

## AN ESTIMATION OF THE CONTROLLABILITY TIME FOR SINGLE-INPUT SYSTEMS ON COMPACT LIE GROUPS \*

ANDREI AGRACHEV<sup>1</sup> AND THOMAS CHAMBRION<sup>2</sup>

**Abstract.** Geometric control theory and Riemannian techniques are used to describe the reachable set at time  $t$  of left invariant single-input control systems on semi-simple compact Lie groups and to estimate the minimal time needed to reach any point from identity. This method provides an effective way to give an upper and a lower bound for the minimal time needed to transfer a controlled quantum system with a drift from a given initial position to a given final position. The bounds include diameters of the flag manifolds; the latter are also explicitly computed in the paper.

**Mathematics Subject Classification.** 22E46, 93B03.

Received September 4, 2004. Revised May 8, 2005.

### 1. INTRODUCTION

#### 1.1. Physical context

Some recent papers [9–12, 14, 15, 22, 32] have studied the problem of transferring a population of quantum particles from a quantum state to another, using an external field (laser). In many applications (*e.g.*, MNR) a crucial issue is to achieve this quantum transfer in the shortest possible time, in order to avoid relaxation effects. In mathematical terms, the problem is to control the evolution of the solution of the  $n$ -dimensional Schrödinger equation:

$$i \frac{d\psi}{dt} = H\psi(t), \quad (1.1)$$

where  $\psi : \mathbb{R} \rightarrow \mathbb{C}^n$  is the wave function and  $H : \mathbb{C}^n \rightarrow \mathbb{C}^n$  a (time varying) Hermitian linear operator with vanishing trace. We will restrict ourselves to the case where  $H$  has the form  $H = h_0 + uh_1$  where  $h_0$  and  $h_1$  are two fixed hermitian linear operators of  $\mathbb{C}^n$  with vanishing trace, and  $u$  is a (one dimensional) *a priori* not bounded real control. The aim is to transfer the system from a given initial situation  $\psi_{init}$  to a given final situation  $\psi_{final}$  minimizing the time of transfer. The fact that  $u$  is a real function (not a complex one) means that only the amplitude (not the phase) of the laser can be controlled. The hypothesis that  $u$  is not *a priori* bounded can be regarded as a limit case where the amplitude of the laser can be made arbitrarily large with respect to the energy constants of the system.

---

*Keywords and phrases.* Control systems, semi-simple Lie groups, Riemannian geometry.

\* *Second author was partially supported by Marie Curie Fellowships HPMT-GH-01-00278-11 and HPMT-GH-01-00278-47.*

<sup>1</sup> SISSA, Via Beirouth 2-4, 34013 Trieste, Italy; agrachev@sissa.it

<sup>2</sup> SYSTeMS Group, University of Ghent, Technologiepark 914, 9052 Zwijnaarde, Belgium; Thomas.Chambrion@UGent.be

© EDP Sciences, SMAI 2006

A very classical and efficient way to study linear differential equations is to consider the so-called resolvent of the problem, that is to say to lift the problem from  $\mathbb{C}^n$  to a matrix group of linear transformations of  $\mathbb{C}^n$ . In the case of the Schrödinger equation with the special form of  $H$ ,  $H = h_0 + uh_1$ , this leads to a control-affine system *i.e.*, with drift, on  $SU(n)$ .

In [32], systems with a drift and  $k$  real *a priori* unbounded controls (that is to say, when  $H$  takes the form  $H = h_d + u_1h_1 + \dots + u_kh_k$ ) have been extensively studied, under the extra hypothesis that the Lie algebra generated by  $h_1, \dots, h_k$  is a special part of a Cartan decomposition of  $\mathfrak{su}(n)$ . In our case, when  $k = 1$  and  $n \geq 3$ , this property is never verified and this method cannot be applied.

In this paper, we apply classical techniques of geometric control theory to describe the so-called *attainable set* (see precise definitions below) of the control system. This description and the use of more or less classical Riemannian techniques allow us to estimate the time needed to transfer a population from any initial state to any final state using only one unbounded real control. Although we are primarily interested in quantum control problems modeled with control systems on  $SU(n)$ , it appears that the methods we develop here are also valid in any semi-simple compact real Lie group and hence can be applied to many other fields, such as electrical networks or robotic systems, modeled with control systems on  $SO(n)$ .

**Remark 1.1.** When dealing with invariant control problems on Lie groups, it is usual to treat only the *left* invariant case (the right invariant case is very similar, it suffices to replace “ $L$ ” by “ $R$ ” in most of the statements to obtain equivalent results, see [3] for further discussion). On the other hand, the lifting of the classical Schrödinger equation (1.1) with  $n$  levels gives a *right* invariant system on  $SU(n)$ . Since most of the methods presented here can be applied in a very wide range of problems on semi-simple compact Lie groups, we decided to keep the left invariant presentation. Obviously, with minor modification, all results remain true for right invariant problems, just replacing formally each “ $L$ ” translation by an “ $R$ ”.

## 1.2. Statement of the problem

The dynamic is governed by the following equation: let  $h_0$  and  $h_1$  be two elements of the compact semi-simple Lie algebra  $\mathfrak{u}$  of the semi-simple compact Lie group  $U$ , with dimension  $N_{\mathfrak{u}}$  and rank  $r_{\mathfrak{u}}$ . Consider the control problem on  $U$  governed by the following (left invariant) differential system  $\Sigma_1$

$$(\Sigma_1) \begin{cases} \dot{g} &= dL_g(h_0 + u h_1) \\ g(0) &= Id, \end{cases}$$

where  $u : \mathbb{R}^+ \rightarrow \mathbb{R}$  is any  $L^\infty$  function. First we recall some basic definitions. To every time  $t$  greater than 0, we can associate the corresponding *attainable set*.

**Definition 1.2.** A point  $g$  of  $U$  is called *attainable* from identity in time exactly  $t$  for  $\Sigma_1$  if and only if there exists some control function  $u \in L^\infty(\mathbb{R}^+, \mathbb{R})$  such that the corresponding solution  $\Phi_{\Sigma_1}^u : [0, t] \rightarrow U$  of  $\Sigma_1$  verifies

$$\Phi_{\Sigma_1}^u(0) = Id_U \text{ and } \Phi_{\Sigma_1}^u(t) = g.$$

**Definition 1.3.** The set  $a_{\Sigma_1}(t)$  of all attainable points from identity in time exactly  $t$  for  $\Sigma_1$  is called *the attainable set in time exactly  $t$*  for  $\Sigma_1$ .

**Definition 1.4.** The set  $\mathcal{A}_{\Sigma_1}(t) = \bigcup_{0 \leq s \leq t} a_{\Sigma_1}(s)$  is called *the attainable set in time  $t$*  for  $\Sigma_1$ .

Many results about controllability of systems of the type of  $\Sigma_1$  (sometimes called “invariant single-input nonhomogeneous” systems) have been obtained in the last thirty years [2, 5–8, 13, 17, 20, 21, 23, 25, 26, 30, 33]. The following corollary of the orbit theorem has been known for some time (see [37, 38] for a proof and further discussion).

**Theorem 1.5.**  $\mathcal{A}_{\Sigma_1}(t) = U$  for  $t$  large enough if and only if the Lie algebra generated by  $h_0$  and  $h_1$  is equal to  $\mathfrak{u}$ . This condition on  $h_0$  and  $h_1$  is open and dense.

**Remark 1.6.** If the semi-simple real Lie group  $U$  is no more assumed to be compact, the situation is more complicated. El Assoudi, Gauthier and Kupka (1996) have given explicit (but quite technical) sufficient conditions for controllability. For precise assumptions on  $h_0$  and  $h_1$  in this rather difficult result, please refer to [37] and to the papers [18, 19].

The problem is to find  $t_0 = \inf \mathcal{T}_{\Sigma_1}$ , where  $\mathcal{T}_{\Sigma_1} = \{t > 0; \overline{\mathcal{A}_{\Sigma_1}(t)} = U\}$ . In other words, what is the minimal time needed to go from  $Id$  to (any neighborhood of) any other point using the system  $\Sigma_1$ ? This time  $t_0$  will be called the *diameter of  $U$  under  $\Sigma_1$*  and will be denoted by  $\mathcal{D}(U, \Sigma_1)$ .

**Remark 1.7.** Because of the left invariance of the system, the assumption that the starting point is the identity of the group is absolutely not restrictive. Actually, if a trajectory  $t \mapsto \Phi_{\Sigma_1}^u(t)$  verifies  $\Phi_{\Sigma_1}^u(0) = Id_U$  and  $\Phi_{\Sigma_1}^u(t_0) = g$ , then for any  $g_0$  in  $U$ , the curve  $t \mapsto g_0 \Phi_{\Sigma_1}^u(t)$  is a solution of the equation  $\dot{g} = dL_g(h_0 + u h_1)$  satisfying  $g_0 \Phi_{\Sigma_1}^u(0) = g_0$  and  $dL_{g_0} \dot{\Phi}_{\Sigma_1}^u(t_0) = dL_{g_0} \dot{g}$ .

**Definition 1.8.** Given a point  $x_0$  of a metric space  $(M, d_M)$ , we call *diameter of  $M$  for the metric  $d_M$  from the point  $x_0$*  and denote by  $\mathcal{D}(M, d_M, x_0)$  the supremum, if it exists, of the function  $d_M^{x_0}$  defined by

$$\begin{aligned} d_M^{x_0} : M &\rightarrow \mathbb{R} \\ x &\mapsto d_M(x, x_0). \end{aligned}$$

In many cases, the diameter  $\mathcal{D}(M, d_M, x_0)$  will not depend on the point  $x_0$ . By abuse, in this case, we define the *diameter of  $M$  for  $d_M$*  and note  $\mathcal{D}(M, d_M)$ . This notation is consistent with the preceding definition of  $\mathcal{D}(U, \Sigma_1)$ .

### 1.3. Scope of the paper

We are not able to compute exactly the diameter  $\mathcal{D}(U, \Sigma_1)$ . However, by comparison with the diameter of  $U$  and certain quotient spaces  $U/T$  for some canonical metrics, we give an effective and explicit estimation of  $\mathcal{D}(U, \Sigma_1)$ . Actually, one of the main task will be to compute the diameter  $\mathcal{D}(U/T, d_T)$ , where  $T$  is a maximal torus of  $U$ , and  $d_T$  the canonical projection of the Killing metric of  $U$  (see precise definitions below).

In this paper, we shall give:

- (1) (for generic controllable control systems in any semi-simple compact Lie group) an effective way to construct an extension of a left invariant single input control system;
- (2) (for generic controllable control systems in any semi-simple compact Lie group) an easily computable non trivial upper bound for the diameter;
- (3) (for the four series of classical compact groups of matrices, for generic controllable control systems only) an estimation of the diameter  $\mathcal{D}(U/T, d_T)$ , from which we deduce more accurate easily computable upper and lower bounds for the diameter  $\mathcal{D}(U, \Sigma_1)$ ;
- (4) an explicit computation of the diameter  $\mathcal{D}(SU(n)/T, d_T)$ ;
- (5) (in  $SU(n)$ , for generic controllable control systems only) a quite good approximation of the system by a sequence of Riemannian problems, the expression of the associated Ricci tensors and an upper bound for the diameter  $\mathcal{D}(SU(n), \Sigma_1)$  using Index theory and Ricci curvature.

### 1.4. Content of the paper

The paper is organized as follows. In Section 2, we give precise statements of the main results of the paper. For this task, we need to recall some well known properties of semi-simple Lie algebras. In Section 3, we give a general method to construct an equivalent system  $\Sigma_{EQ}$  whose closure of the attainable set at each instant is the same as the one of  $\Sigma_1$ . In Section 4, we approximate the system obtained in Section 3 by a Riemannian problem on the quotient of a semi-simple compact Lie group by a maximal torus, endowed with the metric induced by the standard (bi-invariant) Killing form. Complete computations are done in the case where  $U = SU(n)$ .

In Section 5, we define a sequence of (only left invariant) Riemannian systems close to the system obtained in Section 3 and we use the special structure of  $\mathfrak{su}(n)$  to apply well-known Riemannian techniques (such as relation between Ricci curvature and the length of optimal geodesics) to get an upper bound for the diameter in the case where  $U = \text{SU}(n)$ . In Section 6, the methods and the results presented in the paper are illustrated by an example in  $\text{SU}(3)$ .

In Appendix A, we recall some basic properties of compact semi-simple Lie algebras. In Appendix B, we give some computation rules valid only in  $\mathfrak{su}(n)$ . These rules allow us to compute the solutions of a certain class of equations on  $\mathfrak{su}(n)$  in Appendix C.

## 2. MAIN RESULTS

### 2.1. Compact Subalgebra of real semi-simple algebra

The aim of this section is to recall some basic facts about Lie algebras and to fix the notations. A more detailed presentation of these properties and the basic facts about the structure of semi-simple Lie algebras are given in Appendix A. An effective decomposition of the Lie algebra  $\mathfrak{su}(3)$  is presented in Section 6.1.

Let  $\mathfrak{u}$  be a real semi-simple compact Lie algebra, *i.e.*, a real Lie algebra with a strictly negative definite Killing form, see [4, 24, 39]. We assume that  $U$  has dimension  $N_{\mathfrak{u}}$  and rank  $r_{\mathfrak{u}}$ . In the case where  $U = \text{SU}(n)$ , one has  $N_{\mathfrak{u}} = n^2 - 1$  and  $r_{\mathfrak{u}} = n - 1$ . We denote by  $\langle \cdot, \cdot \rangle$  the opposite of the Killing form on  $\mathfrak{u}$ . Let  $\mathfrak{g}$  be the complexification of  $\mathfrak{u}$ . Then  $\mathfrak{g}$  is a (complex) semi-simple Lie algebra, because of the semi-simplicity of  $\mathfrak{u}$ . If  $h_1$  is a regular element of  $\mathfrak{g}$ , that is, if  $\text{ad}_{h_1}$  has maximal rank among all  $\text{ad}_h$  for  $h$  in  $\mathfrak{g}$ , then  $\ker(\text{ad}_{h_1})$  is a Cartan subalgebra  $\mathfrak{h}$  of  $\mathfrak{g}$ . The root space decomposition of  $\mathfrak{g}$  with respect to  $\mathfrak{h}$  writes

$$\mathfrak{g} = \mathfrak{h} + \sum_{\alpha \in \Delta} \mathfrak{g}^{\alpha} + \sum_{\alpha \in \Delta} \mathfrak{g}^{-\alpha},$$

where  $\Delta$  is the set of the positive roots of  $\mathfrak{g}$  (with respect to the Cartan subalgebra  $\mathfrak{h}$  and an arbitrary order) and  $\mathfrak{g}^{\alpha}$  represents the root space associated to the root  $\alpha$ . The Killing form being not degenerate over  $\mathfrak{h}$ , it provides an isomorphism between  $\mathfrak{h}$  and its dual space  $\mathfrak{h}^*$ . Thus, to each root  $\alpha$  in  $\Delta$ , we can associate  $H_{\alpha} \in \mathfrak{h}$ . For all (positive or negative) root  $\alpha$ , it is possible to choose  $X_{\alpha}$  in  $\mathfrak{g}^{\alpha}$  such that, for all root  $\beta$

$$\begin{aligned} [X_{\alpha}, X_{-\alpha}] &= H_{\alpha}, \\ [H, X_{\alpha}] &= \alpha(H)X_{\alpha} \text{ for any } H \in \mathfrak{h}, \\ [X_{\alpha}, X_{\beta}] &= 0 \text{ if } \alpha + \beta \neq 0 \text{ and } \alpha + \beta \notin \Delta, \\ [X_{\alpha}, X_{\beta}] &= N_{\alpha, \beta}X_{\alpha + \beta} \text{ if } \alpha + \beta \in \Delta, \end{aligned}$$

where  $N_{\alpha, \beta}$  are some constants. Applying this notation, a compact real form  $\mathfrak{g}_k$  of  $\mathfrak{g}$  is given by (see [24], Th. 6.3, Chap. 3)

$$\mathfrak{g}_k = \sum_{\alpha \in \Gamma} \mathbb{R}(iH_{\alpha}) + \sum_{\alpha \in \Delta} \mathbb{R}(X_{\alpha} - X_{-\alpha}) + \sum_{\alpha \in \Delta} \mathbb{R}(i(X_{\alpha} + X_{-\alpha})), \tag{2.1}$$

where  $\Gamma$  is the set of the simple roots of  $\Delta$ .

Up to an inner isomorphism, we may assume without loss of generality that  $\mathfrak{g}_k = \mathfrak{u}$ . In the following, we define  $\mathfrak{t}$  to be the real vector space spanned by vectors  $iH_{\alpha}$ , and we note  $\partial_{\alpha} = X_{\alpha} - X_{-\alpha}$ ,  $\partial_{\overline{\alpha}} = i(X_{\alpha} + X_{-\alpha})$  for every  $\alpha$  in  $\Delta$  and  $h_{\alpha} = iH_{\alpha}$  for every  $\alpha$  in  $\Gamma$ . We get a basis  $\mathcal{B}$  of  $\mathfrak{u}$  constituted by vectors in three subfamilies: the  $\mathcal{B}_{\mathfrak{t}}$ -subfamily, containing the vectors of the form  $h_{\alpha}$  for  $\alpha$  in  $\Gamma$ , the  $\mathcal{B}_{\Delta}$ -subfamily containing vectors of the form  $\partial_{\alpha}$ , and the  $\mathcal{B}_{\overline{\Delta}}$ -subfamily containing vector of the form  $\partial_{\overline{\alpha}}$  for  $\alpha$  in  $\Delta$ . Union  $\mathcal{B}_{\Delta} \cup \mathcal{B}_{\overline{\Delta}}$  is denoted by  $\mathcal{B}_{\perp}$ .

**Remark 2.1.** When the numbers  $N_{\alpha, \beta}$  are integers, the basis  $\mathcal{B}$  of  $\mathfrak{u}$  is sometimes called *Weyl-Chevalley basis* of  $\mathfrak{u}$ . Existence of such basis for semi-simple Lie algebras is essential in the classification of semi-simple Lie algebras. Actually, the Dynkin diagram of a semi-simple Lie algebra can be deduced from the angles between the  $H_{\alpha}$  vectors appearing in basis  $\mathcal{B}$ .

Some technical computations in the following will be shorted by use of the following *non standard* definitions.

**Definition 2.2.** For any  $\alpha$  in  $\Delta$ , the *root modulus* of  $\alpha$  or  $\bar{\alpha}$  is  $[\alpha] = [\bar{\alpha}] = \alpha$ . By extension, for any  $\alpha$  in  $\Delta$ , the *modulus of a vector*  $\partial_\alpha$  or  $\partial_{\bar{\alpha}}$  of  $\mathcal{B}_{\mathfrak{t}^\perp}$  is the positive root attached to it. For any  $\alpha$  in  $\Delta$ ,  $[\partial_\alpha] = [\partial_{\bar{\alpha}}] = \alpha$ .

**Remark 2.3.** This modulus has not to be confused with the usual complex modulus  $|\cdot|$  or with any norm on  $\mathfrak{u}$  or  $\mathfrak{t}^*$ .

We have already introduced the “ $\bar{\cdot}$ ” notation. We define  $\bar{\Delta} = \{\bar{\alpha}, \alpha \in \Delta\}$ , the set of the barred roots. In the following, it will be convenient to use the double bar notation: for every root  $\alpha$  in  $\Delta$ , we define  $\bar{\bar{\alpha}} = \alpha$ .

The basis  $\mathcal{B}$  is usually not orthonormal for the usual Killing form, because the vectors  $h_\alpha$  are usually pairwise not orthogonal, and because vectors  $\partial_\alpha$  and  $\partial_{\bar{\alpha}}$  do not have norm 1. Nevertheless, the  $\mathcal{B}_{\mathfrak{t}}$ -subfamily, the  $\mathcal{B}_{\Delta}$ -subfamily and the  $\mathcal{B}_{\bar{\Delta}}$ -subfamily are always pairwise orthogonal for the usual Killing form.

In the following, we denote by  $d_\alpha^0$ ,  $a_\alpha^0$  and  $a_{\bar{\alpha}}^0$  the coordinates of  $h_0$  in the basis  $\mathcal{B}$ .

$$h_0 = \sum_{\alpha \in \Gamma} d_\alpha^0 h_\alpha + \sum_{\alpha \in \Delta} a_\alpha^0 \partial_\alpha + \sum_{\alpha \in \Delta} a_{\bar{\alpha}}^0 \partial_{\bar{\alpha}}. \quad (2.2)$$

The classical exponential mapping provides a useful link between a Lie group and its Lie algebra. We denote by  $T = \{\exp(h), h \in \mathfrak{t}\}$  the image set of  $\mathfrak{t}$  under this application.

We use the following *non standard* definition to characterize vectors generating a maximal torus.

**Definition 2.4.** A vector  $h$  in  $\mathfrak{u}$  is called *very strongly regular* if  $h$  is regular and if the one parameter sub-group  $\{\exp(t h), t \in \mathbb{R}\}$  is dense in the subgroup  $T_h = \exp(\mathfrak{t}_h)$  where  $\mathfrak{t}_h$  is the Cartan subalgebra defined by  $h$ .

The Killing form on  $\mathfrak{u}$  provides a natural Euclidian structure on  $\mathfrak{u}$  and also the standard associated measure. We use this measure to explain in which sense the very strong regularity is a generic property.

**Proposition 2.5.** *The set  $\mathcal{VSR}$  of the very strongly regular elements of  $\mathfrak{u}$  is residual with full measure, hence dense, but not open.*

*Proof.* It is a classical fact (see [24]) that the set  $\mathcal{Reg}$  of the regular elements of a semi-simple Lie algebra is open with full measure. For any regular  $h$ ,  $\exp: (\mathfrak{t}_h, +) \rightarrow (T_h, \cdot)$  is a homomorphism between abelian groups, with discrete kernel  $\mathbf{k}$ , see [1].  $\mathbf{k}$  is a free Abelian group, and

$$T_h \sim \sum_{i=1}^n g_i (\mathbb{R}/\mathbb{Z}),$$

where  $g_1, g_2, \dots, g_n$  are  $\mathbb{R}$  (linearly independent) generators of  $\mathbf{k}$  and  $\sim$  denotes a group isomorphism. This is a classical way to show that  $T_h$  is actually a torus. Expressing an element  $f$  of  $\mathfrak{t}_h$  in the basis  $(g_1, \dots, g_n)$ ,

$$f = \sum_{i=1}^n f^i g_i,$$

we see that  $f$  is very strongly regular if and only if the  $f^i$  are (globally)  $\mathbb{Q}$ -linearly independent, see [31]. It is known that this condition defines a residual set  $\mathcal{RC}$  with full measure in  $\mathfrak{t}_h$ . Denoting by  $\mathfrak{C}_{\mathfrak{t}_h}$  the complement of a subset of  $\mathfrak{t}_h$  in  $\mathfrak{t}_h$ , one has  $\mathfrak{C}_{\mathfrak{t}_h} \mathcal{VSR} \subset \mathfrak{C}_{\mathfrak{t}_h} \mathcal{Reg} \cup \mathfrak{C}_{\mathfrak{t}_h} \mathcal{RC}$ . And  $\mathfrak{t}_h \cap \mathcal{VSR}$  has full measure in  $\mathfrak{t}_h$ .  $\square$

2.2. Statements of the main results

Using ideas developed by Jurdjevic and Kupka in [27, 28], we provide a general method to associate to any control system of the form of  $\Sigma_1$  another “nicer” system having for any strictly positive time the same closure of the attainable set. Such a system is called an *extension* of the first one. Precisely, we enounce the theorem:

**Theorem 2.6.** *Note  $\mathcal{Y} = \text{conv}(\mathfrak{t} + \text{Ad}_T(h_0))$  the convex hull of the sum of the Cartan algebra  $\mathfrak{t}$  and the set of all  $\text{Ad}_\tau(h_0)$  where  $\tau$  is an element of the torus  $T$ , and define the left invariant control system*

$$(\Sigma_{EQ}) \begin{cases} \dot{g} & = dL_g(w) \\ g(0) & = Id, \end{cases}$$

where  $w : \mathbb{R}^+ \rightarrow \mathcal{Y}$  is any measurable bounded map. If  $h_1$  is very strongly regular, then, for any time  $t$  strictly greater than zero, the closure of the attainable sets at time  $t$  of the systems  $\Sigma_1$  and  $\Sigma_{EQ}$  are the same.

The main interest of the previous result is that the one dimensional control  $u$  of  $\Sigma_1$  is replaced by a control in  $\mathcal{Y}$  for  $\Sigma_{EQ}$ . The set  $\mathcal{Y}$  is usually full dimensional, as shown in the following proposition.

**Proposition 2.7.** *If  $h_1$  is very strongly regular, and if for all root  $\alpha$  in  $\Delta$ , the number  $|a_\alpha^0|^2 + |a_{-\alpha}^0|^2$  is strictly positive, then  $\mathcal{Y}$  contains the Killing ball centered in  $0_{\mathfrak{u}}$  with radius  $\rho > 0$  where*

$$\rho = \frac{\prod_{x \in \mathcal{B}_{\mathfrak{t}^\perp}} \sqrt{\langle x, x \rangle (a_{[x]}^0)^2 + (a_{[-x]}^0)^2}}{\sqrt{\sum_{x \in \mathcal{B}_{\mathfrak{t}^\perp}} \prod_{\substack{y \in \mathcal{B}_{\mathfrak{t}^\perp} \\ y \neq x}} \langle y, y \rangle (a_{[y]}^0)^2 + (a_{[-y]}^0)^2}}. \tag{2.3}$$

It is possible to simplify the expression of Proposition 2.7 to obtain the following slightly weaker result.

**Proposition 2.8.** *If  $h_1$  is very strongly regular, and if  $(a_{[x]}^0)^2 + (a_{[-x]}^0)^2 > 0$  for all  $x$  in  $\mathcal{B}_{\mathfrak{t}^\perp}$ , then  $\mathcal{Y}$  contains the ball centered in  $0_{\mathfrak{u}}$  with radius  $\rho_2$  where*

$$\rho_2 = \frac{\min_{x \in \mathcal{B}_{\mathfrak{t}^\perp}} \left( \sqrt{\langle x, x \rangle (a_{[x]}^0)^2 + (a_{[-x]}^0)^2} \right)}{\sqrt{N_{\mathfrak{u}} - r_{\mathfrak{u}}}}.$$

Using now classical Riemannian techniques, we get a first estimation of  $\mathcal{D}(U, \Sigma_1)$ .

**Proposition 2.9.** *If  $h_1$  is very strongly regular, and if  $\sqrt{\langle x, x \rangle (a_{[x]}^0)^2 + (a_{[-x]}^0)^2} > 0$  for all  $x$  in  $\mathcal{B}_{\mathfrak{t}^\perp}$ , then the diameter  $\mathcal{D}(U, \Sigma_1)$  is less or equal than  $2\pi \frac{\sqrt{N_{\mathfrak{u}} - 1}}{\rho}$  (and hence less than  $2\pi \frac{\sqrt{N_{\mathfrak{u}} - 1}}{\rho_2}$ ).*

The estimation given in Proposition 2.9 is valid in any semi-simple compact Lie group. This includes the classical groups  $SU(n)$ ,  $SO(n)$  and  $Sp(n)$ , but also the so-called *exceptional* groups. This bound is also easy to compute.

We can get a better upper bound and a lower bound for  $\mathcal{D}(U, \Sigma_1)$  by use of the projection of the bi-invariant Killing metric on the quotient  $U/T$ :

**Theorem 2.10.** *If  $h_1$  is very strongly regular, and if for all root  $\alpha$  in  $\Delta$ , the number  $|a_\alpha^0|^2 + |a_{-\alpha}^0|^2$  is strictly positive, then*

$$\frac{1}{\sqrt{\langle h_0, h_0 \rangle}} \mathcal{D}(U/T, d_T) \leq \mathcal{D}(U, \Sigma_1) \leq \frac{1}{\rho} \mathcal{D}(U/T, d_T), \tag{2.4}$$

where  $\mathcal{D}(U, \Sigma_1)$  and  $\mathcal{D}(U/T, d_T)$  respectively denote the diameter of  $U$  under  $\Sigma_1$  and the diameter of the homogeneous space  $U/T$  under the projection  $d_T$  of the (opposite of the) standard Killing metric of  $U$ .

To take full advantage of this result, the value of  $\mathcal{D}(U/T, d_T)$  is needed. Unfortunately, it seems that no explicit expression for  $\mathcal{D}(U/T, d_T)$  is known. In this paper, we compute it in the case where  $U = \text{SU}(n)$ .

**Theorem 2.11.**

$$\mathcal{D}(\text{SU}(n)/T, d_T) = \frac{\pi}{3} \sqrt{6(n+1)(n-1)}.$$

Replacing now  $\mathcal{D}(\text{SU}(n)/T, d_T)$  by its expression in Theorem 2.10, we obtain an effective estimation of  $\mathcal{D}(\text{SU}(n), \Sigma_1)$ .

**Corollary 2.12.** *If  $U = \text{SU}(n)$ , if  $h_1$  is very strongly regular, and if  $a$  and  $b$  are the two real numbers  $a = \sqrt{\langle h_0, h_0 \rangle}$  and  $b = \min_{x \in \mathcal{B}_{\mathfrak{t}^\perp}} \sqrt{\langle x, x \rangle (a_{[x]}^0)^2 + a_{[x]}^0}$ , then the diameter of  $\text{SU}(n)$  under  $\Sigma_1$  is more than  $\frac{\pi}{3a} \sqrt{6(n+1)(n-1)}$  and less than  $\frac{\pi}{3b} (n-1) \sqrt{6(n+1)n}$ .*

Approximating now  $\mathcal{Y}$  by balls, not for the standard bi-invariant Killing form, but for an only left invariant metric, we describe a general method to get a last upper bound of the diameter  $\mathcal{D}(U, \Sigma_1)$ . The very special root structure of  $\text{SU}(n)$  allows us to do explicit computations in full generality. The result can be stated in the following way:

Fix any strictly positive real numbers  $\{d_\alpha, \alpha \in \Delta\}$  such that the convex hull of the points  $\pm \sqrt{a_\alpha^0{}^2 + a_\alpha^0} \partial_\alpha$  and  $\pm \sqrt{a_\alpha^0{}^2 + a_\alpha^0} \partial_{\bar{\alpha}}$  for all  $\alpha$  in  $\Delta$  contains the ellipsoid in  $\mathfrak{t}^\perp$  defined by the equation  $\sum_\alpha \frac{x_\alpha^2 + x_\alpha^2}{d_\alpha^2} \leq 1$ , where  $x_\alpha$ , resp.  $x_{\bar{\alpha}}$ , denotes the coordinate along  $\partial_\alpha$ , resp.  $\partial_{\bar{\alpha}}$ . Then define the bilinear symmetric form  $\langle \cdot, \cdot \rangle_D$  on  $\mathfrak{u}$  in the following way:

- $\mathfrak{t}^\perp$  and  $\mathfrak{t}$  are orthogonal for  $\langle \cdot, \cdot \rangle_D$ ;
- the restriction to  $\mathfrak{t}$  of  $\langle \cdot, \cdot \rangle$  and  $\langle \cdot, \cdot \rangle_D$  coincide;
- for any  $x, y \in \mathcal{B}_{\mathfrak{t}^\perp}$ ,  $\langle x, y \rangle_D = 0$  if  $x \neq y$ ;
- for any  $x \in \mathcal{B}_{\mathfrak{t}^\perp}$ ,  $\langle x, x \rangle_D = d_{[x]}^2$ .

Define a function  $C : \mathcal{B} \times \mathcal{B} \rightarrow \mathbb{R}$  in the following way:

- if  $[x, y] = 0$ ,  $C_{x,y} = 0$ ;
- if  $[x, y] \neq 0$ ,
  - if  $x \in \mathfrak{t}^\perp$  and  $y \in \mathfrak{t}$ ,  $C_{x,y} = 0$ ;
  - if  $x, y \in \mathfrak{t}^\perp$  and  $x = \bar{y}$ ,  $C_{x,y} = 1/2$ ;
  - if  $x, y \in \mathfrak{t}^\perp$  and  $x \neq \bar{y}$ ,  $C_{x,y} = \frac{\langle [x,y], [x,y] \rangle_D + \langle y, [x, [x,y]] \rangle_D + \langle x, [y, [x,y]] \rangle_D}{2 \langle [x,y], [x,y] \rangle_D}$ .

And finally, define a function  $\tilde{\mathcal{R}} : \mathcal{B}_{\mathfrak{t}^\perp} \rightarrow \mathbb{R}$  such that:

$$\tilde{\mathcal{R}}(x) = \sum_{z \in \mathcal{B}} \langle C_{z,x} C_{x,[x,z]} [x, [x, z]] + C_{[x,z],z} [[x, z], z], z \rangle_D. \tag{2.5}$$

**Theorem 2.13.** *If  $r = \min_{x \in \mathcal{B}_{\mathfrak{t}^\perp}} \tilde{\mathcal{R}}(x)/d_{[x]}^2 > 0$ , then the diameter of  $\text{SU}(n)$  under  $\Sigma_1$  is less than  $\pi \frac{\sqrt{n^2-2}}{r}$ .*

### 3. ATTAINABLE SET OF AN AFFINE SYSTEM ON A COMPACT LIE GROUP

#### 3.1. “Variation of the constant”

The first step is to apply the technique of “variation of the constant” to system  $\Sigma_1$ . Fix a strictly positive real  $t_0$  and a measurable bounded function  $u : [0, t_0] \rightarrow \mathbb{R}$ . We denote by  $g : [0, t_0] \rightarrow U$  the corresponding solution of  $\Sigma_1$ . This solution is defined on the whole interval  $[0, t_0]$ . For every  $t \in [0, t_0]$  we define  $q(t)$  by

$$g(t) = q(t) \exp \left( \int_0^t u(\tau) d\tau \ h_1 \right). \tag{3.1}$$

The mapping  $q : [0, t_0] \rightarrow U$  is well defined and has the same regularity as  $g$ . Obviously,  $q(0) = Id$ . Differentiating equation (3.1), it is easy to obtain the differential equation for  $q$

$$\dot{q} = dL_{q(t)} \exp\left(\int_0^t u(\tau) d\tau \operatorname{ad} h_1\right) h_0,$$

which writes

$$\dot{q} = dL_{q(t)} \exp(v(t) \operatorname{ad} h_1) h_0, \tag{3.2}$$

when noting

$$v(t) = \int_0^t u(\tau) d\tau.$$

The equation for  $q$  is left invariant.

Since  $Ad = \exp(\operatorname{ad})$  is an isometry for the Killing form, the norm  $\sqrt{\langle \dot{q}, \dot{q} \rangle}$  of  $\dot{q}$  is equal to the norm of  $h_0$ , and hence bounded. This fact will be used in the proof of Proposition 3.2.

Equation (3.2) with  $q(0) = Id$  as initial condition and every Lipschitzian functions  $v$  such that  $v(0) = 0$  as admissible controls defines a control system  $\Sigma_2$

$$(\Sigma_2) \begin{cases} \dot{q} & = dL_{q(t)} \exp(v(t) \operatorname{ad} h_1) h_0 \\ q(0) & = Id. \end{cases}$$

**Remark 3.1.** As integral of a bounded measurable function,  $v$  has to be Lipschitzian. Moreover,  $v(0) = 0$ . We shall come back to the regularity of  $v$  in the next subsection.

Let us note respectively  $\mathcal{A}_1(t)$  and  $\mathcal{A}_2(t)$  the attainable sets of  $\Sigma_1$  and  $\Sigma_2$  at time  $t$ . The next proposition gives a link between  $\mathcal{A}_1(t)$  and  $\mathcal{A}_2(t)$ .

**Proposition 3.2.** *If  $h_1$  is a very strongly regular element of  $\mathfrak{t}$ , then for every  $t \in ]0, t_0]$ ,  $\overline{\mathcal{A}_1(t)} = \overline{\mathcal{A}_2(t)}$   $T$ .*

*Proof.* We prove this set equality in two steps.

- $\overline{\mathcal{A}_1(t)} \subset \overline{\mathcal{A}_2(t)}$   $T$  is obvious by using the definition of  $q$  in equation (3.1).
- $\overline{\mathcal{A}_1(t)} \supset \overline{\mathcal{A}_2(t)}$   $T$  is more complicated. It suffices to prove that  $\overline{\mathcal{A}_1(t)} \supset \mathcal{A}_2(t)$   $T$ . Take any point  $q(t)$  in  $\mathcal{A}_2(t)$ , and  $u : [0, t] \rightarrow \mathbb{R}$  the corresponding (measurable bounded) control, and any  $h_t$  in  $T$ . We are looking for a sequence  $g_n(t)$  in  $\mathcal{A}_1(t)$  such that  $g_n(t)$  converges to  $q(t) h_t$  as  $n$  goes to infinity. Using  $u$  in  $\Sigma_1$ , we obtain a point  $g_u(t)$  of  $\mathcal{A}_1(t)$ . Using equation (3.1) we can express  $g_u(t) = q_u(t) h_u(t)$ . Obviously,  $q_u(t) = q(t)$ . If  $h_t = h_u(t)$ , we need go no further. Otherwise, let  $h_n$  be a sequence of points of  $\{\exp(\mathbb{R} h_1)\} \subset T$  which tends to  $h_u(t)^{-1} h_t$ . This is possible because of the density of the one parameter group generated by  $h_1$  in the maximal torus  $T$ . (Notice that this property depends on the very strong regularity of  $h_1$ .) For each positive integer  $n$ , define the real number  $r_n$  by  $h_n = \exp(r_n h_1)$ . This definition is consistent.  $r_n$  is a real number whose modulus may tend to infinity as  $n$  grows.

For any integer  $n > 0$ , we define the sequence of control functions  $u_n : [0, t] \rightarrow \mathbb{R}$  by  $u_n(s) = n r_n$  if  $t - \frac{1}{n} < s < t$  and  $u_n(s) = u(s)$  else. We note  $g_{u_n}$  the trajectory of  $\Sigma_1$  obtained with this control function  $u_n$ . For every  $s$  in  $[0, t]$ ,  $g_{u_n}(s) = q_{u_n}(s) h_{u_n}(s)$ . Convergence of  $g_{u_n}(t)$  to  $q(t)h_u(t)$  is the consequence of the two following facts:

- $q_{u_n}(t)$  tends to  $q_u(t)$ , because  $q_{u_n}$  and  $q_u$  follow the same differential equation for time  $s$  such that  $0 < s < t - \frac{1}{n}$ , and  $q_{u_n}(0) = q_u(0) = Id$ , hence, for any  $n$ ,  $q_{u_n}(t - \frac{1}{n}) = q_u(t - \frac{1}{n})$ . The fact that  $\dot{q}$  is bounded guarantees the uniform convergence of the sequence  $(q_{u_n})_{n \in \mathbb{N}}$  to  $q_u$  on the interval  $[0, t]$ .
- $h_{u_n}$  tends to  $h_t$  because of the particular choice of  $r_n$ .

This completes the proof. □



**Remark 3.3.** The fact that  $\exp(\{uh_1, u \in \mathbb{R}\})$  is dense in  $T$  is absolutely crucial. This is a consequence of the fact that  $u$  has no *a priori* bound and also that the Lie algebra  $\mathfrak{u}$  is finite dimensional.

**Remark 3.4.** The proof above is valid in every compact real semi-simple Lie group. Actually, the only requirement for the control part is that the closure of the subgroup generated by it is a maximal torus, not necessary single-input.

From now on, we make the following generic assumption.

**Assumption 1.**  $h_1$  will be assumed to be very strongly regular.

**Remark 3.5.**  $\mathcal{A}_2(t)$  can actually be thought as a subset of the quotient  $U/T$ . More precisely, we are just interested in the image  $\pi_T(\mathcal{A}_2(t))$  of  $\mathcal{A}_2(t)$  on  $U/T$  under the canonical projection  $\pi_T : U \rightarrow U/T$ . This point of view will be used later in Section 4.

Proposition 3.2 proves that it is enough to compute  $\overline{\mathcal{A}_2(t)}$ . The rest of this section is dedicated to the derivation of a good enough approximation of this set.

### 3.2. Proof of the extension results

In this section, we prove Theorem 2.6 and construct a problem on  $U$  that approximates the initial problem. This approximation will be used in Sections 4 and 5 to estimate the diameter  $\mathcal{D}(U, \Sigma_1)$ . An example of the practical use in  $SU(3)$  of the results presented here can be found in Section 6.1.

First of all, we use the fact that  $\{\exp(t.h_1), t \in \mathbb{R}\}$  is a dense subset of  $T$  to relax the regularity hypothesis on  $v$ .

We define a new system  $\Sigma_3$

$$(\Sigma_3) \begin{cases} \dot{q} & = dL_q \exp(ad v)h_0 \\ q(0) & = Id, \end{cases}$$

where  $v : [0, t] \rightarrow \mathfrak{t}$  is any bounded measurable function. We denote by  $\mathcal{A}_3(t)$  the attainable set at time  $t$  from  $Id$  of  $\Sigma_3$ .

**Proposition 3.6.**  $\overline{\mathcal{A}_2(t)} = \overline{\mathcal{A}_3(t)}$  for any  $t$  greater than 0.

*Proof.* It is clear that  $\mathcal{A}_2(t) \subset \mathcal{A}_3(t)$ . Hence, it is enough to prove that  $\mathcal{A}_3(t) \subset \overline{\mathcal{A}_2(t)}$  for every  $t$  greater than zero. Take any point  $q(t)$  in  $\mathcal{A}_3(t)$ , and the associated control function  $v$ . If  $v$  is Lipschitzian such that  $v(0) = 0$  and  $t \mapsto \exp(v(t)h_1)$  takes value in  $\{\exp(u h_0), u \in \mathbb{R}\}$ , we need go no further. Otherwise, because of the density of  $\{\exp(u h_0), u \in \mathbb{R}\}$  in  $T$ , and because of the density of the set of absolutely continuous functions vanishing at 0 in the set of measurable bounded functions from  $[0..t]$  to  $\mathfrak{t}$  for the  $L^1$  norm, we can find a sequence  $q_n(t)$  of points of  $\mathcal{A}_2(t)$  converging to  $q(t)$ .  $\square$

Denoting by  $h_0^\mathfrak{t}$  the orthogonal (with respect to the Killing form) projection of  $h_0$  on  $\mathfrak{t}$ , and using the density of  $\{\exp(t h_1 + h_0^\mathfrak{t}), t \in \mathbb{R}\}$  in  $\exp(\mathfrak{t})$ , we can assume that  $h_0$  is in  $\mathfrak{t}^\perp$  (orthogonal subspace to  $\mathfrak{t}$  with respect to  $\langle, \rangle$ ) without loss of generality. From now on, we shall make the following non restrictive assumption.

**Assumption 2.** Vector  $h_0$  of  $\mathfrak{u}$  will be assumed to have zero orthogonal projection on  $\mathfrak{t}$ . That is to say, we shall assume that  $h_0$  belongs to  $\mathfrak{t}^\perp$ , the orthogonal subspace to  $\mathfrak{t}$  for the standard Killing form.

The idea is now to use the relaxation principle (see [28]) in the particular context of left invariant systems on Lie groups. Let  $S_1$  and  $S_2$  be the two left invariant control systems on  $U$ :

$$(S_1) \begin{cases} \dot{x} & = dL_x v, \text{ with } v \in \mathcal{U} \\ x(0) & = x_0, \end{cases} \quad (S_2) \begin{cases} \dot{x} & = dL_x v, \text{ with } v \in \text{conv}(\mathcal{U}) \\ x(0) & = x_0, \end{cases}$$

where “conv” is a notation for the convex hull of a subset of a vector space, and  $\mathcal{U}$  is a subset of  $\mathfrak{u}$ . Then the attainable sets at time  $t$  of the two systems  $S_1$  and  $S_2$  have the same closure for every  $t > 0$ .

We introduce the control system  $\Sigma_4$ .

$$(\Sigma_4) \begin{cases} \dot{q} &= dL_q w \\ q(0) &= Id, \end{cases}$$

where  $w$  is any measurable function from  $\mathbb{R}^+$  to  $\mathcal{X} = \text{conv}(\{e^{ad v}(h_0), v \in \mathfrak{t}\})$ .

For every  $t > 0$ , we will denote by  $\mathcal{A}_4(t)$  the attainable set at time  $t$  from  $Id$  for the system  $\Sigma_4$ . The relaxation principle asserts that the two systems  $\Sigma_3$  and  $\Sigma_4$  are equivalent.

**Lemma 3.7.** *For every  $t > 0$ ,  $\overline{\mathcal{A}_4(t)} = \overline{\mathcal{A}_3(t)}$ .*

Let us finally introduce the system  $\Sigma_{EQ}$

$$(\Sigma_{EQ}) \begin{cases} \dot{g} &= dL_g w \\ g(0) &= Id, \end{cases}$$

where  $w$  is any measurable bounded application  $w : \mathbb{R}^+ \rightarrow \mathcal{Y}$  with  $\mathcal{Y} = \text{conv}\{\mathcal{X} + \mathfrak{t}\}$ , and denote by  $\mathcal{A}_{EQ}(t)$  its attainable set at time  $t$ . The systems  $\Sigma_1$  and  $\Sigma_{EQ}$  are equivalent.

**Theorem 2.6** For any strictly positive time  $t$ ,  $\overline{\mathcal{A}_{EQ}(t)} = \overline{\mathcal{A}_1(t)}$ .

*Proof.* To any measurable bounded function  $u : [0, t_0] \rightarrow \mathcal{Y}$ , we associate the two measurable bounded functions  $u_{\mathfrak{t}} : [0, t_0] \rightarrow \mathfrak{t}$  and  $u_{\mathfrak{t}^\perp} : [0, t_0] \rightarrow \mathfrak{t}^\perp$  in the following way: for each  $t$  in  $[0, t_0]$ ,  $u_{\mathfrak{t}}(t) = \pi_{\mathfrak{t}}(u(t))$  and  $u_{\mathfrak{t}^\perp}(t) = u(t) - u_{\mathfrak{t}}(t) = \pi_{\mathfrak{t}^\perp}(u(t))$ , where  $\pi_{\mathfrak{t}} : \mathfrak{u} \rightarrow \mathfrak{t}$  and  $\pi_{\mathfrak{t}^\perp} : \mathfrak{u} \rightarrow \mathfrak{t}^\perp$  are respectively the orthogonal projections for the standard Killing form on  $\mathfrak{t}$  and on  $\mathfrak{t}^\perp$ . For each trajectory  $g_u : [0, t_0] \rightarrow U$  of the system  $\Sigma_{EQ}$  associated with the control function  $u : [0, t_0] \rightarrow \mathcal{Y}$ , we define a mapping  $q_u : [0, t_0] \rightarrow U$  in the following way

$$g_u(t) = q_u(t) \exp \left( \int_0^t u_{\mathfrak{t}}(s) ds \right). \tag{3.3}$$

The evolution equation for  $q_u$  is obtained by derivation of the equation (3.3) with respect to  $t$

$$\dot{q} = dL_{q(t)} \left( \exp \left( \text{ad} \int_0^t u_{\mathfrak{t}}(\tau) d\tau \right) \right) u_{\mathfrak{t}^\perp}. \tag{3.4}$$

Equation (3.4) with the initial condition  $q(0) = Id$  and with the set of admissible controls equal to the set of all measurable bounded mapping  $u_{\mathfrak{t}}$  from  $\mathbb{R}^+$  to  $\mathfrak{t}$ , defines a control system  $\Sigma_5$ , whose attainable set at time  $t$  is denoted by  $\mathcal{A}_5(t)$ . A slight adaptation of the proof of Proposition 3.2 proves that  $\overline{\mathcal{A}_{EQ}(t)} = \overline{\mathcal{A}_5(t)T}$  for any strictly positive  $t$ . Using now Lemma 3.7 and the fact that  $\mathcal{X} \subset \mathfrak{t}^\perp \perp \mathfrak{t}$ , one sees that  $\overline{\mathcal{A}_5(t)} = \overline{\mathcal{A}_2(t)}$  for any strictly positive  $t$ . Hence  $\overline{\mathcal{A}_{EQ}(t)} = \overline{\mathcal{A}_2(t)T} = \overline{\mathcal{A}_1(t)}$ .  $\square$

### 3.3. Geometric properties of $\mathcal{Y}$

The principal benefit of the introduction of the control system  $\Sigma_{EQ}$  is that the convex set  $\mathcal{Y}$  has nice geometric properties.

**Proposition 3.8.** *The set  $\mathcal{Y}$  is a closed convex neighbourhood of  $0_{\mathfrak{u}}$ . Its orthogonal projection (with respect to the Killing form) on  $\mathfrak{t}^\perp$  is the closed convex set  $\mathcal{X}$ . The set  $\mathcal{X}$  is a compact neighborhood of  $0_{\mathfrak{u}}$  in  $\mathfrak{t}^\perp$  bounded by the Killing norm of  $h_0$ . Moreover, for any root  $\alpha_0$ , the four points  $\pm \sqrt{(a_{\alpha_0}^0)^2 + (a_{-\alpha_0}^0)^2} \partial_{\alpha_0}$  and  $\pm \sqrt{(a_{\alpha_0}^0)^2 + (a_{-\alpha_0}^0)^2} \partial_{-\alpha_0}$  belong to  $\mathcal{X}$ .*

The first statements are obvious. The proof of the last statement requires some intermediate results:

**Lemma 3.9.** *If  $x \in \mathcal{X}$ , then  $\exp(\text{adv})x \in \mathcal{X}$  for any  $v$  in  $\mathfrak{t}$ .*

*Proof.* If  $x$  is in  $\mathcal{X}$ ,  $x$  writes  $x = \sum_{j=1}^p \lambda_j e^{\text{ad } v_j} h_0$  with  $\sum_{i=1}^p \lambda_j \leq 1$  and  $\lambda_j \geq 0$  for all  $j$ . Then  $e^{\text{ad } v}(x) = \sum_{j=1}^p \lambda_j (e^{\text{ad } v})(e^{\text{ad } v_j}) h_0$ . Since  $e^{\text{ad } v_1} e^{\text{ad } v_2} = e^{\text{ad } (v_1+v_2)}$  as soon as  $[v_1, v_2] = 0$ , for all  $j$ , there exists  $\tilde{v}_j$  in  $\mathfrak{t}$  such that  $e^{\text{ad } v}(x) = \sum_{j=1}^p \lambda_j e^{\text{ad } \tilde{v}_j} h_0$ . □

**Lemma 3.10.** *Let  $\alpha_0$  and  $\alpha_1$  be two different positive roots and  $u$  be a real number. If  $x = \sum_{\alpha \in \Delta} a_\alpha^x \partial_\alpha + \sum_{\alpha \in \Delta} a_\alpha^x \partial_{\bar{\alpha}}$  is in  $\mathcal{X}$ , then there exists  $x' = \sum_{\alpha \in \Delta} a_\alpha^{x'} \partial_\alpha + \sum_{\alpha \in \Delta} a_\alpha^{x'} \partial_{\bar{\alpha}}$  in  $\mathcal{X}$  such that  $(a_\alpha^x)^2 + (a_{\bar{\alpha}}^x)^2 = (a_\alpha^{x'})^2 + (a_{\bar{\alpha}}^{x'})^2$  for all  $\alpha$  in  $\Delta$ ,  $a_{\alpha_0}^{x'} = \cos(u)a_{\alpha_0}^x + \sin(u)a_{\bar{\alpha}_0}^x$ ,  $a_{\bar{\alpha}_0}^{x'} = -\sin(u)a_{\alpha_0}^x + \cos(u)a_{\bar{\alpha}_0}^x$ ,  $a_{\alpha_1}^{x'} = a_{\alpha_1}^x$  and  $a_{\bar{\alpha}_1}^{x'} = a_{\bar{\alpha}_1}^x$ .*

*Proof.* Since  $\alpha_0$  and  $\alpha_1$  are two different roots, they are not colinear as elements of  $\mathfrak{t}^*$ . Thus, there exists  $v$  in  $\mathfrak{t}$  such that  $\alpha_0(v) = i u$  and  $\alpha_1(v) = 0$ . Let  $x'$  be defined by  $x' = e^{\text{ad } v}(x)$ . By Lemma 3.9,  $x'$  is in  $\mathcal{X}$ . Now, it suffices to use the computation rules (A.14), (A.15) and (A.16) to get the result. □

*Proof of Proposition 3.8.* We define an arbitrary complete order  $\prec$  in  $\Delta$  such that  $\alpha_0$  is the smallest element of  $\Delta$ . Up to a reindexation of the element of  $\Delta$ , we may assume that  $\alpha_0 \prec \alpha_1 \prec \dots \prec \alpha_n$ . We want to construct recursively a sequence  $x_n, x_{n-1}, \dots, x_0$  of elements of  $\mathcal{X}$  such that:

- $x_n = h_0$ ;
- $x_0 = a_{\alpha_0}^0 \partial_{\alpha_0} + a_{\bar{\alpha}_0}^0 \partial_{\bar{\alpha}_0}$ ;
- for any  $k$ , the coordinates of  $x_k$  along  $\partial_{\alpha_j}$  and  $\partial_{\bar{\alpha}_j}$  are zero for all  $j > k$ ;
- for any  $k$ , the coordinates of  $x_k$  along  $\partial_{\alpha_0}$  and  $\partial_{\bar{\alpha}_0}$  are exactly  $a_{\alpha_0}^0$  and  $a_{\bar{\alpha}_0}^0$ .

Assume  $x_{k+1}$  is constructed. Using Lemma 3.10 with the roots  $\alpha_k$  and  $\alpha_0$ , and  $u = \pi$ , we can construct  $y_{k+1}$  from  $x_{k+1}$  such that the  $\partial_{\alpha_0}$  and  $\partial_{\bar{\alpha}_0}$  components of  $x_{k+1}$  and  $y_{k+1}$  are the same, and the  $\partial_{\alpha_{k+1}}$  and  $\partial_{\bar{\alpha}_{k+1}}$  components of  $x_{k+1}$  and  $y_{k+1}$  are opposite. Then  $x_k$  defined by  $x_k = \frac{x_{k+1} + y_{k+1}}{2}$  satisfies the required properties.

Now, it is enough to apply again the Lemma 3.10 to rotate the  $\partial_{\alpha_0}$  and  $\partial_{\bar{\alpha}_0}$  components of  $x_0$  to get the result. □

## 4. APPROXIMATION BY THE KILLING FORM

### 4.1. The Killing approximation

The aim of this section is to get an upper and a lower bound for the diameter of the compact semi-simple Lie group  $U$  under a control system of the type  $\Sigma_1$ . For this purpose, we approximate  $\mathcal{Y}$  by balls for the Killing form. That is to say, we will find a ball  $B_1(0, r_1)$  of  $\mathfrak{u}$  centered in  $0_{\mathfrak{u}}$  having radius  $r_1$  such that  $B_1 \subset \mathcal{Y}$ . Then, the diameter of  $\Sigma_1$  is less than  $1/r_1$  times the diameter of  $U$  endowed with  $d_U$ , the metric induced by (the opposite of) the standard Killing form. The first task is to find a radius  $r_1$  as large as possible satisfying  $B_1(0, r_1) \subset \mathcal{Y}$ . Then, we shall compute an upper bound for the diameter of  $U$  equipped with the standard Killing form. This can be done in a very general way, using classical Myers theorem. An upper bound for the diameter of system  $\Sigma_1$  is given in Section 4.2. This easily computable upper bound is valid in any compact real semi-simple Lie group (Prop. 2.9).

As already said in Remark 3.5, direction  $\mathfrak{t}$  is free. A natural idea is to approximate  $\mathcal{X}$ , the orthogonal projection of  $\mathcal{Y}$  along  $\mathfrak{t}$ , instead of  $\mathcal{Y}$ . Indeed, this projection has lower dimension than  $\mathcal{Y}$ , and it is usually easier to approximate small dimensional sets. Furthermore, such an approximation in  $U/T$  is closer to the geometry of the problem than our previous approximation in  $U$ . For example, not only can we expect to get a better upper bound, but also a lower bound for the diameter of  $\Sigma_1$ , since, as already said, the velocity of  $\pi_T(q)$  in  $U/T$  is bounded by  $k_2 = \sqrt{\langle h_0, h_0 \rangle}$ , where  $\pi_T : U \rightarrow U/T$  is the canonical projection. The diameter of  $U$  under  $\Sigma_1$  is greater than  $\frac{1}{k_2}$  times the diameter of the quotient space  $U/T$  under the projection of  $d_U$ .

The problem is to compute the diameter of the quotient of a semi-simple compact Lie group by a maximal torus for the metric induced by the standard Killing metric. This is not a simple matter, and it seems that no answer is known, except in some particular low dimensional cases. Computations have recently been made for  $SU(3)$  in [36]. For the classical groups  $U = SO(n)$  and  $U = Sp(n)$  (corresponding to the case where  $\mathfrak{u}$  is a compact real form of one of the classical algebra  $\mathfrak{b}_n, \mathfrak{c}_n, \text{ or } \mathfrak{d}_n$ ), we shall give a lower bound for the diameter  $U/T$ , just exhibiting a particular point of  $U/T$ . In Section 4.4, we compute in a complete and explicit way the diameter of  $U/T$  for the projection of  $\langle, \rangle$  in the case where  $U = SU(n)$ . This allows us to obtain both a lower and an upper approximation of the diameter of  $\Sigma_1$ .

### 4.2. Upper bound of the diameter in the general case

#### 4.2.1. Result

**Proposition 2.9.** Let us define  $\rho$  as in Proposition 2.7. The diameter of  $U$  under  $\Sigma_1$  is less than  $2\pi \frac{\sqrt{N_{\mathfrak{u}}-1}}{\rho}$ .

An example of application in  $SU(3)$  of this result is presented in Section 6.2.

**Remark 4.1.** Although it is not generic, it may happen that  $\rho$  vanishes. In this case,  $\rho_2$  vanishes too, since  $\rho_2 \leq \rho$ . This happens when at least one of the numbers  $a_{\alpha}^{0^2} + a_{\bar{\alpha}}^{0^2}$  vanishes. In this case, the system  $\Sigma_1$  can be controllable or not. That is to say, its diameter can be defined (and finite), or not. In any case, if  $\rho$  is zero, what follows remains true, but will always give  $+\infty$  as an upper bound for the diameter of  $\Sigma_1$ .

The first task is to find a radius  $r_1$ , as large as possible, satisfying  $\mathcal{B}_1(0, r_1) \subset \mathcal{Y}$ . Then, we estimate the diameter of  $U$  endowed with the standard Killing form.

#### 4.2.2. A big ball contained in $\mathcal{X}$

**Proposition 2.7.** If  $h_1$  is very strongly regular, and if for all root  $\alpha$  in  $\Delta$ , the number  $|a_{\alpha}^0|^2 + |a_{\bar{\alpha}}^0|^2$  is strictly positive, then  $\mathcal{Y}$  contains the Killing ball centered in  $0_{\mathfrak{u}}$  with radius  $\rho > 0$  where

$$\rho = \frac{\prod_{x \in \mathcal{B}_{\perp}} \sqrt{\langle x, x \rangle (a_{[x]}^0)^2 + a_{[x]}^0{}^2}}{\sqrt{\sum_{x \in \mathcal{B}_{\perp}} \prod_{\substack{y \in \mathcal{B}_{\perp} \\ y \neq x}} \langle y, y \rangle (a_{[y]}^0)^2 + a_{[y]}^0{}^2}}$$

*Proof.* From Section 3, we know that for all positive roots  $\alpha$ , the four points  $\pm \sqrt{a_{\alpha}^0{}^2 + a_{\bar{\alpha}}^0{}^2} \partial_{\alpha}$  and  $\pm \sqrt{a_{\alpha}^0{}^2 + a_{\bar{\alpha}}^0{}^2} \partial_{\bar{\alpha}}$  belong to  $\mathcal{X}$ . All these points have norm  $\sqrt{\langle \partial_{\alpha}, \partial_{\alpha} \rangle (a_{\alpha}^0{}^2 + a_{\bar{\alpha}}^0{}^2)}$ . Let us stress again that the norm of the vectors  $\partial_{\alpha}$  or  $\partial_{\bar{\alpha}}$  is *not* always one. Since  $\mathcal{Y}$  is convex, any point of the form  $\sum_{y \in \mathcal{B}_{\perp}} x_y \sqrt{(a_{[y]}^0)^2 + a_{[y]}^0{}^2} y$  with  $x_y \geq 0$  for all  $y$  in  $\mathcal{B}_{\perp}$  and  $\sum_{y \in \mathcal{B}_{\perp}} x_y \leq 1$  belongs to  $\mathcal{Y}$  too. The mapping  $\mathcal{L} : (x_y)_{y \in \mathcal{B}_{\perp}} \mapsto \sqrt{\sum_y x_y^2 (a_{[y]}^0)^2 + a_{[y]}^0{}^2} \langle y, y \rangle$  associates to any family of  $N_{\mathfrak{u}} - r_{\mathfrak{u}}$  real numbers in  $[0, 1]$  the norm of the corresponding convex combination in  $\mathcal{X}$ . To determine a radius  $\rho$ , such that the ball centered in  $0_{\mathfrak{u}}$  with radius  $\rho$  is contained in  $\mathcal{Y}$ , it suffices to find the maximum of  $\mathcal{L}$  on the set  $\{(x_y)_y : x_y \geq 0 \text{ and } \sum_y x_y = 1\}$ . The classical theorem on constrained optimization asserts that  $\mathcal{L}$  reaches its maximum  $\rho$  at the point  $(x_z)_{z \in \mathcal{B}_{\perp}}$  with

$$x_z = \frac{\prod_{x \in \mathcal{B}_{\perp}} \langle x, x \rangle (a_{[x]}^0)^2 + a_{[x]}^0{}^2}{\langle z, z \rangle (a_{[z]}^0)^2 + a_{[z]}^0{}^2} \sum_{x \in \mathcal{B}_{\perp}} \prod_{\substack{y \in \mathcal{B}_{\perp} \\ y \neq x}} \langle y, y \rangle (a_{[y]}^0)^2 + a_{[y]}^0{}^2}. \quad \square$$

By denoting by  $a$  the minimum of all  $\sqrt{(a_{[x]}^0)^2 + a_{[x]}^0} \langle x, x \rangle$  for  $x$  in the basis  $\mathcal{B}_{\perp}$ ,  $\mathcal{Y}$  contains any vector of the form  $\frac{a}{\sqrt{\langle x, x \rangle}} x$ , for  $x$  in the basis  $\mathcal{B}$ . One obtains a slightly weaker but more practical result.

**Proposition 2.8.**  $\mathcal{Y}$  contains the ball centered in  $0_{\mathfrak{u}}$  with radius  $\frac{\min_{x \in \mathcal{B}} \left\{ \sqrt{(a_{[x]}^0)^2 + a_{[x]}^0} \langle x, x \rangle \right\}}{\sqrt{N_{\mathfrak{u}} - r_{\mathfrak{u}}}}$ .

4.2.3. *Upper bound of the diameter of  $U$  endowed with the Killing form*

The idea is to use the classical Myers theorem which gives a link between the Ricci curvature and the length of optimal geodesics (for a precise definition of Ricci curvature and a detailed proof of Myers theorem, see [16] or [34]). In the particular context of left invariant metrics on a connected semi-simple compact Lie group, the result can be stated as follows.

**Theorem 4.2.** *Let  $M$  be an  $N_M$ -dimensional compact semi-simple Lie group, endowed with a left invariant Riemannian structure  $(\cdot, \cdot)$ . We denote by  $K$  its Ricci curvature. Let  $v$  be a unitary vector of the Lie algebra  $T_{Id}M$ , and assume that  $K(v, v) > 0$ . Then the optimal geodesic starting from the unit element of the group and directed by  $v$  has a length less or equal to  $\pi \sqrt{\frac{N_M - 1}{K(v, v)}}$ . In particular, if there exists some strictly positive real number  $r$  such that  $K(v, v) \geq r$  for all unit vectors  $v$  in the Lie algebra  $T_{Id}M$ , then  $M$  has a diameter not greater than  $\pi \sqrt{\frac{N_M - 1}{r}}$ .*

A classical computation (see [16, 24, 34]) gives:

**Proposition 4.3.** *The Ricci curvature of the standard Killing form on a Lie group verifies  $K(v, v) = \frac{1}{4} \langle v, v \rangle$  for any vector  $v$  in the Lie algebra.*

We deduce from it an upper bound for the diameter of  $U$  endowed with the standard Killing metric.

**Corollary 4.4.** *The diameter of  $U$  endowed with the standard Killing metric is less than  $2\pi \sqrt{N_{\mathfrak{u}} - 1}$ .*

Proposition 2.9 follows.

**Remark 4.5.** Note that this easily computable upper bound is valid in *any* semi-simple compact Lie group.

4.3. **Lower bound of the diameter of  $U/T$ , in the case of the classical groups**

In this section, we give lower bounds for the diameter of the quotients of the classical matrix groups by a maximal torus. Details for the representations that we choose here, including the associated maximal tori and the expressions of the Killing form, can be found with proofs and further discussion in [24].

4.3.1.  $U = \text{SO}(2n)$

$\text{SO}(2n)$  is the group of the  $n$ -by- $n$  real matrices, which preserves the oriented canonical Euclidian structure of  $\mathbb{R}^{2n}$ . Such matrices are the solutions of the equation  ${}^tMM = Id$  with determinant equal to 1. The maximal torus  $T$  we choose is the set of all block diagonal matrices

$$\begin{pmatrix} R_{\theta_1} & 0_{2,2} & \dots & 0_{2,2} \\ 0_{2,2} & R_{\theta_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0_{2,2} \\ 0_{2,2} & \dots & 0_{2,2} & R_{\theta_n} \end{pmatrix},$$

where matrices  $R_{\theta_k}$  are 2-by-2 matrices defined by

$$R_{\theta_k} = \begin{pmatrix} \cos(\theta_k) & \sin(\theta_k) \\ -\sin(\theta_k) & \cos(\theta_k) \end{pmatrix},$$

for any real number  $\theta_k$ , and  $0_{k,l}$  is a notation for the zero matrix with  $k$  rows and  $l$  columns. The Lie algebra  $\mathfrak{so}(2n)$  of  $\text{SO}(2n)$  is the tangent space at the unit element of the smooth manifold  $\text{SO}(2n)$ . It is a normal compact real form of the Lie algebra  $\mathfrak{d}_n$ . The Riemannian metric induced by the standard Killing form is  $\langle X, Y \rangle = -(2n - 2) \text{tr}(XY)$  for any  $X, Y$  in  $\mathfrak{so}(2n)$ .

Let us introduce the matrix  $M$  of  $\text{SO}(2n)$  with entries  $m_{j,k}$  defined by  $m_{k+2,k} = 1$  for  $1 \leq k \leq n - 2$ ,  $m_{1,2n-1} = m_{2,2n} = 1$  and  $m_{j,k} = 0$  otherwise.  $M$  is the block matrix

$$M = \begin{pmatrix} 0_{2,2} & 0_{2,2} & \dots & 0_{2,2} & I_2 \\ I_2 & 0_{2,2} & \ddots & \vdots & 0_{2,2} \\ 0_{2,2} & I_2 & \ddots & 0_{2,2} & \vdots \\ \vdots & \ddots & \ddots & 0_{2,2} & \vdots \\ 0_{2,2} & \dots & 0_{2,2} & I_2 & 0_{2,2} \end{pmatrix},$$

where  $I_2$  is a notation for the identity matrix of order 2.

Then, the left translate  $M T$  of  $T$  by  $M$  is the set of matrices of  $\text{SO}(2n)$  of the form

$$M^{(\theta_1, \dots, \theta_n)} = \begin{pmatrix} 0_{2,2} & 0_{2,2} & \dots & 0_{2,2} & R_{\theta_n} \\ R_{\theta_1} & 0_{2,2} & \ddots & \vdots & 0_{2,2} \\ 0_{2,2} & R_{\theta_2} & \ddots & 0_{2,2} & \vdots \\ \vdots & \ddots & \ddots & 0_{2,2} & \vdots \\ 0_{2,2} & \dots & 0_{2,2} & R_{\theta_{n-1}} & 0_{2,2} \end{pmatrix}.$$

Matrix  $M^{(\theta_1, \dots, \theta_n)}$  has the characteristic polynomial

$$P_{M^{(\theta_1, \dots, \theta_n)}}(X) = X^{2n} - 2 \cos(\theta) X^n + 1,$$

where  $\theta = \sum_{k=1}^n \theta_k$ . That is to say,  $M^{(\theta_1, \dots, \theta_n)}$  has eigenvalues

$$\lambda_k = \exp\left(i \frac{2\pi k + \theta}{n}\right) \quad \text{and} \quad \mu_k = \exp\left(i \frac{2\pi k - \theta}{n}\right), \quad k \in \{0, \dots, n - 1\}.$$

We assume without loss of generality that  $\theta$  belongs to  $[0, 2\pi[$  if  $n$  is even, and that  $\theta$  belongs to  $[-\pi, \pi[$  if  $n$  is odd. Hence,

$$d_U(M^{(\theta_1, \dots, \theta_n)}, Id)^2 = \begin{cases} 2(2n - 2) \left( \sum_{k=-n/2}^{n/2-1} \frac{(2k\pi + \theta)^2}{n^2} \right) & \text{if } n \text{ is even} \\ 2(2n - 2) \left( \sum_{k=-(n-1)/2}^{(n-1)/2} \frac{(2k\pi + \theta)^2}{n^2} \right) & \text{if } n \text{ is odd,} \end{cases}$$

is minimum when  $\theta = \pi$  if  $n$  is even and  $\theta = 0$  if  $n$  is odd. Hence,  $n$  being odd or even,

$$d_T(\pi_T(Id), \pi_T(M)) = \frac{\pi}{3} \sqrt{\frac{6(n^2 - 1)(2n - 2)}{n}}. \tag{4.1}$$

This provides a lower bound for the diameter of the quotient of  $\text{SO}(2n)$  by a maximal torus.

4.3.2.  $U = \text{SO}(2n+1)$ 

In the case where  $U = \text{SO}(2n+1)$ , the Lie algebra  $\mathfrak{u} = \mathfrak{so}(2n+1)$  is a compact normal form of the classical Lie algebra  $\mathfrak{b}_n$ . The maximal torus  $T$  we choose is the set of  $(2n+1)$ -by- $(2n+1)$  block diagonal matrices of the form

$$\begin{pmatrix} R_{\theta_1} & 0_{2,2} & \cdots & 0_{2,2} & 0_{2,1} \\ 0_{2,2} & R_{\theta_2} & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0_{2,2} & 0_{2,1} \\ 0_{2,2} & \cdots & 0_{2,2} & R_{\theta_n} & \\ 0_{1,2} & \cdots & \cdots & 0_{1,2} & 1 \end{pmatrix},$$

where  $(\theta_1, \dots, \theta_n)$  belongs to  $\mathbb{R}^n$ . The Riemannian structure induced by the standard Killing form is given by  $\langle X, Y \rangle = -(2n-1) \text{tr}(XY)$  for any  $x, y$  in  $\mathfrak{so}(2n+1)$ . Let us introduce the  $(2n+1)$ -by- $(2n+1)$  block matrix  $N$  with entries  $n_{j,k}$  defined by  $n_{k+2,k} = 1$  for  $1 \leq k \leq n-2$ ,  $n_{1,2n-1} = n_{2,2n} = n_{2n+1,2n+1} = 1$  and  $n_{j,k} = 0$  otherwise.  $N$  is the block matrix

$$N = \begin{pmatrix} 0_{2,2} & 0_{2,2} & \cdots & 0_{2,2} & I_2 & 0_{2,1} \\ I_2 & 0_{2,2} & \ddots & \vdots & 0_{2,2} & 0_{2,1} \\ 0_{2,2} & I_2 & \ddots & 0_{2,2} & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0_{2,2} & \vdots & \vdots \\ 0_{2,2} & \cdots & 0_{2,2} & I_2 & 0_{2,2} & 0_{2,1} \\ 0_{1,2} & \cdots & \cdots & \cdots & 0_{1,2} & 1 \end{pmatrix}.$$

Then  $N T$ , the left translate of  $T$  by  $N$ , is the set of matrices of  $\text{SO}(2n+1)$  of the form

$$N^{(\theta_1, \dots, \theta_n)} = \begin{pmatrix} 0_{2,2} & 0_{2,2} & \cdots & 0_{2,2} & R_{\theta_n} & 0_{2,1} \\ R_{\theta_1} & 0_{2,2} & \ddots & \vdots & 0_{2,2} & 0_{2,1} \\ 0_{2,2} & R_{\theta_2} & \ddots & 0_{2,2} & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0_{2,2} & \vdots & \vdots \\ 0_{2,2} & \cdots & 0_{2,2} & R_{\theta_{n-1}} & 0_{2,2} & 0_{2,1} \\ 0_{1,2} & \cdots & \cdots & \cdots & 0_{1,2} & 1 \end{pmatrix}.$$

The computations are essentially the same as in the case of  $\text{SO}(2n)$ . Matrix  $N^{(\theta_1, \dots, \theta_n)}$  has the characteristic polynomial

$$P_{N^{(\theta_1, \dots, \theta_n)}}(X) = (1-X) \cdot (X^{2n} - 2 \cos(\theta) X^n + 1), \quad (4.2)$$

where  $\theta = \sum_{k=1}^n \theta_k$ . That is to say,  $N^{(\theta_1, \dots, \theta_n)}$  has eigenvalues 1 and

$$\lambda_k = \exp\left(i \frac{2\pi k + \theta}{n}\right) \quad \text{and} \quad \mu_k = \exp\left(i \frac{2\pi k - \theta}{n}\right), \quad k \in \{0, \dots, n-1\}. \quad (4.3)$$

We assume without loss of generality that  $\theta$  belongs to  $[0, 2\pi[$  if  $n$  is even, and that  $\theta$  belongs to  $[-\pi, \pi[$  if  $n$  is odd. Hence,

$$d_U(M^{(\theta_1, \dots, \theta_n)}, Id)^2 = \begin{cases} 2(2n-2) \left( \sum_{k=-n/2}^{n/2-1} \frac{(2k\pi+\theta)^2}{n^2} \right) & \text{if } n \text{ is even} \\ 2(2n-2) \left( \sum_{k=-(n-1)/2}^{(n-1)/2} \frac{(2k\pi+\theta)^2}{n^2} \right) & \text{if } n \text{ is odd,} \end{cases}$$

is minimum when  $\theta = \pi$  if  $n$  is even and  $\theta = 0$  if  $n$  is odd. Hence,  $n$  being odd or even,

$$d_T(\pi_T(Id), \pi_T(N)) = \frac{\pi}{3} \sqrt{\frac{6(n^2 - 1)(2n - 2)}{n}}. \tag{4.4}$$

This provides a lower bound for the diameter of the quotient of  $SO(2n + 1)$  by a maximal torus.

### 4.3.3. $U = Sp(2n)$

The classical symplectic Lie group  $Sp(2n)$  is the set of the linear unitary operators which preserve the canonical symplectic form of  $\mathbb{C}^{2n}$ . Its Lie algebra  $\mathfrak{sp}(n)$  can be seen as the set of the  $2n$ -by- $2n$  matrices with entries of the form

$$\begin{pmatrix} U & V \\ -\bar{V} & \bar{U} \end{pmatrix},$$

where  $U$  and  $V$  are two  $n$ -by- $n$  square matrices such that  $U$  is skew Hermitian and  $V$  is symmetric. The maximal torus we choose is the set of diagonal matrices with purely imaginary entries  $t_{j,j}$  with trace zero. The Euclidean form induced by the canonical Killing form is  $\langle X, Y \rangle = (2n + 2)\text{tr}(XY)$ , for any  $X, Y$  in  $\mathfrak{sp}(n)$ . Let us introduce matrix  $M$  defined by

$$\begin{cases} M_{j,k} = 1 & \text{if } 1 \leq j \leq n - 1 \text{ and } k = j + 1 \text{ or } (j, k) = (n, 1) \\ M_{j,k} = 0 & \text{otherwise.} \end{cases}$$

By a computation formally identical to the ones we have done in  $SO(2n)$  and  $SO(2n + 1)$  before, one sees that  $d_T(\pi_T(Id), \pi_T(M)) = \frac{\pi}{3} \sqrt{\frac{6(n+1)(2n+1)(2n+2)}{n}}$ . This provides a lower bound for the diameter of the quotient of  $Sp(2n)$  by a maximal torus.

## 4.4. The case of $SU(n)$

### 4.4.1. Result

**Theorem 2.11** *The diameter of  $SU(n)/T$  endowed with the projection of the standard Killing metric is*

$$\mathcal{D}(SU(n), d_T) = \frac{\pi}{3} \sqrt{6(n + 1)(n - 1)}.$$

*This diameter is the length of the optimal geodesics joining the identity class  $\pi_T(Id)$  to the images of matrices of  $n$ -cycles under the canonical projection  $\pi_T$ .*

### 4.4.2. The link with $PSU(n)$

Let us define  $PSU(n)$  by  $PSU(n) = SU(n)/\sim$  where  $\sim$  is the equivalence relation defined by:  $A \sim B$  if and only if there exists some  $\lambda \in \mathbb{C}$  such that  $A = \lambda B$ . Obviously, such a  $\lambda$  must be a root of 1 of order  $n$ . We denote by  $\Lambda$  the set of the  $n$  diagonal matrices  $\lambda Id$  where  $\lambda$  is a root of 1 of order  $n$ , and by  $\pi_P$  the canonical projection  $\pi_P : SU(n) \rightarrow PSU(n)$ . We define a mapping  $\phi : PSU(n) \rightarrow SU(n)/T$  in the following way: to any  $\tilde{x} = \pi_P(x)$  in  $PSU(n)$ , we associate  $\phi(\tilde{x}) = \pi_T(x)$ . This definition of  $\phi$  is consistent. The standard Killing metric of  $SU(n)$  provides a bi-invariant Riemannian structure with metric  $d_P$  on  $PSU(n)$ . Actually, for any  $\tilde{x} = \pi_P(x), \tilde{y} = \pi_P(y)$  in  $PSU(n)$ ,  $d_P(\tilde{x}, \tilde{y}) = d_U(x\Lambda, y\Lambda) = \inf\{d_U(x\lambda_1, y\lambda_2); \lambda_1, \lambda_2 \in \Lambda\}$ .



For any  $\bar{x} = \pi_T(x), \bar{y} = \pi_T(y)$  in  $T$ ,  $d_T(\bar{x}, \bar{y}) = \inf\{d_U(xt_1, yt_2); t_1, t_2 \in T\} \leq d_U(x\Lambda, y\Lambda)$  (since  $\Lambda \subset T$ ). This proves that  $\phi$  is a contraction from  $(\text{PSU}(n), d_P)$  to  $(\text{SU}(n)/T, d_T)$ .

**Proposition 4.6.** *The diameter of  $\text{SU}(n)/T$  is smaller than or equal to the diameter of  $\text{PSU}(n)$ .*

The situation is summarised in the following diagram.

$$\begin{array}{ccc} \text{SU}(n) & \xrightarrow{\text{Id}} & \text{SU}(n) \\ \downarrow \pi_P & & \downarrow \pi_T \\ \text{PSU}(n) & \xrightarrow{\phi} & \text{SU}(n)/T \end{array}$$

#### 4.4.3. The diameter of $\text{PSU}(n)$

##### Result

By  $n$ -cycles, we mean a permutation of the finite set with  $n$  elements  $\{1, \dots, n\}$  whose decomposition in cycles with disjoint supports consists in only one cycle (whose length is  $n$ ). Let us call  $n$ -cycle matrix of  $\text{SU}(n)$  any matrix  $M$  of  $\text{SU}(n)$  for which there exists an  $n$ -cycle  $\sigma$  such that entries  $m_{j,k}$  of  $M$  verify  $m_{j,k} = 0$  if and only if  $j \neq \sigma(k)$ . Let us denote by  $\mathcal{P}$  the set of all  $n$ -cycle matrices of  $\text{SU}(n)$ .

**Proposition 4.7.** *Let  $p_1$  in  $\mathcal{P}$  be any cyclic matrix and let  $\tilde{x}$  be any element of  $\text{PSU}(n)$ . Then,  $d_P(\pi_P(\text{Id}), \tilde{x}) \leq d_P(\pi_P(\text{Id}), \pi_P(p_1))$ .*

*Proof.* The proof of this apparently obvious statement is surprisingly intricate. We split it into three parts. The first one gives a rough idea of the proof, the second one introduces some notions needed in the following. The third one is the proof itself.

##### Idea of the proof

Let  $\tilde{x}$  be any point of  $\text{PSU}(n)$  and  $x$  be any point of  $\text{SU}(n)$  such that  $\pi_P(x) = \tilde{x}$ . There exist some diagonal matrix  $D$  and some matrix  $Q$  in  $\text{SU}(n)$  such that  $x = Q^{-1}DQ$ .  $D = \text{diag}(e^{i\theta_1}, \dots, e^{i\theta_n})$  with  $\theta_j \in [-\pi, \pi[$  for all  $j \in \{1, \dots, n\}$ . And  $d_U(\text{Id}, x) = \sqrt{2n \sum_1^n \theta_i^2}$  (we recall that  $\text{Kill}(\cdot, \cdot) = 2n \text{tr}(\cdot \cdot)$ ). As already said,  $d_P(\pi_P(\text{Id}), \tilde{x}) = \inf\{d_U(\text{Id}, \lambda x), \lambda \in \Lambda\}$ . For any  $\lambda \in \Lambda$ ,  $\lambda x = Q^{-1}\lambda DQ$ . The multiplication by  $\lambda$  can be seen as a rotation of the unit circle in  $\mathbb{C}$  of angle  $\frac{2k\pi}{n}$  for some integer  $k$ . The point is to see that the maximum of  $\min \left\{ \sum_j \left( \theta_j + \frac{2k\pi}{n} \right)^2, k \in \{1, \dots, n\} \right\}$  over all matrices  $x$  is reached when  $x$  is an  $n$ -cycle matrices. First of all, notice that for  $n$ -cycle matrices, multiplication by an element of  $\Lambda$  does not change anything. Actually, the set of eigenvalues of a  $n$ -cycle matrix is invariant under multiplication by a root of 1 of order  $n$ . (Alternatively, we can say that  $\mathcal{P}$  is invariant under multiplication by  $\Lambda$ , and all elements of  $\mathcal{P}$  are at the same distance of  $\text{Id}$ .)

We give now a rough sketch of the proof. Fix any element  $x$  of  $\text{SU}(n)$  which is not a  $n$ -cycle matrix. We shall exhibit another element  $y$  in  $\text{SU}(n)$ , arbitrarily close to  $x$ , such that  $d_P(\pi_P(\text{Id}), \pi_P(x)) < d_P(\pi_P(\text{Id}), \pi_P(y))$ . As already done, let us call  $\theta_1, \theta_2, \dots, \theta_n$  the arguments of the  $n$  eigenvalues of  $x$  such that  $-\pi \leq \theta_1 \leq \theta_2 \leq \dots \leq \theta_n \leq \pi$ . Since  $x$  is not an  $n$ -cycle matrix, the differences  $\theta_{k+1} - \theta_k$  and  $\theta_1 + 2\pi - \theta_n$  are not all equal to  $\frac{2\pi}{n}$ . At least one is less than  $\frac{2\pi}{n}$ . Hence, there exist some integers  $k_0, k_1$  such that  $\theta_{k_0}$  and  $\theta_{k_0+1}$  both belong to the closed set  $[(2k_1 + 1)\frac{\pi}{n}, (2k_1 + 3)\frac{\pi}{n}]$ , and at least one of the two points  $\theta_{k_0}$  and  $\theta_{k_0+1}$  belongs to the open set  $](2k_1 + 1)\frac{\pi}{n}, (2k_1 + 3)\frac{\pi}{n}[$ . In "the general case", whatever  $\lambda$  in  $\Lambda$  is, eigenvalues  $\lambda \exp(i\theta_{k_0})$  and  $\lambda \exp(i\theta_{k_0+1})$  of  $\lambda x$  will never have arguments differing from more than  $|\theta_{k_0+1} - \theta_{k_0}|$  - roughly speaking,  $-1$  will never be "between" the two eigenvalues of  $\lambda x$ , one argument being close to  $\pi$  and smaller, other arguments being close to  $-\pi$ , and bigger. Now, let us define  $y$  such that  $Q^{-1}yQ$  is diagonal with the same eigenvalues as  $x$  except  $\exp(i\theta_{k_0})$  and  $\exp(i\theta_{k_0+1})$  which are changed respectively into  $\exp(i(\theta_{k_0} - \epsilon))$  and  $\exp(i(\theta_{k_0+1} + \epsilon))$  for a small

enough strictly positive real number  $\epsilon$ . Then it is easy to check that  $d(Id, \Lambda y) > d(Id, \Lambda x)$ , which proves that  $\pi_P(x)$  was not the farthest point from  $\pi_P(Id)$  for distance  $d_P$ .

As can be seen at this point, a crucial issue is to understand how eigenvalues rotate around  $-1$ , where their argument jumps from  $\pi$  to  $-\pi$ .

Preliminaries

We introduce now some notions needed in the following.

**Definition 4.8.** Let  $M$  be a matrix in  $T$ , the maximal torus of  $SU(n)$  consisting of diagonal matrices with determinant 1 with all diagonal entries of modulus 1. We call *position* of  $M$  the set of the eigenvalues of  $M$ . We shall write an argument of a position  $\{\exp(i\mu_1), \dots, \exp(i\mu_n)\}$  in the form  $\{-\pi \leq \theta_1 \leq \dots \leq \theta_n < \pi\}$ , where the  $\exp(i\theta_j)$  are the diagonal entries of matrix  $M$ , ordered by increasing argument.

**Remark 4.9.** From a Lie group point of view, choosing this increasing rewriting of the arguments of the eigenvalues of a diagonal element of  $SU(n)$  is just a particular choice of a system of primitive roots in  $\mathfrak{t}$ .

The center  $Z$  of  $SU(n)$  acts by matrix multiplication on  $T$ , and on the set of eigenvalues of elements of  $T$ .

**Definition 4.10.** We call *configurations* the points of the quotient set of the set of the positions by the group action of  $Z$ .

As already said, the particular value  $r = \frac{2\pi}{n}$  will play an important role. In the following, it will be denoted by  $r$ . Let  $\{-\pi \leq \theta_1 \leq \dots \leq \theta_{n-p-1} < \pi - r \leq \theta_{n-p} \leq \dots \leq \theta_{n-p+p} < \pi\}$  be the arguments of a position, where  $p$  is a integer taking value in the set  $\{-1, \dots, n\}$ . By iteration of the action of the matrix  $e^{ir} Id$  (with  $r = \frac{2\pi}{n}$ ), we get successively the positions with arguments  $\theta_j^{(2)}$ :  $\{-\pi \leq \theta_1^{(2)} = \theta_{n-p} + r - 2\pi \leq \dots \leq \theta_{p+1}^{(2)} = \theta_n + r - 2\pi \leq \theta_{p+2}^{(2)} = \theta_1 + r \leq \theta_2 + r \leq \dots \leq \theta_n^{(2)} = \theta_{n-p+1} + r < \pi\}$ , and so on with argument  $\theta_j^{(3)}$  of the form  $\theta_j^{(3)} = \theta_k^{(2)} + r$ , except for the first ones  $\theta_j^{(3)} = \theta_k^{(2)} + r - 2\pi$ .

**Definition 4.11.** Let  $\{-\pi \leq \theta_1 \leq \dots \leq \theta_n < \pi\}$  be the arguments of a position  $P$ . We call *weight* of this position and denote by  $\mathcal{W}(P)$  the real positive number  $\sum_{j=1}^n \theta_j^2$ . We call *weight* of a configuration the minimum of the weights of the associated positions.

The weight of a configuration is nothing other than  $(1/2n \text{ times})$  the square of the distance between the associated point in  $U/Z$  and  $\pi_P(Id)$ .

**Remark 4.12.** This last definition just uses the classical fact that every point of a compact Lie group is conjugated by an adjoint transformation to an element of a fixed maximal torus  $T$ . Since the adjoint transformation preserves the Killing metric (bi-invariance of the Killing form), distances can be evaluated in a torus. The point here is that the action of  $Z$  commutes with the action of  $T$ , and it is easy to compute the weights of the different positions of a configuration.

As already said, the “density” of the points of a configuration will play an important role.

**Definition 4.13.** Let  $\{-\pi \leq \theta_1 \leq \dots \leq \theta_n < \pi\}$  be the arguments of a given position. We call *intervals* the couples  $\{(\theta_n, \theta_1 + 2\pi), (\theta_1, \theta_2), (\theta_2, \theta_3), \dots, (\theta_{n-1}, \theta_n)\}$ . An interval  $(a, b)$  of length  $\delta = b - a$  will be called *short* if  $\delta < r$ , *medium* if  $\delta = r$  and *large* if  $\delta > r$ .

The set of the lengths of the intervals does not depend on the position. It just depends on the configuration.

Two configurations may have the same set of values for the lengths of their intervals, but be different, when the intervals are not in the same order.

There is only one configuration for which every interval is medium. This configuration is obtained with  $n$ -cycle matrices. Indeed, a position of such a configuration must have arguments  $\{\theta - k_1 r, \dots, \theta - r, \theta, \theta + r, \dots, \theta + k_2 r\}$ , for some real  $\theta$  in  $[0, r[$ , and two positive integers  $k_1, k_2$  such that  $k_1 = k_2$  if  $n$  is odd, and  $k_1 = k_2 + 1 = n/2$  if  $n$  is even. The sum of the arguments is  $n\theta$  if  $n$  is odd and  $n\theta - k_1 r$  if  $n$  is even. This sum must be equal to  $2p\pi$  for a certain (positive or negative, possibly zero) integer  $p$ . Hence,  $\theta = 2p\pi/n$  if  $n$  is odd, and  $\theta = \pi/n + pr$  if  $n$  is even, which gives the result.

The idea is now that if a certain point  $x$  of  $SU(n)$  does not have the same eigenvalues as the  $n$ -cycle matrices, then the configuration associated with  $x$  contains at least one short (and hence one large) interval. Making this short interval bigger will *usually* give a configuration more distant from  $\pi Id$ .

**Definition 4.14.** A short interval is said to be *both-sided* if the corresponding segment  $\theta_j, \theta_{j+1}$  or  $\theta_n, \theta_1 + 2\pi$  contains  $\pi$  or  $-\pi$ . A short interval will be said to be *potentially both-sided* if there is a position of the same configuration such that the (image of the) interval is both-sided.

The “general” (good) case is when no interval is potentially both-sided. In this case, the rough proof presented in the previous paragraph is valid. We introduce now a useful tool to decide whether a position contains potentially both-sided intervals or not.

**Definition 4.15.** Let  $\{-\pi \leq \theta_1 \leq \dots \leq \theta_n < \pi\}$  be the arguments of a position  $P$ . We call *height* of this position and denote by  $\mathcal{H}(P)$  the real number  $\sum_{j=1}^n \theta_j$ .

The following lemma describes the variation of the height for a given configuration.

**Lemma 4.16.** *The set of the heights taken by the various positions of a given configuration has the form  $2\pi I$  where  $I$  is a segment of  $\mathbb{Z}$ , containing 0.*

**Remark 4.17.** A configuration associated with  $n$ -cycle matrices has constant zero height, in any position. There exist configurations not associated with  $n$ -cycle matrices such that the height is constantly equal to zero for all their positions.

*Proof.* First note that, since we are dealing with matrices of  $SU(n)$ , the height of a position is an integer multiple of  $2\pi$ . Now, fix an arbitrary configuration, and choose an arbitrary position  $P_0$  of this configuration with arguments  $\{-\pi \leq \theta_1 \leq \dots \leq \theta_{n-p-1} < \pi - r \leq \theta_{n-p} \leq \dots \leq \theta_n < \pi\}$  ( $p$  is an integer which can take value  $-1$  if there is no eigenvalue with argument between  $\pi - r$  and  $\pi$ ). An iteration of the elementary rotation of angle  $r$  changes the value of the height in the following way:

- if  $p = -1$ , there is no point with argument between  $\pi - r$  and  $\pi$  in  $P_0$ , the height increases by  $2\pi$ ;
- if  $p = 0$ , there is one point with argument between  $\pi - r$  and  $\pi$  in  $P_0$ , the height does not change;
- if  $p \geq 1$ , there are  $p + 1$  points with argument between  $\pi - r$  and  $\pi$  in  $P_0$ , the height decreases by  $(p - 1)2\pi$ .

Hence the height can decrease by a maximum of  $2\pi(n - 1)$  but can only increase by  $2\pi$  after a rotation of the position of angle  $r$ . After  $n$  iterations of this elementary rotation, we find again the initial position, hence the height returns to its original level. This proves that the set of the heights taken by the various positions of a configuration has the form  $2\pi I$  where  $I$  is a segment of  $\mathbb{Z}$ . Let us now compute the average height, that is the sum of all values taken by the heights of a configuration, divided by  $n$ .

**If  $n$  is odd:** for each argument  $\theta_j$  of a point  $j$  of the initial position, there is an integer  $k_j$  such that  $\mu_j = \theta_j + k_j r \in [-\frac{r}{2}, \frac{r}{2}]$ . The list of all values for the argument corresponding to point  $j$  under iterations of the elementary rotation is  $\{\mu_j - E(\frac{n}{2})r, \dots, \mu_j - r, \mu_j, \mu_j + r, \dots, \mu_j + E(\frac{n}{2})r\}$ , where  $E$  denotes the integer part function. The average of this list is then  $\mu_j$ .

**If  $n$  is even:** for each argument  $\theta_j$  of a point  $j$  of the initial position, there is an integer  $k_j$  such that  $\mu_j = \theta_j + k_j r \in [0, r[$ . The list of all values for the argument corresponding to point  $j$  is now  $\{\mu_j - \frac{n}{2}r, \dots, \mu_j - r, \mu_j, \mu_j + r, \dots, \mu_j + E(\frac{n-1}{2})r\}$ . The average of this list is  $\mu_j - \frac{r}{2}$ .

In both case, the average of the heights belongs to segment  $[-\frac{r}{2}, \frac{r}{2}]$ . Let us proceed by contradiction, and let us assume now that a configuration has only non zero height - for instance only strictly positive height. Then the average height has to be at least  $2\pi > \frac{r}{2} = \frac{\pi}{n}$ . This gives a contradiction with the preceding computation, and proves that 0 belongs to  $I$ . For configurations having an interval of the form  $\pi - r \leq \theta_k \leq \theta_{k+1} < \pi$ , the height is not constant and  $I$  contains at least 1 or  $-1$ . But any small enough deformation of the configuration associated with  $n$ -cycle matrices contains a small interval, and has constant height equal to zero. This completes the proof of Lemma 4.16.  $\square$

From the proof of Lemma 4.16, we can deduce that a configuration with non constant height has maximum height just before a decrease at a position with arguments of the form  $\{-\pi \leq \theta_1 \leq \dots \leq \theta_{n-p-1} < \pi - r \leq \theta_{n-p} \leq \dots \leq \theta_n < \pi\}$  with  $p \geq 1$ . Interval  $\theta_{n-p}, \theta_{n-p+1}$  is short. It is also potentially both-sided if and only if  $\theta_{n-p} = \pi - r$ .

The minimum is reached just after a decrease in the height, that is to say at a position with arguments of the type  $\{-\pi \leq \theta_1 \leq \theta_2 \dots \leq \theta_p < -\pi + r \leq \theta_{p+1} \leq \dots \leq \theta_{n-1} < \pi - r \leq \theta_n < \pi\}$  with interval  $\theta_1, \theta_2$  short, and potentially both-sided if and only if  $\theta_1 = -\pi$ .

*Proof of Proposition 4.7.* Let  $\mathcal{C}$  be a configuration containing some short interval, and let one of its position  $P$  be given. We will exhibit another configuration arbitrarily close to the previous one, but with bigger weight. The general case is very easy:

If there is one not potentially both-sided short interval  $(\theta_j, \theta_{j+1})$  in  $P$ , then the configuration admitting position  $\{\theta_1, \dots, \theta_{j-1}, \theta_j - \epsilon, \theta_{j+1} + \epsilon, \dots, \theta_n\}$ , with  $\epsilon$  a small enough strictly positive number, has bigger weight than  $\mathcal{C}$ . Indeed, the fact that  $(\theta_j, \theta_{j+1})$  is not potentially both-sided means  $\pi - kr < \theta_j \leq \theta_{j+1} < \pi - (k-1)r$  for some integer  $k$ . Let us choose a strictly positive number  $\epsilon$  small enough. Then  $\pi - kr < \theta_j - \epsilon < \theta_{j+1} + \epsilon < \pi - (k-1)r$ . Consider now the configuration  $\mathcal{C}_\epsilon$  admitting argument  $\{\theta_1 \leq \theta_2 \leq \dots \leq \theta_j - \epsilon \leq \dots \leq \theta_{j+1} + \epsilon \leq \dots \leq \theta_n\}$ . Any position  $P_\epsilon$  of  $\mathcal{C}_\epsilon$  has now a bigger weight than the corresponding position  $P$  of  $\mathcal{C}$ . Hence configuration  $\mathcal{C}_\epsilon$  has a bigger weight than configuration  $\mathcal{C}$ .

This simple proof could be extended to the case where all short intervals are potentially both-sided, if we were able to find an interval such that the position where it is both-sided had weight strictly bigger than the weight of the configuration. By reason of the continuity of the weight on the set of the configuration (minimum of a finite set of polynomial functions), for  $\epsilon$  small enough, the weight of position  $\mathcal{C}_\epsilon$  will not be reached in the position where the interval is both-sided, and the preceding method applies.

Hence, it suffices to check that, if  $\mathcal{C}$  is a configuration where all short intervals are potentially both-sided, there is at least one short interval such that the weight of the position  $P_0$ , where it is both-sided, is strictly bigger than the weight of the configuration  $\mathcal{C}$ .

For the proof, we distinguish two cases, since the configuration  $\mathcal{C}$  may have constant or not constant height.

If  $\mathcal{C}$  has constant zero height then any of its position admits arguments of the form  $\{\theta_k, k = 1 \dots n\}$  with  $-\pi + (k-1)r \leq \theta_k < -\pi + kr$  for all  $k$  in  $1 \dots n$ . Let us define  $\{\mu_k, k = 1 \dots n\}$  by  $\mu_k = -\pi + \pi/n + (k-1)r$  and  $\epsilon_k = \theta_k - \mu_k$  for all  $k$  in  $1 \dots n$ . The set  $\{\mu_k, k = 1 \dots n\}$  is nothing other than the list of arguments of any position of the configuration associated with  $n$ -cycle matrices.  $\sum_k \theta_k = \sum_k \mu_k = 0$  so  $\sum_k \epsilon_k = 0$ . Since  $\mathcal{C}$  is not associated to an  $n$ -cycle matrix, there is at least one  $\epsilon_{k_0}$  which is not zero, and at least one strictly positive  $\epsilon_{k_1}$  such that  $\epsilon_{k_1+1}$  is less than or equal to zero. (We have committed the obvious abuse  $k_{n+1} = k_1$ .) Hence we have found a position  $P_0$  of  $\mathcal{C}$  admitting arguments  $\{-\pi \leq \theta_1 \leq -\pi + r/2 < -\pi + r \leq \theta_2 < -\pi + 2r \leq \dots \leq \pi - r < \pi - r/2 \leq \theta_n < \pi\}$ . Let us note  $P_1$  the position obtained from  $P_0$  by a rotation of angle  $r$ .  $P_1$  admits arguments  $\{-\pi < -\pi + r/2 \leq \theta_n + r - 2\pi < -\pi + r \leq \theta_1 + r \leq \dots \leq \theta_{n-1} + r\}$ . Let us compare the weights of these two positions  $P_0$  and  $P_1$ .

$$\mathcal{W}(P_0) = \sum_{k=1}^n \theta_k^2, \tag{4.5}$$

$$\mathcal{W}(P_1) = \sum_{k=1}^{n-1} (\theta_k + r)^2 + (\theta_n - 2\pi + r)^2, \tag{4.6}$$

$$\mathcal{W}(P_1) - \mathcal{W}(P_0) = 2r \sum_{k=1}^n \theta_k + nr^2 + 4\pi(\pi - (\theta_n + r)), \tag{4.7}$$

$$\mathcal{W}(P_1) - \mathcal{W}(P_0) = \underbrace{\frac{4\pi^2}{n}}_a + \underbrace{4\pi(\pi - (\theta_n + r))}_b. \tag{4.8}$$

But  $\pi - r/2 < \theta_n < \pi$  so  $b < -2\pi r = -\frac{4\pi^2}{n} = -a$ . Hence  $\mathcal{W}(P_0) > \mathcal{W}(P_1)$ , and position  $P_0$ , where couple  $(\theta_n, \theta_1)$  is both-sided, has weight strictly greater than  $\mathcal{W}(\mathcal{C})$ .

*If  $\mathcal{C}$  has non constant height* From the proof of Proposition 4.16, we can deduce that the maximum for the height for this configuration is reached just before a decrease at a position with arguments of the type  $\{-\pi \leq \theta_1 \leq \dots \leq \theta_{n-p-1} < \pi - r = \theta_{n-p} = \dots = \theta_{n-1} \leq \theta_n < \pi\}$  with  $p \geq 1$ , and interval  $(\theta_{n-p}, \theta_{n-p+1})$  is short and potentially both-sided. The minimum is reached just after a decrease in the height, that is to say at a position with arguments of the type  $\{-\pi = \theta_1 \leq \theta_2 \leq \dots \leq \theta_p < -\pi + r \leq \theta_{p+1} \leq \dots \leq \theta_{n-1} < \pi - r \leq \theta_n < \pi\}$  with interval  $\theta_1, \theta_2$  short and both-sided.

Let  $P_0$  be a position of  $\mathcal{C}$  with height  $\mathcal{H}(P_0)$  reaching the extremum of the modulus of the height of the various positions of  $\mathcal{C}$ . We have to distinguish two cases, since  $\mathcal{H}(P_0)$  may be negative or positive.

$\mathcal{H}(\mathbf{P}_0) < \mathbf{0}$ . Then  $P_0$  is a minimum for the height.  $P_0$  has argument  $\{-\pi = \theta_1 \leq \theta_2 \dots \leq \theta_p < -\pi + r \leq \theta_{p+1} \leq \dots \leq \theta_n < \pi\}$ . Then, position  $P_1$  obtained by rotation of angle  $r$  from position  $P_0$  has arguments  $\theta_{n-p+1} + r - 2\pi \leq \dots \theta_n + r - 2\pi < -\pi + r = \theta_1 + r \leq \dots \theta_{n-p} + r < \pi$ . Denoting as usual with  $\mathcal{W}(\cdot)$  the weight of a position

$$\mathcal{W}(P_0) = \sum_{j=1}^n \theta_j^2, \tag{4.9}$$

$$\mathcal{W}(P_1) = \sum_{j=1}^n \theta_j^2 + nr^2 + 2r \sum_{j=1}^n \theta_j + 4\pi(\pi - \theta_n - r). \tag{4.10}$$

Hence

$$\mathcal{W}(P_1) - \mathcal{W}(P_0) = \underbrace{r(n.r + 2 \sum_{j=0}^n \theta_j)}_a + \underbrace{4\pi(\pi - \theta_n - r)}_b, \tag{4.11}$$

but  $nr + 2 \sum_{j=0}^n \theta_j < 0$  so  $a < 0$  and  $\theta_n + r > \pi$  so  $b < 0$  and position  $P_1$  has a lower weight than position  $P_0$  where the interval  $(\theta_n, \theta_1)$  is both-sided.

$\mathcal{H}(\mathbf{P}_0) > \mathbf{0}$ . Then  $P_0$  is a maximum for the height.  $P_0$  has argument  $\{-\pi + r \leq \theta_1 \leq \dots \leq \theta_{n-p-1} < \pi - r = \theta_{n-p} = \dots = \theta_{n-1} \leq \theta_n < \pi\}$ . Position  $P_1$  obtained by rotation of angle  $r$  from position  $P_0$  has arguments  $\{-\pi = \theta_{n-p} + r - 2\pi = \dots = \theta_{n-1} + r - 2\pi \leq \theta_n + r - 2\pi < -\pi + r \leq \theta_1 + r \leq \dots \leq \theta_{n-p-1} + r < \pi\}$  and the interval  $(\theta_{n-1} + r - 2\pi, \theta_n + r - 2\pi)$  is short and both-sided. One has

$$\mathcal{W}(P_0) = \sum_{j=1}^n \theta_j^2, \tag{4.12}$$

$$\mathcal{W}(P_1) = \sum_{j=1}^{n-p-1} (\theta_j + r)^2 + \sum_{k=n-p}^n (\theta_k + r - 2\pi)^2. \tag{4.13}$$

Hence,

$$\mathcal{W}(P_1) - \mathcal{W}(P_0) = \underbrace{2r \sum_{j=1}^n \theta_j}_a + 4\pi \underbrace{\left( (p+1)(\pi - r) - \sum_{k=n-p}^n \theta_k \right)}_b. \tag{4.14}$$

But  $a \geq 4\pi\frac{2\pi}{n} = 8\frac{\pi^2}{n}$  and  $(p+1)(\pi-r) \leq \sum_{k=n-p}^n \theta_k < p(\pi-r) + \pi$ , so  $b > -4\pi(\frac{2\pi}{n}) = -8\frac{\pi^2}{n}$  and  $W(P_1) - W(P_0) > 0$ . Hence position  $P_0$  has a lower weight than position  $P_1$ , where the couple  $\theta_{n-1} + r - 2\pi, \theta_n + r - 2\pi$  is both-sided. This completes the proof of Proposition 4.7.  $\square$

4.4.4. Diameter of  $SU(n)/T$

**Lemma 4.18.** *For any  $p_1$  in  $\mathcal{P}$ ,  $d_P(\pi_P(Id), \pi_P(p_1)) = d_T(\pi_T(Id), \pi_T(p_1))$*

*Proof.* It suffices to see that if  $p_1$  is some  $n$ -cycle matrix, then  $d_P(\pi_P(Id), \pi_P(x)) = d_T(\pi_T(Id), \pi_T(x)) = d_U(Id, x)$ .

Each  $p_1$  in  $\mathcal{P}$  admits the same characteristic polynomial  $X^n + (-1)^n$ , and hence has eigenvalues  $\{\exp(i(-\pi + \pi/n + k.r)), k \in [0 \dots n - 1]\}$ . As already said, for any  $\lambda$  in  $T$ , and hence for any  $\lambda$  in  $\Lambda$ ,  $p_1$  and  $\lambda p_1$  have the same spectrum. Hence every position associated with an  $n$ -cycle matrix have the same weight

$$\mathcal{W}(p_1) = \sum_{k=0}^{n-1} (-\pi + \pi/n + k.r), \tag{4.15}$$

$$\mathcal{W}(p_1) = \frac{8\pi^2}{n^2} \sum_{k=0}^{\frac{n-1}{2}} k^2 \text{ if } n \text{ is odd,} \tag{4.16}$$

$$\mathcal{W}(p_1) = \frac{2\pi^2}{n^2} \sum_{k=0}^{\frac{n}{2}} (2k+1)^2 \text{ if } n \text{ is even.} \tag{4.17}$$

In both cases,

$$\mathcal{W}(p_1) = \frac{\pi^2}{3} \frac{(n+1)(n-1)}{n}. \tag{4.18}$$

Hence, in both cases,  $d_U(Id, p_1) = \sqrt{2n\mathcal{W}(p_1)} = \frac{\pi}{3} \sqrt{6(n+1)(n-1)}$ .  $\square$

*Proof of Theorem 2.11.* For all points of  $\mathcal{P}$ ,  $d_U(Id, x) = d_P(\pi_P(Id), \pi_P(x)) = d_T(\pi_T(Id), \pi_T(x))$ , and it proves that the diameter  $\mathcal{D}(SU(n)/T, d_T)$  is greater than the diameter  $\mathcal{D}(PSU(n), d_P)$ . Hence the diameter  $\mathcal{D}(PSU(n), d_P)$  and the diameter  $\mathcal{D}(SU(n)/T, d_T)$  are the same, since  $d_P(\pi_P(Id), \pi_P(x)) \geq d_T(\pi_T(Id), \pi_T(x))$  for any point  $x$  of  $SU(n)$ .  $\square$

*Proof of Corollary 2.12.* Assume now that the projection  $\pi_{\mathfrak{t}^\perp}(\mathcal{X})$  of  $\mathcal{X}$  on  $\mathfrak{t}^\perp$  contains a ball for the Killing norm centered on  $0_{\mathfrak{u}}$ , with radius  $a/\sqrt{n}$ . This is the case, if for each  $x$  in  $\mathcal{B}^\perp$ ,  $a/\sqrt{2} x$  belongs to  $\mathcal{X}$ . We are in this situation for example when  $\min_{\alpha \in \Delta} \sqrt{a_\alpha^{0^2} + a_\alpha^2} \geq a$ . Then  $\mathcal{D}(U, \Sigma_1)$  is less than  $a/\sqrt{n}$  times the diameter of  $SU(n)/T$  endowed with the projection of the standard Killing metric.

On the other hand,  $\mathcal{D}(U, \Sigma_1)$  is more than  $1/\sqrt{\langle h_0, h_0 \rangle}$  times the diameter of  $SU(n)/T$  endowed with the projection of the standard Killing metric.  $\square$

An example of application in  $SU(3)$  of Corollary 2.12 is given in Section 6.3.

## 5. THE DEFORMED KILLING METRIC AND ITS CURVATURE TENSOR

### 5.1. The deformed Killing metric

In Section 3, we have seen that, for generic  $h_0$  and  $h_1$ , the system  $\Sigma_1$  was equivalent to a left invariant control system in  $U$ , with a full dimensional control set  $\mathcal{Y}$ . In Section 4, we have approximated  $\mathcal{Y}$  from above and from below by balls for the Euclidian metric induced by the standard Killing form. This approach has given some results, but remains quite far from the geometry of the problem. Indeed, the tangent space  $\mathfrak{u}$  splits in a free

part  $\mathfrak{t}$ , in whose direction the velocity is arbitrarily large, and the orthogonal part  $\mathfrak{t}^\perp$ , in which the control set  $\mathcal{X}$  is a compact convex neighbourhood of  $0_{\mathfrak{t}^\perp}$ . It is important to distinguish between the free part (along  $\mathfrak{t}$ ) and the non free part (along  $\mathfrak{t}^\perp$ ), but it is also important to distinguish between directions  $\partial_\alpha$  in  $\mathcal{X}$  with high cost (when  $a_\alpha^{0^2} + a_{\bar{\alpha}}^{0^2}$  is low), and directions with low cost (when  $a_\alpha^{0^2} + a_{\bar{\alpha}}^{0^2}$  is high).

In the following, we define a sequence  $\langle \cdot, \cdot \rangle_{D_\lambda}$  of Riemannian structures on  $\mathfrak{su}(n)$ , for which the square of the norm  $\langle h, h \rangle_{D_\lambda}$  of any fixed vector  $h$  in  $\mathfrak{t}$  will tend to zero as  $\lambda$  tends to infinity, and which distinguishes between the different directions in  $\mathcal{X}$ . The highest lower bound of the sequence of associated Ricci curvatures will provide an estimation of the diameter of the group.

**Definition 5.1.** Let  $U$  be a given semi-simple compact Lie group, with Lie algebra  $\mathfrak{u}$ , and let  $\mathfrak{t}$  be a Cartan subalgebra. We will use the notation of the decomposition (2.1). Let  $D_\lambda = \{\lambda, d_\alpha, \alpha \in \Delta\}$  be a set of strictly positive real numbers. We call  $D_\lambda$ -deformed Killing form and denote by  $\langle \cdot, \cdot \rangle_{D_\lambda}$  the unique symmetric bilinear form on  $\mathfrak{u}$  such that  $\langle \cdot, \cdot \rangle_{D_\lambda}$  is  $-\lambda$  times the usual Killing form on  $\mathfrak{t}$ , and the orthogonal for  $\mathfrak{t}$  under  $\langle \cdot, \cdot \rangle_{D_\lambda}$  and the usual Killing form are the same,  $\langle \partial_\alpha, \partial_\alpha \rangle_{D_\lambda} = \langle \partial_{\bar{\alpha}}, \partial_{\bar{\alpha}} \rangle_{D_\lambda} = d_\alpha^2$  and  $\langle \partial_\alpha, \partial_\beta \rangle_{D_\lambda} = \langle \partial_\alpha, \partial_{\bar{\beta}} \rangle_{D_\lambda} = \langle \partial_\alpha, \partial_{\bar{\beta}} \rangle_{D_\lambda} = \langle \partial_{\bar{\alpha}}, \partial_{\bar{\beta}} \rangle_{D_\lambda} = 0$  for any  $\alpha$  and  $\beta$  in  $\Delta$  such that  $\alpha \neq \beta$ .

**Remark 5.2.** This is the same as saying that a deformed Killing form is a bilinear symmetric form for which the basis obtained in equation (2.1) is orthogonal and whose restriction to  $\mathfrak{t}$  is a multiple of the usual Killing form.

**Definition 5.3.** A deformed Killing form being given, we call *deformed Killing metric* the associated Riemannian structure.

Although these definitions make sense in any real semi-simple compact Lie group, from now on, we restrict ourselves to the case where  $U = \text{SU}(n)$ , in order to use the very special structure of the root pattern in this case. The following can be generalized to other cases, but the final expression of the result may be not so nice. Recall that  $U = \text{SU}(n)$  is the general frame of quantum mechanics.

Let us fix some constants  $d_1, \dots, d_n$ , and let us choose an arbitrary strictly positive real number  $\lambda$ . We will compute the Riemannian connection (or rather the associate covariant derivative)  $\nabla_X^{D_\lambda} Y$  and the Ricci curvature  $\mathcal{R}_{D_\lambda}$  associated with these constants. Then, we will apply Myers theorem to get an estimation of the diameter of  $\text{SU}(n)$  for the  $D_\lambda$  Killing metric. This expression will depend on  $\lambda$ . Finally, we will let  $\lambda$  tend to infinity.

## 5.2. The Riemannian connection of the $\lambda$ -deformed Killing metric

This section is dedicated to the computation of the Riemannian connection and the covariant derivation  $\nabla^{D_\lambda}$  of the  $D_\lambda$ -deformed Killing form. For precise definitions of the Riemannian connection and the associated covariant derivation of Riemannian structure, see for instance [16, 24, 34]. The result is as follows:

**Proposition 5.4.** For any  $j$  and  $k$  in  $\mathcal{B}$ , there exists some (finite) real numbers  $C_{j,k}^{D_\lambda}$  such that  $\nabla_j^{D_\lambda} k = C_{j,k}^{D_\lambda} [j, k]$ . If  $[j, k] \neq 0$ , then  $C_{j,k}^{D_\lambda} = \frac{1}{2\langle [j, k], [j, k] \rangle_{D_\lambda}} (\langle [j, k], [j, k] \rangle_{D_\lambda} - \langle k, [j, [j, k]] \rangle_{D_\lambda} - \langle j, [k, [j, k]] \rangle_{D_\lambda})$ , otherwise  $C_{j,k}^{D_\lambda} = 0$ .

The proof is just an adaptation of the classical computation of the Riemannian connection of a left invariant metric on a Lie group (see [34] or [35] for example). Using classical properties of the connection, we get an expression depending only on Lie bracket, and then we use the special structure of  $\mathfrak{su}(n)$ .

**Lemma 5.5.** For any  $j, k, l$  in  $\mathcal{B}$ ,  $\langle \nabla_j^{D_\lambda} k, l \rangle_{D_\lambda} = \frac{1}{2} (\langle l, [j, k] \rangle_{D_\lambda} - \langle k, [j, l] \rangle_{D_\lambda} - \langle j, [k, l] \rangle_{D_\lambda})$ .

*Proof.* It is known (see [34]) that for any vector fields  $X, Y$  and  $Y'$

$$X_p \langle Y, Y' \rangle_{D_\lambda} = \langle \nabla_{X_p}^{D_\lambda} Y, Y' \rangle_{D_\lambda} + \langle Y_p, \nabla_{X_p}^{D_\lambda} Y' \rangle_{D_\lambda}. \quad (5.1)$$

The connection being symmetric, one has moreover

$$\nabla_X^{D_\lambda} Y - \nabla_Y^{D_\lambda} X = [X, Y]. \tag{5.2}$$

Replacing  $X, Y, Y'$  respectively by  $j, k, l$  in (5.1), we obtain

$$\langle \nabla_j^{D_\lambda} k, l \rangle_{D_\lambda} + \langle \nabla_j^{D_\lambda} l, k \rangle_{D_\lambda} = 0. \tag{5.3}$$

By permutation on the set of indices  $\{j, k, l\}$

$$\langle \nabla_k^{D_\lambda} l, j \rangle_{D_\lambda} + \langle \nabla_k^{D_\lambda} j, l \rangle_{D_\lambda} = 0, \tag{5.4}$$

$$\langle \nabla_l^{D_\lambda} j, k \rangle_{D_\lambda} + \langle \nabla_l^{D_\lambda} k, j \rangle_{D_\lambda} = 0. \tag{5.5}$$

The standard method is now to apply equality (5.2) to these equations to get a linear system in the variables  $\langle \nabla_j^{D_\lambda} k, l \rangle_{D_\lambda}$ ,  $\langle \nabla_k^{D_\lambda} l, j \rangle_{D_\lambda}$  and  $\langle \nabla_l^{D_\lambda} j, k \rangle_{D_\lambda}$  only.

After this transformation, one has

$$\begin{cases} \langle \nabla_j^{D_\lambda} k, l \rangle_{D_\lambda} + \langle \nabla_l^{D_\lambda} j, k \rangle_{D_\lambda} &= \langle [l, j], k \rangle_{D_\lambda} \\ \langle \nabla_j^{D_\lambda} k, l \rangle_{D_\lambda} + \langle \nabla_k^{D_\lambda} l, j \rangle_{D_\lambda} &= \langle [j, k], l \rangle_{D_\lambda} \\ \langle \nabla_k^{D_\lambda} l, j \rangle_{D_\lambda} + \langle \nabla_l^{D_\lambda} j, k \rangle_{D_\lambda} &= \langle [k, l], j \rangle_{D_\lambda} \end{cases}$$

and after elimination, we deduce that for every vector  $j, k, l$  of the basis  $\mathcal{B}$

$$\langle \nabla_j^{D_\lambda} k, l \rangle_{D_\lambda} = \frac{1}{2}(\langle [l, j], k \rangle_{D_\lambda} + \langle [j, k], l \rangle_{D_\lambda} - \langle [k, l], j \rangle_{D_\lambda}). \tag{5.6}$$

This completes the proof of Lemma 5.5. □

*Proof of Proposition 5.4.* The result of Appendix B proves that the scalar product  $\langle \nabla_j^{D_\lambda} k, l \rangle_{D_\lambda}$  is zero if  $l$  is not colinear to  $[j, k]$ . Dividing equation (5.6) by the square of the norm of  $[j, k]$  gives an expression for  $C_{j,k}^{D_\lambda}$ . This completes the proof of Proposition 5.4. □

Let us note that if  $w$  and  $y$  are two vectors of  $\mathfrak{u}$  colinear respectively to two elements  $j$  and  $k$  of the basis  $\mathcal{B}_{\mathfrak{t}^\perp}$ , then we still have  $\nabla_w^{D_\lambda} y = C_{j,k}^{D_\lambda}[w, y]$ . By extension, if  $w \neq 0_{\mathfrak{u}}$  and  $y \neq 0_{\mathfrak{u}}$ , we shall note indifferently  $C_{j,k}^{D_\lambda}$  or  $C_{w,y}^{D_\lambda}$ . Because of the homogeneity of the coefficients  $C_{j,k}^{D_\lambda}$ , this notation is coherent with what precedes (and especially with the result of Prop. 5.4), and expression of the form of  $C_{x,[y,z]}^{D_\lambda}$  makes sense when  $x, y, z$  belongs to  $\mathcal{B}$ . Indeed, if  $[x, [y, z]] = 0$ ,  $C_{x,[y,z]}^{D_\lambda} = 0$ , otherwise we use the preceding convention. This notation will be useful when computing the Ricci curvature of the  $D_\lambda$  deformed Killing metric.

The behavior of the coefficients  $C_{j,k}^{D_\lambda}$  as  $\lambda$  tends to infinity can easily be obtained from the expression of Lemma 5.4.

**Lemma 5.6.** *For any  $j, k$  in  $\mathcal{B}$ ,  $C_{j,k}^{D_\lambda}$  admits a finite limit  $C_{j,k}^D$  when  $\lambda$  tends to  $\infty$ .*

*Proof.* If  $[j, k] = 0$ , the result is obvious. Else, we can write  $C_{j,k}^{D_\lambda} = \frac{1}{d}(a - b - c)$  with

$$a = \langle [j, k], [j, k] \rangle_{D_\lambda} \tag{5.7}$$

$$b = \langle k, [j, [j, k]] \rangle_{D_\lambda} \tag{5.8}$$

$$c = \langle j, [k, [j, k]] \rangle_{D_\lambda} \tag{5.9}$$

$$d = 2\langle [j, k], [j, k] \rangle_{D_\lambda} \tag{5.10}$$



$a/d$  is a constant equal to  $1/2$ . Now, it is enough to study the behavior of each of the terms  $b$ ,  $c$  and  $d$  as  $\lambda$  grows to infinity.

- $j$  in  $\mathfrak{t}^\perp$  and  $k$  in  $\mathfrak{t}$ ,  $b$  tends to zero as  $\lambda$  tends to infinity, and  $d$  is constant.  $a - c$  is constant equal to zero. So  $C_{j,k}^{D_\lambda}$  tends to zero as  $\lambda$  tends to infinity, and is equivalent to  $-b/d$ .
- $j$  in  $\mathfrak{t}$  and  $k$  in  $\mathfrak{t}^\perp$   $-b$  is constant equal to  $a$ ,  $d$  is constant.  $c$  takes the form  $\frac{c_1}{\lambda^2}$ , where  $c_1$  is independent of  $\lambda$ , and  $C_{j,k}^{D_\lambda}$  tends to one as  $\lambda$  grows to infinity.
- $j$  and  $k$  in  $\mathfrak{t}^\perp, j \neq \bar{k}$   $a, b, c$  and  $d$  are constant. Hence  $C_{j,k}^{D_\lambda}$  is constant.
- $j$  and  $k$  in  $\mathfrak{t}^\perp, j = \bar{k}$   $a$  has the form  $\frac{a_1}{\lambda^2}$ .  $b$  and  $c$  are constant and non zero, and their sum is zero. Hence  $C_{j,k}^{D_\lambda}$  tends to  $a/d = 1/2$  as  $\lambda$  tends to infinity.

□

### 5.3. The Ricci curvature of the $\lambda$ -deformed Killing metric

Let us keep the notations of the last paragraph. We first recall some standard definitions.

**Definition 5.7.** The curvature tensor of the Riemannian structure  $\langle \cdot, \cdot \rangle_{D_\lambda}$  is defined as the tri-linear application  $R_{D_\lambda}$  from  $T_p \text{SU}(n)^3$  to  $T_p \text{SU}(n)$  such that  $R_\lambda(X, Y)Z = -\nabla_X^{D_\lambda}(\nabla_Y^{D_\lambda} Z) + \nabla_Y^{D_\lambda}(\nabla_X^{D_\lambda} Z) + \nabla_{[X, Y]}^{D_\lambda} Z$ .

**Definition 5.8.** The Ricci curvature  $\mathcal{R}_{D_\lambda}$  of the  $D_\lambda$ -deformed Killing metric  $\langle \cdot, \cdot \rangle_{D_\lambda}$  at a point  $p$  of  $\text{SU}(n)$  is defined as the bilinear application from  $T_p \text{SU}(n)^2$  to  $\mathbb{R}$  such that  $\mathcal{R}_{D_\lambda}(X, Y)$  is the trace of the linear application  $\bullet \mapsto R_{D_\lambda}(X, \bullet)Y$  from  $T_p \text{SU}(n)$  to itself.

In the following, we compute the Ricci curvature of the  $D_\lambda$ -deformed Killing metric.

**Proposition 5.9.** *The restriction to  $\mathfrak{t}^\perp$  of the Ricci curvature  $\mathcal{R}_{D_\lambda}$  of the  $D_\lambda$  deformed Killing metric is diagonal in the canonical basis  $\mathcal{B}^{\mathfrak{t}^\perp}$  and tends to a finite limit  $\mathcal{R}_D$  as  $\lambda$  tends to infinity.  $\mathfrak{t}$  and  $\mathfrak{t}^\perp$  are orthogonal for  $\mathcal{R}_D$ , and the restriction of  $\mathcal{R}_D$  to  $\mathfrak{t}$  is a positive definite bilinear form.*

*Proof.* For all  $x, y, z$  in  $\mathcal{B}_{\mathfrak{t}^\perp}$ , we can write:

$$R_{D_\lambda}(x, z)y = \underbrace{-\nabla_x^{D_\lambda}(\nabla_z^{D_\lambda} y)}_a + \underbrace{\nabla_z^{D_\lambda}(\nabla_x^{D_\lambda} y)}_b + \underbrace{\nabla_{[x, z]}^{D_\lambda} y}_c. \tag{5.11}$$

Hence,

$$\langle R_\lambda(x, z)y, z \rangle_{D_\lambda} = \langle a, z \rangle_{D_\lambda} + \langle b, z \rangle_{D_\lambda} + \langle c, z \rangle_{D_\lambda}, \tag{5.12}$$

with

$$\langle a, z \rangle_{D_\lambda} = -C_{z, y}^{D_\lambda} C_{x, [z, y]}^{D_\lambda} \langle [x, [z, y]], z \rangle_{D_\lambda}, \tag{5.13}$$

$$\langle b, z \rangle_{D_\lambda} = C_{x, y}^{D_\lambda} C_{z, [x, y]}^{D_\lambda} \langle [z, [x, y]], z \rangle_{D_\lambda}, \tag{5.14}$$

$$\langle c, z \rangle_{D_\lambda} = C_{[x, z], y}^{D_\lambda} \langle [[x, z], y], z \rangle_{D_\lambda}. \tag{5.15}$$

Using now Jacobi identity  $[Z, [X, Y]] = -[X, [Y, Z]] + [Y, [X, Z]]$ , we can rewrite equation (5.14):

$$\langle b, z \rangle_{D_\lambda} = C_{x, y}^{D_\lambda} C_{z, [x, y]}^{D_\lambda} (-\langle [x, [y, z]], z \rangle_{D_\lambda} + \langle [y, [x, z]], z \rangle_{D_\lambda}).$$

Proposition C.1 asserts that the matrix of the restriction of  $\mathcal{R}_{D_\lambda}$  to  $\mathfrak{t}^\perp$  is diagonal in basis  $\mathcal{B}_{\mathfrak{t}^\perp}$ , with eigenvalues

$$\mu_y^{D_\lambda} = \sum_{z \in \mathcal{B}} \frac{1}{\langle z, z \rangle_{D_\lambda}} \left( -C_{z, y}^{D_\lambda} C_{y, [z, y]}^{D_\lambda} \langle [y, [z, y]], z \rangle_{D_\lambda} + C_{[y, z], y}^{D_\lambda} \langle [[y, z], y], z \rangle_{D_\lambda} \right),$$

for any  $y$  in  $\mathcal{B}_{\mathfrak{t}^\perp}$ . As  $\lambda$  tends to infinity, all  $C_{j,k}^{D_\lambda}$  tend to a finite limit  $C_{j,k}^D$ , and the restriction to  $\mathfrak{t}^\perp$  of  $\langle \cdot, \cdot \rangle_\lambda$  is fixed. The only things we have to control are the values of the scalar products  $\langle a, z \rangle_{D_\lambda}$  and  $\langle c, z \rangle_{D_\lambda}$ , when  $z$  belongs to  $\mathfrak{t}$ . Actually, since  $a$  and  $c$  are homogeneous of degree 1 in  $z$  (whose norm for  $\langle \cdot, \cdot \rangle_{D_\lambda}$  takes the form  $1/\lambda$ ), these expressions are constant as  $\lambda$  varies. So, we may remove the  $\lambda$  index and the restriction of  $\mathcal{R}_{D_\lambda}$  to  $\mathfrak{t}^\perp$  tends to a finite bilinear symmetric form, whose matrix in basis  $\mathcal{B}_{\mathfrak{t}^\perp}$  is diagonal with eigenvalues

$$\mu_y^D = \sum_{z \in \mathcal{B}} \frac{1}{\langle z, z \rangle_D} (-\langle C_{z,y}^D C_{y,[y,z]}^D \langle [y, [z, y]], z \rangle_D + C_{[y,z],y}^D \langle [[y, z], y], z \rangle_D),$$

for any  $y$  in  $\mathcal{B}_{\mathfrak{t}^\perp}$ .

The restriction to  $\mathfrak{t}$  of the Ricci curvature  $\mathcal{R}_{D_\lambda}$  still takes the form

$$\mathcal{R}_{D_\lambda}(y, y) = \sum_{z \in \mathcal{B}} \frac{1}{\langle z, z \rangle_{D_\lambda}} (-C_{z,y}^{D_\lambda} C_{y,[y,z]}^{D_\lambda} \langle [y, [z, y]], z \rangle_{D_\lambda} + C_{[y,z],y}^{D_\lambda} \langle [[y, z], y], z \rangle_{D_\lambda}), \tag{5.16}$$

for any  $x$  in  $\mathcal{B}_{\mathfrak{t}}$ . One remarks that if  $z$  belongs to  $\mathfrak{t}$ , then

$$-\langle C_{z,y}^{D_\lambda} C_{y,[y,z]}^{D_\lambda} \langle [y, [z, y]], z \rangle_{D_\lambda} + C_{[y,z],y}^{D_\lambda} \langle [[y, z], y], z \rangle_{D_\lambda}, \tag{5.17}$$

so it is enough to consider only the case where  $z$  belongs to  $\mathcal{B}_{\mathfrak{t}^\perp}$ . Now, for any  $z$  in  $\mathcal{B}_{\mathfrak{t}^\perp}$ ,  $C_{z,y}^{D_\lambda}$  tends to zero as  $\lambda$  tends to infinity, and  $C_{[y,z],z}^{D_\lambda}$  tends to 0 if  $[y, z] = 0$  and  $1/2$  otherwise. Hence, in any case,

$$\frac{1}{\langle z, z \rangle_{D_\lambda}} (-\langle C_{z,y}^{D_\lambda} C_{y,[y,z]}^{D_\lambda} \langle [y, [z, y]], z \rangle_{D_\lambda} + C_{[y,z],y}^{D_\lambda} \langle [[y, z], y], z \rangle_{D_\lambda}) \tag{5.18}$$

tends to  $1/2[y] (\frac{z}{i})^2$ . Also  $\mathcal{R}_{D_\lambda}(y, y)$  tends to  $\frac{1}{2} \sum_{z \in \mathcal{B}_{\mathfrak{t}^\perp}} [y] (\frac{z}{i})^2$ . Hence  $\mathcal{R}_D(y, y) > 0$  for any non zero  $y$  in  $\mathfrak{t}$ .

This complete the proof of Proposition 5.9. □

#### 5.4. Upper bound for the length of the geodesics of $\langle \cdot, \cdot \rangle_\lambda$ with initial velocity in $\mathfrak{t}^\perp$

For any non zero  $x$  in  $\mathfrak{u}$ , we denote by  $x_\lambda$  the unique unitary vector for the  $D_\lambda$ -deformed Killing metric having the same direction and versus as  $x$ .

$$x_\lambda = \frac{1}{\sqrt{\langle x, x \rangle_{D_\lambda}}} x. \tag{5.19}$$

If  $x$  belongs to  $\mathfrak{t}^\perp$ ,  $x_\lambda$  is constant when  $\lambda$  varies. If  $x$  belongs to  $\mathfrak{t}$ , then  $x_\lambda = \lambda x_1$  tends to infinity for the (opposite of the) standard Killing form (or for any fixed metric on  $\mathfrak{u}$ ).

Geodesics  $\gamma_x$  of the  $D_\lambda$  deformed Killing metric starting form  $Id_{\text{SU}(n)}$  whose tangent at  $Id$  is  $x_\lambda$  have the equation:

$$\begin{aligned} \gamma_x : \mathbb{R} &\rightarrow \text{SU}(n) \\ t &\mapsto \exp(tx_\lambda). \end{aligned} \tag{5.20}$$

Such a geodesic minimizing the length between  $Id$  and the left translate of the maximal torus  $yT$  (for a given point  $y$  of  $U$ ) has to respect the (Riemannian) transversality condition:

$$x_\lambda \perp \mathfrak{t} \Leftrightarrow x_\lambda \in \mathfrak{t}^\perp.$$

The idea is now to use the classical Myers theorem. Fix any  $\lambda$  strictly positive big enough. Then if the minimum  $k_{D_\lambda}$  of the eigenvalues of  $\mathcal{R}_{D_\lambda}$  on  $\mathfrak{t}^\perp$  is strictly positive, then any geodesic for  $\langle \cdot, \cdot \rangle_{D_\lambda}$  tangent to  $\mathfrak{t}^\perp$  at identity has a length less than  $\pi \frac{\sqrt{N_{\mathfrak{u}}-1}}{k_{D_\lambda}}$ .

Assume now that the restriction of the limit Ricci curvature  $\mathcal{R}$  to  $\mathfrak{t}^\perp$  is positive definite, with least eigenvalue  $r$ . Then, the classical Myers theorem asserts that no length minimizing geodesic has a length longer than  $\pi \frac{\sqrt{N_u-1}}{r}$ . This completes the proof of Theorem 2.13.  $\square$

An example of application in  $SU(3)$  is presented in Section 6.4.

## 6. EXAMPLES OF APPLICATION

In this section, we show how to apply the techniques that we have developed in this paper.

### 6.1. Notations and statement of the problem

We define the following basis of  $SU(3)$  :

$$\begin{aligned} \partial_{\alpha_1} &= \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \partial_{\alpha_2} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} & \partial_{\alpha_3} &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \\ \partial_{\bar{\alpha}_1} &= \begin{pmatrix} 0 & i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \partial_{\bar{\alpha}_2} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & i \\ 0 & i & 0 \end{pmatrix} & \partial_{\bar{\alpha}_3} &= \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \\ h_{\alpha_1} &= 2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} & h_{\alpha_2} &= 2 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \end{aligned}$$

Recall that, on  $\mathfrak{su}(3)$ , The Killing form is given by  $Kill(x, y) = -\langle x, y \rangle = 6 \operatorname{tr}(xy)$  for all matrices  $x, y$  in  $\mathfrak{su}(3)$ . The family  $(\partial_{\alpha_1}, \partial_{\alpha_2}, \partial_{\alpha_3}, \partial_{\bar{\alpha}_1}, \partial_{\bar{\alpha}_2}, \partial_{\bar{\alpha}_3})$  is orthogonal. All these vectors have the same Killing norm:  $\langle \partial_{\alpha_1}, \partial_{\alpha_1} \rangle = \langle \partial_{\alpha_2}, \partial_{\alpha_2} \rangle = \langle \partial_{\alpha_3}, \partial_{\alpha_3} \rangle = \langle \partial_{\bar{\alpha}_1}, \partial_{\bar{\alpha}_1} \rangle = \langle \partial_{\bar{\alpha}_2}, \partial_{\bar{\alpha}_2} \rangle = \langle \partial_{\bar{\alpha}_3}, \partial_{\bar{\alpha}_3} \rangle = 12$ . The  $\mathfrak{t}$ -subspace spanned by vectors  $h_{\alpha_1}$  and  $h_{\alpha_2}$  is a Cartan subalgebra of the Lie algebra  $\mathfrak{su}(3)$ .

For all non zero complex numbers  $(a_1, a_2, a_3)$  and non zero purely imaginary numbers  $(d_1, d_2, d_3)$  such that  $d_1 + d_2 + d_3 = 0$  and such that  $d_1 - d_2$  and  $d_2 - d_3$  are rationally independent, we denote with  $S^{(a_1, a_2, a_3)}$  the control problem in  $SU(3)$  whose dynamic is given by:

$$(S^{(a_1, a_2, a_3)}) \begin{cases} \dot{g}(t) &= g(t) \left( \begin{pmatrix} 0 & a_1 & a_3 \\ -\bar{a}_1 & 0 & a_2 \\ -\bar{a}_3 & -\bar{a}_2 & 0 \end{pmatrix} + u \begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{pmatrix} \right) \\ g(0) &= Id, \end{cases}$$

where the admissible controls are the measurable bounded functions from  $\mathbb{R}^+$  to  $\mathbb{R}$ . We may rewrite the system  $S^{(a_1, a_2, a_3)}$  as:

$$(S^{(a_1, a_2, a_3)}) \begin{cases} \dot{g}(t) &= g(h_0 + u h_1) \\ g(0) &= Id, \end{cases}$$

with

$$h_0 = \sum_{j=1}^3 x_j \partial_{\alpha_j} + \sum_{j=1}^3 y_j \partial_{\bar{\alpha}_j},$$

where  $x_k$  and  $y_k$  are respectively the real and imaginary parts of  $a_k$ , and

$$h_1 = -\frac{i}{2} (d_1 h_{\alpha_1} - d_3 h_{\alpha_2}).$$

The problem is to give an estimation of  $\mathcal{D}(Id, S^{(a_1, a_2, a_3)})$ .

Using the notations of Theorem 2.6,  $\mathcal{Y} = \text{conv}(\mathfrak{t} + \{\text{Ad}_{\exp(\mathfrak{t})}(h_0)\})$  is given by

$$\mathcal{Y} = \text{conv} \left( \mathfrak{t} + \left\{ \begin{pmatrix} 0 & a_1 e^{i\theta_1} & a_3 e^{i(\theta_1+\theta_2)} \\ -\overline{a_1} e^{-i\theta_1} & 0 & a_2 e^{i\theta_2} \\ -\overline{a_3} e^{-i(\theta_1+\theta_2)} & -\overline{a_2} e^{-i\theta_2} & 0 \end{pmatrix}, (\theta_1, \theta_2) \in \mathbb{R}^2 \right\} \right),$$

and  $\mathcal{X} = \pi_{\mathfrak{t}^\perp}(\mathcal{Y})$  is given by

$$\mathcal{X} = \text{conv} \left\{ \begin{pmatrix} 0 & a_1 e^{i\theta_1} & a_3 e^{i(\theta_1+\theta_2)} \\ -\overline{a_1} e^{-i\theta_1} & 0 & a_2 e^{i\theta_2} \\ -\overline{a_3} e^{-i(\theta_1+\theta_2)} & -\overline{a_2} e^{-i\theta_2} & 0 \end{pmatrix}, (\theta_1, \theta_2) \in \mathbb{R}^2 \right\}.$$

For instance, the matrices

$$\begin{pmatrix} 0 & a_1 & a_3 \\ -\overline{a_1} & 0 & a_2 \\ -\overline{a_3} & -\overline{a_2} & 0 \end{pmatrix}, \begin{pmatrix} 0 & a_1 & -a_3 \\ -\overline{a_1} & 0 & -a_2 \\ \overline{a_3} & \overline{a_2} & 0 \end{pmatrix}, \begin{pmatrix} 0 & -a_1 & -a_3 \\ \overline{a_1} & 0 & a_2 \\ \overline{a_3} & -\overline{a_2} & 0 \end{pmatrix},$$

belong to  $\mathcal{X}$ . But, if  $a_1 \cdot a_2 \cdot a_3 \neq 0$ , the matrices

$$\begin{pmatrix} 0 & -a_1 & a_3 \\ \overline{a_1} & 0 & a_2 \\ -\overline{a_3} & -\overline{a_2} & 0 \end{pmatrix}, \begin{pmatrix} 0 & a_1 & a_3 \\ -\overline{a_1} & 0 & -a_2 \\ -\overline{a_3} & \overline{a_2} & 0 \end{pmatrix}, \begin{pmatrix} 0 & -a_1 & -a_3 \\ \overline{a_1} & 0 & -a_2 \\ \overline{a_3} & \overline{a_2} & 0 \end{pmatrix},$$

do not belong to  $\mathcal{X}$ .

### 6.2. Approximation by the standard Killing form on the group

$\mathcal{Y}$  contains the ball (for the standard Killing metric) with center  $0_{\mathfrak{u}}$  and radius  $\frac{2\sqrt{3} \min_k(|a_k|)}{\sqrt{6}} = \sqrt{2} \min_k(|a_k|)$ . Proposition 2.9 gives:

$$\mathcal{D}(\text{SU}(3), S^{(a_1, a_2, a_3)}) \leq \frac{4\pi}{\min_k(|a_k|)}.$$

In the particular case where  $a_1 = a_2 = 1$  and  $a_3 = 1/2$ , we obtain

$$\mathcal{D}(\text{SU}(3), S^{(1, 1, 1/2)}) \leq 8\pi \approx 25, 133.$$

### 6.3. Approximation by the projection of the standard Killing form on the quotient $\text{SU}(3)/T$

The Killing norm of the drift term is

$$\sqrt{\left\langle \left\langle \begin{pmatrix} 0 & a_1 & a_3 \\ -\overline{a_1} & 0 & a_2 \\ -\overline{a_3} & -\overline{a_2} & 0 \end{pmatrix}, \begin{pmatrix} 0 & a_1 & a_3 \\ -\overline{a_1} & 0 & a_2 \\ -\overline{a_3} & -\overline{a_2} & 0 \end{pmatrix} \right\rangle \right\rangle} = 2\sqrt{3(|a_1|^2 + |a_2|^2 + |a_3|^2)}.$$

Theorem 2.11 asserts that  $\mathcal{D}(\text{SU}(3)/T, d_T) = \frac{\pi}{3} \sqrt{6.4.2} = \frac{4\sqrt{3}\pi}{3}$ . From Theorem 2.10, we get

$$\frac{2\pi}{3(|a_1|^2 + |a_2|^2 + |a_3|^2)} \leq \mathcal{D}(\text{SU}(3), S^{(a_1, a_2, a_3)}) \leq \sqrt{6} \frac{2\pi}{3 \min |a_k|}. \tag{6.1}$$

In particular, if  $a_1 = a_2 = 1$  and  $a_3 = 1/2$ , then

$$0, 931 \approx \frac{8}{27}\pi \leq \mathcal{D}(\text{SU}(3), S^{(1,1,1/2)}) \leq \frac{4}{3}\sqrt{6}\pi \approx 10, 260.$$

### 6.4. Approximation by a deformed Killing form

Let us consider the  $D_\lambda$ -deformed Killing form with  $D_\lambda = (\lambda, l_1, l_2, l_3)$ . The orthogonal projection on  $\mathfrak{t}$  of the unit sphere for the  $D_\lambda$ -deformed Killing metric is the ellipsoid (for the standard Killing form):

$$l_1^2(x_1^2 + y_1^2) + l_2^2(x_2^2 + y_2^2) + l_3^2(x_3^2 + y_3^2) = 1. \tag{6.2}$$

Proposition 5.4 gives coefficients  $C_{j,k}^D$  in the basis  $(\partial_{\alpha_1}, \partial_{\alpha_2}, \partial_{\alpha_3}, \partial_{\bar{\alpha}_1}, \partial_{\bar{\alpha}_2}, \partial_{\bar{\alpha}_3}, h_{\alpha_1}, h_{\alpha_2})$ , from which we deduce an expression for the Ricci curvature in  $\mathcal{B}_{\mathfrak{t}^\perp}$ :

$$\begin{aligned} \mathcal{R}_D(\partial_{\alpha_1}, \partial_{\alpha_1}) &= \mathcal{R}_D(\partial_{\bar{\alpha}_1}, \partial_{\bar{\alpha}_1}) = \frac{-l_1^2 l_3^4 + l_1^6 - l_1^2 l_2^4 + 6l_3^2 l_2^2 l_1^2 + 2l_0^2 l_3^2 l_2^2}{l_1^2 l_2^2 l_3^2}, \\ \mathcal{R}_D(\partial_{\alpha_2}, \partial_{\alpha_2}) &= \mathcal{R}_D(\partial_{\bar{\alpha}_2}, \partial_{\bar{\alpha}_2}) = \frac{-l_2^2 l_3^4 + l_2^6 - l_2^2 l_1^4 + 6l_3^2 l_2^2 l_1^2 + 2l_0^2 l_3^2 l_1^2}{l_1^2 l_2^2 l_3^2}, \\ \mathcal{R}_D(\partial_{\alpha_3}, \partial_{\alpha_3}) &= \mathcal{R}_D(\partial_{\bar{\alpha}_3}, \partial_{\bar{\alpha}_3}) = \frac{l_3^6 - l_3^2 l_2^4 - l_3^2 l_1^4 + 6l_3^2 l_2^2 l_1^2 + 2l_0^2 l_2^2 l_1^2}{l_1^2 l_2^2 l_3^2}. \end{aligned}$$

**Remark 6.1.** For certain choices of  $(l_1, l_2, l_3)$ , some curvatures may be negative. In this case, Theorem 2.13 is no more valid.

In the special case where  $a_1 = a_2 = 1$ , we choose  $l_1 = l_2 = \sqrt{2}$  and  $l_3 = \frac{\sqrt{2}}{|a_3|}$ .

$$\begin{aligned} \frac{1}{\langle \partial_{\alpha_1}, \partial_{\alpha_1} \rangle_{D_\lambda}} \mathcal{R}(\partial_{\alpha_1}, \partial_{\alpha_1}) &= \frac{1}{\langle \partial_{\bar{\alpha}_1}, \partial_{\bar{\alpha}_1} \rangle_{D_\lambda}} \mathcal{R}(\partial_{\bar{\alpha}_1}, \partial_{\bar{\alpha}_1}) = 3 - \frac{1}{2|a_3|^2}, \\ \frac{1}{\langle \partial_{\alpha_2}, \partial_{\alpha_2} \rangle_{D_\lambda}} \mathcal{R}(\partial_{\alpha_2}, \partial_{\alpha_2}) &= \frac{1}{\langle \partial_{\bar{\alpha}_2}, \partial_{\bar{\alpha}_2} \rangle_{D_\lambda}} \mathcal{R}(\partial_{\bar{\alpha}_2}, \partial_{\bar{\alpha}_2}) = 3 - \frac{1}{2|a_3|^2}, \\ \frac{1}{\langle \partial_{\alpha_3}, \partial_{\alpha_3} \rangle_{D_\lambda}} \mathcal{R}(\partial_{\alpha_3}, \partial_{\alpha_3}) &= \frac{1}{\langle \partial_{\bar{\alpha}_3}, \partial_{\bar{\alpha}_3} \rangle_{D_\lambda}} \mathcal{R}(\partial_{\bar{\alpha}_3}, \partial_{\bar{\alpha}_3}) = 2|a_3|^2 + \frac{1}{2|a_3|^2}. \end{aligned}$$

From Theorem 2.13, we get

$$\mathcal{D}(\text{SU}(3), S^{(1,1,a_3)}) \leq \pi r, \tag{6.3}$$

with

$$r = \sqrt{\frac{7}{\min \left\{ 3 - \frac{1}{2|a_3|^2}, 2|a_3|^2 + \frac{1}{2|a_3|^2} \right\}}}.$$

With  $a_1 = a_2 = 1$  and  $a_3 = 1/2$ ,  $r = \sqrt{7}$ , equation (6.3) reads

$$\mathcal{D}(\text{SU}(3), S^{(1,1,1/2)}) \leq \sqrt{7}\pi \approx 8, 312.$$

## APPENDIX A. SOME COMPUTATION RULES IN REAL SEMI-SIMPLE COMPACT LIE ALGEBRAS

In this appendix, we intend to recall some identities used in the paper. For proofs and more advanced results, see [24], Chap. 3. We keep the notations of the paper. Let  $\mathfrak{u}$  be a compact real semi-simple Lie algebra and define  $\mathfrak{g} = \mathfrak{u} \otimes_{\mathbb{R}} \mathbb{C}$  its complexification. The Lie algebra  $\mathfrak{u}$  is then identified with its image in  $\mathfrak{g}$  under the map  $x \mapsto x \otimes 1$  (see [24, 39]). An element  $h$  of  $\mathfrak{g}$  or  $\mathfrak{u}$  is said to be regular if the mapping  $ad_h$  has maximal rank. The regular elements of  $\mathfrak{u}$  are exactly the regular elements of  $\mathfrak{g}$  that are elements of  $\mathfrak{u}$ . Choose  $h_1$  a regular element of  $\mathfrak{g}$ . Then  $\ker(ad_{h_1}) = \mathfrak{h}$  is a nilpotent subalgebra of  $\mathfrak{g}$  that is its own normalizer in  $\mathfrak{g}$ . Such an algebra  $\mathfrak{h}$  is called a Cartan subalgebra. The mappings  $ad_h$  acts on  $\mathfrak{g}$ . There exists a basis  $(H_\gamma, X_\alpha)$  of  $\mathfrak{g}$  such that  $(H_\gamma)$  is a basis of  $\mathfrak{h}$  and all these mappings  $ad_h$  are diagonal. For any  $\alpha$ , we define an element of  $\mathfrak{h}^*$  that to any  $H$  in  $\mathfrak{h}$  associates the number  $\lambda$  such that  $[H, X_\alpha] = \lambda X_\alpha$ . We still denote this application, called a root, with  $\alpha$ . We

choose an arbitrary order  $<$  on the space  $\mathfrak{t}^*$ , from which we deduce an order on the roots  $\{\alpha\}$ , and we denote with  $\Delta$  the set of the positive roots. It is possible to find a basis  $\{H_\alpha, X_\alpha\}$  of  $\mathfrak{g}$  such that

$$[X_\alpha, X_{-\alpha}] = H_\alpha \quad (\text{A.1})$$

$$[H, X_\alpha] = \alpha(H)X_\alpha \text{ for } H \in \mathfrak{t} \quad (\text{A.2})$$

$$[X_\alpha, X_\beta] = 0 \text{ if } \alpha + \beta \neq 0 \text{ and } \alpha + \beta \notin \Delta \quad (\text{A.3})$$

$$[X_\alpha, X_\beta] = N_{\alpha,\beta}X_{\alpha+\beta} \text{ if } \alpha + \beta \in \Delta, \quad (\text{A.4})$$

where the coefficients  $N_{\alpha,\beta}$  are some constants. Up to an inner automorphism, the Lie algebra  $\mathfrak{u}$  writes

$$\mathfrak{u} = \sum_{\alpha \in \Delta} \mathbb{R}(iH_\alpha) + \sum_{\alpha \in \Delta} \mathbb{R}(X_\alpha - X_{-\alpha}) + \sum_{\alpha \in \Delta} \mathbb{R}(i(X_\alpha + X_{-\alpha})). \quad (\text{A.5})$$

Defining  $h_\alpha = iH_\alpha$ ,  $\partial_\alpha = X_\alpha - X_{-\alpha}$  and  $\partial_{\bar{\alpha}} = i(X_\alpha + X_{-\alpha})$ , or equivalently  $X_\alpha = \frac{\partial_\alpha - i\partial_{\bar{\alpha}}}{2}$  and  $X_{-\alpha} = \frac{-\partial_\alpha - i\partial_{\bar{\alpha}}}{2}$ , we get

$$\mathfrak{u} = \sum_{\alpha \in \Delta} \mathbb{R}(h_\alpha) + \sum_{\alpha \in \Delta} \mathbb{R} \partial_\alpha + \sum_{\alpha \in \Delta} \mathbb{R} \partial_{\bar{\alpha}}. \quad (\text{A.6})$$

The commutation rules for these vectors  $h_\alpha$ ,  $\partial_\alpha$  and  $\partial_{\bar{\alpha}}$  can be easily deduced from equations (A.1), (A.2), (A.3), (A.4)

$$[\partial_\alpha, \partial_\beta] = 0 \text{ if } \alpha + \beta \notin \Delta \text{ and } \alpha - \beta \notin \Delta, \quad (\text{A.7})$$

$$[\partial_\alpha, \partial_\beta] = N_{\alpha,\beta}\partial_{\alpha+\beta} \text{ if } \alpha + \beta \in \Delta \text{ and } \alpha - \beta \notin \Delta, \quad (\text{A.8})$$

$$[\partial_\alpha, \partial_\beta] = N_{\alpha,-\beta}\partial_{\alpha-\beta} \text{ if } \alpha + \beta \notin \Delta \text{ and } \alpha - \beta \in \Delta \text{ and } \alpha - \beta \geq 0, \quad (\text{A.9})$$

$$[\partial_\alpha, \partial_\beta] = N_{\alpha,\beta}\partial_{\alpha+\beta} + N_{\alpha,-\beta}\partial_{\alpha-\beta} \text{ if } \alpha + \beta \in \Delta \text{ and } \alpha - \beta \in \Delta \text{ and } \alpha - \beta \geq 0, \quad (\text{A.10})$$

$$[h_\beta, \partial_\alpha] = \beta(H_\beta)\partial_{\bar{\alpha}}. \quad (\text{A.11})$$

We define  $\mathfrak{t} = \text{span}_{\mathbb{R}}(h_\alpha)$  the real linear space spanned by the vectors  $h_\alpha$ . For any root  $\alpha$  and all  $v$  in  $\mathfrak{t}$ ,  $\alpha(v)$  is a purely imaginary number. Using these notations, for any  $\alpha$  in  $\Delta$  and any  $v$  in  $\mathfrak{t}$

$$(e^{\text{ad } v})\partial_\alpha = (e^{\text{ad } v})(X_\alpha - X_{-\alpha}) = (e^{\text{ad } v})X_\alpha - (e^{\text{ad } v})X_{-\alpha}. \quad (\text{A.12})$$

Recalling that  $\text{ad}(v)(X_\beta) = \beta(v)X_\beta$  for every root  $\beta$  in  $\Delta$ , and defining the real  $\theta$  by  $i\theta = \alpha(v)$ ,

$$(e^{\text{ad } v})\partial_\alpha = e^{\alpha(v)}X_\alpha - e^{-\alpha(v)}X_{-\alpha} = e^{i\theta} \left( \frac{\partial_\alpha - i\partial_{\bar{\alpha}}}{2} \right) + e^{-i\theta} \left( \frac{\partial_\alpha + i\partial_{\bar{\alpha}}}{2} \right) = \cos(\theta)\partial_\alpha + \sin(\theta)\partial_{\bar{\alpha}}. \quad (\text{A.13})$$

Hence for any  $\alpha$  in  $\Delta$  and any  $v$  in  $\mathfrak{t}$

$$(e^{\text{ad } v})\partial_\alpha = \cos(i\alpha(v))\partial_\alpha - \sin(i\alpha(v))\partial_{\bar{\alpha}}, \quad (\text{A.14})$$

$$(e^{\text{ad } v})\partial_{\bar{\alpha}} = \sin(i\alpha(v))\partial_\alpha + \cos(i\alpha(v))\partial_{\bar{\alpha}}, \quad (\text{A.15})$$

and obviously

$$(e^{\text{ad } v})h_\alpha = h_\alpha. \quad (\text{A.16})$$

## APPENDIX B. SOME COMPUTATION RULES IN $\mathfrak{su}(n)$

This appendix recalls some classical computational rules in  $\mathfrak{su}(n)$ . While Appendix A only uses very general semi-simple Lie algebra properties, and mainly the existence of roots, this appendix essentially uses the very

particular structure of  $\mathfrak{su}(n)$ , and mostly the very special expression of the highest root. All the results presented here are proved in [24], and we just intend to rewrite these results with our notations.

Calling  $\mathfrak{t}$  a Cartan subalgebra of  $\mathfrak{u} = \mathfrak{su}(n)$ , choosing an arbitrary order on  $\mathfrak{t}^*$  and denoting by  $\alpha_1, \dots, \alpha_{n-1}$ , with an appropriate indexation, the associated simple roots, every positive root writes:

$$\alpha = \sum_{i=p}^q \alpha_i, \tag{B.1}$$

with  $p$  and  $q$  two integers such that  $1 \leq p < q \leq n - 1$ .

The following properties are now obvious:

**Proposition B.1.** *Let  $\alpha$  and  $\beta$  be two roots. Then  $\alpha + \beta$  and  $\alpha - \beta$  cannot be simultaneously roots.*

**Proposition B.2.** *For any roots  $\alpha$  and  $\beta$  in  $\Delta$ , such that  $\alpha \geq \beta$ ,  $[\partial_\alpha, \partial_\beta]$  is either equal to*

- 0 if neither  $\alpha + \beta$  nor  $\alpha - \beta$  lies in  $\Delta$ ,
- $N_{\alpha,\beta} \partial_{\alpha+\beta}$  if  $\alpha + \beta \in \Delta$ ,
- $-N_{\alpha,-\beta} \partial_{\alpha-\beta}$  if  $\alpha - \beta \in \Delta$ .

The following result is the key point of the curvature computations of the fifth section.

**Proposition B.3.** *For any  $D_\lambda$  deformation of the Killing form, and for any elements  $j, k, l$  of the basis  $\mathcal{B}$ , a sufficient condition for the real number  $\langle j, [l, k] \rangle_{D_\lambda} + \langle [j, k], l \rangle_{D_\lambda} + \langle k, [l, j] \rangle_{D_\lambda}$  to be zero is that  $l$  is not colinear to  $[j, k]$ .*

*Proof.* In the following, we will denote by  $a, b$  and  $c$  respectively the three real numbers  $\langle k, [l, j] \rangle_{D_\lambda}$ ,  $\langle [j, k], l \rangle_{D_\lambda}$  and  $\langle j, [l, k] \rangle_{D_\lambda}$  respectively. The inequality  $a + b + c \neq 0$  implies that at least one of the three real numbers  $a, b$  and  $c$  is non zero. It is clear it is enough to prove that  $a + c = 0$  as soon as  $l$  is not colinear to  $[j, k]$ . We just have to check the result for the different possible configurations in  $j$  and  $k$ .

- **$j$  and  $k$  in  $\mathfrak{t}$**  Then  $a = b = c = 0$  for any choice of  $l$ .
- **$j$  and  $k$  in  $\mathfrak{t}^\perp$  and  $[j] \neq [k]$**  Then  $k$  colinear to  $[l, j]$  implies that  $[l, k]$  is colinear to  $j$  and  $l$  is colinear to  $[k, j]$ . Hence  $a$  and  $b$  are both non zero if  $l$  is colinear to  $[j, k]$  and both zero else.
- **$j$  and  $k$  in  $\mathfrak{t}^\perp$  and  $j = \bar{k}$ .** Then  $a = b = c = 0$  if  $l \notin \mathfrak{t}$ . For any  $l$  in  $\mathfrak{t}$ ,  $a = \pm [j](l) \langle j, j \rangle_{D_\lambda}$  and  $c = -a$ . Hence, for any  $l$  in  $\mathcal{B}$ ,  $a + c = 0$ .
- **$j$  in  $\mathfrak{t}$  and  $k$  in  $\mathfrak{t}^\perp$**  Then  $a \neq 0$  or  $c \neq 0$  implies  $l = \bar{k}$ , which is colinear to  $[j, k]$ .
- **$j$  in  $\mathfrak{t}^\perp$  and  $k$  in  $\mathfrak{t}$**  Then  $a \neq 0$  or  $c \neq 0$  implies  $l = \bar{k}$ , which is colinear to  $[j, k]$ . □

### APPENDIX C. A CONDITION ON $j$ AND $k$ FOR $\text{AD}_j \circ \text{AD}_k$ TO BE DIAGONAL ON THE WEYL-CHEVALLEY BASIS $\mathcal{B}$ OF $\mathfrak{su}(n)$

The aim of this appendix is to prove the following result:

**Proposition C.1.** *For any  $j$  and any  $k$  in  $\{\mathcal{B} \times \mathcal{B}\} \setminus \{\mathcal{B}^\perp \times \mathcal{B}^\perp\}$ , for any  $l$  in  $\mathcal{B}$ , if  $\text{ad}_j \circ \text{ad}_k(l)$  is non zero and colinear to  $l$ , then  $j = k$ .*

*Proof.* It is enough to check the result for the various possible configuration of  $j, k$  and  $l$ .

- **if  $j \in \mathfrak{t}$  and  $k \in \mathfrak{t}^\perp$** 
  - if  $l \in \mathfrak{t}$ , then  $[k, l]$  is colinear to  $\bar{k}$ . Hence  $[j, [k, l]]$  is colinear to  $k$  and so zero or non-colinear to  $l$ ;
  - if  $l \in \mathfrak{t}^\perp$ , then  $[k, l] = 0$  or  $[[k, l]] \neq [l]$ . In this case,  $[\text{ad}_j([k, l])] = [[k, l]] \neq [l]$  and  $\text{ad}_j(\text{ad}_k(l))$  is not colinear to  $l$ ,
- **if  $j \notin \mathfrak{t}$  and  $k \in \mathfrak{t}$** 
  - if  $l \in \mathfrak{t}$ , then  $[k, l] = 0$ ;
  - if  $l \notin \mathfrak{t}$ , then  $\text{ad}_j \circ \text{ad}_k(l) = 0$  or  $[\text{ad}_j \circ \text{ad}_k(l)] \neq [l]$ ,

- if  $j \notin \mathfrak{t}$  and  $k \notin \mathfrak{t}$ 
  - if  $l \in \mathfrak{t}$ , then  $[k, l]$  is colinear to  $\bar{k}$ . Hence  $[j, [k, l]]$  is in  $\mathfrak{t}$  and non zero if and only if  $j = k$ ;
  - if  $l \notin \mathfrak{t}$ , the proof is more complicated. The idea of the proof is first to use the root modulus to show that  $j$  and  $k$  are not in  $\mathfrak{t}$  and that  $[j] = [k] \neq 0$ , and then to check that  $j \neq \bar{k}$ . First, the modulus of  $\text{ad}_j \circ \text{ad}_k(l)$  has one of the 7 following forms:  $[j] + [k] + [l]$ ,  $[k] + [l] - [j]$ ,  $[j] - [k] - [l]$ ,  $[k] - [l] + [j]$ ,  $[k] - [l] - [j]$ ,  $[j] - [k] + [l]$ ,  $[l] - [k] - [j]$ . Hence,  $[\text{ad}_j \circ \text{ad}_k(l)] = [l]$  implies that  $[j]$  and  $[k]$  satisfy (at least) one of the following 4 equations:

$$[j] + [k] = 0 \quad (\text{C.1})$$

$$[k] - [j] = 0 \quad (\text{C.2})$$

$$[j] - [k] = 2[l] \quad (\text{C.3})$$

$$[j] + [k] = 2[l]. \quad (\text{C.4})$$

The particular root structure of  $\mathfrak{su}(n)$  implies that the equations (C.1), (C.3) and (C.4) have no solution. Hence we have  $[j] = [k]$ , so  $j = k$  or  $j = \bar{k}$ . If  $j = \bar{k}$ , then  $\text{ad}_j \circ \text{ad}_k(l)$  is colinear to  $\bar{l} \neq l$ . This completes the proof of Proposition C.1. □

*Acknowledgements.* Second author is grateful to Ivan Kupka for many suggestions.

## REFERENCES

- [1] J.F. Adams, *Lectures on Lie groups*. W.A. Benjamin, Inc., New York-Amsterdam (1969).
- [2] A.A. Agrachev, Introduction to optimal control theory, in *Mathematical control theory, Part 1, 2 (Trieste, 2001)*, ICTP Lect. Notes, VIII, Abdus Salam Int. Cent. Theoret. Phys., Trieste (2002) 453–513 (electronic).
- [3] A.A. Agrachev and Y.L. Sachkov, *Control theory from the geometric viewpoint*, Encyclopaedia of Mathematical Sciences. **87** Springer-Verlag, Berlin (2004). Control Theory and Optimization, II.
- [4] A.O. Barut and R. Raczka, *Theory of group representations and applications*. World Scientific Publishing Co., Singapore, second edn. (1986).
- [5] B. Bonnard, V. Jurdjevic, I. Kupka and G. Sallet, Systèmes de champs de vecteurs transitifs sur les groupes de Lie semi-simples et leurs espaces homogènes, in *Systems analysis (Conf., Bordeaux, 1978)* **75 Astérisque, Soc. Math. France**, Paris (1980) 19–45.
- [6] B. Bonnard, V. Jurdjevic, I. Kupka and G. Sallet, Transitivity of families of invariant vector fields on the semidirect products of Lie groups. *Trans. Amer. Math. Soc.* **271** (1982) 525–535.
- [7] B. Bonnard, Couples de générateurs de certaines sous-algèbres de Lie de l’algèbre de Lie symplectique affine, et applications. *Publ. Dép. Math. (Lyon)* **15** (1978) 1–36.
- [8] B. Bonnard, Contrôlabilité de systèmes mécaniques sur les groupes de Lie. *SIAM J. Control Optim.* **22** (1984) 711–722.
- [9] U. Boscain, T. Chambrion and J.-P. Gauthier, On the K + P problem for a three-level quantum system: optimality implies resonance. *J. Dynam. Control Syst.* **8** (2002) 547–572.
- [10] U. Boscain, G. Charlot and J.-P. Gauthier, Optimal control of the Schrödinger equation with two or three levels, in *Nonlinear and adaptive control (Sheffield 2001)*, Springer, Berlin, *Lect. Not. Control Inform. Sci.* **281** (2003) 33–43.
- [11] U. Boscain, G. Charlot, J.-P. Gauthier, S. Guérin and H.-R. Jauslin, Optimal control in laser-induced population transfer for two- and three-level quantum systems. *J. Math. Phys.* **43** (2002) 2107–2132.
- [12] U. Boscain and G. Charlot, Resonance of minimizers for  $n$ -level quantum systems with an arbitrary cost. *ESAIM: COCV* **10** (2004) 593–614.
- [13] U. Boscain and Y. Chitour, On the minimum time problem for driftless left-invariant control systems on  $\text{SO}(3)$ . *Commun. Pure Appl. Anal.* **1** (2002) 285–312.
- [14] R. Brockett, New issues in the mathematics of control, in *Mathematics unlimited — 2001 and beyond*. Springer, Berlin (2001), pp. 189–219.
- [15] D. D’Alessandro and M. Dahleh, Optimal control of two-level quantum systems. *IEEE Trans. Automat. Control* **46** (2001) 866–876.
- [16] M.P. do Carmo, *Riemannian geometry, Mathematics: Theory & Applications*. Birkhäuser Boston Inc., Boston, MA (1992). Translated from the second Portuguese edition by Francis Flaherty.
- [17] R. El Assoudi and J.-P. Gauthier, Controllability of right invariant systems on real simple Lie groups of type  $F_4$ ,  $G_2$ ,  $C_n$ , and  $B_n$ . *Math. Control Signals Syst.* **1** (1988) 293–301.



- [18] R. El Assoudi and J.-P. Gauthier, Controllability of right-invariant systems on semi-simple Lie groups, in *New trends in nonlinear control theory* (Nantes, 1988). Springer, Berlin, *Lect. Notes Control Inform. Sci.* **122** (1989) 54–64.
- [19] R. El Assoudi, J.P. Gauthier and I.A.K. Kupka, Controllability of right invariant systems on semi-simple Lie groups, in *Geometry in nonlinear control and differential inclusions (Warsaw, 1993)*. Banach Center Publ., Polish Acad. Sci., Warsaw **32** (1995) 199–208.
- [20] R. El Assoudi, J.P. Gauthier and I.A.K. Kupka, On subsemigroups of semisimple Lie groups. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **13** (1996) 117–133.
- [21] R. El Assoudi and J.-P. Gauthier, Contrôlabilité sur l'espace quotient d'un groupe de Lie par un sous-groupe compact. *C. R. Acad. Sci. Paris Sér. I Math.* **311** (1990) 189–191.
- [22] A.L. Fradkov and A.N. Churilov, Eds. *Proceedings of the conference "Physics and Control" 2003* IEEE. August (2003).
- [23] J.-P. Gauthier, I. Kupka and G. Sallet, Controllability of right invariant systems on real simple Lie groups. *Syst. Contr. Lett.* **5** 187–190 (1984).
- [24] S. Helgason, Differential geometry, Lie groups, and symmetric spaces **80**, *Pure Appl. Math.*, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York (1978).
- [25] V. Jurdjevic, Optimal control problems on Lie groups: crossroads between geometry and mechanics, in *Geometry of feedback and optimal control*. Dekker, New York, *Monogr. Textbooks Pure Appl. Math.* **207** (1998) 257–303.
- [26] V. Jurdjevic, Optimal control, geometry, and mechanics, in *Mathematical control theory*. Springer, New York (1999) 227–267.
- [27] V. Jurdjevic and I. Kupka, Control systems on semisimple Lie groups and their homogeneous spaces. *Ann. Inst. Fourier (Grenoble)* **31** (1981) 151–179.
- [28] V. Jurdjevic and I. Kupka, Control systems subordinated to a group action: accessibility. *J. Differ. Equ.* **39** (1981) 186–211.
- [29] V. Jurdjevic, *Geometric control theory, Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge **52** (1997).
- [30] V. Jurdjevic, Lie determined systems and optimal problems with symmetries, in *Geometric control and non-holonomic mechanics* (Mexico City, 1996), Providence, RI. *CMS Conf. Proc., Amer. Math. Soc.* **25** (1998) 1–28.
- [31] A. Katok and B. Hasselblatt, *Introduction to the modern theory of dynamical systems*, Encyclopedia of Mathematics and its Applications. **54** Cambridge University Press, Cambridge (1995). With a supplementary chapter by Katok and Leonardo Mendoza.
- [32] N. Khaneja, S.J. Glaser and R. Brockett, Sub-Riemannian geometry and time optimal control of three spin systems: quantum gates and coherence transfer. *Phys. Rev. A* **65** (2002) 032301, 11.
- [33] I. Kupka, Applications of semigroups to geometric control theory, in *The analytical and topological theory of semigroups de Gruyter Exp. Math.* de Gruyter, Berlin **1** (1990) 337–345.
- [34] J. Milnor, *Morse theory*. Based on lecture notes by M. Spivak and R. Wells. Annals of Mathematics Studies, No. **51**. Princeton University Press, Princeton, N.J. (1963).
- [35] J. Milnor, Curvatures of left invariant metrics on Lie groups. *Advances Math.* **21** (1976) 293–329.
- [36] T. Püttmann, Injectivity radius and diameter of the manifolds of flags in the projective planes. *Math. Z.* **246** (2004) 795–809.
- [37] Y.L. Sachkov, Controllability of invariant systems on Lie groups and homogeneous spaces. *J. Math. Sci.* **100** (2000) 2355–2427  
Dynamical systems, 8.
- [38] H.J. Sussmann and V. Jurdjevic, Controllability of nonlinear systems. *J. Differ. Equ.* **12** (1972) 95–116.
- [39] V.S. Varadarajan, *Lie groups, Lie algebras, and their representations*. Prentice-Hall Inc., Englewood Cliffs, N.J. (1974).  
Prentice-Hall Series in Modern Analysis.