The foundations of Runge–Kutta–Munthe–Kaas methods

Brynjulf Owren

July 30, 2015

Introduction. This lecture can be divided into four parts

1. A generic format for differential equations on a manifold via a Lie group action
2. The derivative of the exponential map
3. Relatedness of vector fields
4. Runge–Kutta–Munthe–Kaas (RKMK) methods

Generic format for differential equations on manifolds. A significant part of the material presented here is adapted from [3], but see also [1, 2, 4]. Suppose M is a manifold, and G a Lie group acting on M, its Lie algebra is $\mathfrak{g}$. We use the notation $\Lambda : G \times M \to M$ for the action which should satisfy

$$\Lambda(e, m) = m, \forall m \in M \quad \text{and} \quad \Lambda(g, \Lambda(h, m)) = \Lambda(gh, m) \text{ (left Lie group action)}$$

Here $e \in G$ is the identity element of the group and g, h are arbitrary elements of G. We shall also assume that G is locally transitive, by this we mean that for every $m \in M$ there is an open neighborhood, $M \supset U \ni M$, containing m such that $\Lambda_m := \Lambda(\cdot, m)$ maps G onto U.

Let us pick an arbitrary $v_m \in T_m M$. This tangent vector is tangent to a curve $\gamma(t)$ through m, i.e. $\gamma(0) = m$, $\dot\gamma(0) = v_m$, and some segment of it can be represented by a curve $g(t)$ on G, i.e. $\gamma(t) = \Lambda(g(t), m)$, where $g(0) = e$ and $\dot g(0) = \xi \in \mathfrak{g}$. Therefore we have

$$v_m = T_e \Lambda_m(\xi), \quad \xi \in G.$$

The ξ need not be unique, but the local transitivity property guarantees that we can find at least one such element. Note in particular that the construction will also give us a whole vector field on M, namely

$$^1 F_\xi|_m = T_e \Lambda_m(\xi), \quad \text{we write } F_\xi = \lambda_*(\xi)$$

so we have defined a map that sends elements of $\mathfrak{g}$ to vector fields on M.

[1] In this note, we use the following notation: If $X \in \mathcal{X}(M)$ (smooth vector field on M), then $X|_m \in T_m M$ is the value of X at m, i.e. a vector in the tangent space $T_m M$ at m.

Example 1 You might recall a similar construction when $M = G$ is the Lie group itself. Then one can use right multiplication as a left Lie group action, i.e. $\Lambda(g, m) = g \cdot m = \Lambda_m(g) = R_m(g)$. In this case, the vector field $X_\xi|_m = \lambda_*(\xi)(m) = T_e R_m(\xi)$ is nothing else than a right invariant vector field on G. □

The vector fields F_ξ defined above are of a special type, they constitute a finite dimensional linear space. The space $\mathcal{X}(M)$ is infinite dimensional so an arbitrary $F \in \mathcal{X}$ cannot be expected to be of such a form as the F_ξ. However, if we can replace $\xi \in \mathfrak{g}$ by some (smooth) function, $\xi(m)$, i.e. $\xi : M \to \mathfrak{g}$ we can in principle "reach" any element of $\mathcal{X}(M)$, by

$$F|_m = \lambda_*(\xi(m))(m)$$

This way of representing a vector field on M by the Lie group action is what we shall call our *generic format*.

The derivative of the exponential map. Recall the exponential map $\exp : \mathfrak{g} \to G$, defined e.g. via the flow of left invariant vector fields on G. Here we shall find a formula for its derivative or tangent map, i.e. we shall consider, for any $u \in \mathfrak{g}$, $v \in \mathfrak{g}$

$$T_u \exp(v)$$

Thinking of the curves approach to tangent maps, we can take the tangent of the image under exp of curves $\gamma(s) = u + sv$, i.e.

$$\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \exp(u + sv) \in T_{\exp(u)}G$$

But now we recall that all tangent spaces of a Lie group can be identified with the tangent space at the identity element (i.e. the Lie algebra $\mathfrak{g}$) via left or right translations. For our purposes here, it turns out that right translation is the better choice. So composing $TR_{\exp(-u)}$ with $T_u \exp(v)$ we get an element of $T_e G \simeq \mathfrak{g}$. This leads to the important definition of $\mathrm{dexp}_u(v)$

$$\mathrm{dexp}_u(v) := TR_{\exp(-u)} \circ T_u \exp(v)$$

We shall avoid complicated notation in this note and it will be sufficient for our purposes to think of Lie group and Lie algebra elements as matrices, although the results hold in a more general setting.

We begin by defining the ad-operator and its powers, we have, for any u, v in $\mathfrak{g}$

$$\mathrm{ad}_u(v) := [u, v] = uv - vu, \quad \mathrm{ad}_u^2(v) = [u, [u, v]] \text{ etc.}$$

This is a linear operator on a finite dimensional space, and linear combinations are well defined, so we define

$$\exp(\mathrm{ad}_u)(v) = \sum_{k=0}^{\infty} \frac{1}{k!} \mathrm{ad}_u^k(v) \quad (\mathrm{ad}_u^0 v = v).$$

We have the following result to start with

Lemma 1.

$$T_e L_{\exp(v)} \circ T_e R_{\exp(-v)}(w) = \exp(v) w \exp(-v) = \exp(\mathrm{ad}_v)(w)$$

Proof. Let

$$y(t) = \exp(tv) w \exp(-tv)$$

$$z(t) = \exp(t\mathrm{ad}_v)(w)$$

We differentiate these two to find

$$\dot{y}(t) = v \exp(tv) w \exp(-tv) - \exp(tv) w \exp(-tv) v = [v, y], \quad y(0) = w$$

$$\dot{z}(t) = \mathrm{ad}_v(z) = [v, z], \quad z(0) = w$$

So $y(t)$ and $z(t)$ satisfy the same differential equation with the same initial value. We conclude that $y(t) \equiv z(t)$ and in particular that $y(1) = z(1)$. □

Lemma 2. *(Tangent map of* $\exp : \mathfrak{g} \to G$*). Let* $u \in \mathfrak{g}$*,* $v \in T_u\mathfrak{g} \simeq \mathfrak{g}$*. Then*

$$T_u \exp(v) = \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \exp(u + sv) = TR_{\exp(u)} \circ \mathrm{dexp}_u v = \mathrm{dexp}_u(v) \cdot \exp(u)$$

where

$$\mathrm{dexp}_u(v) = \int_0^1 \exp(r\mathrm{ad}_u)(v)\,\mathrm{d}r = \left.\frac{e^z - 1}{z}\right|_{z=\mathrm{ad}_u}(v)$$

Proof. Let $y_s(t) = \exp(t(u + sv))$ such that

$$T_u \exp(v) = \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} y_s(1)$$

But for now we differentiate with respect to t to obtain

$$\dot{y}_s := \frac{\mathrm{d}}{\mathrm{d}t} y_s(t) = (u + sv) y_s(t) \tag{1}$$

We also note that $y_s(t) = \exp(tu) + \mathcal{O}(s)$ as $s \to 0$. From (1) we then get

$$\dot{y} - u y_s = s v e^{tu} + \mathcal{O}(s^2)$$

and the integrating factor e^{-tu} yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-tu} y_s\right) = s e^{-tu} v e^{tu} + \mathcal{O}(s^2)$$

Integrating, and using that $y_s(0) = \mathrm{Id}$, we get

$$y_s(t) = e^{tu} + s \int_0^t e^{ru} v e^{-ru} e^{tu}\,\mathrm{d}r + \mathcal{O}(s^2)$$

and so

$$\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} y_s(1) = \int_0^1 e^{ru} v e^{-ru}\,\mathrm{d}r\, e^u = \int_0^1 e^{r\,\mathrm{ad}_u}(v)\,\mathrm{d}r\; e^u$$

where we have used Lemma 1 in the last equality. Formally, we can write

$$\int_0^1 e^{r\,\mathrm{ad}_u}(v)\,\mathrm{d}r = \int_0^1 e^{rz}|_{z=\mathrm{ad}_u}(v)\,\mathrm{d}r = \left.\frac{e^z-1}{z}\right|_{z=\mathrm{ad}_u}(v)$$

□

It is often useful to consider the dexp-map as an infinite series of nested commutators

$$\mathrm{dexp}_u(v) = (I + \frac{1}{2!}\mathrm{ad}_u + \frac{1}{3!}\mathrm{ad}_u^2 + \cdots)(v) = v + \frac{1}{2}\,[u,v] + \frac{1}{6}\,[u,[u,v]] + \cdots$$

In the integrators we consider, it turns out to be the inverse of the dexp-map which is most important. Note that the function

$$\phi_1(z) = \frac{e^z-1}{z}$$

is entire, i.e. analytic in all of $\mathbb{C}$, this means that its reciprocal

$$\frac{z}{e^z-1}$$

is analytic where $\phi_1(z) \neq 0$. In particular this means that $\frac{1}{\phi_1(z)}$ has a converging Taylor series about $z=0$ in the open disk $|z| < 2\pi$. This series expansion can be shown to be

$$\frac{z}{e^z-1} = 1 - \frac{z}{2} + \sum_{k=1}^{\infty} \frac{B_{2k}}{(2k)!} z^{2k}$$

where B_{2k} are the *Bernoulli numbers*, the first few of them are: $B_2 = \frac{1}{6}$, $B_4 = -\frac{1}{30}$, $B_6 = \frac{1}{42}$, $B_8 = -\frac{1}{30}$, $B_{10} = \frac{5}{66}$. The map

$$v = \mathrm{dexp}_u^{-1}(w) \qquad (\text{whenever } w = \mathrm{dexp}_u(v))$$

is given precisely as

$$\mathrm{dexp}_u^{-1}(w) = \left.\frac{z}{e^z-1}\right|_{z=\mathrm{ad}_u}(w) = w - \frac{1}{2}[u,w] + \frac{B_2}{2!}[u,[u,w]] + \cdots \tag{2}$$

ϕ-relatedness of vector fields. Let $\phi : M \to N$ be a smooth map between the manifolds M and N. Let $X \in \mathcal{X}(M)$ and $Y \in \mathcal{X}(N)$ be vector fields on these two manifolds respectively. If

$$X[f \circ \phi] = Y[f] \circ \phi, \quad \forall f \in \mathcal{F}(N) \tag{3}$$

we say that X and Y are ϕ-related, and we write $X \sim_\phi Y$. Here $\mathcal{F}(N)$ is the ring of smooth functions from N to $\mathbb{R}$. Recall that when we use the square bracket, it means that we apply the vector field as a differential operator as explained earlier, $X[f] = \mathrm{d}f(X)$. The *pull-back* ϕ^* applied to a function on N

transforms it to a function on M and is defined as $\phi^* f = f \circ \phi$, so this means we can write simply

$$X[\phi^* f] = \phi^* Y[f] \quad \forall f \in \mathcal{F}(N).$$

A less elegant, but to us, more useful formulation of ϕ-relatedness is obtained by rewriting (3) as

$$\mathrm{d}(f \circ \phi)(X) = \mathrm{d}f \circ T\phi \circ X = \mathrm{d}f \circ Y \circ \phi, \quad \forall f \in \mathcal{F}(N)$$

which is true if and only if

$$T\phi \circ X = Y \circ \phi$$

The purpose for us in introducing ϕ-relatedness lies in the fact that if $X \sim_\phi Y$, then ϕ maps the flow of X to the flow of Y. By just defining

$$y(t) = \phi(x(t)) = \phi \circ \varphi_{t,X}(x_0) = \phi \circ \exp(tX)x_0$$

we compute

$$\dot{y}(t) = T\phi \circ \dot{x}(t) = T\phi \circ X(x(t)) = Y \circ \phi \circ x(t) = Y(y(t))$$

so that $y(t) = \phi(x(t))$ is indeed the flow of Y.

Runge–Kutta–Munthe–Kaas (RKMK) methods. We now put together what we have learned so far in this note to uncover the main principle behind the RKMK methods. In the next lemma, we shall define locally a map between $\mathfrak{g}$ and M called λ_m and then construct a vector field η on the Lie algebra $\mathfrak{g}$ of the acting group G, such that $\eta \sim_{\lambda_m} F$ where F is the ODE vector field on the manifold. The we can apply standard methods, in this case Runge–Kutta methods to the ODE on $\mathfrak{g}$ and map it back to M via λ_m. We make this rigorous through the following lemma

Lemma 3. *[3] Let M be a smooth manifold with a left Lie group action $\Lambda : G \times M \to M$, let $\mathfrak{g}$ be the Lie algebra of $\mathfrak{g}$. Let $\lambda_* : \mathfrak{g} \to \mathcal{X}(M)$ be defined by*

$$\lambda_*(\xi)|_m = T_e\Lambda_m(\xi), \quad \Lambda_m(g) = \Lambda(g, m).$$

Suppose $F \in \mathcal{X}(M)$ is of the (generic) form

$$F|_m = \lambda_*(\xi(m))(m), \quad \text{for some } \xi : M \to \mathfrak{g}.$$

Fix $m \in M$ and set $\lambda_m(u) = \Lambda(\exp(u), m)$. Then there is an open set $U \subseteq \mathfrak{g}$ containing 0 such that the vector field $\eta \in \mathcal{X}(U)$ defined as

$$\eta|_u = \mathrm{dexp}_u^{-1}(\xi \circ \lambda_m(u))$$

is λ_m-related to F.

Proof. We need to show that $T\lambda_m \circ \eta = F \circ \lambda_m$. To simplify the writing a little, we write $x := \lambda_m(u) \in M$. So we first compute

$$\begin{aligned} T\lambda_m \circ \eta|_u &= T(\Lambda_m \circ \exp) \circ \eta = T\Lambda_m \circ TR_{\exp(u)} \circ \operatorname{dexp}_u \circ \operatorname{dexp}_u^{-1} \circ \xi(x) \\ &= T\Lambda_m \circ TR_{\exp(u)} \circ \xi(x) \end{aligned}$$

where we have used Lemma 2. Then we compute for any u, v in $\mathfrak{g}$

$$\begin{aligned} \lambda_*(v)|_x &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Lambda(\exp(tv), \Lambda(\exp(tu), m)) \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} = T(\Lambda_m \circ R_{\exp(u)})(v) = T\Lambda_m \circ TR_{\exp(u)}(v) \end{aligned}$$

This formula is used to compute

$$\begin{aligned} (F \circ \lambda_m)|_u &= \lambda_*(\xi(m) \circ \lambda_m(u))(\lambda_m(u)) \\ &= T\Lambda_m \circ TR_{\exp(u)}(\xi(m) \circ \lambda_m(u)) = T\lambda_m \circ \eta|_u \end{aligned}$$

□

We now present the general idea of RKMK methods through 3 steps.

1. Given $y_n \in M$, set $m := y_n$, thus $m = \lambda_m(0)$.
2. Take one step with a standard Runge–Kutta method on the problem

   $$\dot{u} = \eta|_u = \operatorname{dexp}_u^{-1}(\xi \circ \lambda_m(u)), \quad u(0) = u_0 = 0.$$

   this results in $u_1 \in \mathfrak{g}$.
3. Update by setting $y_{n+1} = \lambda_m(u_1)$.

Remarks.

A. $\operatorname{dexp}_u^{-1} v$ does not have to be computed exactly. Since $u_0 = 0$ the method only applies the function to quantities $u = \mathcal{O}(h)$. This means that

$$\operatorname{ad}_u^k(v) = [\mathcal{O}(h), [\mathcal{O}(h), \cdots [\mathcal{O}(h), v] \cdots] = \mathcal{O}(h^k)$$

and therefore the series for $\operatorname{dexp}_u^{-1}(v)$ can be truncated after p terms for a pth order method without compromising the order of convergence. Note also that any truncation will belong to the Lie algebra $\mathfrak{g}$. The truncated version of $\operatorname{dexp}_u^{-1(v)}$ after p terms is denoted $\operatorname{dexpinv}(u, v, p)$.

B. If the underlying classical Runge-Kutta method has order p then the method will also be of order p for Lie groups. The reason is that the map λ_m is smooth and therefore

$$\lambda_m(\varphi_{\eta,h}(u_0) + \mathcal{O}(h^{p+1})) = \lambda_m(\varphi_{h,\eta}(u_0)) + \mathcal{O}(h^{p+1}) = \varphi_{F,h}(m) + \mathcal{O}(h^{p+1})$$

the last equality owing to the fact that $\eta \sim_{\lambda_m} F$.

C. Let us now write the RKMK algorithm in a form that is possible to implement in a computer program. The following algorithm takes one step with a Runge-Kutta methods having coefficients (a_{ij}, b_i).

$m = y_0$
for $i = 1, 2, \ldots, s$ **do**
$u_i = h \sum_{j=1}^{i-1} a_{ij} \tilde{k}_j$
$k_i = \xi(\lambda_m(u_i))$
$\tilde{k}_i = \text{dexpinv}(u_i, k_i, p)$
end
$v = h \sum_{j=1}^{s} b_j \tilde{k}_j$
$y_1 = \lambda_m(v)$

References

[1] Elena Celledoni, Håkon Marthinsen, and Brynjulf Owren. An introduction to Lie group integrators—basics, new developments and applications. *J. Comput. Phys.*, 257(part B):1040–1061, 2014.

[2] Snorre H. Christiansen, Hans Z. Munthe-Kaas, and Brynjulf Owren. Topics in structure-preserving discretization. *Acta Numer.*, 20:1–119, 2011.

[3] Hans Munthe-Kaas. High order Runge-Kutta methods on manifolds. In *Proceedings of the NSF/CBMS Regional Conference on Numerical Analysis of Hamiltonian Differential Equations (Golden, CO, 1997)*, volume 29, pages 115–127, 1999.

[4] Hans Munthe-Kaas and Brynjulf Owren. Computations in a free Lie algebra. *R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci.*, 357(1754):957–981, 1999.