BNL-99365-2013-TECH
C-A/AP/215;BNL-99365-2013-IR

Thoughts on Healys Symplectification Algorithm

W. MacKay. W.

August 2005

Collider Accelerator Department
Brookhaven National Laboratory

U.S. Department of Energy
USDOE Office of Science (SC)

DISCLAIMER

C-A/AP/#215
August 2005

Thoughts on Healy's Symplectification Algorithm

Waldo MacKay

Collider-Accelerator Department
Brookhaven National Laboratory
Upton, NY 11973

8 Aug., 2005

Thoughts on Healy's Symplectification Algorithm

Waldo MacKay

A nice algorithm for tweaking an almost symplectic matrix into a symplectic matrix has been given by Healy in his thesis[1] according to the literature. (I have not seen a copy of his thesis, so I am not sure of his actual wording.) Before discussing the algorithm and its limitations a lemma and theorem will first be proven.

Lemma: Given two square matrices $\mathbf{S}$ and $\mathbf{W}$ of the same rank with $\mathbf{S}^2 = -\mathbf{I}$ where $\mathbf{I}$ is the identity matrix then

$$(\mathbf{I}-\mathbf{WS})\mathbf{S}(\mathbf{I}+\mathbf{SW}) = (\mathbf{I}+\mathbf{WS})\mathbf{S}(\mathbf{I}-\mathbf{SW}). \tag{1}$$

Proof:

$$\begin{aligned}(\mathbf{I}-\mathbf{WS})\mathbf{S}(\mathbf{I}+\mathbf{SW}) &= (\mathbf{S}+\mathbf{W})(\mathbf{I}+\mathbf{SW})\\ &= \mathbf{S}+\mathbf{W}+\mathbf{S}^2\mathbf{W}+\mathbf{WSW})\\ &= \mathbf{S}-\mathbf{W}-\mathbf{S}^2\mathbf{W}+\mathbf{WSW})\\ &= (\mathbf{S}-\mathbf{W})(\mathbf{I}-\mathbf{SW})\\ &= (\mathbf{I}+\mathbf{WS})\mathbf{S}(\mathbf{I}-\mathbf{SW}).\end{aligned}$$

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

$$\mathbf{M} = (\mathbf{I}+\mathbf{SW})(\mathbf{I}-\mathbf{SW})^{-1}, \tag{2}$$

if and only if $\mathbf{W}$ is a symmetric matrix, and where $\mathbf{S}$ is the metric for the selected representation of the symplectic group. This statement must be qualified with the requirement that

$$|\mathbf{I}-\mathbf{SW}| \neq 0.$$

A discussion of the restrictions will be given at the end.

In accelerator physics we usually require $\mathbf{S}$ to be a block diagonal $2n\times 2n$ matrix with

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

in the diagonal blocks. It is worth noting that $\mathbf{S}$ has the properties

$$\mathbf{S}^{\mathrm{T}} = \mathbf{S}^{-1} = -\mathbf{S}.$$

Proof:

Show that if $\mathbf{W}$ is symmetric, then $\mathbf{M}$ is symplectic:

$$\begin{aligned}
\mathbf{M}^{\mathrm{T}}\mathbf{S}\mathbf{M} &= (\mathbf{I}-\mathbf{W}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}})^{-1}(\mathbf{I}+\mathbf{W}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}})\mathbf{S}(\mathbf{I}+\mathbf{S}\mathbf{W})(\mathbf{I}-\mathbf{S}\mathbf{W})^{-1}\\
&= (\mathbf{I}+\mathbf{W}\mathbf{S})^{-1}(\mathbf{I}-\mathbf{W}\mathbf{S})\mathbf{S}(\mathbf{I}+\mathbf{S}\mathbf{W})(\mathbf{I}-\mathbf{S}\mathbf{W})^{-1}\\
&= (\mathbf{I}+\mathbf{W}\mathbf{S})^{-1}(\mathbf{I}+\mathbf{W}\mathbf{S})\mathbf{S}(\mathbf{I}-\mathbf{S}\mathbf{W})(\mathbf{I}-\mathbf{S}\mathbf{W})^{-1}\\
&= \mathbf{S}.
\end{aligned}$$

Therefore $\mathbf{M}$ is symplectic if $\mathbf{W}$ is symmetric.

Now let us assume that $\mathbf{W}$ is not symmetric, so it can be written as the sum of symmetric and antisymmetric matrices:

$$\mathbf{W} = \mathbf{P} + \mathbf{Q}$$

where $\mathbf{P} = \mathbf{P}^{\mathrm{T}}$ and $\mathbf{Q} = -\mathbf{Q}^{\mathrm{T}}$.

$$\begin{aligned}
\mathbf{M}^{\mathrm{T}}\mathbf{S}\mathbf{M} = \mathbf{S} &= (\mathbf{I}-\mathbf{W}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}})^{-1}(\mathbf{I}+\mathbf{W}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}})\mathbf{S}(\mathbf{I}+\mathbf{S}\mathbf{W})(\mathbf{I}-\mathbf{S}\mathbf{W})^{-1}\\
&= (\mathbf{I}-\mathbf{W}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}})^{-1}(\mathbf{S}+\mathbf{P}-\mathbf{Q})(\mathbf{I}+\mathbf{S}\mathbf{P}+\mathbf{S}\mathbf{Q})(\mathbf{I}-\mathbf{S}\mathbf{W})^{-1}\\
&= (\mathbf{I}-\mathbf{W}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}})^{-1}(\mathbf{S}+\mathbf{P}-\mathbf{Q}-\mathbf{P}-\mathbf{Q}+\mathbf{W}^{\mathrm{T}}\mathbf{S}\mathbf{W})(\mathbf{I}-\mathbf{S}\mathbf{W})^{-1}\\
&= (\mathbf{I}-\mathbf{W}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}})^{-1}[(\mathbf{S}-\mathbf{P}+\mathbf{Q}+\mathbf{P}+\mathbf{Q}+\mathbf{W}^{\mathrm{T}}\mathbf{S}\mathbf{W})-4\mathbf{Q}](\mathbf{I}-\mathbf{S}\mathbf{W})^{-1}\\
&= (\mathbf{I}-\mathbf{W}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}})^{-1}[(\mathbf{S}-\mathbf{W}^{\mathrm{T}}-\mathbf{S}^2\mathbf{W}+\mathbf{W}^{\mathrm{T}}\mathbf{S}\mathbf{W})-4\mathbf{Q}](\mathbf{I}-\mathbf{S}\mathbf{W})^{-1}\\
&= (\mathbf{I}-\mathbf{W}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}})^{-1}[(\mathbf{S}-\mathbf{W}^{\mathrm{T}})(\mathbf{I}-\mathbf{S}\mathbf{W})-4\mathbf{Q}](\mathbf{I}-\mathbf{S}\mathbf{W})^{-1}\\
&= (\mathbf{I}-\mathbf{W}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}})^{-1}[(\mathbf{I}-\mathbf{W}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}})\mathbf{S}(\mathbf{I}-\mathbf{S}\mathbf{W})-4\mathbf{Q}](\mathbf{I}-\mathbf{S}\mathbf{W})^{-1}\\
&= \mathbf{S}-4(\mathbf{I}-\mathbf{W}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}})^{-1}\mathbf{Q}(\mathbf{I}-\mathbf{S}\mathbf{W})^{-1}.
\end{aligned}$$

So assuming that the inverses in the last line exist then $\mathbf{Q} = 0$. (Actually the inverse $(\mathbf{I}-\mathbf{W}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}})^{-1}$ must exist if its transpose $(\mathbf{I}-\mathbf{S}\mathbf{W})^{-1}$ exists.) This proves the theorem.

Given the symplectic matrix $\mathbf{M}$, form a new matrix

$$\mathbf{V} = \mathbf{S}(\mathbf{I}-\mathbf{M})(\mathbf{I}+\mathbf{M})^{-1} \qquad (3).$$

Then

$$\begin{aligned}
\mathbf{V}+\mathbf{V}\mathbf{M} &= \mathbf{S}-\mathbf{S}\mathbf{M}.\\
(\mathbf{S}+\mathbf{V})\mathbf{M} &= \mathbf{S}-\mathbf{V}\\
\mathbf{M} &= (\mathbf{S}+\mathbf{V})^{-1}(\mathbf{S}-\mathbf{V}).
\end{aligned}$$

Taking the transpose gives

$$\mathbf{M}^{\mathrm{T}} = (\mathbf{S}^{\mathrm{T}}-\mathbf{V}^{\mathrm{T}})(\mathbf{S}^{\mathrm{T}}+\mathbf{V}^{\mathrm{T}})^{-1},$$

and we can calculate the inverse via

$$\begin{aligned}\mathbf{M}^{-1} = \mathbf{S}\mathbf{M}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}} &= \mathbf{S}(\mathbf{S}^{\mathrm{T}} - \mathbf{V}^{\mathrm{T}})(\mathbf{S}^{\mathrm{T}} + \mathbf{V}^{\mathrm{T}})^{-1}\mathbf{S}^{\mathrm{T}} \\ &= (\mathbf{I} - \mathbf{S}\mathbf{V})(\mathbf{I} + \mathbf{S}\mathbf{V})^{-1}.\end{aligned}$$

Inverting this yields

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

which is identical in form to Eq. 2.

The symplectification algorithm

The symplectification algorithm for an almost symplectic matrix $\mathbf{M}$ is to calculate $\mathbf{V}$ by the above Eq. 3 (assuming that $|\mathbf{I} - \mathbf{M}| \neq 0$), then create a symmetric matrix

$$\mathbf{W} = \frac{\mathbf{V} + \mathbf{V}^{\mathrm{T}}}{2} \tag{4}$$

which then may be used to calculate a new matrix

$$\mathbf{M}' = (\mathbf{I} + \mathbf{S}\mathbf{W})(\mathbf{I} - \mathbf{S}\mathbf{W})^{-1},$$

assuming that $|\mathbf{I} - \mathbf{S}\mathbf{W}| \neq 0$. This new matrix $\mathbf{M}'$ must be symplectic by the previous theorem, and it should be close to the original matrix $\mathbf{M}$.

Problems with the method occur in constructing $\mathbf{V}$ when

$$|\mathbf{I} + \mathbf{M}| = 0.$$

This will happen when $\mathbf{M}$ has at least one eigenvalue equal to -1. If $|\mathbf{I} - \mathbf{M}| \neq 0$, then we can define the new almost symmetric matrix by

$$\widehat{\mathbf{V}} = \mathbf{S}(\mathbf{I} + \mathbf{M})(\mathbf{I} - \mathbf{M})^{-1} \tag{3$'$}$$

with

$$\widehat{\mathbf{W}} = \frac{\widehat{\mathbf{V}} + \widehat{\mathbf{V}}^{\mathrm{T}}}{\mathbf{2}}. \tag{4$'$}$$

$$\mathbf{M}' = -(\mathbf{I} + \mathbf{S}\widehat{\mathbf{W}})(\mathbf{I} - \mathbf{S}\widehat{\mathbf{W}})^{-1}. \tag{2$'$}$$

Now we must have $|\mathbf{I} - \mathbf{S}\widehat{\mathbf{W}}| \neq 0$, and $|\mathbf{I} - \mathbf{M}| \neq 0$.

If $\mathbf{M}$ has at least one eigenvalue equal to $+1$, and another equal to -1 then

$$|\mathbf{I} - \mathbf{M}| = |\mathbf{I} + \mathbf{M}| = 0.$$

and this method may not work.

For an example of this, we must be considering a matrix for at least two planes with at least four eigenvalues, since the symplectic matrix must have pairs of eigenvalues equal to 1 and -1. The matrix

$$\mathbf{M} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

is symplectic and obviously has $|\mathbf{I} \pm \mathbf{M}| = 0$, so we cannot hope to construct a symmetric $\mathbf{V}$ in this case.

Consider the perturbation of this matrix

$$\mathbf{M} = \begin{pmatrix} 1+\delta & 0 & 0 & 0 \\ 0 & 1-\delta & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

then

$$\begin{aligned} \widehat{\mathbf{V}} &= \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2+\delta & 0 & 0 & 0 \\ 0 & 2-\delta & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -\delta & 0 & 0 & 0 \\ 0 & \delta & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 0 & \delta-2 & 0 & 0 \\ \delta+2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \frac{1}{2\delta} \begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & \delta & 0 \\ 0 & 0 & 0 & \delta \end{pmatrix} \\ &= \frac{1}{\delta} \begin{pmatrix} 0 & \delta-2 & 0 & 0 \\ -\delta-2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

While for nonzero values of δ this exists and is symmetric, the limit of $\widehat{\mathbf{V}}$ blows up as δ

goes to zero, however

$$\widehat{\mathbf{W}} = \frac{2}{\delta}\begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\mathbf{S}\widehat{\mathbf{W}} = \frac{2}{\delta}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\mathbf{M}' = -\begin{pmatrix} \frac{\delta+2}{\delta} & 0 & 0 & 0 \\ 0 & \frac{\delta-2}{\delta} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} \frac{\delta-2}{\delta} & 0 & 0 & 0 \\ 0 & \frac{\delta+2}{\delta} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}^{-1}$$

$$\mathbf{M}' = \begin{pmatrix} \frac{2+\delta}{2-\delta} & 0 & 0 & 0 \\ 0 & \frac{2-\delta}{2+\delta} & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad \text{and}$$

$$\lim_{\delta\to 0}\mathbf{M}' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

as expected.

Consider a different perturbation of the matrix $\mathbf{M}$:

$$\mathbf{M} = \begin{pmatrix} 1+\delta & 0 & 0 & 0 \\ 0 & 1+\delta & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

then

$$\begin{aligned} \widehat{\mathbf{V}} &= \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 2+\delta & 0 & 0 & 0 \\ 0 & 2+\delta & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} -\delta & 0 & 0 & 0 \\ 0 & -\delta & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 0 & -\delta-2 & 0 & 0 \\ \delta+2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\frac{1}{2\delta}\begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & \delta & 0 \\ 0 & 0 & 0 & \delta \end{pmatrix} \\ &= \frac{1}{\delta}\begin{pmatrix} 0 & \delta+2 & 0 & 0 \\ -\delta-2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Again for nonzero values of δ this exists, but is antisymmetric so that

$$\widehat{\mathbf{W}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

This leads to $\mathbf{M}' = -\mathbf{I}$, so

$$\lim_{\delta\to 0} \mathbf{M}' = \mathbf{I} \neq \mathbf{M}$$

which might be unexpected, and is quite different from the original unperturbed matrix.

Comment on an error in Ref. [3].

In Eq. 14.13 of Ref. [3], Iselin states that a symplectic matrix $\mathbf{F} = \exp(\mathbf{SG})$ with a symmetric matrix $\mathbf{G}$ can be written in the form

$$\mathbf{F} = [\mathbf{I} + \tanh(\mathbf{SG}/2)][\mathbf{I} - \tanh(\mathbf{SG}/2)]^{-1} = (\mathbf{I} + \mathbf{W})(\mathbf{I} - \mathbf{W})^{-1}, \qquad \text{(I14.13)}$$

where $\mathbf{W}$ is symmetric if and only if F is symplectic. This far right-hand side is incorrect, and is probably a typo. He should have replaced $\mathbf{W}$ by $\mathbf{SW}$ in this equation. The middle part of the equation is correct in most cases and basically comes from

$$e^x = \frac{\cosh\frac{x}{2} + \sinh\frac{x}{2}}{\cosh\frac{x}{2} - \sinh\frac{x}{2}} = \left(1 + \tanh\frac{x}{2}\right)\left(1 - \tanh\frac{x}{2}\right)^{-1} \qquad (5)$$

and the fact that Hamilton's equations may be written in the form

$$\frac{d\mathbf{X}}{ds} = -\mathbf{SCX} = \mathbf{SGX},$$

where

$$C_{ij} = C_{ji} = \frac{\partial^2 H}{\partial X_i \partial X_j}.$$

Hamilton's equations give the general form of the generators for this matrix representation of the symplectic group Sp(2n, r) with the metric $\mathbf{S}$. For real x, Eq. 5 is analytic since $|\tanh(x/2)| < 1$, however for complex x the hyperbolic tangent can take on values of 1, so that Eq. 5 has poles. In the case where $x = \mathbf{SG}$ is a generator of a symplectic matrix, then the modified equation becomes

$$e^{\mathbf{SG}} = [\mathbf{I} + \tanh(\mathbf{SG}/2)][\mathbf{I} - \tanh(\mathbf{SG}/2)]^{-1}, \qquad (6)$$

and this factorization will not work when the matrix $\tanh(\mathbf{SG}/2)$ has an eigenvalue equal to 1. We should also note that since $\tanh(x) = -\tanh(-x)$ is an odd function it can be expanded as

$$\tanh(x) = \sum_{j=0}^{\infty} A_j x^{2j+1},$$

so that

$$
\begin{aligned}
\tanh\left(\frac{\mathbf{SG}}{2}\right)\mathbf{S} &= \sum_{j=0}^{\infty} A_j \frac{(\mathbf{SG})^{2j+1}\mathbf{S}}{2} = \mathbf{S}\tanh\left(\frac{\mathbf{GS}}{2}\right) \\
&= \sum_{j=0}^{\infty} A_j \frac{(\mathbf{SG})^{2j+1}\mathbf{S}}{2}(-1)^{2j+2} \\
&= \left[\tanh\left(\frac{\mathbf{SG}}{2}\right)\mathbf{S}\right]^{\mathrm{T}} = \left[\mathbf{S}\tanh\left(\frac{\mathbf{GS}}{2}\right)\right]^{\mathrm{T}}
\end{aligned}
$$

From this it should be obvious that the last part of Eq. I14.13 should have been written as

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

for symmetric $\mathbf{W}$.

Acknowledgments:

I would like to thank Christoph Montag, Yun Luo, and Alfredo Luccio for helpful discussions.

References

1. L. M. Healy, *Lie Algebraic Methods for Treating Lattice Parameter Errors in Particle Accelerators*, Doctoral thesis, University of Maryland, unpublished (1986).
2. David Sagan, *The Bmad Reference Manual*, Rev. 3.6, (2004)
3. F. Christoph Iselin, *The MAD Program Physical Methods Manual*, CERN/SL/92-?? (AP), unfinished report (1994).