions

$$[\mathfrak{p}, \mathfrak{k}_v] \subseteq \mathfrak{p}, [\mathfrak{p}, \mathfrak{p}] \subseteq \mathfrak{k}_v, [\mathfrak{k}_v, \mathfrak{k}_v] \subseteq \mathfrak{k}_v.$$

So it suffices to show that $\mathfrak{p} \subset \mathcal{L}$.

Since $i(A_z \otimes A_z)$ is in $\mathfrak{p}$,

$$[i(A_z \otimes A_z), X \otimes I_2 + I_2 \otimes Y] = i[A_z, X] \otimes A_z + A_z \otimes i[A_z, Y]$$

is in $\mathcal{L}$ for any X and Y in $\mathfrak{su}(2)$. Therefore both $i[A_z, X] \otimes A_z$ and $i(A_z \otimes i[A_z, Y])$ are in $\mathcal{L}$, which then implies that $i(X \otimes A_z)$ and $i(A_z \otimes Y)$ are in $\mathcal{L}$ for any X, Y in $\mathfrak{su}(2)$ (because $i(A_z \otimes A_z)$ is in $\mathcal{L}$).

But then $[i(X \otimes A_z), I_2 \otimes Y] = X \otimes i[A_z, Y]$ and $[i(X \otimes A_z), Y \otimes I_2] = i([X, Y] \otimes A_z$ yields that $i(X \otimes Y)$ is in $\mathcal{L}$ for any X and Y in $\mathfrak{su}(2)$.

Corollary 5. *The reachable set from the identity is equal to $SU(4)$.*

The following lemma reveals the connection to the appropriate symmetric Riemannian space.

Lemma 3. *Let* $h = \sqrt{2}\begin{pmatrix} -A_z & A_y \\ A_x & -\frac{1}{2}I_2 \end{pmatrix}$. *Since* $h^* = \bar{h}^T = \sqrt{2}\begin{pmatrix} A_z & -A_x \\ -A_y & -\frac{1}{2}I_2 \end{pmatrix} = h^{-1}$, *and* $Det(h) = 1$, *h belongs to $SU(4)$. Then*

$$Ad_h(A \otimes I_2) = \frac{1}{2}\begin{pmatrix} 0 & -a_1 & -a_2 & -a_3 \\ a_1 & 0 & -a_3 & a_2 \\ a_2 & a_3 & 0 & -a_1 \\ a_3 & -a_2 & a_1 & 0 \end{pmatrix}, Ad_h(I_2 \otimes B) = \frac{1}{2}\begin{pmatrix} 0 & -b_1 & b_2 & -b_3 \\ b_1 & 0 & b_3 & b_2 \\ -b_2 & -b_3 & 0 & b_1 \\ b_3 & -b_2 & -b_1 & 0 \end{pmatrix}.$$

Also, $Ad_h(i(A \otimes B)) = \frac{1}{4}i\begin{pmatrix} C_1 & C_2 \\ C_2^T & C_3 \end{pmatrix}$, $C_1 = \begin{pmatrix} -a_1b_1 + a_2b_2 - a_3b_3 & a_3b_2 + a_2b_3 \\ a_3b_2 + a_2b_3 & -a_1b_1 - a_2b_2 + a_3b_3 \end{pmatrix}$,

$$C_2 = \begin{pmatrix} a_3b_1 - a_1b_3 & -a_1b_2 - a_2b_1 \\ a_1b_2 - a_2b_1 & -a_1b_3 - a_3b_1 \end{pmatrix}, C_3 = \begin{pmatrix} a_1b_1 + a_2b_2 + a_3b_3 & a_3b_2 - a_2b_3 \\ a_3b_2 - a_2b_3 & a_1b_1 - a_2b_2 - a_3b_3 \end{pmatrix}$$

for any matrices $A = \frac{1}{2}\begin{pmatrix} ia_3 & a \\ -\bar{a} & -ia_3 \end{pmatrix}$ *and* $B = \frac{1}{2}\begin{pmatrix} ib_3 & b \\ -\bar{b} & -ib_3 \end{pmatrix}$, $a = a_1 + ia_2$ *and* $b = b_1 + ib_2$. *We leave these verifications to the reader.*

It then follows that

$$Ad_h(\mathfrak{k}_v) = \mathfrak{so}(4), Ad_h(\mathfrak{p}) = \{iS : S \in \mathfrak{sl}(4), S^T = S\} \tag{5.15}$$

which then yields that the quotient space $SU(4)/K_v$ is isomorphic to the symmetric space $SU(4)/SO(4)$. The above formulas also show that the two-spin system with $m = 2$ is conjugate to

$$\frac{dg}{dt} = \frac{1}{4}g(t)\begin{pmatrix} i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \end{pmatrix} + \frac{1}{2}g(t)\begin{pmatrix} 0 & -U_1 & -V_2 & 0 \\ U_1 & 0 & 0 & V_1 \\ V_2 & 0 & 0 & -U_2 \\ 0 & -V_1 & U_2 & 0 \end{pmatrix}$$

where

$$U_1 = u_1 + u_2, U_2 = u_1 - u_2, V_1 = v_1 + v_2, V_2 = v_1 - v_2.$$

For $m = 1$ the controls are reduced to $U = U_1 = U_2$ and $V = V_1 = V_2$.

Corollary 6. *The time optimal solutions for the two-spin chains are given by the same formulas as in Proposition 16.*

5.3. The three-spin chains

Let us now consider the three-spin systems

$$\frac{dg}{dt} = g(t)(-i\sum_{j=2}^{3} J_{(j-1)j}A_{(j-1)z}A_{jz} + \sum_{i=1}^{m}(u_i(t)A_{ix} + v_i(t)A_{iy}), \ m \le 3 \tag{5.16}$$

in $G = SU(8)$.

It follows that $A_{1z}A_{2z} = (A_z \otimes I_2 \otimes I_2)(I_2 \otimes A_z \otimes I_2) = (A_z \otimes A_z) \otimes I_2$. Similarly, $A_{2z}A_{3z} = I_2 \otimes (A_z \otimes A_z)$. So the drift Hamiltonian H_d is of the form

$$H_d = ai(A_z \otimes I_z) \otimes I_2 + bI_2 \otimes i(A_z \otimes I_z),$$

where a and b are arbitrary non-zero constants. In the case that $m = 3$, the controlled Hamiltonians are given by

$$H_1 = A_x \otimes I_2 \otimes I_2, H_2 = A_y \otimes I_2 \otimes I_2, H_3 = I_2 \otimes A_x \otimes I_2,$$
$$H_4 = I_2 \otimes A_y \otimes I_2, H_5 = I_2 \otimes I_2 \otimes A_x, H_6 = I_2 \otimes I_2 \otimes A_y.$$

It is easy to verify that the vertical algebra $\mathfrak{k}_v$ generated by the controlled Hamiltonians is equal to

$$\mathfrak{su}(2) \otimes I_2 \otimes I_2, m = 1,$$
$$\mathfrak{su}(2) \otimes I_2 \otimes I_2 + I_2 \otimes \mathfrak{su}(2) \otimes I_2, m = 2,$$
$$\mathfrak{su}(2) \otimes I_2 \otimes I_2 + I_2 \otimes \mathfrak{su}(2) \otimes I_2 + I_2 \otimes I_2 \otimes \mathfrak{su}(2), m = 3.$$

Case $m = 1$ is similar to its two spin analogue and will be omitted. The remaining cases $m = 2$ and $m = 3$, however, show new phenomena that take their solutions outside the general framework described earlier in the paper.

The following lemma highlights some of the calculations in $m = 2$.

Lemma 4. *Let $\mathfrak{k} = \mathfrak{k}_v + \mathfrak{k}_h$ where $\mathfrak{k}_v = \mathfrak{su}(2) \otimes I_2 \otimes I_2 + I_2 \otimes \mathfrak{su}(2) \otimes I_2$ and $\mathfrak{k}_h = \mathfrak{su}(2) \otimes \mathfrak{su}(2) \otimes A_z$. Then $\mathfrak{k}$ is a Lie subalgebra in $\mathfrak{su}(8)$, $\langle \mathfrak{k}_v, \mathfrak{k}_h \rangle = 0$ and*

$$[\mathfrak{k}_h, \mathfrak{k}_v] \subseteq \mathfrak{k}_h, [\mathfrak{k}_h, \mathfrak{k}_h] \subset \mathfrak{k}_v.$$

The proof follows by simple calculations which we leave to the reader..

Proposition 20. *For $m = 2$, the Lie algebra $\mathcal{L}$ generated by H_d and the controlled Hamiltonians H_1, H_2, H_3, H_4 contains the Lie algebra $\mathfrak{k}$ in the preceding lemma. If $\mathfrak{p}$ denotes the orthogonal complement of $\mathfrak{k}$ in $\mathcal{L}$ then $\mathcal{L} = \mathfrak{k} + \mathfrak{p}$ and $[\mathfrak{p}, \mathfrak{k}] \subseteq \mathfrak{p}, [\mathfrak{p}, \mathfrak{p}] \subseteq \mathfrak{k}, [\mathfrak{k}, \mathfrak{k}] \subseteq \mathfrak{k}$.*

Proof. For $m = 2$, $\mathfrak{k}_v = \mathfrak{su}(2) \otimes I_2 \otimes I_2 + I_2 \otimes \mathfrak{su}(2) \otimes I_2$ is a subalgebra in $\mathcal{L}$. If X_1 and X_2 are any elements in $\mathfrak{su}(2)$ let $\tilde{X}_1 = X_1 \otimes I_2 \otimes I_2$ and $\tilde{X}_2 = X_2 \otimes I_2 \otimes I_2$. Then,

$$ad\tilde{X}_1(H_d) = a([X_1, A_z] \otimes A_z \otimes iI_2),$$
$$ad\tilde{X}_2 ad\tilde{X}_1(H_d) = a[X_2, [X_1, A_z]] \otimes A_z \otimes iI_2.$$

Therefore $\mathfrak{su}(2) \otimes A_z \otimes iI_2$ is in $\mathcal{L}$ since X_1, X_2 are arbitrary and $a \neq 0$. In particular, $-a(A_z \otimes A_z \otimes iI_2) \subseteq \mathcal{L}$, and consequently $b(iI_2 \otimes A_z \otimes A_z) \subseteq \mathcal{L}$.

Let now $\tilde{Y}_1 = I_2 \otimes Y_1 \otimes I_2$ and $\tilde{Y}_2 = I_2 \otimes Y_2 \otimes I_2$ with Y_1 and Y_2 arbitrary elements in $\mathfrak{su}(2)$. Then

$$ad\tilde{Y}_1(A_z \otimes A_z \otimes iI_2) = A_z \otimes [Y_1, A_z] \otimes iI_2,$$
$$ad\tilde{Y}_2 \mathrm{ad}\tilde{Y}_1(A_z \otimes A_z \otimes iI_2) = A_z \otimes [Y_2, [Y_1, A_z]] \otimes iI_2$$

show that $A_z \otimes \mathfrak{su}(2) \otimes iI_2$ is in $\mathcal{L}$. Similar calculation with $iI_2 \otimes A_z \otimes A_z$ in place of $A_z \otimes A_z \otimes iI_2$ shows that $iI_2 \otimes \mathfrak{su}(2) \otimes A_z$ is also in $\mathcal{L}$. But then

$$[iI_2 \otimes X \otimes A_z, A_z \otimes Y \otimes iI_2] = A_z \otimes [X, Y] \otimes A_z.$$

Hence $A_z \otimes \mathfrak{su}(2) \otimes A_z$ is in $\mathcal{L}$. Finally,

$$ad\tilde{X}_2 ad\tilde{X}_1(A_z \otimes X \otimes A_z) = [X_2, [X_1, A_z]] \otimes X \otimes A_z, X \in \mathfrak{su}(2),$$

shows that $\mathfrak{su}(2) \otimes \mathfrak{su}(2) \otimes A_z$ is in $\mathcal{L}$. Therefore $\mathfrak{k}$ of the preceding lemma in $\mathcal{L}$.

Let now $\mathfrak{p} = \mathfrak{su}(2)\otimes\mathfrak{su}(2)\otimes iI_2 + iI_2\otimes\mathfrak{su}(2)\otimes A_z + \mathfrak{su}(2)\otimes iI_2\otimes iA_z$. We showed above that $iI_2\otimes\mathfrak{su}(2)\otimes A_z$ is in $\mathcal{L}$. Since $[iI_2 \otimes \mathfrak{su}(2) \otimes A_z, \mathfrak{k}_h]$ is in $\mathcal{L}$, $[iI_2 \otimes Z \otimes A_z, X \otimes Y \otimes A_z] = -\frac{1}{4}X \otimes [Z, Y]] \otimes iI_2$ is in $\mathcal{L}$ for any X, Y, and Z in $\mathfrak{su}(2)$. That is, $\mathfrak{su}(2) \otimes \mathfrak{su}(2) \otimes iI_2$ is in $\mathcal{L}$.

An easy calculation with $[\mathfrak{su}(2) \otimes \mathfrak{su}(2) \otimes iI_2, \mathfrak{k}_h]$ shows that $\mathfrak{su}(2) \otimes iI_2 \otimes A_z$ belongs to $\mathcal{L}$. Therefore $\mathfrak{p} \subset \mathcal{L}$.

It follows from above that both $\mathfrak{p}$ and $\mathfrak{k}$ are in $\mathcal{L}$. Since $\mathfrak{p}$ and $\mathfrak{k}$ are orthogonal, $\mathfrak{p} \cap \mathfrak{k} = \{0\}$, and $[\mathfrak{p}, \mathfrak{k}] \subseteq \mathfrak{p}$. The reader can readily show that $[\mathfrak{p}, \mathfrak{p}] \subseteq \mathfrak{k}$. Therefore $\mathfrak{k}$ and $\mathfrak{p}$ satisfy Cartan's conditions (1.4), and consequently $\mathfrak{k} + \mathfrak{p}$ is a Lie algebra. Since $\mathcal{L} \subseteq (\mathfrak{k} + \mathfrak{p}) \subseteq \mathcal{L}$, $\mathcal{L} = \mathfrak{k} + \mathfrak{p}$.

Proposition 21. *$\mathcal{L}$ is isomorphic to $\mathfrak{su}(4) \times \mathfrak{su}(4)$, and $\mathfrak{k}$ is isomorphic to $\mathfrak{su}(4)$.*

Proof. First, let us note that $\mathfrak{k}$ and $\mathfrak{su}(4)$ are isomorphic under the isomorphism

$$F(X \otimes Y \otimes A_z + Z \otimes I_2 \otimes I_2 + I_2 \otimes W \otimes I_2) = i(X \otimes Y) + Z \otimes I_2 + I_2 \otimes W.$$

Indeed $F([U, V]) = [F(U), F(V)]$ for any U and V in $\mathfrak{k}_v$ by a straightforward calculation. If U and V are in $\mathfrak{k}_h$ then $U = X_1\otimes X_2\otimes A_z$ and $V = Y_1\otimes Y_2\otimes A_z$. It follows that $[U, V] = \frac{1}{4}(\langle X_2, Y_2\rangle[X_1, X_2]\otimes I_2 + \langle X_1, Y_1\rangle I_2\otimes [X_2, Y_2])\otimes)I_2$, and hence $F([U, V]) = \frac{1}{4}(\langle X_2, Y_2\rangle[X_1, X_2] \otimes I_2 + \langle X_1, Y_1\rangle I_2 \otimes [X_2, Y_2]) = [F(U), F(V)]$. The remaining case $U \in \mathfrak{k}_v$, $V \in \mathfrak{k}_h$ also yields $F([U, V]) = [F(U), F(V)]$ which shows that F is an isomorphism whose range is $\mathfrak{su}(4)$. Thus $\mathfrak{k}$ is isomorphic to $\mathfrak{su}(4)$.

Then $\mathfrak{p}$ can be identified with the Hermitian matrices in $\mathfrak{sl}(4, \mathbb{C})$ via the identification

$$X \otimes Y \otimes iI_2 + Z \otimes iI_2 \otimes A_z + iI_2 \otimes W \otimes A_z \cong X \otimes Y + i(Z \otimes I_2 + I_2 \otimes W),$$

Now $\mathfrak{su}(4)$ is a compact real form for $\mathfrak{sl}(4, \mathbb{C})$ ($\mathfrak{sl}(4, \mathbb{C}) = \mathfrak{su}(4) + i\mathfrak{su}(4)$). It follows that $\mathcal{L}$ and the real Lie algebra generated by $\mathfrak{sl}(4, \mathbb{C})$ are isomorphic, (since $\mathfrak{sl}(4, \mathbb{C})$ is the complexification of $\mathfrak{su}(4)$).

The above calculations show that the horizontal systems associated with three-spin systems starting with $m = 2$ exhibit notable differences from the horizontal systems associated with two-spin systems that considerably complicate the time-optimal solutions. As demonstrated above, the reachable set G is isomorphic to $SU(4) \times SU(4)$ and K is isomorphic to $SU(4)$, hence $M = SU(4) \times SU(4)/SU(4)$ is the associated symmetric Riemannian space. However, the Lie algebra generated by the controlled vector fields is a proper subalgebra of the isotropy algebra $\mathfrak{k}$ ($\mathfrak{k}_v = \mathfrak{su}(2) \times \mathfrak{su}(2)$ and $\mathfrak{k} = \mathfrak{su}(4)$), and therefore the associated homogeneous manifold G/K_v does not admit a natural metric compatible with the decomposition $\mathfrak{k}_v^\perp + \mathfrak{k}_v$. As a consequence, the time optimal solutions of the horizontal system

$$\frac{dg}{dt} = g(t)Ad_{h(t)}(a(A_z \otimes A_z \otimes iI_2) + b(iI_2 \otimes A_z \otimes A_z)), h(t) \in K_v$$

are no longer given by the exponentials of matrices in $\mathfrak{p}$ mainly because K is no longer the symmetry group for the horizontal system.

The same phenomena occur in the three-spin chains with $m = 3$. For then

$$\mathfrak{k}_v = \mathfrak{su}(2) \otimes I_2 \otimes I_2 + I_2 \otimes \mathfrak{su}(2) \otimes I_2 + I_2 \otimes I_2 \otimes \mathfrak{su}(2)$$

is contained in the Lie algebra $\mathfrak{k}$ equal to the linear span of $\mathfrak{k}_v$ and matrices of the form $X \otimes Y \otimes Z$ where each of X, Y, Z range over the matrices in $\mathfrak{su}(2)$. A simple count shows that $dim(\mathfrak{k}) = 36$. Then $\mathfrak{p}$, the linear span of matrices $X \otimes Y \otimes Z$, where one of the matrices X, Y, Z is equal to iI_2 and the remaining two are in $\mathfrak{su}(2)$, is orthogonal to $\mathfrak{k}$. Since $dim(\mathfrak{p})) = 27$, $dim(\mathfrak{p} + \mathfrak{k}) = 63 = dim(\mathfrak{su}(8))$. Hence $\mathfrak{su}(8) = \mathfrak{p} \oplus \mathfrak{k}$.

Proposition 22. *The preceding decomposition $\mathfrak{p} \oplus \mathfrak{k}$ is a Cartan decomposition of* **Type AII** *associated with the symmetric space $SU(8)/Sp(4)$.*

Proof. Let us recall $h = \sqrt{2}\begin{pmatrix} -A_z & A_y \\ A_x & -\frac{1}{2}I_2 \end{pmatrix}$ from Proposition 3. Since h is a point in $SU(4)$, $\Psi = \begin{pmatrix} h & 0 \\ 0 & h \end{pmatrix}$ is a point in $SU(8)$ and hence Ad_Ψ is an isomorphism on $\mathfrak{su}(8)$.

Let $Ad_\Psi(X \otimes Y \otimes Z) = M = \begin{pmatrix} M_{11} & M_{12} \\ -M_{12}^* & M_{22} \end{pmatrix}$ where M_{ij} are 4×4 matrices. To show that $Ad_\Psi(\mathfrak{k})$ and $Ad_\Psi(\mathfrak{p})$ correspond to a Cartan pair of type **AII** we need to show that $Ad_\Psi(\mathfrak{k})$ satisfies $M_{11} = \bar{M}_{22}$ and $M_{12} = M_{12}^T$, and $Ad_\Psi(\mathfrak{p})$ satisfies $M_{11} = -\bar{M}_{22}, Tr(M_{11}) = 0,$ and $M_{12}^T = -M_{12}$.

When $X = \frac{1}{2}\begin{pmatrix} ix_3 & x \\ -\bar{x} & -ix_3 \end{pmatrix}$, $Y = \frac{1}{2}\begin{pmatrix} iy_3 & y \\ -\bar{y} & -iy_3 \end{pmatrix}$, $Z = \frac{1}{2}\begin{pmatrix} iz_3 & z \\ -\bar{z} & -iz_3 \end{pmatrix}$, $X \otimes Y \otimes Z$ belongs to $\mathfrak{k}$ and

$$Ad_\Psi(X \otimes Y \otimes Z) = \begin{pmatrix} ix_3 Ad_h(Y \otimes Z) & x Ad_h(Y \otimes Z) \\ -\bar{x} Ad_h(Y \otimes Z) & -ix_3 Ad_h(Y \otimes Z) \end{pmatrix}$$

The formulas in Lemma 3 show that $Ad_h(Y \otimes Z)$ is a symmetric matrix with real entries. Hence $\bar{M}_{22} = M_{11}$ and $M_{12}^T = M_{12}$.

If one of X, Y, Z is equal to iI_2 then $X \otimes Y \otimes Z$ belongs to $\mathfrak{p}$. When $X = iI_2$ then

$$Ad_\Psi(iI_2 \otimes Y \otimes Z) = \begin{pmatrix} iAd_h(Y \otimes Z) & 0 \\ 0 & iAd_h(Y \otimes Z) \end{pmatrix} = \begin{pmatrix} M_{11} & 0 \\ 0 & M_{22} \end{pmatrix}.$$

Evidently, $\bar{M}_{22} = -M_{11}$.

In the complementary case when Y or Z is iI_2 and $X = \frac{1}{2}\begin{pmatrix} ix_3 & x \\ -\bar{x} & -ix_3 \end{pmatrix}$, $M_{11} = ix_3 Ad_h(Y \otimes Z)$, $M_{22} = -ix_3 Ad_h(Y \otimes Z)$, and $M_{12} = x Ad_h(Y \otimes Z)$. It follows that $Ad_h i(Y \otimes Z)$ is a skew-symmetric matrix and therefore, $\bar{M}_{22} = -M_{11}$ and $M_{12}^T = -M_{12}$.

In the remaining cases two elements in $X \otimes Y \otimes Z$ are equal to I_2 and $X \otimes Y \otimes Z$ belongs to $\mathfrak{k}$. If $Y = Z = I_2$ then $M_{11} = ix_3 I_4$, $M_{22} = -ix_3 I_4$ and $M_{12} = xI_4$. Evidently $M_{11} = M_{22}$ and $M_{12}^T = M_{12}$

When $X = I_2$ then either Y or Z is equal to I_2. But then $Ad_h(Y \otimes Z)$ is a skew-symmetric matrix, and therefore $M_{11} = iAd_h(Y \otimes Z) = M_{22} = -iAd_h(Y \otimes Z)$, and $M_{12} = 0$. Hence $Ad_\Psi(\mathfrak{k})$ and $Ad_\Psi(\mathfrak{p})$ correspond to the Cartan factors of **Type AII**.

Proposition 23. *For $m = 3$ the three spin system (56) is controllable in $SU(8)$.*

Proof. Let $\mathcal{L}$ denote the Lie algebra generated by H_d and $\mathfrak{k}_v$. Then, $[H_d, \mathfrak{su}_2 \otimes I_2 \otimes_2] = a(A_z^\perp \otimes A_z \otimes iI_2)$, and $[A_z^\perp \otimes A_z \otimes iI_2, I_2 \otimes \mathfrak{su}_(2) \otimes I_2] = A_z^\perp \otimes A_z^\perp \otimes I_2$, where $A_z^\perp$ denotes the orthogonal complement of A_z in $\mathfrak{su}(2)$.

Similarly, $[H_d, I_2 \otimes I_2 \otimes \mathfrak{su}(2)] = b(iI_2 \otimes A_z \otimes A_z^{\perp})$, and $[iI_2 \otimes A_z \otimes A_z^{\perp}, I_2 \otimes \mathfrak{su}(2) \otimes I_2] = iI_2 \otimes A_z^{\perp} \otimes A_z^{\perp}$. Therefore, both $iI_2 \otimes A_z^{\perp} \otimes A_z^{\perp}$ and $A_z^{\perp} \otimes A_z^{\perp} \otimes iI_2$ belong to $\mathcal{L}$. In particular $A_x \otimes A_x \otimes iI_2$, $A_y \otimes A_y \otimes iI_2$, $iI_2 \otimes A_x \otimes A_x$, and $iI_2 \otimes A_y \otimes A_y$ all belong to $\mathcal{L}$.

Analogous calculations with $A_x \otimes A_x \otimes iI_2$, $iI_2 \otimes A_x \otimes A_x$, $A_y \otimes A_y \otimes iI_2$, and $iI_2 \otimes A_y \otimes A_y$ show that $A_x^{\perp} \otimes A_x^{\perp} \otimes iI_2$, $iI_2 \otimes A_x^{\perp} \otimes A_x^{\perp}$ belong to $\mathcal{L}$, as well as $A_y^{\perp} \otimes A_y^{\perp} \otimes iI_2$ and $iI_2 \otimes A_y^{\perp} \otimes A_y^{\perp}$.

Therefore, $\mathfrak{su}(2) \otimes \mathfrak{su}(2) \otimes iI_2$ and $iI_2 \otimes \mathfrak{su}(2) \otimes \mathfrak{su}(2)$ belong to $\mathcal{L}$. But then $[\mathfrak{su}(2) \otimes \mathfrak{su}(2) \otimes iI_2, iI_2 \otimes \mathfrak{su}(2) \otimes \mathfrak{su}(2)] = \mathfrak{su}(2) \otimes \mathfrak{su}(2) \otimes \mathfrak{su}(2)$. Hence $\mathfrak{k} \subset \mathcal{L}$. But then, $\mathfrak{su}(2) \otimes iI_2 \otimes \mathfrak{su}(2)$ is contained in

$$[\mathfrak{su}(2) \otimes \mathfrak{su}(2) \otimes iI_2 + iI_2 \otimes \mathfrak{su}(2) \otimes \mathfrak{su}(2), \mathfrak{su}(2) \otimes \mathfrak{su}(2) \otimes \mathfrak{su}(2)],$$

and therefore, $\mathfrak{p} \subset \mathcal{L}$.

The above suggests that one cannot expect time optimal solutions of three-spin chains to have a simple and computable form. However, there are some solvable cases that shed light on the general situation. One such case is a three-spin chain defined by the drift $H_d = 2(J_{12}(I_z \otimes I_z \otimes I_2) + J_{21}(I_2 \otimes I_z \otimes I_z))$ controlled by a single Hamiltonian $H_c = I_{2y} = I_2 \otimes I_y \otimes I_2$. This system first appeared in studies on nuclear magnetic resonance spectroscopy ([3]), ([19]), ([4]).

Let us first make some introductory remarks on the results presented in ([3], [4]). The aforementioned studies begin with the density equation

$$\frac{d\rho}{dt} = -i[H_d + uH_c, \rho] \tag{5.17}$$

associated with a right-invariant affine system

$$\frac{dg}{dt} = -i(H_d + u(t)H_c)g(t), \tag{5.18}$$

with $H_d = 2(J_{12} I_z \otimes I_z \otimes I_2 + J_{21} I_2 \otimes I_z \otimes I_z)$ and $H_c = I_2 \otimes I_x \otimes I_2$.

The density equation is assumed to evolve in the Hilbert space $\mathcal{H}$ of Hermitian matrices in $i\mathfrak{su}(8)$ endowed with its natural scalar product $\langle X, Y\rangle = \frac{1}{2}Tr(XY)$. Recall that iX is Hermitian for each $X \in \mathfrak{su}(n)$.

Rather than studying the density equation directly, the above papers consider instead the time-optimal evolution of the expectation values of certain elements in $\mathcal{H}$, where the expectation value of an element X along a solution $\rho(t)$ is defined by $\langle X, \rho(t)\rangle$. It then follows that the expectation value of X evolves in time according to

$$\frac{d}{dt}\langle X, \rho(t)\rangle = -\langle X, i[H_d + u(t)H_c), \rho]\rangle = -\langle [X, i(H_d + u(t)H_c)], \rho(t)\rangle.$$

In particular when $X = X_1 = (I_x \otimes I_2 \otimes I_2)$, then $\langle [X_1, i(H_d + u(t)H_c)], \rho(t)\rangle = -J_{12}\langle 2(I_x \otimes I_z \otimes I_2), \rho\rangle$. Hence the expected value $x_1 = \langle X_1, \rho\rangle$ evolves according to

$$\frac{dx_1}{dt} = -J_{12}\langle 2(I_y \otimes I_z \otimes I_2), \rho\rangle = -J_{12}x_2(t)$$

where $x_2(t)$ is the expected value of $X_2 = 2(I_y \otimes I_z \otimes I_2)$. Continuing this way one obtains new elements X_3 and X_4 whose expectation values $x_3(t)$ and $x_4(t)$ together with $x_1(t)$ and $x_2(t)$ satisfy a closed differential system

$$\frac{dx}{dt} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & -u & 0 \\ 0 & u & 0 & -k \\ 0 & 0 & k & 0 \end{pmatrix} x(t), k = \frac{J_{23}}{J_{12}}, \tag{5.19}$$

with the time rescaled by a factor J_{12}, where $x(t)$ is the column vector in R^4 with the coordinates x_1, x_2, x_3, x_4. In fact, $x_3 = -\langle 2I_x \otimes I_y \otimes I_2, \rho\rangle$, and $x_4 = \langle 4iI_x \otimes I_x \otimes I_z, \rho\rangle$ ([4]). The above authors then pose the time-optimal problem of reaching $(0,0,0,1)^T$ from $(1,0,0,0)^T$ in the least amount of time. We will refer to this problem as the Yuan's optimal problem since it was originated in ([3]).

Rather than tackling this problem directly, the papers ([3]), ([19]), ([4]) concentrate on certain lower dimensional approximations and then show that these approximations are integrable in terms of elliptic functions. As far as I know, the original problem remained open.

We will show that Yuan's problem and the time optimal problem associated with the affine system (5.18) are essentially the same and both can be integrated in terms of elliptic functions.

5.4. Symmetric three-spin systems

For the sake of consistency with the rest of the paper we will formulate (5.18) in the left-invariant way as

$$\frac{dg}{dt} = g(t)(i(H_d + u(t)H_c)), \tag{5.20}$$

with $iH_d = 2(J_{12}(I_z \otimes I_z \otimes iI_2) + J_{21}(iI_2 \otimes I_z \otimes I_z))$ and $H_c = I_{2y}$, which we will write as $iH_d = 2a(A_z \otimes A_z \otimes iI_2) + 2b(iI_2 \otimes A_z \otimes A_z)$, $a = -J_{12}, b = -J_{23}$, and $iH_c = iI_{2y} = I_2 \otimes A_x \otimes I_2$. We will refer to the above system as a *symmetric three-spin system.*

Proposition 24. *If $\mathcal{L}$ denotes the Lie algebra generated by iH_d and iH_c then $\mathcal{L}$ is the vector space spanned by*

$$U_1 = I_2 \otimes A_x \otimes I_2, U_2 = 2(A_z \otimes A_y \otimes iI_2), U_3 = 2(A_z \otimes A_z \otimes iI_2),$$
$$V_1 = -4(A_z \otimes A_x \otimes A_z), V_2 = 2(iI_2 \otimes A_y \otimes A_z), V_3 = 2(iI_2 \otimes A_z \otimes A_z).$$

Proof.

$$[iH_d, iH_c] = [2a(A_z \otimes A_z \otimes iI_2) + 2b(iI_2 \otimes A_z \otimes A_z), I_2 \otimes A_x \otimes I_2)] = -2a(A_z \otimes A_y \otimes iI_2) - 2b(iI_2 \otimes A_y \otimes A_z).$$

Therefore $H_2 = a(A_z \otimes A_y \otimes iI_2) + b(iI_2 \otimes A_y \otimes A_z)$ is in $\mathcal{L}$. Then

$$[\frac{1}{2}H_d, H_2] = \frac{1}{4}(a^2 + b^2)(I_2 \otimes A_x \otimes I_2 + 2ab(A_z \otimes A_x \otimes A_z),$$

hence $H_3 = A_z \otimes A_x \otimes A_z$ belongs to $\mathcal{L}$. Continuing,

$$H_4 = [H_2, H_3] = -\frac{1}{4}(a(iI_2 \otimes A_z \otimes A_z) + b(A_z \otimes A_z \otimes iI_2),$$

is in $\mathcal{L}$. But then

$$4aH_4 + \frac{b}{2} iH_d = (b^2 - a^2)(iI_2 \otimes A_z \otimes A_z), \text{ and } 4bH_4 + \frac{a}{2} iH_d = (a^2 - b^2)(A_z \otimes A_z \otimes iI_2),$$

and hence, $H_5 = A_z \otimes A_z \otimes iI_2$, and $H_6 = iI_2 \otimes A_z \otimes A_z$ are in $\mathcal{L}$.

Finally, $[H_5, H_c] = [A_z \otimes A_z \otimes iI_2, I_2 \otimes A_x \otimes I_2] = A_z \otimes A_y \otimes iI_2$, which it turn implies that $iI_2 \otimes A_y \otimes A_z$ is in $\mathcal{L}$. We have now shown that

$$A_z \otimes A_z \otimes iI_2, iI_2 \otimes A_z \otimes A_z, I_2 \otimes A_x \otimes I_2, A_z \otimes A_y \otimes iI_2, iI_2 \otimes A_y \otimes A_z, A_z \otimes A_x \otimes A_z$$

are contained in $\mathcal{L}$.

Let now

$$U_1 = I_2 \otimes A_x \otimes I_2, U_2 = 2(A_z \otimes A_y \otimes iI_2), U_3 = 2(A_z \otimes A_z \otimes iI_2),$$
$$V_1 = -4(A_z \otimes A_x \otimes A_z), V_2 = 2(iI_2 \otimes A_y \otimes A_z), V_3 = 2(iI_2 \otimes A_z \otimes A_z).$$

It is now easy to verify that the above matrices satisfy the following Lie bracket table

Table 1

[,]	U_1	U_2	U_3	V_1	V_2	V_3
U_1	0	$-U_3$	U_2	0	$-V_3$	V_2
U_2	U_3	0	$-U_1$	V_3	0	$-V_1$
U_3	$-U_2$	U_1	0	$-V_2$	V_1	0
V_1	0	$-V_3$	V_2	0	$-U_3$	U_2
V_2	V_3	0	$-V_1$	U_3	0	$-U_1$
V_3	$-V_2$	V_1	0	$-U_2$	U_1	0

Let $\mathcal{L}_0$ denote the linear span of matrices $U_i, V_i, i = 1, 2, 3$. It follows from the above table that $\mathcal{L}_0$ is a Lie subalgebra of $\mathfrak{su}(8)$. Since iH_d and iH_c belong to $\mathcal{L}_0$, $\mathcal{L} \subseteq \mathcal{L}_0$. But then $\mathcal{L}_0 \subseteq \mathcal{L}$ by our construction. Therefore $\mathcal{L}_0 = \mathcal{L}$.

Corollary 7. *$\mathcal{L}$ is isomorphic to* $\mathfrak{so}(4)$.

Proof. Let $\hat{U}_1 = e_4 \wedge e_3, \hat{U}_2 = e_2 \wedge e_4, \hat{U}_3 = e_2 \wedge e_3, \hat{V}_1 = e_2 \wedge e_1, \hat{V}_2 = e_3 \wedge e_1, \hat{V}_3 = e_4 \wedge e_1$. Then $\hat{U}_i, \hat{V}_i, i = 1, 2, 3$ is a standard basis in $\mathfrak{so}(4)$ that conforms to the same Lie bracket table as displayed in Table 1.

Proposition 25. *The set of points reachable from the identity by the trajectories of*

$$\frac{dg}{dt} = g(t)i((H_d + u(t)H_c)), \; g(0) = I_8,$$

is a six dimensional subgroup G of $SU(8)$ isomorphic to $SO(4)$.

Proof. $\mathcal{L}$ is a Lie algebra isomorphic to $\mathfrak{so}(4)$, which is also isomorphic to $\mathfrak{su}(2) \times \mathfrak{su}(2)$. In fact if $\mathfrak{g}_1$ is the linear span of $\frac{1}{2}(U_1 + V_1), \frac{1}{2}(U_2 + V_2), \frac{1}{2}(U_3 + V_3)$, and $\mathfrak{g}_2$ is the linear span of $\frac{1}{2}(U_1 - V_1), \frac{1}{2}(U_2 - V_2), \frac{1}{2}(U_3 - V_3)$, then $\mathcal{L} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$, $[\mathfrak{g}_1, \mathfrak{g}_2] = 0$ and each factor $\mathfrak{g}_i$ is isomorphic to $\mathfrak{su}(2)$.

Since $\mathcal{L}$ is isomorphic to $\mathfrak{su}(2) \times \mathfrak{su}(2)$ there is a subgroup $\tilde{G}$ in $SU(8)$ which is isomorphic to $SU(2) \times SU(2)$ (Lie algebras are in one to one correspondence with simply connected Lie groups ([6])). But then $SU(2) \times SU(2)$ is a double cover of $SO(4)$ and $SO(4)$ is the connected component of $SU(2) \times SU(2)$ that contains the group identity (see for instance [11]). Therefore the reachable set of (5.20) is a subgroup G of $\tilde{G}$ isomorphic to $SO(4)$.

In terms of the notations introduced above (5.20) can be rewritten as

$$\frac{dg}{dt} = g(t)((aU_3 + bV_3) + u(t)U_1), g(0) = I, \tag{5.21}$$

or as

$$\frac{dg}{dt} = g(t)((kU_3 + V_3) + u(t)U_1), k = \frac{a}{b}, g(0) = I, \tag{5.22}$$

after suitable reparametrizations ($t \to \frac{t}{b}, u \to \frac{u}{b}$).

We will now reformulate Yuan's problem as a variational problem on the sphere S^4 realized as the quotient $SO(4)/K, K = \{1\} \times SO(3)$ under the right action $(g, x) \to g^{-1}x$. Then equation (5.19) can be recast as

$$\frac{dg}{dt} = -g(t)\begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & -u & 0 \\ 0 & u & 0 & -k \\ 0 & 0 & k & 0 \end{pmatrix}, g(0) = I, x(t) = g^{-1}(t)e_1$$

or as

$$\frac{dg}{dt} = -g(t)(\tilde{V}_1 + k\tilde{U}_1 + u\tilde{U}_3), x(t) = g^{-1}(t)e_1 \tag{5.23}$$

in terms of the basis $\hat{U}_1 = e_4 \wedge e_3, \hat{U}_2 = e_2 \wedge e_4, \hat{U}_3 = e_2 \wedge e_3, \hat{V}_1 = e_2 \wedge e_1, \hat{V}_2 = e_3 \wedge e_1, \hat{V}_3 = e_4 \wedge e_1$ introduced in the preceding corollary.

Proposition 26. *Yuan's differential system* (5.23) *is isomorphic to the affine-symmetric system* (5.22)*.*

Proof. Let $R = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$. Then $R \in SO(4)$ and hence, $R^{-1} = R^T$. If $\tilde{g}(t) = Rg(t)R^{-1}$ then $\tilde{g}(t)$ is a solution curve of

$$\frac{d\tilde{g}}{dt} = \tilde{g}(t)(\hat{V}_3 + k\hat{U}_3 + u(t)\hat{U}_1) \tag{5.24}$$

for any solution $g(t)$ of equation (5.23). The correspondence $U_i \to \tilde{U}_i, V_i \to \tilde{V}_i$ is a Lie algebra isomorphism from $\mathcal{L}$ onto $\mathfrak{so}(4, \mathbb{R})$. So (5.23) and (5.24) are isomorphic and (5.24) and (5.22) are isomorphic.

It follows that the time optimal solutions of (5.23) and (5.22) are qualitatively the same, apart from the fact that in Yuan's problem time optimality is relative to the cosets gK. We will come back to this point later on in the text. Let us now come to the horizontal three-spin symmetric system

$$\frac{dg}{dt} = g(t)Ad_{h(t)}(kU_3 + V_3), \tag{5.25}$$

where $h(t)$ is a solution of $\frac{dh}{dt} = u(t)h(t)U_1)$. Since $(2U_1)^2 = -I_8$ where I_8 is the identity in $SU(8)$,

$$e^{2U_1 t} = I_8(1 - \frac{t^2}{2} + \frac{t^4}{4!} - \cdots) + 2U_1(t - \frac{t^3}{3!} + \frac{t^5}{5!} - \cdots) = I_8 \cos t + 2U_1 \sin t,$$

or $e^{U_1 t} = I_8 \cos \frac{t}{2} + 2U_1 \sin \frac{t}{2}$. Let now $\theta(t) = \int_0^t u(s)\,ds + \theta_0$. Then

$$h(t) = e^{\theta(t)U_1} = I \cos \frac{1}{2}\theta(t) + 2U_1 \sin \frac{1}{2}\theta(t).$$

Easy calculations show that

$$U_1U_3U_1 = \frac{1}{4}U_3, U_1V_3U_1 = \frac{1}{4}V_3, \text{ and } [U_1, (kU_3 + V_3)] = -(kU_2 + V_2).$$

Therefore,

$$h(t)(kU_3 + V_3)h^{-1}(t) = (kU_3 + V_3)\cos\theta - (kU_2 + V_2)\sin\theta.$$

It follows that (5.25) is of the form

$$\frac{dg}{ds} = g(t)(V_3 + kU_3)u_1(s) + (V_2 + kU_2)u_2(s)), g(0) = I, \tag{5.26}$$

where $u_1(s) = \cos\theta(s), u_2(s) = -\sin\theta(s)$. To pass to its convex extension it is sufficient to enlarge the controls to the ball $u_1^2 + u_2^2 \leq 1$.

We will now consider the time optimal problem in the reachable group G in $SU(8)$ associated with the above convex system.

We remind the reader that $\langle\, ,\, \rangle$ is the scalar product on $\mathfrak{su}(8)$ given by $\langle A, B\rangle = -\frac{1}{2}Tr(AB)$. This scalar product is a multiple of the Killing form and hence satisfies $\langle [A, B], C\rangle = \langle A, [B, C]\rangle$ for any matrices A, B, C in $\mathfrak{su}(8)$. Relative to $\langle\, ,\, \rangle$ matrices $U_1, U_2, U_3, V_1, V_2, V_3$ constitute an orthonormal basis. Then G with the left-invariant metric induced by the above scalar product becomes a Riemannian manifold as well as a sub-Riemannian manifold with the sub-Riemannian length defined over the horizontal curves by

$$\int_0^T \|u_1(t)(V_3 + kU_3) + u_2(t)(V_2 + kU_2)\|\,dt = \sqrt{1 + k^2} \int_0^T \sqrt{u_1^2(t) + u_2^2(t)}\,dt.$$

Thus a horizontal curve $g(t)$ that connects $g_0 = I$ to a point $g_1 \in G$ in T units of time is a curve of minimal length if and only if $\int_0^T \sqrt{u_1^2(t) + u_2^2(t)}\,dt$ is minimal. As expected the non-stationary time optimal horizontal curves coincide with the sub-Riemannian geodesics of shortest length.

The sub-Riemannian metric induces a Riemannian metric on the quotient space $M = G/K_v$ with the geodescs on M equal the projections of the sub-Riemannian geodescs in G that connect the initial coset K_v to the terminal coset g_1K_v. It is important to note that the above sub-Riemannian metric is not of contact type, that is, $[\Gamma, \Gamma] \neq \mathcal{L}$ where Γ denotes the vector space spanned by $V_2 + kU_2$ and $V_3 + kU_3$. Instead,

$$\Gamma + [\Gamma, \Gamma + [\Gamma, [\Gamma, \Gamma]] = \mathcal{L}, k \neq 1.$$

Secondly, it may be important to note that the induced metric on G/K_v is not symmetric.

Let us now use the maximum principle to get the extremal curves associated with the above time optimal problem.

5.5. The extremal curves

We will follow the formalism outlined in Section 3, in which the cotangent bundle T^*G is trivialized by the left-translations and represented as $G \times \mathfrak{g}^*$, where $\mathfrak{g}^*$ denotes the dual of $\mathfrak{g}$., Then $\mathfrak{g}^*$ will be identified with $\mathfrak{g}$ via $\langle\, ,\, \rangle$ with $\ell \in \mathfrak{g}^*$ identified with $L \in \mathfrak{g}$ through the formula $\langle L, X\rangle = \ell(X)$ for any $X \in \mathfrak{g}$. Every $L \in \mathfrak{g}$ admits a representation $L = \sum_{i=1}^3 P_i B_i + M_i A_i$ where $P_i = \ell(B_i)$ and $M_i = \ell(A_i)$.

Then the Hamiltonian lift of the horizontal system (5.26) is given by

$$H(\ell) = \ell((V_3 + kU_3)u_1 + (V_2 + kU_2)u_2) = \langle L, (V_3 + kU_3)u_1 + (V_2 + kU_2)u_2\rangle =$$
$$(P_3 + kM_3)u_1 + (P_2 + kM_2)u_2,$$

where $P_i = \langle L, V_i\rangle$, and $M_i = \langle L, U_i\rangle$, $i = 1, 2, 3$.

We recall that the Hamiltonian equations associated with H are given by the equations

$$\frac{dg}{dt} = g(t)dH_{\ell(t)}, \frac{d\ell}{dt}(t) = -ad^*(dH_{\ell(t)})(\ell(t))$$

where $dH = (V_3 + kU_3)u_1(t) + (V_2 + kU_2)u_2(t)$, or, dually by $\frac{dL}{dt} = [dH, L]$. In the coordinates, P_i, M_i the preceding equations take on the following form

$$\begin{aligned} \dot{M}_1 &= (P_2 + kM_2)u_1 - (P_3 + kM_3)u_2, \\ \dot{M}_2 &= -(P_1 + kM_1)u_1, \\ \dot{M}_3 &= (P_1 + kM_1)u_2, \\ \dot{P}_1 &= (M_2 + kP_2)u_1 - (M_3 + kP_3)u_2, \\ \dot{P}_2 &= -(M_1 + kP_1)u_1, \\ \dot{P}_3 &= (M_1 + kP_1)u_2. \end{aligned} \tag{5.27}$$

According to the maximum principle time optimal trajectories are the projections of the extremal curves which can be abnormal and normal. In the abnormal case the maximum principle results in the constraints

$$P_2(t) + kM_2(t) = 0, P_3(t) + kM_3(t) = 0, \tag{5.28}$$

while in the normal case the maximum principle singles out the Hamiltonian

$$H = \frac{1}{2}(P_3 + kM_3)^2 + (P_2 + kM_2)^2,$$

generated by the extremal controls $u_1 = P_3 + kM_3$, $u_2 = P_2 + kM_2$, whose integral curves on energy level $H = \frac{1}{2}$ coincide with the normal extremal curves. Let us begin with the abnormal extremals.

Proposition 27. *Abnormal extremal curves associated with the time optimal curves $g(t)$ are generated by the controls*

$$u_1(t) = c_1 \cos\omega t + c_2 \sin\omega t, u_2(t) = c_1 \sin\omega t - c_2 \cos\omega t, c_1^2 + c_2^2 = 1$$

and are confined to the manifold

$$P_2(t) + kM_2(t) = P_3 + kM_3(t) = M_1(t) + kP_1(t) + k(P_1(t) + kM_1(t)) = 0.$$

In addition, $M_1(t)$ and $P_1(t)$ are constant. On $M_1 = 0$, both u_1 and u_2 are constant, hence $g(t)$ is a Riemannian geodesic in G.

Proof. As stated above, abnormal extremal curves satisfy

$$P_2(t) + kM_2(t) = 0, P_3(t) + kM_3(t) = 0,$$

and when they correspond to a time optimal curve, then they satisfy another constraint, known as the Goh condition, namely

$$\{P_2 + kM_2, P_3 + kM_3\} = 0,$$

which yields

$$M_1 + kP_1 + k(P_1 + kM_1) = 0. \tag{5.29}$$

Since $\dot{M}_1 = \{H, M_1\} = (P_2 + kM_2)u_1 - (P_3 + kM_3)u_2 = 0$, M_1 is constant, and hence P_1 must be constant also.

Upon differentiating (5.29) along the extremal curve we get

$$2k(-(M_2 + kP_2)u_1 + (M_3 + kP_3)u_2) = 0,$$

which implies that

$$u_1(t) = M_3(t) + kP_3(t), u_2(t) = M_2(t) + kP_2(t),$$

since time optimality demands that $u_1^2 + u_2^2 = 1$ whenever $u \neq 0$. Then

$$\dot{u}_1(t) = \dot{M}_3(t) + k\dot{P}_3(t) = -(P_1 + kM_1 + k(M_1 + kP_1))u_2(t)) = -\omega\, u_2(t),$$
$$\dot{u}_2(t) = \dot{M}_2(t) + k\dot{P}_2(t) = (P_1 + kM_1 + k(M_1 + kP_1))u_1(t)) = \omega\, u_1(t),$$

hence

$$u_1(t) = c_1 \cos\omega t + c_2 \sin\omega t, u_2(t) = c_1 \sin\omega t - c_2 \cos\omega t.$$

On $M_1 = 0$, $P_1 = 0$, and $\omega = 0$.

We now come to the normal extremals. Let us first note that the Poisson equation $\frac{dL}{dt} = [dH, L]$ that governs the normal extremals is completely integrable on each coadjoint orbit in $\mathfrak{so}(4)$ for the following reasons: $\mathfrak{so}(4)$ is of rank two, and hence admits two universal conservation laws (Casimirs)

$$I_1 = ||M||^2 + ||P||^2,\ I_2 = M_1P_1 + M_2P_2 + M_3P_3.$$

Therefore, generic coadjoint orbits are four dimensional, and since coadjoint orbits are symplectic, they admit at most two independent integrals of motion functionally independent from the Casimirs. In the present case, $I_3 = M_1$ and $H = \frac{1}{2}(P_2 + kM_2)^2 + (P_3 + kM_3)^2)$ are the required integrals. The fact that M_1 is constant was clear from the very beginning since $K_v = \{e^{\varepsilon U_1}, \varepsilon \in \mathbb{R}\}$ is a symmetry for (5.26).

We will now show that the normal extremals can be integrated by quadrature in terms of elliptic functions on the manifold

$$c_1 = 2(H - 2kI_2), c_2 = M_1, c_3 = I_1 - M_1^2, c_4 = I_2$$

Then,

$$c_1 = 2(H - kI_2) = (P_2 + kM_2)^2 + (P_3 + kM_3)^2 - 2k(P_1M_1 + P_2M_2 + P_3M_3) =$$
$$P_2^2 + P_3^2 + k^2(M_2^2 + M_3^2) - 2kP_1M_1,\ \text{and}$$

$$P_2^2 + P_3^2 + M_2^2 + M_3^2 = I_1 - P_1^2 - M_1^2 = c_3 - P_1^2.$$

It follows that

$$(1-k^2)(P_2^2 + P_3^2) = c_1 + 2kP_1c_2 - k^2(c_3 - P_1^2) = c_1 - k^2c_3 - 2kc_2P_1 + k^2P_1^2,$$
$$(1-k^2)(M_2^2 + M_3^2) = c_3 - P_1^2 - (c_1 + 2kP_1c_2) = c_3 - c_1 - 2kc_2P_1 - P_1^2.$$

We now have

$$\frac{1}{(1-k^2)^2}(\tfrac{dP_1}{dt})^2 = (P_2M_3 - P_3M_2)^2 = P_2^2M_3^2 + P_3^2M_2^2 - 2P_2P_3M_2M_3 =$$
$$P_2^2M_3^2 + P_3^2M_2^2 - (P_2M_2 + P_3M_3)^2 + P_2^2M_2^2 + P_3^2M_3^2 =$$
$$(P_2^2 + P_3^2)(M_2^2 + M_3^2) - (I_2 - P_1M_1)^2 =$$
$$\frac{1}{(1-k^2)^2}(c_1 - k^2c_3 + 2kc_2P_1M_1 + k^2P_1^2))(c_3 - c_1 - 2kc_2P_1 - P_1^2) - (I_2 - P_1M_1)^2.$$

Hence,

$$(\tfrac{dP_1}{dt})^2 = (c_1 - k^2c_3 + 2kc_2P_1 + k^2P_1^2))(c_3 - c_1 - 2kc_2P_1) - P_1^2) - (1-k^2)^2(I_1 - M_1P_1)^2$$
$$= -k^2P_1^4 - 2kM_1P_1^3(k^2+1) + \alpha P_1^2 + \beta P_1 + \gamma,$$

where

$$\alpha = 2k^2c_3 - c_1(1+k^2) - 4k^2M_1^2 - (1-k^2)^2M_1^2I_1^2,$$
$$\beta = (2kc_2(k^2+1)c_3 - 2c_1) + 2(1-k^2)^2c_4c_2, \gamma = (c_1 - k^2c_3)(c_3 - c_1) - (1-k^2)^2c_4^2.$$

It is well known that the solutions of $\frac{dz}{dt} = \sqrt{P(z)}$ with P a fourth degree polynomial can be solved in terms of elliptic integrals (for instance, see ([20])).

The remaining variables can be integrated by quadrature through the representation

$$u_1(t) = \cos\theta(t), u_2(t) = \sin\theta(t). \tag{5.30}$$

Then

$$-u_2(t)\dot{\theta}(t) = \dot{u}_1(t) = \dot{P}_3 + k\dot{M}_3 = -(M_1 + kP_1) + k(P_1 + kM_1))u_2(t)$$

yields

$$\theta(t) = \theta(0) - \int_0^t (c_2(1+k^2) + 2kP_1(s))\,ds. \tag{5.31}$$

Hence the extremal controls are now specified and the projected curve $g(t)$ is obtained as a solution of a fixed ordinary differential equation.

In the presence of the transversality conditions, $M_1 = 0$, and the above equation simplifies. For when $M_1 = 0$,

$$(\frac{dP_1}{dt})^2 = -k^2P_1^4 + \alpha P_1^2 + \gamma$$

Then $\xi = P_1^2$ is a solution of

$$(\frac{1}{2}\frac{d\xi}{dt})^2 = P_1^2(\frac{dP_1}{dt})^2 = -k^2\xi^3 + \alpha\xi^2 + \gamma\xi. \tag{5.32}$$

The preceding equation can be put in its canonical form $\frac{d\xi}{dt} = \sqrt{4\xi^3 - g_2\xi - g_1}$ and then can be solved in terms of the Weierstrass' $\wp$ function ([7]), page 113).

The solutions of Yuan's optimal problem satisfy additional transversality conditions, namely, the extremal curve $L(t)$ is orthogonal to $\mathfrak{k}$ at the initial and the terminal time, where $\mathfrak{k}$ is the Lie algebra spanned by U_1, U_2, U_3. That means that $M_i(0) = 0$ and $M_i(T) = 0$ for $i = 1, 2, 3$. Such extremal curves reside on $I_2 = 0$.

Use of AI tools declaration

The author declares not having used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

I am grateful to Fatima Silva Leite for her constructive criticisms of an earlier version of the paper as well as for her help with various technical requirements imposed by the publisher.

Conflict of interest

The author declares there is no conflict of interest.

References

1. R. Brockett, S. J. Glaser, N. Khaneja, Time optimal control in spin systems, *Phys. Rev. A*, **63** (2001), 032308. https://doi.org/10.1103/PhysRevA.63.032308
2. N. Khaneja, S. J. Glazer, R. Brockett, Sub-Riemannian geometry and time optimal control of three spin systems: quantum gates and coherence transfer, *Phys. Rev. A*, **65** (2002), 032301. https://doi.org/10.1103/PhysRevA.65.032301
3. H. Yuan, Geometry, Optimal Control and Quantum Computing, Ph.D thesis, Harvard, 2006.
4. B. Bonnard, T. Combot, L. Jassionnesse, Integrability methods in the time minimal coherence transfer for Ising chains of three spins, *Discrete Contin. Dyn. Syst.*, **35** (2015), 4095–4114. https://doi.org/10.3934/dcds.2015.35.4095
5. P. Eberlein, Geometry of nonpositively curved manifolds. Chicago Lectures in Mathematics, The University of Chicago Press, Chicago, 1996.
6. S. Helgason, *Differential Geometry, Lie Groups and Symmetric Spaces*, Academic Press, New York, 1978.
7. V. Jurdjevic, *Optimal Control and Geometry: Integrable systems*, Cambridge University Press, Cambridge Studies in Advanced Mathematics, (2016), Cambridge, UK. https://doi.org/10.1017/CBO9781316286852
8. A. Agrachev, T. Chambrion, An estimation of the controllability time for single input time on compact Lie groups, *ESAIM: Control, Optimization and Calculus of Variations*, **12** (2006), 409–441. https://doi.org/10.1051/cocv:2006007

9. J. P. Gauthier, F. Rossi, A universal gap for non-spin quantum control systems, *Proc. Amer. Math. Soc.*, **149** (2021), 1203–1214. https://doi.org/10.1090/proc/15301
10. F. Albertini, D. D'Alessandro, B. Sheller, Sub-Riemannian geodesics in $SU(n)/S(U(n-1)\times U(1))$ and Optimal Control of Three Level Quantum Systems, *IEEE Trans. Automat. Contr.*, **65** (2020), 1176–1191. https://doi.org/10.1109/TAC.2019.2950559
11. V. Jurdjevic, *Geometric Control Theory*, Cambridge University Press, Cambridge Studies in Advanced Mathematics, New York, 1987. https://doi.org/10.1017/CBO9780511530036
12. A. A. Kirillov, *Lectures on the orbit method.*, Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2004. https://doi.org/10.1090/gsm/064
13. R. Montgomery Isoholonomic Problems and Some Applications, *Commun. Math. Phys*, **128** (1990), 565–592. https://doi.org/10.1007/BF02096874
14. V. Jurdjevic, H. J. Sussmann, Control systems on Lie groups, *Jour. Diff. Equations*, **12** (1972), 313–329. https://doi.org/10.1016/0022-0396(72)90035-6
15. F. Wellstood, Notes about Bloch sphere, Lecture notes, 2016, University of Maryland.
16. N. Khaneja, S. J. Glaser, Efficient transfer of coherence through Ising spin chains, *Phys. Rev. A*, **66** (2002), 060301. https://doi.org/10.1103/PhysRevA.66.060301
17. H. Yuan, S. J. Glaser, N. Khaneja, Geodesics for efficient creation and propagation of order along Ising spin chains, *Phys. Rev. A*, **76** (2007), 012316. https://doi.org/10.1103/PhysRevA.76.012316
18. U. Boscain, T. Chambrion, J. P. Gauthier, On the K+P problem for a three-level quantum system: optimality implies resonance, *J. Dynam. Control Systems*, **8** (2002), 547–572. https://doi.org/10.1023/A:1020767419671
19. H. Yuan, R. Zeier, N. Khaneja, Elliptic functions and efficient control of Ising spin chains with unequal couplings, *Phys. Rev. A*, **77** (2008), 032340. https://doi.org/10.1103/PhysRevA.77.032340
20. R. Appell, E. Lacour Principes de la théorie des fonctions elliptiques, Gauthier Villars (1922).

AIMS Press