More formal symplectic integration

In the neighborhood of a reference trajectory $\vec{\mathbf{X}} = \widehat{\mathbf{X}}(s)$, we can expand the equation of motion about $\widehat{\mathbf{X}}(s)$. Equation of the reference trajectory becomes:

$$\frac{d\widehat{X}_i}{ds} = \sum_{j=1}^{6} S_{ij}\frac{\partial H}{\partial X_j}(\widehat{\mathbf{X}}), \qquad \text{or in matrix notation:} \quad \frac{d\widehat{\mathbf{X}}}{ds} = \mathbf{S}\,\nabla_6 H.$$

Expanding both sides in Taylor series yields

$$\begin{aligned}\frac{d}{ds}(\widehat{X}_i + \Delta X_i) &= \sum_{j=1}^{6} S_{ij}\frac{\partial H}{\partial X_j}(\widehat{\mathbf{X}} + \Delta\mathbf{X}) \\ &= \sum_{j=1}^{6} S_{ij}\left[\frac{\partial H}{\partial X_j}(\widehat{\mathbf{X}}) + \sum_{k=1}^{6}\frac{\partial^2 H}{\partial X_j \partial X_k}(\widehat{X})\Delta\mathbf{X}_k + \cdots\right].\end{aligned}$$

$$\frac{d\Delta X_i}{ds} = \sum_{j=1}^{6}\sum_{k=1}^{6} S_{ij}\frac{\partial^2 H}{\partial X_j \partial X_k}(\widehat{X})\Delta\mathbf{X}_k + \cdots.$$

The matrix of second derivatives $C_{jk} = \frac{\partial^2 H}{\partial X_j \partial X_k}(\widehat{X}) = C_{kj}$.

In matrix notation, our 1st order equation is

$$\frac{d\Delta\widehat{\mathbf{X}}}{ds} = \mathbf{SC}\,\Delta\widehat{\mathbf{X}}. \qquad \text{(Type of equation Sophus Lie invented his algebra for.)}$$

So for an infinitesimal step (from s_ν to $s_{\nu+1}$), we have

$$\Delta\widehat{\mathbf{X}}_{\nu+1} = \Delta\widehat{\mathbf{X}}_\nu + \Delta\widehat{\mathbf{X}}_\nu\,\mathbf{SC}\,ds = (\mathbf{I} + \mathbf{G}\,ds)\,\Delta\widehat{\mathbf{X}}_\nu.$$

For the case where $\mathbf{C}$ is constant, the integration gives

$$\mathbf{M}(s) = \lim_{n\to\infty}\left(\mathbf{I} + \mathbf{SC}\,\frac{s}{n}\right)^n = e^{\mathbf{SC}\,s}.$$

If $\mathbf{C}$ is not constant then we must have something more like

$$\mathbf{M}(s) = \lim_{ds\to 0} e^{\mathbf{SC}(s-ds)ds}\cdots e^{\mathbf{SC}(2ds)ds}e^{\mathbf{SC}(ds)ds}e^{\mathbf{SC}(0)ds}.$$

How do we approximate this?

A general $2n\times 2n$-symmetric matrix has

$$\frac{(2n)^2-2n}{2}+2n=(2n+1)n$$

degrees of freedom. Since a $n\times n$-symplectic matrix can be written as the exponentiation of $\mathbf{SC}$, the symplectic matrices also have $(2n+1)n$ free parameters.

n	$2n$	d.o.f.
1	2	3
2	4	10
3	6	21
4	8	36

For example any 4×4 symmetric real matrix can be written as

$$\mathbf{S}=\sum_{j=1}^{10}\alpha_j\mathbf{c}_j,$$

where the α_j are real coefficients and the 10 $\mathbf{c}_j$ form a *basis set* of the 4×4 symmetric matrices.

For the 4×4 symmetric matrices, one possible basis is

$$\mathbf{c}_j : \begin{pmatrix} 0&0&0&0\\0&1&0&0\\0&0&0&0\\0&0&0&1 \end{pmatrix}, \begin{pmatrix} 0&0&0&0\\0&1&0&0\\0&0&0&0\\0&0&0&-1 \end{pmatrix}, \begin{pmatrix} 1&0&0&0\\0&0&0&0\\0&0&1&0\\0&0&0&0 \end{pmatrix}, \begin{pmatrix} 1&0&0&0\\0&0&0&0\\0&0&-1&0\\0&0&0&0 \end{pmatrix},$$

$$\begin{pmatrix} 0&1&0&0\\1&0&0&0\\0&0&0&0\\0&0&0&0 \end{pmatrix}, \begin{pmatrix} 0&0&1&0\\0&0&0&0\\1&0&0&0\\0&0&0&0 \end{pmatrix}, \begin{pmatrix} 0&0&0&1\\0&0&0&0\\0&0&0&0\\1&0&0&0 \end{pmatrix}, \begin{pmatrix} 0&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&0 \end{pmatrix},$$

$$\begin{pmatrix} 0&0&0&0\\0&0&0&1\\0&0&0&0\\0&1&0&0 \end{pmatrix}, \begin{pmatrix} 0&0&0&0\\0&0&0&0\\0&0&0&1\\0&0&1&0 \end{pmatrix}, \quad \text{in the order of } j \text{ from 1 to 10.}$$

The ten products $G_j = \mathbf{S}\mathbf{c}_j$ form a set of generators 4×4 symplectic matrices which may be written in the form

$$\exp\left(\sum_{j=1}^{10} \mathbf{G}_j \alpha_j \right) .$$

The product of two exponentials of different matrices can be combined into a single exponential of a third matrix:

$$e^{\mathbf{X}}e^{\mathbf{Y}} = e^{\mathbf{Z}}.$$

If $[\mathbf{X},\mathbf{Y}] = 0$, then we have simply

$$e^{(\mathbf{X}+\mathbf{Y})} = e^{\mathbf{X}}\, e^{\mathbf{Y}}.$$

However if the matrices $\mathbf{X}$ and $\mathbf{Y}$ do not commute, then $\mathbf{Z}$ can be calculated from the the Baker-Campbell-Hausdorff (BCH) formula:

$$\begin{aligned}\mathbf{Z} &= \log\left(e^{\mathbf{X}}e^{\mathbf{Y}}\right)\\ &= \mathbf{X}+\mathbf{Y}+\frac{1}{2}[\mathbf{X},\mathbf{Y}]+\frac{1}{12}([\mathbf{X},\mathbf{X},\mathbf{Y}]+[\mathbf{Y},\mathbf{Y},\mathbf{X}])+\frac{1}{24}[\mathbf{X},\mathbf{Y},\mathbf{Y},\mathbf{X}]+\mathcal{O}(5),\end{aligned}$$

where the extended commutator notation indicates multiple commutators nested to the right:

$$[a,b,c] = [a,[b,c]], \qquad [a,b,c,d] = [a,[b,[c,d]]].$$

The Zassenhaus formula (sort of like the reverse process of the BCH formula) splits a single exponentiation of the sum of two matrices **A** and **B** into a product of two exponentials of the two matrices times higher order exponentials of commutators:

$$e^{(\mathbf{A}+\mathbf{B})h} = e^{\mathbf{A}h}\, e^{\mathbf{B}h}\, e^{-[\mathbf{A},\mathbf{B}]h^2/2}\, e^{(2[\mathbf{B},\mathbf{A},\mathbf{B}]+[\mathbf{A},\mathbf{A},\mathbf{B}])h^3/6}\, e^{\mathcal{O}(h^4)\cdots}\cdots,$$

where the parameter h is a small integration step.

To second order in h, this can be written as (see Problem 3–10)

$$e^{(\mathbf{A}+\mathbf{B})h} = e^{\mathbf{A}h/2} e^{\mathbf{B}h} e^{\mathbf{A}h/2} + \mathcal{O}(h^3). \tag{ZH2}$$

To second order an integration step $e^{\mathbf{SC}h}$ can be expanded into

$$\exp\left(\sum_{j=1}^{n} \alpha_j \mathbf{G}_j h\right) = \left(\prod_{j=1}^{n} \exp(\mathbf{G}_j h/2)\right)\left(\prod_{j=n}^{1} \exp(\mathbf{G}_j h/2)\right) + \mathcal{O}(h^3),$$

by successive applications of Eq. (ZH2).

If we partition the sum in $\exp\left(\sum_{j=1}^{10} \mathbf{G}_j \alpha_j\right)$ into

$$\mathbf{A} = \alpha_1 \mathbf{G}_1, \quad \text{and} \quad \mathbf{B} = \sum_{n=2}^{10} \alpha_n \mathbf{G}_n,$$

then the factor

$$e^{\mathbf{A}h/2} = \begin{pmatrix} 1 & \alpha_1 h/2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \alpha_1 h/2 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \mathcal{O}(h^3),$$

produces a drift matrix of length $\alpha_1 h/2$ up to second order in h, and the factor $e^{\mathbf{B}h}$ corresponds to a thin element kick.

The end result is a thin kick sandwiched between to drifts.

- Drift-Kick codes are symplectic. (e. g. Teapot by Talman and Schachinger)
    - Kicks can even be of higher order (nonlinear).
- Nonlinear kicks with drifts are a standard way to treat sextupoles, octopoles, etc. even in codes with thick lens elements.

“Time-symmetric” integrators

If we apply the BCH formula twice to the time-symmetric* product

$$e^{\mathbf{W}} = e^{\mathbf{X}h}\, e^{\mathbf{Y}h}\, e^{\mathbf{X}h},$$

then it must be that

$$\mathbf{W} = (2\mathbf{X}+\mathbf{Y})h + \frac{1}{6}\left([\mathbf{Y},\mathbf{Y},\mathbf{X}] - [\mathbf{X},\mathbf{X},\mathbf{Y}]\right)h^3 + \mathcal{O}(h^5).$$

If an integrator formula $I(h) = e^{\mathbf{g}_1 h + \mathbf{g}_2 h^2 + \mathbf{g}_3 h^3 + \cdots}$, with matrices $\mathbf{g}_j$ is “time reversible”, then we must have $I(h)I(-h) = 1$. The BCH formula gives to lowest order:

$$I(h)I(-h) = e^{-\mathbf{g}_1 h/2} e^{\mathbf{g}_2 h^2} e^{-\mathbf{g}_1 h/2} \quad e^{\mathbf{g}_1 h/2} e^{\mathbf{g}_2 h^2} e^{\mathbf{g}_1 h/2} + \mathcal{O}(h^3) = 1.$$

$$e^{\mathbf{g}_1 h/2} e^{-\mathbf{g}_1 h/2} = 1 = e^{\mathbf{g}_2 h^2} + \mathcal{O}(h^3).$$

So we must have $\mathbf{g}_2 = 0$.

---

* This symmetry is typically referred to as time-symmetric even when the integration variable may be the s-coordinate rather than time, since s is the independent time-like parameter of the Hamiltonian.

Repeating this with $\mathbf{g}_2 = 0$, now the lowest term would require that $\mathbf{g}_4 = 0$.

By induction, all the even powers of h in $I(h)$ vanish: $\mathbf{g}_{2j} = 0$.

I. e. $\log(I(h))$ must be an odd function of h, if $I(h)$ is time-reversible.

Higher order integrators may be constructed from the second order integration function,

$$I_2(h) = e^{\mathbf{A}h/2}\, e^{\mathbf{B}h}\, e^{\mathbf{A}h/2} = e^{\mathbf{g}_1 h + \mathbf{g}_3 h^3 + \cdots}.$$

Yoshida constructed a 4^{th}-order integrator from a time-symmetric product of second order integrators:

$$I_4(h) = I_2(ah)\, I_2(bh)\, I_2(ah),$$

where a and b are parameters to be determined.

$$I_4(h) = e^{(2a+b)\mathbf{g}_1 h + (2a^3+b^3)\mathbf{g}_3 h^3 + \cdots}, = e^{(\mathbf{A}+\mathbf{B})h + \mathcal{O}(h^5)}.$$

To 4^{th}-order this requires that $\quad 1 = 2a + b, \quad$ and $\quad 0 = 2a^3 + b^3.$

$$a = \frac{1}{2 - 2^{1/3}}, \qquad \text{and} \qquad b = -\frac{2^{1/3}}{2 - 2^{1/3}}.$$ (Note typo in book on p. 61.)

Groups

A *group* is a set G with a binary operation on its elements having the properties:

i. for any two elements $a, b \in G$, then $ab \in G$;
ii. if $a,\ b,\ c \in G$, then $a(bc) = (ab)c$;
iii. there is a unique element $e \in G$ such that $ea = a = ae$ for any element $a \in G$;
iv. for each $a \in G$ there is an element $a^{-1} \in G$ such that $a^{-1}a = e = aa^{-1}$.

Familiar examples:

1. the integers Z with the addition operator.
2. *general linear group* $\mathrm{Gl}(n, \mathbb{R})$ of $n \times n$-square real matrices with nozero determinant.
3. *special linear group* $\mathrm{Sl}(n, \mathbb{R})$ of $n \times n$-square real matrices with unit determinant.
4. *orthogonal group* of rotations $\mathrm{O}(n, \mathbb{R})$ (Includes reflections.)
5. *special unitary group* $\mathrm{SU}(n, \mathbb{C})$.
6. *symplectic group* $\mathrm{Sp}(2n, \mathbb{R})$. (Representation depends on choice of **S**.)

Lie algebras

A *Lie algebra* is a vector space over a field (for our case either the real numbers $\mathbb{R}$ or complex numbers $\mathbb{C}$) with an additional binary operation $[\cdot,\cdot]$ called the *Lie bracket* or *commutator*. The Lie bracket operator satisfies the following properties for any elements x, y, z in the Lie algebra and a, b in the field:

1. Bilinearity:
   $$[ax+by, z] = a[x,z] + b[y,z], \qquad \text{(correction to “muddlement” in §3.8.1)}$$
   $$[z, ax+by] = a[z,x] + b[z,y];$$
2. Anticommutativity:
   $$[x,y] = -[y,x];$$
3. Jacobi Identity:
   $$[x,[y,z]] + [y,[z,x]] + [z,[x,y]] = 0,$$

   with 0 here being the identity element in the vector space.

Lie group

A *Lie group* is basically a group which is also a *differentiable manifold.*

For an example consider the 6-d surface defined by a conserved Hamiltonian:

$$\mathcal{H}(x, x', y, y', z, \delta; s) = -\frac{qA_s}{p_0} - \frac{x}{\rho} - \frac{x\delta}{\rho} + \frac{1}{2}(w_x^2 + w_y^2) + \cdots = \text{ a constant.}$$

Given a particle with initial position and momentum lying on this surface, the particle's trajectory will remain on this surface.

As we have seen, the transport matrices are elements of the group Sp(6,$\mathbb{R}$).

Quadratic Lie Groups

A *quadratic Lie group* G_c defined by a group representation of $n\times n$ nonsingular (i. e., the inverse must exist) complex matrices:

$$G_c = \{\mathbf{M} \in \mathrm{GL}_n(\mathbb{C}) : \mathbf{MJM}^\dagger = \mathbf{J}\},$$

where $\mathrm{GL}_n(\mathbb{C})$ is general linear group of complex $n\times n$ matrices, and where $\mathbf{J}$ is any particular matrix in $\mathrm{GL}_n(\mathbb{C})$.

Note that the dagger indicates the Hermitian conjugate: $\mathbf{M}^\dagger = (\mathbf{M}^*)^\mathrm{T}$, i. e. the transpose of the complex conjugate of the matrix.

The corresponding *Lie algebra* can be represented by

$$\mathfrak{g}_c = \{\mathbf{A} \in \mathbb{C}^{n\times n} : \mathbf{AJ} + \mathbf{JA}^\dagger = \mathbf{0}\},$$

where $\mathbb{C}^{n\times n}$ is the set of all $n\times n$ complex matrices.

If we consider only real matrices then

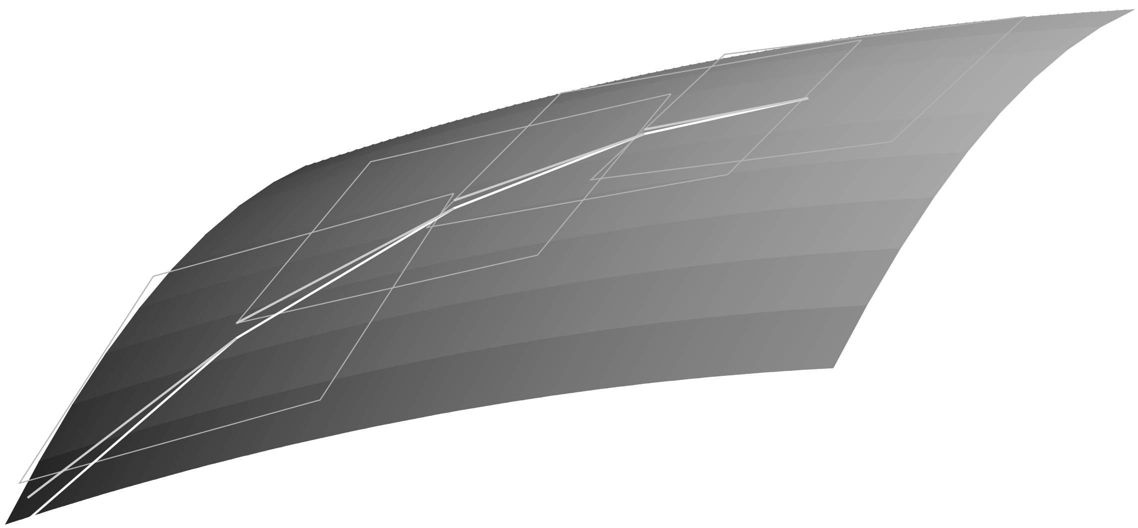
$$G_{\rm r} = \{\mathbf{M} \in \mathrm{GL}_n(\mathbb{R}) : \mathbf{MJM}^{\rm T} = \mathbf{J}\}, \quad \text{and}$$
$$\mathfrak{g}_{\rm r} = \{\mathbf{A} \in \mathbb{R}^{n\times n} : \mathbf{AJ} + \mathbf{JA}^{\rm T} = \mathbf{0}\},$$

where $\mathrm{GL}_n(\mathbb{R})$ is the general linear group of $n\times n$ nonsingular real matrices and $\mathbf{J}$ is a particular matrix in $\mathrm{GL}_n(\mathbb{R})$.

Some examples other than the symplectic group $\mathrm{Sp}(2n, \mathbb{R})$:

- The unitary group: $\mathrm{U}(n)$ with $\mathbf{J} = \mathbf{I}$.
  - The special unitary group restricted to have $|\mathbf{M}| = 1$.
- The orthogonal group: $\mathrm{O}(n) \in U(n)$ restricted to real matrices with $\mathbf{J} = \mathbf{I}$.
  - The special orthogonal group restricted to have $|\mathbf{M}| = 1$.
- The Lorentz group $\mathrm{SO}(3, 1, \mathbb{R})$ with $\mathbf{J} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$.

- $\mathrm{Sp}(2n, \mathbb{R}\mathbf{S})$ describes the symplectic geometry of trajectories flowing on the manifold's surface.
- The corresponding Lie algebra $\mathfrak{sp}(2n,\mathbb{R};\mathbf{S})$ describes the geometry of a plane tangent to the surface at a point.
  - It approximates a linear neighborhood of the surface around the point.



Four integration steps for a trajectory with four integration steps in tangent planes and the resulting integrated trajectory projected back onto the manifold of the Hamiltonian by a *lift* function Φ.

Lift function:

$$\Phi \quad : \mathfrak{sp}(2n,\mathbb{R};\mathbf{S}) \to \mathrm{Sp}(2n, \mathbb{R})$$

$$\Phi^{-1} : \mathrm{Sp}(2n, \mathbb{R}) \to \mathfrak{sp}(2n,\mathbb{R};\mathbf{S})$$

The obvious choice for Φ is an exponential map.

Cayley transforms

Cayley showed that an orthogonal matrix $\mathbf{Q}$ could be factored as

$$\mathbf{Q} = (\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A})^{-1},$$

where $\mathbf{A}$ is antisymmetric $\mathbf{A} = -\mathbf{A}^{\mathrm{T}}$.

$$\mathbf{A} = (\mathbf{I} - \mathbf{Q})(\mathbf{I} + \mathbf{Q})^{-1}.$$

Orthogonalization of an almost orthogonal matrix

Suppose we have an almost (but not quite) orthogonal matrix $\mathbf{Q}_0$.

1. Calculate $\mathbf{A}_0 = (\mathbf{I} - \mathbf{Q}_0)(\mathbf{I} + \mathbf{Q}_0)^{-1}$. $\mathbf{A}_0$ will not be quite antisymmetric.
2. Calculate $\mathbf{A}_1 = (\mathbf{A}_0 - \mathbf{A}_0^{\mathrm{T}})/2$. Now $\mathbf{A}_1^{\mathrm{T}} = -\mathbf{A}_1$.
3. Calculate $\mathbf{Q}_1 = (\mathbf{I} - \mathbf{A}_1)(\mathbf{I} + \mathbf{A}_1)^{-1}$. $\mathbf{Q}_1$ is orthogonal and close to $\mathbf{Q}_0$.

Healy's symplectification algorithm

A symplectic matrix $\mathbf{M}$ may be written in the form*

$$\mathbf{M} = (\mathbf{I} + \mathbf{SW})(\mathbf{I} - \mathbf{SW})^{-1},$$

if and only if $\mathbf{W}$ is a symmetric matrix.

Inverse transform: $\mathbf{W} = \mathbf{S}(\mathbf{I} + \mathbf{M})^{-1}(\mathbf{I} - \mathbf{M}).$

The symplectification algorithm

If $\mathbf{M}_0$ is almost symplectic, then calculate

1. $\mathbf{W}_0 = \mathbf{S}(\mathbf{I} + \mathbf{M}_0)^{-1}(\mathbf{I} - \mathbf{M}_0),$
2. $\mathbf{W}_1 = (\mathbf{W}_0 + \mathbf{W}_0^{\mathrm{T}})/2,$
3. $\mathbf{M}_1 = (\mathbf{I} + \mathbf{SW}_1)(\mathbf{I} - \mathbf{SW}_1)^{-1}.$

$\mathbf{M}_1$ will be an approximation of $\mathbf{M}_0$ which is now symplectic.

---

* Provided that $|\mathbf{I} - \mathbf{SW}| \neq 0$ and $|\mathbf{I} + \mathbf{M}| \neq 0$.