COMMENT ON HEALY'S SYMPLECTIFICATION ALGORITHM*

W. W. MacKay†, BNL, Upton, NY 11973, USA

Abstract

For long-term tracking, it is important to have symplectic maps for the various electromagnetic elements in an accelerator ring. While many standard elements are handled well by modern tracking programs, new magnet configurations (e.g., a helical dipole with a superimposed solenoid[1]) are being used in real accelerators. Transport matrices and higher terms may be calculated by numerical integration through model-generated or measured field maps. The resulting matrices are most likely not quite symplectic due to numerical errors in the integrators as well as the field maps. In his thesis[2], Healy presented a simple algorithm to symplectify a matrix. While the method is quite robust, this paper presents a discussion of its limitations.

INTRODUCTION

A nice algorithm for tweaking an almost symplectic matrix into a symplectic matrix has been given by Healy in his thesis[2]. In order to understand the limitations of the method, it is worthwhile to present a derivation of the method, particularly since it has sometimes been quoted incorrectly[3].

INVERSION FORMULAE

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

$$\begin{aligned}(\mathbf{I}-\mathbf{WS})\mathbf{S}(\mathbf{I}+\mathbf{SW}) &= (\mathbf{S}+\mathbf{W})(\mathbf{I}+\mathbf{SW}) \\ &= (\mathbf{S}-\mathbf{W})(\mathbf{I}-\mathbf{SW}) \\ &= (\mathbf{I}+\mathbf{WS})\mathbf{S}(\mathbf{I}-\mathbf{SW}). \end{aligned} \quad (1)$$

A square $2n\times 2n$ real matrix, $\mathbf{M}$, is symplectic in a particular representation of the group $\mathrm{Sp}(2n,r)$ with respect to the metric $\mathbf{S}$, if it satisfies

$$\mathbf{M}^{\mathrm{T}}\mathbf{S}\mathbf{M} = \mathbf{S}. \quad (2)$$

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

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad (3)$$

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

$$\mathbf{S}^{\mathrm{T}} = \mathbf{S}^{-1} = -\mathbf{S}, \quad \text{and} \quad \mathbf{S}^2 = -\mathbf{I}. \quad (4)$$

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

$$\mathbf{M} = (\mathbf{I}+\mathbf{SW})(\mathbf{I}-\mathbf{SW})^{-1}, \quad (5)$$

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

$$|\mathbf{I}-\mathbf{SW}| \neq 0. \quad (6)$$

A discussion of the restrictions will be given later.

Proof:

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

$$\begin{aligned}&\mathbf{M}^{\mathrm{T}}\mathbf{S}\mathbf{M} \\ &= (\mathbf{I}-\mathbf{W}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}})^{-1}(\mathbf{I}+\mathbf{W}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}})\mathbf{S} \\ &\quad \times (\mathbf{I}+\mathbf{SW})(\mathbf{I}-\mathbf{SW})^{-1} \\ &= (\mathbf{I}+\mathbf{WS})^{-1}(\mathbf{I}-\mathbf{WS})\mathbf{S}(\mathbf{I}+\mathbf{SW})(\mathbf{I}-\mathbf{SW})^{-1} \\ &= (\mathbf{I}+\mathbf{WS})^{-1}(\mathbf{I}+\mathbf{WS})\mathbf{S}(\mathbf{I}-\mathbf{SW})(\mathbf{I}-\mathbf{SW})^{-1} \\ &= \mathbf{S}. \end{aligned} \quad (7)$$

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:

$$\mathbf{W} = \mathbf{P} + \mathbf{Q}, \quad (8)$$

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

$$\begin{aligned}&\mathbf{M}^{\mathrm{T}}\mathbf{S}\mathbf{M} = \mathbf{S} \\ &= (\mathbf{I}-\mathbf{W}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}})^{-1}(\mathbf{I}+\mathbf{W}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}})\mathbf{S}(\mathbf{I}+\mathbf{SW})(\mathbf{I}-\mathbf{SW})^{-1} \\ &= (\mathbf{I}-\mathbf{W}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}})^{-1}(\mathbf{S}+\mathbf{P}-\mathbf{Q}-\mathbf{P}-\mathbf{Q}+\mathbf{W}^{\mathrm{T}}\mathbf{S}\mathbf{W}) \\ &\quad \times (\mathbf{I}-\mathbf{SW})^{-1} \\ &= (\mathbf{I}-\mathbf{W}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}})^{-1}[(\mathbf{S}-\mathbf{W}^{\mathrm{T}})(\mathbf{I}-\mathbf{SW})-4\mathbf{Q}] \\ &\quad \times (\mathbf{I}-\mathbf{SW})^{-1} \\ &= \mathbf{S} - 4(\mathbf{I}-\mathbf{W}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}})^{-1}\mathbf{Q}(\mathbf{I}-\mathbf{SW})^{-1}. \end{aligned} \quad (9)$$

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 $(\mathbf{I}-\mathbf{SW})^{-1}$ exists.) This proves the theorem.

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

$$\mathbf{V} = \mathbf{S}(\mathbf{I}-\mathbf{M})(\mathbf{I}+\mathbf{M})^{-1}. \quad (10)$$

Then

$$\begin{aligned}\mathbf{V}+\mathbf{VM} &= \mathbf{S}-\mathbf{SM}. \\ (\mathbf{S}+\mathbf{V})\mathbf{M} &= \mathbf{S}-\mathbf{V} \\ \mathbf{M} &= (\mathbf{S}+\mathbf{V})^{-1}(\mathbf{S}-\mathbf{V}). \end{aligned} \quad (11)$$

* Work performed under the United States Department of Energy Contract No. DE-AC02-98CH1-886, and Renaissance Technologies Corp. (USA)

† waldo@bnl.gov

Taking the transpose gives

$$\mathbf{M}^{\mathrm{T}} = (\mathbf{S}^{\mathrm{T}} - \mathbf{V}^{\mathrm{T}})(\mathbf{S}^{\mathrm{T}} + \mathbf{V}^{\mathrm{T}})^{-1}, \tag{12}$$

and we can calculate the inverse via

$$\mathbf{M}^{-1} = \mathbf{S}\mathbf{M}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}} = \mathbf{S}(\mathbf{S}^{\mathrm{T}} - \mathbf{V}^{\mathrm{T}})(\mathbf{S}^{\mathrm{T}} + \mathbf{V}^{\mathrm{T}})^{-1}\mathbf{S}^{\mathrm{T}} = (\mathbf{I} - \mathbf{S}\mathbf{V})(\mathbf{I} + \mathbf{S}\mathbf{V})^{-1}. \tag{13}$$

Inverting this yields

$$\mathbf{M} = (\mathbf{I} + \mathbf{S}\mathbf{V})(\mathbf{I} - \mathbf{S}\mathbf{V})^{-1}, \tag{14}$$

which is identical in form to Eq. 5, so $\mathbf{V} = \mathbf{W}$.

THE SYMPLECTIFICATION ALGORITHM

The symplectification algorithm for an almost symplectic matrix $\mathbf{M}$ is to calculate $\mathbf{V}$ by the above Eq. 10 (assuming that $|\mathbf{I} - \mathbf{M}| \neq 0$), then create a symmetric matrix

$$\mathbf{W} = \frac{\mathbf{V} + \mathbf{V}^{\mathrm{T}}}{2} \tag{15}$$

which then may be used to calculate a new matrix

$$\mathbf{M}' = (\mathbf{I} + \mathbf{S}\mathbf{W})(\mathbf{I} - \mathbf{S}\mathbf{W})^{-1}, \tag{16}$$

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 arise in constructing $\mathbf{V}$ when M is an "exceptional" matrix[4], i. e., $|\mathbf{I} + \mathbf{M}| = 0$. This will happen when $\mathbf{M}$ has at least one eigenvalue equal to -1. If $|\mathbf{I} - \mathbf{M}| \neq 0$, then since $-\mathbf{M}$ must also be symplectic, we can define the new almost symmetric matrix by

$$\widehat{\mathbf{V}} = \mathbf{S}(\mathbf{I} + \mathbf{M})(\mathbf{I} - \mathbf{M})^{-1} \tag{17}$$

with

$$\widehat{\mathbf{W}} = \frac{\widehat{\mathbf{V}} + \widehat{\mathbf{V}}^{\mathrm{T}}}{2}. \tag{18}$$

$$\mathbf{M}' = -(\mathbf{I} + \mathbf{S}\widehat{\mathbf{W}})(\mathbf{I} - \mathbf{S}\widehat{\mathbf{W}})^{-1}. \tag{19}$$

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

BREAKDOWN OF THE METHOD

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

$$|\mathbf{I} - \mathbf{M}| = |\mathbf{I} + \mathbf{M}| = 0. \tag{20}$$

and this method may not work.

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

$$\mathbf{M} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{21}$$

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 following perturbation of this matrix

$$\mathbf{M} = \begin{pmatrix} 1+\delta & 0 & 0 & 0 \\ 0 & 1-\delta & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{22}$$

Then

$$\widehat{\mathbf{V}} = \mathbf{S} \begin{pmatrix} 2+\delta & 0 & 0 & 0 \\ 0 & 2-\delta & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -\delta & 0 & 0 & 0 \\ 0 & \delta & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}^{-1} = \frac{1}{\delta} \begin{pmatrix} 0 & \delta-2 & 0 & 0 \\ -\delta-2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{23}$$

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

$$\widehat{\mathbf{W}} = \frac{2}{\delta} \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \text{and} \tag{24}$$

$$\mathbf{M}' = \begin{pmatrix} \frac{2+\delta}{2-\delta} & 0 & 0 & 0 \\ 0 & \frac{2-\delta}{2+\delta} & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{25}$$

Taking the limit gives back the unperturbed matrix

$$\lim_{\delta \to 0} \mathbf{M}' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \tag{26}$$

as expected.

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

$$\mathbf{M} = \begin{pmatrix} 1+\delta & 0 & 0 & 0 \\ 0 & 1+\delta & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \tag{27}$$

then

$$\widehat{\mathbf{V}} = \mathbf{S} \begin{pmatrix} 2+\delta & 0 & 0 & 0 \\ 0 & 2+\delta & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -\delta & 0 & 0 & 0 \\ 0 & -\delta & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}^{-1} = \frac{1}{\delta} \begin{pmatrix} 0 & \delta+2 & 0 & 0 \\ -\delta-2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{28}$$

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

$$\widehat{\mathbf{W}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (29)$$

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

$$\lim_{\delta \to 0} \mathbf{M}' = \mathbf{I} \neq \mathbf{M}, \quad (30)$$

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

$$\begin{aligned} \mathbf{F} &= [\mathbf{I} + \tanh(\mathbf{SG}/2)][\mathbf{I} - \tanh(\mathbf{SG}/2)]^{-1} \\ &= (\mathbf{I} + \mathbf{W})(\mathbf{I} - \mathbf{W})^{-1}, \quad \Leftarrow \text{Incorrect} \end{aligned} \quad (31)$$

where $\mathbf{W}$ is symmetric if and only if F is symplectic. This far second line is incorrect, and is probably just a typo. (Ref. [3] is only available in an unfinished draft, so we must be careful when using it as a source.) 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

$$e^x = \frac{\cosh\frac{x}{2} + \sinh\frac{x}{2}}{\cosh\frac{x}{2} - \sinh\frac{x}{2}} = \left(1 + \tanh\frac{x}{2}\right)\left(1 - \tanh\frac{x}{2}\right)^{-1} \quad (32)$$

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

$$\frac{d\mathbf{X}}{ds} = -\mathbf{SCX} = \mathbf{SGX}, \quad (33)$$

where

$$C_{ij} = C_{ji} = \frac{\partial^2 H}{\partial X_i \partial X_j}. \quad (34)$$

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. 32 is analytic since $|\tanh(x/2)| < 1$, however for complex x the hyperbolic tangent can take on values of 1, so that Eq. 32 has poles. In the case where $x = \mathbf{SG}$ is a generator of a symplectic matrix, then the modified equation becomes

$$e^{\mathbf{SG}} = [\mathbf{I} + \tanh(\mathbf{SG}/2)][\mathbf{I} - \tanh(\mathbf{SG}/2)]^{-1}, \quad (35)$$

and this factorization will not work when $\tanh(\mathbf{SG}/2)$ 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

$$\tanh(x) = \sum_{j=0}^{\infty} A_j x^{2j+1}, \quad (36)$$

so that

$$\begin{aligned} \tanh\left(\frac{\mathbf{SG}}{2}\right)\mathbf{S} &= \sum_{j=0}^{\infty} A_j \frac{(\mathbf{SG})^{2j+1}\mathbf{S}}{2} = \mathbf{S}\tanh\left(\frac{\mathbf{GS}}{2}\right) \\ &= \sum_{j=0}^{\infty} A_j \frac{(\mathbf{SG})^{2j+1}\mathbf{S}}{2}(-1)^{2j+2} \\ &= \left[\tanh\left(\frac{\mathbf{SG}}{2}\right)\mathbf{S}\right]^{\mathrm{T}} = \left[\mathbf{S}\tanh\left(\frac{\mathbf{GS}}{2}\right)\right]^{\mathrm{T}} \end{aligned} \quad (37)$$

From this it should be obvious that the last part of Eq. 31 should have been written as

$$(\mathbf{I} + \mathbf{SW})(\mathbf{I} - \mathbf{SW})^{-1} \quad (38)$$

for symmetric $\mathbf{W}$.

SU(N) AND SO(N) MATRICES

As an aside, it is perhaps worth mentioning similar inversion formulae for special unitary and orthogonal matrices. A special unitary matrix $\mathbf{M}$ must satisfy the formula

$$\mathbf{M}^\dagger \mathbf{M} = \mathbf{I}, \quad (39)$$

where the dagger represents the complex conjugate of the matrix. The corresponding inversion formulae are

$$\mathbf{V} = (\mathbf{I} - \mathbf{M})(\mathbf{I} + \mathbf{M})^{-1} \quad \text{and} \quad (40)$$

$$\mathbf{M} = (\mathbf{I} + \mathbf{V})^{-1}(\mathbf{I} - \mathbf{V}), \quad (41)$$

where $\mathbf{V}$ is antihermitian: $\mathbf{V}^\dagger = -\mathbf{V}$. An almost antihermitian matrix $\mathbf{V}$ may be tweaked into an antihermitian matrix via the equation

$$\widehat{\mathbf{V}} = \frac{\mathbf{V} - \mathbf{V}^\dagger}{2}. \quad (42)$$

Special orthogonal matrices must satisfy the same formulae since they form a real subgroup of the special unitary matrices; in this case, the dagger just becomes the transpose operator.

ACKNOWLEDGMENTS

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

REFERENCES

[1] Erich Willen et al., “Superconducting Helical Snake Magnet for the AGS”, PAC2005, 2935 (2005).

[2] L. M. Healy, ”Lie Algebraic Methods for Treating Parameter Errors in Particle Accelerators”, Doctoral Thesis. University of Maryland, unpublished (1986).

[3] F. Christoph Iselin, *The MAD Program Physical Methods Manual*, CERN/SL/92-?? (AP), unfinished report (1994).

[4] Hermann Weyl, *The Classical Groups*, Princeton, NJ (1946).

[5] David Sagan, *The Bmad Reference Manual*, Rev. 3.6, (2004).