24 March, 2010

Calculating a closed orbit position and a response matrix

W. W. MacKay

1. Closed orbit calculation

Consider a ring with a kick at a corrector and a position measured elsewhere at a BPM such as indicated in Fig. 1. Note that I use $\mathbf{M}$ for my transport matrices rather than $\mathbf{R}$ as is used in MAD. For simplicity I ignore any transverse coupling or path-length changes in the ring due to the applied kick and just work with a simple linear model employing 2×2 dimensional matrices.

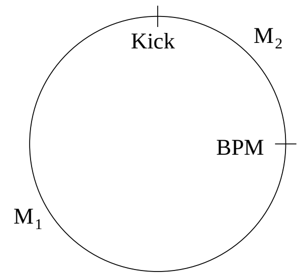

Figure 1. Schematic of a ring with a corrector (kick) and BPM located as shown. Propagating in a clockwise direction, $\mathbf{M}_1$ and $\mathbf{M}_2$ are the respective transport matrices from the BPM to the to the kick and the kick back around to the BPM.

The transport matrix for a complete turn starting at the BPM, I label as

$$\mathbf{M} = \mathbf{M}_2\mathbf{M}_1 = \begin{pmatrix} \cos\mu + \alpha_x \sin\mu & \beta_x \sin\mu \\ -\gamma_x \sin\mu & \cos\mu - \alpha_x \sin\mu \end{pmatrix},$$

with the Twiss parameters at the BPM and phase advance $\mu = 2\pi Q_h$ where $\mu = 2\pi Q_h$ with Q_h being the horizontal betatron tune. For a kick of angle θ applied by the dipole steering magnet, we can determine the effect on a closed orbit at the BPM from the equation

$$\mathbf{M}_2 \left[\mathbf{M}_1 \begin{pmatrix} x_c \\ x_c' \end{pmatrix} + \begin{pmatrix} 0 \\ \theta \end{pmatrix} \right] = \begin{pmatrix} x_c \\ x_c' \end{pmatrix}.$$

Solving for the closed orbit, we can write

$$\begin{pmatrix} x_c \\ x_c' \end{pmatrix} = [\mathbf{I} - \mathbf{M}]^{-1} \mathbf{M}_2 \begin{pmatrix} 0 \\ \theta \end{pmatrix}.$$

For the special case on 2×2 matrices (uncoupled transverse planes)

$$\begin{aligned} \begin{pmatrix} x_c \\ x_c' \end{pmatrix} &= \frac{1}{2(1-\cos\mu)} \left[\mathbf{I} - \mathbf{M}_1^{-1}\mathbf{M}_2^{-1} \right] \mathbf{M}_2 \begin{pmatrix} 0 \\ \theta \end{pmatrix} \\ &= \frac{1}{2(1-\cos\mu)} \left[\mathbf{M}_2 - \mathbf{M}_1^{-1} \right] \begin{pmatrix} 0 \\ \theta \end{pmatrix}, \end{aligned}$$

I realize that there was perhaps a big leap there from the previous equation, so let me show a neat little theorem about 2×2 matrices.

Writing a general 2×2 matrix $\mathbf{N}$ as $\mathbf{N} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, the inverse of $\mathbf{I} - \mathbf{N}$ is

$$\begin{aligned} (\mathbf{I} - \mathbf{N})^{-1} &= \begin{pmatrix} 1-a & -b \\ -c & 1-d \end{pmatrix}^{-1} = \frac{1}{(1-a)(1-d) - bc} \begin{pmatrix} 1-d & b \\ c & 1-a \end{pmatrix} \\ &= \frac{1}{1 + |\mathbf{N}| - \mathrm{tr}(\mathbf{N})} \left(\mathbf{I} - |\mathbf{N}|\, \mathbf{N}^{-1} \right), \end{aligned}$$

where $\mathrm{tr}(\mathbf{N}) = a + d$ is the trace of the matrix. In our case, we are using symplectic matrices so $|\mathbf{N}| = 1$ and we get

$$(\mathbf{I} - \mathbf{M})^{-1} = \frac{1}{2 - \mathrm{tr}(\mathbf{M})}(\mathbf{I} - \mathbf{M}^{-1}) = \frac{1}{2(1 - \cos\mu)}(\mathbf{I} - \mathbf{M}^{-1}).$$

2. Response matrix

Consider making m measurements of the difference orbit (written as the vector $\mathbf{Y}$) and using n correctors (written as the vector $\boldsymbol{\theta}$) to flatten the orbit. The response matrix $\mathbf{A}$ is given by

$$A_{ij} = \frac{\partial Y_i}{\partial \theta_j}.$$

In matrix form we have the equation

$$\mathbf{Y} = \mathbf{A}\boldsymbol{\theta}.$$

When the position measurements have weights, a diagonal matrix $\mathbf{W}$ of weights σ_j^{-1} for BPM j should give the correct normalization for χ^2 Provided that there are at least as many measurements as correctors ($m \geq n$), we may try to find a least squares solution to minimize the orbit distortion:

$$\begin{aligned}
\chi^2 &= (\mathbf{Y} - \mathbf{A}\boldsymbol{\theta})^{\mathrm{T}}\mathbf{W}(\mathbf{Y} - \mathbf{A}\boldsymbol{\theta}) \\
&= \mathbf{Y}^{\mathrm{T}}\mathbf{W}\mathbf{Y} + \boldsymbol{\theta}^{\mathrm{T}}(\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{A})\boldsymbol{\theta} - \mathbf{Y}^{\mathrm{T}}\mathbf{W}\mathbf{A}\boldsymbol{\theta} - \boldsymbol{\theta}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{Y} \\
&= \mathbf{Y}^{\mathrm{T}}\mathbf{W}\mathbf{Y} + \boldsymbol{\theta}^{\mathrm{T}}(\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{A})\boldsymbol{\theta} - 2\mathbf{Y}^{\mathrm{T}}\mathbf{W}\mathbf{A}\boldsymbol{\theta}. \\
\nabla\chi^2 &= 2(\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{A}\widehat{\boldsymbol{\theta}} - \mathbf{Y}^{\mathrm{T}}\mathbf{W}\mathbf{A}) = 0. \\
\widehat{\boldsymbol{\theta}} &= (\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{Y} \\
&= (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{Y}, \quad \text{if all weights } w_j \text{ are identical,} \\
&= \mathbf{A}^{-1}\mathbf{Y}, \quad \text{if } \mathbf{A} \text{ is a square matrix, i. e. if } n = m.
\end{aligned}$$

Provided that $(\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{A})^{-1}$ exists, then the minimum value of K is

$$\begin{aligned}
{\chi^2}_{\min} &= \left[\mathbf{Y} - \mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{Y}\right]^{\mathrm{T}}\mathbf{W}\left[\mathbf{Y} - \mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{Y}\right] \\
&= \mathbf{Y}^{\mathrm{T}}\left[\mathbf{W} - \mathbf{W}\mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\right]\mathbf{W}\left[\mathbf{W} - \mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{W}\right]\mathbf{Y} \\
&= \mathbf{Y}^{\mathrm{T}}\left[\mathbf{W} + \mathbf{W}\mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{W} - 2\mathbf{W}\mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{W}\right]\mathbf{Y} \\
&= \mathbf{Y}^{\mathrm{T}}\left[\mathbf{W} + \mathbf{W}\mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{W} - 2\mathbf{W}\mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{W}\right]\mathbf{Y} \\
&= \mathbf{Y}^{\mathrm{T}}\left[\mathbf{W} + \mathbf{W}\mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{W} - 2\mathbf{W}\mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{W}\right]\mathbf{Y} \\
&= \mathbf{Y}^{\mathrm{T}}\left[\mathbf{W} - \mathbf{W}\mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{W}\right]\mathbf{Y} \\
&= \mathbf{Y}^{\mathrm{T}}\mathbf{W}\mathbf{Y} - \mathbf{Y}^{\mathrm{T}}\mathbf{W}\mathbf{A}\widehat{\boldsymbol{\theta}} \\
&= \mathbf{Y}^{\mathrm{T}}\mathbf{W}\left(\mathbf{Y} - \mathbf{A}\widehat{\boldsymbol{\theta}}\right).
\end{aligned}$$

Of course if the matrix $\mathbf{A}$ is a big matrix, then numerically this algorithm may not work well, i. e. $(\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}$ may be hard to compute numerically. For large matrices other techniques, such as a three-bump algorithm or QR or SVD decomposition may work better.

3. Comments on correction via SVD

Singular value decomposition (SVD) is frequently used for orbit correction. There appear to be two different, but similar conventions for SVD. Namely an $m \times n$ matrix $\mathbf{A}$ with $m \geq n$ can be decomposed into a set of three matrices

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{T}} = \mathbf{D}\mathbf{E}\mathbf{T}^{\mathrm{T}},$$

with

- $\mathbf{U}$: an $m \times m$ othogonal matrix,
- $\mathbf{S}$: an $m \times n$ matrix with an upper diagonal $n\times n$ block and zeros in the remaining $m-n$ rows,
- $\mathbf{V}$: an $n\times n$ orthogonal matrix,

or

- $\mathbf{D}$: an $m \times n$ matrix with orthonormal column vectors,
- $\mathbf{E}$: an $n\times n$ diagonal matrix,
- $\mathbf{T}$: an $n\times n$ orthogonal matrix as before.

The elements of $\mathbf{S}$ and $\mathbf{E}$ are nonzero. The orthogonality conditions mean that

$$\begin{aligned}
\mathbf{U}^{\mathrm{T}}\mathbf{U} &= \mathbf{U}\mathbf{U}^{\mathrm{T}} = \mathbf{I}_m,\\
\mathbf{V}^{\mathrm{T}}\mathbf{V} &= \mathbf{V}\mathbf{V}^{\mathrm{T}} = \mathbf{I}_n,\\
\mathbf{T}^{\mathrm{T}}\mathbf{T} &= \mathbf{T}\mathbf{T}^{\mathrm{T}} = \mathbf{I}_n,\\
\mathbf{D}^{\mathrm{T}}\mathbf{D} &= \mathbf{I}_n.
\end{aligned}$$

In the accelerator handbook[1], Ref. [2] uses the convention with $\mathbf{D}$, $\mathbf{E}$, and $\mathbf{V}$, whereas the `svd()` function in `Octave` returns matrices of $(\mathbf{U}, \mathbf{S}, \mathbf{V})$ convention. The $(\mathbf{D}, \mathbf{E}, \mathbf{T})$ convention is more compact.

Write $\mathbf{U}$ in block form

$$\mathbf{U} = \begin{pmatrix} \mathbf{U}_1 & \mathbf{U}_2 \\ \mathbf{U}_3 & \mathbf{U}_4 \end{pmatrix},$$

with the block sizes for $\mathbf{U}_1$, $\mathbf{U}_2$, $\mathbf{U}_3$, and $\mathbf{U}_4$ being respectively $n\times n$, $n\times(m-n)$, $(n-m)\times n$, and $(n-m)\times n$. The matrix $\mathbf{D}$ may similarly be written as

$$\mathbf{D} = \begin{pmatrix} \mathbf{D}_1 \\ \mathbf{D}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{U}_1 \\ \mathbf{U}_3 \end{pmatrix}.$$

The orthogonality of $\mathbf{U}$ requires that

$$\begin{pmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{m-n} \end{pmatrix} = \mathbf{U}^{\mathrm{T}}\mathbf{U} = \begin{pmatrix} {\mathbf{U}_1}^{\mathrm{T}} & {\mathbf{U}_3}^{\mathrm{T}} \\ {\mathbf{U}_2}^{\mathrm{T}} & {\mathbf{U}_4}^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} \mathbf{U}_1 & \mathbf{U}_2 \\ \mathbf{U}_3 & \mathbf{U}_4 \end{pmatrix} = \begin{pmatrix} {\mathbf{U}_1}^{\mathrm{T}}\mathbf{U}_1 + {\mathbf{U}_3}^{\mathrm{T}}\mathbf{U}_3 & {\mathbf{U}_1}^{\mathrm{T}}\mathbf{U}_2 + {\mathbf{U}_3}^{\mathrm{T}}\mathbf{U}_4 \\ {\mathbf{U}_2}^{\mathrm{T}}\mathbf{U}_1 + {\mathbf{U}_4}^{\mathrm{T}}\mathbf{U}_3 & {\mathbf{U}_2}^{\mathrm{T}}\mathbf{U}_2 + {\mathbf{U}_4}^{\mathrm{T}}\mathbf{U}_4 \end{pmatrix},$$

or

$$\begin{aligned}
\mathbf{I}_n &= {\mathbf{U}_1}^{\mathrm{T}}\mathbf{U}_1 + {\mathbf{U}_3}^{\mathrm{T}}\mathbf{U}_3,\\
\mathbf{I}_{m-n} &= {\mathbf{U}_2}^{\mathrm{T}}\mathbf{U}_2 + {\mathbf{U}_4}^{\mathrm{T}}\mathbf{U}_4,\\
\mathbf{0} &= {\mathbf{U}_1}^{\mathrm{T}}\mathbf{U}_2 + {\mathbf{U}_3}^{\mathrm{T}}\mathbf{U}_4,
\end{aligned}$$

and we see that this representation of $\mathbf{D}$ indeed satisfies $\mathbf{D}^{\mathrm{T}}\mathbf{D} = \mathbf{I}_n$. Similarly, may identify $\mathbf{T} = \mathbf{V}$ and $\mathbf{S} = \begin{pmatrix} \mathbf{E} \\ \mathbf{0} \end{pmatrix}$ with the zero block having dimensions $(n-m)\times n$.

Our orbit response matrix is

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{T}} = \begin{pmatrix} \mathbf{U}_1 & \mathbf{U}_2 \\ \mathbf{U}_3 & \mathbf{U}_4 \end{pmatrix} \begin{pmatrix} \mathbf{E} \\ \mathbf{0} \end{pmatrix} \mathbf{V}^{\mathrm{T}} = \begin{pmatrix} \mathbf{U}_1\mathbf{E}\mathbf{V}^{\mathrm{T}} \\ \mathbf{U}_3\mathbf{E}\mathbf{V}^{\mathrm{T}} \end{pmatrix} = \mathbf{D}^{\mathrm{T}}\mathbf{E}\mathbf{V}^{\mathrm{T}}.$$

The least-squares solution to $\mathbf{Y} = \mathbf{A}\boldsymbol{\theta}$ (with equal weights, i. e. $\mathbf{W} = w\mathbf{I}_n$) was

$$\begin{aligned}
\widehat{\boldsymbol{\theta}} &= (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{Y}\\
&= \left[\begin{pmatrix} \mathbf{V}\mathbf{E}{\mathbf{U}_1}^{\mathrm{T}} & \mathbf{V}\mathbf{E}{\mathbf{U}_3}^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} \mathbf{U}_1\mathbf{E}\mathbf{V}^{\mathrm{T}} \\ \mathbf{U}_3\mathbf{E}\mathbf{V}^{\mathrm{T}} \end{pmatrix}\right]^{-1} \begin{pmatrix} \mathbf{V}\mathbf{E}{\mathbf{U}_1}^{\mathrm{T}} & \mathbf{V}\mathbf{E}{\mathbf{U}_3}^{\mathrm{T}} \end{pmatrix}\mathbf{Y}\\
&= \left[\mathbf{V}\mathbf{E}{\mathbf{U}_1}^{\mathrm{T}}\mathbf{U}_1\mathbf{E}\mathbf{V}^{\mathrm{T}} + \mathbf{V}\mathbf{E}{\mathbf{U}_3}^{\mathrm{T}}\mathbf{U}_3\mathbf{E}\mathbf{V}^{\mathrm{T}}\right]^{-1}\mathbf{V}\mathbf{E}\mathbf{D}^{\mathrm{T}}\mathbf{Y}\\
&= \left[\mathbf{V}\mathbf{E}\mathbf{E}\mathbf{V}^{\mathrm{T}}\right]^{-1}\mathbf{V}\mathbf{E}\mathbf{D}^{\mathrm{T}}\mathbf{Y}\\
&= \mathbf{V}(\mathbf{E}^{-1})^2\mathbf{V}^{\mathrm{T}}\mathbf{V}\mathbf{E}\mathbf{D}^{\mathrm{T}}\mathbf{Y}\\
&= \mathbf{V}\mathbf{E}^{-1}\mathbf{D}^{\mathrm{T}}\mathbf{Y}
\end{aligned}$$

The minimum value of the χ^2 for this solution is

$$\begin{aligned}
{\chi^2}_{\min} &= \mathbf{Y}^{\mathrm{T}}\mathbf{W}\left(\mathbf{Y}-\mathbf{A}\widehat{\boldsymbol{\theta}}\right) \\
&= \mathbf{Y}^{\mathrm{T}}\mathbf{W}\mathbf{Y}-\mathbf{Y}^{\mathrm{T}}\mathbf{W}\mathbf{D}\mathbf{E}\mathbf{V}^{\mathrm{T}}\,\mathbf{V}\mathbf{E}^{-1}\mathbf{D}^{\mathrm{T}}\mathbf{Y} \\
&= \mathbf{Y}^{\mathrm{T}}\mathbf{W}[\mathbf{I}-\mathbf{D}\mathbf{D}^{\mathrm{T}}]\mathbf{Y}.
\end{aligned}$$

So a figure of merit for the solution could be $\chi^2/(m-n)$.

Just as an example, I have taken the BPM offset errors we had in the RHIC Blue ring and calculated corrector strengths to match twice the error. The results are shown in Fig. 2.

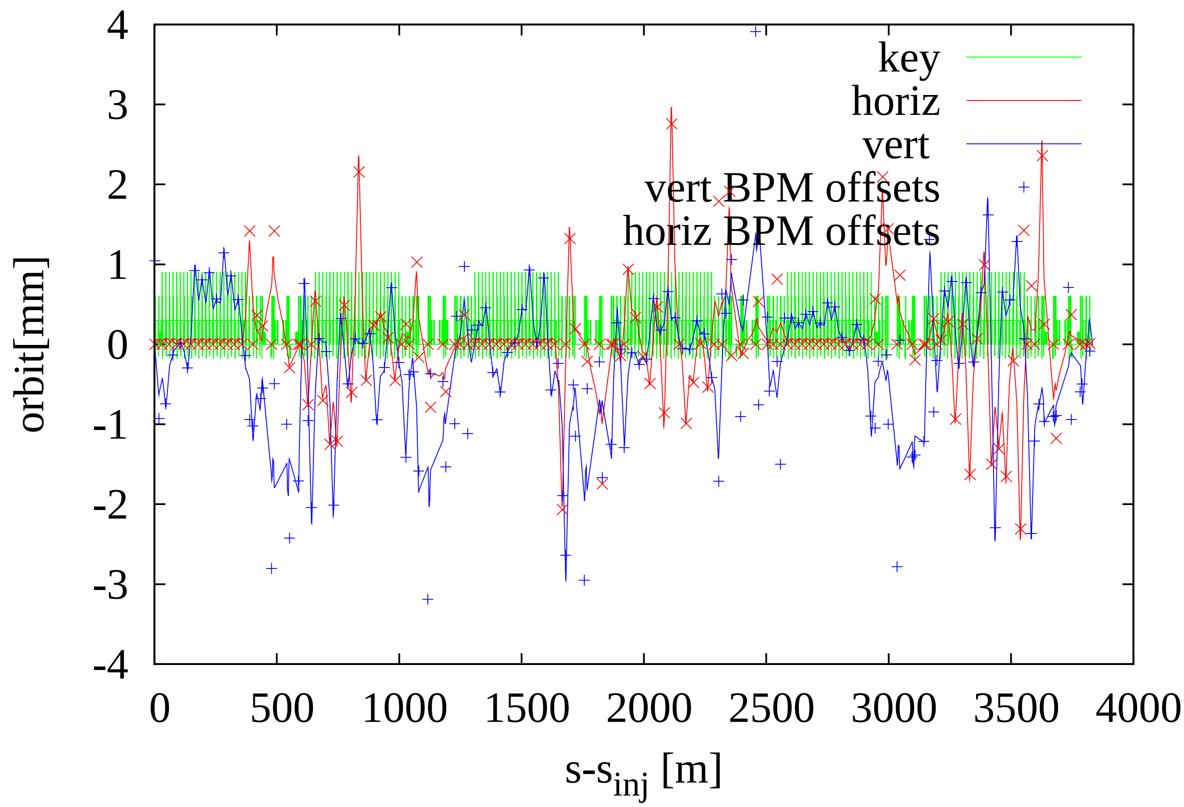

Figure 2. Simulation orbit correction using this SVD algorithm to calculate corrector strengths to match the erroneous (100 GeV protons using the 2009 lattice) BPM offsets in the RHIC ring. The orbits were simulated with the program `blspin`.

4. References

1. Eds. A. W. Chao and M. Tigner, "Handbook of Accelerator Physics and Engineering", World Sci., Singapore (1998).
2. S. Krinsky, "Orbit Correction", § 4.5.3 in [1].
3. John W. Eaton and others, `Octave` programming language, `http://www.gnu.org/software/octave`.