

OPTIMAL TWISS PARAMETERS FOR EMITTANCE MEASUREMENT IN PERIODIC TRANSPORT CHANNELS

V.Balandin, W.Decking, N.Golubeva*
DESY, Hamburg, Germany

Abstract

This paper is the continuation of the paper [1], where we started the study of the impact of the errors in the beam size measurements on the precision of the reconstruction of the beam parameters. Our objective in this paper is to describe an invariant optimality criterion and then apply it to the procedure of emittance measurement in periodic beam transport channels. We use, without further explanations, all definitions and notations given in [1], and refer to the equations of that paper in the form (1.*).

OPTIMALITY CRITERION

Let us assume that, by using the least squares approach (1.21)-(1.22), we have obtained the estimate $\mathbf{m}_\zeta(r)$ of the beamlike vector $\mathbf{m}_0(r) = \varepsilon_0 \mathbf{t}_0(r)$ actually matched to the measurements system, and let our estimate be itself a beamlike vector. Then the estimates ε_ζ and $\mathbf{t}_\zeta(r)$ for the beam parameters can be calculated as follows:

$$\varepsilon_\zeta^2 = \mathbf{m}_\zeta^\top(r) S \mathbf{m}_\zeta(r), \quad \mathbf{t}_\zeta(r) = \mathbf{m}_\zeta(r) / \varepsilon_\zeta. \quad (1)$$

Using the error vector (1.26), the equation for ε_ζ^2 can be rewritten in the form

$$\varepsilon_\zeta^2 = \varepsilon_0^2 + (\mathbf{m}_0^\top S \tilde{\mathbf{m}}_\zeta + \tilde{\mathbf{m}}_\zeta^\top S \mathbf{m}_0) + \tilde{\mathbf{m}}_\zeta^\top S \tilde{\mathbf{m}}_\zeta, \quad (2)$$

and for the mismatch between \mathbf{t}_0 and \mathbf{t}_ζ one obtains

$$\varepsilon_\zeta m_p(\mathbf{t}_0, \mathbf{t}_\zeta) = \varepsilon_0 + (\mathbf{m}_0^\top S \tilde{\mathbf{m}}_\zeta + \tilde{\mathbf{m}}_\zeta^\top S \mathbf{m}_0) / (2\varepsilon_0). \quad (3)$$

Averaging of both sides of the formulas (2) and (3) with respect to the measurement statistics gives us

$$\langle \varepsilon_\zeta^2 \rangle = \varepsilon_0^2 + \langle \tilde{\mathbf{m}}_\zeta^\top S \tilde{\mathbf{m}}_\zeta \rangle = \varepsilon_0^2 + \text{tr}(V_m S), \quad (4)$$

$$\langle \varepsilon_\zeta m_p(\mathbf{t}_0, \mathbf{t}_\zeta) \rangle = \varepsilon_0. \quad (5)$$

These formulas are quite remarkable. They are exact, they involve only first (which are equal to zero in our case) and second statistical moments of the error vector $\tilde{\mathbf{m}}_\zeta$, and they do not depend on the matched Twiss vector \mathbf{t}_0 . Besides that, one sees the appearance of the invariant of the covariance matrix V_m in the right hand side of the equation (4). Due to (1.34) this invariant is negative if the matrix V_ζ is diagonal. It means that if errors in the beam size measurements are uncorrelated, then the ‘‘typical reconstructed emittance’’ ε_ζ is an underestimation of the real emittance ε_0 , that was already observed a number of times by different authors in their Monte-Carlo simulations.

To go further, behind the exact relations (4) and (5), we do need to make some approximation. Let us assume that the measurement errors are not too large and let us denote

$$\mathcal{A} = \frac{\mathbf{m}_0^\top S \tilde{\mathbf{m}}_\zeta + \tilde{\mathbf{m}}_\zeta^\top S \mathbf{m}_0}{2\varepsilon_0^2}, \quad \mathcal{B} = \frac{\tilde{\mathbf{m}}_\zeta^\top S \tilde{\mathbf{m}}_\zeta}{\varepsilon_0^2}. \quad (6)$$

With these notations we obtain from (2) and (3) that

*nina.golubeva@desy.de

$$\frac{\varepsilon_\zeta - \varepsilon_0}{\varepsilon_0} = \sqrt{(1 + \mathcal{A})^2 - (\mathcal{A}^2 - \mathcal{B})} - 1, \quad (7)$$

$$d_H(\mathbf{t}_\zeta, \mathbf{t}_0) = \text{arccosh} \left(\frac{1 + \mathcal{A}}{\sqrt{(1 + \mathcal{A})^2 - (\mathcal{A}^2 - \mathcal{B})}} \right). \quad (8)$$

Squaring both sides of these formulas and then expanding the right hand sides of the obtained equalities with respect to the variables \mathcal{A} and $\mathcal{A}^2 - \mathcal{B}$, we obtain

$$((\varepsilon_\zeta - \varepsilon_0) / \varepsilon_0)^2 = \mathcal{A}^2 + \dots, \quad (9)$$

$$d_H^2(\mathbf{t}_\zeta, \mathbf{t}_0) = \mathcal{A}^2 - \mathcal{B} + \dots, \quad (10)$$

where the omitted terms are of the orders three and higher in the components of the error vector $\tilde{\mathbf{m}}_\zeta$.

We take the functions

$$\psi_1 = \mathcal{A}^2 \quad \text{and} \quad \psi_2 = \mathcal{A}^2 - \mathcal{B} \quad (11)$$

as the basic components for the construction of the vector-valued optimality criterion. Both these functions can be written as quadratic forms with respect to the vector $\tilde{\mathbf{m}}_\zeta$:

$$\psi_1 = (1 / \varepsilon_0^2) \cdot \tilde{\mathbf{m}}_\zeta^\top (S \mathbf{t}_0 \mathbf{t}_0^\top S) \tilde{\mathbf{m}}_\zeta, \quad (12)$$

$$\psi_2 = (1 / \varepsilon_0^2) \cdot \tilde{\mathbf{m}}_\zeta^\top (S \mathbf{t}_0 \mathbf{t}_0^\top S - S) \tilde{\mathbf{m}}_\zeta. \quad (13)$$

Using this representation, one finds that each of these quadratic forms is positive semidefinite, but is not positive definite. Moreover, it is possible to show that ψ_1 and ψ_2 are incomparable (there are points where ψ_1 is equal to zero, but ψ_2 is not, and vice versa), i.e., really, properties of both these functions have to be reflected in the optimality criterion.¹ The values of the functions ψ_1 and ψ_2 do not depend on the positioning of the reconstruction point, and as the components of the optimality criterion we take the averages

$$\langle \psi_1 \rangle = \mathcal{F} / \varepsilon_0^2, \quad \langle \psi_2 \rangle = (\mathcal{F} - \text{tr}(V_m S)) / \varepsilon_0^2, \quad (14)$$

where the function \mathcal{F} is defined in (1.35).

If the errors in the beam size determination at different measurement states are uncorrelated, then one can write

$$\begin{aligned} \langle \psi_1 \rangle &= \frac{1}{2\varepsilon_0^2 \Delta_\zeta} \sum_{i,j=1}^n \frac{a_{12}^2(s_i, s_j)}{\sigma_i \sigma_j} \cdot \frac{\beta_0(s_i)}{\sigma_i} \cdot \frac{\beta_0(s_j)}{\sigma_j} \\ &\quad - \frac{1}{2\varepsilon_0^2 \Delta_\zeta} \sum_{i,j=1}^n \left(\frac{a_{12}^2(s_i, s_j)}{\sigma_i \sigma_j} \right)^2 = \\ &\quad \frac{1}{8\varepsilon_0^2 \Delta_\zeta} \sum_{i,j=1}^n \left(\sin(2\mu_0(s_i, s_j)) \cdot \frac{\beta_0(s_i)}{\sigma_i} \cdot \frac{\beta_0(s_j)}{\sigma_j} \right)^2, \quad (15) \end{aligned}$$

¹If only measurement of emittance is of interest, then, looking at the equation (9), one may conclude that the optimization of ψ_1 alone could be sufficient. In general, it is not true. The function ψ_2 , if not controlled, may become very large and can spoil the optimization result through the high order terms which are not shown in (9).

$$\langle \psi_2 \rangle = \frac{1}{2\varepsilon_0^2 \Delta_\zeta} \sum_{i,j=1}^n \frac{a_{12}^2(s_i, s_j)}{\sigma_i \sigma_j} \cdot \frac{\beta_0(s_i)}{\sigma_i} \cdot \frac{\beta_0(s_j)}{\sigma_j} = \frac{1}{2\varepsilon_0^2 \Delta_\zeta} \sum_{i,j=1}^n \left(\sin(\mu_0(s_i, s_j)) \cdot \frac{\beta_0(s_i)}{\sigma_i} \cdot \frac{\beta_0(s_j)}{\sigma_j} \right)^2, \quad (16)$$

where β_0 is the betatron function matched to the measurement system and μ_0 is the corresponding phase advance.

The other useful forms of $\langle \psi_1 \rangle$ and $\langle \psi_2 \rangle$ can be obtained if we will use the notations

$$\frac{C_m(r)}{S_m(r)} = \sum_{k=1}^n \left(\frac{\beta_0(s_k)}{\sigma_k} \right)^2 \cdot \frac{\cos(m\mu_0(r, s_k))}{\sin(m\mu_0(r, s_k))} \quad (17)$$

With these notations:

$$\langle \psi_1 \rangle = (C_0^2 - C_4^2 - S_4^2) / (16\varepsilon_0^2 \Delta_\zeta), \quad (18)$$

$$\langle \psi_2 \rangle = (C_0^2 - C_2^2 - S_2^2) / (4\varepsilon_0^2 \Delta_\zeta), \quad (19)$$

where Δ_ζ can now be expressed as follows

$$16 \Delta_\zeta = C_0 \cdot [C_0^2 - C_4^2 - S_4^2] - 2 [(C_0 - C_4) \cdot C_2^2 - 2 S_4 \cdot C_2 S_2 + (C_0 + C_4) \cdot S_2^2]. \quad (20)$$

Let us introduce vector $\tilde{\zeta} = V_\zeta^{-1/2} \zeta$, where ζ is the vector of the measurement errors. The vector $\tilde{\zeta}$ has covariance matrix equal to the identity matrix, that allows us to say that its components are “better balanced in the order of magnitude” than the components of the original vector of the measurement errors. Using the vector $\tilde{\zeta}$, the expression for the functions ψ_m ($m = 1, 2$) can be written in the form

$$\varepsilon_0^2 \psi_m = \tilde{\zeta}^\top (V_\zeta^{-1/2} [W - (m-1)U] V_\zeta^{-1/2}) \tilde{\zeta}, \quad (21)$$

where the matrices U and W are defined in the equations (1.37) and (1.38), respectively. One can calculate, that

$$\text{tr}(V_\zeta^{-1/2} [W - (m-1)U] V_\zeta^{-1/2}) = \varepsilon_0^2 \langle \psi_m \rangle, \quad (22)$$

and, therefore, one can obtain the estimates

$$0 \leq \psi_m \leq \langle \psi_m \rangle \cdot \tilde{\zeta}^\top \tilde{\zeta}, \quad m = 1, 2. \quad (23)$$

The upper estimates of this type are usually called worst-case estimates, and we see that the minimization of the statistical averages of the functions ψ_m improve also their upper worst-case estimates.²

So we have introduced the optimality criterion, which is independent from the position of the reconstruction point. Unfortunately, this criterion includes two objective functions. To reduce a number of objectives to a single one we suggest to use the additive convolution and take the average of the weighted sum

$$\psi_0 = \kappa_1 \psi_1 + \kappa_2 \psi_2, \quad \kappa_{1,2} > 0, \quad (24)$$

as the single valued optimality criterion:

$$\langle \psi_0 \rangle = ((\kappa_1 + \kappa_2) \mathcal{F} - \kappa_2 \text{tr}(V_m S)) / \varepsilon_0^2. \quad (25)$$

²The more precise upper estimate of a positive semidefinite quadratic form can be obtained if one will use not the trace (sum of eigenvalues), but the largest eigenvalue of its matrix. Unfortunately, it is not easy to find these largest eigenvalues analytically for the quadratic forms (21), while the traces can easily be calculated.

Note that, for arbitrary positive κ_1 and κ_2 , the function ψ_0 is a positive definite quadratic form with respect to the error vector $\tilde{\mathbf{m}}_\zeta$. It is also a leading term of the expansion of the function

$$\Psi_0 = \kappa_1 (\varepsilon_\zeta - \varepsilon_0)^2 / (\varepsilon_\zeta \varepsilon_0) + \kappa_2 d_H^2(\mathbf{t}_\zeta, \mathbf{t}_0) \quad (26)$$

with respect to the components of the same vector $\tilde{\mathbf{m}}_\zeta$. Thus, as an output of Monte-Carlo simulations, one may use the average $\langle \Psi_0 \rangle$ and compare it with the analytical predictions given by the function $\langle \psi_0 \rangle$.

The choice of the weights κ_1 and κ_2 (as usual in the area of the multicriteria analysis) should reflect the specific of the problem under study, but, as concerning general situation, we do not see currently any clear theoretical reasons to take them unequal.

OPTIMAL TWISS PARAMETERS

The Twiss parameters matched to the measurement system enter the optimality criterion in two different ways, directly through the function \mathcal{F} and indirectly, when the measurement errors depend on the measured beam sizes (i.e when the matrix V_ζ depend on the vector \mathbf{b}_0).

If the matrix V_ζ depend on the vector \mathbf{b}_0 , then there is not much that one can say besides the trivial statement that the optimal Twiss parameters are the Twiss parameter which, for the fixed transport matrices between measurement states, minimize the optimality criterion.

So, let us assume that the measurement errors are independent from the measured beam sizes. In this situation one can prove that

$$\min_{\mathbf{t}_0 \in \mathcal{T}_s} \mathcal{F} = 1 / \min_{\mathbf{t}_0 \in \mathcal{T}_s} \mathcal{G} = \lambda_1, \quad (27)$$

where \mathcal{G} is defined in (1.36). From this it follows that, if the Twiss parameters are optimal, then

$$\langle \psi_1 \rangle = \lambda_1 / \varepsilon_0^2 \quad \text{and} \quad \langle \psi_2 \rangle = |\lambda_2 + \lambda_3| / \varepsilon_0^2. \quad (28)$$

Let us assume additionally that the matrix V_ζ is diagonal. Then the second equality in (27) takes on the form

$$\frac{1}{\lambda_1} = \min_{\mathbf{t}_0 \in \mathcal{T}_s} \sum_{i=1}^n \left(\frac{\beta_0(s_i)}{\sigma_i} \right)^2, \quad (29)$$

i.e. the optimal Twiss parameters minimize the weighted sum of squares of the betatron function.

Because both minimums in (27) are achieved in the same point, the optimal Twiss parameters (and only they) satisfy

$$\mathcal{F} = C_0^{-1}. \quad (30)$$

Using the formulas (18) and (20), the equality (30) can be transformed into the following equivalent form

$$(C_0 - C_4) \cdot C_2^2 - 2S_4 \cdot C_2 S_2 + (C_0 + C_4) \cdot S_2^2 = 0. \quad (31)$$

The left hand side of (31) is a positive definite quadratic form in the variables C_2 and S_2 , and therefore it can be equal to zero if and only if

$$C_2 = S_2 = 0, \quad (32)$$

which is the characteristic property of the optimal Twiss parameters.

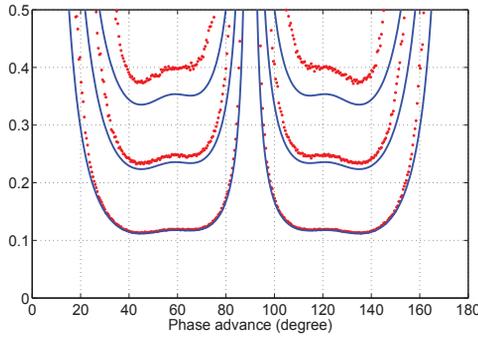


Figure 1: Solid blue and dotted red curves show $\sqrt{\langle \psi_0 \rangle}$ and $\sqrt{\langle \Psi_0 \rangle}$, respectively, as functions of the cell phase advance. The measurement errors are relative and are equal to 10%, 20% and 30% (from lower to upper pairs of curves).

EMITTANCE MEASUREMENT IN PERIODIC TRANSPORT CHANNELS

In this section we consider the question of optimal phase advance choice for the procedure of emittance measurement in periodic beam transport channels. We assume that we have a measurement system with n measurement states s_1, \dots, s_n and with the property that the matrices propagating particle coordinates from the state s_m to the state s_{m+1} are all equal to each other and allow periodic beam transport with the periodic betatron function β_p and with the corresponding phase advance μ_p being not a multiple of 90° . We assume also that this periodic betatron function is the betatron function matched to the measurement system and that the errors in the beam size determination at different measurement states are uncorrelated and have the same rms magnitude σ_p (i.e. $V_\zeta = \sigma_p^2 I$).

Under these assumptions the formulas (18)-(19) for the components of the optimality criterion can be rewritten as follows:

$$\langle \psi_1 \rangle = \frac{1}{n} \left(\frac{\sigma_p}{\varepsilon_0 \beta_p} \right)^2 \cdot \varrho_n(\mu_p), \quad (33)$$

$$\langle \psi_2 \rangle = \frac{4}{n} \left(\frac{\sigma_p}{\varepsilon_0 \beta_p} \right)^2 \cdot \varpi_n(\mu_p), \quad (34)$$

where

$$\varrho_n(\mu_p) = \frac{1 + \mathfrak{a}_n(2\mu_p)}{1 + \mathfrak{a}_n(2\mu_p) - 2\mathfrak{a}_n^2(\mu_p)}, \quad (35)$$

$$\varpi_n(\mu_p) = \frac{1 - \mathfrak{a}_n^2(\mu_p)}{1 - \mathfrak{a}_n^2(2\mu_p)} \cdot \varrho_n(\mu_p), \quad (36)$$

and \mathfrak{a}_n is given by

$$\mathfrak{a}_n(\mu_p) = \frac{\sin(n\mu_p)}{n \sin(\mu_p)}. \quad (37)$$

For an arbitrary $n \geq 3$ the function $\varrho_n(\mu_p)$ is 180° -periodic and can be extended by continuity for all μ_p inside a period, and becomes unbounded as one approaches 0° and 180° . It is never smaller than one and is equal to one (reaches its minimum) only in the points

$$\mu_p = k \cdot 180^\circ / n \pmod{180^\circ}, \quad (38)$$

where $k = 1, \dots, n-1$ with the exception of the value $n/2$

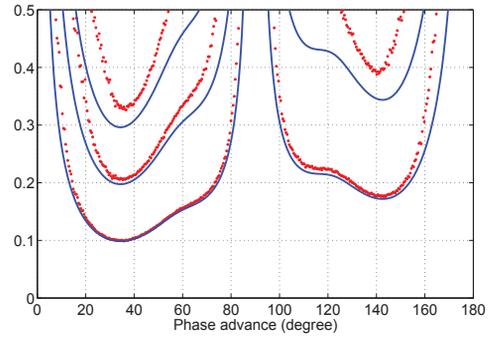


Figure 2: Solid blue and dotted red curves show $\sqrt{\langle \psi_0 \rangle}$ and $\sqrt{\langle \Psi_0 \rangle}$, respectively, as functions of the cell phase advance. The measurement errors are beam size independent and are normalized to coincide with the relative errors of 10%, 20% and 30% for $\mu_p = 45^\circ$ (from lower to upper pairs of curves).

if n is even. The same properties hold also for the function $\varpi_n(\mu_p)$ with the addition that this function becomes unbounded also near 90° .

Looking at the formulas (33) and (34) one sees that the choice of μ_p in accordance with the rule (38) brings the second multipliers in the right hand sides of these formulas to the minimal possible values. But, in general, it does not guarantee that the products of the two multipliers in (33) and (34) are also minimized. So the answer to the question, if a n -cell measurement system reaches an optimal performance when its cell phase advance is a multiple of 180° divided by n , depends on the behavior of the ratio $\sigma_p / (\varepsilon_0 \beta_p)$ during the scan of the phase advance. If this ratio stays unchanged, then it is certainly true. It includes, for example, the case when σ_p is proportional to $\varepsilon_0 \beta_p$, i.e. when errors in the beam size determination are relative to the beam size itself. But if the ratio $\sigma_p / (\varepsilon_0 \beta_p)$ changes, then the optimality of the phase advances (38) is not guaranteed.

To be more specific, let us consider a transport channel constructed from four identical thin lens FODO cells of the unit length and let us assume that four measurement stations are placed in the middle of drift spaces separating the defocusing and focusing lenses. For this particular system we made a series of Monte-Carlo simulations using as an output the square root of the average of the functions (26) with $\kappa_1 = \kappa_2 = 1$, and then compared them with the analytical predictions given by the square root of the function (25). These simulations and comparisons were made for two cases: measurement errors are relative to the beam size and measurement errors are beam size independent. The results are presented in Fig.1 and Fig.2, respectively. One sees, that while for the relative errors the optimal values of μ_p are 45° and 135° (as expected), the effect of the absolute errors gives a single optimal point $\mu_p \approx 34.5^\circ$.

REFERENCES

- [1] V.Balandin, W.Decking and N.Golubeva, Invariant Criterion for the Design of Multiple Beam Profile Emittance and Twiss Parameters Measurement Sections, Proceedings of IPAC13, Shanghai, China, 12-17 May 2013, TUPWO011.