Theory and Praxis of Map Analysis in CHEF
Part 1: Linear Normal Form

Leo Michelotti

October 14, 2008

> Every other man spoke a language entirely his own, which he had figured out by private thinking; he had his own ideas and peculiar ways. If you wanted to talk about a glass of water, you had to start back with God creating the heavens and the earth; the apple; Abraham; Moses and Jesus; Rome; the Middle Ages; gunpowder; the Revolution; back to Newton; up to Einstein; then war and Lenin and Hitler. After reviewing this and getting it all straight again you could proceed to talk .... "I'm fainting, please get me a little water." You were lucky even then to make yourself understood.
>
> — Saul Bellow
> *Seize the Day*

1 Introduction

This memo begins a series which, put together, could comprise the "CHEF Documentation Project" if there were such a thing. The first - and perhaps only - three will telegraphically describe theory, algorithms, implementation and usage of the normal form map analysis procedures encoded in CHEF's collection of libraries. [1] This one will begin the sequence by explaining the linear manipulations that connect the Jacobian matrix of a symplectic mapping to its normal form. It is a "Reader's Digest" version of material I wrote in *Intermediate Classical Dynamics* (ICD) [2] and randomly scattered across technical memos, seminar viewgraphs, and lecture notes for the past quarter century. Much of its content is old, well known, and in some places borders on the trivial.[1] Nevertheless, completeness requires their inclusion. The primary objective is the "fundamental theorem" on normalization written on page 8. I plan to describe the nonlinear procedures in a subsequent memo and devote a third to laying out algorithms and lines of code, connecting them with equations written in the first two. Originally this was to be done in one short paper, but I jettisoned that approach after its first section exceeded a dozen pages.

The organization of this document is as follows. A brief description of notation is followed by a section containing a general treatment of the linear problem. After the "fundamental theorem" is proved, two further subsections discuss the generation of equilibrium distributions and issue of "phase." The final major section reviews parametrizations – that is,

[1] Notwithstanding, it is worth remembering that a few decades ago (almost?) everyone in accelerator physics routinely preferred a language carried over from the early days of cyclotrons.

lattice functions – in two and four dimensions with a passing glance at the six-dimensional version. Appearances to the contrary, for the most part I have tried to restrict consideration to matters needed to understand the code in CHEF's libraries.

2 Preliminaries on notation

I shall use the notation of ICD [2]: a collection of phase space coordinates is indicated with an underlined letter, as in "$\underline{z}$," and a matrix with double underlines, as in "$\underline{\underline{M}}$." This is *not* meant to imply that phase space is a vector space. It is not, but locally it looks like one, because it is coordinatized[2] as an open subset of R^N, for some N, containing the origin. In any case, the linear part of a mapping subsists in a vector space, the tangent space attached to some point of the manifold.

For reasons that might become more evident later, I prefer to arrange phase space coordinates so that the "position" and "momentum" sectors are separated. (As a result, this is how they have always been ordered internally in CHEF's software.) Most others prefer to keep individual "position" and "momentum" coordinates together, which in accelerator physics leads to grouping according to "horizontal," "vertical," and "longitudinal" sectors. For example, for a four-dimensional, R^4 phase space,

$$\textbf{I would use}: \quad \underline{z} = \begin{pmatrix} x \\ y \\ p_x \\ p_y \end{pmatrix} \quad \textbf{while many others use}: \quad \underline{z} = \begin{pmatrix} x \\ p_x \\ y \\ p_y \end{pmatrix} \quad . \tag{1}$$

Both choices possess advantages in different contexts: the former is convenient for carrying out matrix multiplications in 2×2 blocks; the latter, for performing horizontal and vertical projections. When it is necessary to refer to this convention, and a few others to be given later, I shall write that the coordinates are "sensibly ordered."[3]

We deal with discrete dynamical systems, whose orbits are represented by indexed sequences,

$$\{ \cdots, \underline{z}_{-3}, \underline{z}_{-2}, \underline{z}_{-1}, \underline{z}_0, \underline{z}_1, \underline{z}_2, \underline{z}_3, \cdots \} \ .$$

Each $\underline{z}_k$ is an array of coordinates on a phase space chart. They need not be real and generally are not. Nonetheless, we shall assume we *start with* real coordinates, i.e. $\underline{z}_k \in R^N$, for some $N = 2n$.[4] If the system is deterministic, then there is an indexed collection of "transition functions," or "mappings," T_k, such that $\underline{z}_{k+1} = T_k(\underline{z}_k)$; if it is also autonomous, then all T_k are the same, and we write

$$\underline{z}_{k+1} = T(\underline{z}_k), \quad \text{or} \quad T : \underline{z}_k \mapsto \underline{z}_{k+1}, \quad \text{or sometimes just} \quad T : Z \to Z, \ \ \underline{z} \mapsto T(\underline{z}) \ .$$

Of course, the application to accelerator physics is implicit: the transition represents a single turn in a periodic machine.

3 The linear sector

The goal of this section is to fix the definition of normal coordinates for linear systems, with special attention paid to the normalization appropriate when the transition is symplectic, such as one arising from a Hamiltonian flow. Our key result is the "fundamental theorem" written in Eq.(10). Following subsections treat quadratic statistics for invariant distributions and the calculation of "phase advance" in a ring.

[2] If there is such an abominable word.

[3] This is a biassed characterization, of course, but I have to call it *something*.

[4] The restriction to an even-dimensional phase space could be lifted for greater generality, e.g. to include spin, but not here and not now.

3.1 Generic linear normal form coordinates

Regardless of symplecticity, if the system is autonomous, then its linear part is described by a square $N \times N$ matrix, $\underline{\underline{M}}$, such that

$$\underline{z}_{k+1} = \underline{\underline{M}} \cdot \underline{z}_k + O(\underline{z}_k^2) \quad \textbf{for all } k \ .$$

I'll immediately drop the "$O(\underline{z}_k^2)$" notation from all following equations and write more simply,

$$\underline{z}_{k+1} = \underline{\underline{M}} \cdot \underline{z}_k \quad \text{or} \quad \underline{z} \mapsto \underline{\underline{M}} \cdot \underline{z} \ .$$

If there exists a square matrix $\underline{\underline{B}}$ and a diagonal matrix $\underline{\underline{\Lambda}}$ such that

$$\underline{\underline{M}} \cdot \underline{\underline{B}} = \underline{\underline{B}} \cdot \underline{\underline{\Lambda}} \ , \tag{2}$$

then the phase space decomposes into N independent, one-dimensional, possibly complex, invariant subspaces.[5] The columns of $\underline{\underline{B}}$ contain eigenvectors of $\underline{\underline{M}}$, with each Λ_{ii} being the corresponding eigenvalue. We define the "(linear) normal form coordinates," $\underline{a}$, according to

$$\underline{z} = \underline{\underline{B}} \cdot \underline{a} \ . \tag{3}$$

Which brings us to the eternal question of which "picture" to adopt. To abuse slightly familiar terminology from quantum mechanics, in the "Schrödinger picture," states evolve while the operators remain fixed; in the "Heisenberg picture," states remain fixed – at their initially specified values – and operators evolve. The choice here is,

$$\begin{aligned} \underline{\underline{M}} \cdot \underline{z} \ = \ \underline{\underline{M}} \cdot (\underline{\underline{B}} \cdot \underline{a}) \ &= \ (\underline{\underline{M}} \cdot \underline{\underline{B}}) \cdot \underline{a} \\ &= \ (\underline{\underline{B}} \cdot \underline{\underline{\Lambda}}) \cdot \underline{a}, \quad \text{Heisenberg} \\ &= \ \underline{\underline{B}} \cdot (\underline{\underline{\Lambda}} \cdot \underline{a}), \quad \text{Schrodinger} \ . \end{aligned}$$

Put another way, do we interpret $\underline{z} \mapsto \underline{\underline{M}} \cdot \underline{z}$ as $\underline{a} \mapsto \underline{\underline{\Lambda}} \cdot \underline{a}$ (i.e. $a_i \mapsto \Lambda_{ii} a_i$) or as $\underline{\underline{B}} \mapsto \underline{\underline{B}} \cdot \underline{\underline{\Lambda}}$? We shall eschew the choice and adopt "both-and" rather than "either-or" in different circumstances as convenience dictates. In fact, a third choice, the "interaction picture," is frequently used. We shall take it up in Section 3.4.

If $\underline{\underline{M}}$ is not orthogonal, its eigenvalues, eigenvectors and the normal form coordinates generally are complex. Further, because $\underline{\underline{M}}$ is real, eigenvalues must come in complex conjugate pairs. We extend the notion of "sensible ordering" to the columns of $\underline{\underline{B}}$ and $\underline{\underline{\Lambda}}$: they are "sensibly ordered" if $\Lambda_{ii}^* = \Lambda_{i+n\ i+n}$ whenever $\Im \Lambda_{ii} \neq 0$. (Note: index arithmetic is done modulo $2n$.) This induces a corresponding order on the columns of $\underline{\underline{B}}$. Eigenvectors associated with complex conjugate eigenvalues need not themselves be complex conjugates of each other, as an eigenvector can be multiplied by any complex number and remain an eigenvector (with the same eigenvalue). That is, for any diagonal matrix $\underline{\underline{W}}$,

$$\underline{\underline{M}} \cdot \underline{\underline{B}} \cdot \underline{\underline{W}} \ = \ \underline{\underline{B}} \cdot \underline{\underline{\Lambda}} \cdot \underline{\underline{W}} \ = \ \underline{\underline{B}} \cdot \underline{\underline{W}} \cdot \underline{\underline{\Lambda}}$$

so that $\underline{\underline{B}} \cdot \underline{\underline{W}}$ also satisfies Eq.(2) and could be used instead of $\underline{\underline{B}}$. So, we enforce the condition and assert that the columns of $\underline{\underline{B}}$ have been scaled in such a way that, like $\underline{\underline{\Lambda}}$, it satisfies $B_{ii}^* = B_{i+n\ i+n}$ whenever $\Im B_{ii} \neq 0$. As a consequence, half of the the normal form coordinates will be complex conjugates of the other half. These conditions can be written as matrix equations by introducing the extended "Pauli matrix":

$$\underline{\underline{\sigma}}_1 \equiv \begin{pmatrix} \underline{\underline{0}} & \underline{\underline{1}} \\ \underline{\underline{1}} & \underline{\underline{0}} \end{pmatrix} ; \quad \underline{\underline{B}}^* = \underline{\underline{B}} \cdot \underline{\underline{\sigma}}_1 ; \quad \underline{\underline{\Lambda}}^* = \underline{\underline{\Lambda}} \cdot \underline{\underline{\sigma}}_1 \ , \tag{4}$$

[5] Were we to require phase space coordinates to remain real, then some invariant subspaces would be two-dimensional: i.e. one complex coordinate subsumes two real coordinates. Enforcing such a requirement produces no advantage. The notation used here is the one introduced in Ref. [3]

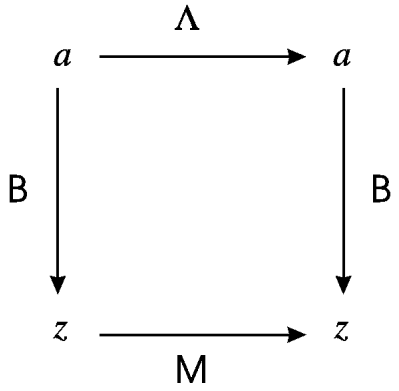

Figure 1: Commutative diagram connecting real and normal form coordinates.

where $\underline{\underline{\sigma}}_1$ is written in 2×2 block form: $\underline{\underline{0}}$ is the $n \times n$ zero matrix and $\underline{\underline{1}}$, the $n \times n$ identity matrix.

NOTE: This equation does not exhaust the requirements on preordering the columns of $\underline{\underline{B}}$. Permutations could still be carried out within the $2n \times n$ submatrices: i.e. the first and second blocks of n columns. There are two more conditions which will be introduced later, when they are needed.[6]

The attentive reader already has noticed some difficulties in defining normal form coordinates via Eq.(3). We have tacitly assumed that the dynamical system is reversible, which means that $\underline{\underline{B}}$ is invertible: no eigenvalue is zero, and the null space of $\underline{\underline{M}}$ is trivial. Further, degeneracy – two or more identical eigenvalues – would produce additional ambiguity; corresponding eigenvectors can be superposed to produce different ones. Most importantly for us, a final normalization must be chosen. In principle, this is not important as long as one remains consistent within a particular calculation, but I shall specify a preferred normalization for symplectic systems in the next section.

COMMENTS:

1. All of this is very basic and requires no further explanation. Both eigenvalues and eigenvectors are generally complex, so normal form coordinates will be too. In principle, the original "phase space" could have been complex by defining orbits as sequences over C^N.

2. Not all matrices can be diagonalized. In such cases, proving the non-existence of $\underline{\underline{B}}$ is an ill-posed problem. We'll say no more about that; let's move on.

3. Algebraists express Eq.(2) graphically as a "commutative diagram," like the one shown in Figure 1. As we shall see in the next memo, it can be extended surprisingly naturally to the nonlinear sector, where it helps to organize one's thinking.

[6] If you want to peek ahead, see Eq.(9) and Eq.(26).

3.2 Symplectic systems

While this is a somewhat restricted[7] definition, we shall understand the linear system, $\underline{z} \mapsto \underline{\underline{M}} \cdot \underline{z}$, to be "symplectic" when $\underline{\underline{M}}$ satisfies

$$\underline{\underline{M}} \cdot \underline{\underline{J}} \cdot \underline{\underline{M}}^T = \underline{\underline{J}}, \quad \textbf{where} \quad \underline{\underline{J}} \equiv \begin{pmatrix} \underline{\underline{0}} & \underline{\underline{1}} \\ -\underline{\underline{1}} & \underline{\underline{0}} \end{pmatrix} \ . \tag{5}$$

Here, $\underline{\underline{J}}$ is written in block form; it looks as it does because dynamical coordinates are "sensibly ordered," as in Eq.(1). Before continuing, it is worth noting Eq.(5) could also be written,

$$\underline{\underline{M}}^T \cdot \underline{\underline{J}} \cdot \underline{\underline{M}} = \underline{\underline{J}} \ , \tag{6}$$

as the two expressions are interchangeable. Since this is not frequently asserted, we shall pause to do so, and even generalize the observation to complex matrices, using the usual notation, $\underline{\underline{M}}^\dagger = \underline{\underline{M}}^{*T}$

ASSERTION: Let $\underline{\underline{O}}$ be a unitary matrix: that is, $\underline{\underline{O}}^\dagger \cdot \underline{\underline{O}} = \underline{\underline{O}} \cdot \underline{\underline{O}}^\dagger = \underline{\underline{1}}$. Let $\underline{\underline{M}}$ be a matrix satisfying $\underline{\underline{M}} \cdot \underline{\underline{O}} \cdot \underline{\underline{M}}^\dagger = \underline{\underline{O}}$. Then, $\underline{\underline{M}}^\dagger \cdot \underline{\underline{O}} \cdot \underline{\underline{M}} = \underline{\underline{O}}$.

PROOF: Clearly, both $\underline{\underline{O}}$ and $\underline{\underline{M}}$ are square and not singular. Thus $\underline{\underline{M}}^{-1}$ exists; in fact, $\underline{\underline{M}}^{-1} = \underline{\underline{O}} \cdot \underline{\underline{M}}^\dagger \cdot \underline{\underline{O}}^\dagger$, for

$$\underline{\underline{M}} \cdot \underline{\underline{O}} \cdot \underline{\underline{M}}^\dagger \cdot \underline{\underline{O}}^\dagger = \underline{\underline{O}} \cdot \underline{\underline{O}}^\dagger = \underline{\underline{1}} \ .$$

Then, since $(\underline{\underline{M}}^\dagger)^{-1} = (\underline{\underline{M}}^{-1})^\dagger$, we have

$$\begin{aligned} \underline{\underline{O}} &= (\underline{\underline{M}}^{-1}) \cdot (\underline{\underline{M}} \cdot \underline{\underline{O}} \cdot \underline{\underline{M}}^\dagger) \cdot (\underline{\underline{M}}^{-1})^\dagger \\ &= (\underline{\underline{O}} \cdot \underline{\underline{M}}^\dagger \cdot \underline{\underline{O}}^\dagger) \cdot (\underline{\underline{O}}) \cdot (\underline{\underline{O}} \cdot \underline{\underline{M}}^\dagger \cdot \underline{\underline{O}}^\dagger)^\dagger \\ &= \underline{\underline{O}} \cdot (\underline{\underline{M}}^\dagger \cdot \underline{\underline{O}} \cdot \underline{\underline{M}}) \cdot \underline{\underline{O}}^\dagger \ . \end{aligned}$$

Finally, multiplying on the left by $\underline{\underline{O}}^\dagger$ and right by $\underline{\underline{O}}$ provides the result

$$\underline{\underline{M}}^\dagger \cdot \underline{\underline{O}} \cdot \underline{\underline{M}} = \underline{\underline{O}} \ .$$

□

The application to Eq.(6) is immediate upon noting that $\underline{\underline{J}}$ is a real, orthogonal matrix and, therefore, unitary.

As already noted, eigenvalues and eigenvectors of any real $\underline{\underline{M}}$ must come in complex conjugate pairs. If $\underline{\underline{M}}$ is symplectic, they also can be associated in reciprocal pairs.

ASSERTION: If λ is an eigenvalue of a symplectic matrix, $\underline{\underline{M}}$, then so are λ^* and $1/\lambda$.

PROOF HINT: This simple proof can be found in many books, but as a hint for the second part: notice that $\underline{\underline{M}}^{-1} = -\underline{\underline{J}} \cdot \underline{\underline{M}}^T \cdot \underline{\underline{J}}$ when $\underline{\underline{M}}$ is symplectic, and manipulate the characteristic polynomial for $\underline{\underline{M}}^{-1}$, $\det(\underline{\underline{M}}^{-1} - \lambda \underline{\underline{1}}) = 0$.

□

[7] The restriction comes from assuming the coordinates are canonical. It is sometimes useful to exploit symplecticity with noncanonical coordinates.

COMMENTS:

4. Obviously the first assertion above works as well in the other direction: simply substitute $\underline{\underline{M}}^{\dagger}$ for $\underline{\underline{M}}$ throughout its proof.

5. The reciprocal of an eigenvalue is generally not its complex conjugate. However, if orbits are bounded (i.e. linearly stable), then all eigenvalues must lie on the unit circle — $|\lambda| = 1$ for all λ — and therefore satisfy $\lambda^* = 1/\lambda$. Though it is not always essential, *in everything that follows we shall tacitly assume this to be the case when it either simplifies or validates a calculation.* In the language of accelerator physics, this means that linear (resonance) stopbands, such as $\nu_x = \nu_y$ or $2\nu_x =$ integer, have been avoided.

6. For $N = 2$, the symplectic condition, Eq.(5), is tantamount to $\mathbf{det}(\underline{\underline{M}}) = 1$, because $\underline{\underline{M}} \cdot \underline{\underline{J}} \cdot \underline{\underline{M}}^T = \mathbf{det}(\underline{\underline{M}})\ \underline{\underline{J}}$. In higher dimensions, its implications are more involved (see below), but clearly $\mathbf{det}(\underline{\underline{M}}) = \pm 1$. In fact, as the determinant is the product of the eigenvalues, any symplectic matrix must have determinant $+1$, because of the reciprocal property asserted above.

7. The symplectic condition is equivalent to invariance of Poincaré's differential two-form, ω, which is evaluated using a canonical coordinate chart as follows. Let $\mathbf{u}$ and $\mathbf{v}$ be arbitrary tangent vectors.

$$\begin{aligned} \omega(\mathbf{u},\mathbf{v}) &= \left(\sum_{k=1}^{n} dx_k \wedge dp_k\right)(\mathbf{u},\mathbf{v}) &= \sum_{k=1}^{n} dx_k(\mathbf{u})dp_k(\mathbf{v}) - dp_k(\mathbf{u})dx_k(\mathbf{v}) \\ &&= (d\underline{x}\ \ d\underline{p})_{(\mathbf{u})} \cdot \underline{\underline{J}} \cdot \begin{pmatrix} d\underline{x} \\ d\underline{p} \end{pmatrix}_{(\mathbf{v})} \\ &&= d\underline{z}^T(\mathbf{u}) \cdot \underline{\underline{J}} \cdot d\underline{z}(\mathbf{v}) \ . \end{aligned} \tag{7}$$

8. As is well known, symplectic matrices form a group. We'll say no more about that, *except* to mention in passing a lovely theorem which states that the entire group can be generated by the two subgroups of matrices of the form

$$\begin{pmatrix} \underline{\underline{1}} & s\underline{\underline{1}} \\ \underline{\underline{0}} & \underline{\underline{1}} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \underline{\underline{1}} & \underline{\underline{0}} \\ \underline{\underline{K}} & \underline{\underline{1}} \end{pmatrix} \ ,$$

where s is any real number and $\underline{\underline{K}}$ is any symmetric, $n \times n$ matrix. Again using the language of accelerator physics, the first subgroup corresponds to drift spaces, and the second comprise linear kicks, a generalization of thin quadrupoles. If you are interested in reading a proof, one can be found in Guillemin and Sternberg [4].

9. Symplecticity does not guarantee diagonalizability. The most trivial (one parameter) symplectic subgroups that cannot be diagonalized are first learned in high school and were already written above:

$$\underline{\underline{M}}(t) = \begin{pmatrix} \underline{\underline{1}} & t\underline{\underline{1}} \\ \underline{\underline{0}} & \underline{\underline{1}} \end{pmatrix}, \quad t \in R.$$

This is the time-evolution operator of unaccelerated motion, particles travelling in straight lines at constant velocities. So, in what follows, we must *assume* that Eq.(2) is satisfied.

The eigenvector matrix, $\underline{\underline{B}}$, defined in Eq.(2), which transforms real canonical coordinates into normal form coordinates, does not satisfy Eq.(5) and, therefore, is not symplectic. What remains of this section is devoted to proving

that, when orbits are bounded, the columns of $\underline{\underline{B}}$, can always be normalized to satisfy the analogous condition, $B^T \cdot \underline{\underline{J}} \cdot \underline{\underline{B}} = i\,\underline{\underline{J}}$. This fixes the normalization of normal coordinates so that they can be associated unambiguously with action-angle coordinates and, eventually, with *at least one* definition of the emittance of a bunch. Because this result does not normally appear in textbooks, I'll write a proof that steps through a sequence of lemmas. If you have no interest in proofs, skip what follows and begin again with the comments following the "fundamental theorem" on page 8.

LEMMA: Let $\underline{\underline{M}} \cdot \underline{\underline{B}} = \underline{\underline{B}} \cdot \underline{\underline{\Lambda}}$, where all matrices are square and sensibly ordered, $\underline{\underline{M}}$ is real and symplectic, and $\underline{\underline{\Lambda}}$ is diagonal, non-degenerate, with all non-zero elements on the unit circle. Then $\underline{\underline{B}}^T \cdot \underline{\underline{J}} \cdot \underline{\underline{B}} \cdot \underline{\underline{J}}$ is a diagonal matrix.

PROOF: Using the previous lemma, symplecticity of $\underline{\underline{M}}$ is written $\underline{\underline{M}}^T \cdot \underline{\underline{J}} \cdot \underline{\underline{M}} = \underline{\underline{M}} \cdot \underline{\underline{J}} \cdot \underline{\underline{M}}^T = \underline{\underline{J}}$, from which we obtain,

$$\begin{aligned}
\underline{\underline{B}}^T \cdot \underline{\underline{J}} \cdot \underline{\underline{B}} &= \underline{\underline{B}}^T \cdot (\underline{\underline{M}}^T \cdot \underline{\underline{J}} \cdot \underline{\underline{M}}) \cdot \cdot \underline{\underline{B}} \\
&= \left(\underline{\underline{M}} \cdot \underline{\underline{B}}\right)^T \cdot \underline{\underline{J}} \cdot \left(\underline{\underline{M}} \cdot \underline{\underline{B}}\right) \\
&= \left(\underline{\underline{B}} \cdot \underline{\underline{\Lambda}}\right)^T \cdot \underline{\underline{J}} \cdot \left(\underline{\underline{B}} \cdot \underline{\underline{\Lambda}}\right) \\
&= \underline{\underline{\Lambda}}^T \cdot \left(\underline{\underline{B}}^T \cdot \underline{\underline{J}} \cdot \underline{\underline{B}}\right) \cdot \underline{\underline{\Lambda}}
\end{aligned}$$

We note in passing that $\underline{\underline{\Lambda}}^T = \underline{\underline{\Lambda}}$, as $\underline{\underline{\Lambda}}$ is diagonal. In terms of components, this is written:

$$\textbf{for all } i,j: \quad (\,\underline{\underline{B}}^T \cdot \underline{\underline{J}} \cdot \underline{\underline{B}})_{ij} = (\Lambda_{ii}\Lambda_{jj})\,(\,\underline{\underline{B}}^T \cdot \underline{\underline{J}} \cdot \underline{\underline{B}})_{ij}$$

which means

$$\textbf{either } \; (\,\underline{\underline{B}}^T \cdot \underline{\underline{J}} \cdot \underline{\underline{B}})_{ij} = 0 \quad \textbf{or} \quad \Lambda_{ii}\Lambda_{jj} = 1.$$

Because of the hypotheses about $\underline{\underline{\Lambda}}$, the only non-zero elements of $\underline{\underline{B}}^T \cdot \underline{\underline{J}} \cdot \underline{\underline{B}}$ must be in the same position as the non-zero elements of $\underline{\underline{J}}$. Thus, $\underline{\underline{B}}^T \cdot \underline{\underline{J}} \cdot \underline{\underline{B}} \cdot \underline{\underline{J}}$ is a diagonal matrix.

□

Without loss of generality, we express this matrix as $-i\underline{\underline{D}}$, where $\underline{\underline{D}}$ is diagonal, so that we can write

$$\underline{\underline{B}}^T \cdot \underline{\underline{J}} \cdot \underline{\underline{B}} = i\,\underline{\underline{D}} \cdot \underline{\underline{J}}, \quad \text{where } \underline{\underline{D}} \text{ is diagonal.}$$

Our goal now is reduced to proving that the columns of $\underline{\underline{B}}$ can be normalized such that $\underline{\underline{D}} = \underline{\underline{1}}$. To that end, I offer a final intermediate lemma for purposes of meditation: namely, $[\underline{\underline{D}}, \underline{\underline{\sigma}}_1] = \underline{\underline{0}}$.

LEMMA: If $\underline{\underline{B}}$ is symplectic and sensibly ordered, then

$$\underline{\underline{B}}^T \cdot \underline{\underline{J}} \cdot \underline{\underline{B}} \cdot \underline{\underline{J}} = -i \begin{pmatrix} \underline{\underline{D}}' & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{D}}' \end{pmatrix} ,$$

where the $n \times n$ block $\underline{\underline{D}}'$ is a real, diagonal matrix.

PROOF: Because $\underline{\underline{B}}$ is sensibly ordered, we can write it in block form as follows.

$$\underline{\underline{B}} = \begin{pmatrix} \underline{\underline{B}}_1 & \underline{\underline{B}}_1^* \\ \underline{\underline{B}}_2 & \underline{\underline{B}}_2^* \end{pmatrix}$$

A simple calculation then gives us,

$$\underline{\underline{B}}^T \cdot \underline{\underline{J}} \cdot \underline{\underline{B}} \cdot \underline{\underline{J}} = \begin{pmatrix} \underline{\underline{B}}_2^T \cdot \underline{\underline{B}}_1^* - \underline{\underline{B}}_1^T \cdot \underline{\underline{B}}_2^* & \underline{\underline{B}}_1^T \cdot \underline{\underline{B}}_2 - \underline{\underline{B}}_2^T \cdot \underline{\underline{B}}_1 \\ \underline{\underline{B}}_2^\dagger \cdot \underline{\underline{B}}_1^* - \underline{\underline{B}}_1^\dagger \cdot \underline{\underline{B}}_2^* & \underline{\underline{B}}_1^\dagger \cdot \underline{\underline{B}}_2 - \underline{\underline{B}}_2^\dagger \cdot \underline{\underline{B}}_1 \end{pmatrix} \tag{8}$$

We have seen already that this is diagonal, so the the diagonal blocks are themselves diagonal matrices, and the off-diagonal blocks must vanish. This means that

$$\underline{\underline{B}}_1^T \cdot \underline{\underline{B}}_2 = (\underline{\underline{B}}_1^T \cdot \underline{\underline{B}}_2)^T \ .$$

That is, $\underline{\underline{B}}_1^T \cdot \underline{\underline{B}}_2$ is a symmetric matrix, which is interesting but not to the point. More to the point is that the (diagonal) diagonal blocks are anti-Hermitian and negative complex conjugates of each other. This means

$$\underline{\underline{B}}^T \cdot \underline{\underline{J}} \cdot \underline{\underline{B}} \cdot \underline{\underline{J}} = \begin{pmatrix} -i\underline{\underline{D}}' & \underline{\underline{0}} \\ \underline{\underline{0}} & -i\underline{\underline{D}}' \end{pmatrix} \equiv -i\underline{\underline{D}} \ ,$$

where $\underline{\underline{D}}'$ is a real, diagonal matrix.

□

We have shown not only that $\underline{\underline{D}}$ is real and diagonal, but that its upper left and lower right blocks are identical. That will be useful in establishing the final result, which follows below. Before entering into its proof, we must introduce another preordering condition on the columns of $\underline{\underline{B}}$: referring to the notation of Eq.(8),

$$\textbf{for all } i: \ \Im(\underline{\underline{B}}_1^T \cdot \underline{\underline{B}}_2^*)_{ii} > 0 \ . \tag{9}$$

Effectively, this means that the diagonal elements of $\underline{\underline{D}}$ will all be positive. A few moments reflection will convince you that such an ordering is always possible. First, no D_{ii} will vanish, as that would imply $\det \underline{\underline{D}} = 0$, which is impossible. Second, if it happens that Eq.(9) is violated for some i, switching columns i and $i+n$ of $\underline{\underline{B}}$ will fix it, as $\underline{\underline{B}}$ already satisfies Eq.(4). With that done, we can proceed.

FUNDAMENTAL THEOREM: The columns of $\underline{\underline{B}}$ can be normalized so that

$$\underline{\underline{B}}^T \cdot \underline{\underline{J}} \cdot \underline{\underline{B}} = i\underline{\underline{J}} \ . \tag{10}$$

PROOF: We multiply $\underline{\underline{B}}$ on the right by a diagonal matrix, $\underline{\underline{W}}$, and use the previous lemmas to calculate.

$$\begin{aligned} (\underline{\underline{B}} \cdot \underline{\underline{W}})^T \cdot \underline{\underline{J}} \cdot (\underline{\underline{B}} \cdot \underline{\underline{W}}) &= \underline{\underline{W}}^T \cdot (\underline{\underline{B}}^T \cdot \underline{\underline{J}} \cdot \underline{\underline{B}}) \cdot \underline{\underline{W}} \\ &= \underline{\underline{W}} \cdot (i\underline{\underline{D}} \cdot \underline{\underline{J}}) \cdot \underline{\underline{W}} \\ &= i\underline{\underline{D}} \cdot \underline{\underline{W}} \cdot \underline{\underline{J}} \cdot \underline{\underline{W}} \end{aligned}$$

Again, do this calculation in 2×2 block form.

$$\begin{aligned} i\underline{\underline{D}} \cdot \underline{\underline{W}} \cdot \underline{\underline{J}} \cdot \underline{\underline{W}} &= i \begin{pmatrix} \underline{\underline{D}}' & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{D}}' \end{pmatrix} \cdot \begin{pmatrix} \underline{\underline{W}}_1 & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{W}}_2 \end{pmatrix} \cdot \begin{pmatrix} \underline{\underline{0}} & \underline{\underline{W}}_2 \\ -\underline{\underline{W}}_1 & \underline{\underline{0}} \end{pmatrix} \\ &= i \begin{pmatrix} \underline{\underline{D}}' & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{D}}' \end{pmatrix} \cdot \begin{pmatrix} \underline{\underline{0}} & \underline{\underline{W}}_1 \cdot \underline{\underline{W}}_2 \\ -\underline{\underline{W}}_2 \cdot \underline{\underline{W}}_1 & \underline{\underline{0}} \end{pmatrix} \\ &= i \begin{pmatrix} \underline{\underline{0}} & \underline{\underline{D}}' \cdot \underline{\underline{W}}_1 \cdot \underline{\underline{W}}_2 \\ -\underline{\underline{D}}' \cdot \underline{\underline{W}}_1 \cdot \underline{\underline{W}}_2 & \underline{\underline{0}} \end{pmatrix} \end{aligned}$$

To maintain Eq.(4), we must have $\underline{\underline{W}}_2 = \underline{\underline{W}}_1^*$. Thus, all that is necessary to obtain the desired result is to choose diagonal components for the submatrix $\underline{\underline{W}}_1$ so that

$$\textbf{for all } i, \;\; (\underline{\underline{D}}')_{ii} \, |\, (\underline{\underline{W}}_1)_{ii} \,|^2 = 1 \;\; .$$

This is always possible because, by our preordering assertion, Eq.(9), all $(\underline{\underline{D}}')_{ii}$ are positive.

□

COMMENTS:

10. There remains a phase ambiguity. Let $\underline{\underline{\Omega}}$ be any unitary matrix satisfying $\underline{\underline{\Omega}} \cdot \underline{\underline{\sigma}}_1 = \underline{\underline{\sigma}}_1 \cdot \underline{\underline{\Omega}}^*$. That is,

$$\underline{\underline{\Omega}} = \begin{pmatrix} \exp(-i\underline{\underline{\psi}}) & \underline{\underline{0}} \\ \underline{\underline{0}} & \exp(i\underline{\underline{\psi}}) \end{pmatrix} \;\; , \tag{11}$$

where $\underline{\underline{\psi}}$ is an arbitrary real, symmetric, $n \times n$ matrix. Then, $\underline{\underline{\Omega}}$ is itself symplectic – i.e. $\underline{\underline{\Omega}}^T \cdot \underline{\underline{J}} \cdot \underline{\underline{\Omega}} = \underline{\underline{J}}$ – so that

$$(\underline{\underline{B}} \cdot \underline{\underline{\Omega}})^T \cdot \underline{\underline{J}} \cdot (\underline{\underline{B}} \cdot \underline{\underline{\Omega}}) \;=\; \underline{\underline{\Omega}}^T \cdot \underline{\underline{B}}^T \cdot \underline{\underline{J}} \cdot \underline{\underline{B}} \cdot \underline{\underline{\Omega}} \;=\; \underline{\underline{\Omega}}^T \cdot i\underline{\underline{J}} \cdot \underline{\underline{\Omega}} \;=\; i\underline{\underline{J}} \;\; .$$

However, unless there are degeneracies, to keep Eq.(2) intact, $\underline{\underline{\psi}}$ should be diagonal, not just symmetric. Even with degeneracy, it must be diagonal in order to avoid mixing the eigenvectors.

In words, the normalization is not changed by multiplying every column in $\underline{\underline{B}}$ by a phase while retaining its sensible ordering. Thus, to finish specifying $\underline{\underline{B}}$ unambiguously requires one final and more or less arbitrary convention. The two most obvious ones are:

"β-convention": for $k = 1 \ldots n$, $B_{x_k a_k}$ is real and positive
"γ-convention": for $k = 1 \ldots n$, $B_{p_k a_k}$ is real and positive

The first is popular, the second is not, for a reason that will be clear soon, if it is not already.

11. Referring back to Eq.(7), notice that with the correct normalization,

$$d\underline{z}^T(\mathbf{u}) \cdot \underline{\underline{J}} \cdot d\underline{z}(\mathbf{v}) \;=\; d\underline{a}^T(\mathbf{u}) \cdot (i\underline{\underline{J}}) \cdot d\underline{a}(\mathbf{v}) \;\; .$$

Written in the language of two-forms, and using the fact that $\underline{\underline{B}}$ is sensibly ordered, this becomes

$$\sum_{k=1}^{n} dx_k \wedge dp_k \;=\; i \sum_{k=1}^{n} da_k \wedge da_{k+n} \;=\; i \sum_{k=1}^{n} da_k \wedge da_k^* \tag{12}$$

12. When phase space is two-dimensional and Eq.(10) is satisfied, then Eq.(8) provides the simple numerical condition, $B_2 B_1^* - B_1 B_2^* = 2i\Im(B_2 B_1^*) = -i$. If the phase is chosen such that B_1 is real, then this becomes $B_1 \Im(B_2) = -1/2$. We shall see two examples of this in the next section.

13. With this normalization, the connection between normal coordinates and action-angle coordinates, (φ, I), was written in the second chapter of ICD [2] – esp. p.24; Eq.(2.9) – and developed further in subsequent chapters: viz.

$$\textbf{for } k = 1, \ldots n: \quad a_k = i\sqrt{I_k}\, e^{-i\varphi_k}, \quad a_{k+n} = a_k^* = -i\sqrt{I_k}\, e^{i\varphi_k} \;\; . \tag{13}$$

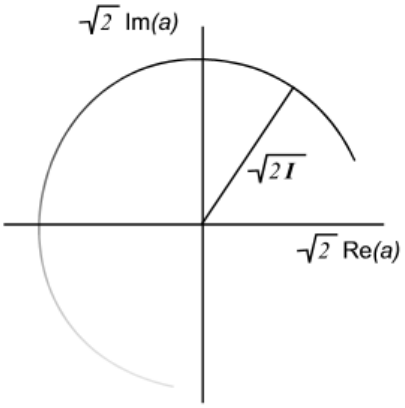

Figure 2: Normal coordinates rescaled so their real and imaginary parts present the visually correct phase space area. The angle variable, φ, is measured clockwise from the $\Im a$ axis.

We can partly justify this identification here by extending the chain of equivalences begun in Eq.(12). Within each two dimensional-invariant subspace, $da_k/a_k = dI_k/2I_k - i\,d\varphi_k$, so that

$$\begin{aligned} da_k \wedge da_k^* &= |a_k|^2 (dI_k/2I_k - i\,d\varphi_k) \wedge (dI_k/2I_k + i d\varphi_k) \\ &= (|a_k|^2/2I_k)(-i\,d\varphi_k \wedge dI_k + i\,dI_k \wedge d\varphi_k) \\ &= -i\,d\varphi_k \wedge dI_k \ . \end{aligned}$$

Putting this into Eq.(12) and summing extends the chain correctly.

$$\sum_{k=1}^{n} dx_k \wedge dp_k \ = \ i \sum_{k=1}^{n} da_k \wedge da_k^* \ = \ \sum_{k=1}^{n} d\varphi_k \wedge dI_k$$

14. A complex normal form coordinate can be separated into its real and imaginary parts, say $a_k = u_k + iv_k$, where u_k and v_k are real coordinates for the two-dimensional invariant subspace. However, the (oriented) phase space area element on this subspace is *not* $du_k \wedge dv_k$. Instead,

$$dx_k \wedge dp_k \ = \ i da_k \wedge da_k^* \ = \ i(du_k + i dv_k) \wedge (du_k - i dv_k) \ = \ 2\,du_k \wedge dv_k \ = \ d(\sqrt{2}u_k) \wedge d(\sqrt{2}v_k)$$

Put another way, real coordinates that preserve phase space area are $(\sqrt{2}\Re(a), \sqrt{2}\Im(a))$, and the corresponding polar "radius" coordinate is $\sqrt{2}\,|a| = \sqrt{2I}$. (See Figure 2.)

3.3 Equilibrium distributions of particles

Quadratic statistics for a distribution of particles subsist in the covariance matrix.

$$\underline{\underline{C}} = \langle\, \underline{z}\cdot\underline{z}^T \,\rangle - \langle\, \underline{z}\,\rangle \cdot \langle\, \underline{z}^T \,\rangle$$

If the system is linear and the distribution is in equilibrium, then $\langle\, \underline{z}\,\rangle = \underline{0}$, and this expression is reduced to its first term; if there are nonlinearities or the distribution is not in equilibrium, then the second term must be retained. In this section, we shall assume the former and write,

$$\underline{\underline{C}} \;=\; \langle\, \underline{z}\cdot\underline{z}^T \,\rangle \;=\; \underline{\underline{B}}\cdot\langle\, \underline{a}\cdot\underline{a}^\dagger \,\rangle\cdot\underline{\underline{B}}^\dagger \quad , \tag{14}$$

where I have used Eq.(3) and the obvious equality $(\underline{\underline{B}}\cdot\underline{a})^T = (\underline{\underline{B}}\cdot\underline{a})^\dagger$.

We now evaluate the central matrix. Equilibrium is achieved by distributing particles so that (1) action and angle coordinates are independent and (2) angle coordinates are distributed uniformly. That this works is obvious if one envisions the ensemble rotating in normalized phase space coordinates. Under these conditions,

$$\langle\, \underline{a}\cdot\underline{a}^\dagger \,\rangle_{kj} \;=\; \langle\, a_k a_j^* \,\rangle \;=\; \langle\, \sqrt{I_k I_j}\; e^{i(\varphi_j-\varphi_k)} \,\rangle \;=\; \langle\, \sqrt{I_k I_j}\,\rangle\langle\, e^{i(\varphi_j-\varphi_k)} \,\rangle \;=\; \langle\, I_k \,\rangle\,\delta_{kj} \quad .$$

Now put this back into Eq.(14) and write in terms of components.

$$C_{ij} \;=\; \sum_{k=1}^{2n} B_{ik}B_{jk}^*\,\langle\, I_k \,\rangle \;=\; \sum_{k=1}^{n} \Re(B_{ik}B_{jk}^*)\,\langle\, 2I_k \,\rangle \tag{15}$$

The diagonal elements are particularly simple.

$$C_{ii} \;=\; \sigma_i^2 \;=\; \sum_{k=1}^{n} |\, B_{ik}\,|^2\,\langle\, 2I_k \,\rangle \tag{16}$$

COMMENTS:

15. From what was said in previous sections about phase space area, we can make an obvious identification between the expectation value of actions and "emittances" for the distribution.

$$\varepsilon_k = \pi\,\langle 2I_k\rangle, \;\; \text{so that} \;\; C_{ij} \;=\; \sum_{k=1}^{n} \Re(B_{ik}B_{jk}^*)\,\varepsilon_k/\pi \tag{17}$$

This is valid even when the motion is coupled in all degrees of freedom, although the "emittances" are easier to interpret geometrically without coupling.

16. These expressions are not restricted to uncoupled motion, to Gaussian distributions, or by the number of degrees of freedom. The quadratic statistics of an equilibrium bunch in $2n$-dimenional phase space are completely determined by n numbers characterizing the bunch's "size," contained in the values of ε_k, and n^2 real numbers comprising the elements of the eigenvector matrix, $\underline{\underline{B}}$. If the system is symplectic, then only $n^2 - n(n-1)/2 = n(n+1)/2$ of the latter are independent. These become the "lattice parameters," or "lattice functions," of the machine, whose definitions are somewhat arbitrary. However, such parametrizations are *fundamentally* unnecessary, albeit sometimes convenient. The separation between "beam parameters" and "machine parameters" has already taken place; indeed it was already inherent in the definition of normal coordinates.

17. The phase ambiguity in the definition of $\underline{\underline{B}}$ is irrelevant for this calculation. If $\langle\, \underline{a}\cdot\underline{a}^\dagger \,\rangle$ is diagonal, then changing $\underline{\underline{B}}$ to

$\underline{\underline{B}} \cdot \underline{\underline{\Omega}}$ in Eq.(14), where $\underline{\underline{\Omega}}$ is a diagonal unitary matrix, results in the same covariance matrix *provided* the distribution is in equilibrium.

18. Using normal coordinates – or, equivalently, action-angle coordinates – to populate an invariant distribution is easily generalized to nonlinear dynamics. Normal coordinates can be calculated to any order desirable, after which one populates them just as in the linear case: uniformly in angle and using whatever distribution seems appropriate in the action coordinate. We shall return to this in a subsequent memo.

3.4 Phase

This brings us to the ever obnoxious issue of "phase advance." Let $\underline{\underline{M}}_{c \leftarrow b}$ represent the (linearized) transition "from point b to a different point c" in the ring, and let $\underline{\underline{M}}_b$ and $\underline{\underline{M}}_c$ be the one turn transit matrices at those two points. Because the system is assumed to be autonomous, these obey the master equation,

$$\underline{\underline{M}}_c \cdot \underline{\underline{M}}_{c \leftarrow b} = \underline{\underline{M}}_{c \leftarrow b} \cdot \underline{\underline{M}}_b \quad . \tag{18}$$

If this is not clear, think about it for twenty seconds, and it will become obvious. Also so obvious as to be barely worth mentioning is the crucial fact that, from this equation, $\underline{\underline{M}}_b$ and $\underline{\underline{M}}_c$ must possess the same eigenvalues.

Now, let $\underline{\underline{B}}_b$ and $\underline{\underline{B}}_c$ be the eigenvector matrices at those points, *with all normalization and ordering conventions intact.* Multiply Eq.(18) on the right by $\underline{\underline{B}}_b$ to get,

$$\underline{\underline{M}}_c \cdot (\underline{\underline{M}}_{c \leftarrow b} \cdot \underline{\underline{B}}_b) \;=\; \underline{\underline{M}}_{c \leftarrow b} \cdot \underline{\underline{M}}_b \cdot \underline{\underline{B}}_b \;=\; (\underline{\underline{M}}_{c \leftarrow b} \cdot \underline{\underline{B}}_b) \cdot \underline{\underline{\Lambda}} \quad .$$

If we assume no degeneracy – i.e., all eigenvalues are distinct, and therefore the eigenvectors are not rearranged or superposed – then it must be that $\underline{\underline{M}}_{c \leftarrow b} \cdot \underline{\underline{B}}_b = \underline{\underline{B}}_c \cdot \underline{\underline{D}}$, where $\underline{\underline{D}}$ is some diagonal matrix. However, because $\underline{\underline{M}}_{c \leftarrow b}$ is itself symplectic, $\underline{\underline{D}}$ does not change the normalization: viz.

$$(\underline{\underline{M}}_{c \leftarrow b} \cdot \underline{\underline{B}}_b)^T \cdot \underline{\underline{J}} \cdot (\underline{\underline{M}}_{c \leftarrow b} \cdot \underline{\underline{B}}_b) \;=\; \underline{\underline{B}}_b^T \cdot (\underline{\underline{M}}_{c \leftarrow b}^T \cdot \underline{\underline{J}} \cdot \underline{\underline{M}}_{c \leftarrow b}) \cdot \underline{\underline{B}}_b \;=\; \underline{\underline{B}}_b^T \cdot \underline{\underline{J}} \cdot \underline{\underline{B}}_b \;=\; i\underline{\underline{J}} \quad .$$

Therefore, it must be that $\underline{\underline{M}}_{c \leftarrow b} \cdot \underline{\underline{B}}_b$ and $\underline{\underline{B}}_c$ are related by a phase matrix,

$$\underline{\underline{M}}_{c \leftarrow b} \cdot \underline{\underline{B}}_b = \underline{\underline{B}}_c \cdot \underline{\underline{\Omega}}_{c \leftarrow b} \quad , \tag{19}$$

where $\underline{\underline{\Omega}}$ is of the form given in Eq.(11), with $\underline{\underline{\psi}}$ diagonal.

The (diagonal) matrix elements of $\underline{\underline{\Omega}}$ provide only the fractional part of the "phase advance." To obtain the integral part, Eq.(19) must be applied repeatedly at intervals small enough that $\Delta\psi_k < 2\pi$ for all k, and the incremental (fractional) phases accumulated and stored. The integral part increases by one every time an accumlated phase crosses the 2π boundary.

We have yet to connect the normal coordinates at b and c.

$$\underline{\underline{M}}_{c \leftarrow b} \cdot \underline{z}_b \;=\; \underline{z}_c \quad \Longrightarrow \quad \underline{\underline{M}}_{c \leftarrow b} \cdot \underline{\underline{B}}_b \cdot \underline{a}_b \;=\; \underline{\underline{B}}_c \cdot \underline{\underline{\Omega}}_{c \leftarrow b} \cdot \underline{a}_b \;=\; \underline{\underline{B}}_c \cdot \underline{a}_c \quad \Longrightarrow \quad \underline{\underline{\Omega}}_{c \leftarrow b} \cdot \underline{a}_b \;=\; \underline{a}_c$$

This is correct but violates the sometimes useful science fiction that orbits continuously "behave like a harmonic oscillator," which would require $\underline{a}_k \mapsto \underline{a}_k e^{-i2\pi\nu_k s/C} = \underline{a}_k e^{-i\nu_k \theta}$, not $\underline{a}_k \mapsto \underline{a}_k e^{-i\psi_k}$. (Note: C = path length of closed orbit.) This is recovered by rewriting Eq.(19),

$$\underline{\underline{M}}_{c \leftarrow b} \cdot \underline{\underline{B}}_b \;=\; \underline{\underline{B}}_c \cdot \underline{\underline{\Omega}} \cdot \exp(i\underline{\underline{\nu}}\theta) \cdot \exp(-i\underline{\underline{\nu}}\theta) \;=\; \underline{\underline{B}}_c \cdot \exp(-i(\underline{\underline{\psi}} - \underline{\underline{\nu}}\theta)) \cdot \exp(-i\underline{\underline{\nu}}\theta) \quad ,$$

and redefining $\underline{\underline{\tilde{B}}}_c \equiv \underline{\underline{B}}_c \cdot \exp(-i(\underline{\underline{\psi}} - \underline{\underline{\nu}}\theta)) \equiv \underline{\underline{B}}_c \cdot \exp(-i\underline{\underline{\tilde{\psi}}})$ with $\underline{\tilde{a}} \equiv \exp(-i\underline{\underline{\nu}}\theta) \cdot \underline{a}$ as the normal coordinate. This restores the desired science fiction by providing a third "picture," intermediate between Schrödinger's and Heisenberg's, in which both the state and the operators evolve. Again abusing quantum mechanical terminology, we can call it the "interaction picture."

COMMENTS:

19. The numbers stored in the matrix elements of $\underline{\underline{\Omega}}_{c\leftarrow b}$ *will depend on the conventions used to normalize* $\underline{\underline{B}}$. Referring to the nomenclature of Comment 20, below Eq.(11), the β-convention and the γ-convention – and anything in between – lead to different but equally legitimate values for "phase advance." Obviously, the popular one arises from the β-convention. We will briefly return to this point later.

20. Eq.(18) has the form, $\underline{\underline{F}} \cdot \underline{\underline{X}} - \underline{\underline{X}} \cdot \underline{\underline{G}} = \underline{\underline{C}}$, with $\underline{\underline{C}} = \underline{\underline{0}}$. Under a certain condition on $\underline{\underline{F}}$ and $\underline{\underline{G}}$, there is a unique solution for $\underline{\underline{X}}$ given any $\underline{\underline{C}}$.[8] One-turn transit matrices do not satisfy that condition, which is just as well, since $\underline{\underline{X}} = \underline{\underline{M}}_{c\leftarrow b} = \underline{\underline{0}}$ would then be the only solution. It is obvious from physical considerations that $\underline{\underline{M}}_{c\leftarrow b}$ cannot be determined from $\underline{\underline{M}}_b$ and $\underline{\underline{M}}_c$ alone.

21. Eq.(19) is contingent upon Eq.(18), which means the transition matrix, $\underline{\underline{M}}_{c\leftarrow b}$, must model a segment of the periodic structure in which both $\underline{\underline{M}}_b$ and $\underline{\underline{M}}_c$ are embedded as one-turn matrices. If $\underline{\underline{M}}_{c\leftarrow b}$ is an arbitrary symplectic matrix – e.g. linearly modeling transit between two points in a transfer line – then $\underline{\underline{\Omega}}_{c\leftarrow b}$ need not be unitary. Put another way, the columns of $\underline{\underline{B}}$, which are no longer eigenvectors,[9] can lose their normalization.

If $\underline{\underline{M}}_{c\leftarrow b}$ is the *complete* transition matrix associated with a transfer line connecting one ring to another, and if Eq.(19) is satisfied, with $\underline{\underline{M}}_b$ and $\underline{\underline{M}}_c$ being the one-turn matrices of the two rings at the extraction and injection points, and $\underline{\underline{\Omega}}$ unitary, then the transfer line is "matched." The task of finding such a line is the "matching problem" (normally expressed as "matching the lattice functions").

4 Parametrizations: "lattice functions"

For the case of linear(ized) dynamics, decompositions like Eq.(3) or Eq.(15) effect a separation between machine parameters, subsisting in $\underline{\underline{B}}$ and $\underline{\underline{\Lambda}}$, and the state information contained in $\underline{a}$. Any physically meaningful calculation relating observables can be formulated and done without introducing additional notation. For example, the answer to "What is the contribution of 'vertical emittance' to σ_x?" is "$|B_{xa_2}|^2\varepsilon_2/\pi$." (This begs the question: in any situation where $B_{xa_2} \neq 0$, do you understand what ε_2 means?)

"Parametrization" begins from the observation that symplecticity creates $N(N-1)/2$ quadratic relationships among the elements of $\underline{\underline{M}}$ or of $\underline{\underline{B}}$. Their N^2 elements can thus (generally) be expressed in terms of $N(N+1)/2 = n(2n+1)$ real variables, n of which are the phase angles (mod 2π). Within the context of accelerator physics these are called "lattice functions."

4.1 Harmonic oscillator

Before looking at parametrizations in accelerator physics, I pause to consider the harmonic oscillator, because (a) it is the universal archetype of all stable physical systems, and (b) to explain why I use the symbols I do. Finding the harmonic

[8] Proofs can be found in textbooks on linear algebra and matrices, such as Bellman [5]. A loosely stated, insufficient physical interpretation of the required condition is that matrices $\underline{\underline{F}}$ and $\underline{\underline{G}}$ must contract phase space, making the origin an attractor.

[9] At least not legitimate eigenvectors. Given any diagonal matrix, $\underline{\underline{\Lambda}}$, one can always artificially *define* a matrix $\underline{\underline{M}}$ according to $M \equiv \underline{\underline{B}} \cdot \underline{\underline{\Lambda}} \cdot \underline{\underline{B}}^{-1}$ so that the columns of $\underline{\underline{B}}$ become eigenvectors of $\underline{\underline{M}}$.

oscillator's eigenvector matrix, $\underline{\underline{B}}$, was done in ICD [2], Sec.2.1.1, pp.22-24. From there we shall lift only the two expressions relevant here.

The correspondence between (x,p) and the complex (a,a^*) coordinates is given by

$$\begin{pmatrix} x \\ p \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2m\omega} & 1/\sqrt{2m\omega} \\ -i\sqrt{m\omega/2} & i\sqrt{m\omega/2} \end{pmatrix} \cdot \begin{pmatrix} a \\ a^* \end{pmatrix} = \underline{\underline{B}} \cdot \begin{pmatrix} a \\ a^* \end{pmatrix} \ ,$$

where m and ω are, respectively, the mass and radial frequency parameters associated with the oscillator. It is trivially verified that Eq.(10) is satisfied. The inverse form of this is written,

$$\begin{pmatrix} a \\ a^* \end{pmatrix} = \underline{\underline{B}}^{-1} \cdot \begin{pmatrix} x \\ p \end{pmatrix} = \begin{pmatrix} \sqrt{m\omega/2} & i/\sqrt{2m\omega} \\ \sqrt{m\omega/2} & -i/\sqrt{2m\omega} \end{pmatrix} \cdot \begin{pmatrix} x \\ p \end{pmatrix} \ ,$$

validating that these are correctly normalized so that, as classical mechanics gives way to quantum mechanics, they become (apart from factors of $\hbar$) the annihilation and creation operators, $\mathbf{a}$ and $\mathbf{a}^\dagger$, of the harmonic oscillator system.

4.2 Transverse accelerator coordinates: one degree of freedom

For accelerator problems in one degree of freedom, the traditional parametrization connects matrix elements of $\underline{\underline{B}}$ with Courant-Snyder "lattice functions." There are several ways to proceed; I choose one here that begins from the identification given in Eq.(13) and the definition of Eq.(3) and ends with (something like) familiar "Courant-Snyder" expressions.

As already mentioned in Comment 22 below Eq.(12), for a two-dimensional phase space, the correct normalization follows from requiring the product,

$$\Im(B_{21}B_{11}^*) = -1/2 \ . \tag{20}$$

The most general matrix satisfying this can be written as follows in terms of two real parameters and a phase, not surprisingly notated here as α, β, and $e^{-i\tilde{\psi}}$.

$$\underline{\underline{B}} \equiv \frac{1}{\sqrt{2\beta}} \begin{pmatrix} \beta & \beta \\ -i-\alpha & i-\alpha \end{pmatrix} \begin{pmatrix} e^{-i\tilde{\psi}} & 0 \\ 0 & e^{i\tilde{\psi}} \end{pmatrix}$$

It is easy to verify directly that that Eq.(10) is satisfied for any choice of phase, $\tilde{\psi}$.

We now justify this notation by establishing the connection with familiar expressions. First use Eq.(3) and Eq.(13) to relate the normal coordinates with the original coordinates.

$$\begin{pmatrix} x \\ X \end{pmatrix} = \frac{1}{\sqrt{2\beta}} \begin{pmatrix} \beta & \beta \\ -i-\alpha & i-\alpha \end{pmatrix} \begin{pmatrix} e^{-i\tilde{\psi}} & 0 \\ 0 & e^{i\tilde{\psi}} \end{pmatrix} \cdot \begin{pmatrix} i\sqrt{I}e^{-i\varphi} \\ -i\sqrt{I}e^{i\varphi} \end{pmatrix}$$

I am using here a notation employed regularly in celestial mechanics, whereby the canonical momenta to coordinates (x,y,z) are written (X,Y,Z). This temporarily leaves it vague whether X refers to $x' = dx/ds = p_x/p_z = \tan\theta$ or to $p_x/p_{\rm ref} = \sin\theta$. Within the eikonal (paraxial) approximation *required* for the linear sector – i.e. $\theta^2 \ll 1$ – the two are equivalent. We shall return to this point when discussing the nonlinear sector in a subsequent memo.

Expanding the top row provides the following.

$$x \ = \ \sqrt{\beta/2}\ i\sqrt{I}\ \left(e^{-i(\tilde{\psi}+\varphi)} - e^{i(\tilde{\psi}+\varphi)}\right) \ = \ \sqrt{2\beta I}\ \sin(\tilde{\psi}+\varphi) \ = \ \sqrt{\beta}\ \left(\sqrt{2}\,\Re(e^{-i\tilde{\psi}}a)\right) \tag{21}$$

Anticipating our final result, instead of using the second row as written, we expand $\alpha x+\beta X$.

$$\begin{aligned}
\alpha x+\beta X &= (\alpha\ \beta)\cdot\begin{pmatrix} x \\ X \end{pmatrix} \\
&= \frac{1}{\sqrt{2\beta}}\,(-i\beta\ \ i\beta)\cdot\begin{pmatrix} e^{-i\tilde{\psi}}a \\ e^{i\tilde{\psi}}a^* \end{pmatrix} \\
&= \sqrt{\beta}\,\left(\sqrt{2}\,\Im(e^{-i\tilde{\psi}}a)\right) \\
&= \sqrt{2\beta I}\ \cos(\tilde{\psi}+\varphi)
\end{aligned}$$

All of which can be encapsulated into a single phasor equation.

$$x+i(\alpha x+\beta X)\ =\ \sqrt{2\beta}\ e^{-i\tilde{\psi}}a \tag{22}$$

This completes our justification of this notation and provides the expressions used by CHEF to calculate "lattice functions" from matrix elements of the eigenvector matrix, $\underline{\underline{B}}$. A "meta-algorithm" for doing so is written below; the matrix elements of $\underline{\underline{B}}$ are expressed, in pseudo-Dirac fashion, by the coordinates they connect rather than the (arbitrary) integers 1 and 2.

$$\begin{aligned}
\beta &= 2|B_{xa}|^2 \\
\alpha &= -2\,\mathrm{Re}(B_{xa^*}B_{Xa}) \\
e^{-i\tilde{\psi}} &= B_{xa}/|\,B_{xa}\,|
\end{aligned} \tag{23}$$

As a consistency condition (or test of correctness), the following must be satisfied as well.

$$1 = -2\,\mathrm{Im}(B_{xa^*}B_{Xa})$$

This merely recaptures the symplectic condition, as summarized in Eq.(20).

COMMENTS:

22. Different, equivalent parametrizations follow upon redefining the phase, $\tilde{\psi}$. As one (and only one) example, if we choose to make the second row real and positive, rather than the first, we could write,

$$\underline{\underline{B}}\ =\ \frac{1}{\sqrt{2\beta}}\begin{pmatrix} \beta & \beta \\ -i-\alpha & i-\alpha \end{pmatrix}\begin{pmatrix} e^{-i\tilde{\psi}} & 0 \\ 0 & e^{i\tilde{\psi}} \end{pmatrix}\ =\ \frac{1}{\sqrt{2\gamma}}\begin{pmatrix} i-\alpha & -i-\alpha \\ \gamma & \gamma \end{pmatrix}\begin{pmatrix} e^{-i(\tilde{\psi}+\xi)} & 0 \\ 0 & e^{i(\tilde{\psi}+\xi)} \end{pmatrix}\ ,$$

where $e^{i\xi} = (i-\alpha)/|\,i-\alpha\,|$, and $\beta\gamma-\alpha^2 = 1$. Instead of Eq.(22), we then would get,

$$X - i(\alpha X+\gamma x)\ =\ \sqrt{2\gamma}\ e^{-i\eta}a\ , \tag{24}$$

with $\eta = \tilde{\psi}+\xi$. The interpretation of this choice is too obvious to dwell upon here.

23. In just one degree of freedom, Eq.(16) and Eq.(17) are trivially connected with the more familiar relation,

$$\sigma_x^2\ =\ |B_{xa_1}|^2\varepsilon_1/\pi\ =\ \beta\varepsilon_1/2\pi\ . \tag{25}$$

This expression relating a directly measurable beam width to the emittance and the form of Eq.(21), as opposed Eq.(24) (X cannot be measured directly) are, with virtual certainty, why the β-convention is overwhelmingly preferred.

4.3 Transverse accelerator coordinates: two degrees of freedom

In going beyond one degree of freedom, we should make one last assertion regarding the order of the columns of $\underline{\underline{B}}$. It is not at all essential, but we (at least, I) would like coordinates $(a_1, a_3) = (a_1, a_1*)$ to refer to "mostly horizontal" motion and $(a_2, a_4) = (a_2, a_2*)$, to "mostly vertical."[10] This arrangement can be accomplished in several ways. For example, one could assert

$$|\, B_{z_k a_k} \,| + |\, B_{z_{k+n} a_{k+n}} \,| = \max_j (\, |\, B_{z_k a_j} \,| + |\, B_{z_{k+n} a_{j+n}} \,|\,) \quad . \tag{26}$$

The reader can come up with many other variations, all of which would lead to the same result *provided the coupling is weak.* For strong coupling, this ambiguity may be much more difficult, if not impossible, to resolve.

In the four dimensional phase space of transverse motion, two parametrizations found favor at Fermilab: one devised many years ago by Edwards and Teng [6], the other, more recently, by Lebedev and Bogacz [7]. The latter is written as follows.

$$\underline{\underline{B}} \equiv \sqrt{\frac{1}{2}} \begin{pmatrix} \beta_{1x}^{1/2} & e^{i\nu_2}\,\beta_{2x}^{1/2} & \beta_{1x}^{1/2} & e^{-i\nu_2}\,\beta_{2x}^{1/2} \\ e^{i\nu_1}\,\beta_{1y}^{1/2} & \beta_{2y}^{1/2} & e^{-i\nu_1}\,\beta_{1y}^{1/2} & \beta_{2y}^{1/2} \\ (-i(1-u)-\alpha_{1x})\beta_{1x}^{-1/2} & e^{i\nu_2}\,(-iu-\alpha_{2x})\beta_{2x}^{-1/2} & (i(1-u)-\alpha_{1x})\beta_{1x}^{-1/2} & e^{-i\nu_2}\,(iu-\alpha_{2x})\beta_{2x}^{-1/2} \\ e^{i\nu_1}\,(-iu-\alpha_{1y})\beta_{1y}^{-1/2} & (-i(1-u)-\alpha_{2y})\beta_{2y}^{-1/2} & e^{-i\nu_1}\,(iu-\alpha_{1y})\beta_{1y}^{-1/2} & (i(1-u)-\alpha_{2y})\beta_{2y}^{-1/2} \end{pmatrix}$$

This uses the notation of Ref.[7] but the normalization of Eq.(10), which introduces an extra $\sqrt{1/2}$. Eq.(9) is satisfied, the β-convention is used in both planes, and the notation is devised to mimic the two-dimensional case as closely as possible. The eleven parameteric symbols in this matrix are three too many: $4 \cdot 5/2 = 10 = 8 + 2$ column phases, which have been suppressed. Thus, the symplectic condition provides three additional equations that can be used to eliminate ν_1, ν_2, and u, leaving $[\beta|\alpha]_{[1|2][x|y]}$ as the preferred set of eight independent parameters. Explicit equations relating the three dependent parameters to this set can be found in Ref.[7]. CHEF does not use those expressions; if, for some reason, the values are desired, they can be obtained directly from $\underline{\underline{B}}$.

$$\begin{aligned} e^{i\nu_1} &= B_{ya_1}/|B_{ya_1}| \\ e^{i\nu_2} &= B_{xa_2}/|B_{xa_2}| \\ u &= 1 + \Im(B_{xa_1} B_{Xa_1}) \end{aligned}$$

The other eight are calculated in manners similar to what was done in the two-dimensional case, Eq.(23).

The Edwards-Teng approach did not use the eigenvector matrix directly but decomposed the one-turn transit matrix into three factors. Their description is made easier by rearranging phase space coordinates. Rather than splitting along "position" and "momentum" sectors as we have been, we shall group the coordinates into "horizontal" and "vertical" sectors, that is, $(z_1, Z_1; z_2, Z_2)^T$ instead of $(z_1, z_2; Z_1, Z_2)^T$. Then, in place of our starting point, Eq.(2), the Edwards-Teng decomposition is written,[11]

$$\underline{\underline{M}} \cdot \begin{pmatrix} \cos\phi\,\underline{1} & -\sin\phi\,\underline{\underline{D}}^{-1} \\ \sin\phi\,\underline{\underline{D}} & \cos\phi\,\underline{1} \end{pmatrix} = \begin{pmatrix} \cos\phi\,\underline{1} & \sin\phi\,\underline{\underline{D}}^{-1} \\ -\sin\phi\,\underline{\underline{D}} & \cos\phi\,\underline{1} \end{pmatrix} \cdot \begin{pmatrix} \underline{\underline{H}} & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{V}} \end{pmatrix} \ ,$$

where $\underline{\underline{D}}$ is a 2×2 unimodular matrix (three parameters) and $\underline{\underline{H}}$ and $\underline{\underline{V}}$ have the form of uncoupled two-dimensional one-turn transit matrices (three parameters each, including a phase). This has a structure similar to Eq.(2), but all matrices

[10] To anticipate, this notation might produce confusion when we get around to discussing CHEF's code, as programming indices begin with zero, not one. Thus, (a_1, a_3) will become (a[0],a[2]), and so forth.

[11] This is not the form taken in the original paper. I have shifted one matrix to the left hand side to highlight the correspondence with Eq.(2).

are real. It provides a similarity transformation relating $\underline{\underline{M}}$ to its block-diagonal form, operating on two *real* two-dimensional invariant subspaces, the ones spanned by complex conjugate normal coordinates.

Reference [6] contains a procedure for obtaining the parameters, but they also can be generated from the matrix elements of $\underline{\underline{B}}$ by combining complex conjugate columns to form real bases for the invariant subspaces. CHEF still provides them, but its usage has been deprecated.

4.4 Accelerator coordinates: three degrees of freedom

If a non-accelerating cavity is used to confine longitudinal motion, then the transit matrix of the linearized system is 6×6, and there are twenty-one independent parameters, of which three are phase angles and the other eighteen can be interpreted as "lattice functions." CHEF currently does not provide such a parametric set. It nonetheless may be worth comparing the familiar relation,

$$\sigma_x^2 = \beta_x \varepsilon_x / 2\pi + D^2 \sigma_{\delta p/p}^2 \ , \tag{27}$$

which relates horizontal beamwidth to horizontal emittance, dispersion and momentum spread, with Eq.(16) and Eq.(17).

$$\begin{aligned} \sigma_{z_1}^2 &= |B_{z_1 a_1}|^2 \varepsilon_1/\pi + |B_{z_1 a_2}|^2 \varepsilon_2/\pi + |B_{z_1 a_3}|^2 \varepsilon_3/\pi \\ \sigma_{z_3}^2 &= |B_{z_3 a_1}|^2 \varepsilon_1/\pi + |B_{z_3 a_2}|^2 \varepsilon_2/\pi + |B_{z_3 a_3}|^2 \varepsilon_3/\pi \end{aligned}$$

As was done in Eq.(25), we define $\beta_x/2 \equiv |B_{z_1 a_1}|^2$. If we assume $\varepsilon_2 = 0$, then ε_3 can be eliminated to get,

$$\left(\beta_x/2 \ - \ |B_{z_3 a_1} / B_{z_3 a_3}|^2 \right) \cdot \varepsilon_1/\pi \ + \ \left(|B_{z_1 a_3}|^2 / |B_{z_3 a_3}|^2 \right) \cdot \sigma_{z_3}^2 \ . \tag{28}$$

A direct comparison between Eq.(28) and Eq.(27) would be a subject in its own right, because the physical situations are different. To begin with, the value of $\beta_x/2$ may not be the same, as could the values of the "horizontal emittance." Further, Eq.(27) assumes that $\delta p/p$ remains constant for each particle in the bunch, so that the dispersion, D, is well defined as the closed orbit at that momentum offset. Eq.(28) assumes $\delta p/p$ oscillates at something close to the synchrotron frequency and "dispersion" has not even been defined. I shall postpone discussing this to another day, after we have been forced to pin down physical interpretations for transverse and longitudinal coordinates.

References

[1] Leo Michelotti and Jean-François Ostiguy. CHEF: a framework for accelerator optics and simulation. In *Proceedings of 9th International Computational Accelerator Physics Conference*, 2006 (ICAP 2006: TUAPMP02). Chamonix Mont-Blanc, France, October 2-6, 2006. (See also: FERMILAB-CONF-06-373-AD.)

[2] Leo Michelotti. *Intermediate Classical Dynamics with Applications to Beam Physics*. John Wiley & Sons, Inc., New York, 1995.

[3] J. A. Holt, M. A. Martens, L. Michelotti, and G. Goderre. Calculating luminosity for a coupled tevatron lattice. In *Proceedings of the IEEE Particle Accelerator Conference. Dallas, TX.*, 1995.

[4] Victor Guillemin and Shlomo Sternberg. *Symplectic Techniques in Physics*. Cambridge University Press, New York, 1984.

[5] Richard Bellman. *Introduction to Matrix Analysis*. McGraw-Hill, New York, 1970.

[6] D. A. Edwards and L. C. Teng. Parametrization of linear coupled motion in periodic systems. *IEEE Transactions on Nuclear Science*, NS-20(3):885, 1973. Proceedings of the 1973 IEEE Particle Accelerator Conference. (March 5-7).

[7] V. Lebedev and S. Bogacz. Betatron motion with coupling of horizontal and vertical degrees of freedom. Technical report, JLAB, 1999. JLAB-ACC-99-19.