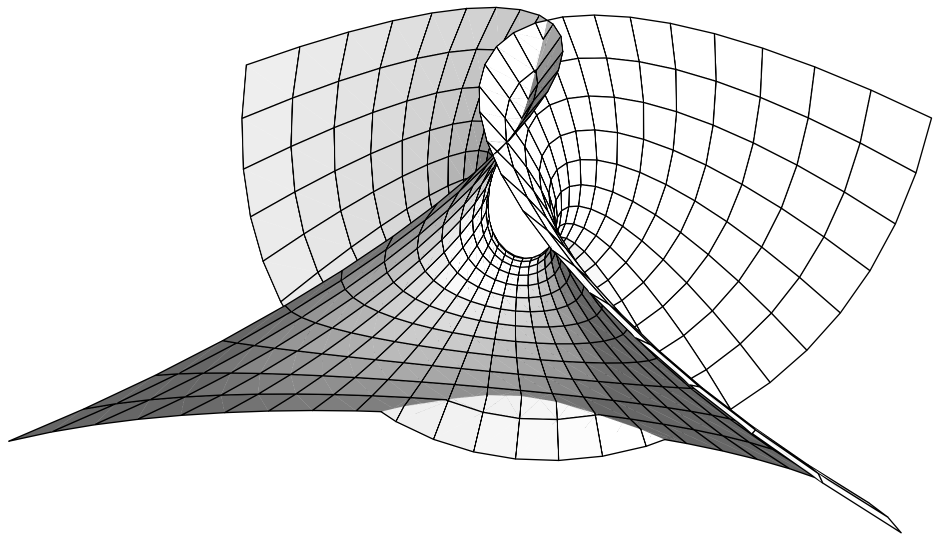
Optimization on Lie Groups: Applications in NMR Spectroscopy

6th Workshop on Dynamics and Computation

From Robotics to Quantum Control
Royal Academy of Sciences, Belgium, July 1-2, 2002

U. Helmke
University of Würzburg

Joint work with K. Hüper



The following notes cover the Brussels Lectures 2002. For questions concerning the material please contact:

Uwe Helmke

Lehrstuhl II
Mathematisches Institut
Universität Würzburg
Am Hubland
D-97074 Würzburg, Germany

helmke@mathematik.uni-wuerzburg.de

Abstract

The goal of these lectures is to demonstrate the usefulness and power of dynamical systems methods for the analysis and the design of numerical algorithms. Special emphasis is on applications relevant for quantum control and computation.

Further details can be found in the book, U. Helmke and J.B. Moore, *Optimization and Dynamical Systems*, Springer, London, 1994, as well as in the Ph.D thesis by K. Hüper, *Structure and Convergence of Jacobi-type Methods for Matrix Computations*, Technical University of Munich, 1996.

Outline

§1 Introduction

§2 Tutorial: Optimization on Lie Groups

§3 Optimization of NMR Spin Systems

§4 NMR and Representations

Design of Algorithms

Optimization on Lie groups/manifolds

$\Downarrow$

Gradient–like Flows

$\Downarrow$

Discretization of Flows

$\Downarrow$

Efficient numerical algorithms

§1 Introduction

Matrix Eigenvalue Problems

Given $X \in \mathbb{R}^{n \times n}$, find the eigenvalues and eigenvectors of X by computing a similarity transformation bringing X to some normal form.

Choose a suitable cost function

$$f : GL(n, \mathbb{R}) \to \mathbb{R}$$

whose global minima (maxima) have the desired normal form.

Examples

(a) *The "off-norm"-function*
(Jacobi 1846)

$f : SO(n, \mathbb{R}) \to \mathbb{R}$

$$f(\Theta) = \|\Theta X \Theta^\top - \mathrm{diag}\,(\Theta X \Theta^\top)\|^2$$

Facts:

(i) All minima Θ_{min} of f are global.

(ii) Θ_{min} are characterized as

$$\Theta_{\mathrm{min}} X \Theta_{\mathrm{min}}^\top = \text{diagonal}$$

Examples cont'd

(1) **Real symmetric eigenvalue problem**
Given an $X = X^\top \in \mathbb{R}^{n\times n}$.

(b) *A trace function*
(Brockett 1988)

Let $N = \mathrm{diag}\,(1, \ldots, n)$.

$$f : SO(n, \mathbb{R}) \;\to\; \mathbb{R}$$

$$f(\Theta) \;=\; \|N - \Theta X \Theta^\top\|^2$$

Facts:

(i) All minima are global.

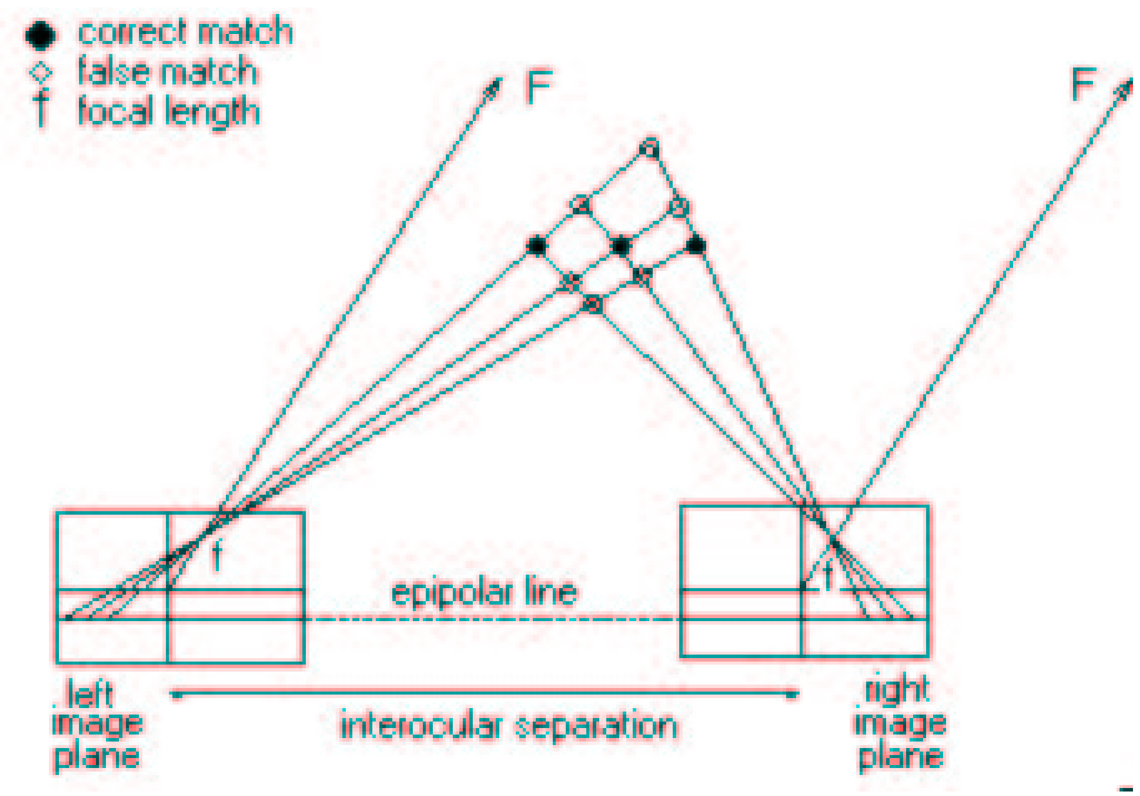
(ii) Θ_{min} are characterized as

$$\Theta_{\mathrm{min}} X \Theta_{\mathrm{min}}^\top = \mathrm{diag}\,(\lambda_1, \ldots, \lambda_n)$$

with $\lambda_1 \leq \cdots \leq \lambda_n$ the eigenvalues of X.

Examples cont'd

(2) **Stereo Matching Problem**
(without correspondence)



Examples cont'd

Given:

Positive definite symmetric $A_1, A_2 \in \mathbb{R}^{3\times 3}$

$$A_1 = \sum_{i=1}^{k} \begin{bmatrix} x_{1,i} \\ y_{1,i} \\ 1 \end{bmatrix} \begin{bmatrix} x_{1,i} \\ y_{1,i} \\ 1 \end{bmatrix}^\top, \; A_2 = \sum_{i=1}^{k} \begin{bmatrix} x_{2,i} \\ y_{2,i} \\ 1 \end{bmatrix} \begin{bmatrix} x_{2,i} \\ y_{2,i} \\ 1 \end{bmatrix}^\top .$$

Task:

Solve

$$A_2 - \Theta A_1 \Theta^\top = 0_3$$

for $\Theta \in G$ over the Lie group

$$G := \left\{ \Theta \in \mathbb{R}^{3\times 3} \middle| \Theta = I_3 + \frac{e^{b_1} - 1}{b_1} \begin{bmatrix} b_1 & b_2 & b_3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right\}$$

Solution:

Find minimizer of

$$f : G \to \mathbb{R}, \quad f(\Theta) = \|A_2 - \Theta A_1 \Theta^\top\|^2.$$

Examples cont'd

(3) $C-$**Numerical Range Computations**

Let $C, A \in \mathbb{C}^{n\times n}$ be arbitrary. The $C-$**numerical range** of A is defined as

$$W(C,A) := \{\text{tr}(C^\dagger U A U^\dagger) | U \in U(n,\mathbb{C})\}.$$

In general, the shape of $W(C,A) \subset \mathbb{C}$ is unknown.

$\Rightarrow$ Develop numerical methods to find good bounds on size of $W(C,A)$.

Cost Function:

$$f : U(n,\mathbb{C}) \to \mathbb{R}, f(U) = Re(\text{tr}(C^* U A U^*))$$

Has applications in

(i) Nuclear Magnetic Resonance (NMR) Spectroscopy

(ii) Quantum Computing

§2: Tutorial

Optimization on Lie Groups

1. Lie Groups and Lie Algebras

2. Riemannian Metrics

3. Optimization via Gradient Flows

4. Discrete–time Gradient Optimization

1. Lie Groups and Lie Algebras

A basic example of a Lie group is the general linear group of invertible $n \times n$ matrices

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

More generally

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

(a) The **real orthogonal group**

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

(b) The **unitary group**

$$U(n,\mathbb{C}) := \{X \in \mathbb{C}^{n\times n}|\ XX^{\dagger} = I_n\}$$

(c) The **Euclidean group**

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

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

A beautifully simple criterion for matrix Lie groups is

Cartan's Criterion
A subgroup $G \subset GL(n, \mathbb{R})$ is a matrix Lie group if and only if G is a closed subset of $GL(n, \mathbb{R})$.

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

(a) *The tangent space $\mathfrak{g} := T_I G$ at the identity matrix is a Lie algebra with commutator as the Lie bracket:*

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

(b) *The matrix exponential function*

$$\exp : \mathfrak{g} \to G, \quad \exp(X) = e^X$$

maps $\mathfrak{g}$ into G. It is surjective if G is a compact connected Lie group.

Examples

(a) The Lie algebra of $O(n,\mathbb{R})$ is

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

(b) The Lie algebra of $U(n,\mathbb{C})$ is

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

(c) The Lie algebra of $E(n,\mathbb{R})$ is

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

(d) The Lie algebra of $GL(n,\mathbb{R})$ is

$$\mathfrak{gl}(n,\mathbb{R}) := \mathbb{R}^{n\times n}.$$

Lie Group Actions

Definition

An **action** of a Lie group G on a manifold M is a smooth map $\alpha : G \times M \to M$, $(g, x) \to g \cdot x$, with

(i) $g \cdot (h \cdot x) = (gh) \cdot x \qquad \forall g, h \in G,\ x \in M$

(ii) $e \cdot x = x \qquad \forall x \in M$

An **orbit** is the set

$$G \cdot x := \{g \cdot x |\, g \in G\}.$$

Examples

(a) **Similarity action:** $G = GL(n, \mathbb{R}),\ M = \mathbb{R}^{n \times n}$

$$(S, X) \to SXS^{-1}.$$

(b) **Equivalence:** $G = O(n, \mathbb{R}) \times O(m, \mathbb{R}),\ M = \mathbb{R}^{n \times m}$

$$((U, V), X) \to UXV^{-1}.$$

Examples

(a) Let $X \in \mathbb{R}^{n\times n}$. Every **similarity orbit**

$$GL(n,\mathbb{R}) \cdot X = \{SXS^{-1} | S \in GL(n,\mathbb{R})\}$$

is a smooth submanifold of $\mathbb{R}^{n\times n}$.

(b) Let $X \in \mathbb{R}^{n\times m}$. Every **equivalence orbit**

$$(O(n,\mathbb{R}) \times O(m,\mathbb{R})) \cdot X$$

$$= \{UXV^\top | U \in O(n,\mathbb{R}), V \in O(m,\mathbb{R})\}$$

is a smooth submanifold of $\mathbb{R}^{n\times m}$.

Examples cont'd

(d) **Procrustes Problem on Orbits**

Let $\alpha : G \times V \to V$ be a Lie group action on a vector space V.

Definition
An orbit

$$G \cdot x = \{g \cdot x | g \in G\}$$

is called **closed** if it is a closed subset of V.

Consider any norm $\| \cdot \|$ on V. For any $a \in V$ consider the **distance function**

$$f_a : G \cdot x \to \mathbb{R}, \quad f_a(x) = \|a - g \cdot x\|^2.$$

Lemma
A minimum exists for all $a \in V$ iff $G \cdot x$ is closed.

2. Riemannian metrics

A **Riemannian metric** on a submanifold $M \subset \mathbb{R}^n$ is an inner product

$$\langle\ ,\ \rangle_x :\ T_xM \times T_xM \to \mathbb{R}$$

on each tangent space that varies smoothly with x.

A Riemannian metric is thus a smooth map

$$Q : M \to \mathbb{R}^{n\times n}$$

with

(i) $Q(x) = Q(x)^\top \ \forall x \in M$

(ii) $Q(x) > 0$ on $T_xM \times T_xM$.

Basic Fact
Every manifold M has a Riemannian metric.

Definition

Let

$$\langle\ ,\ \rangle_x :\ T_xM \times T_xM \to \mathbb{R}$$

be a Riemannian metric on M. The **gradient** of a smooth function $f : M \to \mathbb{R}$ is the unique vector field $\operatorname{grad} f$ on M such that

(i) $\operatorname{grad} f(x) \in T_xM \ \forall x \in M$.

(ii) $\mathrm{D}\, f(x)\xi = \langle \operatorname{grad} f(x), \xi \rangle_x \quad \forall \xi \in T_xM, x \in M$.

3. Optimization via Gradient Flows

Consider the optimization task

Minimize a smooth cost function $f : M \to \mathbb{R}$ on a Riemannian manifold.

Existence of Minima: *Minima of a smooth function $f : M \to \mathbb{R}$ exist if the* **sublevel sets**

$$f(-\infty, c] = \{x \in M |\, f(x) \leq c\}, \quad c \in \mathbb{R}$$

are always compact.

This is e.g. the case when M is compact.

Example

Let $M \subset \mathbb{R}^n$ be a subset. For $a \notin M$ consider

$$f : M \to \mathbb{R}, \quad f(x) = \|x - a\|^2.$$

There exists minima if M is closed.

Convergence Theorem A:
Let $f : M \to \mathbb{R}$ be a smooth function on a Riemannian manifold with compact sublevel sets. The solutions of the gradient flow

$$\boxed{\dot{x}(t) = -\,\mathsf{grad}\, f(x(t))}$$

exist for all $t \geq 0$ and converge to a connected component of the set of critical points of f.

Problem

If $f : M \to \mathbb{R}$ has **infinitely many** critical points, then solutions $x(t)$ may not converge to a **single** equilibrium point. $\longrightarrow$ **Mexican hat example.**

1 equilibrium point   infinitely many

Mexican Hat

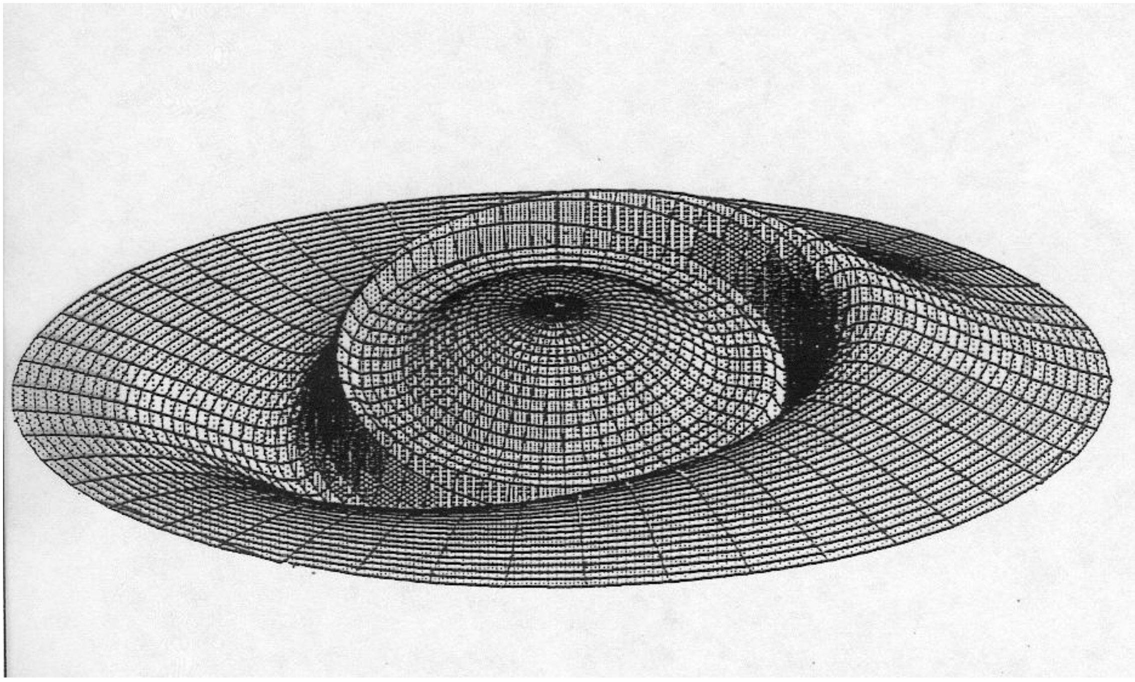


Figure 5

There is a beautiful and extremely useful result for real analytic functions.

Convergence Theorem B (Łojasiewicz)
Let $f : M \to \mathbb{R}$ be a **real analytic function** *on a real analytic Riemannian manifold with real analytic Riemannian metric. If f has compact sublevel sets, then every solution of the gradient flow*

$$\dot{x} = -\operatorname{grad} f(x)$$

converges for $t \to +\infty$ to a single equilibrium point.

Proof rests on (Łojasiewicz)

Lemma
Let $f : \mathbb{R}^n \to \mathbb{R}$ be a real analytic function and let $x^* \in \mathbb{R}^n$ be a critical point. Then there exists a neighborhood $U \subset \mathbb{R}^n$ of $x^* \in U$ and real numbers $\mu \in (0, 1)$, $C > 0, s.t.$

$$\|\operatorname{grad} f(x)\| \geq C|f(x) - f(x^*)|^\mu, \quad x \in U.$$

4. Discrete–time Gradient

Optimization

Various possibilities for minimizing a smooth function $f : M \to \mathbb{R}$ on a Riemannian manifold

(i) Discretization of gradient flow.

(ii) Conjugate gradient method.

(iii) Jacobi–type method (Gauss-Southwell, Coordinate descent).

4.1 Geodesic Approximation Scheme (GAS)
For any initial condition $x_0 \in M$ the GAS generates a sequence of points $(x_k)_{k\in\mathbb{N}}$ in M with limit points being critical points of f.

- Convergence to a single critical point is not guaranteed.
- Convergence to saddle points or local minima possible.

GAS–Recursion

Given $a := x_k \in M$ *with* $v := -\operatorname{grad} f(a) \in T_a M$ *let* $t \to \exp(tv)$ *be the geodesic through* a *with initial velocity* v. *Set*

$$x_{k+1} := \exp(t_* v)$$

$$t_* := \arg\min f(\exp(tv)).$$

- t_* and x_{k+1} can often not be computed exactly $\rightsquigarrow$ replace by approximations via e.g. Newton–method

- Ph.D. Theses by Smith (1993), Mahony (1994).

- **Bottleneck:** Computation of geodesics $\exp(tv)$ can be hard.

Geodesic Approximations

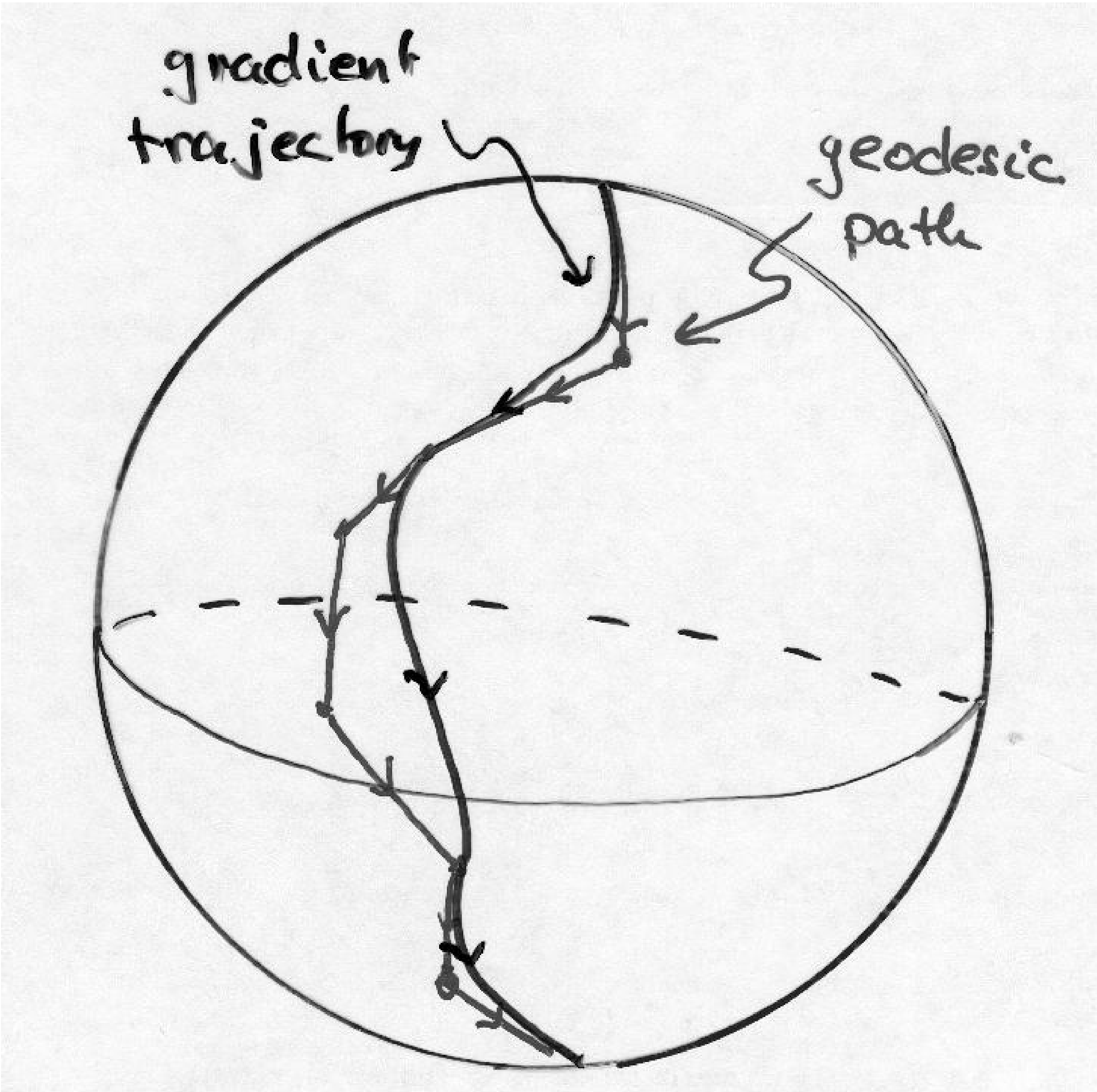


Figure 6

Exponential Map

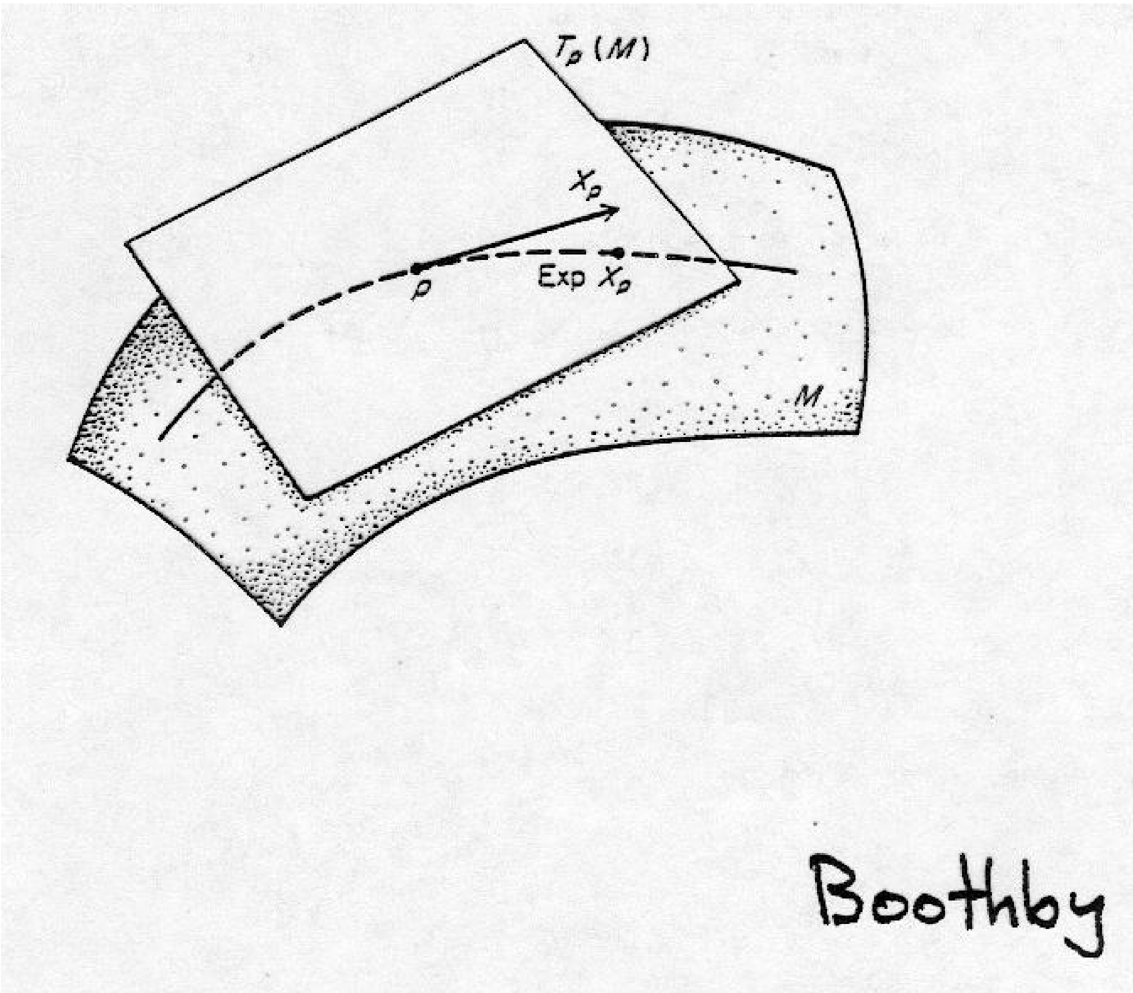


Figure 7

4.2 Conjugate Gradient Method (CGM)
Variant of the above. See Edelman, Smith.

Bottleneck: Computation of geodesics can be hard.

4.3 Jacobi–type method
Here we do not minimize along geodesics in gradient direction, but rather in pre–chosen directions. The choice of these directions depends on computational simplicity and the attempt to achieve fast convergence rates. This approach usually outperforms the previous ones.

Jacobi Algorithm on Homogeneous Spaces

Let $\alpha : G \times M \to M$ be a Lie group action, $G \cdot x$ an orbit, and $f : G \cdot x \to \mathbb{R}$ be a smooth function with compact sublevel sets.

Choose a basis $v_1, \ldots, v_N$ of the Lie algebra $\mathfrak{g}$ of G, with associated 1–parameter groups $g_i : \mathbb{R} \to G$,

$$g_i(t) = \exp(tv_i), \quad i = 1, \ldots, N.$$

Jacobi Sweep

$$\begin{aligned}
x_k^{(1)} &:= g_1(t_*^{(1)}) \cdot x_k \\
x_k^{(2)} &:= g_2(t_*^{(2)}) \cdot x_k^{(1)} \\
&\vdots \\
x_k^{(N)} &:= g_N(t_*^{(N)}) \cdot x_k^{(N-1)}
\end{aligned}$$

$x_k^{(i)}$:= Global minimum of f restricted to the closure of $\{g_i(t) \cdot x_k^{(i-1)} | \, t \in \mathbb{R}\}$.

Jacobi Algorithm on Homogeneous Spaces

- Let $x_0, \ldots, x_k \in G \cdot x$ be given for $k \in \mathbb{N}_0$.

- Define the recursive sequence $x_k^{(1)}, \ldots, x_k^{(N)}$ (sweep).

- Set $x_{k+1} := x_k^{(N)}$. Proceed with the next sweep.

Jacobi Iterations

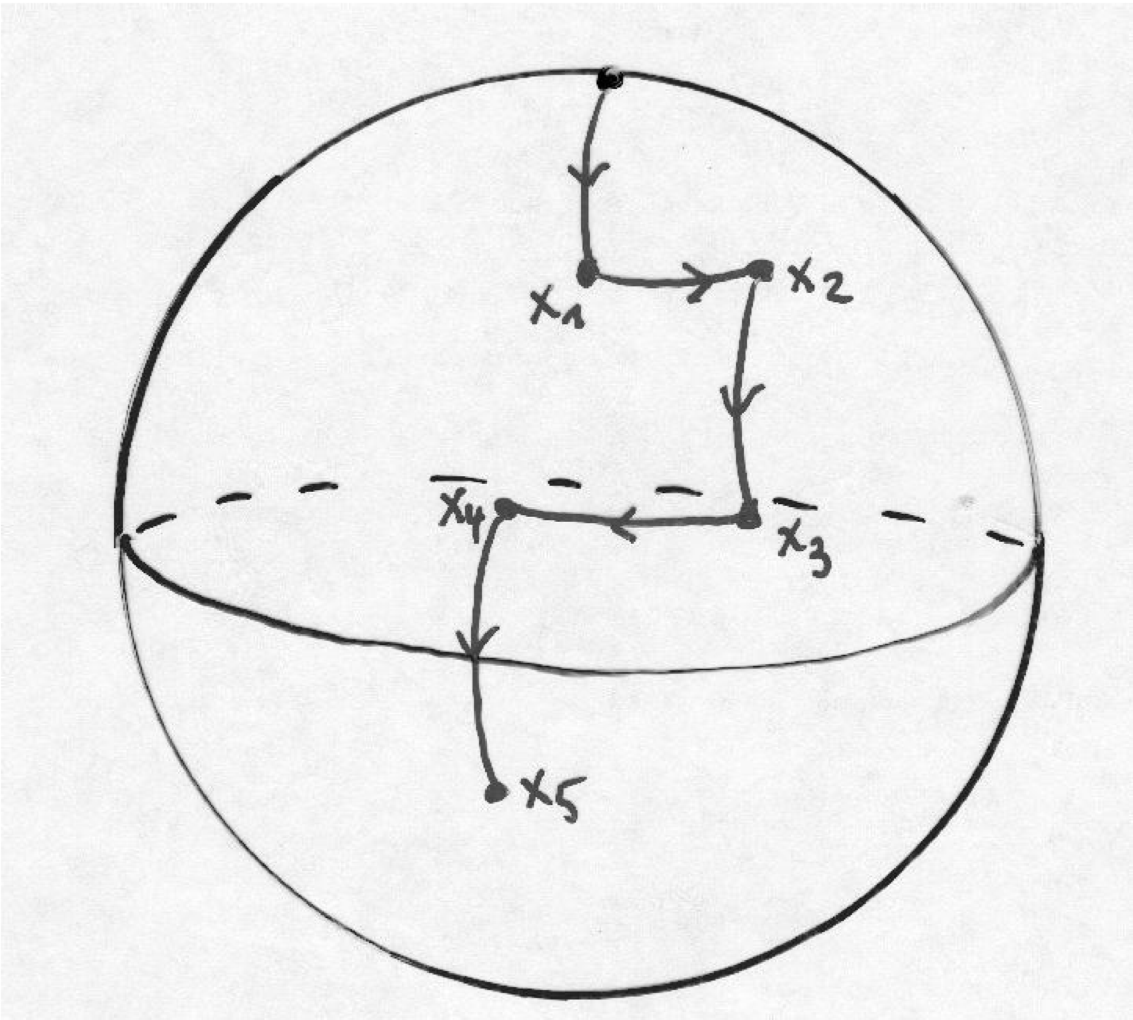


Figure 8

Deriving general convergence results for discrete time gradient recursions is difficult. For simplicity: $M = \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$

Assumption A: $f : \mathbb{R}^n \to \mathbb{R}$ is real analytic, $f \geq 0$, $f^{-1}(0) \neq \emptyset$.

Assumption B: Let $Q_k = Q_k^\top > 0$, $k \in \mathbb{N}$, be positive definite $n \times n$ matrices, such that the condition numbers $c(Q_k)$ are uniformly bounded from above.

Assumption C: The gradient descent algorithm

$$\boxed{x_{k+1} = x_k - \alpha_k Q_k^{-1} \nabla f(x_k)}, \ \alpha_k > 0$$

satisfies the two conditions

(i) $f(x_{k+1}) - f(x_k) \leq \rho \, \mathrm{D}\, f(x_k)(x_{k+1} - x_k) \quad \forall k$ for some $0 < \rho < 1$.

(ii) $\sigma \, \mathrm{D}\, f(x_k)(x_{k+1} - x_k) \leq \mathrm{D}\, f(x_{k+1})(x_{k+1} - x_k) \quad \forall k$ and $\rho < \sigma < 1$.

Theorem (R.Mahony/1999)
Under assumptions A–C, the gradient algorithm

$$x_{k+1} = x_k - \alpha_k Q_k^{-1} \nabla f(x_k)$$

converges, starting from any initial condition in a neighborhood of $f^{-1}(0)$*, to a single critical point* $x_* \in f^{-1}(0)$.

Extension to Riemannian Manifolds: C. Lageman, Masters Thesis, 2002, University of Würzburg

Example: Brockett Function

The gradient flow for the trace function

$$\begin{aligned} f : SO(n, \mathbb{R}) &\rightarrow \mathbb{R} \\ f(\Theta) &= \|N - \Theta X \Theta^\top\|^2 \end{aligned}$$

is

$$\dot{\Theta} = \Theta[N, \Theta^\top X \Theta].$$

The solutions converge to an orthonormal eigenbasis of X.

§3 Optimization of NMR Spin Systems

Single spin $\frac{1}{2}$ system

Hilbert space: $\mathbb{C}^2$

N coupled spin $\frac{1}{2}$ system

Hilbert space: $\mathbb{C}^{2^N} = \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2$

Hilbert spaces are **finite-dimensional**!

Evolution of N coupled spin $\frac{1}{2}$:

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

Schrödinger equation on $SU(2^N)$

Drift term: H_d (Hermitian matrix)

Control Hamiltonians: $H_1, \ldots, H_{2N}$ (Hermitian)

Controls: $u_1(t), \ldots, u_{2N}(t)$

Explicit Forms

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\sigma_{kx} = \mathbb{1} \otimes \cdots \otimes \mathbb{1} \otimes \boldsymbol{\sigma} \otimes \mathbb{1} \otimes \cdots \otimes \mathbb{1}$$

Weak Coupling Terms:

$$\sigma_{kz}\sigma_{lz} = \mathbb{1} \otimes \cdots \otimes \mathbb{1} \otimes \boldsymbol{\sigma_z} \otimes \mathbb{1} \otimes \cdots \otimes \mathbb{1} \otimes \boldsymbol{\sigma_z} \otimes \mathbb{1} \otimes \cdots \otimes \mathbb{1}$$

Drift term H_d: Linear combination of $\sigma_{kz}\sigma_{lz}$

Control Hamiltonians:

$$H_j = \sigma_{kx} \text{ or } H_j = \sigma_{ky} \qquad k = 1, \ldots, 2N$$

Time-optimal control problem:

Given **initial state** $X_0 = I$
and **final state** $X_{max} \in SU(2^N)$.
Find **controls** $u_1(t), \ldots, u_{2N}(t)$
and **smallest time** $T > 0$ such that for solution of Schrödinger equation

$$X(0) = X_0, X(T) = X_{max}$$

Need to solve at least 3 problems:

Subproblem A: Controllability

Theorem (Ph.D.thesis Schulte-Herbrüggen, 1998)
The Pauli operators $\sigma_{kx}, \sigma_{ky}, k = 1, \ldots, N$, together with weak coupling terms

$$\{\sigma_{kz}\sigma_{lz} \mid 1 \leq k < l \leq N\}$$

generate the Lie algebra $su(2^N)$.
The Schrödinger equation is controllable on the Lie group $SU(2^N)$.

Subproblem B: Find (time-opt.) controls!

Basically open problem:

Given initial state X_0 and final state X_1 **find controls** $u_1, \ldots u_{2N}$ **that steer from** X_0 to X, **in minimal finite time** $T > 0$.

Preliminary work: N.Khaneja, R.Brockett, S.Glaser: 'Time optimal control in spin systems', Phys. Review A, Vol. 63 (2001)

Subproblem C: Find final state X_{max}!

$\to$ Solve optimization problem!

$I_{N-1}S$ system:

Given:

$$C := \sigma_- \otimes \mathbb{1} \otimes \cdots \otimes \mathbb{1}$$

$$A := \mathbb{1} \otimes (\sigma_- \otimes \mathbb{1} \otimes \cdots \otimes \mathbb{1} + \cdots + \mathbb{1} \otimes \cdots \otimes \mathbb{1} \otimes \sigma_-)$$

$$\sigma_- = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

find unitary matrix $U_{opt} \equiv X_{max}$ that maximizes

$$f : U(2N) \to \mathbb{R}$$

$$f(U) = \mathrm{Re}\,\mathrm{tr}(CUAU^*)$$

NMR Matrices

For $n \in \mathbb{N}$ consider the recursively defined nilpotent $(2^{n+1} \times 2^{n+1})$-matrices

$$C_n := \begin{bmatrix} 0 & 0 \\ I_{2^n} & 0 \end{bmatrix}, \qquad C_0 := \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},$$

$$A_n := \begin{bmatrix} N_n & 0 \\ 0 & N_n \end{bmatrix}, \tag{1}$$

$$N_n := \begin{bmatrix} N_{n-1} & 0 \\ I_{2^{n-1}} & N_{n-1} \end{bmatrix}, \quad N_0 := 0.$$

Theorem
For $C = \begin{bmatrix} 0 & 0 \\ I & 0 \end{bmatrix}$ and A arbitrary the C-numerical range $W(C, A)$ is a **circular disc around the origin**.

NMR Matrices ($n = 1, 2, 3$)

$$C_1 = \left[\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right], \quad A_1 = \left[\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right],$$

$$C_2 = \left[\begin{array}{c|c} 0 & 0 \\ \hline I_4 & 0 \end{array}\right], \quad A_2 = \left[\begin{array}{c|c} \begin{matrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{matrix} & 0 \\ \hline 0 & \begin{matrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{matrix} \end{array}\right],$$

$$C_3 = \left[\begin{array}{c|c} 0 & 0 \\ \hline I_8 & 0 \end{array}\right],$$

$$A_3 = \left[\begin{array}{c|c|c|c} \begin{matrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{matrix} & 0 & 0 & 0 \\ \hline I_4 & \begin{matrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{matrix} & 0 & 0 \\ \hline 0 & 0 & \begin{matrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{matrix} & 0 \\ \hline 0 & 0 & I_4 & \begin{matrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{matrix} \end{array}\right]$$

Gradient Flows and Critical Points

Theorem
The gradient flow of $f : U(n,\mathbb{C}) \to \mathbb{R}$, $f(U) = \operatorname{Re}\operatorname{tr}(C^*UAU^*)$ with respect to the bi-invariant Riemannian metric is given from the gradient

$$\operatorname{grad} f(U) = 2[C^*, UAU^*]_- U \tag{2}$$

$$\dot{U} = 2(C^*UA + CUA^* - UAU^*C^*U - UA^*U^*CU). \tag{3}$$

Every solution of (3) exists in $U(n,\mathbb{C})$ for all $t \in \mathbb{R}$ and converges for $t \to \pm\infty$ to a critical point. The critical points of f are characterized as

$$[C^*, UAU^*] = [C^*, UAU^*]^*. \tag{4}$$

Theorem
U is a critical point for f if and only if

$$B = UAU^* = \begin{bmatrix} B_1 & B_2 \\ H & B_1 \end{bmatrix} \tag{5}$$

with $H = H^*$ Hermitian.

Numerical Optimization Algorithm

Theorem

Let

$$\alpha = \frac{1}{\sqrt{n2^{n+1}}}$$

The algorithm

$$U_{k+1} = \mathrm{e}^{\alpha [C^*, U_k A U_k^*]_-} U_k$$

converges to the set of critical points of

$$\mathrm{Re}(\mathrm{tr}(C_n^* U A_n U^*)).$$

Lemma

For A_n, C_n and for all $U \in U(n, \mathbb{C})$ it holds

$$\mathrm{Re}(\mathrm{tr}(C_n^* U A_n U^*)) \leq 2(n-m)\binom{n}{m}$$

where $m = n/2$ if n is even and $m = (n-1)/2$ if n is odd.

Numerical Optimization Algorithm

For $n = 1, 2$ the upper bounds of the Lemma are sharp.

Conjecture
For A_n, C_n the gradient flow of the cost function $\text{Re}(\text{tr}(C_n^* U A_n U^*))$ converges to the following maximal values

n	1	2	3	4	5
max	2	4	$4(1+\sqrt{3})$	$8(1+\sqrt{3})$	$16(1+\sqrt{3})+4\sqrt{5}$

§4 NMR and Representations

Let $\mathfrak{g}$ be a finite dimensional Lie algebra. A linear map

$$\rho : \mathfrak{g} \to \mathbb{R}^{N \times N}$$

is called a **representation** if for all $x, y \in \mathfrak{g}$:

$$\rho([x, y]) = [\rho(x), \rho(y)] := \rho(x)\rho(y) - \rho(y)\rho(x).$$

Two representations $\rho_i : \mathfrak{g} \to \mathbb{R}^{N \times N}$, $i = 1, 2$, are **isomorphic** if there exists $S \in GL_N(\mathbb{R})$ with

$$\rho_2(x) = S\rho_1(x)S^{-1}, \quad \forall \quad x \in \mathfrak{g}.$$

(**unitarily equivalent** if S can be chosen unitary)
A representation is called **irreducible**, if it is not isomorphic to a direct sum of nontrivial representations.

Irreducible Representations of $\mathfrak{sl}_2(\mathbb{C})$

Let

$$\mathfrak{sl}_2(\mathbb{C}) = \{X \in \mathbb{C}^{2\times 2} | \operatorname{tr} X = 0\}$$

with basis

$$e = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad f = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad h = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Define

$$E := \begin{bmatrix} 0 & n & 0 & \cdots & 0 \\ 0 & 0 & n-1 & \cdots & 0 \\ \vdots & & \ddots & & \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}, \quad E := \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 2 & \cdots & 0 & 0 \\ \vdots & & \ddots & & \\ 0 & 0 & \cdots & n & 0 \end{bmatrix}$$

$$H := \operatorname{diag}(n, n-2, \ldots, -n).$$

Then ($n := 2l$)

$$\rho_l : \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{sl}_{2l+1}(\mathbb{C}), \quad l = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$$

$$\begin{aligned} e &\mapsto E \\ f &\mapsto F \\ h &\mapsto H \end{aligned}$$

defines an irreducible representation.

Tensor Products

The **tensor product** of two representations

$$\rho_i : \mathfrak{g} \to \mathbb{R}^{N_i \times N_i},\ i = 1, 2,$$

is the representation

$$\rho_1 \otimes \rho_2 : \mathfrak{g} \to \mathbb{R}^{N_1 N_2 \times N_1 N_2} \cong \mathrm{End}(\mathbb{R}^{N_1} \otimes \mathbb{R}^{N_2})$$

defined by

$$(\rho_1 \otimes \rho_2)(x)(v \otimes w) = \rho_1(x)v \otimes w + v \otimes \rho_2(x)w.$$

Thus, using **Kronecker Products Notation**

$$(\rho_1 \otimes \rho_2)(x) = \rho_1(x) \otimes I + I \otimes \rho_2(x).$$

Clebsch-Gordan Formula: For the irreducible representations ρ_l, ρ_m of $\mathfrak{sl}_2(\mathbb{C})$ one has

$$\rho_l \otimes \rho_m \cong \rho_{l+m} \oplus \rho_{l+m-1} \oplus \cdots \oplus \rho_{|l-m|}.$$

Example for n-fold Tensor Product

$$\rho_{\frac{1}{2}} \otimes \cdots \otimes \rho_{\frac{1}{2}} = \bigoplus_{\nu=0}^{n/2} c_\nu^{(n)} \rho_\nu$$

Clebsch-Gordan decomposition. The coefficients can be explicitly computed:

$$p_n(t) := \sum_{\nu=0}^{n/2} c_\nu^{(n)} t^\nu \in \mathbb{Z}[t].$$

Then

$$p_n(t) = e_1^\top A(t)^{n-1}(te_1 + e_2), \quad n \in \mathbb{N}$$

with

$$A(t) := \begin{bmatrix} t & 1 & & 0 \\ 1 & 0 & \ddots & \\ & \ddots & \ddots & 1 \\ 0 & & 1 & 0 \end{bmatrix}$$

Least Squares Matching of Representations

Fix **generators** $x_1, \ldots, x_r$ of a Lie algebra $\mathfrak{g}$ and let

$$\langle A, B \rangle := \mathrm{tr}(AB^*), \quad \|A\|_F^2 := \mathrm{tr}(AA^*)$$

be the **Frobenius Norm**. For representations

$$\rho_1, \rho_2 : \mathfrak{g} \to \mathbb{C}^{N \times N}$$

define

$$\|\rho_1 - \rho_2\|^2 := \sum_{i=1}^{r} \|\rho_1(x_i) - \rho_2(x_i)\|_F^2$$

.

Unitary least squares matching: Find a unitary matrix $U_{opt} \in \mathbb{C}^{N \times N}$ with

$$U_{opt} := \mathrm{argmin}\, \|\rho_1 - U\rho_2 U^*\|^2$$

Example 1 (Abelian Case)

Consider two r-tuples of commuting $N\times N$-matrices

$$(A_1,\dots A_r),\quad [A_i,A_j]=0\quad \forall i,j.$$
$$(B_1,\dots B_r),\quad [B_i,B_j]=0\quad \forall i,j.$$

This defines representations of the **abelian Lie algebra** $\mathfrak{g}=\mathbb{C}^r$ via

$$\rho_1:\mathbb{C}^r\to\mathbb{C}^{N\times N},\quad e_i\mapsto A_i,\quad i=1,\dots,r.$$
$$\rho_2:\mathbb{C}^r\to\mathbb{C}^{N\times N},\quad e_i\mapsto B_i,\quad i=1,\dots,r.$$

We see:

Two r-tuples $(A_1,\dots,A_r)$ and $(B_1,\dots,B_r)$ of commuting matrices are **simultaneously unitary equivalent**, i.e.,

$$(B_1,\dots,B_r)=(UA_1U^*,\dots,UA_rU^*),$$

if and only if ρ_2 is **unitary equivalent** to ρ_1.

Example 2 ($\mathsf{sl}(2,\mathbb{C})$**)**

Jacobson-Morosov Lemma: For any nilpotent matrix $A \in \mathsf{sl}(N,\mathbb{C})$ there exists a representation $\rho : \mathsf{sl}(2,\mathbb{C}) \to \mathsf{sl}(N,\mathbb{C})$ with

$$\rho\left(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right) = A.$$

Kostant's Theorem: For representations $\rho_1, \rho_2 :$ $\mathsf{sl}(2,\mathbb{C}) \to \mathsf{sl}(N,\mathbb{C})$ with

$$\rho_1\left(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right) = \rho_2\left(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right)$$

there exists a $S \in \mathbb{C}^{N\times N}$, $\det S = 1$ with

$$\rho_2(x) = S\rho_1(x)S^{-1} \quad \forall x \in \mathsf{sl}(2,\mathbb{C}).$$

Corollary: Two nilpotent matrices $A, B \in \mathbb{C}^{N\times N}$ are similar if and only if the associated representations are isomorphic.

Example 3 ($\mathfrak{su}(2,\mathbb{C})$)

Fact 1: Every representation $\rho : \mathfrak{su}(2,\mathbb{C}) \to \mathfrak{su}(N,\mathbb{C})$ is unitarily equivalent to a direct sum of irreducible representations, i.e.,

$$\rho \;\cong\; c_1\rho_{\nu_1} \oplus \cdots \oplus c_r\rho_{\nu_r},$$

$$\rho_l \quad : \quad \mathfrak{su}(2,\mathbb{C}) \to \mathfrak{su}(2l+1,\mathbb{C})$$

standard irreducible representations.

Fact 2: The **stabilizer in** $U(N,\mathbb{C})$ of ρ has dimension

$$c_1^2 + \cdots + c_r^2.$$

In particular, the **dimension of the unitary orbit** of ρ is

$$N^2 - c_1^2 - \cdots - c_r^2.$$

Representations and NMR-Optimization)

After a simple equivalent reformulation we arrive at

$$C_n = \rho \otimes \mathbb{1}^n$$

$$A_n = \mathbb{1} \otimes (\rho \otimes \mathbb{1}^{n-1} + \cdots + \mathbb{1}^{n-1} \otimes \rho)$$

with

$$\mathbb{1} = I_2, \rho = \rho_{1/2}\left(\left[\begin{array}{cc} 0 & -1 \\ 1 & 0 \end{array}\right]\right),$$

$$\rho_{1/2} = \mathfrak{su}(2, \mathbb{C}) \to \mathfrak{su}(2, \mathbb{C})$$

standard irreducible representation. Thus for

$$\tau, \sigma : \mathfrak{su}(2, \mathbb{C}) \to \mathfrak{su}(2^{n+1}, \mathbb{C})$$

$$\sigma := \rho_{1/2} \otimes (0 \otimes \cdots \otimes 0)$$

$$\tau := 0 \otimes (\rho_{1/2} \otimes \cdots \otimes \rho_{1/2})$$

we have

$$C_n = \sigma\left(\left[\begin{array}{cc} 0 & -1 \\ 1 & 0 \end{array}\right]\right), A_n = \tau\left(\left[\begin{array}{cc} 0 & -1 \\ 1 & 0 \end{array}\right]\right).$$

Remark

$$
\begin{aligned}
(\rho_{1/2} \otimes 0)(x)(v \times w) &= \rho_{1/2}(x)v \otimes w + v \otimes 0(x)w \\
&= \rho_{1/2}(x)v \otimes w \\
&= (\rho_{1/2}(x) \otimes \mathbb{1})(v \otimes w)
\end{aligned}
$$

More generally,

$$(\rho_{1/2} \otimes 0 \otimes \cdots \otimes 0)(x)(v, w_1, \ldots, w_n)$$

$$= \rho_{1/2}(x)(v) \otimes (w_1 \otimes \cdots \otimes w_n)$$

$$= (\rho_{1/2}(x) \otimes \mathbb{1} \otimes \cdots \otimes \mathbb{1})(v, w_1, \ldots, w_n).$$

Clebsch-Gordan Formula

1. Unitary equivalence ('C_n−Matrix')

$$\rho = \rho_{1/2} \otimes (0 \otimes \cdots \otimes 0)$$

$$\cong \rho_{1/2} \otimes (2\rho_0 \otimes \cdots \otimes 2\rho_0)$$

$$= 2^n \rho_{1/2}$$

2. Unitary equivalence ('A-Matrix')

$$\tau = 0 \otimes (\rho_{1/2} \otimes \cdots \otimes \rho_{1/2})$$

$$\cong 2c_0\rho_0 \oplus 2c_1\rho_1 \oplus \cdots \oplus 2c_{\frac{n}{2}}\rho_{\frac{n}{2}}$$

where c_i are Clebsch-Gordan coefficients of

$$\rho_{1/2} \otimes \cdots \otimes \rho_{1/2} \cong \bigoplus_{j=0}^{n/2} c_j \rho_j .$$

In the **Fermionic case** ($n = odd$) we obtain:

Theorem 1
There exists a $U_n \in U(2^{n+1}, \mathbb{C})$ such that

$$\mathrm{tr}\left(C_n^* U_n A_n U_n^*\right) =$$

$$2 \sum_{k=0}^{\frac{n-1}{2}} c_{\frac{2k+1}{2}} \sum_{j=0}^{k} \sqrt{(2j+1)(2k+1-2j)}.$$

Here c_i are the Clebsch-Gordan coefficients of $\rho_{1/2} \otimes \cdots \otimes \rho_{1/2}$.

Theorem 2(Doubling Argument) Let $n \in \mathbb{N}$ be odd. and $U_n \in U(2^{n+1}, \mathbb{C})$ chosen as in Theorem 1. Then there exists a unitary transformation $U_{n+1} \in U(2^{n+2}, \mathbb{C})$ such that

$$\mathrm{tr}(C_{n+1} U_{n+1} A_{n+1} U_{n+1}^*) = 2\,\mathrm{tr}(C_n U_n A_n U_n^*)$$

The following values for the cost function are assumed:

n	1	2	3	4	5
max	2	4	$4(1+\sqrt{3})$	$8(1+\sqrt{3})$	$16(1+\sqrt{3})+4\sqrt{5}$