

Control Theoretic Aspects of Matrix Factorizations

U. Helmke

University of Würzburg, Mathematical Institute, Germany

http://www.mathematik.uni-wuerzburg.de/RM2

Joint work with G. Dirr and M. Kleinsteuber.

DAAD-Project: PPP Hong Kong, D/0122045,

partially supported by Marie-Curie CTS Fellowship



Contents

- Motivation
- Lie Groups and Lie Algebras
- Control on Lie Groups
- Time Optimal Control on Lie Groups
- Time Optimal Factorizations



Motivation

- Quantum Computing
- Quantum Control, Control of Spin Systems
- Control of Numerical Algorithms
- Constructive Controllability, Motion Planning in Robotics





Time-optimal Factorization Problem

- ullet G compact connected Lie group with Lie Algebra ${\mathfrak g}$
- $\omega := \{\Omega_1^+, ..., \Omega_r^+, \Omega_1^-, ..., \Omega_s^-\}$ finite set of LA generators of \mathfrak{g}
- Ω_i^+ : "slow, cost expensive" directions Ω_i^- : "fast, cheap" directions
- Given $X \in G$, define

$$T_{\min}(X) = \inf\left\{\sum_{i} |t_{i}^{+}| \mid X = \prod_{\text{finite}} e^{t_{i}^{\pm}\Omega_{i}^{\pm}}\right\}$$

Problem:

• Is $T_{\min} < \infty$ always? Compute T_{\min} !

When does there exist a *finite, time-optimal* factorization?



Example

Optimal Condition Numbers

- G = GL(n) general linear group of invertible matrices
- $\omega := \{\Omega_1^+, ..., \Omega_r^+, \Omega_1^-, ..., \Omega_s^-\}$ finite set of LA generators of $\mathfrak{gl}(\mathfrak{n})$
- Ω_i^+ : "hyperbolic Jacobi rotations" Ω_i^- : "standard Jacobi directions"
- Given $X \in G$, define (κ denotes the condition number)

$$T_{\min}(X) = \inf\left\{\sum_{i} \kappa(\mathrm{e}^{t_{i}^{+}\Omega_{i}^{+}}) | \mid X = \prod_{\text{finite}} \mathrm{e}^{t_{i}^{\pm}\Omega_{i}^{\pm}}\right\}$$

Problem:

- This factorization task with minimal total condition number!
- Does there exists factorization with better condition numbers than for X?
 Edinburgh '04, Helmke – p.5/51







Edinburgh '04, Helmke – p.6/51



Example. General linear group of invertible $n \times n$ matrices

```
GL(n,\mathbb{R}) := \{ X \in \mathbb{R}^{n \times n} | \det X \neq 0 \}.
```

Definition. A matrix *Lie group* is any subgroup $G \subset GL(n, \mathbb{R})$ that is also

a (locally closed) submanifold of $\mathbb{R}^{n \times n}$.





Examples, cont'd:

(a) The real orthogonal group

$$O(n) := \{ X \in \mathbb{R}^{n \times n} | X X^\top = I_n \}$$

(b) The *special unitary group*

$$SU(n) := \{ X \in \mathbb{C}^{n \times n} | XX^* = I_n, \det X = 1 \}$$

(c) The *Euclidean group*

$$E(n) := \left\{ \begin{bmatrix} R & p \\ 0 & 1 \end{bmatrix} \middle| R \in O(n), p \in \mathbb{R}^n \right\}.$$

The first two examples are compact groups, while the third is not.



Definition. A vector space V with a bilinear operation $[\;,\;]:V\times V\to V$ satisfying

(i)
$$[x, y] = -[y, x]$$

(ii) [x, [y, z]] + [z, [x, y]] + [y, [z, x]] = 0 (Jacobi Identity)

is called a Lie Algebra.



- Lie algebras are the tangent spaces of Lie groups.
- Theorem. Let $G \subset GL(n, \mathbb{R})$ be a matrix Lie group. Then the tangent space $\mathfrak{g} := T_I G$ at the identity matrix is a Lie algebra with commutator as the Lie bracket:

$$[X,Y] = XY - YX.$$





Examples

(a) The Lie algebra of O(n) is

$$\mathfrak{o}(n) := \{ \Omega \in \mathbb{R}^{n \times n} | \ \Omega^{\top} = -\Omega \}.$$

(b) The Lie algebra of SU(n) is

$$\mathfrak{su}(n) := \{ \Omega \in \mathbb{C}^{n \times n} | \ \Omega^* = -\Omega, \operatorname{tr}\Omega = 0 \}$$

(c) The Lie algebra of E(n) is

$$\mathbf{e}(n) := \left\{ \begin{bmatrix} \Omega & v \\ 0 & 0 \end{bmatrix} \middle| \Omega^{\top} = -\Omega, v \in \mathbb{R}^n \right\}.$$







Edinburgh '04, Helmke – p.12/51



Control on Lie Groups

- G Lie Group with Lie Algebra \mathfrak{g} .
- \blacksquare Bilinear control system on G

(
$$\Sigma$$
) $\dot{X}(t) = \left(A_d + \sum_{j=1}^m u_j(t)A_j\right) X(t), \ X(0) = I,$

where $A_d, A_1, ..., A_m \in \mathfrak{g}$.

• Reachable Set at time T > 0

$$\mathcal{R}(T) = \{ X_F \in G | \exists u_1, ..., u_m \text{ and } s \le T : X(s) = X_F \}$$

Reachable Set

$$\mathcal{R} = \cup_T \mathcal{R}(T)$$



Definition

• Accessibility: The reachable set $\mathcal{R}(T)$ has an interior point

• Local Controllability: The identity $I \in \mathcal{R}(T)$ is an interior point

• Controllability: For any $X_F \in G$ there exist controls $u_1(\cdot), ..., u_m(\cdot)$ and T > 0 s.t. the solution of (Σ) satisfies $X(0) = I, X(T) = X_F$.



Problem 1 (Accessibility)

Definition (System Lie Algebra)

 $\mathcal{L} :=$ smallest Lie subalgebra of \mathfrak{g} , containing $A_1, ..., A_m, A_d$

Generators: ([A, B] = AB - BA)

 $A_d, A_1, ..., A_m, [A_d, A_i], [A_i, A_j], [A_d, [A_i, A_j]], ...$

• Theorem. (Σ) is accessible if and only if the system Lie algebra is $\mathcal{L} = \mathfrak{g}$.



Control on Lie Groups

- Theorem (Lian et al. 1994) Suppose
 - (i) For some constant controls $u_1, ..., u_m$

$$(\Sigma_{const})$$
 $\dot{X} = (A_d + \sum_j u_j A_j) X$

is weakly positively Poisson stable. (ii) The system Lie algebra \mathcal{L} satisfies $\mathcal{L} = \mathfrak{g}$. Then the bilinear control system is controllable.

 $Accessability + Poisson \ Stability \Rightarrow Controllability$



Definition (Poisson Stability)

Flow of (Σ_{const}) : $\Phi: G \times \mathbb{R} \to G; \ (z,t) \mapsto \Phi(z,t)$

• (Σ_{const}) is Weakly Positively Poisson Stable if for all $z \in G$, any neighborhood B(z) of z and all T > 0, there exists t > T such that $\Phi(U_z, t) \cap B(z) \neq \emptyset$.

Examples: a swing (no damping), satellite attitude, ball rolling in a bowl.



- Theorem (Jurdjevic-Sussmann) Assume:
 - (i) There exist constant controls such that $A_d + \sum_j u_j A_j$ lies in a **compact** subalgebra \mathfrak{k} of \mathfrak{g} .
 - (ii) The system Lie algebra \mathcal{L} satisfies $\mathcal{L} = \mathfrak{g}$.

Then the system (Σ) is controllable.



Corollary

Let G be a **compact** connected Lie group. Then (Σ) is controllable if and only if

 $\mathcal{L} = \mathfrak{g}.$







Edinburgh '04, Helmke - p.20/51



General Notation:

• Let G be a compact Lie Group with Lie algebra \mathfrak{g} ; $K \subset G$ a compact connected Lie subgroup with LA \mathfrak{k} . Consider the bilinear control system on G

(
$$\Sigma$$
) $\dot{X} = \left(A_d + \sum_{j=1}^m u_j A_j\right) X, \quad X(0) = I$

with $A_d \in \mathfrak{g}, A_1, ..., A_m \in \mathfrak{k}$.

- Assumption:
 - Σ is controllable, i.e. $\mathfrak{g} = \mathsf{LA}$ generated by $A_d, A_1, ..., A_m$
 - $\mathfrak{k} = \mathsf{LA}$ generated by $A_1, ..., A_m$



- Given: Initial state $X_0 = I$, Final state $X_F \in G$
- Problem 1. Find controls $u_1(\cdot), ..., u_m(\cdot)$ s.t. the corresponding solution X(t) of (Σ) satisfies

$$X(0) = X_0, \ X(T) = X_F$$
 for some $T > 0$

- Problem 2.If problem 1 has at least one solution, then find a time-optimal one, i.e. one with minimal $T = T_{opt}(X_F)$.
- Problem 1 is always solvable, provided (Σ) is controllable!



Fast versus slow directions

- A_d is called the *drift term*, $A_1, ..., A_m$ the *fast directions*
- Fact 1. If $A_d = 0$ and (Σ) controllable, then can control to X_F in *arbitrarily small time*: $T_{opt}(X_F) = 0$, always!
- Fact 2. The presence of drift term $A_d \neq 0$ is responsible for $T_{\text{opt}} > 0$.
- Idea: Factor out fast directions!



Quotient System and Equivalence Principle

Consider the quotient space

$$G/K := \{Kg \mid g \in G\}$$

of left co-sets Kg, $K = \exp(\mathfrak{k})$ Lie Group generated by fast controls.

• G/K is a smooth manifold



Example: (NMR)

For the NMR Schrödinger Equation on $G = SU(2^N)$

$$\dot{X} = -i\left(H_d + \sum_{j=1}^{2N} u_j H_j\right) X, \quad X(0) = I$$

 $\mathfrak{k} := \mathsf{LA}$ generated by $iH_1, ..., iH_{2N}$ $K := \exp(\mathfrak{k})$ compact, connected Lie subgroup of $SU(2^N)$, generated by $\exp(itH_j), t \in \mathbb{R}, j = 1, ..., 2N$.

One verifies $K = SU(2) \otimes ... \otimes SU(2)$

- For N = 1 : K = SU(2) = G
- For N=2 : $K=SU(2)\otimes SU(2)\simeq SO(4)\subset SU(4)$



Quotient System and Equivalence Principle

The quotient system of

(
$$\Sigma$$
) $\dot{X} = \left(A_d + \sum_{j=1}^m u_j A_j\right) X, \quad X(0) = I, \quad X(T) = X_F$

is the control system on ${\cal G}/{\cal K}$

 (Σ/K) $\dot{P} = \operatorname{Ad}_{U(t)}(A_d)P, \quad P(0) = K, \quad P(T) = KX_F$

 $\operatorname{Ad}_g(A_d) = gA_dg^{-1}, \ g \in K.$ The control functions for (Σ/K) are arbitrary L^1_{loc} functions $t \mapsto U(t) \in K.$



Quotient System and Equivalence Principle

Theorem (Equivalence Principle).

 (Σ) is controllable on G iff (Σ/K) is controllable on G/K. Moreover, the optimal times on G and G/K coincide.

$$T_{\rm opt}^G(X_F) = T_{\rm opt}^{G/K}(KX_F)$$

Proof: PhD thesis by Khaneja

• The optimal time $T_{opt}^{G/K}$ has an interpretation within Sub-Riemannian Geometry.



Sub-Riemannian Geometry

Let M be a Riemannian manifold, E ⊂ TM a constant dimensional subbundle that satisfies the Hörmander Condition For any p ∈ M, the LA of the sections of E evaluated in p is equal to T_pM (controllability cond.)

• For any two points $x,y\in M$, the Sub-Riemannian distance is

$$d(x,y) := \inf \left\{ \int_0^1 ||\dot{\alpha}(t)| |dt | \alpha(0) = x, \alpha(1) = y, \dot{\alpha}(t) \in E_{\alpha(t)} \right\}.$$

- Example: $M = G/K, E_p := \operatorname{span}\{kA_dk^{-1} \mid k \in K\}P, P \in M$ satisfies the Hörmander Cond. (Equivalence principle)
- NMR: $M = SU(2^N)/SU(2) \otimes ... \otimes SU(2)$ Sub-Riemannian space



Sub-Riemannian Geometry

Theorem.

$$T_{\text{opt}}^{G/K}(KX_F) = d(K, KX_F)$$

Sub-Riemannian distance

• Remark. The Sub-Riemannian distance d(x, y) is greater than or equal the Riemannian distance on G/K:

 $d(x,y) \geq$ geodesic distance between x,y

There is one case where these distances are equal: Riemannian symmetric spaces.



Sub-Riemannian Geometry

• Theorem. If G/K is a Riemannian Symmetric Space, then

 $T_{\text{opt}}(X_F) = \text{length of a geodesic in } G/K \text{ that connects } K \text{ with } KX_F$

Main Advantage: Riemannian distances (i.e. lengths of geodesics) are much easier to compute than Sub-Riemannian distances.



• Theorem. The homogenous space G/K is a Riemannian symmetric space, provided $(\mathfrak{g}, \mathfrak{k})$ is a Cartan-pair, i.e. \mathfrak{g} is semisimple and

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}, \quad \mathfrak{p} := \mathfrak{k}^{\perp}$$

satisfies

 $[\mathfrak{k},\mathfrak{k}]\subset\mathfrak{k},\quad [\mathfrak{k},\mathfrak{p}]\subset\mathfrak{p},\quad [\mathfrak{p},\mathfrak{p}]\subset\mathfrak{k}$





Riemannian Symmetric Spaces

- SU(n)/SO(n) is a Riemannian Symmetric Space
- $SU(4)/SU(2) \otimes SU(2)$ is a Riemannian Symmetric Space (good! 2-Spin Case)
- $SU(8)/SU(2) \otimes SU(2) \otimes SU(2)$ is *NOT* a Riemannian Symmetric Space (bad!)







Edinburgh '04, Helmke - p.33/51



- Let G be a connected, compact Lie group with Lie algebra \mathfrak{g} .
- Let $K \subset G$ be a connected compact subgroup with Lie algebra \mathfrak{k} .
- Let $\Delta \in \mathfrak{g}$ be a drift term s.t. $\langle \Delta, \mathfrak{k} \rangle_L = \mathfrak{g}$.
- Consider the discrete control System:

 $(\Sigma_d) X_{n+1} = K_n e^{t_n \Delta} L_n X_n, X_0 = I K_n, L_n \in K, t_n \ge 0.$ For $X \in G$ let $T^d_{opt}(X) :=$

$$\inf \left\{ \sum_{n=1}^{\infty} t_n \mid \exists (K_n, L_n, t_n) : \prod_{n=1}^{\infty} K_n e^{t_n \Delta} L_n = X \right\}.$$



Problem:

- Is (Σ_d) controllable, i.e. does $T^d_{opt}(X) < \infty$ hold for all $X \in G$?
- Determine the "minimal" time $T^d_{opt}(X)$ for $X \in G$.



Generalized Version (multiple drifts)

- G compact connected Lie group with LA \mathfrak{g}
- $\omega := \{\Omega_1^+, ..., \Omega_r^+, \Omega_1^-, ..., \Omega_s^-\}$ finite set of LA generators of \mathfrak{k}
- Ω_i^+ : "slow, cost expensive" directions Ω_i^- : "fast, cheap" directions

• Given $X \in G$, define

$$T_{\min}(X) = \inf\left\{\sum_{i} |t_i^+| \mid X = \prod_{\text{finite}} e^{t_i^{\pm}\Omega_i^{\pm}}\right\}$$



Problem

- Is $T_{\min} < \infty$ always? Compute T_{\min} !
- When does there exist a *finite*, *time-optimal* factorization?



Time-optimal Factorization

Example 1 (Euler Angles)

•
$$SO(3), \ \omega = \{\Omega_1^+, \Omega_1^-\},\$$

 $\Omega_1^+ := \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \ \Omega_1^- := \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$



$$X = e^{\theta_1 \Omega_1^-} e^{\theta_2 \Omega_1^+} e^{\theta_3 \Omega_1^-}, \quad \theta_i \in [-\pi, \pi]$$

We will show: Euler Angles are time-optimal and

$$T_{\min} = |\theta_2| \in [0,\pi]$$



Example 2 (Euler Angles)

• SO(3), $\omega = \{\Omega_1^+, \Omega_2^+\}$, $\Omega_2^+ := \Omega_1^-$

Then Euler angles are i.g. NOT time-optimal:

 $T_{\min} < \theta_1 + \theta_2 + \theta_3$! (Mittenhuber)





Time-optimal Factorization

Equivalence Principle

- Let G be a connected, compact Lie group with Lie algebra \mathfrak{g} .
- Let $\mathfrak{k} := \langle A_1, ..., A_m \rangle_L$, $K := \exp \mathfrak{k}$.
- Let $\Delta \in \mathfrak{g}$ be a drift term such that $\langle \Delta, \mathfrak{k}
 angle_L = \mathfrak{g}$.
- Theorem.
 - (a) The discrete control system (Σ_d) on G is controllable and thus $T^d_{\mathrm{opt}}(X) < \infty$
 - (b) For any $X \in G$ the minimal times $T^d_{opt}(X) = T_{opt}(X)$ coincide, where $T_{opt}(X)$ is the minimal time for the control problem

$$\dot{X} = \left(\Delta + \sum_{j=1}^{m} u_j A_j\right) X, \quad X(0) = I, X(T) = X$$



• Problem: I.g. time optimal factorizations are infinite

Under what conditions on the drift term Δ are they finite?

• Definition [Haselgrove, Nielsen, Osborne]: A drift term Δ is called *lazy*, if there exists $\varepsilon > 0$ such that

$$T_{\text{opt}}(e^{t\Delta}) < t$$
 for all $t \in (0, \varepsilon)$. (**)

If Δ is not lazy, we call it *fast*.



• Theorem. If Δ is lazy, there are no finite, time optimal factorizations for any element $X \in G - K$.





- Conjecture 1: There exists a finite, time optimal factorization for all $X \in G$ iff Δ is fast.
- Conjecture 2: Δ fast $\iff [\Delta, \Delta^{\perp}] = 0.$
- Remark: Conjecture 2 implies Conjecture 1.



Computation of Optimal Time

Theorem (Khaneja). Let $(\mathfrak{g}, \mathfrak{k})$ be a Cartan pair. Let Δ^{\perp} be the orthogonal projection of Δ onto \mathfrak{p} and let \mathfrak{a} be a maximal abelian subalgebra of \mathfrak{p} that contains Δ^{\perp} . Then:

• Each $X \in G$ has a decomposition of the form

 $X = U\Sigma V \qquad \text{with } U, V \in K \text{ and } \Sigma \in \exp \mathfrak{a}.$

The minimal time is given by

$$T_{\rm opt}(X) = \min\left\{t \ge 0 \ \Big| \ \left(t \cdot \operatorname{conv} \mathcal{W}(\Delta^{\perp})\right) \cap \exp^{-1}(\Sigma) \neq \emptyset\right\},\$$

where $X = U\Sigma V$ is an arbitrary factorization of the above type and $\mathcal{W}(\Delta^{\perp})$ denotes the Weyl orbit of Δ^{\perp} .



Computation of Optimal Time



Convex hull of the Weyl Orbit of a ''symmetric'' drift term Δ

Convex hull of the Weyl Orbit of an arbitrary Δ .



Computation of Optimal Time

Example 1, cont'd:

•
$$G := \mathrm{SO}(3)$$
 and $\mathfrak{g} := \mathfrak{so}(3)$,

$$\Omega_1 := \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \Omega_2 := \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

•
$$\Delta := \alpha \Omega_1 + \beta \Omega_2$$
, $\mathfrak{k} := \langle \Omega_2 \rangle$

Euler Angles: $X = e^{\theta_1 \Omega_2} e^{\theta_2 \Omega_1} e^{\theta_3 \Omega_2}$, $\theta_i \in [-\pi, \pi]$

•
$$T_{\text{opt}}(X) = \alpha^{-1} |\theta_2|$$
,

•
$$\Delta$$
 fast $\iff \beta = 0.$



Computation of minimal time

Example: (NMR cont'd)

• NMR-Schrödinger equation on SU(4)

$$\dot{X} = -2\pi i \Big(H_d + \sum_{i=1}^4 u_i H_i \Big), \quad X(0) = I,$$

where $H_d := \sigma_z \otimes \sigma_z$, $H_1 := I_2 \otimes \sigma_x$, $H_2 := I_2 \otimes \sigma_y$, $H_3 := \sigma_x \otimes I_2$, and $H_4 := \sigma_y \otimes I_2$.

• $K = \mathrm{SU}(2) \otimes \mathrm{SU}(2).$

• $\Delta = -2\pi \mathrm{i} H_d$ and $\mathfrak{a} := \mathrm{i} \langle \sigma_x \otimes \sigma_x, \sigma_y \otimes \sigma_y, \sigma_z \otimes \sigma_z \rangle$.



Computation of minimal time

Example: (NMR cont'd)

Theorem. For all $X = U\Sigma V \in SU(4)$ and $U, V \in K$, and $\Sigma \in \exp \mathfrak{a}$ fixed it holds

•
$$T(X) = \min\left\{\sum_{n=1}^{3} |t_n| \left| e^{t_1 2\pi i (\sigma_x \otimes \sigma_x)} e^{t_2 2\pi i (\sigma_y \otimes \sigma_y)} e^{t_3 2\pi i (\sigma_z \otimes \sigma_z)} = \Sigma\right\}\right\}$$

•
$$T(X) \leq \frac{3}{2}$$



Computation of minimal time

Optimization Algorithm (NMR cont'd)

Let $X(t, u) = U(u_1, ..., u_6) \Sigma(t_1, t_2, t_3) V(u_7, ..., u_{12})$,

 $U(u_1, \dots, u_6) = e^{-i2\pi u_1 H_1} e^{-i2\pi u_2 H_2} e^{-i2\pi u_3 H_1} e^{-i2\pi u_4 H_3} e^{-i2\pi u_5 H_4} e^{-i2\pi u_6 H_3}$ $V(u_7, \dots, u_{12}) = e^{-i2\pi u_7 H_1} e^{-i2\pi u_8 H_2} e^{-i2\pi u_9 H_1} e^{-i2\pi u_{10} H_3} e^{-i2\pi u_{11} H_4} e^{-i2\pi u_{12} H_3}$ $\Sigma = e^{t_1 2\pi i (\sigma_x \otimes \sigma_x)} e^{t_2 2\pi i (\sigma_y \otimes \sigma_y)} e^{t_3 2\pi i (\sigma_z \otimes \sigma_z)}$

To compute the minimal time T(X), we combine simulated annealing with gradient methods to solve the nonlinear optimization problem: $\min \quad f(t, u) := |t_1| + |t_2| + |t_3|,$ subject to $g(t, u) := 4 - \operatorname{Retr}(X_F^*X(t, u)) = 0$

where $t = [t_1, t_2, t_3], u = [u_1, u_2, ..., u_{12}] \in [-1, 1]^{12}$



Computation of Time-optimal Pulse Sequences

Consists of two sub-problems:

• Given $T \ge 0$, solve

$$\min_{\substack{t,u}} g(t,u),$$

subject to $f(t,u) \leq T,$
 $t \geq 0.$

• Let V(T) be the global optimal value of g(t, u), associated with a given $T \ge 0$.

Minimize
$$T$$

subject to $V(T) = 0$,
 $T \ge 0$.



Computation of Time-optimal Pulse Sequences

٠

Example

$$X_F = e^{-\frac{i\pi}{4}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

 $T(X_F) = 1.499996$ t = [0.499993 | 0.500017 | 0.499986]

