Lie Group Methods

Geometric Numerical Integration

Seminar WS 05/06

Contents

Lie Group Methods

Consider a differential equation

$$\dot{Y} = A(Y)Y, \qquad Y(0) = Y_0$$

on a matrix Lie Group G :

$$Y_0 \in G, \qquad A(Y) \in \mathcal{G} \qquad \forall \mathrm{Y} \in \mathrm{G}.$$

Since this is a special case of **differential equations on manifold**,

- **projection methods** as well as
- **methods based on local coordinates**

are well suited for their numerical treatment.

Now we study further approaches which also yield approximations that lie on the manifold.

Crouch-Grossman Methods

The numerical approximation of **explicit Runge-Kutta methods** is obtained by a composition of the following two basic operations:

- an "evaluation of the vector field $f(Y) = A(Y)Y$" and
- a "computation of an update of the form $Y + haf(Z)$".

In the context of differential equations on Lie groups, these methods have the **disadvantage** that,

even when Y and $Z \in G$, the update $Y + haA(Z)Z$ is in general $\notin G$.

The idea of Crouch-Grossman is to replace the "update" operation with $exp(haA(Z))Y$:

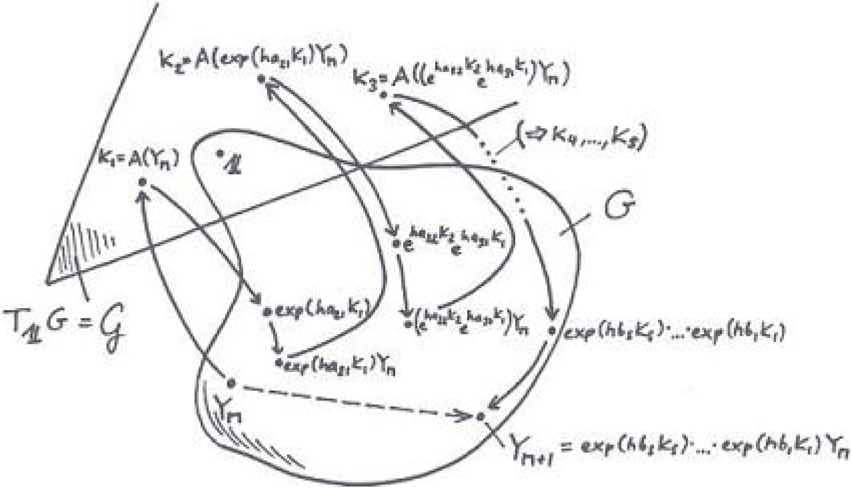
Definition (explicit s-stage Crouch-Grossman Method (1993))

Let b_i, a_{ij} $(i, j = 1, ..., s)$ be real numbers.
Then, the step $Y_n \mapsto Y_{n+1}$ is defined as follows:

$$Y^{(i)} = exp(ha_{i,i-1}K_{i-1}) \cdot ... \cdot exp(ha_{i1}K_1)Y_n, \qquad K_i = A(Y^{(i)}),$$

$$Y_{n+1} = exp(hb_sK_s) \cdot ... \cdot exp(hb_1K_1)Y_n.$$

By construction, the methods of Crouch-Grossman give rise to approximations Y_n which **lie exactly on the manifold** defined by the Lie group.


K_2 = A(exp(ha_{21}K_1)Y_n)
K_3 = A((e^{ha_{32}K_2} e^{ha_{31}K_1})Y_n)
(⇒ K_4, ..., K_s)
K_1 = A(Y_n)
𝟙
G
e^{ha_{32}K_2} e^{ha_{31}K_1}
exp(ha_{21}K_1)
(e^{ha_{32}K_2} e^{ha_{31}K_1})Y_n
exp(hb_sK_s)·...·exp(hb_1K_1)
exp(ha_{21}K_1)Y_n
T_𝟙G = 𝔤
Y_n
Y_{n+1} = exp(hb_sK_s)·...·exp(hb_1K_1)Y_n

Theorem

Let $c_i = \sum_j a_{ij}$. A Crouch-Grossman method has **order** p *($p \leq 3$) if the following* **order conditions** *are satisfied:*

- *order 1:* $\sum_i b_i = 1$
- *order 2:* $\sum_i b_i c_i = 1/2$
- *order 3:* $\sum_i b_i c_i^2 = 1/3$
  $\sum_{ij} b_i a_{ij} c_j = 1/6$
  $\sum_i b_i^2 c_i + 2 \sum_{i<j} b_i c_i b_j = 1/3.$

Proof:
The order conditions can be found by comparing the Taylor series expansions of the exact and the numerical solution.

$$Y(0) = Y_n,$$
$$Y(h) = Y_n + h \cdot A(Y_n) Y_n + O(h^2),$$
$$Y_{n+1} = Y_n + h \cdot (\textstyle\sum_i b_i) A(Y_n) Y_n + O(h^2).$$

The theory of order conditions for Runge-Kutta methods has been extended to Crouch-Grossman methods.

It turns out that the order conditions for classical Runge-Kutta methods form a subset of those for Crouch-Grossman methods.

Example (Two Crouch-Grossman methods of order 3)

0			
-1/24	-1/24		
17/24	161/24	-6	
	1	-2/3	2/3

0			
3/4	3/4		
17/24	119/216	17/108	
	13/51	-2/3	24/17

Munthe-Kaas Methods

Idea:
Write the solution as $Y(t) = exp(\Omega(t))Y_0$
and solve numerically the differential equation for $\Omega(t)$.

We replace the differential equation $\dot{Y} =^{(*)} A(Y)Y$ by a more complicated one.
However, the nonlinear invariants $g(Y) = 0$ of $(*)$ defining the Lie group are replaced with linear invariants $g'(I)(\Omega) = 0$ defining the Lie algebra.
We know that essentially all numerical methods (for example all explicit and implicit Runge-Kutta methods) automatically conserve linear invariants.

Now we need two Lemmata from section III.4:

Lemma

The derivative of $\exp\Omega = \sum_{k\geq 0} \frac{1}{k!}\Omega^k$ *is given by*

$$\left(\frac{d}{d\Omega}\exp\Omega\right)H = (dexp_{\Omega}(H))\exp\Omega,$$

where

$$dexp_{\Omega}(H) = \sum_{k\geq 0} \frac{1}{(k+1)!} ad_{\Omega}^k(H),$$

and $ad_{\Omega}(A) = [\Omega, A] = \Omega A - A\Omega$.
(Convention: $ad_{\Omega}^0(A) = A$*)*

The series $dexp_{\Omega}(H)$ *converges for all matrices* Ω.

Lemma (Baker (1905))

If the eigenvalues of the linear operator ad_Ω *are different from* $2l\pi i$ *with* $l \in \{\pm 1, \pm 2, \ldots\}$, *then* $dexp_\Omega$ *is invertible.*
Furthermore, we have for $\|\Omega\| < \pi$ *that*

$$dexp_\Omega^{-1}(H) = \sum_{k\geq 0} \frac{B_k}{k!} ad_\Omega^k(H),$$

where B_k *are the Bernoulli numbers, defined by* $\sum_{k\geq 0}(B_k/k!)x^k = x/(e^x - 1)$.

$$B_0 = 1,\ B_1 = -\frac{1}{2},\ B_2 = \frac{1}{6},\ \text{for odd } k > 1:\ B_k = 0,\ B_4 = -\frac{1}{30}, \ldots$$

The following Theorem from section IV.7 is important:

Theorem (Magnus (1954))

The solution of the differential equation $\dot{Y} = A(t)Y$ (apart from continuous dependence on t, no assumption on the matrix $A(t)$ is made) can be written as $Y(t) = exp(\Omega(t))Y_0$ with $\Omega(t)$ defined by

$$(**) \quad \dot{\Omega} = dexp_{\Omega}^{-1}(A(t)), \qquad \Omega(0) = 0.$$

Proof:
Comparing the derivative of $Y(t) = exp(\Omega(t))Y_0$,

$$\dot{Y}(t) = (\frac{d}{d\Omega} exp\Omega(t))\dot{\Omega}(t)Y_0 = (dexp_{\Omega(t)}(\dot{\Omega}(t)))exp(\Omega(t))Y_0,$$

with the differential equation we obtain $A(t) = dexp_{\Omega(t)}(\dot{\Omega}(t))$. Applying the inverse operator $dexp_{\Omega}^{-1}$ to this relation yields the differential equation $(**)$ for $\Omega(t)$.

Now we apply these results to our situation:
$\dot{Y} = A(Y)Y, \; Y(0) = Y_0$ differential equation on the matrix Lie group G.

Idea: Write the solution as $Y(t) = exp(\Omega(t))Y_0$ and solve numerically the differential equation for $\Omega(t)$.

We write $Y(t) = exp(\Omega(t))Y_0$, **where** $\Omega(t)$ **is the solution of** $\dot{\Omega} = dexp_{\Omega}^{-1}(A(Y(t))), \quad \Omega(0) = 0.$

Since it is not practical to work with the operator $dexp_{\Omega}^{-1}$, we truncate the series

$$dexp_{\Omega}^{-1}(H) = \sum_{k \geq 0} \frac{B_k}{k!} ad_{\Omega}^{k}(H),$$

suitably and consider the differential equation:

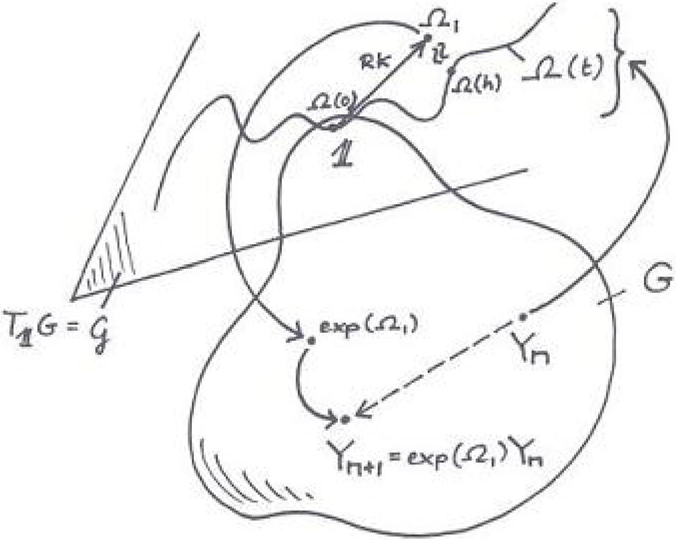
$$(***) \ \dot{\Omega} = A(exp(\Omega)Y_0) + \sum_{k=1}^{q} \frac{B_k}{k!} ad_{\Omega}^{k}(A(exp(\Omega)Y_0)), \quad \Omega(0) = 0.$$

This leads to the following method:

Definition (Munthe-Kaas Method (1999))

The step $Y_n \mapsto Y_{n+1}$ is defined as follows:

- Consider $(***)$ with Y_n instead of Y_0 and apply a Runge-Kutta method (explicit or implicit) to get an approximation $\Omega_1 \approx \Omega(h)$.
- Then define the numerical solution by $Y_{n+1} = exp(\Omega_1)Y_n$.



Ω_1
RK
$\Omega(h)$
$\Omega(t)$
$\Omega(0)$
$\mathbb{1}$
$T_{\mathbb{1}}G = \mathfrak{g}$
G
$\exp(\Omega_1)$
Y_n
$Y_{n+1} = \exp(\Omega_1) Y_n$

Theorem

The numerical solution of the Munthe-Kaas method lies in G, i.e., $Y_n \in G \ \forall n = 0, 1, 2, \ldots$.

Proof:

It is sufficient to prove that for $Y_0 \in G$ the numerical solution Ω_1 of the Runge-Kutta method applied to $(***)$ lies in $\mathcal{G}$.

Since the Lie bracket $[\Omega, A]$ is an operation $\mathcal{G} \times \mathcal{G} \to \mathcal{G}$, and since $exp(\Omega)Y_0 \in G$ for $\Omega \in \mathcal{G}$, the right-hand expression of $(***)$ is in $\mathcal{G}$ for $\Omega \in \mathcal{G}$.

Hence, $(***)$ **is a differential equation on the vector space** $\mathcal{G}$ **with solution** $\Omega(t) \in \mathcal{G}$.

All operations in a Runge-Kutta method give results in $\mathcal{G}$, so that the numerical approximation Ω_1 also lies in $\mathcal{G}$.

Theorem

If the Runge-Kutta method is of order p and if the truncation index in $(***)$ satisfies $q \geq p-2$, then the Munthe-Kaas method is of order p.

Proof:

For sufficiently smooth $A(Y)$ we have (Taylor):

$$\Omega(t) = \underbrace{\Omega(0)}_{=0} + \underbrace{\dot{\Omega}(0)}_{=A(Y_0)+\underbrace{\sum_{k=1}^{q} \frac{B_k}{k!} ad^k_{\Omega(0)}(A(Y_0))}_{=0}} \cdot t + O(t^2) = tA(Y_0) + O(t^2)$$

$$Y(t) = Y(0) + t\dot{Y}(0) + O(t^2) = Y_0 + O(t)$$

$$[\Omega(t), A(Y(t))] = \underbrace{\underbrace{\Omega(t)}_{=tA(Y_0)+O(t^2)} \cdot A(\overbrace{\overbrace{Y(t)}^{=Y_0+O(t)})}^{=A(Y_0)+O(t)}}_{=tA(Y_0)A(Y_0)+O(t^2)}$$

$$- \underbrace{\overbrace{A(\overbrace{Y(t)}^{=Y_0+O(t)})}^{=A(Y_0)+O(t)} \cdot \underbrace{\Omega(t)}_{=tA(Y_0)+O(t^2)}}_{=A(Y_0)tA(Y_0)+O(t^2)} = O(t^2)$$

This implies that $ad^k_{\Omega(t)}(A(Y(t))) = O(t^{k+1})$, so that the truncation of the series in $(***)$ induces an error of size $O(h^{q+2})$ for $|t| \leq h$:

$$\sum_{k\geq 0} \frac{B_k}{k!} ad^k_{\Omega}(A(exp(\Omega)Y_0)) - \sum_{k=0}^{q} \frac{B_k}{k!} ad^k_{\Omega}(A(exp(\Omega)Y_0)) = O(h^{q+2})$$

Since we take a Runge-Kutta method (for $(***)$) of order p, we get:

$$\Omega_1 = \underbrace{\Omega(0)}_{=0} + h\sum_i b_i k_i, \quad k_i = f(\underbrace{\underbrace{\Omega(0)}_{=0} + h\sum_j a_{ij}k_j}_{=O(h)}),$$

where $f(\Omega) = \sum_{k=0}^{q} \frac{B_k}{k!} ad^k_{\Omega}(A(exp(\Omega)Y_0))$.

$$f^{\infty}(\Omega) = \sum_{k\geq 0} \frac{B_k}{k!} ad_{\Omega}^{k}(A(exp(\Omega)Y_0))$$

$$k_i^{\infty} = f^{\infty}(\underbrace{\underbrace{\Omega(0)}_{=0} + h\sum_j a_{ij}k_j^{\infty}}_{=O(h)})$$

$$\Omega_1 = h\textstyle\sum_i b_i(k_i^{\infty} + O(h^{q+2})) = \displaystyle\underbrace{h\sum_i b_i k_i^{\infty}}_{=\Omega^{exact}(h)+O(h^{p+1})} + O(h^{q+3}) =$$

$$\Omega^{exact}(h) + O(h^{p+1}) + O(h^{q+3})$$

For $q + 2 \geq p$ we get:
$Y_1 = exp(\Omega_1)Y_0 = exp(\Omega^{exact}(h) + O(h^{p+1}))Y_0 =_{(Taylor)}$
$(exp(\Omega^{exact}(h)) + O(h^{p+1}))Y_0 = Y^{exact}(h) + O(h^{p+1})$.

The most simple Lie group method is obtained if we take

- the explicit Euler method as basic discretization and
- $q = 0$ in $(***)$.

This leads to the so-called **Lie-Euler method**

$$\boxed{Y_{n+1} = exp(hA(Y_n))Y_n}$$

This is also a special case of the Crouch-Grossman methods.

- Taking the implicit midpoint rule ($y_{n+1} = y_n + h(\frac{y_n+y_{n+1}}{2})$) as the basic discretization and
- $q = 0$ in $(***)$,

we obtain the **Lie midpoint rule**

$$\boxed{Y_{n+1} = exp(\Omega)Y_n, \quad \Omega = hA(exp(\Omega/2)Y_n)}$$

(This is an implicit equation in Ω and has to be solved by fixed point iteration or by Newton-type methods.)

Example

Example: Motion of a free rigid body, whose centre of mass is at the origin.

This problem with

- the angular momentum $y = (y_1, y_2, y_3)^T$ in the body frame and
- the principal moments of inertia I_1, I_2, I_3

can be written (Euler equations) as:

$$\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \\ \dot{y}_3 \end{pmatrix} = \begin{pmatrix} 0 & y_3/I_3 & -y_2/I_2 \\ -y_3/I_3 & 0 & y_1/I_1 \\ y_2/I_2 & -y_1/I_1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix},$$

which is of the form $\dot{Y} = A(Y)Y$ with a skew-symmetric matrix $A(Y)$.

There are two invariants:

- $y_1^2 + y_2^2 + y_3^2$
- $\frac{1}{2}(\frac{y_1^2}{I_1} + \frac{y_2^2}{I_2} + \frac{y_3^2}{I_3})$ (kinetic energy)

We take $I_1 = 2,\ I_2 = 1,\ I_3 = 2/3$ and the initial condition: $y_0 = (cos(1.1), 0, sin(1.1))^T$.

As coefficients for the 3rd order Munthe-Kaas and Crouch-Grossman methods we take:

0			
3/4	3/4		
17/24	119/216	17/108	
	13/51	-2/3	24/17

For the computation of the matrix exponential we use the **Rodrigues formula:**

$$exp(\Omega) = I + \frac{sin\alpha}{\alpha}\Omega + \frac{1}{2}(\frac{sin(\alpha/2)}{\alpha/2})^2\Omega^2$$

for $\Omega = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}$ and $\alpha = \sqrt{\omega_1^2 + \omega_2^2 + \omega_3^2}$.

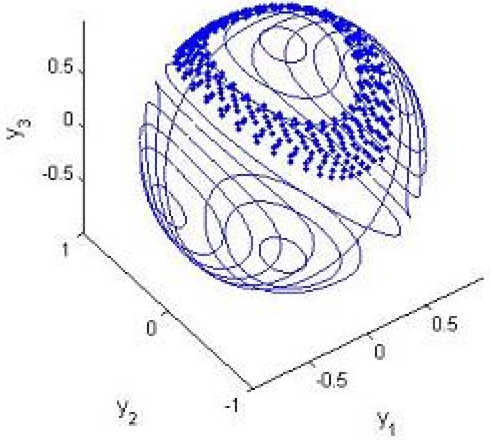


Figure: CG

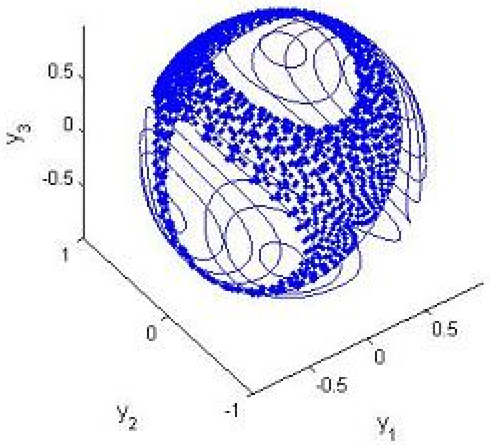


Figure: MKq=0

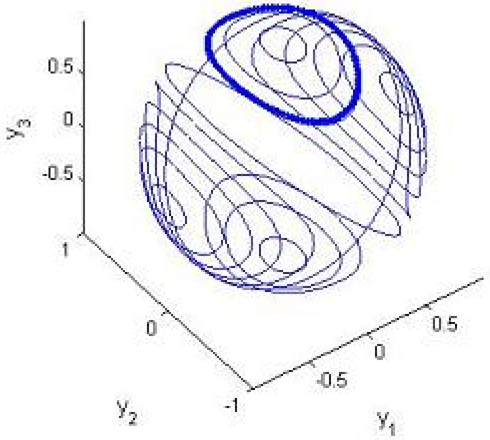


Figure: MKq=1

Now, we want to compare Munthe-Kaas and Crouch-Grossman methods.

If the CG and MK methods are based on the same set of Runge-Kutta coefficients:

	CG	MK
evaluations of $A(Y)$	s	s
computation of matrix exponentials	$s(s+1)/2$	s
computation of commutators	no	yes (if $q \geq 1$)

Every classical Runge-Kutta method defines a Munthe-Kaas method of the same order,
but Crouch-Grossman methods of high order are very difficult to obtain, and need more stages for the same order (if $p \geq 4$).