Time-optimal decompositions in $SU(2)$.
Yuly Billig[1]

Abstract. A connected Lie group G is generated by its two 1-parametric subgroups $\exp(tX)$, $\exp(tY)$ if and only if the Lie algebra of G is generated by $\{X, Y\}$. We consider decompositions of elements of G into a product of such exponentials with times $t > 0$ and study the problem of minimizing the total time of the decompositions for a fixed element of G. We solve this problem for the group SU_2 and describe the structure of the time-optimal decompositions.

0. Introduction.

In the axiomatics of quantum mechanics a state of a quantum system is a unit vector in a complex hermitian vector space and its evolution is given by unitary transformations. In quantum computing the space of states is finite-dimensional and the quantum computation is an element of SU_N, which is a compact connected real Lie group. In order to carry out a quantum computation in physical reality, we need to be able to transform an initial state ψ of a quantum processor into a state $g\psi$, where g is the element of SU_N representing the quantum computation. In the classical (non-quantum) case, a computation can be viewed as a boolean function in several variables. A processor can not implement an arbitrary function directly, but instead decomposes the computation into elementary steps, which is reflected in the process of programming. An efficient algorithm for a classical computation is a program that minimizes the number of elementary steps, which results in a shorter run time.

Likewise, an implementation of a quantum computation g is a factorization of g into certain elementary factors, called the quantum gates (in analogy with the classical logical gates used for the decomposition of an arbitrary boolean function of several variables in a disjunctive normal form, for example). Some quantum gates are the direct analogues of the classical logical gates and are discrete, while others depend on a continuous time parameter.

Time evolution of a quantum system is governed by a Hamiltonian H and is given by the exponential $\exp(itH) \in SU_N$. Here $X = iH$ belongs to the Lie algebra su_N. We are going to assume that all available quantum gates are continuous, since the discrete gates are physically realized as continuous gates applied for a specific finite time.

This leads to the following quantum control problem: given a set $S \subset su_N$ of quantum controls (gates), decompose an element $g \in SU_N$ into a product

$$g = \exp(t_1 C_1) \times \ldots \times \exp(t_n C_n), \tag{0.1}$$

where $C_i \in S$, $t_i \in \mathbb{R}$. This quantum control problem admits a solution for every g precisely when the set S generates the Lie algebra su_N (see Theorem 1.1 below). This

[1] School of Mathematics and Statistics, Carleton University, 1125 Colonel By Drive, Ottawa, K1S 5B6, Canada. E-mail: billig@math.carleton.ca

result is proved using topological methods and does not provide an effective procedure for finding such decompositions.

Minimal sets of quantum gates for quantum computing were discussed in [2] and [6], where it is shown that a set of two generic Hamiltonians is sufficient for controllability. A review of quantum control methods in physical chemistry is given in [7].

The problem of finding explicit factorizations of type (0.1) goes back to Euler [3], who studied it for the group SO_3 of rotations in $\mathbb{R}^3$. The Lie algebra so_3 consists of skew-symmetric 3×3 matrices and the exponentials $\exp(tC_i)$ corresponding to

$$
C_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad C_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}
$$

are the rotations in angle t around X- and Y-axis respectively. Euler proved that every $g \in SO_3$ can be factored as $g = \exp(t_1 C_1) \exp(t_2 C_2) \exp(t_3 C_1)$ for some $t_1, t_2, t_3 \in [0, 2\pi)$. The parameters t_1, t_2, t_3 are called Euler angles. This problem has applications in robotics when an object needs to be oriented in space in a prescribed way.

For the quantum applications the parameters t_i in (0.1) represent time, and thus must be positive. It is natural to pose an optimization problem in this context: among all factorizations (0.1) with positive times t_i, find the one that minimizes the total time $t_1 + \ldots + t_n$. This corresponds to finding an implementation of a quantum algorithm with the shortest running time.

As we shall see in this paper, this optimization problem, as posed, might not have a solution – there may be a sequence of factorizations of type (0.1) with the number of factors n going to infinity, while the total time going to infinum. For this reason we slightly modify the above optimization problem and ask for the infinum of total times for all factorizations of a given $g \in SU_N$

In this paper we solve this problem for the group SU_2 with the set of controls S consisting of two elements, $S = \{X, Y\}$.

We show that the infinum time does not change if we replace $S = \{X, Y\}$ with its convex closure $\overline{S} = \{\tau X + (1 - \tau)Y \mid 0 \leq \tau \leq 1\}$. In the latter case there will be in fact an optimal decomposition (0.1) with a finite number of factors. It turns out that to get an optimal decomposition it is sufficient to add to S at most one element $W \in \overline{S}$, the one that is orthogonal to $X - Y$. In this paper we give explicit descriptions of the optimal decompositions in SU_2. Since there is a surjective homomorphism $SU_2 \to SO_3$, our results are also applicable to the group SO_3.

Our solution is based on the method of Lagrange multipliers adopted for the set-up of Lie groups. Alternatively one could use the geometric theory developed in [4], which is based on the Pontryagin maximum principle. The general methods, however, give only necessary conditions for optimality, which in practice are not sufficient. To get the desired results we supplement these general methods with some explicit calculations in SU_2 which allow us to get stronger and more explicit optimality conditions.

We hope that our methods and results will help to solve this problem in greater generality.

1. Time-optimal decompositions in Lie groups.

Let G be a compact connected real Lie group, and let $\mathfrak{g}$ be its Lie algebra. An element $X \in \mathfrak{g}$ defines a 1-parametric subgroup $\{\exp(tX)|t \in \mathbb{R}\}$. In this paper we study an optimal control problem on G, describing optimal decompositions of an arbitrary given element $g \in G$ into a product of exponentials $\exp(tX)$ with X belonging to a fixed set S of controls, $S \subset \mathfrak{g}$ and positive times t.

The following well-known criterion describes when the group G is controllable by a set S:

Theorem 1.1. ([4], Theorem 6.1) A real connected Lie group G is generated by its subgroups $\{\exp(tX)\}$, $X \in S$, if and only if S generates the Lie algebra $\mathfrak{g}$.

In Theorem 1.1 the parameters in the subgroups $\{\exp(tX)\}$ could be both positive and negative. In applications to controllability of quantum systems, the parameters t represent time, and must be positive. If we restrict the question to controllability with positive time parameters, then the analogue of Theorem 1.1 still holds for compact Lie groups:

Theorem 1.2. ([4], Theorem 6.3) Let G be a compact connected real group. If the set S generates the Lie algebra of G then every element $g \in G$ admits a factorization

$$g = \exp(t_1 C_1) \cdot \ldots \cdot \exp(t_n C_n)$$

for some $n \geq 0$ with $C_i \in S$ and $t_i > 0$.

These theorems are proved using topological methods, and do not provide an effective way of finding such decompositions. In this context it is natural to pose the problem of describing decompositions that are time-optimal:

Problem. For a given $g \in G$ determine

$$\inf \left\{ t_1 + \ldots + t_n \,\middle|\, g = \exp(t_1 C_1) \cdot \ldots \cdot \exp(t_n C_n), \ t_i \geq 0, C_i \in S \right\}. \tag{1.1}$$

A compact Lie group G is isomorphic to a Lie subgroup in a general linear group ([5], Corollary 4.22), so we assume that $G \subset GL_d(\mathbb{F})$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. The Lie algebra $\mathfrak{g}$ is then a real subalgebra in the matrix Lie algebra $M_d(\mathbb{F})$.

Consider an arbitrary $\mathbb{R}$-bilinear positive-definite scalar product $\langle \cdot, \cdot \rangle$ on $M_d(\mathbb{F})$. We can make it left and right G-invariant by integration over G using the Haar measure on G ([5], Section IV.2):

$$(A, B) = \iint\limits_{G \times G} \langle gAh, gBh \rangle \, dg dh.$$

Since the Haar measure on G is left- and right-invariant, this averaging procedure will yield a G-bi-invariant scalar product $(\cdot, \cdot)$ on $M_d(\mathbb{F})$ which is still positive-definite. This scalar product defines a norm $|\cdot|$ on $M_d(\mathbb{F})$. We can rescale the above scalar product to achieve

$$|AB| \leq |A| \cdot |B| \quad \text{for all} \quad A, B \in M_d(\mathbb{F}).$$

3

The restriction of this scalar product to $\mathfrak{g}$ gives a positive-definite bilinear form on $\mathfrak{g}$, which is invariant under the conjugation action of G on its Lie algebra:

$$(gAg^{-1}, gBg^{-1}) = (A, B) \quad \text{for} \quad A, B \in \mathfrak{g}, \ g \in G.$$

Note that the Killing form on $\mathfrak{g}$ yields a proportional scalar product when $\mathfrak{g}$ is simple, but in case of a general compact Lie group G the Killing form has a disadvantage that it is only (negative) semi-definite ([5], Corollary 4.26), and may be degenerate, whereas the form constructed above is strictly positive-definite.

We get the following estimate on the norm of a Lie bracket:

$$\|[A, B]\| = |AB - BA| \leq 2 \cdot |A| \cdot |B|.$$

The exponential map

$$\exp : \ \mathfrak{g} \to G, \quad \exp(A) = \sum_{k=0}^{\infty} \frac{1}{k!} A^k$$

is surjective for compact G ([5], Corollary 4.48), but is not injective. However locally, it maps bijectively a neighbourhood of $0 \in \mathfrak{g}$ to a neighbourhood of $1 \in G$. The logarithm, which is a local inverse of the exponential map and assumes values in $\mathfrak{g}$, is defined for $g \in G$ satisfying $|1 - g| < 1$ using the standard formula

$$\ln(g) = -\sum_{k=1}^{\infty} \frac{1}{k} B^k, \quad \text{where} \quad B = 1 - g.$$

The following Lemma is a straightforward norm estimation with power series of operators, and we omit its proof:

Lemma 1.3. (a) $|\exp(A)| \leq e^{|A|}$,
(b) $|\exp(A + B) - \exp(A)| \leq e^{|A|} \left(e^{|B|} - 1\right)$,
(c) $|\exp(A + B) - \exp(A)| \leq e^{|A|}|B| + o(|B|)$, as $|B| \to 0$.

Let us recall the Campbell-Hausdorff formula ([1], II.6.4):

$$\exp(A)\exp(B) = \exp\left(\sum_{r,s\geq 0} H_{r,s}(A, B)\right), \tag{1.2}$$

where

$$H_{r,s}(A, B) = \frac{1}{r+s} \sum_{m\geq 1} \frac{(-1)^{m-1}}{m} \sum_{\substack{r_1+\ldots r_m=r \\ s_1+\ldots s_m=s \\ r_i+s_i\geq 1}} \left(\prod_{i=1}^{m} \frac{1}{r_i!s_i!}\right) [A^{r_1}B^{s_1}\ldots A^{r_m}B^{s_m}]$$

and $H_{0,0}(A, B) = 0$. Here for a monomial $U = A^{r_1} B^{s_1} \ldots A^{r_m} B^{s_m}$, we denote by $[U]$ the iterated commutator, defined inductively for $C \in \{A, B\}$ as $[C] = C$, $[CU] = [C, [U]]$. The iterated commutator $[A^{r_1} B^{s_1} \ldots A^{r_m} B^{s_m}]$ is zero unless either $s_m = 1$ or $s_m = 0$ and $r_m = 1$.

The first few terms of this series are:

$$\sum_{r,s \geq 0} H_{r,s}(A, B) = A + B + \frac{1}{2}[A, B] + \frac{1}{12}[A, [A, B]] - \frac{1}{12}[B, [A, B]] + \ldots.$$

Let us discuss convergence of the Campbell-Hausdorff series $\sum\limits_{r,s \geq 0} H_{r,s}(A, B)$. Consider a function in two variables

$$f(u, v) = -\frac{1}{2} \ln(2 - e^{2u + 2v}).$$

We can expand this function in a Taylor series

$$f(u, v) = \sum_{r,s \geq 0} \eta_{r,s} u^r v^s,$$

which is absolutely convergent when $|u| + |v| \leq \frac{\ln 2}{2}$.

By Lemma II.7.1 in [1], we can estimate the terms of the Campbell-Hausdorff series:

$$|H_{r,s}(A, B)| \leq \eta_{r,s} |A|^r |B|^s.$$

Thus the Campbell-Hausdorff series is convergent in $\mathfrak{g}$ and the equality (1.2) holds when $|A| + |B| \leq \frac{\ln 2}{2}$. In addition, a tail of the Campbell-Hausdorff series may be estimated with the corresponding tail of the series $f(|A|, |B|)$.

We will also use the following well-known formula ([5], Proposition 1.93):

$$\exp(A) B \exp(-A) = \exp(\mathrm{ad} A) B,$$

where $\mathrm{ad}(A) B = [A, B]$.

Theorem 1.4. Optimization problem (1.1) in SU_2 with $S = \{X, Y\}$ is equivalent to the problem with the set of controls $\overline{S} = \{\tau X + (1 - \tau) Y \mid \tau \in [0, 1]\}$.

Proof. We need to show that for every $\tau \in [0, 1]$ the infimum of time in (1.1) with $S = \{X, Y\}$ for $g = \exp(\tau X + (1 - \tau) Y)$ is less or equal to 1. It is well-known that g can be approximated with an arbitrary precision in time 1 (see e.g., Theorem 3.7 in [4]). Indeed, let $A = \tau X$, $B = (1 - \tau) Y$. Then

$$g = \exp(A + B) = \lim_{N \to \infty} \left[\exp\left(\frac{A}{N}\right) \exp\left(\frac{B}{N}\right) \right]^N. \tag{1.3}$$

To see this, we apply the Campbell-Hausdorff formula,

$$\exp\left(\frac{A}{N}\right) \exp\left(\frac{B}{N}\right) = \exp\left(\frac{A}{N} + \frac{B}{N} + \frac{1}{2N^2}[A, B] + o(1/N^2)\right),$$

5

from which (1.3) follows:

$$\left[\exp\left(\frac{A}{N}\right)\exp\left(\frac{B}{N}\right)\right]^N = \exp(A + B + O(1/N)).$$

We need to prove a stronger result, that g itself can be decomposed in a product of exponentials with control set S with a total time not exceeding $1 + \epsilon$ for an arbitrarily small $\epsilon > 0$.

A direct computation with the Campbell-Hausdorff formula shows that we can improve the approximation (1.3) in the following way:

$$h_N = \left[\exp\left(\frac{A}{2N}\right)\exp\left(\frac{B}{N}\right)\exp\left(\frac{A}{2N}\right)\right]^N = \exp\left(A + B + O(1/N^2)\right). \qquad (1.4)$$

Let us show that we can attain $\exp(A + B)$ by a small variation of time parameters in (1.4), while keeping these parameters positive. Let $\mathcal{B}(\epsilon) = \{g \in G \,|\, |1 - g| \leq \epsilon\}$. Since the norm is G-invariant, we get that the ball in G of radius ϵ with the center in $h \in G$ can be written as $h\mathcal{B}(\epsilon) = \mathcal{B}(\epsilon)h$. We are going to show that by varying three of the parameters by $N^{-\frac{3}{2}}$, we can cover the set $h_N\mathcal{B}\left(cN^{-\frac{3}{2}}\right)$, which will contain $\exp(A+B)$ for large N by Lemma 1.3(c). In order to show that a variation of certain three parameters covers a ball around h_N, we need to establish that the corresponding jacobian is non-zero. The computation of the jacobian takes place in the Lie algebra su_2. Since $\{A, B\}$ is a generating set for this Lie algebra, the set $\{A, B, [A, B]\}$ forms a basis of su_2. The tangent vector to $\exp(t \cdot \mathrm{ad}(A + B))B$ at $t = 0$ is $[A, B]$, which implies that for some $M \in \mathbb{N}$, the Lie algebra su_2 is spanned by $\left\{A, B, \exp\left(M^{-1}\mathrm{ad}(A + B)\right)B\right\}$. Fix M, and assume that N is a multiple of M, $N = MK$. We consider the following variation of (1.4):

$$\exp\left(\frac{A}{2N} + \epsilon_1 A\right)\exp\left(\frac{B}{N} + \epsilon_2 B\right)\left[\exp\left(\frac{A}{N}\right)\exp\left(\frac{B}{N}\right)\right]^{K-1}\exp\left(\frac{A}{N}\right)$$

$$\times \exp\left(\frac{B}{N} + \epsilon_3 B\right)\left[\exp\left(\frac{A}{N}\right)\exp\left(\frac{B}{N}\right)\right]^{N-K-1}\exp\left(\frac{A}{2N}\right).$$

The differential of this variation is

$$\exp\left(\frac{A}{2N}\right)\left\{\epsilon_1 A + \epsilon_2 B + \epsilon_3\left[\exp\left(\frac{B}{N}\right)\exp\left(\frac{A}{N}\right)\right]^K B\left[\exp\left(\frac{B}{N}\right)\exp\left(\frac{A}{N}\right)\right]^{-K}\right\}$$

$$\times \left[\exp\left(\frac{B}{N}\right)\exp\left(\frac{A}{N}\right)\right]^{N-1}\exp\left(\frac{B}{N}\right)\exp\left(\frac{A}{2N}\right).$$

Since

$$\lim_{N\to\infty}\left[\exp\left(\frac{B}{N}\right)\exp\left(\frac{A}{N}\right)\right]^K B\left[\exp\left(\frac{B}{N}\right)\exp\left(\frac{A}{N}\right)\right]^{-K} = \exp\left(M^{-1}\mathrm{ad}(A + B)\right)B,$$

we conclude that for large N the jacobian of this variation is non-zero. Thus by varying the corresponding time parameters with $|\epsilon_1|, |\epsilon_2|, |\epsilon_3| \leq N^{-\frac{3}{2}}$, we can cover the set $h_N \mathcal{B}\left(cN^{-\frac{3}{2}}\right)$, for some constant $c > 0$ that depends on A and B. Since h_N approximates $g = \exp(A + B)$ with precision $O(N^{-2})$, this set contains g for large N. This completes the proof of the Theorem.

Since a uniform rescaling of the controls $X' = cX$, $Y' = cY$, gives an equivalent optimization problem (with optimal time rescaled by a factor c^{-1}), we may assume without loss of generality that $|X| = 1$ and $|Y| \geq |X|$. Let $\kappa = 1/|Y| \leq 1$. We may pass to the normalized set of controls $\{X, Y/|Y|\}$ by replacing the total time $t_1 + \ldots + t_n$ for the decomposition $\exp(t_1 C_1) \cdot \ldots \cdot \exp(t_n C_n)$ with the cost function

$$\sum_{i=1}^{n} \kappa_i t_i, \quad \text{where } \kappa_i = \left\{ \begin{array}{ll} 1, & \text{if } C_i = X, \\ \kappa, & \text{if } C_i = Y/|Y|. \end{array} \right. \tag{1.5}$$

From now on we assume that $S = \{X, Y\}$ with $|X| = |Y| = 1$ and consider the optimization problem with cost function (1.5) where the *cost factor* $\kappa \leq 1$. This will allow us to consider the limiting case $\kappa = 0$, when there is no cost associated with control Y.

Let us introduce some terminology.

An *admissible word of length n* is an expression $\exp(t_1 C_1) \cdot \ldots \cdot \exp(t_n C_n)$ with $n \geq 0$, $t_i \geq 0$ and $C_i \in S$. A word of zero length is the identity element of G.

Every admissible word can be written in a *reduced form*, where $t_i > 0$ and $C_i \neq C_{i+1}$ for all i.

A decomposition of $g \in G$ as an admissible word of length n is called *n-optimal* if it has the minimum cost among all admissible words of length n that are equal to g.

A decomposition of $g \in G$ as an admissible word of length n is called *optimal* if it has the minimum cost among all admissible words of arbitrary lengths that are equal to g.

For a given $g \in G$, an optimal decomposition may not exist since there might be a sequence of decompositions of g of increasing lengths and with cost going to infinum. On the other hand, n-optimal decompositions exist, as we show in the following Lemma.

Lemma 1.5. Let G be a connected Lie group, and suppose that the set of controls S is finite and satisfies the following condition: every generator $C \in S$ with a zero cost factor $\kappa = 0$ has a periodic exponential, i.e., $\exp(TC) = 1$ for some $T > 0$. If $g \in G$ has a decomposition as an admissible word of length n then it has an n-optimal decomposition.

Proof. Let $t_{\inf}$ be the infinum cost over the set of all decompositions of g in admissible words of length n. Let

$$\left\{ \exp(t_1^{(j)} C_1^{(j)}) \cdot \ldots \cdot \exp(t_n^{(j)} C_n^{(j)}) \right\}_{j=1,2,\ldots}$$

be a sequence of admissible words of length n with cost converging to $t_{\inf}$. Since the set of controls S is finite, we can choose a subsequence in which the generators $\{C_k^{(j)}\}$ are independent of j.

All times $t_k^{(j)}$ are bounded by the same constant (for the generators with non-zero cost factors a bound is obtained from the bound on the total cost, and for other generators

7

from the periodicity assumption). Thus there is a subsequence where all times converge to some values, $\lim_{j\to\infty} t_k^{(j)} = t_k$. Then by continuity, $\exp(t_1 C_1) \cdot \ldots \cdot \exp(t_n C_n) = g$ and the cost of the word $\exp(t_1 C_1) \cdot \ldots \cdot \exp(t_n C_n)$ is $t_{\inf}$.

The following Lemma is obvious:

Lemma 1.6. (a) If a word of length n is optimal then it is n-optimal.

(b) If the word

$$\exp(t_1 C_1) \cdot \ldots \cdot \exp(t_n C_n) \tag{1.6}$$

is optimal (resp. n-optimal), then its subword

$$\exp(t_p C_p) \cdot \ldots \cdot \exp(t_k C_k)$$

with $1 \le p \le k \le n$ is also optimal (resp. $k - p + 1$-optimal).

(c) If the word (1.6) is optimal then

$$\exp(s_1 C_1) \exp(t_2 C_2) \cdot \ldots \cdot \exp(t_{n-1} C_{n-1}) \exp(s_n C_n)$$

with $0 \le s_1 \le t_1$, $0 \le s_n \le t_n$, is also optimal.

The last claim of the lemma suggests that for the optimal words there are stronger constraints on time parameters t_i with $2 \le i \le n - 1$. We will call these the *middle time parameters*.

2. Optimal words in SU_2.

For the rest of the paper we will focus on the case $G = SU_2$. We find it convenient to use the realization of SU_2 as the unit sphere in the quaternion algebra $\mathbb{H}$:

$$SU_2 = \left\{ a1 + bi + cj + dk \,|\, a^2 + b^2 + c^2 + d^2 = 1 \right\}.$$

The Lie algebra su_2 in this realization is the tangent space at 1 and has basis $\{i, j, k\}$. This basis is orthonormal relative to the invariant bilinear form. The norm in $\mathbb{H}$ satisfies $|xy| = |x| \cdot |y|$ and is SU_2-bi-invariant.

The isomorphism with the standard matrix construction of su_2 is given by Pauli matrices:

$$i \mapsto \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad j \mapsto \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad k \mapsto \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

The group SU_2 acts on its Lie algebra by the conjugation automorphisms. Since $-I$ acts trivially, this action factors through $SO_3 \cong SU_2/\{\pm I\}$, and is given by the natural action of SO_3 on $\mathbb{R}^3$.

The action of SO_3 on the unit sphere is transitive, so without loss of generality we may assume that $X = i$, while $Y = i \cos\alpha + j \sin\alpha$, where α is the angle between the vectors X and Y, $0 < \alpha < \pi$. This identification will allow us to carry out certain calculations in an explicit form. If C is an element of su_2 of norm 1 then $\exp(tC) = \cos(t) + C \sin(t)$.

Since $\exp(\pi X) = \exp(\pi Y) = -1$ is a central element, we see that an n-optimal word satisfies the following

π-**Condition:** at most one time parameter may be greater or equal to π; without loss of generality we may assume that this parameter corresponds to the generator Y, as it has a lower cost, and is not a middle time parameter.

We begin by describing 4-optimal words.

Proposition 2.1. (a) Let

$$g = \exp(t_1 X)\exp(t_2 Y)\exp(t_3 X)\exp(t_4 Y) \tag{2.1}$$

be a 4-optimal word with $t_1, t_2, t_3 < \pi$. Then either

$$\frac{\tan(t_2)}{\tan(t_3)} = \frac{\kappa - \cos(\alpha)}{1 - \kappa\cos(\alpha)} \tag{2.2}$$

or (2.1) is not reduced.

(b) The same condition holds for a 4-optimal word

$$\exp(t_4 Y)\exp(t_3 X)\exp(t_2 Y)\exp(t_1 X).$$

Proof. We will apply a version of the Lagrange multipliers method. Consider an infinitesimal variation of (2.1):

$$\exp((t_1 + \epsilon_1)X)\exp((t_2 + \epsilon_2)Y)\exp((t_3 + \epsilon_3)X)\exp((t_4 + \epsilon_4)Y), \tag{2.3}$$

with the constraint that the product still equals g. Carrying out calculations to the first order in ϵ_i, we have $\exp(\epsilon_i C_i) \approx 1 + \epsilon_i C_i$. We will collect all terms with ϵ_i in the middle of the word (2.3), using the relations:

$$\epsilon_1 X \exp(t_2 Y) = \exp(t_2 Y)\exp(-t_2 Y)\epsilon_1 X \exp(t_2 Y)$$

$$= \exp(t_2 Y)\left[(\cos(t_2) - \sin(t_2)Y)\epsilon_1 X (\cos(t_2) + \sin(t_2)Y)\right]$$

$$= \exp(t_2 Y)\epsilon_1 \left[\cos^2(t_2)X - \sin^2(t_2)YXY + \sin(t_2)\cos(t_2)[X, Y]\right].$$

Similarly,

$$\exp(t_3 X)\epsilon_4 Y = \epsilon_4 \left[\cos^2(t_3)Y - \sin^2(t_3)XYX + \sin(t_3)\cos(t_3)[X, Y]\right]\exp(t_3 X).$$

Setting $Z = \frac{1}{2}[X, Y] = \sin(\alpha)k$, we can get the following expressions in the basis $\{X, Y, Z\}$:

$$XYX = Y - 2\cos(\alpha)X, \quad YXY = X - 2\cos(\alpha)Y. \tag{2.4}$$

Using the above relations, we see that to the first order in ϵ_i, (2.3) can be written as

$$\exp(t_1 X)\exp(t_2 Y)\left[(\epsilon_1 \cos(2t_2) + \epsilon_3 + 2\epsilon_4 \cos(\alpha)\sin^2(t_3))X\right.$$

$$\left. + (2\epsilon_1 \cos(\alpha)\sin^2(t_2) + \epsilon_2 + \epsilon_4 \cos(2t_3))Y + (\epsilon_1 \sin(2t_2) + \epsilon_4 \sin(2t_3))Z\right]\exp(t_3 X)\exp(t_4 Y).$$

9

Equating the middle factor to zero, we get the system:

$$\begin{cases} \epsilon_1 \cos(2t_2) + \epsilon_3 + 2\epsilon_4 \cos(\alpha) \sin^2(t_3) = 0, \\ 2\epsilon_1 \cos(\alpha) \sin^2(t_2) + \epsilon_2 + \epsilon_4 \cos(2t_3) = 0, \\ \epsilon_1 \sin(2t_2) + \epsilon_4 \sin(2t_3) = 0. \end{cases} \tag{2.5}$$

If this system has rank 3 then by the implicit function theorem there exists a smooth curve in the space of parameters $\{(t_1, t_2, t_3, t_4)\}$ such that the product (2.1) is equal identically to g on the curve. The tangent vector to this curve is a non-zero solution of the system (2.5). Since by assumption the word (2.1) is 4-optimal, the differential of the cost function

$$\epsilon_1 + \kappa\epsilon_2 + \epsilon_3 + \kappa\epsilon_4$$

must be zero on the tangent vector of the curve. Thus the determinant

$$\begin{vmatrix} \cos(2t_2) & 0 & 1 & 2\cos(\alpha)\sin^2(t_3) \\ 2\cos(\alpha)\sin^2(t_2) & 1 & 0 & \cos(2t_3) \\ \sin(2t_2) & 0 & 0 & \sin(2t_3) \\ 1 & \kappa & 1 & \kappa \end{vmatrix}$$

must be zero, since the corresponding homogeneous system of equations has a non-trivial solution. In case when the system (2.5) has rank less than 3, the above determinant is still equal to zero. Evaluating this determinant we get:

$$4\sin(t_2)\sin(t_3)\left((1 - \kappa\cos(\alpha))\sin(t_2)\cos(t_3) - (\kappa - \cos(\alpha))\cos(t_2)\sin(t_3)\right) = 0.$$

Since $0 < t_2, t_3 < \pi$, we can cancel the factor $4\sin(t_2)\sin(t_3)$ and obtain the claim of the proposition.

Part (b) is completely analogous.

Remark 2.2. The denominator of $\frac{\kappa - \cos(\alpha)}{1 - \kappa\cos(\alpha)}$ is non-zero since $0 \leq \kappa \leq 1$ and $|\cos(\alpha)| < 1$. If the numerator of this fraction vanishes, the condition (2.2) should be replaced with

$$(1 - \kappa\cos(\alpha))\sin(t_2)\cos(t_3) = (\kappa - \cos(\alpha))\cos(t_2)\sin(t_3). \tag{2.6}$$

Using Proposition 2.1 and Lemma 1.6 we get a description of n-optimal words:

Corollary 2.3. Suppose $\kappa \neq \cos(\alpha)$. Let $\exp(t_1 C_1) \exp(t_2 C_2) \cdot \ldots \cdot \exp(t_n C_n)$ be a reduced n-optimal word with $n \geq 4$. Then

$$t_p = t_{p+2} \text{ for all } 2 \leq p \leq n - 3. \tag{2.7}$$

Since in an n-optimal word all middle parameters corresponding to the same control are equal, we will denote by t_x (resp. t_y) the middle time parameters corresponding to X (resp. Y).

Corollary 2.4. Under the assumptions of the previous corollary,

$$\frac{\tan(t_y)}{\tan(t_x)} = \frac{\kappa - \cos(\alpha)}{1 - \kappa \cos(\alpha)}. \tag{2.8}$$

We see from Corollaries 2.3 and 2.4 that reduced n-optimal words are described with at most three independent time parameters for all n. Since the group SU_2 is three-dimensional, we conclude that for each n there exists only a finite number of n-optimal words representing a given $g \in SU_2$ (for the case $\kappa = \cos(\alpha)$ see Theorem 2.12 below).

Next we shall investigate optimality of words of length 3.

Proposition 2.5. Let $\cos(t) > 0$ and let $\epsilon > 0$ be a small parameter. Then

$$(i) \quad \exp(\epsilon X)\exp(tY)\exp(\epsilon X) = \exp(\tau Y)\exp(\mu X)\exp(\tau Y) \tag{2.9}$$

and

$$(ii) \quad \exp(\epsilon Y)\exp(tX)\exp(\epsilon Y) = \exp(\tau X)\exp(\mu Y)\exp(\tau X), \tag{2.10}$$

where

$$\tau = t/2 + \epsilon \cos(\alpha)(1 - \cos(t)) + o(\epsilon) \tag{2.11}$$

and

$$\mu = 2\epsilon \cos(t) + o(\epsilon). \tag{2.12}$$

Proof. Let us write both sides of (i) in the form $a + bX + cY + dZ$, using the relations

$$XY = -\cos(\alpha) + Z, \quad YX = -\cos(\alpha) - Z, \tag{2.13}$$

together with (2.4). We immediately see that $d = 0$, while

$$a = \cos(2\epsilon)\cos(t) - \cos(\alpha)\sin(2\epsilon)\sin(t) = \cos(2\tau)\cos(\mu) - \cos(\alpha)\sin(2\tau)\sin(\mu), \tag{2.14}$$

$$b = \sin(2\epsilon)\cos(t) - 2\cos(\alpha)\sin^2(\epsilon)\sin(t) = \sin(\mu), \tag{2.15}$$

$$c = \sin(t) = \sin(2\tau)\cos(\mu) - 2\cos(\alpha)\sin^2(\tau)\sin(\mu). \tag{2.16}$$

We can determine μ from (2.15), getting (2.12).

From the equality

$$(c + b\cos(\alpha))\cos(\mu) - a\cos(\alpha)\sin(\mu) = \sin(2\tau)(1 - \sin^2(\alpha)\sin^2(\mu))$$

we determine τ and obtain (2.11). It can be seen that this indeed gives a solution of (i). The condition $\cos(t) > 0$ is required to ensure the positivity of μ. The claim for part (ii) follows by symmetry.

Corollary 2.6. Let $0 < t < \frac{\pi}{2}$ and let $\epsilon > 0$ be a small parameter. Then
(i) The word $\exp(\epsilon X)\exp(tY)\exp(\epsilon X)$ is not 3-optimal.
(ii) If $\kappa > \cos(\alpha)$ then the word $\exp(\epsilon Y)\exp(tX)\exp(\epsilon Y)$ is not 3-optimal.

11

Proof. We are going to show that the cost of the right hand sides of (2.9) and (2.10) are smaller than for the left hand sides. Let us evaluate the differences in these costs:

$$2\epsilon + \kappa t - 2\kappa\tau - \mu = 2\epsilon + \kappa t - \kappa t - 2\kappa\epsilon\cos(\alpha)(1 - \cos(t)) - 2\epsilon\cos(t) + o(\epsilon)$$

$$= 2\epsilon(1 - \cos(t))(1 - \kappa\cos(\alpha)) + o(\epsilon) > 0,$$

which proves (i), while for (ii) we have:

$$t + 2\kappa\epsilon - 2\tau - \kappa\mu = t + 2\kappa\epsilon - t - 2\epsilon\cos(\alpha)(1 - \cos(t)) - 2\kappa\epsilon\cos(t) + o(\epsilon)$$

$$= 2\epsilon(1 - \cos(t))(\kappa - \cos(\alpha)) + o(\epsilon) > 0.$$

Corollary 2.7. Suppose $\kappa > \cos(\alpha)$. A reduced optimal word $\exp(t_1 C_1)\ldots\exp(t_n C_n)$ of length $n > 1$ satisfies the *strong π-condition*, which is the π-condition with an additional restriction $t_1, t_n \leq \pi$.

Proof. It is sufficient to show that for a reduced optimal word $\exp(t_1 Y)\exp(t_2 X)$ of length 2, we have $t_1 \leq \pi$. Suppose $t_1 > \pi$ and let $0 < \tau < \min(t_2, \frac{\pi}{2})$. Then

$$\exp(t_1 Y)\exp(t_2 X) = \exp((t_1 - \pi)Y)\exp(\tau X)\exp(\pi Y)\exp((t_2 - \tau)X),$$

which is not optimal by Corollary 2.6(ii). This contradiction implies $t_1 \leq \pi$.

Now we can describe possible optimal decompositions in SU_2. We will consider several cases.

Theorem 2.8. Suppose $\cos(\alpha) < \kappa \leq 1$. Then the infinum of the cost of admissible decompositions of a given element of SU_2 is attained either

(a) on a reduced n-optimal word

$$\exp(t_1 C_1)\exp(t_2 C_2) \cdot \ldots \cdot \exp(t_n C_n)$$

satisfying the strong π-condition, (2.7) with $\frac{\pi}{2} \leq t_x < \pi$, $\frac{\pi}{2} \leq t_y < \pi$, and when $n \geq 4$ the condition (2.8).

or

(b) on a word

$$\exp(t_1 C_1)\exp(t_2 W)\exp(t_3 C_3),$$

where $C_1, C_3 \in \{X, Y\}$, $t_1, t_2, t_3 \geq 0$ and

$$W = (1 - \kappa\cos(\alpha))X + (\kappa - \cos(\alpha))Y$$

with the cost of $\exp(t_2 W)$ equal to $(\kappa^2 - 2\kappa\cos(\alpha) + 1)t_2$.

Proof. It is clear that the infinum of the cost is either attained on an n-optimal word, or is a limit of costs for a sequence of reduced words of length k which are k-optimal with $k \to \infty$.

In the first case, the n-optimal word must satisfy the conditions in (a) above by Corollaries 2.3, 2.4, 2.6 and 2.7.

Alternatively, if for $g \in SU_2$ the infinum of the cost is the limit of costs for a sequence of k-optimal words of increasing length,

$$g = \exp(t_1^{(j)} C_1^{(j)}) \exp(t_2^{(j)} C_2^{(j)}) \cdot \ldots \cdot \exp(t_{k_j}^{(j)} C_{k_j}^{(j)}), \quad j = 1, 2, 3 \ldots, \tag{2.17}$$

there will be a subsequence with the fixed first and last generators, $C_1^{(j)} = C', C_{k_j}^{(j)} = C''$. Since all time parameters belong to the compact set $[0, 2\pi]$, there is a subsequence where $\{t_1^{(j)}\}$ and $\{t_{k_j}^{(j)}\}$ converge,

$$\lim_{j \to \infty} t_1^{(j)} = t', \quad \lim_{j \to \infty} t_{k_j}^{(j)} = t''.$$

Time parameters in k-optimal words satisfy the conditions of Corollaries 2.3 and 2.4. Denote by $t_x^{(j)}$ the middle X-times and by $t_y^{(j)}$ the middle Y-times in (2.17). Then

$$g = \lim_{j \to \infty} \exp(t'C') \left(\exp(t_x^{(j)} X) \exp(t_y^{(j)} Y) \right)^{\left[\frac{k_j}{2} \right]} \exp(t''C'').$$

Since the total cost is bounded, we can choose a subsequence for which the sequence $\left[\frac{k_j}{2} \right] t_x^{(j)}$ converges, and let

$$\lim_{j \to \infty} \left[\frac{k_j}{2} \right] t_x^{(j)} = t_x.$$

By Corollary 2.3

$$\frac{\tan(t_y^{(j)})}{\tan(t_x^{(j)})} = \lambda, \quad \lambda = \frac{\kappa - \cos(\alpha)}{1 - \kappa \cos(\alpha)},$$

and using the fact that $\arctan(\lambda \tan(t)) = \lambda t + o(t)$ as $t \to 0$, we get that $\lim_{j \to \infty} \left[\frac{k_j}{2} \right] t_y^{(j)} = \lambda t_x$.

By (1.3) we get that

$$g = \exp(t'C') \exp(t_x(X + \lambda Y)) \exp(t''C'').$$

Finally, we note that $t_x(X + \lambda Y) = \tilde{t}W$, where $W = (1 - \kappa \cos(\alpha))X + (\kappa - \cos(\alpha))Y$ and $\tilde{t} = t_x/(1 - \kappa \cos(\alpha))$.

Remark 2.9. The vector W is orthogonal to the line passing through X and Y/κ.

Remark 2.10. Time-optimal decompositions that appear in Theorem 2.8 involve at most three independent time parameters. Since the group SU_2 is 3-dimensional, there will be a finite number of such decompositions of each length. Moreover, since middle times in the decomposition (a) above are at least $\frac{\pi}{2}$, any given decomposition of g gives a bound on the length of an optimal decomposition of type (a).

Theorem 2.11. Suppose $0 < \kappa < \cos(\alpha)$. Then the infinum of the cost of admissible decompositions of a given element of SU_2 is attained on a reduced n-optimal word

$$\exp(t_1 C_1) \exp(t_2 C_2) \cdot \ldots \cdot \exp(t_n C_n)$$

satisfying the π-condition, (2.7) with $\frac{\pi}{2} \leq t_y < \pi$, and when $n \geq 4$ the conditions $0 < t_x \leq \frac{\pi}{2}$ and (2.8).

Proof. In this case the infinum of the cost can not be realized as a limit cost of a sequence of n-optimal words of increasing lengths since by Corollary 2.4 in an n-optimal word we have either $t_x \geq \frac{\pi}{2}$ or $t_y \geq \frac{\pi}{2}$, which implies that the cost would go to infinity as $n \to \infty$.

Applying Proposition 2.1 and Corollary 2.6, we see that for a reduced optimal word of length n the claim of the theorem holds.

Theorem 2.12. Suppose $\kappa = \cos(\alpha) > 0$. Then the infinum of the cost of admissible decompositions of a given element of SU_2 is attained on a reduced n-optimal word

$$\exp(t_1 C_1) \exp(t_2 C_2) \cdot \ldots \cdot \exp(t_n C_n)$$

satisfying the π-condition, (2.7) with $\frac{\pi}{2} \leq t_y < \pi$, and when $n \geq 4$ the condition $t_x = \frac{\pi}{2}$.

Proof. Applying Proposition 2.1 (see also Remark 2.2), n-optimal words with $n \geq 4$ have middle X-times equal to $\frac{\pi}{2}$. This implies that the infinum of cost is attained on an n-optimal word and is not a limit for a sequence of n-optimal words of increasing length. It remains to show that the middle Y-times are all equal to each other. This can be done using the same method as in the proof of Lemma 2.6, by considering a the following variation of a 5-optimal word:

$$\exp((t_1 + \epsilon_1)X) \exp((t_2 + \epsilon_2)Y) \exp\left(\frac{\pi}{2}X\right) \exp((t_3 + \epsilon_3)Y) \exp((t_4 + \epsilon_4)X).$$

Evaluating the resulting 4×4 determinant, we get

$$8 \cos(\alpha) \sin^2(\alpha) \sin(t_2) \sin(t_3) \sin(t_2 - t_3) = 0,$$

which implies $t_2 = t_3$.

Finally let us consider the case when the cost associated with the generator Y is zero. In this case we have the freedom of replacing the generator Y with $-Y$ since $\exp(-tY) = \exp((2\pi - t)Y)$. Thus without loss of generality we may assume that $\cos(\alpha) \leq 0$. The case when $\cos(\alpha) < 0$ is then covered by Theorem 2.8.

Theorem 2.13. Let $\cos(\alpha) = 0$ and $\kappa = 0$. For any $g \in SU_2$ the infinum cost is attained on a word of length at most 3.

Proof. By Proposition 2.1, in an n-optimal word of length $n \geq 4$, the middle X-times are equal to $\frac{\pi}{2}$. Thus the infinum cost is attained on an n-optimal word for some n. Let for a given g, n be the smallest length such that the infinum cost is attained on an n-optimal

14

word. Suppose the optimal word has a middle factor of $\exp(\frac{\pi}{2}X)$. Applying the following identity

$$\exp(t_1 Y)\exp\left(\frac{\pi}{2}X\right)\exp(t_2 Y) = \exp((t_1 - t_2)Y)\exp\left(\frac{\pi}{2}X\right) = \exp\left(\frac{\pi}{2}X\right)\exp((t_2 - t_1)Y),$$

we see that such a word is not optimal. This implies $n \leq 3$.

In conclusion, let us consider a particular case.

Example 2.14. Let $X = i$, $Y = j$ and $\kappa = 1$. Then an optimal decomposition of any $g \in SU_2$ is given by words of the following types:

$$(a)\ \exp(t_1 C_1), \quad \text{where}\quad 0 \leq t_1 < 2\pi;$$

$$(b)\quad \exp(t_1 C_1)\exp(t_2 C_2), \quad \text{where}\quad 0 < t_1, t_2 \leq \pi;$$

$$(c)\quad \exp(t_1 C_1)\exp(t_2 C_2)\exp(t_3 C_1), \quad \text{where}\quad t_2 \geq \frac{\pi}{2},\ t_1, t_2, t_3 \leq \pi;$$

$$(d)\quad \exp(t_1 C_1)\exp(t_2 C_2)\exp(t_2 C_1)\exp(t_3 C_2), \quad \text{where}\quad t_2 \geq \frac{\pi}{2},\ t_1, t_2, t_3 \leq \pi;$$

or

$$(e)\quad \exp(t_1 C_1)\exp\left(t_2 \frac{i + j}{2}\right)\exp(t_3 C_3),$$

where $C_i \in \{X, Y\}$, $C_1 \neq C_2$, with the infinum time $\sum_k t_k$. Here $\exp\left(t\frac{i+j}{2}\right)$ may be viewed as

$$\exp\left(t\frac{i + j}{2}\right) = \lim_{N \to \infty} \left[\exp\left(\frac{ti}{2N}\right)\exp\left(\frac{tj}{2N}\right)\right]^N.$$

We obtain this result by applying Theorem 2.8. The only thing we need to show is that there are no optimal words of length $n \geq 5$. Suppose such a word is indeed optimal. Then it is also n-optimal and by Theorem 2.8, all middle times satisfy $\frac{\pi}{2} \leq t_x = t_y < \pi$. However,

$$\exp\left(\frac{\pi}{2}i\right)\exp(t_y j)\exp\left(\frac{\pi}{2}i\right) = \exp(\pi i)\exp(-t_y j) = \exp(\pi j)\exp(-t_y j) = \exp((\pi - t_y)j),$$

which has a lower total time. The analogous equality holds if i and j are switched. This shows that if the infinum time is attained on a word of length n then $n \leq 4$.

It is fairly straightforward to write down explicit formulas for t_1, t_2, t_3 in (a)-(e) above for a given $g \in SU_2$.

References:

[1] N. Bourbaki, Lie groups and Lie algebras. Chapters 1-3. Elements of Mathematics. Springer-Verlag, Berlin, 1989.

[2] D. Deutsch, A. Barenco, A. Ekert, Universality in quantum computation, Proc.R.Soc. Lond. A **449** (1995), 669-677.

[3] L. Euler, Formulae generales pro translatione quacunque corporum rigidorum, Novi Commentarii Academiae Scientiarum Petropolitanae **20** (1776), 189-207.

[4] V. Jurdjevic, Geometric control theory, Cambridge studies in advanced mathematics **51**, Cambridge University Press, New York, 1997.

[5] A. W. Knapp, Lie groups beyond an introduction, Progress in Mathematics **140**, Birkhäuser, Boston, 1996.

[6] S. Lloyd, Almost any quantum logic gate is universal, Phys.Rev.Let. **75** (1995), 346-349.

[7] H. Rabitz, R. de Vivie-Riedle, M. Motzkus, K. Kompa, Whither the future of controlling quantum phenomena? Science **288** (2000), 824-828.