An Extrinsic Approach to Sub-Riemannian Geodesics on the Orthogonal Group

Knut Hüper[1(✉)], Irina Markina[2], and Fátima Silva Leite[3,4]

[1] Institute for Mathematics, Julius-Maximilians-Universität, Würzburg, Germany
hueper@mathematik.uni-wuerzburg.de
[2] Mathematical Institute, University of Bergen, Bergen, Norway
irina.markina@uib.no
[3] Department of Mathematics, University of Coimbra, 3001-501 Coimbra, Portugal
fleite@mat.uc.pt
[4] Institute of Systems and Robotics, DEEC-UC, 3030-290 Coimbra, Portugal

Abstract. In this paper we use a variational approach, combining holonomic and nonholonomic constraints, to find an equation for sub-Riemannian geodesics on the orthogonal group. This approach is extrinsic in nature and makes the paper fully self-contained and possibly more accessible for a wide audience. The problem is formulated in the vector space of real square matrices, subject to two side conditions, and solved using a Langrange multiplier approach. The nonholonomic constraint corresponds to the requirement that the curves are tangent to a left-invariant distribution. This distribution is defined by the vector space that shows up in a Cartan decomposition of the Lie algebra associated to the orthogonal group.

Keywords: Calculus of variations · Cartan decomposition · Lagrangian multipliers · Nonholonomic constraints · Orthogonal Group · Sub-Riemannian geodesics

1 Introduction

The control of nonholonomic mechanical systems, i.e., systems whose motions are subject to nonintegrable differential constraints, has attracted growing attention over the years. We refer to [2,3,6], and [7] for details concerning the geometry of these problems and to [4] and [11] for an engineering point of view. One can find many examples of nonholonomic systems in physics, mechanical engineering and robotics. These nonholonomic constraints arise naturally in the presence of underactuated systems, which are characterised by having less inputs than degrees of freedom and include robot manipulators with flexible links and joints, mobile cars, robotic cars with trailers, aircrafts and underwater vehicles, among

J. A. Gonçalves et al. (Eds.): CONTROLO 2020, LNEE 695, pp. 274–283, 2021.
https://doi.org/10.1007/978-3-030-58653-9_26

others. The configuration space of most mechanical systems has components that are Lie groups, and in this context the orthogonal group plays an important role.

One important objective in the study of underactuated systems is the possibility of controlling their motion, so that, in spite of the reduced dimension of the input space, it is still possible to reach any configuration by properly steering the input. This corresponds to the ability of connecting an initial state to a final state by a trajectory of the system. For the problem studied here, all possible trajectories of the system correspond to curves that are tangent to the distribution defined by the nonholonomic constraints. In spite of that, the bracket generating property of the system ensures that one can move continuously between any two given admissible configurations. Among the connecting curves we distinguish an optimal one which minimises a certain cost functional induced by the natural sub-Riemannian metric on the nonintegrable distribution.

More precisely, in this paper we consider the group of orthogonal matrices, endowed with the trace metric, and look for curves that minimise the length functional and have their velocity vector restricted to be tangent to a distinguished subspace of the tangent space at each point. This problem is known as a sub-Riemannian problem on the Lie group O_n of orthogonal matrices and can be considered as a particular interesting case usually tackled by geometric optimal control theory methods involving the Pontryagin Maximum Principle and dealing with the nonholonomic constraints. However, in this paper we propose a simpler alternative approach which is extrinsic and variational in nature and does not require additional knowledge besides some variational and matrix calculus. For that, we embed the orthogonal group O_n into the vector space $\mathbb{R}^{n\times n}$ of square matrices and reformulate the minimising problem on this bigger space with additional constraints to ensure that the solution will stay in the group of orthogonal matrices. This variational view point is combined with a Lagrange multiplier approach, where the multipliers are associated to the holonomic and nonholonomic constraints.

The organisation of the paper is the following. In Sect. 2 we set the terminology and notations that will be used later. The statement of the problem and some convenient reformulations appear at the beginning of Sect. 3. In Subsect. 3.1 we derive necessary conditions for a curve in O_n to minimise the given energy functional and to satisfy the prescribed nonholonomic constraints, subject to some initial conditions. Finally, we comment about the importance of the sub-Riemannian geodesics in the orthogonal group to derive geodesics on homogeneous spaces that are quotients of O_n by certain subgroups of O_n, such as Stiefel and Grassmann manifolds.

2 Terminology and Notations

In what follows, $\mathbb{R}^{n\times n}$ stands for the Lie algebra of $n \times n$ real matrices, the Lie bracket being the matrix commutator, i.e., $[A, B] := AB - BA$. The orthogonal group O_n in its standard representation is denoted by

$$\mathrm{O}_n := \{X \in \mathbb{R}^{n\times n} \mid X^\top X = I_n\}, \tag{1}$$

where I_n is the $n \times n$-identity matrix. The Lie algebra of O_n is the set of real skewsymmetric $n \times n$-matrices, denoted by $\mathfrak{so}_n$, and so defined by

$$\mathfrak{so}_n := \{A \in \mathbb{R}^{n\times n} \mid A = -A^\top\}. \tag{2}$$

The tangent space of O_n at X is then

$$T_X\mathrm{O}_n \cong \mathfrak{so}_n\, X \cong X\,\mathfrak{so}_n\,. \tag{3}$$

In the sequel we will mainly use the second isomorphism in (3). In particular, to each $A \in \mathfrak{so}_n$ one associates a left-invariant vector field in O_n defined by XA, with $X \in \mathrm{O}_n$. We also need the vector space of symmetric $(n \times n)$ matrices

$$\mathrm{sym}_n := \{A \in \mathbb{R}^{n\times n} | A = A^\top\}. \tag{4}$$

The vector space $\mathbb{R}^{n\times n}$ is also considered as a Riemannian manifold endowed with the Euclidean metric, i.e., with the usual Frobenius inner product $\langle\cdot,\cdot\rangle$ on each tangent space $T_X\mathbb{R}^{n\times n} \cong \mathbb{R}^{n\times n}$

$$\langle\cdot,\cdot\rangle\colon \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n} \to \mathbb{R}, \quad \langle A, B\rangle := \mathrm{tr}(A^\top B). \tag{5}$$

It is well known that $\mathbb{R}^{n\times n}$ is a direct sum of $\mathfrak{so}_n$ and sym_n and, moreover, $\langle \mathfrak{so}_n, \mathrm{sym}_n\rangle = 0$. Consequently, we write:

$$\mathbb{R}^{n\times n} = \mathfrak{so}_n \oplus_\perp \mathrm{sym}_n. \tag{6}$$

Accordingly, any matrix $A \in \mathbb{R}^{n\times n}$ can be uniquely decomposed as the sum of a skewsymmetric and a symmetric matrix, here denoted by A_{skew} and A_{sym} respectively, that is,

$$A = A_{\mathrm{skew}} + A_{\mathrm{sym}}, \tag{7}$$

where

$$A_{\mathrm{skew}} := \tfrac{A-A^\top}{2}, \quad \text{and} \quad A_{\mathrm{sym}} := \tfrac{A+A^\top}{2}. \tag{8}$$

Furthermore, we consider any Cartan decomposition of $\mathfrak{so}_n$:

$$\mathfrak{so}_n = \mathfrak{k} \oplus \mathfrak{p}, \tag{9}$$

where $\mathfrak{k}$ is a Lie subalgebra of $\mathfrak{so}_n$ and $\mathfrak{p}$ is a vector space satisfying the additional Lie algebraic relations

$$[\mathfrak{k}, \mathfrak{k}] \subseteq \mathfrak{k}, \quad [\mathfrak{p}, \mathfrak{k}] \subseteq \mathfrak{p}, \quad \mathfrak{k} \subseteq [\mathfrak{p}, \mathfrak{p}]. \tag{10}$$

The restriction of $\langle\cdot,\cdot\rangle$ to $\mathfrak{so}_n$ is a scalar multiple of the Killing form and it also happens that $\mathfrak{p}^\perp = \mathfrak{k}$. So, instead of (9), we write

$$\mathfrak{so}_n = \mathfrak{k} \oplus_\perp \mathfrak{p}. \tag{11}$$

Putting together (6) and (11), we have

$$\mathbb{R}^{n\times n} = \mathfrak{k} \oplus_\perp \mathfrak{p} \oplus_\perp \mathrm{sym}_n, \tag{12}$$

i.e., every matrix $A \in \mathbb{R}^{n\times n}$ can be uniquely decomposed as

$$A = A_{\mathfrak{k}} + A_{\mathfrak{p}} + A_{\mathrm{sym}}. \tag{13}$$

The inner product $\langle \cdot, \cdot \rangle$ satisfies

$$\langle A, [B, C] \rangle = \langle B, [C, A] \rangle, \quad A, B, C \in \mathfrak{so}_n\,. \tag{14}$$

The space $\mathfrak{p}$ induces a family $\mathcal{H}_{\mathfrak{p}}$ of left-invariant vector fields, and the associated distribution $\mathcal{H}_{\mathfrak{p}}(X) = \{XA \in T_X\mathrm{O}_n | X \in \mathrm{O}_n,\ A \in \mathfrak{p}\}$ is called horizontal. The left-invariant vector fields XB, with $B \in \mathfrak{k}$, are called vertical. A smooth curve $t \mapsto X(t) \in \mathrm{O}_n$ that satisfies $X^\top(t)\dot{X}(t) \in \mathfrak{p}$ is called horizontal.

Conditions (10) imply that $\mathfrak{p} + [\mathfrak{p}, \mathfrak{p}] = \mathfrak{so}_n$, which in turn implies that $\mathcal{H}_{\mathfrak{p}}$ is a two step bracket generating distribution and, consequently, any two points in the same connected component of O_n can be joined by a horizontal curve [1,5,13].

3 A Sub-Riemannian Geodesic Problem on O_n

We are interested in the solutions of the following optimal control problem:

$$\min_{u_i \in L^2([0,1],\mathbb{R}^m)} \ \tfrac{1}{2}\int_0^1 \sum_{i=1}^m u_i^2(t)\,\mathrm{d}t \tag{15}$$

subject to:

$$\dot{X}(t) = X(t)\left(\sum_{i=1}^m u_i(t) A_i\right), \quad X \in \mathrm{O}_n, \quad m < n, \tag{16}$$

$$X(0) = X_0,\ X(1) = X_1,$$

where $\{A_1, \ldots, A_m\}$ is an orthonormal basis for the subspace $\mathfrak{p} \subset \mathfrak{so}_n$, and X_0 and X_1 are given matrices belonging to the same connected component of O_n.

The above optimal control problem can be reformulated as the following sub-Riemannian geodesic problem on O_n.

$$\min_{X \in \mathrm{O}_n} \tfrac{1}{2}\int_0^1 \mathrm{tr}\left(\dot{X}^\top(t)\dot{X}(t)\right)\mathrm{d}t \tag{17}$$

subject to

$$\begin{array}{ll} X^\top(t)\dot{X}(t) \in \mathfrak{p}, & \text{(nonholonomic constraint);} \\ X(0) = X_0,\ X(1) = X_1, & \text{(boundary conditions).} \end{array} \tag{18}$$

Solutions of this problem are called *sub-Riemannian geodesics*. These problems can be tackled using the geometric optimal control theory involving the Pontryagin Maximum Principle and consequently Hamiltonian methods. This approach has been addressed before, for instance in [8] and [10], but here we propose an alternative variational approach that is extrinsic in nature.

The sub-Riemannian geodesic problem on O_n is formulated as a Riemannian problem on $\mathbb{R}^{n\times n}$ over the space of smooth curves $\mathcal{S} := C^\infty([0,1], \mathbb{R}^{n\times n})$, subject to holonomic and nonholonomic constraints:

$$\min_{X\in\mathcal{S}} \tfrac{1}{2}\int_0^1 \operatorname{tr}\left(\dot{X}^\top(t)\dot{X}(t)\right) \mathrm{d}t \tag{19}$$

subject to:

$$\begin{array}{ll} X^\top(t)X(t) = I_n, & \text{(holonomic constraint);} \\ X^\top(t)\dot{X}(t) \in \mathfrak{p}, & \text{(nonholonomic constraint);} \\ X(0) = X_0,\ X(1) = X_1, & \text{(boundary conditions).} \end{array} \tag{20}$$

3.1 A Lagrange Multiplier Approach to Solve the Sub-Riemannian Geodesic Problem

To solve this problem we use a Langrange multiplier approach, by extending the Lagrangian in (19) by adding extra terms corresponding to the holonomic and nonholonomic constraints. So, instead of minimising the functional (19) we minimise the functional defined by

$$\begin{aligned} &F\colon \widetilde{\mathcal{S}} \to \mathbb{R}, \\ &(X,S,Q) \mapsto \tfrac{1}{2}\int_0^1 \operatorname{tr}\left(\dot{X}^\top\dot{X} + S^\top\left(X^\top X - I_n\right) + Q^\top X^\top\dot{X}\right)\mathrm{d}t =: F(X,S,Q). \end{aligned} \tag{21}$$

Here $\widetilde{\mathcal{S}} := C^\infty([0,1], V)$, where $V := \mathbb{R}^{n\times n} \times \mathrm{sym}_n \times (\mathfrak{k} \times \mathrm{sym}_n)$. The first summand in (21) is the energy functional, i.e., the squared norm of the velocity of the smooth curve $X(t) \in \mathbb{R}^{n\times n}$ integrated over the interval $[0,1]$. The second summand in (21) incorporates the holonomic constraint in (20). Here the matrix $S \in \mathrm{sym}_n$ serves as a matrix valued Langrange multiplier to ensure that the curve X will stay on the orthogonal group. Clearly, by taking derivatives this will imply $X^\top(t)\dot{X}(t) = -\dot{X}^\top(t)X(t) \in \mathfrak{so}_n$ for all $t \in [0,1]$. The third summand in (21) will ensure the nonholonomic constraint in (20). Here the matrix $Q \in \mathfrak{k} \oplus_\perp \mathrm{sym}_n$ will serve as a second Lagrange multiplier, ensuring that the velocity $\dot{X}$, when left-translated back to the identity $I_n \in \mathrm{O}_n$, will stay in the subspace $\mathfrak{p}$.

According to the theory of calculus of variations, a neccessary condition for a critical point of the functional (21) is the vanishing of the first derivative of the following smooth function, evaluated at 0.

$$\begin{aligned} &F_\varepsilon\colon [-\delta,\delta] \times \widetilde{\mathcal{S}} \to \mathbb{R}, \\ &\big(\varepsilon, (X_\varepsilon, S_\varepsilon, Q_\varepsilon)\big) \mapsto \tfrac{1}{2}\int_0^1 \operatorname{tr}\left(\dot{X}_\varepsilon^\top\dot{X}_\varepsilon + S_\varepsilon\left(X_\varepsilon^\top X_\varepsilon - I_n\right) + Q_\varepsilon^\top X_\varepsilon^\top\dot{X}_\varepsilon\right)\mathrm{d}t, \end{aligned} \tag{22}$$

where $\delta > 0$, and $X_\varepsilon, S_\varepsilon, Q_\varepsilon$ are admissible variations of X, S, Q, respectively, defined by

$$X_\varepsilon := X + \varepsilon Y \in \mathbb{R}^{n\times n}, \qquad Y(0) = Y(1) = 0, \tag{23}$$

$$S_\varepsilon := S + \varepsilon T \in \mathrm{sym}_n, \qquad\qquad T(0) = T(1) = 0, \tag{24}$$

$$Q_\varepsilon := Q + \varepsilon R \in \mathfrak{k} \oplus_\perp \mathrm{sym}_n, \qquad\qquad R(0) = R(1) = 0. \tag{25}$$

By computing the first derivative of F_ε at 0, we obtain:

$$\begin{aligned}
F'_\varepsilon(0) &= \int_0^1 \mathrm{tr}\Big(\dot X^\top \dot Y + \tfrac{1}{2}T\big(X^\top X - I_n\big) + SX^\top Y \\
&\qquad + \tfrac{1}{2}R^\top X^\top \dot X + \tfrac{1}{2}Q^\top\big(Y^\top \dot X + X^\top \dot Y\big)\Big)\mathrm{d}t \\
&= \int_0^1 \mathrm{tr}\Big(- \ddot X^\top Y + \tfrac{1}{2}T\big(X^\top X - I_n\big) + (XS)^\top Y \\
&\qquad + \tfrac{1}{2}(X^\top \dot X)^\top R + \tfrac{1}{2}(\dot X Q^\top)^\top Y - \tfrac{1}{2}\big(\tfrac{\mathrm{d}}{\mathrm{d}t}(XQ)^\top\big)Y\Big)\mathrm{d}t \\
&= \int_0^1 \mathrm{tr}\Big(\big(- \ddot X + XS + \tfrac{\dot X Q^\top}{2} - \tfrac{\dot X Q + X\dot Q}{2}\big)^\top Y\Big)\mathrm{d}t \\
&\qquad + \int_0^1 \mathrm{tr}\Big(\big(\tfrac{X^\top X - I_n}{2}\big)T\Big)\mathrm{d}t + \int_0^1 \mathrm{tr}\Big(\big(\tfrac{X^\top \dot X}{2}\big)^\top R\Big)\mathrm{d}t.
\end{aligned} \tag{26}$$

For the second equality in (26) we have integrated by parts, respecting the boundary conditions which Y has to fullfil. Hence, $F'(0) = 0$ holds for all Y, T, R in their respective spaces, if and only if

$$\begin{aligned}
-\ddot X + XS + \tfrac{\dot X(Q^\top - Q)}{2} - \tfrac{X\dot Q}{2} &= 0, \\
X^\top X &= I, \\
\dot X^\top X = -X^\top \dot X &\in \mathfrak{p}.
\end{aligned} \tag{27}$$

The objective now is to simplify the first equality in (27) to get rid of the Lagrange multipliers, eventually to solve the corresponding Euler-Lagrange-type differential equation under the holonomic and nonholonomic constraints. We proceed as follows. First, we rewrite the first equality in (27) exploiting $S = S^\top$, respecting the skewsymmetry of $X^\top \dot X \in \mathfrak{so}_n$ and the holonomic constraint $X^\top X = I_n$, to arrive at

$$S = X^\top \ddot X - \tfrac{X^\top \dot X(Q^\top - Q)}{2} + \tfrac{\dot Q}{2} = \ddot X^\top X - \tfrac{(Q - Q^\top)\dot X^\top X}{2} + \tfrac{\dot Q^\top}{2} = S^\top. \tag{28}$$

Thus,

$$S - S^\top = X^\top \ddot X - \ddot X^\top X + [X^\top \dot X, \tfrac{Q - Q^\top}{2}] + \tfrac{\dot Q - \dot Q^\top}{2} = 0. \tag{29}$$

We see that the latter equality depends only on the skewsymmetric part of the Lagrange multiplier Q. Consequently, exploiting $\frac{Q - Q^\top}{2} = Q_\mathfrak{k}$ and accordingly $\frac{\dot Q - \dot Q^\top}{2} = \dot Q_\mathfrak{k}$, we obtain

$$X^\top \ddot X - \ddot X^\top X + [X^\top \dot X, Q_\mathfrak{k}] = -\dot Q_\mathfrak{k}. \tag{30}$$

Recall the special structure of the terms in (30) and the consequences, i.e.,

$$\begin{aligned} X^\top \ddot{X} - \ddot{X}^\top X &\in \mathfrak{so}_n, \\ X^\top \dot{X} &\in \mathfrak{p}, \\ [X^\top \dot{X}, Q_\mathfrak{k}] &\in \mathfrak{p}. \end{aligned} \tag{31}$$

The holonomic constraint, i.e., the second condition in (27), implies, by taking derivatives

$$X^\top X = I_n \Longrightarrow X^\top \dot{X} + \dot{X}^\top X = 0 \Longrightarrow 2\dot{X}^\top \dot{X} + X^\top \ddot{X} + \ddot{X}^\top X = 0. \tag{32}$$

The nonholonomic constraint, i.e., the third condition in (27), implies, by taking derivatives

$$X^\top \dot{X} \in \mathfrak{p} \Longrightarrow \tfrac{\mathrm{d}}{\mathrm{d}t}\big(X^\top \dot{X}\big) \in \mathfrak{p} \Longleftrightarrow \dot{X}^\top \dot{X} + X^\top \ddot{X} \in \mathfrak{p}. \tag{33}$$

The last condition in (33), respecting $\dot{X}^\top \dot{X} \in \mathrm{sym}_n$ together with (32) implies

$$(X^\top \ddot{X})_\mathfrak{k} = 0. \tag{34}$$

Last but not least, from (30) and (31) we get

$$(X^\top \ddot{X})_\mathfrak{p} = \tfrac{[Q_\mathfrak{k}, X^\top \dot{X}]}{2}. \tag{35}$$

Finally, the last equality in (30) together with (34) imply

$$\dot{Q}_\mathfrak{k} = 0 \quad \Longrightarrow \mathfrak{k} \ni Q_\mathfrak{k} = \mathrm{const}, \tag{36}$$

i.e., the skewsymmetric part $Q_\mathfrak{k}$ of the Lagrange multiplier Q is a constant. In summary, we have proved the following result

Theorem 1. *The necessary condition for the matrix $X \in \mathbb{R}^{n\times n}$ to be a critical point of the functional* (21) *is given by the system of differential-algebraic equations*

$$\begin{aligned} (X^\top \ddot{X})_\mathfrak{p} &= \tfrac{[Q_\mathfrak{k}, X^\top \dot{X}]}{2}, \\ (X^\top \ddot{X})_\mathfrak{k} &= 0, \\ X^\top X &= I_n, \\ X^\top \dot{X} &\in \mathfrak{p}. \end{aligned} \tag{37}$$

Corollary 1. *System* (37) *in Theorem 1 implies that a critical point of the functional* (21) *satisfies the one-parameter family of ordinary differential equations*

$$\ddot{X} = X\big((X^\top \dot{X})^2 + [Q_\mathfrak{k}, X^\top \dot{X}]\big). \tag{38}$$

Proof. According to the decomposition (13),

$$X^\top \ddot{X} = (X^\top \ddot{X})_{\mathrm{sym}} + (X^\top \ddot{X})_{\mathfrak{k}} + (X^\top \ddot{X})_{\mathfrak{p}}. \tag{39}$$

From the last equality in (32) and the skewsymmetry of $X^\top \dot{X}$, the symmetric part of $X^\top \ddot{X}$ can be written as

$$(X^\top \ddot{X})_{\mathrm{sym}} = \tfrac{X^\top \ddot{X} + \ddot{X}^\top X}{2} = -\dot{X}^\top \dot{X} = -\dot{X}^\top X X^\top \dot{X} = (X^\top \dot{X})^2. \tag{40}$$

So, using the first two equations in (37) together with (40), and also absorbing the constant $\frac{1}{2}$ in the value of the constant matrix $Q_{\mathfrak{k}}$, the decomposition (39) reduces to the family of ordinary differential equations (38). □

Let us rewrite Eq. (38) in a way that involves the data from both summands of the Lie algebra $\mathfrak{so}_n = \mathfrak{p} \oplus \mathfrak{k}$. We define

$$U(t) := X^\top(t)\dot{X}(t) \in \mathfrak{p}. \tag{41}$$

Therefore, by making use of (40) and (38) we obtain

$$\begin{aligned} \dot{U} &= \dot{X}^\top \dot{X} + X^\top \ddot{X} \\ &= -(X^\top \dot{X})^2 + X^\top \ddot{X} \\ &= -U^2 + U^2 + [Q_{\mathfrak{k}}, U] = [Q_{\mathfrak{k}}, U]. \end{aligned} \tag{42}$$

That is, the horizontal tangent vector $U(t) \in \mathfrak{p}$ fulfils a linear time-invariant matrix valued ordinary differential equation of first order, with the solution

$$U(t) = \mathrm{e}^{tQ_{\mathfrak{k}}}\, U(0)\, \mathrm{e}^{-tQ_{\mathfrak{k}}}, \quad U(0) = X_0^\top \dot{X}_0, \quad \dot{X}_0 := \dot{X}(0). \tag{43}$$

In the sequel we will simply use Q instead of $Q_{\mathfrak{k}}$, since, as explained earlier, the symmetric part Q_{sym} of the Lagrange multiplier Q could be chosen equal to zero. As the main result of the paper we state the following.

Theorem 2. *For fixed $Q \in \mathfrak{k}$, the unique solution of the initial value problem*

$$\begin{aligned} \ddot{X}(t) &= X(t)\big(U^2(t) + [Q, U(t)]\big) \\ &= X(t)\,\mathrm{e}^{tQ}\,\big(P^2 + [Q, P]\big)\,\mathrm{e}^{-tQ}, \\ P := U(0) &= X_0^\top \dot{X}_0 \in \mathfrak{p}, \\ X_0 &\in O_n, \end{aligned} \tag{44}$$

is the sub-Riemannian geodesic tangent to the left-invariant distribution defined by $\mathfrak{p}$, and starting from the point $X_0 \in O_n$ with initial velocity $\dot{X}_0$. The solution is given by

$$X(t) = X_0\, \mathrm{e}^{t(P+Q)}\, \mathrm{e}^{-tQ}. \tag{45}$$

Proof. The ordinary differential equation in Theorem 2 is linear time *variant* which can be solved by means of the following time variant coordinate transformation

$$Y(t) := X(t)\,\mathrm{e}^{tQ}\,. \tag{46}$$

By making use of the substitution (46) into (44) we reduce the initial system to the Cauchy problem of the linear second order equation with *constant* coefficients

$$\begin{aligned} \ddot{Y} - 2\dot{Y}Q + Y\big(Q^2 - P^2 - [Q,P]\big) &= 0,\\ P &\in \mathfrak{p},\\ Q &\in \mathfrak{k},\\ Y(0) &\in \mathrm{O}_n. \end{aligned} \tag{47}$$

Although the corresponding characteristic second order equation

$$\mathit{\Gamma}^2 - 2\mathit{\Gamma}Q + I_n\big(Q^2 - P^2 - [Q,P]\big) = 0 \tag{48}$$

for the ordinary differential equation in (47) is matrix valued, and therefore not defined over a commutative ring, we solve it by the following Ansatz. Define the matrix $\mathit{\Gamma} := \alpha P + \beta Q$, where α and β are real numbers. Then, by substituting this Ansatz into (48), we find $\alpha = \beta = 1$. It leads to the unique solution of the initial value problem (47), given by

$$Y(t) = Y(0)\,\mathrm{e}^{t(P+Q)}\,. \tag{49}$$

The previous equation, together with (46), finishes the proof of Theorem 2. □

Formula (45) is in accordance to the results in [8] and [10], but here this formula was derived using an approach that is possibly more suitable for a wide audience.

Remark 1. Formula (45) defines a curve in O_n having velocity vector $\dot{X}(t) \in \mathcal{H}_{\mathfrak{p}}(X(t))$, or analogously $X^\top(t)\dot{X}(t) \in \mathfrak{p} \subset \mathfrak{so}_n$. This formula can be used to find the Riemannian geodesics on homogeneous spaces O_n/K, where K is the isotropy subgroup at a point in O_n, having $\mathfrak{k}$ as its Lie algebra. We get a Riemannian metric on O_n/K by pushing forward the bi-invariant metric on O_n, when restricted to the horizontal distribution $\mathcal{H}_{\mathfrak{p}} \subset T\mathrm{O}_n$. The Riemannian geodesic on O_n/K is the projection of (45) under the projection map $\pi : \mathrm{O}_n \to \mathrm{O}_n/K$, when $Q = 0$ and $P \in \mathfrak{p}$. This idea was already exploited when the homogeneous space is the Stiefel manifold $\mathrm{St}_{n,k}$ consisting of all k-orthonormal frames in $\mathbb{R}^n$, or the Grassmann manifold $\mathrm{Gr}_{n,k}$ of all k-dimensional subspaces in $\mathbb{R}^n$. In the Stiefel case $\mathrm{O}_n/K \cong \mathrm{O}_n/\mathrm{O}_{n-k} \cong \mathrm{St}_{n,k}$, and in the Grassmann case $\mathrm{O}_n/K \cong \mathrm{O}_n/(\mathrm{O}_{n-k} \times \mathrm{O}_k) \cong \mathrm{Gr}_{n,k}$. The reader can find more details in [10], Chapter 11, [12], Chapter 11, and in [9].

Acknowledgements. This work was initiated when the first and the third authors were visiting ICMAT in Madrid in November 2019. The hospitality of our hosts, David Martín de Diego and Leonardo Colombo, is greatly appreciated. This work was continued when the authors met in 2020, at the Institute for Mathematics in Würzburg, Germany.

The first author has been supported in part by the German Federal Ministry of Education and Research (BMBF-Projekt 05M20WWA: Verbundprojekt 05M2020 - DyCA).

The second author was partially supported by the project Pure Mathematics in Norway, funded by the Trond Mohn Foundation.

The third author acknowledges Fundação para a Ciência e a Tecnologia (FCT) and COMPETE 2020 program for the financial support to the project UIDB/00048/2020.

References

1. Agrachev, A.A., Sachkov, Y.L.: Control Theory from the Geometric Viewpoint. Springer, Berlin (2004)
2. Bloch, A., Colombo, L., Gupta, R., de Diego, D.M.: A geometric approach to the optimal control of nonholonomic mechanical systems. In: Analysis and Geometry in Control Theory and its Applications. Springer INdAM Series, vol. 11., pp. 35–64. Springer, Cham (2015)
3. Bloch, A.M.: Nonholonomic Mechanics and Control, 2nd edn. Springer, New York (2015)
4. Choukchou-Braham, A., Cherki, B., Djemaï, M., Busawon, K.: Analysis and Control of Underactuated Mechanical Systems. Springer, Cham (2014)
5. Chow, W.L.: Über Systeme von linearen partiellen Differentialgleichungen erster Ordnung. Math. Ann. **117**, 98–105 (1939). in German
6. Cushman, R., Duistermaat, H., Śniatycki, J.: Geometry of Nonholonomically Constrained Systems. World Scientific, Hackensack (2010)
7. Hussein, I.I., Bloch, A.M.: Optimal control of underactuated nonholonomic mechanical systems. IEEE Trans. Automat. Control **53**(3), 668–682 (2008)
8. Jurdjevic, V.: Optimal Control and Geometry: Integrable Systems. Cambridge University Press, Cambridge (2016)
9. Jurdjevic, V., Markina, I., Silva Leite, F.: Extremal curves on Stiefel and Grassmann manifolds. J. Geom. Anal. 1–31 (2019). https://doi.org/10.1007/s12220-019-00223-1
10. Montgomery, R.: A Tour of Sub-Riemannian Geometries, Their Geodesics and Applications. American Mathematical Society, Providence (2002)
11. Moreno-Valenzuela, J., Aguilar-Avelar, C.: Motion Control of Underactuated Mechanical Systems. Springer, Cham (2018)
12. O'Neill, B.: Semi-Riemannian Geometry with Applications to Relativity. Academic Press, New York (1983)
13. Rashevskii, P.K.: About connecting two points of a complete nonholonomic space by an admissible curve. Uch. Zapiski Ped. Inst. K. Liebknecht **2**, 83–94 (1938). (in Russian)