A General Slice Moment Decomposition of RMS Beam Emittance

Chad Mitchell

Lawrence Berkeley National Laboratory, Berkeley, California, 94720, USA

February 24, 2015

Abstract

The square of the horizontal projected (rms) beam emittance is expressed as the sum of four nonnegative contributions, each described using the slice moments of the beam and possessing a natural interpretation in terms of the geometrical properties of the beam in the six-dimensional phase space. The mathematical formalism describing the relationships between projected beam quantities and slice beam quantities is reviewed. The results may be used to reconstruct the moments and emittances of the beam from the moments of its subpopulations, as well as to isolate and better understand a variety of slice and interslice dynamical contributions to the projected beam emittance growth.

1 Introduction

The concepts of slice and projected beam emittance play a central role in the design and optimization of next-generation FEL light sources [1], in the theory of emittance compensation [2]-[3], and in the general theory of beam dynamics in linear and circular accelerators [4]. The slice formalism is particularly useful in the case of long beams with small uncorrelated energy spread, for which the longitudinal slice-to-slice variation in the fields experienced by particles in the beam (e.g., due to wakefields and time-dependent RF fields) can be significant, while there is little particle slippage between slices. In this case, the dynamics of particles within each slice may profitably be separated from the dynamics of the slice centroids.

This separation leads naturally to the question of the relationship between the projected (rms) emittance and the properties of the various beam slices. We demonstrate that the square of the projected emittance can be decomposed as the sum of four contributions, each with a distinct geometrical interpretation in terms of the slice beam moments. Fragments of a treatment along these lines can be found in various forms in the literature [5]-[9]. The purpose of this paper is to consolidate these results in the form of a general and carefully presented mathematical treatment, with proofs of some nontrivial features included. Several tools are presented that point the way toward a similar treatment of other diagnostic beam quantities such as the projected beam energy spread, the 4D and 6D emittances, and the eigenemittances for coupled beams [10].

2 Basic formalism

Let f denote a beam distribution function on the 6D phase space, which we parameterize by canonical coordinates $\mathbf{X} = (x, p_x, y, p_y, z, p_z)$. For convenience, we will assume that all momenta are normalized by the quantity mc. By assumption, f has the properties that $f \geq 0$ and

$$\int f(\mathbf{X}) d\mathbf{X} = 1. \tag{1}$$

Thus, f is a joint probability density in the six random variables x, p_x, y, p_y, z, and p_z that characterizes the distribution of the beam particles. If ϕ is any function defined on the phase space, we define its beam average in the natural way:

$$\langle \phi \rangle = \int \phi(\mathbf{X}) f(\mathbf{X}) d\mathbf{X}. \tag{2}$$

Assume that z is the direction of beam motion, and consider the planes of constant z, which we describe as *longitudinal slices*. Let $\zeta = (x, p_x, y, p_y, p_z)$ denote the remaining phase space coordinates, so we will write $\mathbf{X} = (z, \zeta)$ when we wish to emphasize the separation of the slice coordinate z from the coordinates within a slice ζ. There are two natural functions associated with this decomposition into slices. First, the *longitudinal density profile* of the beam is given by the following marginal probability density:

$$\lambda(z) = \int f(z, \zeta) d\zeta. \tag{3}$$

Second, the *slice distribution function* for the slice z is given by the conditional probability density:

$$f(\zeta|z) = \frac{f(z,\zeta)}{\int f(z,\zeta) d\zeta} = \frac{f(z,\zeta)}{\lambda(z)}, \quad \text{provided} \quad \lambda(z) > 0, \tag{4}$$

with $f(\zeta|z) = 0$ when $\lambda(z) = 0$. It follows that both of these quantities are nonnegative and normalized so that:

$$\int \lambda(z) dz = 1, \qquad \int f(\zeta|z) d\zeta = 1. \tag{5}$$

Given any function ϕ defined on the phase space, we may define its average within slice z by using the slice distribution function:

$$\langle \phi \rangle_z = \int \phi(z, \zeta) f(\zeta|z) d\zeta. \tag{6}$$

We use the notation $f|z$ to refer to the the slice distribution function when we wish to avoid listing its arguments, in order to distinguish this quantity from the full beam distribution function f.

3 Slice and projected moments

We use the notation $\vec{\mu}$ and Σ to denote the centroid vector and covariance matrix of the beam, respectively, whose elements are given by:

$$\mu_a = \langle X_a \rangle, \qquad \Sigma_{ab} = \langle (X_a - \mu_a)(X_b - \mu_b) \rangle \qquad \text{for} \quad a, b = 1, \ldots, 6. \tag{7}$$

Likewise, define the slice centroid vector $\vec{\mu}(z)$ and slice covariance matrix $\Sigma(z)$ for slice z by:

$$\mu_a(z) = \langle X_a \rangle_z, \qquad \Sigma_{ab}(z) = \langle (X_a - \mu_a(z))(X_b - \mu_b(z)) \rangle_z \qquad \text{for} \quad a = 1, \ldots, 6. \tag{8}$$

Note that we have used two distinct forms of averaging $\langle\rangle$ and $\langle\rangle_z$, associated with the probability densities f and $f|z$. Given a quantity ϕ that depends on z, we may also define an average across slices by using the longitudinal density profile λ:

$$\mathrm{E}[\phi] = \int \phi(z)\lambda(z)dz. \tag{9}$$

Likewise, given two $z-$dependent quantities ϕ_1 and ϕ_2, we define their covariance in analogy with (7) and (8):

$$\mathrm{Cov}[\phi_1, \phi_2] = \mathrm{E}[(\phi_1 - \mathrm{E}[\phi_1])(\phi_2 - \mathrm{E}[\phi_2])], \tag{10}$$

and the variance is defined in the usual way as:

$$\mathrm{Var}[\phi] = \mathrm{Cov}[\phi, \phi] = \mathrm{E}[(\phi - \mathrm{E}[\phi])^2]. \tag{11}$$

The projected (7) and slice (8) moments of the beam can now be related using two fundamental results from the theory of conditional probability. The *law of total expectation* gives:

$$\mu_a = \mathrm{E}[\mu_a(z)] \qquad \text{for} \quad a = 1, \ldots, 6, \tag{12}$$

and the *law of total covariance* gives:

$$\Sigma_{ab} = \mathrm{E}[\Sigma_{ab}(z)] + \mathrm{Cov}[\mu_a(z), \mu_b(z)] \qquad \text{for} \quad a, b = 1, \ldots, 6. \tag{13}$$

Note that the notation in (12-13) makes sense, because the slice moments appearing within the brackets $[\cdot]$ on each right-hand side are $z-$dependent quantities. A proof of each of these claims is provided in Appendix A.

The projected (rms) normalized emittance of the beam in the (x, p_x) plane is typically defined by:

$$\epsilon_x^2 = \langle (x - \langle x \rangle)^2 \rangle \langle (p_x - \langle p_x \rangle)^2 \rangle - \langle (x - \langle x \rangle)(p_x - \langle p_x \rangle) \rangle^2, \tag{14}$$

with expressions of the same form for ϵ_y^2 and ϵ_z^2. It is convenient to express these definitions using the 2×2 blocks of the 6D covariance matrix associated with each degree of freedom:

$$\Sigma^X = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \qquad \Sigma^Y = \begin{pmatrix} \Sigma_{33} & \Sigma_{34} \\ \Sigma_{43} & \Sigma_{44} \end{pmatrix} \qquad \Sigma^Z = \begin{pmatrix} \Sigma_{55} & \Sigma_{56} \\ \Sigma_{65} & \Sigma_{66} \end{pmatrix}, \tag{15}$$

so that:

$$\epsilon_x^2 = \det \Sigma^X, \quad \epsilon_y^2 = \det \Sigma^Y, \quad \epsilon_z^2 = \det \Sigma^Z. \tag{16}$$

In the same way, we may express the emittances within slice z as:

$$\epsilon_x^2(z) = \det \Sigma^X(z), \quad \epsilon_y^2 = \det \Sigma^Y(z), \quad \epsilon_z^2 = \det \Sigma^Z(z), \tag{17}$$

where $\Sigma^X(z)$, $\Sigma^Y(z)$, and $\Sigma^Z(z)$ are obtained from the slice covariance matrix $\Sigma(z)$ defined in (8). It follows from our definitions that $\epsilon_z(z) = 0$ for all z, so this quantity will play no role in the discussions to follow.

4 Decomposition of the projected emittance

In this section we will show that the horizontal projected emittance (14) can be written as the sum in quadrature of four contributions, each with a natural interpretation in terms of the slice properties of the beam. Corresponding results apply also for the projected emittances in the other two planes.

Define a 6×6 matrix Λ according to:

$$\Lambda_{ab} = \mathrm{Cov}[\mu_a(z), \mu_b(z)], \qquad \text{for} \quad a, b = 1, \ldots, 6 \tag{18}$$

and likewise define the 2×2 blocks of Λ associated with each degree of freedom:

$$\Lambda^X = \begin{pmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{pmatrix} \qquad \Lambda^Y = \begin{pmatrix} \Lambda_{33} & \Lambda_{34} \\ \Lambda_{43} & \Lambda_{44} \end{pmatrix} \qquad \Lambda^Z = \begin{pmatrix} \Lambda_{55} & \Lambda_{56} \\ \Lambda_{65} & \Lambda_{66} \end{pmatrix}. \tag{19}$$

Then it follows from (13) that we have the 2×2 matrix identities:

$$\Sigma^X = \mathrm{E}[\Sigma^X(z)] + \Lambda^X, \quad \Sigma^Y = \mathrm{E}[\Sigma^Y(z)] + \Lambda^Y, \quad \Sigma^Z = \mathrm{E}[\Sigma^Z(z)] + \Lambda^Z. \tag{20}$$

Therefore according to (16) the horizontal projected emittance is given by:

$$\epsilon_x^2 = \det\left(\mathrm{E}[\Sigma^X(z)] + \Lambda^X\right). \tag{21}$$

We exploit the following useful identity that applies for any sum of 2×2 matrices, which the reader may directly verify:

$$\det(A + B) = \det(A) + \det(B) + \mathrm{tr}(A)\,\mathrm{tr}(B) - \mathrm{tr}(AB). \tag{22}$$

This gives:

$$\begin{aligned} \epsilon_x^2 &= \det(\mathrm{E}[\Sigma^X(z)]) + \det(\Lambda^X) + \mathrm{tr}(\mathrm{E}[\Sigma^X(z)])\,\mathrm{tr}(\Lambda^X) - \mathrm{tr}(\mathrm{E}[\Sigma^X(z)]\Lambda^X) \\ &= \det(\mathrm{E}[\Sigma^X(z)]) + \epsilon_{\mathrm{int}}^2 + \epsilon_{||}^2, \end{aligned} \tag{23}$$

where we have defined:

$$\epsilon_{||}^2 = \det \Lambda^X, \tag{24}$$

$$\epsilon_{\mathrm{int}}^2 = \mathrm{tr}(\mathrm{E}[\Sigma^X(z)])\,\mathrm{tr}(\Lambda^X) - \mathrm{tr}(\mathrm{E}[\Sigma^X(z)]\Lambda^X). \tag{25}$$

The significance of these quantities will soon become apparent. We further decompose the first term appearing on the right-hand side of (23) as follows:

$$\begin{aligned} \det(\mathrm{E}[\Sigma^X(z)]) &= \mathrm{E}[\Sigma_{11}(z)]\,\mathrm{E}[\Sigma_{22}(z)] - \mathrm{E}[\Sigma_{12}(z)]^2 \\ &= \mathrm{E}[\Sigma_{11}(z)\Sigma_{22}(z)] - \mathrm{Cov}[\Sigma_{11}(z), \Sigma_{22}(z)] + \mathrm{Var}[\Sigma_{12}(z)] - \mathrm{E}[\Sigma_{12}(z)^2] \\ &= \mathrm{E}[\epsilon_x^2(z)] - \mathrm{Cov}[\Sigma_{11}(z), \Sigma_{22}(z)] + \mathrm{Var}[\Sigma_{12}(z)] \\ &= \mathrm{E}[\epsilon_x(z)]^2 + \mathrm{Var}[\epsilon_x(z)] - \mathrm{Cov}[\Sigma_{11}(z), \Sigma_{22}(z)] + \mathrm{Var}[\Sigma_{12}(z)] \\ &= \epsilon_\perp^2 + \epsilon_R^2, \end{aligned} \tag{26}$$

where:

$$\epsilon_\perp^2 = \mathrm{E}[\epsilon_x(z)]^2, \tag{27}$$

$$\epsilon_R^2 = \mathrm{Var}[\epsilon_x(z)] - \mathrm{Cov}[\Sigma_{11}(z), \Sigma_{22}(z)] + \mathrm{Var}[\Sigma_{12}(z)]. \tag{28}$$

Thus, we have shown that:

$$\epsilon_x^2 = \epsilon_\perp^2 + \epsilon_R^2 + \epsilon_{\mathrm{int}}^2 + \epsilon_{||}^2. \tag{29}$$

In Appendix B, we prove that each of these four contributions is nonnegative, so that $\epsilon_\perp$, ϵ_R, ϵ_{int} and $\epsilon_{||}$ are always real quantities.

Changing to a more intuitive notation, let us write the first moments of slice z in the form:

$$\mu_x(z) = \mu_1(z) \qquad \mu_{p_x}(z) = \mu_2(z), \tag{30}$$

which denote the horizontal position and momentum of the slice centroid. Similarly, let us write the second moments of the slice z in the form:

$$\sigma_x^2(z) = \Sigma_{11}(z), \qquad \sigma_{p_x}^2(z) = \Sigma_{22}(z), \qquad \langle \Delta x \Delta p_x \rangle_z = \Sigma_{12}(z). \tag{31}$$

Expressing the four emittance contributions in terms of these quantities gives, after writing out (24-25) explicitly:

$$\epsilon_\perp^2 = \mathrm{E}[\epsilon_x(z)]^2, \tag{32a}$$

$$\epsilon_R^2 = \mathrm{Var}[\epsilon_x(z)] - \mathrm{Cov}[\sigma_x^2(z), \sigma_{p_x}^2(z)] + \mathrm{Var}[\langle \Delta x \Delta p_x \rangle_z], \tag{32b}$$

$$\epsilon_{\mathrm{int}}^2 = \mathrm{E}[\sigma_x^2(z)]\,\mathrm{Var}[\mu_{p_x}(z)] + \mathrm{E}[\sigma_{p_x}^2(z)]\,\mathrm{Var}[\mu_x(z)] - 2\,\mathrm{E}[\langle x p_x \rangle_z]\,\mathrm{Cov}[\mu_x(z), \mu_{p_x}(z)], \tag{32c}$$

$$\epsilon_{||}^2 = \mathrm{Var}[\mu_x(z)]\,\mathrm{Var}[\mu_{p_x}(z)] - \mathrm{Cov}[\mu_x(z), \mu_{p_x}(z)]^2. \tag{32d}$$

Let us consider each of these contributions in turn.

1. The quantity

$$\epsilon_\perp = \mathrm{E}[\epsilon_x(z)]$$

is the *mean slice emittance*. It is determined by computing the emittance within each longitudinal beam slice, and then taking the mean among slices, with each slice weighted by its corresponding longitudinal density. It vanishes if and only if every slice of the beam has zero emittance.

2. The quantity ϵ_R is the *mismatch emittance*, and its value is determined by the size of variation in the beam second moments (the covariance matrices $\Sigma^X(z)$) from slice to slice. When $\epsilon_x(z) \neq 0$, it is typical to define the slice Twiss functions $\alpha(z)$, $\beta(z)$, and $\gamma(z)$ such that[1]:

$$\Sigma^X(z) = \epsilon_x(z) \begin{pmatrix} \beta(z) & -\alpha(z) \\ -\alpha(z) & \gamma(z) \end{pmatrix} \quad \text{with} \quad \beta(z)\gamma(z) - \alpha(z)^2 = 1. \tag{33}$$

If the slice Twiss functions $\alpha(z)$, $\beta(z)$, and $\gamma(z)$ are identical for every slice z, then ϵ_R vanishes.

[1]Each mention of the Twiss functions in this work refers to the horizontal Twiss functions, so the subscripts in α_x, β_x, and γ_x have been suppressed; there is no possiblity of confusion with the relativistic factors β and γ, since these appear only once in (66). Since the emittance in (33) is normalized, the Twiss functions here differ from the geometric Twiss functions by the appropriate relativistic factor.

3. The quantity ϵ_{int} is the *linear misalignment emittance*, and its value is determined by the size of variation in the first moments (slice centroids) from slice to slice. This is a "cross term", as the contributions from the centroid variations are themselves weighted by the mean second moments. If the centroids of all beam slices are aligned, so that $\mu_x(z)$ and $\mu_{p_x}(z)$ are identical for every slice, then ϵ_{int} vanishes.

4. The quantity $\epsilon_{||}$ is the *nonlinear misalignment emittance*, and its value is also determined by the size of variation in the first moments (slice centroids) from slice to slice. Unlike ϵ_{int}, it depends exclusively on the slice centroid coordinates. In particular, $\epsilon_{||}$ is the beam emittance that would result if all charge within each slice were concentrated at the location of the corresponding slice centroid. If all slice centroids are aligned, then $\epsilon_{||}$ vanishes.

Appendix B describes the conditions under which each of these contributions vanishes. The reason for separating the terms ϵ_{int} and $\epsilon_{||}$ will become more clear in the following section.

5 Special cases

The decomposition of the previous section is valid for a general beam distribution. However, there are several idealized special cases that arise frequently in the literature. We will consider four of these in this section, with the hope that this will clarify the significance of the four contributions given in (32).

5.1 Uniform slice emittance

First, suppose that the emittance of every longitudinal beam slice is identical. More precisely, assume that:

$$\epsilon_x(z) = \mathrm{E}[\epsilon_x(z)] = \epsilon_\perp \tag{34}$$

for every z with $\lambda(z) \neq 0$. Provided $\epsilon_\perp \neq 0$, it is natural to define the slice Twiss functions $\alpha(z)$, $\beta(z)$, and $\gamma(z)$ as in (33), so that

$$\sigma_x^2(z) = \epsilon_\perp \beta(z), \quad \sigma_{p_x}^2(z) = \epsilon_\perp \gamma(z), \quad \langle \Delta x \Delta p_x \rangle_z = -\epsilon_\perp \alpha(z). \tag{35}$$

In this case the contribution $\epsilon_\perp$ is given simply by (34), and the contribution $\epsilon_{||}$ is unchanged from its original form (32). We can re-express the remaining two contributions in terms of the Twiss functions as:

$$\epsilon_R^2 = \epsilon_\perp^2 \left\{ -\operatorname{Cov}[\beta(z), \gamma(z)] + \operatorname{Var}[\alpha(z)] \right\}, \tag{36a}$$

$$\epsilon_{\text{int}}^2 = \epsilon_\perp \left\{ \mathrm{E}[\beta(z)] \operatorname{Var}[\mu_{p_x}(z)] + \mathrm{E}[\gamma(z)] \operatorname{Var}[\mu_x(z)] + 2\,\mathrm{E}[\alpha(z)] \operatorname{Cov}[\mu_x(z), \mu_{p_x}(z)] \right\}. \tag{36b}$$

Here we have used the facts that for any constant a and any $z-$dependent functions ϕ, ϕ_1, and ϕ_2:

$$\mathrm{E}[a\phi] = a\,\mathrm{E}[\phi], \quad \operatorname{Cov}[a\phi_1, a\phi_2] = a^2 \operatorname{Cov}[\phi_1, \phi_2], \quad \operatorname{Var}[a\phi] = a^2 \operatorname{Var}[\phi]. \tag{37}$$

Also, we used the fact that due to the assumption in (34):

$$\operatorname{Var}[\epsilon_x(z)] = \mathrm{E}[(\epsilon_x(z) - \mathrm{E}[\epsilon_x(z)])^2] = 0. \tag{38}$$

The expression (36a) indicates that the mismatch emittance ϵ_R is determined by the size of slice-to-slice variations in the beam Twiss functions. However, it is convenient (and somewhat surprising) that ϵ_R can also be written directly in terms of the *mean* Twiss functions. This can be seen by noting from (33) that:

$$1 = \mathrm{E}[\beta(z)\gamma(z) - \alpha^2(z)] = \mathrm{E}[\beta(z)\gamma(z)] - \mathrm{E}[\alpha^2(z)] \tag{39}$$

$$= \mathrm{Cov}[\beta(z), \gamma(z)] + \mathrm{E}[\beta(z)]\,\mathrm{E}[\gamma(z)] - \mathrm{Var}[\alpha(z)] - \mathrm{E}[\alpha(z)]^2, \tag{40}$$

so:

$$\epsilon_R^2 = \epsilon_\perp^2 \left\{ \mathrm{E}[\beta(z)]\,\mathrm{E}[\gamma(z)] - \mathrm{E}[\alpha(z)]^2 - 1 \right\}. \tag{41}$$

When $\beta(z)$, $\gamma(z)$, and $\alpha(z)$ are each independent of z, we see that this quantity indeed vanishes as a consequence of (33).

5.2 Uniform slice Twiss functions

Suppose that the Twiss functions of every longitudinal beam slice are identical. More precisely, assume that $\epsilon_x(z) \neq 0$ and:

$$\alpha(z) = \mathrm{E}[\alpha(z)], \quad \beta(z) = \mathrm{E}[\beta(z)], \quad \gamma(z) = \mathrm{E}[\gamma(z)], \tag{42}$$

for every z with $\lambda(z) \neq 0$. We let α_0, β_0, γ_0 denote the uniform values given in (42). The mismatch emittance in (32) then becomes:

$$\begin{aligned} \epsilon_R^2 &= \mathrm{Var}[\epsilon_x(z)] - \mathrm{Cov}[\epsilon_x(z)\beta_0, \epsilon_x(z)\gamma_0] + \mathrm{Var}[-\epsilon_x(z)\alpha_0] \\ &= \mathrm{Var}[\epsilon_x(z)] - \beta_0\gamma_0\,\mathrm{Cov}[\epsilon_x(z), \epsilon_x(z)] + \alpha_0^2\,\mathrm{Var}[\epsilon_x(z)] \\ &= \mathrm{Var}[\epsilon_x(z)] - (\beta_0\gamma_0 - \alpha_0^2)\,\mathrm{Var}[\epsilon_x(z)] = 0, \end{aligned} \tag{43}$$

since $\beta_0\gamma_0 - \alpha_0^2 = 1$. Thus, the misalignment emittance vanishes. The contributions $\epsilon_\perp$ and $\epsilon_{||}$ are independent of the slice second moments, and their form remains unchanged. The remaining contribution becomes:

$$\begin{aligned} \epsilon_{\mathrm{int}}^2 &= \mathrm{E}[\epsilon_x(z)\beta_0]\,\mathrm{Var}[\mu_{p_x}(z)] + \mathrm{E}[\epsilon_x(z)\gamma_0]\,\mathrm{Var}[\mu_x(z)] - 2\,\mathrm{E}[-\epsilon_x(z)\alpha_0]\,\mathrm{Cov}[\mu_x(z), \mu_{p_x}(z)] \\ &= \epsilon_\perp \left\{ \beta_0\,\mathrm{Var}[\mu_{p_x}(z)] + \gamma_0\,\mathrm{Var}[\mu_x(z)] + 2\alpha_0\,\mathrm{Cov}[\mu_x(z), \mu_{p_x}(z)] \right\}. \end{aligned} \tag{44}$$

5.3 Aligned slice centroids

Next, suppose that all slice centroids are aligned. More precisely, assume that:

$$\mu_x(z) = \mathrm{E}[\mu_x(z)], \qquad \mu_{p_x}(z) = \mathrm{E}[\mu_{p_x}(z)], \tag{45}$$

for every z with $\lambda(z) \neq 0$. It immediately follows that:

$$\mathrm{Var}[\mu_x(z)] = \mathrm{Var}[\mu_{p_x}(z)] = \mathrm{Cov}[\mu_x(z), \mu_{p_x}(z)] = 0, \tag{46}$$

and so $\epsilon_{\mathrm{int}} = \epsilon_{||} = 0$. Thus, the only projected emittance contributions are due to the mean slice emittance and the mismatch emittance.

This case is often treated by introducing a *slice mismatch parameter* [11]-[13], given by:

$$\zeta(z) = \frac{1}{2}\left\{\gamma(z)\beta - 2\alpha(z)\alpha + \beta(z)\gamma\right\}, \tag{47}$$

where α, β, and γ are the projected Twiss functions, defined by:

$$\Sigma^X = \epsilon_x \begin{pmatrix} \beta & -\alpha \\ -\alpha & \gamma \end{pmatrix} \quad \text{with} \quad \beta\gamma - \alpha^2 = 1, \tag{48}$$

and ϵ_x is the projected emittance. An alternative expression is given in terms of the projected and slice Twiss matrices, given respectively by:

$$R = \begin{pmatrix} \beta & -\alpha \\ -\alpha & \gamma \end{pmatrix} \qquad R(z) = \begin{pmatrix} \beta(z) & -\alpha(z) \\ -\alpha(z) & \gamma(z) \end{pmatrix}, \tag{49}$$

from which (47) takes the form:

$$\zeta(z) = \frac{1}{2}\operatorname{tr}(R^{-1}R(z)). \tag{50}$$

The slice mismatch parameter has the properties that $\zeta(z) \geq 1$, and $\zeta(z) = 1$ if and only if $R(z) = R$. (A proof is included in Appendix C.) As a result, $\zeta(z)$ is a measure of how far the Twiss functions of slice z vary from the projected Twiss functions of the beam.

Multiplying (47) by the slice emittance at z reveals that:

$$\epsilon_x(z)\zeta(z) = \frac{1}{2}\left\{\Sigma_{22}(z)\beta + 2\Sigma_{12}(z)\alpha + \Sigma_{11}(z)\gamma\right\}. \tag{51}$$

Applying (46) to (13), we see that in the present case the projected and slice covariance matrices are related by:

$$\Sigma_{ab} = \mathrm{E}[\Sigma_{ab}(z)], \qquad a, b = 1, \ldots, 6. \tag{52}$$

Taking the average of (51) across slices using (52) gives immediately the projected emittance:

$$\mathrm{E}[\epsilon_x(z)\zeta(z)] = \frac{1}{2}\left\{\Sigma_{22}\beta + 2\Sigma_{12}\alpha + \Sigma_{11}\gamma\right\} = \epsilon_x\left\{\beta\gamma - \alpha^2\right\} = \epsilon_x. \tag{53}$$

It therefore follows that:

$$\epsilon_R^2 = \mathrm{E}[\epsilon_x(z)\zeta(z)]^2 - \mathrm{E}[\epsilon_x(z)]^2. \tag{54}$$

It is clear from (54) that when $\zeta(z) = 1$ for every slice (the perfectly matched case), then $\epsilon_R = 0$.

5.4 Linear slice centroid misalignment

For the final case, assume a purely linear misalignment of the beam centroids. More precisely, assume that:

$$\mu_x(z) = \mu_x(z_c) + (z - z_c)\mu_x', \qquad \mu_{p_x}(z) = \mu_{p_x}(z_c) + (z - z_c)\mu_{p_x}', \tag{55}$$

for slices with $\lambda(z) \neq 0$. Here $z_c = \mathrm{E}[z]$ denotes the longitudinal location of the beam centroid, and μ_x', μ_{p_x}' are constants. This is a case that occurs regularly in linear optics systems with nonzero transverse-longitudinal coupling.

The quantities $\epsilon_\perp$ and ϵ_R are independent of the slice centroid values, and their form will therefore remain unchanged. To evaluate the remaining two emittance contributions, note from (55) that:

$$\mathrm{Var}[\mu_x(z)] = \mathrm{Var}[z](\mu_x')^2, \qquad \mathrm{Var}[\mu_{p_x}(z)] = \mathrm{Var}[z](\mu_{p_x}')^2, \quad \mathrm{Cov}[\mu_x(z), \mu_{p_x}(z)] = \mathrm{Var}[z]\mu_x'\mu_{p_x}'. \tag{56}$$

An immediate consequence is that the nonlinear misalignment emittance vanishes:

$$\epsilon_{||}^2 = \mathrm{Var}[\mu_x(z)]\,\mathrm{Var}[\mu_{p_x}(z)] - \mathrm{Cov}[\mu_x(z), \mu_{p_x}(z)]^2 = 0. \tag{57}$$

Evaluating the linear misalignment emittance gives:

$$\epsilon_{\mathrm{int}}^2 = \sigma_z^2 \left\{ \mathrm{E}[\sigma_x^2(z)](\mu_{p_x}')^2 + \mathrm{E}[\sigma_{p_x}^2(z)](\mu_x')^2 - 2\,\mathrm{E}[\langle x p_x \rangle_z]\mu_x'\mu_{p_x}' \right\}. \tag{58}$$

To interpret (58), it is helpful to re-express the quantities involved using the projected beam moments. To determine μ_x' and μ_{p_x}', for example, take the covariance of (55) with the slice coordinate z to see that:

$$\mathrm{Cov}[z, \mu_x(z)] = \mathrm{Cov}[z, \mu_x(z_c)] + \mathrm{Cov}[z, z - z_c]\mu_x' = \mathrm{Var}[z]\mu_x', \tag{59}$$

with a corresponding expression involving μ_{p_x}'. However, it also follows from (13) that:

$$\Sigma_{55} = \mathrm{E}[\Sigma_{55}(z)] + \mathrm{Cov}[\mu_5(z), \mu_5(z)] = \mathrm{Var}[z], \tag{60}$$

$$\Sigma_{51} = \mathrm{E}[\Sigma_{51}(z)] + \mathrm{Cov}[\mu_5(z), \mu_1(z)] = \mathrm{Cov}[z, \mu_x(z)], \tag{61}$$

since for each slice z

$$\Sigma_{55}(z) = 0, \qquad \Sigma_{51}(z) = 0, \qquad \mu_5(z) = z. \tag{62}$$

Thus we have:

$$\mu_x' = \frac{\Sigma_{51}}{\Sigma_{55}}, \qquad \mu_{p_x}' = \frac{\Sigma_{52}}{\Sigma_{55}}. \tag{63}$$

These quantities jointly characterize the size of transverse-longitudinal correlations in the beam. Likewise, applying (13) together with (56) gives:

$$\Sigma_{11} = \mathrm{E}[\Sigma_{11}(z)] + \mathrm{Cov}[\mu_1(z), \mu_1(z)] = \mathrm{E}[\sigma_x^2(z)] + \mathrm{Var}[z](\mu_x')^2, \tag{64a}$$

$$\Sigma_{22} = \mathrm{E}[\Sigma_{22}(z)] + \mathrm{Cov}[\mu_2(z), \mu_2(z)] = \mathrm{E}[\sigma_{p_x}^2(z)] + \mathrm{Var}[z](\mu_{p_x}')^2, \tag{64b}$$

$$\Sigma_{12} = \mathrm{E}[\Sigma_{12}(z)] + \mathrm{Cov}[\mu_1(z), \mu_2(z)] = \mathrm{E}[\langle \Delta x \Delta p_x \rangle_z] + \mathrm{Var}[z]\mu_x'\mu_{p_x}'. \tag{64c}$$

Using (64) and (63) to re-express (58) in terms of the projected beam covariance matrix gives:

$$\epsilon_{\mathrm{int}}^2 = \frac{1}{\Sigma_{55}} \left(\Sigma_{11}\Sigma_{52}^2 + \Sigma_{22}\Sigma_{51}^2 - 2\Sigma_{12}\Sigma_{51}\Sigma_{52} \right). \tag{65}$$

Note that this quantity is defined only when $\Sigma_{55} = \sigma_z^2 \neq 0$. This is reasonable, since when $\sigma_z = 0$, the decomposition of the beam into longitudinal slices is no longer meaningful.

6 Generalizations and Related Results

The decomposition given in (32) can be generalized in a number of ways. First, note that the treatment leading up to (32) is independent of the phase space variables used, and the same result may be applied if the conjugate canonical variables (z, p_z) are replaced by the conjugate variables (t, p_t) or $(-ct, \delta)$, as are typically used to describe the beam distribution at a fixed longitudinal location, where for example:

$$p_t = -\gamma mc^2, \qquad \delta = \Delta\gamma/\gamma_0. \tag{66}$$

In this case, the longitudinal beam slices are defined by the temporal coordinate t (or $-ct$) rather than by the spatial coordinate z. Alternatively, one may decompose the beam into slices based on the horizontal coordinate x or the vertical coordinate y. The following two cases are worth describing in some detail.

6.1 Emittance decomposition using 3-D spatial coordinates

The discussion of Section 2 can be modified by replacing the decomposition of the beam into longitudinal slices (using the longitudinal coordinate z) with a decomposition of the beam into spatial points (using the three spatial coordinates x, y, and z). To see how this is done, we write $\mathbf{X} = (\mathbf{r}, \mathbf{p})$ to emphasize the separation of the spatial $\mathbf{r} = (x, y, z)$ and momentum $\mathbf{p} = (p_x, p_y, p_z)$ phase space coordinates.

As in Section 2, the beam distribution function f defines a probability density on the 6-D phase space, and it is natural to define both a marginal and a conditional probability density associated with the decomposition into spatial points. In place of the marginal probability density given in (3), we now have the 3-D spatial density:

$$\rho(\mathbf{r}) = \int f(\mathbf{r}, \mathbf{p}) d\mathbf{p}. \tag{67}$$

Likewise, in place of the conditional probability density given in (4), we now have the distribution function at a point $\mathbf{r}$:

$$f(\mathbf{p}|\mathbf{r}) = \frac{f(\mathbf{r}, \mathbf{p})}{\rho(\mathbf{r})}, \tag{68}$$

provided $\rho(\mathbf{r}) > 0$, with $f(\mathbf{p}|\mathbf{r}) = 0$ otherwise. The averages $\langle \cdot \rangle_z$ and $\mathrm{E}[\cdot]$ are replaced by their analogues defined using densities (67) and (68), so we have the average of a quantity ϕ at a point $\mathbf{r}$ given by:

$$\langle \phi \rangle_{\mathbf{r}} = \int \phi(\mathbf{r}, \mathbf{p}) f(\mathbf{p}|\mathbf{r}) d\mathbf{p}, \tag{69}$$

and the spatial average of any $\mathbf{r}-$dependent quantity ϕ given by:

$$\mathrm{E}[\phi] = \int \phi(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r}. \tag{70}$$

In the derivation leading up to the emittance decomposition (32), the conditional expectation and conditional covariance of Appendix A are now evaluated for a fixed triple of spatial coordinates $\mathbf{r}$, rather than for a fixed slice coordinate z. Using the same theorems, the results of Section 3 carry

over directly when one takes $z \to \mathbf{r}$ and replaces the averages $\langle \cdot \rangle_z$ and $\mathrm{E}[\cdot]$ with (69) and (70). Note that in this case we have:

$$\mu_x(\mathbf{r}) = x, \tag{71}$$

and the covariance matrix at point $\mathbf{r}$ takes the form:

$$\Sigma^X(\mathbf{r}) = \begin{pmatrix} 0 & 0 \\ 0 & \sigma^2_{p_x}(\mathbf{r}) \end{pmatrix}. \tag{72}$$

Of the four contributions to the projected emittance given in (32), it follows from (72) that two of these always vanish:

$$\epsilon^2_\perp = \epsilon^2_R = 0, \tag{73}$$

and the remaining two take the forms:

$$\epsilon^2_{\mathrm{int}} = \mathrm{E}[\sigma^2_{p_x}(\mathbf{r})]\,\mathrm{Var}[\mu_x(\mathbf{r})], \tag{74a}$$

$$\epsilon^2_{||} = \mathrm{Var}[\mu_x(\mathbf{r})]\,\mathrm{Var}[\mu_{p_x}(\mathbf{r})] - \mathrm{Cov}[\mu_x(\mathbf{r}), \mu_{p_x}(\mathbf{r})]^2. \tag{74b}$$

The contribution (74a) is the *intrinsic* beam emittance due to intrinsic variations in the particle momentum at each point, while (74b) is the *correlated* beam emittance due to spatially correlated variations in the particle momentum.

It is convenient to define the transverse beam temperature T_x [4] at a point $\mathbf{r}$ by using the momentum spread at $\mathbf{r}$:

$$\frac{k_B T_x(\mathbf{r})}{mc^2} = \sigma^2_{p_x}(\mathbf{r}), \tag{75}$$

where k_B is Boltzmann's constant. In this way, we see that (74a) is proportional to the spatial average of the transverse beam temperature times the square of the transverse rms beam size. Meanwhile, (74b) is the projected emittance that the beam would have *if it were cold*—that is, if all particles at a given point were given a momentum equal to the mean momentum $\mu_{p_x}(\mathbf{r})$ at that point.

6.2 Emittance decomposition using finite subpopulations

In the previous sections, we described the beam by a continuous density f on the phase space: This is necessary for a rigorous discussion of slice moments and emittances, since these quantities are in general undefined when one considers a finite (discrete) particle distribution together with longitudinal slices of zero width. However, the decomposition (32) is based on mathematical tools that are quite general, which produce corresponding results for the discrete-particle case. These results characterize the relationship between the rms emittances of the beam and the moments of its various subpopulations.

For example, suppose that a finite beam of N particles is divided into M disjoint bins $n = 1, \ldots, M$. These may be longitudinal bins of fixed or varying width, or bins based on some other criterion (such as particle energy); we assume only that the bins divide the phase space into non-overlapping regions. We set $n = 0$ for all points $\mathbf{X}$ in the phase space that lie outside the M bins, so that each phase space point corresponds to a unique index n. Then the value of this index defines a discrete random variable on the phase space, which plays exactly the role previously played by the

continuous random variable z. The conditional expectation and conditional covariance of Appendix A are now evaluated for fixed bin index n, rather than for fixed slice z. Using the same theorems, the results of Section 3 carry over directly when one takes $z \to n$ and replaces the continuous averages $\langle\cdot\rangle$, $\langle\cdot\rangle_z$, and $\mathrm{E}[\cdot]$ with their discrete analogues.

In particular, let $\mathbf{X}_j$, $j = 1, \ldots, N$ denote the particle phase space coordinates, and let N_n denote the number of particles in bin n, with

$$\sum_{n=1}^{M} N_n = N. \tag{76}$$

The averages in (2) and (6) now become:

$$\langle\phi\rangle = \frac{1}{N}\sum_{j=1}^{N}\phi(\mathbf{X}_j), \qquad \langle\phi\rangle_n = \frac{1}{N_n}\sum_{j=1}^{N_n}\phi(\mathbf{X}_{j_n}), \qquad n = 1, \ldots, M \tag{77}$$

where the second sum is taken only over those particles that lie within bin n. Thus, $\langle\phi\rangle_n$ denotes the bin average of the quantity ϕ. We may therefore define the bin centroid vectors $\vec{\mu}_n$, the bin covariance matrices Σ_n, and the bin emittances, in the natural way. Finally, if ϕ_n is a quantity that varies from bin to bin, then we may define its average across bins as:

$$\mathrm{E}[\phi_n] = \frac{N_n}{N}\sum_{j=1}^{M}\phi_n, \tag{78}$$

from which we may also define the covariance of two $n-$dependent quantities as in (10). The horizontal projected emittance of the beam now takes the form:

$$\epsilon_x^2 = \epsilon_\perp^2 + \epsilon_R^2 + \epsilon_{\mathrm{int}}^2 + \epsilon_{||}^2, \tag{79}$$

where the four emittance contributions are given by (32) after taking $z \to n$.

7 Conclusions

The decomposition of a beam into longitudinal slices leads naturally to a corresponding decomposition for each of the projected emittances ϵ_x, ϵ_y, and ϵ_z into distinct nonnegative contributions, separating the effect of particle coordinate/momentum variation *within* each slice from the effect of particle coordinate/momentum variation *between* slices. It is expected that each of these contributions is affected by distinct dynamical processes in a typical accelerator lattice, including nonlinear aberrations due to external focusing fields, transverse-longitudinal coupling due to dispersion, slice mismatch due to time-dependent focusing effects, wakefield-induced emittance growth, and a variety of other effects. As a result, independent consideration of these four contributions may provide helpful diagnostic tools for studying the dynamical sources of projected emittance growth in typical accelerator systems.

8 Acknowledgements

This work was supported by the Director, Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

A Appendix: Mathematical Theorems

In this section, we describe several theorems that are useful in studying the relationships between slice and projected quantities.

A.1 Theorems relating to probability

The following is a fundamental theorem regarding conditional probability [14]-[15]. We give the general result, and then prove the statement that is used in the text (12-13).

Suppose X, Y and Z are three random variables defined on the same probability space. We use standard notation for the expectation $\mathrm{E}[\cdot]$ and covariance $\mathrm{Cov}[\cdot,\cdot]$ of these random variables. The expectation of X and the covariance of X and Y can be evaluated *under the condition that Z takes on some fixed value $Z = z$*. The resulting quantities are written as:

$$\mathrm{E}[X|Z = z], \qquad \mathrm{Cov}[X, Y|Z = z], \tag{80}$$

denoting the conditional expectation and the conditional covariance, respectively. As the value of Z varies, the functions $\mathrm{E}[X|Z]$ and $\mathrm{Cov}[X, Y|Z]$ are themselves random variables that depend on the value of Z. The relationships between these random variables are given by the *law of total expectation*:

$$\mathrm{E}[X] = \mathrm{E}[\mathrm{E}[X|Z]] \tag{81}$$

and the *law of total covariance*:

$$\mathrm{Cov}[X, Y] = \mathrm{E}[\mathrm{Cov}[X, Y|Z]] + \mathrm{Cov}[\mathrm{E}[X|Z], \mathrm{E}[Y|Z]], \tag{82}$$

which hold provided $\mathrm{E}[X]$ and $\mathrm{Cov}[X, Y]$ are each finite, respectively. These results apply independently of whether the random variables X, Y, and Z are continuous, discrete, or a mixture of the two.

Consider now the phase space coordinates $\mathbf{X}$, with joint probability density given by the beam distribution function $f(\mathbf{X})$. Let Z be the random variable corresponding to the coordinate z. Then (81) states that, for any function ϕ on the phase space:

$$\langle\phi\rangle = \mathrm{E}[\langle\phi\rangle_z]. \tag{83}$$

In the special case that $\phi(\mathbf{X}) = X_a$, this corresponds to the result quoted in the main text (12). We now prove (83).

Pf / Taking the mean of $\langle\phi\rangle_z$ across slices gives:

$$\begin{aligned}\mathrm{E}[\langle\phi\rangle_z] &= \int \lambda(z)\langle\phi\rangle_z dz = \int \lambda(z)\left[\int \phi(z,\zeta)f(\zeta|z)d\zeta\right]dz \\ &= \int\left[\int \phi(z,\zeta)f(z,\zeta)d\zeta\right]dz = \int \phi(\mathbf{X})f(\mathbf{X})d\mathbf{X} = \langle\phi\rangle,\end{aligned} \tag{84}$$

where in the third equality we have made use of (4). □

Next, for some indices a and b, let X be the random variable corresponding to the coordinate X_a, let Y be the random variable corresponding to the coordinate X_b, and let Z be the random variable corresponding to the coordinate z. Then we have, using the notation of the main text:

$$\mathrm{Cov}[X,Y|Z] = \Sigma_{ab}(z), \quad \mathrm{E}[X|Z] = \mu_a(z), \quad \mathrm{E}[Y|Z] = \mu_b(z), \tag{85}$$

and (82) corresponds to the result given in the main text as (13), a result that we now prove.

Pf / We have from the definition of the beam covariance matrix Σ and (83) that:

$$\Sigma_{ab} = \langle X_a X_b\rangle - \langle X_a\rangle\langle X_b\rangle = \mathrm{E}[\langle X_a X_b\rangle_z] - \mathrm{E}[\langle X_a\rangle_z]\,\mathrm{E}[\langle X_b\rangle_z]. \tag{86}$$

Applying the definition of the slice centroid coordinates $\mu_a(z)$ and $\mu_b(z)$ gives:

$$= \mathrm{E}[\langle(X_a-\mu_a(z))(X_b-\mu_b(z))\rangle_z + \mu_a(z)\mu_b(z)] - \mathrm{E}[\mu_a(z)]\,\mathrm{E}[\mu_b(z)]. \tag{87}$$

Applying the linearity of the expectation value and the definition of the slice covariance matrix:

$$= \mathrm{E}[\Sigma_{ab}(z)] + \mathrm{E}[\mu_a(z)\mu_b(z)] - \mathrm{E}[\mu_a(z)]\,\mathrm{E}[\mu_b(z)], \tag{88}$$

which gives using (10) that:

$$\Sigma_{ab} = \mathrm{E}[\Sigma_{ab}(z)] + \mathrm{Cov}[\mu_a(z),\mu_b(z)]. \qquad \square \tag{89}$$

The following theorem makes use of the notion of convexity. A *convex subset* of $n-$dimensional space $\mathbb{R}^n$ is a subset C with the property that, for all $\mathbf{x}$ and $\mathbf{y}$ in C and all $0 \le t \le 1$, the point $t\mathbf{x} + (1-t)\mathbf{y} \in C$. (That is, the line segment joining any two points in C is itself contained in C.)

If C is a convex subset of $\mathbb{R}^n$, a *convex function on* C is a function $f : C \to \mathbb{R}$ such that, for all $\mathbf{x}$ and $\mathbf{y}$ in C and all $0 \le t \le 1$,

$$f(t\mathbf{x} + (1-t)\mathbf{y}) \le tf(\mathbf{x}) + (1-t)f(\mathbf{y}). \tag{90}$$

Given a convex function $f : C \to \mathbb{R}$, a convex subset $K \subseteq C$ will be called an *f-affine subset* if and only if, for all $\mathbf{x}$ and $\mathbf{y}$ in K and all $0 \le t \le 1$, equality holds in (90).

Jensen's inequality: Let f be a convex function defined on a convex subset C of $\mathbb{R}^n$, and let $\mathbf{X}$ be a random variable taking values in C. Suppose that the mean $\mathrm{E}[\mathbf{X}]$ exists, in the sense that $\mathrm{E}[|X_j|] < \infty$ for each element $j = 1,\ldots,n$. It then follows that $E[\mathbf{X}] \in C$, the real number $\mathrm{E}[f(\mathbf{X})]$ exists, and

$$f(\mathrm{E}[\mathbf{X}]) \le \mathrm{E}[f(\mathbf{X})]. \tag{91}$$

Furthermore, equality holds in (91) if and only if there exists an f-affine subset $K \subseteq C$ with the property that $\mathbf{X} \in K$ with probability 1 (a.e.). The proof of (91) can be found in [20]-[21]. The condition for equality follows from results described in [22]-[23].

A.2 Theorems relating to matrices

A real $n \times n$ symmetric matrix A is said to be *positive semidefinite* if $\mathbf{x}^T A\mathbf{x} \geq 0$ for all vectors $\mathbf{x} \in \mathbb{R}^n$, which is typically denoted by writing $A \geq 0$. The following results are standard, and the reader may refer to one of the references [16]-[19] for proofs.

Property 1: If A is a positive semidefinite matrix, then the eigenvalues of A are nonnegative and therefore $\det(A) \geq 0$. If $\det(A) \neq 0$, then A^{-1} is also positive semidefinite.

Property 2: If A is the covariance matrix associated with a vector of random variables, then A is a positive semidefinite matrix.

Property 3: The set of $n \times n$ positive semidefinite matrices is a convex subset of the space of real $n \times n$ matrices.

Pf / Let A and B be $n \times n$ positive semidefinite matrices, and let $\mathbf{x} \in \mathbb{R}^n$, so that:

$$\mathbf{x}^T A\mathbf{x} \geq 0, \qquad \mathbf{x}^T B\mathbf{x} \geq 0. \tag{92}$$

For any $0 \leq t \leq 1$, the quantity $tA + (1-t)B$ is a real $n \times n$ symmetric matrix and:

$$\mathbf{x}^T(tA + (1-t)B)\mathbf{x} = t(\mathbf{x}^T A\mathbf{x}) + (1-t)(\mathbf{x}^T B\mathbf{x}) \geq 0, \tag{93}$$

so that $tA + (1-t)B$ is positive semidefinite. $\square$

The following inequalities can be found in [19].

Trace inequality: If A and B are $n \times n$ positive semidefinite matrices, then:

$$n \det{}^{1/n}(AB) \leq \operatorname{tr}(AB) \leq \operatorname{tr}(A)\operatorname{tr}(B). \tag{94}$$

Equality always holds when $n = 1$. Suppose $n > 1$ and $\det(A) \neq 0$. Then equality holds on the left iff $B = cA^{-1}$ for some $c \geq 0$, and equality holds on the right iff $B = 0$, the zero matrix.

Minkowski determinant inequality: If A and B are $n \times n$ positive semidefinite matrices, then

$$\det{}^{1/n}(A + B) \geq \det{}^{1/n} A + \det{}^{1/n} B. \tag{95}$$

Suppose $\det(A) \neq 0$. Then equality holds iff $B = cA$ for some $c \geq 0$.

The following theorem can be viewed as a generalization of (95) to a (possibly infinite) ensemble of matrices.

Jensen's determinant inequality: Let R be a random variable that takes its values in the set of $n \times n$ positive semidefinite matrices. Suppose the expectation value $\mathrm{E}[R]$ exists, in the sense that:

$$\mathrm{E}[|R_{ab}|] < \infty, \quad a, b = 1, \ldots, n. \tag{96}$$

Then the matrix $\mathrm{E}[R]$ is also positive semidefinite, and satisfies:

$$\det{}^{1/n} \mathrm{E}[R] \geq \mathrm{E}[\det{}^{1/n} R]. \tag{97}$$

Furthermore, if the random matrix R has the property that $\det(R) \neq 0$, then equality holds in (97) if and only if there exists a real-valued random variable c and a constant $n \times n$ matrix R_0 such that $R = cR_0$ with probability 1.

Pf / Let g be the function on the set of positive semidefinite matrices given by:

$$g(A) = -\det^{1/n}(A) \tag{98}$$

for each $A \geq 0$. Given any $0 \leq t \leq 1$, it follows from (95) that for any matrices $A \geq 0$ and $B \geq 0$:

$$\begin{aligned} g(tA + (1-t)B) &= -\det^{1/n}(tA + (1-t)B) \leq -\det^{1/n}(tA) - \det^{1/n}((1-t)B) \\ &= -t\det^{1/n}(A) - (1-t)\det^{1/n}(B) = tg(A) + (1-t)g(B). \end{aligned} \tag{99}$$

Thus, the function g is convex on the set of positive semidefinite matrices. It now follows from Jensen's inequality (91) that $\mathrm{E}[R] \geq 0$ and:

$$g(\mathrm{E}[R]) \leq \mathrm{E}[g(R)], \tag{100}$$

from which (97) immediately follows. To prove the equality condition, it follows from (95) that the only $g-$affine subsets in the set of positive semidefinite matrices with $\det(R) \neq 0$ are those of the form $\{cR_0 : c \in [a, b]\}$ for some fixed matrix R_0. Thus, equality holds in (97) if and only if $R = cR_0$ with probability 1. $\square$

B Appendix: Nonnegativity

It follows from (8) and (18) that both $\Sigma^X(z)$ and Λ^X take the form of covariance matrices. By Property 2 of Appendix 6.2, each of these matrices is positive semidefinite, and we may apply all the theorems of that section.

1. *Mean slice emittance* - Since $\Sigma^X(z)$ is positive semidefinite:

$$\epsilon_x(z)^2 = \det \Sigma^X(z) \geq 0. \tag{101}$$

Thus, $\epsilon_x(z)$ is a real quantity for each z, and so is the mean $\mathrm{E}[\epsilon_x(z)]$. Then we must have:

$$\epsilon_\perp^2 = \mathrm{E}[\epsilon_x(z)]^2 \geq 0. \tag{102}$$

Furthermore, since $\epsilon_x(z) \geq 0$ for all z, equality holds in (102) if and only if $\epsilon_x(z) = 0$ a.e.

2. *Mismatch emittance* - Recall from (26) that:

$$\det(\mathrm{E}[\Sigma^X(z)]) = \epsilon_\perp^2 + \epsilon_R^2. \tag{103}$$

The 2×2 covariance matrix $\Sigma^X(z)$ is positive semidefinite for each z. It follows from Jensen's determinant inequality (97) that the matrix $\mathrm{E}[\Sigma^X(z)]$ is also positive semidefinite and satisfies:

$$\det^{1/2}(\mathrm{E}[\Sigma^X(z)]) \geq \mathrm{E}[\det^{1/2}(\Sigma^X(z))] = \mathrm{E}[\epsilon_x(x)] = \epsilon_\perp. \tag{104}$$

Thus,

$$\epsilon_R^2 = \det(\mathrm{E}[\Sigma^X(z)]) - \epsilon_\perp^2 \geq 0. \tag{105}$$

Suppose that the slice emittance is nonzero, $\epsilon_x(z) > 0$. Then equality holds in (105) if and only if $\Sigma^X(z) = c(z)R_0$ a.e. for some real-valued function $c(z)$ and a fixed matrix R_0. We can assume without loss of generality that $\det(R_0) = 1$ (otherwise we may scale c appropriately). Thus, $R_0 = R(z)$ is exactly the slice Twiss matrix and $c(z) = \epsilon_x(z)$ is the slice emittance. Hence, equality holds in (105) if and only if the slice Twiss functions $\alpha(z)$, $\beta(z)$, and $\gamma(z)$ are constant a.e.

3. *Linear misalignment emittance* - Let A and B be the 2×2 matrices defined in (23):

$$A = \mathrm{E}[\Sigma^X(z)], \qquad B = \Lambda^X. \tag{106}$$

Then A and B are each positive semidefinite. By the trace inequality (94), we have:

$$\epsilon_{\mathrm{int}}^2 = \mathrm{tr}(A)\,\mathrm{tr}(B) - \mathrm{tr}(AB) \geq 0. \tag{107}$$

Suppose that the slice emittance is nonzero, $\epsilon_\perp \neq 0$. Then $\det(A) = \epsilon_\perp^2 + \epsilon_R^2 \neq 0$, and equality in (107) holds if and only if $B = \Lambda^X = 0$, the zero matrix. This implies that $\mathrm{Var}[\mu_x(z)] = \mathrm{Var}[\mu_{p_x}(z)] = 0$, so that $\mu_x(z) = \mathrm{E}[\mu_x(z)]$ and $\mu_{p_x}(z) = \mathrm{E}[\mu_{p_x}(z)]$ a.e. That is, the slice centroids are aligned. In this case, we also have $\epsilon_{||}^2 = \det(\Lambda^X) = 0$.

4. The matrix Λ^X is positive semidefinite, and therefore:

$$\epsilon_{||}^2 = \det(\Lambda^X) \geq 0. \tag{108}$$

Suppose that strict inequality holds in (108), so that $\epsilon_{||} \neq 0$. By the results of Section 5.4, it follows that the slice centroids $\mu_x(z)$, $\mu_{p_x}(z)$ are not uniform in z, nor are they described by linear functions of the form (55). That is, one or both of $\mu_x(z)$, $\mu_{p_x}(z)$ possess a nonlinear dependence on the slice coordinate z.

C Appendix: Slice mismatch parameter

Let $\zeta(z)$ be the slice mismatch parameter defined in (50). Then $\zeta(z) \geq 1$, and $\zeta(z) = 1$ if and only if $R(z) = R$.

Pf / Since Σ^X and $\Sigma^X(z)$ are positive semidefinite matrices and we have assumed that $\epsilon_x > 0$, $\epsilon_x(z) > 0$, the matrices $R = \Sigma^X(z)/\epsilon_x$ and $R(z) = \Sigma^X(z)/\epsilon_x(z)$ are also positive semidefinite. By Property 1 of Appendix A, this is also true of R^{-1}. Furthermore, we have:

$$\det(R) = \det(R(z)) = 1. \tag{109}$$

It follows from the trace inequality (94) that:

$$\zeta(z) = \frac{1}{2}\,\mathrm{tr}(R^{-1}R(z)) \geq \det{}^{1/2}(R^{-1}R(z)) = 1. \tag{110}$$

In addition, equality holds if and only if $R(z) = cR$ for some $c \geq 0$. However, we also have:

$$1 = \det(R(z)) = \det(cR) = c^2 \det(R) = c^2, \tag{111}$$

so $c = \pm 1$. Thus, $c = 1$ and $R(z) = R$. $\square$

References

[1] G. Dattoli et al., Nucl. Instrum. Methods Phys. Res., Sect. A **671**, 51 (2012).

[2] B. E. Carlsten, Nucl. Instrum. Methods Phys. Res., Sect. A **285**, 313 (1989).

[3] L. Serafini and J. B. Rosenzweig, Phys. Rev. E **55**, 7565 (1997).

[4] M. Reiser, *Theory and Design of Charged Particle Beams*, 2nd ed., Wiley-VCH, Weinheim (2008).

[5] M. J. Rhee and K. A. Boulais, Phys. Fluids B **3**, 1781 (1991).

[6] M. Ferrario, J. Clendenin, D. Palmer, J. Rosenzweig, and L. Serafini, SLAC-PUB-8400 (2000).

[7] M. Ferrario, V. Fusco, M. Migliorati, and L. Palumbo, Int. J. Mod. Phys. A **22**, 4214 (2007).

[8] D. Filippetto, M. Bellaveglia, M. Castellano et al, Phys. Rev. ST Accel. Beams **14**, 092804 (2011), Appendix.

[9] C. Mitchell, J. Qiang, and P. Emma, Phys. Rev. ST Accel. Beams **16**, 060703 (2013), Appendix B.

[10] B. E. Carlsten et al, Phys. Rev. ST Accel. Beams **14**, 050706 (2011).

[11] P. Emma and W. Spence, Proceedings of the 1991 Particle Accelerator Conference, p. 1549 (1991).

[12] D. H. Dowell et al, Reports No. SLAC-PUB-9540, LCLS-TN-03-2 (2002).

[13] D. A. Edwards and M. J. Syphers, *An Introduction to the Physics of High Energy Accelerators*, Wiley, New York (1993), p. 237.

[14] Sheldon Ross, *A First Course in Probability*, 6th ed., Prentice Hall (2002), p. 392.

[15] R. B. Ash and C. A. Doléans-Dade, *Probability and Measure Theory*, 2nd ed., Academic Press, San Diego, CA (2000), Section 5.5.

[16] K. M. Abadir and J. R. Magnus, *Matrix Algebra*, Cambridge University Press, New York, NY (2005), Section 12.2.

[17] Xingzhi Zhan, *Matrix Theory*, American Mathematical Society, Providence, RI (2013), Section 3.3.

[18] M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, vol. 14, Allyn and Bacon, Inc., Boston, MA (1964), Section II.4, p. 115.

[19] D. S. Bernstein, *Matrix Mathematics: Theory, Facts, and Formulas*, Proposition 8.4.14 and its Corollary, p. 471-472.

[20] T. S. Ferguson, *Mathematical Statistics: A Decision Theoretic Approach*, Academic Press, Inc., New York, NY (1967), Section 2.8, p. 76.

[21] J. R. Schott, *Matrix Analysis for Statistics*, 2nd ed., John Wiley and Sons, Inc., Hoboken, NJ (2005), Section 9.8, p. 378.

[22] E. J. McShane, Bull. Amer. Math. Soc. **43**, 521-527 (1937).

[23] M. D. Perlman, Journal of Multivariate Analysis **4**, 52-65 (1974).

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.