Lie group methods

Elena Celledoni*

July 26, 2015

1 Background material

This is section a summary of the results of the first and second chapter in [13] and chapter 9 in [10] which are particularly relevant for the introduction of Lie group methods.

1.1 Manifolds

Definition 1.1. *[13] An m-dimensional manifold* $\mathcal{M}$ *is a topological space covered by a collection of open subsets* $W_\alpha \subset \mathcal{M}$ *(coordinate charts) and maps* $\mathcal{X}_\alpha : W_\alpha \to V_\alpha \subset \mathbf{R}^m$ *one-to-one and onto, where* V_α *is an open, connected subset of* $\mathbf{R}^m$*.* $(W_\alpha, \mathcal{X}_\alpha)$ *define coordinates on* $\mathcal{M}$*.*
$\mathcal{M}$ *is a smooth manifold if the maps* $\mathcal{X}_{\alpha\beta} = \mathcal{X}_\beta \circ \mathcal{X}_\alpha^{-1}$*, are smooth where they are defined, i.e. on* $\mathcal{X}_\alpha(W_\alpha \cap W_\beta)$ *to* $\mathcal{X}_\beta(W_\alpha \cap W_\beta)$*.*

Example 1.2. $\mathbf{R}^m$ *is a m-dimensional manifold covered with a single chart.*

Example 1.3. *The unit sphere* $\mathbf{S}^{m-1} := \{\mathbf{x} \in \mathbf{R}^m \mid \sum_{i=1}^m x_i^2 = 1\}$ *is a* $m-1$*-dimensional manifold covered with two charts obtained by omitting the north and south poles respectively. The coordinate maps are obtained considering the stereographic projection from the north and south pole respectively.*

Given two smooth manifolds $\mathcal{M}$ and $\mathcal{N}$ we say that $F : \mathcal{M} \to \mathcal{N}$ is a *smooth map* if it is smooth in local coordinates. Introducing local coordinates on the manifolds we get $x \in \mathcal{M}$, $x = (x_1, \ldots, x_m)$, and $y \in \mathcal{N}$, $y = (y_1, \ldots, y_n)$. Assume $y = F(x)$ and $y_i = F_i(x)$ $i = 1, \ldots, n$; if F_i is smooth as a map from an open subset of $\mathbf{R}^m$ to $\mathbf{R}$, then F is smooth also as map between $\mathcal{M}$ and $\mathcal{N}$.

*Department of Mathematical Sciences NTNU

The *rank of a map* $F : \mathcal{M} \to \mathcal{N}$ is the rank of the Jacobian of F, the map is said to be *regular* if its rank is constant.

A subset $\mathcal{N} \subset \mathcal{M}$ of a manifold which is a manifold in its own right is a submanifold.

Definition 1.4. Submanifolds *An immersed submanifold* $\mathcal{N}$ *of a manifold* $\mathcal{M}$ *is a subset* $\mathcal{N} \subset \mathcal{M}$ *and a map* F *smooth and one-to-one* $F : \tilde{\mathcal{N}} \to \mathcal{N} \subset \mathcal{M}$ *with* F *everywhere of maximal rank and* $\tilde{\mathcal{N}}$ *an* n*-dimensional manifold.*

Example 1.5. $\mathcal{M} = \mathbf{R}^3$ *consider the parametrized curve* $\phi(t) = (\cos(t), \sin(t), t)$ *(a circular elix, Figure 2 (left)),* ϕ *is one-to-one and* $\dot{\phi} = (-\sin(t), \cos(t), 1)$ *is never* 0 *so the maximal rank condition is satisfied and the elix is an immersed submanifold of* $\mathbf{R}^3$*.*

1.2 Vector fields

A tangent vector to a manifold $\mathcal{M}$ at a point is the tangent to a smooth curve passing through the point: given $x \in \mathcal{M}$ and $\phi(t) \in \mathcal{M}$ the curve such that $\phi(0) = x$ then

$$\mathbf{v}|_x := \frac{d}{dt}\phi(t)\Big|_{t=0}.$$

The *tangent space* to a m-dimensional manifold $\mathcal{M}$ at the point x is the vector space of dimension m formed by the collection of the tangent vectors at x and is denoted by $T_x\mathcal{M}$. Two curves $\phi(t)$ and $\gamma(t)$ will give the same tangent vector if they both pass through x at $t = 0$ with the same direction, i.e. $\phi(0) = x = \gamma(0)$ and $v = \frac{d}{dt}\phi(t)\big|_{t=0} = \frac{d}{dt}\gamma(t)\big|_{t=0}$.

The *tangent bundle*

$$T\mathcal{M} = \cup_{x \in \mathcal{M}} T_x\mathcal{M}$$

is the collection of all tangent spaces, it can be given the structure of a manifold of dimension $2m$.

The tangent bundle to the circle can be identified with the cartesian product of the circle with $\mathbf{R}$, $TS^1 \simeq S^1 \times \mathbf{R}$. The tangent bundle to the sphere TS^2 cannot be identified with the cartesian product of the sphere and $\mathbf{R}^2$.

A *vector field* on $\mathcal{M}$ is a section of the tangent bundle of $\mathcal{M}$, i.e. is a smoothly varying assignment of tangent vectors: $\mathbf{v} : \mathcal{M} \to T\mathcal{M}$ such that $\mathbf{v}(x) = \mathbf{v}|_x \in T_x\mathcal{M}$. In local coordinates

$$\mathbf{v}(x) = \sum_{i=1}^{m} \xi^i(x) \frac{\partial}{\partial x^i},$$

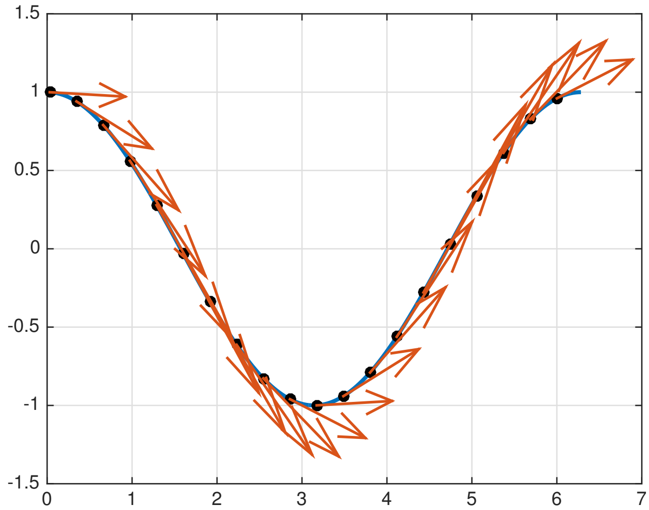


Figure 1: The curve $(t, \cos(t))$ is an immersed submanifold of $\mathbb{R}^2$. For each point $(t, \cos(t))$ of the curve the vector $(1, -\sin(t))$ obtained taking derivatives with respect to t, is the tangent vector at the point $(t, \cos(t))$ to the curve (after translation to the origin $(0, 0)$). We consider here $t \in (0, 2\pi)$.

$\xi^i(x)$ are smooth functions and $\frac{\partial}{\partial x^i}$ denote a basis of the tangent space $T_x\mathcal{M}$.

A curve $\phi : \mathbf{R} \to \mathcal{M}$ is an *integral curve* of the vector field $\mathbf{v}$ if, when $\phi(t) = x$, the tangent to the curve at t coincides with the vector field at x, i.e. $\dot{\phi}(t) = \mathbf{v}(x)$. This means that in local coordinates

$$\frac{dx^i}{dt} = \xi^i(x), \quad x^i = \phi_i(t).$$

Example 1.6. *We consider a vector field on* $\mathbf{R}^2$,

$$\mathbf{v}(x, y) = y\partial_x - x\partial_y,$$

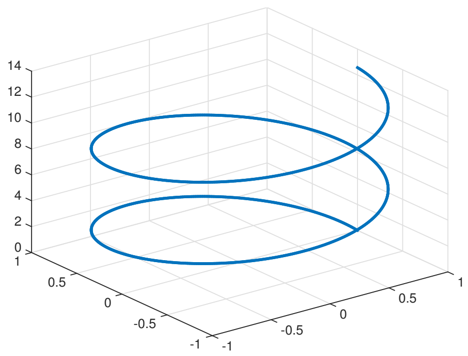
$$\mathbf{v}(x, y) = \begin{pmatrix} \xi^1(x, y) \\ \xi^2(x, y) \end{pmatrix} = \begin{pmatrix} y \\ -x \end{pmatrix},$$

and to find the integral curve one has to solve

$$\begin{aligned} \dot{x} &= y, \\ \dot{y} &= -x, \end{aligned}$$

obtaining

$$\begin{aligned} x(t) &= \cos(t)x_0 + \sin(t)y_0, \\ y(t) &= -\sin(t)x_0 + \cos(t)y_0. \end{aligned}$$



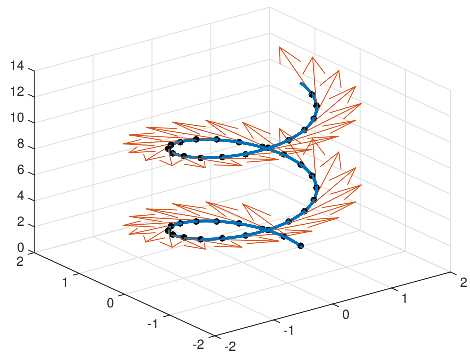


Figure 2: The curve $(\cos(t), \sin(t), t)$ is an immersed submanifold of $\mathbb{R}^2$ (left). The tangent spaces of this manifod are 1-dimensional and they are the lines through the arrows of the vector field tangent to the curve (right). Differentiating the curve we obtain $(-\sin(t), \cos(t), 1)$, i.e. the length and direction of these tangent vectors. We have considered here $t \in (0, 4\pi)$.

If $\phi(t)$ is a maximal integral curve of the vector field we denote it with

$$\phi(t) = \exp(t\mathbf{v})x_0, \quad x_0 = \phi(0),$$

$\exp(t\mathbf{v})x_0$ is the *flow* generated by the vector field $\mathbf{v}$, while $\mathbf{v}$ is called the *infinitesimal generator* of the flow. This notation is justified by some fundamental properties of the flow resembling known properties of the exponential mapping:

$$\exp(t\mathbf{v})\exp(s\mathbf{v})x_0 = \exp((t+s)\mathbf{v})x_0, \quad \exp(0\mathbf{v})x_0 = x_0,$$

$$\exp(t\mathbf{v})^{-1}x_0 = \exp(-t\mathbf{v})x_0, \quad \frac{d}{dt}\exp(t\mathbf{v})x_0 = \mathbf{v}|_{\exp(t\mathbf{v})x_0}$$

and also

$$\mathbf{v}|_{x_0} = \frac{d}{dt}\exp(t\mathbf{v})x_0\Big|_{t=0}, \quad \forall x_0 \in \mathcal{M},$$

i.e. given the flow starting from x_0 we can retrieve the vector field at x_0 by differentiating the flow with respect to t and then setting $t = 0$. The flow of a vector field can be expanded as

$$\exp(t\mathbf{v})x_0 = x_0 + t\,\mathbf{v}|_{x_0} + \mathcal{O}(t^2).$$

If x_0 is such that $\mathbf{v}|_{x_0} = 0$ we say that x_0 is a *singularity or equilibrium point* of the vector field, and this implies

$$\exp(t\mathbf{v})x_0 = x_0, \quad \forall t.$$

Points that are not equilibrium points are called *regular*.

Vector fields can operate on functions as *derivations*.

A derivation is a linear operator defined on an algebra[1] $\mathcal{A}$, $D : \mathcal{A} \to \mathcal{A}$ satisfying the Leibniz rule, $D(ab) = D(a)b + aD(b)$, for all $a, b \in \mathcal{A}$, where ab is the product of a and b in $\mathcal{A}$.

Given $f : \mathcal{M} \to \mathbf{R}$ the result of applying $\mathbf{v}$ as a derivation on f is a new function $\mathbf{v}(f)$ such that

$$\mathbf{v}(f(x)) = \sum_{i=1}^{m} \xi^i(x) \frac{\partial f}{\partial x^i} = \frac{d}{dt} f(\exp(t\mathbf{v})x) \bigg|_{t=0},$$

$\mathbf{v}(f)$ determines the infinitesimal change of f along the flow of $\mathbf{v}$. It is easy to verify that $\mathbf{v}$ acts as a derivation

1. $\mathbf{v}(\lambda f + \mu g) = \lambda \mathbf{v}(f) + \mu \mathbf{v}(g)$,
2. $\mathbf{v}(fg) = f\mathbf{v}(g) + g\mathbf{v}(f)$.

The *Lie series* expansion of $f : \mathcal{M} \to \mathbf{R}$ is an expansion of f evaluated along the flow of $\mathbf{v}$

$$f(\exp(t\mathbf{v})x) = f(x) + t\mathbf{v}(f(x)) + \frac{1}{2} t^2 \mathbf{v}(\mathbf{v}(f(x))) + \ldots,$$

converging for t sufficiently near 0. This is a way of reconstructing f along the flow of $\mathbf{v}$ given $\mathbf{v}$.

The *Lie bracket* of vector fields is an operation on the set of vector fields, given two vector fields $\mathbf{v}$ and $\mathbf{w}$, $[\mathbf{v}, \mathbf{w}]$ is also a vector field. Such vector field is identified by the way it is acting on smooth functions, i.e. for all smooth $f : \mathcal{M} \to \mathbf{R}$,

$$[\mathbf{v}, \mathbf{w}](f) = \mathbf{v}(\mathbf{w}(f)) - \mathbf{w}(\mathbf{v}(f)). \tag{1}$$

In coordinates, assuming

$$\mathbf{v} = \sum_{i=1}^{m} \xi^i \frac{\partial}{\partial x^i}, \quad \mathbf{w} = \sum_{i=1}^{m} \eta^i \frac{\partial}{\partial x^i},$$

[1] An algebra is a vector space equipped with a multiplication operation, "$\cdot$" $: \mathcal{A} \times \mathcal{A} \to \mathcal{A}$. This operation is distributive with respect to the addition of the vector space and is compatible with the product by scalars in an appropriate sense.

we obtain

$$
\begin{aligned}
[\mathbf{v},\mathbf{w}] &= \sum_{i=1}^{m}\xi^i\sum_{j=1}^{m}\frac{\partial\eta^j}{\partial x^i}\frac{\partial}{\partial x^j}-\sum_{i=1}^{m}\eta^i\sum_{j=1}^{m}\frac{\partial\xi^j}{\partial x^i}\frac{\partial}{\partial x^j}\\
&= \sum_{j=1}^{m}\left(\sum_{i=1}^{m}\xi^i\frac{\partial\eta^j}{\partial x^i}-\sum_{i=1}^{m}\eta^i\frac{\partial\xi^j}{\partial x^i}\right)\frac{\partial}{\partial x^j}.
\end{aligned}
$$

One can verify that the following important properties hold for the Lie bracket of vector fields:

1. *bilinearity*: $[\lambda_1\mathbf{v}_1+\lambda_2\mathbf{v}_2,\mathbf{w}]=\lambda_1[\mathbf{v}_1,\mathbf{w}]+\lambda_2[\mathbf{v}_2,\mathbf{w}]$.
2. *skew-symmetry*: $[\mathbf{v},\mathbf{w}]=-[\mathbf{w},\mathbf{v}]$.
3. *Jacobi identity*: $[\mathbf{v},[\mathbf{w},\mathbf{u}]]+[\mathbf{u},[\mathbf{v},\mathbf{w}]]+[\mathbf{w},[\mathbf{u},\mathbf{v}]]=0$.

The *derivative map* or *differential* of a given map $F:\mathcal{M}\to\mathcal{N}$ is a map

$$dF:T\mathcal{M}\to T\mathcal{N},\quad s.t.\quad dF|_x:T_x\mathcal{M}\to T_{F(x)}\mathcal{N}$$

which is such that for any curve $\phi(t)$ such that $\phi(t)|_{t=0}=x$ and correspondingly $F(\phi(t))|_{t=0}=F(x)$, with tangent vectors $\mathbf{v}_x:=\frac{d}{dt}\phi(t)\big|_{t=0}$ and $\mathbf{w}_{F(x)}:=\frac{d}{dt}F(\phi(t))\big|_{t=0}$ we have

$$dF|_x(\mathbf{v}_x)=\mathbf{w}_{F(x)}.$$

The differential is a linear map and in coordinates it is represented by the Jacobian of F. Only when F is one-to-one dF maps vector fields to vector fields, [13].

Assume F is one-to-one and $\mathbf{v}$ a vector field on $\mathcal{M}$ and $d\,F(\mathbf{v})$ a vector field on $\mathcal{N}$, one can prove that

$$F(\exp(t\mathbf{v})x)=\exp(td\,F(\mathbf{v}))F(x). \tag{2}$$

If $\mathbf{w}$ is also a vector field on $\mathcal{M}$ than one can also prove that the Lie bracket of vector fields is invariant under $d\,F$, i.e.

$$d\,F([\mathbf{v},\mathbf{w}])=[d\,F(\mathbf{v}),d\,F(\mathbf{w})].$$

1.3 Lie groups

A Lie group is a manifold G equipped with a smooth product operation "$\cdot$" which gives to G a group structure, e.g. there exists an identity element $e \in G$ and for any $g \in G$ its inverse g^{-1} is in G.

Here follows a list of examples of Lie groups,

- $(\mathbf{R}, +)$
- $\mathrm{GL}(n) = \{A \in M^{n\times n} | \det(A) \neq 0\}$ with the product between $n \times n$ matrices as group product.

Charts

Given any local chart on G we can construct an atlas on the Lie group G by left (or right) multiplication. Suppose (U, φ) is a chart about the identity $e \in G$ and $\varphi : U \to V$, then a chart about $g \in G$, (U_g, φ_g) can be obtained by

$$U_g := L_g(U) := \{L_g(h) \in G \,|\, h \in U\}, \quad \varphi_g = \varphi \circ L_{g^{-1}} : U_g \to V.$$

One can prove that the maps

$$\varphi_{g_1} \circ \varphi_{g_2}^{-1} : \varphi_{g_2}(U_{g_1} \cap U_{g_2}) \to \varphi_{g_1}(U_{g_1} \cap U_{g_2}),$$

are smooth because $\varphi_{g_1} \circ \varphi_{g_2}^{-1} = \varphi \circ L_{g_1}^{-1} \circ L_{g_2} \circ \varphi^{-1}$ is a composition of smooth maps.

Proposition 1.7. *([13] chapter II) Let G be a Lie group. If* $\mathrm{H} \subset \mathrm{G}$ *is a subgroup of* G *and* H *is topologically closed then* H *is a Lie subgroup of G.*

Using the previous proposition it is easily verified that the following sets are Lie subgroups of $GL(n)$.

- $\mathrm{SL}(n) = \{A \in M^{n\times n} | \det(A) = 1\}$ with the product between $n \times n$ matrices as group product,
- $\mathrm{SO}(n) = \{A \in M^{n\times n} | \det(A) = 1, \; A^T A = I\}$ with the product between $n \times n$ matrices as group product,
- $\mathrm{SP}(2r) = \{A \in M^{2r\times 2r} | \; A^T J A = J\}$ with the product between $2r \times 2r$ matrices as group product,

here $M^{n\times n}$ is the set of $n \times n$ real matrices.

1.4 Transformation groups

A transformation group acting on a smooth manifold $\mathcal{M}$ is a Lie group G and a smooth map $\Lambda : \mathrm{G} \times \mathcal{M} \to \mathcal{M}$ such that

- $\Lambda(e, x) = x$ for all $x \in \mathcal{M}$.
- $\Lambda(g, \Lambda(h, x)) = \Lambda(g \cdot h, x)$ for all $x \in \mathcal{M}$ and $g, h \in \mathrm{G}$.

Λ is called a Lie group action. We say that the Lie group action is *global* when $\Lambda(g, x)$ is defined for all $x \in \mathcal{M}$ and $g \in \mathrm{G}$ and local if it is defined on an open subset $\mathcal{V} \subset \mathrm{G} \times \mathcal{M}$ such that $\{e\} \times \mathcal{M} \subset \mathcal{V}$.

Some examples:

- $\mathrm{GL}(n, \mathbf{R})$ (or any of its subgroups) acting on $\mathbf{R}^n$ by matrix-vector multiplication.
- Any Lie group can act on itself by the group multiplication.

The set

$$\mathcal{O}_x = \{m \in \mathcal{M} \,|\, m = \Lambda(g, x),\; g \in \mathrm{G}\}$$

is called *orbit* of the Lie group action.

Example 1.8. *Consider the group* $\mathrm{O}(2)$ *acting on* $\mathbf{R}^2$ *the orbits are circles around the origin of* $\mathbf{R}^2$*. Analogously for* $\mathrm{O}(n)$ *acting on* $\mathbf{R}^n$ *the orbits are spheres:*

$$\{\, x \in \mathrm{R}^n \,|\, \|x\| = C \,\}$$

with C *a constant.*

A Lie group action is said to be *transitive* when there is only one orbit, $\mathcal{O}_x = \mathcal{M}$, i.e.

$$\forall y \in \mathcal{M}\ \exists g\ \in G,\ \ s.t.\, \Lambda(g, x) = y.$$

Example 1.9. *The action of a Lie group* G *on itself by left multiplication is transitive.*

1.5 Homogeneous spaces

Given a Lie group G and a subgroup H we can define an equivalence relation on G:

$$g \sim \tilde{g} \Leftrightarrow \exists\, \tilde{h} \in \mathrm{H}\, \text{s.t.} \tilde{g} = g\tilde{h}.$$

The equivalence classes, $[g] = g \cdot \mathrm{H}$, are called left-cosets

$$[g] = \{gh \,|\, h \in \mathrm{H}\,\}.$$

One can prove that if H is a closed subgroup then the quotient G/H (i.e. G/ $\sim$) is a manifold called homogeneous space.

Recall that for G/H to be a group H needs to be a normal subgroup i.e. $g\mathrm{H}g^{-1} = H$ for all $g \in \mathrm{G}$.

In a homogeneous space the action $\Lambda : \mathrm{G} \times \mathrm{G/H} \to \mathrm{G/H}$, $\Lambda(g, [\tilde{g}]) = [g\tilde{g}]$ is transitive. In fact for any $[g_1]$ and $[g_2]$ in G/H it exists $g \in \mathrm{G}$ such that $g = g_2 g_1^{-1}$ and $\Lambda(g, [g_1]) = [g_2]$.

Example 1.10. Relevant homogeneous spaces. *Prove that the sphere is an homogeneous space* $S^2 = \mathrm{SO}(3)/\mathrm{SO}(2)$.

More in general $\mathrm{SO}(n)/\mathrm{SO}(p)$ *for* $p < n$ *is another interesting homogeneous space called Stiefel manifold and can be identified with the set of all* $n \times p$ *matrices with* p *orthonormal columns.*

Analogously $\mathrm{O}(n)/(\mathrm{O}(p) \times \mathrm{O}(n-p))$ *is the homogeneous space also known as Grassmann manifold.*

Definition 1.11. *Given* $x \in \mathcal{M}$ *and* Λ *a Lie group action on* $\mathcal{M}$ *the isotropy subgroup of* $x \in \mathcal{M}$ *is*

$$\mathrm{G}_x = \{g \in G \,|\, \Lambda(g, x) = x\}.$$

Recall Proposition 1.7, this implies that G_x is a Lie subgroup.

Theorem 1.12. *A Lie group* G *acts globally and transitively on* $\mathcal{M}$ *if and only if* $\mathcal{M} \simeq \mathrm{G/H}$ *is isomorphic to the homogeneous space obtained as* G/H *with* $H = G_x$ *the isotropy subgroup of any chosen* $x \in \mathcal{M}$.

So any transitive Lie group action corresponds to a homogeneous space and viceversa.

1.6 Lie algebra of a Lie group

A Lie algebra $\mathfrak{g}$ ia a vector space with a bracket operation:

- $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ is a bilinear map
- is skew-symmetric: $[u, v] = -[v, u]$, $\forall\, u, v \in \mathfrak{g}$
- satisfies the Jacobi identity: $[u, [v, w]] + [w, [u, v]] + [v, [w, u]] = 0$ $\forall u, v, w \in \mathfrak{g}$.

The Lie algebra of a given Lie group G is the tangent space at the identity element e, $\mathfrak{g} := \mathrm{T}_e G$.

For matrix Lie groups the Lie algebra is typically a linear vector subspace of $M^{n \times n}$.

One can verify that

- the Lie algebra of $(\mathbf{R}, +)$ is $\mathbf{R}$
- the Lie algebra of $\mathrm{GL}(n)$ is $\mathfrak{gl}(n) = M^{n\times n}$,
- the Lie algebra of $\mathrm{SL}(n)$ is $\mathfrak{sl}(n) = \{A \in M^{n\times n} | \mathrm{trace}(A) = 0\}$.

Exercise 1.13. *The Lie algebra of* $\mathrm{Sp}(2r)$ *is*

$$\mathfrak{sp}(2r) = \{A \in M^{2r\times 2r} | AJ + JA^T = O\}.$$

Proof. Consider $A(t) \in \mathrm{Sp}(2r)$ $A(0) = I$, we get $V = \frac{d}{dt}A(t)\big|_{t=0} \in \mathfrak{sp}(2r)$. We differentiate $A(t)JA(t)^T = J$ and obtain

$$\dot{A}JA^T + AJ\dot{A}^T = O.$$

Setting $t = 0$ we obtain

$$VJ + JV^T = O.$$

□

One way to derive the Lie bracket of a Lie algebra $\mathfrak{g}$ of a Lie group G is by the identification of T_eG with the set of left (or right) invariant vector fields on G. This set is naturally a Lie algebra with the Lie bracket of vector fields (1) as bracket operation. This way we will naturally obtain the bracket operation on T_eG.

Consider the left multiplication of the Lie group $L_g : \mathrm{G} \to \mathrm{G}$, $L_g(h) = g\,h$, consider also the derivative mapping or differential of L_g, $d\,L_g|_x : T_xG \to T_{gx}G$.

Definition 1.14. *A vector field on* G, $\mathbf{v}$, *is left invariant if*

$$d\,L_g(\mathbf{v}) = \mathbf{v},$$

and fiberwise

$$d\,L_g|_x\,(\mathbf{v}(x)) = \mathbf{v}(L_g(x)).$$

Since for a left invariant vector field $\mathbf{v}(e) = A$ implies $d\,L_g|_e\,(\mathbf{v}(e)) = \mathbf{v}(g)$, once we know how the vector field is at the identity, e, via left multiplication we know how it is everywhere. For this reason we can identify the tangent space at the identity of the Lie group G, i.e. the Lie algebra $\mathfrak{g}$, with the set of left invariant vector fields. An analogous definition and identification can be given for right invariant vector fields.

The Lie algebra can be also used to describe the tangent space to G at any point. Here is the case of the orthogonal group.

Example 1.15. *Consider $\gamma(t) \in \mathrm{O}(n)$. $\gamma(t)^T\gamma(t) = I$ assume $\gamma(0) = Q$ and $\dot{\gamma}(0) = W$. By differentiating with respect to t and setting $t = 0$ we obtain $W^TQ + Q^TW = O$. Set $A := Q^TW$ and substitute $W = QA$ in the previous equation, obtaining $A^T + A = O$. So we obtain a characterization of the tangent space of $\mathrm{O}(n)$ at Q by means of $\mathfrak{so}(n)$:*

$$T_Q\,\mathrm{O}(n) = \{W = QA \mid A \in \mathfrak{so}(n)\}.$$

Analogous results can be obtained for other matrix Lie groups.

Recall form chapter I in [13]: given a mapping $F : \mathcal{M} \to \mathcal{N}$ the derivative mapping $d\,F$ does not map vector fields to vector fields unless F is one-to-one. Assume F is one-to-one and $\mathbf{v}$ a vector field on $\mathcal{M}$ and $d\,F(\mathbf{v})$ a vector field on $\mathcal{N}$, one can prove that

Proposition 1.16.

$$F(\exp(t\mathbf{v})x) = \exp(td\,F(\mathbf{v}))F(x). \tag{3}$$

If $\mathbf{w}$ is also a vector field on $\mathcal{M}$ than one can also prove that the Lie bracket of vector fields is invariant under $d\,F$, i.e.

Proposition 1.17.

$$d\,F([\mathbf{v}, \mathbf{w}]) = [d\,F(\mathbf{v}), d\,F(\mathbf{w})].$$

If $\mathcal{M} = \mathrm{G}$ (a Lie group) the right multiplication $L_g : \mathrm{G} \to \mathrm{G}$ is one-to-one. Assume $\mathbf{v}$ and $\mathbf{w}$ are left invariant vector fields: $d\,L_g(\mathbf{v}) = \mathbf{v}$ and $d\,L_g(\mathbf{w}) = \mathbf{w}$, then

$$d\,L_g([\mathbf{v}, \mathbf{w}]) = [d\,L_g(\mathbf{v}), d\,L_g(\mathbf{w})] = [\mathbf{v}, \mathbf{w}],$$

so that also the Lie bracket $[\mathbf{v}, \mathbf{w}]$ of the two left invariant vector fields is left invariant. The set of left invariant vector fields on G, $\mathcal{X}_L(G)$, is closed under the Lie-Jacobi bracket of vector fields and is therefore a Lie algebra.

Identification of T_eG with $\mathcal{X}_L(G)$

Let $a \in T_eG$, we consider the vector field on G defined by

$$\mathbf{v}_a(g) := d\,L_g|_e\,(a),$$

and $d\,L_g|_e : T_eG \to T_gG$. The vector field $\mathbf{v}_a$ is left invariant. In fact by the definition $\mathbf{v}_a(L_g(h)) = \mathbf{v}_a(gh)) = d\,L_{gh}|_e\,(a) = d\,(L_g \circ L_h)|_e\,(a) =$

$d\,L_g|_h \circ d\,L_h|_e\,(\mathbf{v}_a(e))$. Substituting $\mathbf{v}_a(e) = a$, and applying the definition of $\mathbf{v}_a$ one more time we get $\mathbf{v}_a(L_g(h)) = d\,L_g|_h\,(\mathbf{v}_a(h))$, which is the definition of left invariance 1.14.

So we have a map $\zeta_1 : T_eG \to \mathcal{X}_L(G)$. On the other hand we can construct the map

$$\zeta_2 : \mathcal{X}_L(G) \to T_eG, \quad \zeta_2(\mathbf{w}) := \mathbf{w}(e).$$

It can be seen that

$$\zeta_1 \circ \zeta_2 = \mathrm{id}, \quad \zeta_2 \circ \zeta_1 = \mathrm{id},$$

and that we so have an isomorphism of vector spaces between T_eG and $\mathcal{X}_L(G)$.

So the Lie bracket of T_eG can be defined to be

$$[a, b] := [\mathbf{v}_a, \mathbf{v}_b](e) \tag{4}$$

where at the left hand side the symbol $[\cdot, \cdot]$ denotes the Lie-Jacobi bracket of vector fields.

Example 1.18. *Consider $\mathfrak{gl}(n)$ and the bracket of vector fields by a similar argument to that used in example 1.15 we obtain that a left invariant vector field $\mathbf{v}_A$ on $\mathrm{GL}(n)$ has coordinates given by the matrix product XA with $A \in \mathfrak{gl}(n)$ and $X \in \mathrm{GL}(n)$. Consider two left invariant vector fields $\mathbf{v}_A$ and $\mathbf{v}_B$ and write them as derivation operators:*

$$\mathbf{v}_A(X) = \sum_{i,j,k=1}^{n} x_{i,k} a_{k,j} \frac{\partial}{\partial x_{i,j}}, \quad \mathbf{v}_B(X) = \sum_{i,j,k=1}^{n} x_{i,k} b_{k,j} \frac{\partial}{\partial x_{i,j}},$$

here we denote with $a_{i,j} = (A)_{i,j}$ the (i, j)-element of the matrix A. Computing the Lie bracket of the two vector fields we obtain

$$\begin{aligned} [\mathbf{v}_A, \mathbf{v}_B](X) &= \sum_{i,j,k=1}^{n} x_{i,k} a_{k,j} \sum_{s=1}^{n} b_{j,s} \frac{\partial}{\partial x_{i,s}} - \sum_{l,r,s=1}^{n} x_{l,r} b_{r,s} \sum_{j=1}^{n} a_{s,j} \frac{\partial}{\partial x_{l,j}} \\ &= \sum_{i,k,s=1}^{n} x_{i,k} (AB - BA)_{k,s} \frac{\partial}{\partial x_{i,s}}. \end{aligned}$$

The Lie-Jacobi bracket of the two vector fields is a vector field with coordinates the matrix commutator of A and B: $[A, B] = AB - BA$ and, as expected,

$$[A, B] = [\mathbf{v}_A, \mathbf{v}_B](I).$$

Here at the left hand side the symbol $[\cdot, \cdot]$ denotes the matrix commutator and at the right hand side the Lie-Jacobi bracket of vector fields. The situation is analogous for the case of right invariant vector fields a part from a change of sign.

1.7 The exponential map

Consider $\mathbf{v}$ a right invariant vector field on G and the right multiplication of the Lie group R_g. Using (3) we obtain

$$R_g(\exp(t\,\mathbf{v})e) = \exp(t\,dR_g(\mathbf{v}))g$$

and further using the right invariance of $\mathbf{v}$ on the right hand side we get

$$(\exp(t\,\mathbf{v})e)g = \exp(t\,\mathbf{v})g,$$

so that the left multiplication of g by the flow through e of $\mathbf{v}$ is equal to the flow through g of $\mathbf{v}$. We therefore can identify the flow of the right invariant vector field to be the corresponding one parameter subgroup[2] of G:

$$\exp(t\,\mathbf{v}) := \exp(t\,\mathbf{v})e.$$

We can also define the exponential map $\exp : \mathfrak{g} \to \mathrm{G}$ as

$$\mathbf{v} \in \mathfrak{g} \mapsto \exp(t\mathbf{v})e|_{t=1} \in \mathrm{G}.$$

Example 1.19. *The flow of a right invariant vector field*

$$\mathbf{v}_A = \sum_{i,j,k} a_{i,j} x_{k,j} \frac{\partial}{\partial x_{i,j}},$$

is $\gamma(t)$ such that $\dot{\gamma} = \mathbf{v}_A(\gamma(t))$, in coordinates

$$\dot{\gamma}_{i,j} = \sum_{i,j=1}^{n} \left(\sum_{k=1}^{n} a_{i,k}\gamma_{k,j}(t) \right), \quad \text{i.e.} \quad \dot{\gamma} = A\gamma,$$

and $\gamma(t) = \exp(tA)\gamma(0)$. For the left invariant vector fields the flow is instead of the type $\eta(t) = \eta(0)\exp(tA)$.

Exercise 1.20. *Show that $\exp(O) = e$, where O is the zero element in $\mathfrak{g}$ and e is the identity element of G. Show that the derivative mapping of $\exp$ at O is the identity mapping in $\mathfrak{g}$.*

The results of the previous exercise guarantee that exp is a local diffeomorphism from a neighborhood of $O \in \mathfrak{g}$ to a neighborhood of $e \in \mathrm{G}$. This follows from the inverse function theorem. (See also [7] chapter IV.6 on this topic.)

The exponential mapping can be used to put local coordinates on the Lie group by means of the Lie algebra.

[2] One parameter subgroup: a subgroup depending on one parameter, in this case t.

Theorem 1.21. *Let* G *be a connected Lie group with Lie algebra* $\mathfrak{g}$. *Every group element can be written as a product of exponentials:*

$$g = \exp(V_1)\exp(V_2)\cdots\exp(V_k),$$

for $V_1, \ldots, V_k \in \mathfrak{g}$.

1.8 Some properties of the exponential in matrix Lie groups

We want to consider

$$\frac{d}{dt}\exp(\sigma(t))e\Big|_{t=0},$$

where $\sigma(t)$ is a curve in $\mathfrak{gl}(n)$. We proceed giving two Lemmas which are used for this aim.

Lemma 1.22. Variation of constants formula.
The solution of the differential equation

$$\dot{u} = Au + w, \quad u(0) = u_0,$$

where A *is a* $m \times m$ *constant matrix and* $u_0, w \in \mathbf{R}^m$ *are fixed, is*

$$u(t) = e^{tA}u_0 + \int_0^t e^{(t-x)A}w dx.$$

Proof. To find the solution of the considered differential equation we use the integrating factor e^{-xA}, we obtain

$$e^{-xA}\dot{u}(x) - Ae^{-xA}u(x) = e^{-xA}w.$$

We now integrate between 0 and t and obtain

$$e^{-tA}u(t) - u(0) = \int_0^t e^{-xA}w dx,$$

and multiplying on both sides with e^{tA} we obtain the result. □

Corollary 1.23. *If* $w \in \mathbf{R}^m$ *and* A *is a* $m \times m$ *matrix we have that*

$$\int_0^t e^{(t-x)A}w dx = \frac{e^{tz} - 1}{z}\Big|_{z=A} w.$$

Proof. We expand the integral at the left hand side of the equality by using the Taylor series of the exponential mapping and we obtain

$$\int_0^t e^{(t-x)A} w dx = \sum_{i=0}^{\infty} \int_0^t \frac{(t-x)^i}{i!} A^i w dx,$$

and since

$$\int_0^t \frac{(t-x)^i}{i!} w dx = \frac{t^{i+1}}{(i+1)!} w,$$

we obtain

$$\int_0^t e^{(t-x)A} w dx = \sum_{i=0}^{\infty} A^i \frac{t^{i+1}}{(i+1)!} w = \sum_{k=1}^{\infty} A^{k-1} \frac{t^k}{k!} w.$$

Now one can verify that

$$\frac{e^{tz}-1}{z} = \sum_{k=1}^{\infty} z^{k-1} \frac{t^k}{k!},$$

(use the expansion for e^{tz}), which implies that

$$\int_0^t e^{(t-x)A} w dx = \left. \frac{e^{tz}-1}{z} \right|_{z=A} w.$$

□

This Lemma is used in the proof of the next Lemma.

Lemma 1.24. *Assume $\sigma(t)$ is a $n \times n$ matrix for each t then we have that*

$$\left(\frac{d}{dt} e^{\sigma(t)} \right) e^{-\sigma(t)} = \left. \frac{e^z - 1}{z} \right|_{z=\mathrm{ad}_\sigma} (\dot{\sigma}), \tag{5}$$

where for two $n \times n$ matrices B and C we have $\mathrm{ad}_B(C) = [B, C] = BC - CB$, where $[\cdot, \cdot]$ is the matrix commutator.

Proof. Consider $B(s,t) = \left(\frac{d}{dt} e^{s\sigma(t)}\right) e^{-s\sigma(t)}$. By differentiating with respect to s we obtain

$$\begin{aligned}
\frac{\partial}{\partial s} B(s,t) &= \left(\frac{d}{dt} \left(\sigma(t) e^{s\sigma(t)} \right) \right) e^{-s\sigma(t)} - \left(\frac{d}{dt} e^{s\sigma(t)} \right) e^{-s\sigma(t)} \sigma(t) \\
&= \dot{\sigma}(t) e^{s\sigma(t)} e^{-s\sigma(t)} + \sigma(t) \left(\frac{d}{dt} e^{s\sigma(t)} \right) e^{-s\sigma(t)} - B(s,t)\sigma(t) \\
&= \dot{\sigma}(t) + [\sigma(t), B(s,t)].
\end{aligned}$$

This means that

$$\frac{\partial}{\partial s}B(s,t) \quad = \quad \mathrm{ad}_\sigma(B)+\dot{\sigma},$$

and we have $B(0,t)=O$. Note that ad_σ is a linear operator acting on $n\times n$ matrices, and can be represented as a $n^2\times n^2$ matrix. Then taking $A=\mathrm{ad}_{\sigma(t)}$, in Lemma 1.22 and Corollary 1.23 we have

$$B(s,t)=\left.\frac{e^{sz}-1}{z}\right|_{z=\mathrm{ad}_{\sigma(t)}}(\dot{\sigma}(t))\,.$$

□

From Lemma 1.24 we have that

$$\frac{d}{dt}e^{\sigma(t)}=\left.\frac{e^{z}-1}{z}\right|_{z=\mathrm{ad}_{\sigma}}(\dot{\sigma})\cdot e^{\sigma(t)},$$

and for ease of notation we define

$$\begin{aligned}\mathrm{dexp}_{\sigma(t)}\left(\dot{\sigma}(t)\right): &= \left.\frac{e^{z}-1}{z}\right|_{z=\mathrm{ad}_{\sigma}}(\dot{\sigma})\\ &= \sum_{k=1}^{\infty}\frac{1}{k!}\mathrm{ad}_{\sigma}^{k-1}\left(\dot{\sigma}(t)\right)\\ &= \dot{\sigma}(t)+\frac{1}{2}[\sigma(t),\dot{\sigma}(t)]+\frac{1}{3!}[\sigma(t),[\sigma(t),\dot{\sigma}(t)]]+\ldots\end{aligned}$$

2 Integration methods on manifolds

2.1 Introduction and motivation

We are interested in deriving *intrinsic* numerical integration methods for the problem

$$\dot{y} \quad = \quad F(y) \tag{6}$$

$$y(t_0) \quad = \quad y_0, \tag{7}$$

with $y_0\in\mathcal{M}$, $\mathcal{M}$ a smooth manifold, and F a vector field on $\mathcal{M}$, i.e. $F(y(t))\in T_{y(t)}\mathcal{M}$ for all t. Using a classical Runge-Kutta or multi-step method to approximate this problem does not make sense unless the manifold is embedded in a vector space of larger dimension where operations like sum and multiplication by scalar are well defined. Even when $\mathcal{M}$ is a sub-manifold of such a vector space a Runge-Kutta method would not produce approximations on $\mathcal{M}$ in general. Our aim is to design numerical methods which are applicable to ODEs on manifolds and by construction produce approximations on the manifold.

Example 2.1. *Consider the following differential equation on the orthogonal group*

$$\dot{Y} = A(Y) \cdot Y, \quad Y(0) = Y_0, \tag{8}$$

where Y and $A(Y)$ are $n \times n$ matrices, $A(Y)$ is skew-symmetric for all Y and Y_0 is an orthogonal matrix. The solution of (8) is an orthogonal matrix in fact if we take the derivative w.r.t. time of $Y(t)^T Y(t)$ we obtain

$$\begin{aligned} \frac{d}{dt} Y(t)^T Y(t) &= \dot{Y}^T Y + Y^T \dot{Y} = Y^T A(Y)^T Y + Y^T A(Y) Y \\ &= -Y^T A(Y) Y + Y^T A(Y) Y = 0, \end{aligned}$$

which means that $Y(t)^T Y(t)$ is constant and therefore

$$Y(t)^T Y(t) = Y_0^T Y_0 = I, \quad \forall t,$$

i.e. $Y(t)$ is an orthogonal matrix for all t. The format (8) is a consequence of the characterization of the tangent space discussed in example 1.15 and is valid in general for $A(Y)$ belonging to the Lie algebra $\mathfrak{g}$ of a Lie group G *and* $Y_0 \in \mathrm{G}$.

2.2 Methods based on frame vector fields

Definition 2.2. *A set of vector fields $\{\mathcal{E}_1, \ldots \mathcal{E}_d\}$ on the manifold $\mathcal{M}$ of dimension $m \leq d$ is a set of frame vector fields if*

$$T_x\mathcal{M} = \mathrm{span}\{\left.\mathcal{E}_1\right|_x, \ldots \left.\mathcal{E}_d\right|_x\}, \quad \forall\, x \in \mathcal{M}.$$

Given any vector field F on $\mathcal{M}$ we have

$$F(y) = \sum_{i=1}^{d} f_i(y) \mathcal{E}_i(y).$$

Definition 2.3. *We denote with F_p the vector field*

$$F_p(x) = \sum_{i=1}^{d} f_i(p) \mathcal{E}_i(x)$$

we say that F_p is the vector field F frosen at the point p.

Given at $\mathcal{M}$ is a manifold with a set of frame vector fields we can define intrinsic Runge-Kutta like methods as follows:

Commutator-free method

for $r = 1 : s$ do
  $Y_r = \exp(\sum_{k=1}^{s} \alpha_{rJ}^k F_k) \cdots \exp(\sum_{k=1}^{s} \alpha_{r1}^k F_k)(p)$
  $F_r = hF_{Y_r} = h \sum_{i=1}^{d} f_i(Y_r)\mathcal{E}_i$
end
$y_1 = \exp(\sum_{k=1}^{s} \beta_J^k F_k) \cdots \exp(\sum_{k=1}^{s} \beta_1^k F_k)p$

Here n counts the number of time steps and h is the step-size of integration. The integrator has s stages and parameters α_{rJ}^k, β_J^k. Each new stage value is obtained as a composition of J exponentials of linear combinations of vector fields frozen at the previously computed stage values.

In the following tableaus we report the coefficients of a method of order 3 and a method of order 4. The method of order 3 requires the computation of one exponential of each internal stage value and the composition of two exponentials for updating the solution. In the order 4 method the first three stage values require one exponential each, while the fourth stage and the solution update require two exponentials.

$$\begin{array}{c|ccc} 0 & & & \\ \frac{1}{3} & \frac{1}{3} & & \\ \frac{2}{3} & 0 & \frac{2}{3} & \\ \hline & \frac{1}{3} & 0 & 0 \\ & -\frac{1}{12} & 0 & \frac{3}{4} \end{array} \qquad \begin{array}{c|cccc} 0 & & & & \\ \frac{1}{2} & \frac{1}{2} & & & \\ \frac{1}{2} & 0 & \frac{1}{2} & & \\ 1 & \frac{1}{2} & 0 & 0 & \\ & -\frac{1}{2} & 0 & 1 & \\ \hline & \frac{1}{4} & \frac{1}{6} & \frac{1}{6} & -\frac{1}{12} \\ & -\frac{1}{12} & \frac{1}{6} & \frac{1}{6} & \frac{1}{4} \end{array}$$

Example 2.4. *Let $\mathcal{M}$ be a manifold acted upon transitively by a Lie group* G. *Denote with $\Lambda : \mathrm{G} \times \mathcal{M} \to \mathcal{M}$ the Lie group action. Suppose $E_1, \ldots, E_d$ a basis of the Lie algebra then $F_{E_1}, \ldots F_{E_d}$ obtained by*

$$F_{E_i}(x) = \frac{d}{dt}\Lambda(\exp(tE_i), x)\Big|_{t=0},$$

are a set of frame vector fields.

In particular for matrix Lie groups consider the vector field $A(Y)Y$ of equation (8). Here $A(y) \in \mathfrak{g}$ and $A(Y) = \sum_{i=1}^{d} a_i(y)E_i$ with $E_1, \ldots, E_d$ a basis of the Lie algebra[3]*. The vector field frosen at a point $P \in \mathrm{G}$ is simply $A(P)Y$.*

References

[1] E. Celledoni and A. Iserles, *Approximating the exponential from a Lie algebra to a Lie group.* Math. Comp. 69 (2000), no. 232, 1457–1480.

[2] E. Celledoni and A. Iserles, *Methods for the approximation of the matrix exponential in a Lie-algebraic setting.* IMA J. Numer. Anal. 21 (2001), no. 2, 463–488.

[3] E. Celledoni, A. Marthinsen and B. Owren, *Commutator-free Lie group methods*, FCGS, *19* (2003), 341–352.

[4] P.E. Crouch, and R. Grossman, *Numerical integration of ordinary differential equations on manifolds*, J. Nonlinear Sci. **3** (1993), 1–33.

[5] F. Diele and L. Lopez and R. Peluso *The Cayley Transform in the Numerical Solution of Unitary Differential Systems*, Journal of Appl. Num. Math., (1998) **8**: 317–334.

[6] E. Hairer, S.P. Nørsett and G. Wanner *Solving Ordinary Differential Equations I*, Springer series in Computational Mathematics, Springer, (2000), second edition.

[7] E. Hairer, C. Lubich and G. Wanner *Geometric Numerical Integration*, Springer series in Computational Mathematics, Springer, (2002), first edition.

[8] A. Iserles and A. Zanna *Efficient computation of the matrix exponential by Generalized Polar Decompositions.* SIAM J. Num. Anal., vol. 42, nr. 5, pp. 2218–2256, (2005).

[9] D. Lewis and J.C. Simo *Conserving algorithms for the dynamics of Hamiltonian systems on Lie groups.* J. Nonlinear Sci. **4** (1994), 253–299.

[10] J.E. Marsden and T.S. Ratiu, *An introduction to mechanics and symmetries*, Texts in Applied Mathematics vol. 17, Springer-Verlag, 1994.

[3]Say for $\mathfrak{so}(n)$ a basis is given by the matrices of rank 2 of the type $\mathbf{e}_i\mathbf{e}_j^T - \mathbf{e}_j\mathbf{e}_i^T$ with $\mathbf{e}_i, \mathbf{e}_j \in \mathbf{R}^n$ canonical vectors, and $i = 1, \ldots, n$, $j \leq i-1$.

[11] H. Munthe-Kaas, *High order Runge-Kutta methods on manifolds*, Appl. Num. Math., **29** (1999), 115–127.

[12] H. Munthe-Kaas and B. Owren, *Computations in a free Lie algebra* R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci. **357** (1999), 957–981.

[13] P. Olver, *Equivalence Invariants and symmetries* Cambridge University Press, Cambridge, (1995).

[14] B. Owren, *Order conditions for commutator-fee Lie group methods* J. of Phys. A: Math. Gen., **39** (2006), 5585–5599.

[15] B. Owren and A. Marthinsen, *Integration methods based on canonical coordinates of the second kind*. Numer. Math. (2001) **87**: 763–790.

[16] V.S. Varadarajan, An introduction to harmonic analysis on semisimple Lie groups. Cambridge Studies in Advanced Mathematics, 16. Cambridge University Press, Cambridge, 1989.