Groups and Symmetries in Numerical Linear Algebra Part 3

Integration on Lie groups and related linear algebra problems

Hans Munthe-Kaas

[1] Department of Mathematics
University of Bergen
Norway
http://hans.munthe-kaas.no
mailto:hans@math.uib.no

CIME-EMS Cetraro 2015

Lie group integration (LGI): WHAT, WHY and HOW?

What?

LGI generalizes classical ODE integrators (Runge–Kutta, Multistep, ..), based on the commutative action of translations on $\mathbb{R}^n$

$$y_{n+1} = y_n + \Phi(h, F, y_n, \ldots),$$

to methods based on Lie group actions on a manifold $\mathcal{M}$

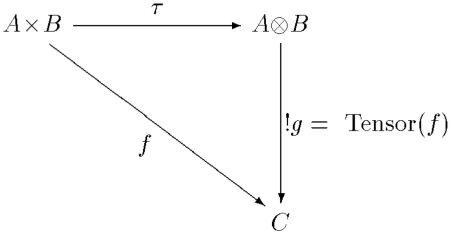
$$y_{n+1} = g(h, f, y_n, \ldots) \cdot y_n, \quad g \in G, y \in \mathcal{M}.$$

LGI - brief history

- P. Crouch & R. Grossman: *Numerical integration of ordinary differential equations on manifolds*, J. Nonlinear Sci. 1993.
- Bergen-Cambridge-Trondheim (Celledoni, Iserles, MK, Nørsett, Owren, Zanna) 1995-...today:
  - MK: *Lie–Butcher theory for Runge–Kutta methods*, BiT 1995.
  - . . .
  - Iserles, MK, Nørsett, Zanna: *Lie group methods*, Acta Numerica 2000.
  - . . .
  - Celledoni, Marthinsen, Owren: *An introduction to Lie group integrators – basics, new developments and applications*, J. Comp. Phys., 2014.
  - . . .

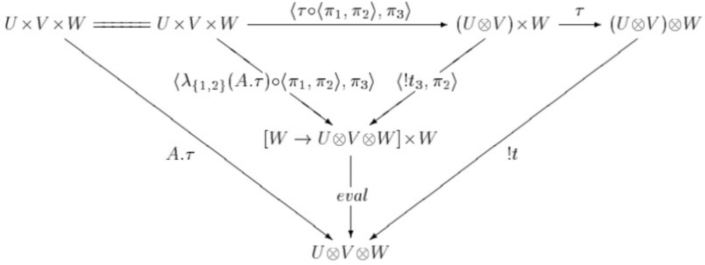
Why? "Coordinate–Free numerics"

- Object orientation:
  Programming is *specification* + *implementation*.
- Can this dichotomy be pursued in computational mathematics/ numerical algorithms?
- Coordinate–Free numerics:
  SOPHUS Library (MK+Haveraaen 1992-1995) pursued this idea for tensor computations.



(Don't read this!)

Example 7 *The ungroup map is the inverse of the above* $t : (U \otimes V) \otimes W \to U \otimes V \otimes W$. *The diagram for this operation is more difficult to draw. It looks like:*



The idea is to work from left to right by completing the arrows indicated by the exclamation marks.

function ungroup(T)
 // input: $T = (U \otimes V) \otimes W$: type information
 // output: $t : (U \otimes V) \otimes W \to U \otimes V \otimes W$: isomorphism
 //
 // from T compute the following type info:
 $A = U \times V \times W$; $B = (U \otimes V) \times W$;
 $t2 = \lambda_{1,2}(A.\tau)$; // : $U \times V \to [W \to U \otimes V \otimes W]$
 $t3 =$ Tensor($t2$); // : $U \otimes V \to [W \to U \otimes V \otimes W]$
 $t4 =$ eval$\circ\langle t3, B.\pi_2 \rangle$ // : $(U \otimes V) \times W \to U \otimes V \otimes W$
 $t =$ Tensor($t4$); // : $(U \otimes V) \otimes W \to U \otimes V \otimes W$
return(t);

Numerical integration (ODEs), abstract view.

Differential equations in 'pure' math:

- Domain: Differential manifold $\mathcal{M}$. Has points (positions $p \in \mathcal{M}$) and tangents (velocities in $v \in T_p\mathcal{M}$ at each base point).
- ODE: $y' = F(y, t)$, where $F\colon \mathcal{M} \to T\mathcal{M}$ specifies the velocity.

Numerical integration needs more information:

- Domain: Differential manifold: position + tangents ($\mathcal{M}$ and $T\mathcal{M}$).
- Computable motions: given by *Lie group action* : $G \times \mathcal{M} \to \mathcal{M}$.
- Rule for transport of tangents from one point to another.

Surprise!

A full theory of numerical integration, including analysis (order theory, structure preservation), and implementation can be developed from these building blocks.

(Matrix) Lie groups. Basic structures.

Examples:

- $\mathrm{GL}(n)$ (general linear)
- $\mathrm{SL}(n)$ (special linear)
- $O(n)$ (orthogonal)
- $\mathrm{SO}(n)$ (special orthogonal)
- $\mathrm{Sp}(n)$ (symplectic)
- $T(n)$ (triangular)
- $E(n)$ (euclidean motion group).
- $\mathrm{Aff}(n)$ (affine group).
- Note: Important Lie groups are not *naturally* described by matrices (some *cannot*).

Basic building blocks:

Tangents, Lie algebra, exponential map and commutators, homogeneous manifolds. (Explained at blackboard.)

Numerical integration

$$\text{ODE:} \quad y'(t) = f(y), \quad y(0) = y_0$$

Analytical: $y(t) = \Phi_{tf}(y_0)$, where $\Phi_{tf}(y_0)$ denotes the exact solution.
Numerical: $y_{n+1} = \Psi_{h,f}(y_n)$, where $\Psi_{h,f}$ is numerical integrator.

- *Order theory:* Find $\Psi_{h,f}$ s.t. $|\Psi_{h,f}(y_0) - \Phi_{tf}(y_0)| = \mathcal{O}(h^{p+1})$.

Example

- Euler method: $y_{n+1} = y_n + h\,f(y_n, t_n)$,
- Runge-Kutta methods

$$\begin{aligned} y_{n+1} &= y_n + h \sum_{i=1}^{\nu} b_i\, f(Y_i, t_n + c_i\, h), \\ Y_i &= y_n + h \sum_{j=1}^{\nu} a_{i,j}\, f(Y_j, t_n + c_j\, h), \quad i = 1, \dots, \nu. \end{aligned}$$

$$A = (a_{i,j})_{i,j=1,\dots,\nu}, \quad \mathbf{b} = [b_1, \dots, b_\nu]^T, \quad \mathbf{c} = [c_1, \dots, c_\nu]^T.$$

Numerical integration

$$\text{ODE:}\quad y'(t) = f(y), \quad y(0) = y_0$$

Analytical: $y(t) = \Phi_{tf}(y_0)$, where $\Phi_{tf}(y_0)$ denotes the exact solution.
Numerical: $y_{n+1} = \Psi_{h,f}(y_n)$, where $\Psi_{h,f}$ is numerical integrator.

- *Structure preservation (Geometric integration):* Given f with certain *geometric properties*, find $\Psi_{h,f}$ exactly preserving this property.
  - Symplectic structure.
  - Energy preservation and other first integrals.
  - Volume preservation.
  - Lyapunov functions.
  - Equations on manifolds.
  - Equivariance and symmetries
  - Fixpoints
  - Correct linearization near fixpoints
  - ...

Numerical integration

$$\text{ODE:}\quad y'(t) = f(y), \quad y(0) = y_0$$

Analytical: $y(t) = \Phi_{tf}(y_0)$, where $\Phi_{tf}(y_0)$ denotes the exact solution.
Numerical: $y_{n+1} = \Psi_{h,f}(y_n)$, where $\Psi_{h,f}$ is numerical integrator.

- Construction of numerical integrators.
  - *Classical methods:* Runge-Kutta, Multistep, General Linear.
  - Exponential integrators.
  - Lie group integrators.
  - Taylor series and modified vector field methods.
  - Methods built from simpler: Composition methods, partitioned methods.

... Lie Group Integrators (How?)

Example: Geometrization of Runge-Kutta methods:

Classic RK solves: $y' = F(y)$, $y, F \in \mathbb{R}^n$

Given $y_n \approx y(t_n)$ and a stepsize h.

We compute $y_{n+1} \approx y(t_n + h)$ as:

for $i = 1, s$

$u_i = \sum_{j=1}^{s} a_{i,j} K_j$

$K_i = h \cdot F(u_i + y_n)$

end

$u = \sum_{j=1}^{s} b_j K_j$

$y_{n+1} = u + y_n$

... Lie Group Integrators (How?)

Example: Geometrization of Runge-Kutta methods:

LGI solves: $y' = f(y)\cdot y$, $y \in \mathcal{M}$, $f : \mathcal{M} \to \mathfrak{g}$

where $\cdot : \mathfrak{g} \times \mathcal{M} \to T\mathcal{M}$ denotes infinitesimal group action

Given $y_n \approx y(t_n)$ and a stepsize h.
We compute $y_{n+1} \approx y(t_n + h)$ as:

for $i = 1, s$

$$u_i = \sum_{j=1}^{s} a_{i,j} K_j$$

$$K_i = h \cdot F(\exp(u_i) \cdot y_n)$$

end

$$u = \sum_{j=1}^{s} b_j K_j$$

$$y_{n+1} = \exp(u) \cdot y_n$$

where $\exp : \mathfrak{g} \to G$, and $\cdot : G \times \mathcal{M} \to \mathcal{M}$.

... Lie Group Integrators (How?)

Example: Geometrization of Runge-Kutta methods:

LGI solves: $y' = f(y) \cdot y$, $y \in \mathcal{M}$, $f : \mathcal{M} \to \mathfrak{g}$

where $\cdot : \mathfrak{g} \times \mathcal{M} \to T\mathcal{M}$ denotes infinitesimal group action

Given $y_n \approx y(t_n)$ and a stepsize h.
We compute $y_{n+1} \approx y(t_n + h)$ as:

for $i = 1, s$

$u_i = \sum_{j=1}^{s} a_{i,j} \tilde{K}_j$

$K_i = h \cdot F(\exp(u_i) \cdot y_n)$

$\tilde{K}_j = d\exp_{u_i}^{-1}(K_j) = K_j - \frac{1}{2}[u_i, K_j] + \frac{1}{12}[u_i, [u_i, K_j]] + \cdots$

end

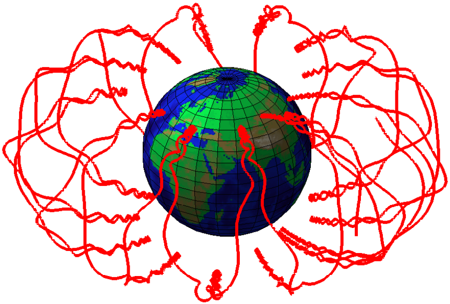
$u = \sum_{j=1}^{s} b_j \tilde{K}_j$

$y_{n+1} = \exp(u) \cdot y_n$

where $\exp : \mathfrak{g} \to G$, and $\cdot : G \times \mathcal{M} \to \mathcal{M}$.

Example: A highly oscillatory problem

Carl Størmer explains Aurorae (1907).



Størmer developed an excellent symplectic integrator.
He 'ran' his scheme on human computers (students), 4500 CPU hours! Three steps per hour ...

Aurora example (cont.)

Equation of motion:

$$\begin{pmatrix} y(t) \\ v(t) \end{pmatrix}' = \begin{pmatrix} v(t) \\ v(t) \times B(y) \end{pmatrix}, \quad B(y): \text{ magnetic field}$$

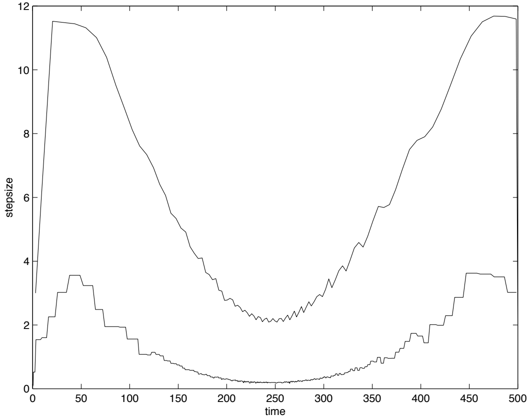
Lie group action by freezing magnetic field ($\Rightarrow$ helical motion):

$$\begin{pmatrix} y(t) \\ v(t) \end{pmatrix}' = \begin{pmatrix} v(t) \\ v(t) \times B(y_0) \end{pmatrix}.$$

Equations rewritten by action of $\mathrm{GL}(6)$ on $\mathbb{R}^6$. $f \colon \mathcal{M} \to \mathfrak{gl}(6)$

$$\begin{pmatrix} y \\ v \end{pmatrix}' = \begin{pmatrix} v \\ v \times B(y) \end{pmatrix} = \begin{pmatrix} 0 & I \\ 0 & -\widehat{B(y)} \end{pmatrix} \cdot \begin{pmatrix} y \\ v \end{pmatrix} = f(y) \cdot \begin{pmatrix} y \\ v \end{pmatrix}.$$

Størmers problem integrated with an LGI



Stepsize in LGI and RK-45.

LGI is more efficient, since it can take 8-10 times longer steps with same accuracy.

Actions for LGI

Actions are for LGI as preconditioners for iterative solvers.

Examples of actions:

- Action of $O(n)$ by similarity transforms on $\mathbb{R}^n$, $g \cdot A = gAg^{-1}$. (Isospectral flows).
- Action of Aff(n) on $\mathbb{R}^n$, $(A, b) \cdot x = Ax + b$. (Info about jacobian can be taken into the action).
- Action by solving simplified differential equations, e.g. equatoin $u_t = \nabla(\mu(x)\nabla u)$ can be solved by using solution of $u_t = c\nabla^2 u + b$ as action (affine action by circulants).
- Co-adjoint action for Lie Poisson systems (geometric mechanics).

Examples of LGI

Crouch–Grossman methods: (CG '93, Owren-Marthinsen '99)

$$
\begin{aligned}
X_1 &= Y_n, & F_1 &= A(t_n, X_1),\\
X_2 &= e^{\frac{3}{4}hF_1} Y_n, & F_2 &= A(t_n + \tfrac{3}{4}h, X_2),\\
X_3 &= e^{\frac{17}{108}hF_2} e^{\frac{119}{216}hF_1} Y_n, & F_3 &= A(t_n + \tfrac{17}{24}h, X_3),\\
Y_{n+1} &= e^{\frac{13}{51}hF_3} e^{-\frac{2}{3}hF_2} e^{\frac{24}{17}hF_1} Y_n.
\end{aligned}
$$

Examples of LGI

MK-type methods: (MK '95, '98, '99)

$$k_1 = hf(y_0),$$
$$k_2 = hf(\exp(\tfrac{1}{2}k_1) \cdot y_0),$$
$$k_3 = hf(\exp(\tfrac{1}{2}k_2 - \tfrac{1}{8}[k_1, k_2]) \cdot y_0),$$
$$k_4 = hf(\exp(k_3) \cdot y_0),$$
$$y_1 = \exp(\tfrac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4 - \tfrac{1}{2}[k_1, k_4])) \cdot y_0.$$

Examples of LGI

Magnus and Fer expansions: (Iserles-Nørsett '99, Iserles '84)

These are special methods for solving equations of the form

$$y' = f(t)y,$$

where $f\colon \mathbb{R} \to \mathfrak{g}$, by iterated integrals.

Examples of LGI

Commutator free methods (Celledoni, Owren, Marthinsen 2003)

By a combination of exponentials and sums one can avoid commutators. Advantage for stiff systems.

$$
\begin{aligned}
Y_1 &= y_0, \\
Y_2 &= \exp(\tfrac{1}{2}k_1)\cdot y_0, \\
Y_3 &= \exp(\tfrac{1}{2}k_2)\cdot y_0 \\
Y_4 &= \exp(k_3 - \tfrac{1}{2}k_1)\cdot Y_2, \\
y_{\frac{1}{2}} &= \exp(\tfrac{1}{12}(3k_1 + 2k_2 + 2k_3 - k_4))\cdot y_0, \\
y_1 &= \exp(\tfrac{1}{12}(-k_1 + 2k_2 + 2k_3 + 3k_4))\cdot y_{\frac{1}{2}}.
\end{aligned}
$$

Examples of LGI

Retraction and coordinate based methods

These are methods based on other mappings (than exp) between $\mathfrak{g}$ and G, ex. Cayley map, generalized polar coordinates etc.

Symplectic Lie-group methods (Bogfjellmo, H. Marthinsen 2014)

These methods are derived from discrete variational principles and are symplectic for Hamiltonian flows on (the cotangent bundle) of Lie groups.

Properties of the (matrix) exponential

BCH equation

Theorem 2.5. (Baker–Campbell–Hausdorff) For sufficiently small $t \geq 0$ we have

$$\exp(tF) \circ \exp(tG) = \exp(tH),$$

where $H = \mathrm{bch}(F, G)$ can be constructed from iterated commutators of F and G. The first few terms are

$$H = F + G + \tfrac{1}{2}t[F, G] + \tfrac{1}{12}t^2\left([F, [F, G]] + [G, [G, F]]\right) + \mathcal{O}(t^3).$$

Properties of the (matrix) exponential

'dexp equation' (Hausdorff 1904)

$$\frac{d}{dt}\exp(A(t)) = d\exp_{A(t)}(A'(t))\exp(A(t))$$

$$d\exp_A(B) = A + \frac{1}{2}[A,B] + \frac{1}{6}[A,[A,B]] + \cdots$$

$$= \sum_{k=0}^{\infty} \frac{1}{(k+1)!}\mathrm{ad}_A^k B = \left.\frac{e^z-1}{z}\right|_{z=\mathrm{ad}_A} B$$

$$d\exp_A^{-1}(B) = \left.\frac{z}{e^z-1}\right|_{z=\mathrm{ad}_A} B = B - \frac{1}{2}[A,B] + \frac{1}{12}[A,[A,B]] - \cdots$$

$$= \sum_{j=0}^{\infty} \frac{B_j}{j!}\mathrm{ad}_A^j(B),$$

B_j Bernoulli numbers.

Properties of the (matrix) exponential

Corollary

The Lie group equation

$$y'(t) = f(y(t)), \quad y(t) \in G, \quad y(0) = y_0$$

has solution

$$y(t) = \exp(\theta(t))y_0, \quad \theta(t) \in \mathfrak{g},$$

where θ *solves the equation*

$$\theta'(t) = d\exp_\theta^{-1}(f(y(t)), \quad \theta(0) = 0$$

on the flat vector space $\mathfrak{g}$.

Generalized polar decompositions and the matrix exponential

(this last part of my talk is done on blackboard).