

Summer 2011

Minimization of Theoretical Minimum Emittance in Storage Rings

Wen Wei Ho*

Dr. Chun-xi Wang†

August 16, 2011

Abstract

In [1], a new minimum-emittance theory was derived for minimizing storage ring emittance, and [2] explored numerically the associated bending profiles for optimal emittance reduction in three cases: theoretical minimum emittance (TME), achromatic minimum emittance (AME) and effective minimum emittance (EME). The bending profiles obtained were relatively simple, and were amenable to approximations by piecewise analytic functions. In this paper, we perform such approximations, and derive (approximate) analytic expressions governing the optimal emittance reduction for the three cases, and numerically evaluate the emittance reduction factor together with various useful quantities as a function of $\kappa \equiv B_{\max}/B_{\text{ref}}$. In addition, we generalize the sandwich dipole studied in [1] to include reverse bending, which may be useful to the case of a wiggler. We also derive the transfer matrix for which the bending profile is linear in nature: $\rho(s) = \rho_0(s - s_0)$. Lastly, we consider a variational approach to the optimization of the minimum-emittance, as yet another attempt at the optimization question. We find that while theoretically plausible such an approach is too complicated for any useful calculations.

1 Introduction

The emittance of a beam is defined as its size in phase space, and is governed primarily by two factors: synchrotron radiation and quantum excitations [3]. Synchrotron radiation reduces the energy of the beam and reduces the emittance, while quantum fluctuations of radiation of the particles in

*Department of Physics, Princeton University, Princeton, New Jersey 08544, USA

†Argonne National Laboratory, 9700 South Cass Avenue, Argonne, Illinois 60439, USA

the beam (in random directions) increases the beam's emittance. There therefore exists a minimal, non-zero emittance of a beam, which is determined by the magnetic profile $B(s) \propto 1/\rho(s)$ of an magnetic element in an accelerator. In this paper we shall only consider dipole magnets with no transverse gradient, i.e. that $\partial_x B(s) = 0$.

Before we discuss the mathematical theory of beam emittance, we discuss qualitatively why having as low a beam emittance as possible is desirable. Accelerators are used for many purposes today - ranging from high energy particle collisions to probe the constituents of matter, to producing rare isotopes for medical research, and to generating hard X-rays for use in various areas of research, to list a few. For example, the Advanced Photon Source (APS) at Argonne National Laboratory has a storage ring that is used for the very last purpose mentioned above. It is desirable to obtain a high intensity of radiation (this is termed brilliance), since certain experiments require radiation with deep penetrating power. The intensity of the X-rays generated is in turn related to the emittance of the beam in the storage ring - a lower emittance creates more brilliant X-rays. The question therefore falls to lowering the beam emittance as far as possible. As another example of why a low beam emittance is desired, the luminosity (rate of collisions) in a particle accelerator depends on the size of the colliding beams - it is intuitively clear that the smaller the beams are, the higher the luminosity. The question once again falls to minimizing the beam emittance for higher luminosities. As a last example on why low emittance is desired, low-emittance bending cells are important for preserving the beam quality in arcs of energy-recovery linac based light sources [1].

The mathematical theory which gives the beam emittance, considering the factors of synchrotron radiation and quantum excitations, is relatively straightforward to derive and can be found in [4]. We shall not repeat the derivation here, but merely state the final results. Assuming negligible insertion-device contributions, the horizontal emittance of a beam in a storage ring is:

$$\epsilon_x = C_q \gamma_L^2 \frac{\langle \mathcal{H}/|\rho^3| \rangle}{J_x \langle 1/\rho^2 \rangle}, \quad (1)$$

while the beam energy spread is:

$$\sigma_\delta^2 = C_q \gamma_L^2 \frac{\langle 1/|\rho^3| \rangle}{J_E \langle 1/\rho^2 \rangle}, \quad (2)$$

where γ_L is the Lorentz factor, J_x and J_E are the horizontal and longitudinal damping partition numbers respectively, and C_q is a particular constant that depends on the nature of the particles of the beam in question. For electrons, $C_q = (55/32\sqrt{3})(\hbar/mc) = 3.83 \times 10^{-13}$ m. ρ is the bending profile of the dipole¹, which as noted before is inversely proportional to the magnetic field profile B , $\langle \dots \rangle$ refer to the averaging over the length of the magnet, while the dispersion action is given by:

$$\mathcal{H} = \gamma\eta^2 + 2\alpha\eta\eta' + \beta\eta'^2, \quad (3)$$

¹Equations 1 and 2 are still valid for any device, not just dipoles in particular. As mentioned, we will be only considering dipoles in this paper, so we will henceforth refer to the devices we are considering as dipoles or magnets interchangeably.

where β, α and γ are the usual Courant-Snyder (CS) parameters. η and η' are the dispersion function and its derivative with respect to the longitudinal direction respectively, whose profile is determined by the bending profile of the magnet. The equation governing η is:

$$\eta'' + \frac{1}{\rho(s)^2}\eta = \frac{1}{\rho(s)}, \quad (4)$$

which is a linear, second order ordinary differential equation. The solution of the differential equation can be cast as:

$$\vec{\eta}(s) = M(s_0|s)\vec{\eta}_{s_0} + \vec{\xi}_{s_0}(s), \quad (5)$$

where $\vec{\eta}(s) \equiv [\eta(s), \eta'(s)]^T$ is the dispersion vector in phase space, $M(s_0|s)$ the transfer matrix from point s_0 to s , $\vec{\eta}_{s_0}$ the initial dispersion vector at the point s_0 , and $\vec{\xi}_{s_0}(s)$ the inhomogeneous solution to the differential equation, termed the dispersion generating vector. Due to the form of the solution, $\vec{\xi}_{s_0}(s)$ depends explicitly on the reference point s_0 , hence the subscript s_0 . In particular, because $M(s_0|s_0) = \mathbb{I}$, the identity matrix, $\vec{\xi}_{s_0}(s)$ is by definition 0 at s_0 .

This therefore gives all the formulas necessary to ascertain the emittance of a beam in a bending magnet. By choosing different bending profiles $\rho(s)$, subject to certain constraints (such as the total bending angle being constant), the functionals governing the emittances, defined by equations 1 and 2, will change accordingly. Hence there might possibly exist an optimal bending profile $\rho(s)$ for which the functionals are minimized. The problem of minimizing the beam emittance is thus one of optimization of $\rho(s)$, and we might employ methods such as the variational method to tackle this problem. In this paper, we will only consider minimizing the horizontal beam emittance for this paper, equation 1.

2 Minimization of emittance

A novel theory behind the minimization of the beam emittance has been given in [1]. In that paper, the author has shown that the minimization of the quantity $\langle\langle \mathcal{H} \rangle\rangle \equiv \langle \mathcal{H} / |\rho|^3 \rangle / \langle 1/\rho^2 \rangle$ is equivalent to the minimization of

$$\langle\langle \mathcal{H} \rangle\rangle = \text{Tr}(G_{s_0}\sigma_{s_0}^+), \quad (6)$$

where $\sigma_{s_0}^+$ is the matrix of CS-parameters

$$\sigma_{s_0}^+ = \begin{bmatrix} \gamma_{s_0} & \alpha_{s_0} \\ \alpha_{s_0} & \beta_{s_0} \end{bmatrix}, \quad (7)$$

and G_{s_0} the matrix

$$G_{s_0} \equiv \check{\rho}\vec{\eta}_{s_0}\vec{\eta}_{s_0}^T + \eta_{s_0}\langle\langle \hat{\xi}_{s_0} \rangle\rangle^T + \langle\langle \hat{\xi}_{s_0} \rangle\rangle\eta_{s_0}^T + \langle\langle \hat{\xi}_{s_0}\hat{\xi}_{s_0}^T \rangle\rangle. \quad (8)$$

The subscript s_0 corresponds only to the quantities at the beginning of the magnet, a point which should not be overlooked - the theory does not hold if one uses a different reference point, which will be discussed in the subsequent section. The variable $\check{\rho}$ is defined to be

$$\check{\rho} \equiv \frac{\langle 1/|\rho|^3 \rangle}{\langle 1/\rho^2 \rangle}, \quad (9)$$

and in similar fashion

$$\langle\langle \dots \rangle\rangle \equiv \frac{\langle \dots / |\rho|^3 \rangle}{\langle 1/\rho^2 \rangle}. \quad (10)$$

The variable $\hat{\xi}_{s_0}(s)$ is defined to be $\hat{\xi}_{s_0}(s) \equiv M(s_0|s)^{-1} \vec{\xi}_{s_0}(s)$. Note that while $\vec{\xi}_{s_0}(s)$ is defined to be $[\xi_{s_0}(s), \xi'_{s_0}(s)]^T$, in general $\hat{\xi}_{s_0}(s) \neq [\hat{\xi}_{s_0}(s), \hat{\xi}'_{s_0}(s)]^T$. We shall term the variable $\hat{\xi}_{s_0}(s)$ as the projected dispersion generating vector.

The minimum emittance is given by the condition $\langle\langle \mathcal{H} \rangle\rangle_{\min} = 2\sqrt{|G_{s_0}|}$ with the optimal CS parameters $\sigma_{s_0} = G_{s_0}/\sqrt{|G_{s_0}|}$ at the beginning of the dipole. The minimum emittance is achieved when the dispersion is proportional to the weighted average of the projected dispersion generating vector,

$$\vec{\eta}_{s_0} = \frac{q \langle\langle \hat{\xi}_{s_0} \rangle\rangle}{\check{\rho}}. \quad (11)$$

Then, the matrix G_{s_0} becomes:

$$\begin{aligned} G_{s_0}(q) &= \langle\langle \hat{\xi}_{s_0} \hat{\xi}_{s_0}^T \rangle\rangle + (q^2 + 2q) \langle\langle \hat{\xi}_{s_0} \rangle\rangle \langle\langle \hat{\xi}_{s_0} \rangle\rangle^T / \check{\rho} \\ &\equiv A_{s_0} + (q^2 + 2q) B_{s_0}. \end{aligned} \quad (12)$$

We thus define the matrices:

$$A_{s_0} \equiv \langle\langle \hat{\xi}_{s_0} \hat{\xi}_{s_0}^T \rangle\rangle \quad (13)$$

and

$$B_{s_0} \equiv \langle\langle \hat{\xi}_{s_0} \rangle\rangle \langle\langle \hat{\xi}_{s_0} \rangle\rangle^T / \check{\rho}. \quad (14)$$

We further define a quantity

$$c_{s_0} \equiv -\frac{\text{Tr}(JA_{s_0}JB_{s_0})}{|A_{s_0}|}, \quad (15)$$

where J is the usual block matrix considered in symplectic matrices (to define $M^+ \equiv -JM^TM$):

$$J \equiv \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \quad (16)$$

For AME, $q = 0$; for TME, $q = -1$, and for EME q is the solution to the cubic equation ($\tau \equiv J_x/J_E$):

$$(1 + \tau)q^3 + 2(2 + \tau)q^2 + [3 + (2 + \tau)/c]q + 2/c_0 = 0. \quad (17)$$

The minimum emittance thus has a nice form:

$$\langle\langle \mathcal{H} \rangle\rangle_{\min} = 2\sqrt{|A_{s_0} + (q^2 + 2q)B_{s_0}|}, \quad (18)$$

which can be further simplified to:

$$\langle\langle \mathcal{H} \rangle\rangle_{\min} = 2\sqrt{|A_{s_0}|}\sqrt{1 + (q^2 + 2q)c}, \quad (19)$$

using the fact that A_{s_0} is invertible and $|B_{s_0}| = 0$.² The minimization of beam emittance is governed by just two numbers, $|A_{s_0}|$ and c_{s_0} , which gives us the characteristics of the dipole magnet in question.

Therefore, for TME,

$$\langle\langle \mathcal{H} \rangle\rangle_{\min} = 2\sqrt{|A_{s_0}|}\sqrt{1 - c_{s_0}}, \quad (20)$$

while for AME,

$$\langle\langle \mathcal{H} \rangle\rangle_{\min} = 2\sqrt{|A_{s_0}|}. \quad (21)$$

3 Addendum to theory

The theory given in section 2 was derived for the specific case for which the projected dispersion vector was evaluated at the beginning of the magnet s_0 , i.e. $\hat{\xi}_{s_0}(s) \equiv M(s_0|s)^{-1}\vec{\xi}_0(s)$. In general the transfer matrix and the dispersion generating vector depend on the reference point (refer to equation 5), and so the averages of the projected dispersion generating functions also depend on the reference point.

Sometimes, in cases where there is symmetry, it is desirable to keep the definitions of A_{s_i} and B_{s_i} , equations 13 and 14, replacing the subscript s_0 (corresponding to the beginning of the magnet) to the subscript s_i (corresponding to any point within the magnet). For example, if the bending profile is symmetric about the center of the magnet, then the dispersion function, the transfer matrix, and the dispersion generating vector and the projected generating vector will be symmetric about the center too. That this is obvious can easily be ascertained from equation 4:

² A_0 has a non-zero determinant by invoking the Cauchy-Schwartz inequality, so $A_{s_0}^{-1}$ exists (See appendix C for the details). Also, $A_{s_0}^{-1}$ is equal to $A_{s_0}^+ / |A_{s_0}|$, a result easily verified. The form of B_{s_0} , equation 14, immediately yields that $|B_{s_0}| = 0$. Thus, the quantity $\sqrt{|A_{s_0} + \tilde{q}B_{s_0}|} = \sqrt{|A_{s_0}| |(\mathbb{I} + \tilde{q}A_{s_0}^{-1}B_{s_0})|} = \sqrt{|A_{s_0}|} \sqrt{|(\mathbb{I} + \tilde{q}A_{s_0}^+ B_{s_0} / |A_{s_0}|)|} = \sqrt{|A_{s_0}|} \sqrt{(1 + \tilde{q}\text{Tr}(JA_{s_0}JB_{s_0}/|A_{s_0}|))}$, since the eigenvalues of $A_{s_0}^+ B_{s_0} / |A_{s_0}|$ are 0 and λ , giving $|(\mathbb{I} + \tilde{q}A_{s_0}^+ B_{s_0} / |A_{s_0}|)| = (1)(1 + \lambda) = (1 + \tilde{q}\text{Tr}(JA_{s_0}JB_{s_0}/|A_{s_0}|))$.

letting the center of the magnet be at the origin, if $\eta(s)$ is a solution of the differential equation, then $\eta(-s)$ automatically is too: this is reflectional symmetry of the system. In that case, $\hat{\xi}_c(s)$ is an even function, and $\hat{\xi}'_c(s)$ an odd function, and since the weighted averages of any odd function is 0, $|A_c|$ and c_c reduce to (the subscript c stands for center):

$$\begin{aligned} |A_c| &= \langle\langle \hat{\xi}_c^2 \rangle\rangle \langle\langle \hat{\xi}'_c \rangle\rangle, \\ c_c &= \frac{\langle\langle \hat{\xi}_c \rangle\rangle^2}{\check{\rho} \langle\langle \hat{\xi}_c^2 \rangle\rangle}. \end{aligned} \quad (22)$$

Calculations are therefore greatly simplified in this form, but there needs to be a way to link the quantities $|A_{s_i}|$ and c_{s_i} with $|A_{s_0}|$ and c_{s_0} .

In what follows, the derivations are written out explicitly, for clarity.

We start with the solution of the dispersion function:

$$\begin{aligned} \vec{\eta}(s) &= M(s_0|s)\vec{\eta}_{s_0} + \vec{\xi}_{s_0}(s) \\ &= M(s_0|s)(\vec{\eta}_{s_0} + \hat{\xi}_{s_0}(s)), \text{ by definition} \\ &= M(s_1|s)M(s_0|s_1)(\vec{\eta}_{s_0} + \hat{\xi}_{s_0}(s)), \text{ using } M_{13} = M_{12}M_{23} \\ &= M(s_1|s)(\vec{\eta}_{s_1} + \hat{\xi}_{s_1}(s)), \text{ by definition.} \end{aligned} \quad (23)$$

We thus have the condition:

$$M(s_0|s_1)(\vec{\eta}_{s_0} + \hat{\xi}_{s_0}(s)) = \vec{\eta}_{s_1} + \hat{\xi}_{s_1}(s). \quad (24)$$

By definition,

$$\vec{\eta}_{s_1} = M(s_0|s_1)(\vec{\eta}_{s_0} + \hat{\xi}_{s_0}(s_1)), \quad (25)$$

and plugging this into the equation above, we obtain the relations:

$$\hat{\xi}_{s_1}(s) = M(s_0|s_1)(\hat{\xi}_{s_0}(s) - \hat{\xi}_{s_0}(s_1)), \quad (26)$$

$$\hat{\xi}_{s_0}(s) = M(s_0|s_1)^{-1}(\hat{\xi}_{s_1}(s) + \vec{\xi}_{s_0}(s_1)). \quad (27)$$

Equation 27 therefore gives us the necessary relation to link $A_0 \equiv A_{s_0}, B_0 \equiv B_{s_0}$ with $A_1 \equiv A_{s_1}, B_1 \equiv B_{s_1}$. By taking the weighted averages, and following the prescription that $\langle\langle k \rangle\rangle = k\check{\rho}$ where k is a constant, one easily proves the follow identity:

$$A_0 - B_0 = M(s_0|s_1)^{-1}(A_1 - B_1)(M(s_0|s_1)^{-1})^T. \quad (28)$$

We wish to inquire more about the nature of the transfer matrix. Explicitly, the matrix elements of $M(s_0|s)$ are:

$$\begin{bmatrix} m_1(s_0|s) & m_2(s_0|s) \\ m'_1(s_0|s) & m'_2(s_0|s) \end{bmatrix}. \quad (29)$$

Also, the functions $m_1(s_0|s)$ and $m_2(s_0|s)$ are the solutions to the homogenous part of the differential equation. That is, they satisfy independently the differential equation:

$$m_i'' + \frac{1}{\rho^2(s)} m_i = 0. \quad (30)$$

The determinant of the matrix $M(s_0|s)$ can be determined: it is a constant, and has unit value. The proof of this can be seen as follows: the differential equation that the m_i satisfy is equation 30, which is a second order ODE that does not contain a term in m'_i . Since the determinant is precisely the Wronskian of m_1 and m_2 , by Abel's Theorem the Wronskian satisfies

$$W[m_1(s), m_2(s)] = ce^{\int p(s)ds} = c, \quad (31)$$

where $p(s)$, the coefficient of m'_i in the differential equation has been noted to be 0. Thus it suffices to calculate the Wronskian at any one point, and the point s_0 yields $m_1(s_0) = 1, m_2(s_0) = 0, m'_1(s_0) = 0, m'_2(s_0) = 1$, which gives $c = 1$. Hence, $|M(s_0|s)| = W[m_1(s), m_2(s)] = 1$.

This fact immediately yields the relation:

$$|A_0 - B_0| = |A_1 - B_1|, \quad (32)$$

an invariant of the system of equations. This corresponds to the value of the TME, equation 18, since $q = -1$, and implies that one can use any reference point at which to define the dispersion generating vector with to calculat the TME. Calculations are greatly simplified for the special case of the TME, if the set-up exhibits some symmetry.

If we wish to consider the AME, then we will need to consider A_0 (Contracting $M(s_0|s_1)$ to M):

$$A_0 = M^{-1}(A_1 - B_1 + MB_0((M^{-1})^T)^{-1})(M^{-1})^T. \quad (33)$$

By the same procedure as above,

$$|A_0| = |A_1 - B_1|(1 + \text{Tr}((A_1 - B_1)^{-1}MB_0((M^{-1})^T)^{-1})). \quad (34)$$

Put in this form, one has a relatively simple relationship linking the matrices A_0, B_0, A_1 and B_1 . However, this form is not amenable to calculations because one still has to solve for the dispersion generating vector for the reference point s_0 , and in addition to that take the weighted averages over the dipole in order to obtain matrix B_0 , which defeats the purpose of using a different reference point for simplifying the calculations.

If instead we start from equation 27, and take weighted averages, we obtain:

$$A_0 = M(s_0|s_1)^{-1}(A_1 + \langle\langle \hat{\xi}_{s_1} \rangle\rangle \vec{\xi}_{s_0}^T(s_1) + \vec{\xi}_{s_0}(s_1) \langle\langle \hat{\xi}_{s_1} \rangle\rangle^T + \check{\rho} \vec{\xi}_{s_0}(s_1) \vec{\xi}_{s_0}^T(s_1))(M(s_0|s_1)^{-1})^T. \quad (35)$$

We still have to find the dispersion generating vector for the reference point s_0 , $\vec{\xi}_{s_0}(s_1)$, but now we do not have to take the weighted average over the dipole, which simplifies the algebra somewhat. If we put in $s = s_0$ into equation 27, and recall that $\hat{\xi}_{s_0}(s_0) = \vec{\xi}_{s_0}(s_0) = 0$ by definition, we obtain the relation:

$$\vec{\xi}_{s_0}(s_1) = -\hat{\xi}_{s_1}(s_0). \quad (36)$$

We can turn equation 35 into a more symmetric form by writing it as

$$A_0 = M(s_0|s_1)^{-1}(A_1 - B_1 + \vec{\zeta}\vec{\zeta}^T)(M(s_0|s_1)^{-1})^T, \quad (37)$$

where $\vec{\zeta}$ is defined as:

$$\vec{\zeta} \equiv \left(\frac{\langle\langle \hat{\xi}_{s_1} \rangle\rangle}{\sqrt{\rho}} - \sqrt{\rho} \hat{\xi}_{s_1}(s_0) \right), \quad (38)$$

which immediately yields the relation:

$$|A_0| = |A_1 - B_1| \left(1 + \text{Tr} \left((A_1 - B_1)^{-1} \vec{\zeta}\vec{\zeta}^T \right) \right). \quad (39)$$

The question remains now to find the dispersion generating vector at our chosen reference point s_1 , $\vec{\xi}_{s_1}(s)$, something we will tackle in the next section. In conclusion for this section though, one sees that the calculation for TME is invariant under choice of different reference points, something which we can use to greatly simplify our calculations.

4 Projected dispersion generating vector

We wish to find a formula to obtain the projected dispersion generating vector $\hat{\xi}_{s_i}(s)$ at some arbitrary reference point s_i , as stated in the previous section. In what follows we consider the case for which $s_i = s_0$, but the results hold similarly for any reference point s_i .

We start from the form of the dispersion equation:

$$\vec{\eta}(s) = M(s_0|s)(\vec{\eta}_{s_0} + \hat{\xi}_{s_0}(s)). \quad (40)$$

Differentiating with respect to s gives:

$$\vec{\eta}'(s) = M(s_0|s)'(\vec{\eta}_{s_0} + \hat{\xi}_{s_0}(s)) + M(s_0|s)\hat{\xi}'_{s_0}(s), \quad (41)$$

and the first term can be written explicitly, using the differential equation that $\eta(s)$ satisfies:

$$\vec{\eta}'(s) \equiv \begin{bmatrix} \eta'(s) \\ \eta''(s) \end{bmatrix} = \begin{bmatrix} \eta'(s) \\ -\frac{\eta(s)}{\rho^2(s)} + \frac{1}{\rho(s)} \end{bmatrix}. \quad (42)$$

Eliminating $\vec{\eta}_{s_0} + \hat{\xi}_{s_0}(s)$ via equation 40 yields:

$$\begin{bmatrix} \eta'(s) \\ -\frac{\eta(s)}{\rho^2(s)} + \frac{1}{\rho(s)} \end{bmatrix} = M(s_0|s)'M(s_0|s)^{-1} \begin{bmatrix} \eta(s) \\ \eta'(s) \end{bmatrix} + M(s_0|s)\hat{\xi}_{s_0}'(s), \quad (43)$$

which can be rearranged for the projected dispersion generating vector:

$$\hat{\xi}_{s_0}'(s) = M(s_0|s)^{-1} \left(\begin{bmatrix} \eta'(s) \\ -\frac{\eta(s)}{\rho^2(s)} + \frac{1}{\rho(s)} \end{bmatrix} - M(s_0|s)'M(s_0|s)^{-1} \begin{bmatrix} \eta(s) \\ \eta'(s) \end{bmatrix} \right). \quad (44)$$

Using the form of $M(s_0|s)$ as was discussed in the previous section, one is able to expand the terms in the parenthesis in equation 44 to obtain the relatively simple and surprising general result, for all profiles:

$$\hat{\xi}_{s_0}'(s) = M(s_0|s)^{-1} \begin{bmatrix} 0 \\ \frac{1}{\rho(s)} \end{bmatrix}. \quad (45)$$

This is thus the desired equation to be solved in order to get the projected dispersion generating vector. Upon integration, the dispersion generating vector is obtained by matrix multiplying the transfer matrix, as per the definition.

5 Approximation for projected dispersion generating vector

We can make quite a general statement regarding the approximation of the equation for the projected dispersion generating vector

$$\hat{\xi}_{s_0}'(s) = \int M(s_0|s)^{-1} \begin{bmatrix} 0 \\ \frac{1}{\rho(s)} \end{bmatrix} ds. \quad (46)$$

First we write the solution of the dispersion function in terms of the matrix elements m_1 and m_2 :

$$\begin{aligned} \eta(s) &= \eta_{s_0} m_1(s_0|s) + \eta'_{s_0} m_2(s_0|s) + \xi_{s_0}(s) \text{ and} \\ \eta'(s) &= \eta_{s_0} m'_1(s_0|s) + \eta'_{s_0} m'_2(s_0|s) + \xi'_{s_0}(s) \end{aligned} \quad (47)$$

By definition, $\vec{\xi}_{s_0}(s_0) = 0$, and so the nature of the solutions $m_1(s_0|s)$ and $m_2(s_0|s)$ must be such that

$$\begin{aligned} m_1(s_0|s) &= 1 + O((s - s_0)^2) \\ m_2(s_0|s) &= (s - s_0) + O((s - s_0)^2) \\ m'_1(s_0|s) &= O(s - s_0) \\ m'_2(s_0|s) &= 1 + O(s - s_0), \end{aligned} \quad (48)$$

in order that $\eta(s_0)$ is η_{s_0} and similarly $\eta'(s_0), \eta'_{s_0}$.

In light of this, the inverse of the matrix $M(s_0|s)$ (recalling that the determinant is 1) can be written as:

$$M(s_0|s)^{-1} = \begin{bmatrix} 1 + O(s - s_0) & -(s - s_0) + O((s - s_0)^2) \\ -O(s - s_0) & 1 + O((s - s_0)^2) \end{bmatrix}. \quad (49)$$

The analysis in the preceding paragraph presupposes that one is able to write a Taylor series for the two functions $m_1(s)$ and $m_2(s)$. Now, the nature of the homogenous solutions m_1 and m_2 depend upon the nature of $\rho(s)$. If $\rho(s)$ is a piecewise continuous function, then the homogenous solutions do not have well-defined second derivatives, since $\lim_{\epsilon \rightarrow 0} \eta''_{x_p + \epsilon} - \eta''_{x_p - \epsilon} = \lim_{\epsilon \rightarrow 0} -(1/\rho^2(x_p + \epsilon))\eta(x_p + \epsilon) + -(1/\rho^2(x_p - \epsilon))\eta(x_p - \epsilon) \neq 0$, at a point x_p where the second derivative is not defined. If worse, $\rho(s)$ is discontinuous, then the homogenous solutions will have not well-defined first derivatives, as can be seen by integrating the differential equation once and taking limits. However, in both cases, $\eta(s)$, and hence $m_1(s)$ and $m_2(s)$, are continuous at any 'bad' points, and we are therefore able to approximate uniformly the solutions by polynomials according to the Weierstrass approximation theorem throughout the entire (compact) interval we are considering. We are then justified in writing the solution for $\vec{\eta}(s)$ as in equation 5, with the matrix elements the polynomial approximations to the actual (not smooth) solutions to the homogenous differential equation. Our analysis above is then quite general, and can be used to obtain a simplification for the formula of the projected dispersion generating vector. Now if in addition $m_1(s)$ and $m_2(s)$ are assumed to be slowly varying functions, then expanding $m_1(s)$ and $m_2(s)$ to first order and correspondingly $m'_1(s)$ and $m'_2(s)$ to zeroth order, the inverse matrix can be approximated to be:

$$M(s_0|s)^{-1} \approx \begin{bmatrix} 1 & -(s - s_0) \\ 0 & 1 \end{bmatrix}. \quad (50)$$

The first order approximation to the projected dispersion generating vector is then given by the integrals:

$$\begin{bmatrix} \hat{\xi}_{1,s_0}(s) \\ \hat{\xi}_{2,s_0}(s) \end{bmatrix} \approx \begin{bmatrix} \int \frac{-(s-s_0)}{\rho(s)} ds \\ \int \frac{1}{\rho(s)} ds \end{bmatrix}. \quad (51)$$

6 Optimal bending profile $\rho(s)$ for minimum TME

Consider the following optimization problem: given a certain bending angle θ , what is the optimal bending profile $\rho(s)$ that will minimize the TME?

Note that this is quite a different question from the minimization theory discussed before. In the previous section, given a bending profile, one can choose CS-parameters in a certain combination at a reference point in the magnet in order to achieve the TME. Now, the problem is about choosing the optimal bending profile that will minimize that TME for a given total bending angle

$$\theta = \int_{s_0}^{s_0+L} \frac{1}{\rho(s)} ds. \quad (52)$$

In [2], an evolutionary program has been written to find the optimal bending profile for TME. In a nutshell, the program divides the length of a magnet into many thin slices, and assigns a value of ρ in each finite length, subject to the constraint. Then, it calculates the TME for this one profile, before changing slightly the values of ρ in some elements and then repeats the calculation again. Comparing the two calculations, the program then chooses the minimal one, and then repeats the process until it believes it has found a global minimum. Figure 1 shows the results, with the optimal TME profile being the red plot.

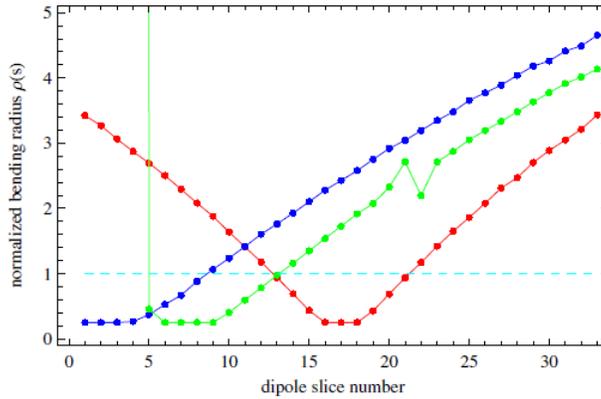


Figure 1: The optimal bending profile as obtained from the evolutionary profile. The red plot corresponds to the optimal TME profile, the blue plot the optimal AME, and finally the green plot the optimal EME profile.

From the plot, it appears that the optimal TME profile can be approximated by a piecewise continuous function consisting of a straight line sloping downwards, a flat portion, and a straight line sloping upwards, with the entire function symmetric about the center of the magnet. If we center the magnet at the origin, then the equation for the bending profile in the positive s region will be given as such:

$$\rho(s) = \begin{cases} \rho_0 & 0 \leq s < L_0 \\ \rho_0 \left(1 + g \left(\frac{s}{L_0} - 1\right)\right) & L_0 < s \leq L_1, \end{cases} \quad (53)$$

where

$$g \equiv \frac{r - 1}{L - 1}, \quad (54)$$

with $r \equiv \rho_1/\rho_0$ and $L \equiv L_1/L_0$. With this definition, $\rho(L_1) = \rho_1$. Also, the function is defined to be symmetric about $s = 0$. Thus, the bending magnet has a total length $2L_1$.

The constraint given by the bending angle is, upon integration:

$$\theta = \frac{2L_0(g + \ln r)}{g\rho_0}. \quad (55)$$

We then perform the weighted averaging as specified in the sections above, to obtain the matrices A_1 and B_1 , and thus $\langle\langle\mathcal{H}\rangle\rangle_{\min} = 2\sqrt{|A_1|}\sqrt{1 - c_1}$. Finally, we normalize our results against the TME of the ideal, uniform dipole of equal length and equal bending angle, which is given by $\theta^3/12\sqrt{15}$.

Before we proceed, we pause to perform some analysis of the form that the function corresponding to the normalized TME will take. In describing the bending profile, we require 4 parameters. In terms of ‘natural’ parameters, they are: L_0, L_1, ρ_0 and ρ_1 , which describe the length and the bending radii respectively. Now there is one constraint on the system, which is that the bending angle must be a constant, equation 55, and this therefore eliminates one of the parameters. In addition to that, it can be shown that the system exhibits scale invariance [1], so we are able to define a new variable $h(L_0, L_1, \rho_0, \rho_1)$ that is scale invariant: $h(L_0, L_1, \rho_0, \rho_1) = h(\lambda L_0, \lambda L_1, \lambda \rho_0, \lambda \rho_1)$. Thus the number of parameters that we will eventually end up with is 2. The set of ‘natural’ parameters is not necessarily the most convenient one to work with, and we will find it more useful to work with the parameters g and r given above. Then, the function for the normalized TME will be of the form $F_{\text{TME}} = F_{\text{TME}}(g, r)$.

Now, the working for the function F_{TME} is straightforward but tedious, and the final expression is messy and cumbersome. It can be found in appendix A. A contour plot of the normalized TME, $F_{\text{TME}}(g, r)$, is given in figure 2.

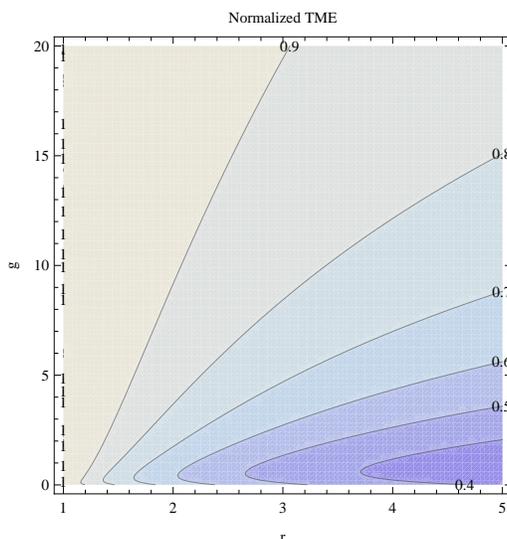


Figure 2: Contour plot of the F_{TME} for the optimal profile.

We now define a new variable, λ , defined as B_{\max}/B_{ref} where ‘ref’ refers to the reference, ideal, uniform dipole of similar bending angle. This is equivalent to the ratio ρ_0/ρ_{ref} , and can be obtained by equating the bending angles of the ideal dipole of equal length with that of our set-

up in consideration. λ has the form $\lambda(g, r)$, which can be solved implicitly for $g(\lambda, r)$. With this substitution, we can express the normalized TME as a function of λ and r only, $\tilde{F}_{\text{TME}} = \tilde{F}_{\text{TME}}(\lambda, r)$. For a fixed value of λ , there exists a minimum for the normalized TME. That point represents the bending profile for which the profile is the best approximation to the optimal bending profile obtained from the computer program. We plot the graphs for the various useful quantities against λ , such as the improvement factor over the uniform dipole, $f \equiv 1/\text{Normalized TME}$, and also r_{\min} , L_{\min} and g_{\min} . These are given in figure 3.

We obtain numerical values for f for $\lambda = 2, 4, 6$, which are approximately 3.14, 5.63, 7.38 respectively, in agreement with the values obtained from the evolutionary program used in [2]. We also note that f reduces to 1, as it should, in the case when $r \rightarrow 1$ or $L \rightarrow 1$, since the set-up reduces to that of the ideal dipole. In the case when $L \rightarrow \infty$, the set-up reduces to that of the symmetric linearly increasing profile, and the resulting algebraic expression agrees with that obtained in the appendix in [2].

7 Optimal bending profile $\rho(s)$ for minimum AME

In section 6, we considered the optimization problem of minimizing the TME of a bending magnet. Sometimes, we wish to impose additional constraints on the system, such as the condition that the bending magnet be achromatic - which is to say that the dispersion function and its derivative are both 0 at the beginning at the entrance:

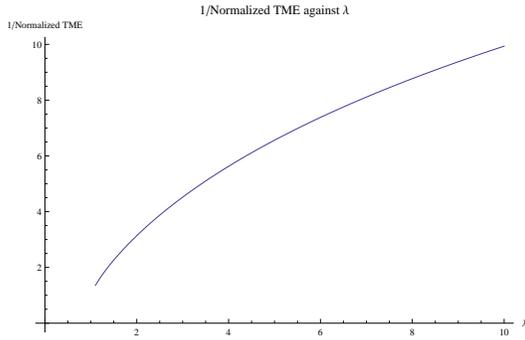
$$\vec{\eta}_{s_0} = 0. \quad (56)$$

In general, the optimal bending profile that will minimize the emittance will not be the same as that for the TME if this constraint is met. We therefore term the minimized emittance for achromatic magnets 'Achromatic minimum emittance', or AME for short.

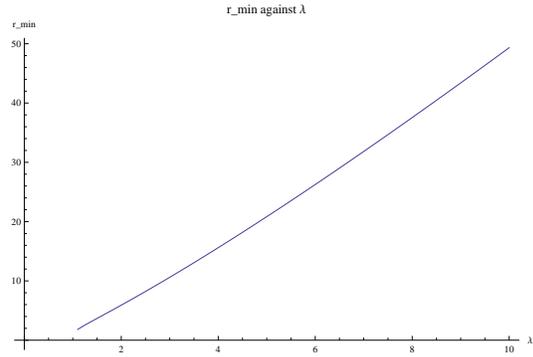
In [2], the optimal bending profile that will achieve AME was also analyzed and obtained via the evolutionary program, and the results can be seen in figure 1. Evidently, the optimal AME profile is not symmetric about the center of the magnet, and is not the same as the optimal TME profile. Instead, it consists of a flat section and then a linear ramp. From this observation, we are led to consider the following bending profile which consists of a linear ramp down, a flat section, followed by a linear ramp up:

$$\rho(s) = \begin{cases} \rho_0 \left((a-1) \frac{s}{L_0} + 1 \right) & 0 \leq s < L_0 \\ \rho_1 & L_0 < s \leq L_1 \\ \rho_1 \left(\frac{b-1}{y-1} \frac{s}{L_1} + \frac{y-b}{y-1} \right) & L_1 < s \leq L_2, \end{cases} \quad (57)$$

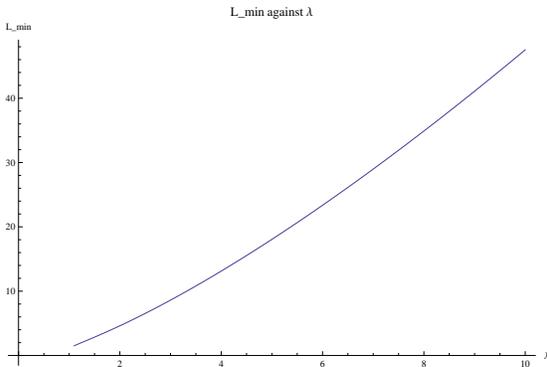
where $a \equiv \rho_1/\rho_0$, $b \equiv \rho_2/\rho_1$ and $y \equiv L_2/L_1$. We will also find it useful to define the dimensionless quantity $d \equiv L_0/L_1$. We also assume that $\rho_0 \geq \rho_1$, $\rho_2 \geq \rho_1$, and that $L_0 \leq L_1 \leq L_2$. These assumptions mean that $a \in (0, 1]$, $b \in [1, \infty)$, $y \in [1, \infty)$, and $d \in [0, 1]$.



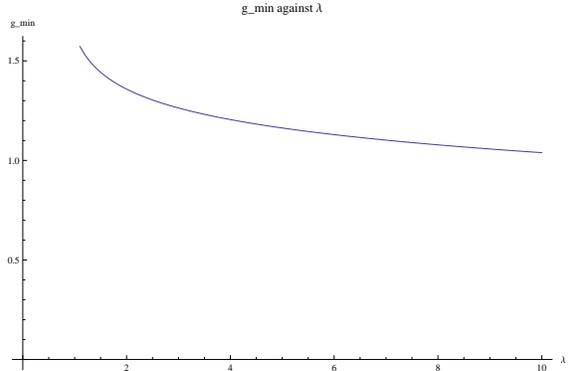
(a) Plot of the optimal improvement factor f against λ . The fourth order polynomial fit is $-1.214 + 2.826\lambda - 0.3971\lambda^2 + 0.03469\lambda^3 - 0.001212\lambda^4$.



(b) Plot of r_{\min} against λ . The fourth order polynomial fit is $-2.608 + 3.895\lambda + 0.1923\lambda^2 - 0.007416\lambda^3 - 0.0001113\lambda^4$.



(c) Plot of L_{\min} against λ . The fourth order polynomial fit is $-1.449 + 2.284\lambda + 0.4162\lambda^2 - 0.02164\lambda^3 + 0.0006167\lambda^4$.



(d) Plot of g_{\min} against λ . The fourth order polynomial fit is $1.867 - 0.3824\lambda + 0.08233\lambda^2 - 0.008394\lambda^3 + 0.0003167\lambda^4$.

Figure 3: Plots of various useful quantities of the approximation of the optimal TME against λ .

The number of parameters of this set-up is six, and the number of constraints on this set-up is one, like in the case of the TME. We are also able to define a dimensionless parameter of the resulting five parameters that reflects the scale invariance of the set-up to reduce the total number of variables needed to describe the set-up to four. It turns out that our astute introduction of the four dimensionless quantities a, b, y and d are indeed enough to describe the set-up at hand.

Once again, the calculations are straightforward but extremely inelegant and messy, and we shall not present the formulae either here or in the appendix. It is sufficient to understand that we end up with a function of the normalized AME (normalized to the AME of the ideal dipole, $\theta^3/4\sqrt{15}$),

$$F_{\text{AME}} = F_{\text{AME}}(a, b, y, d), \quad (58)$$

where F_{AME} is just $2\sqrt{|A_0|}$.

In order to compare the results with those in [2], we eliminate y in favor of $\kappa \equiv B_{\text{max}}/B_{\text{ref}} = \rho_{\text{ref}}/\rho_1$, which is the same quantity as λ in section 6. The substitution in this case is:

$$y = \frac{\kappa(a(-1+b)d \ln(a) - (-1+a)((-1+b)(-1+d) + \ln(b)))}{(-1+a)(-1+b - \kappa \ln(b))}. \quad (59)$$

However, $y \geq 1$. Thus while the range of κ is > 1 , we require simultaneously the domain of the function to be restricted to

$$\frac{\kappa(a(-1+b)d \ln(a) - (-1+a)((-1+b)(-1+d) + \ln(b)))}{(-1+a)(-1+b - \kappa \ln(b))} \geq 1. \quad (60)$$

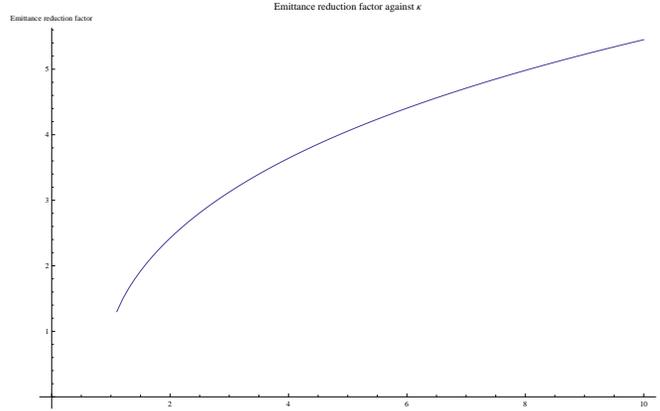
With this in mind, we plug in different values of κ into our expression for the new normalized AME, \tilde{F}_{AME} , and minimize simultaneously a, b and d , consistent with their domains and the restriction above. We use the *FindMinimum* function in Mathematica. Figure 4 shows the resulting plots for $1.1 \leq \kappa \leq 10$.

Notice that the plots of the normalized emittance, b , and y are smooth, but those of a and d are erratic. However, note that the numerical values of d are extremely small - less than 10^{-7} , and it is essentially 0. The fact that it is not exactly 0 could be attributed to small errors of the minimization method that Mathematica uses, but for our purposes we can treat it to be 0. Physically, this means that the value of the starting point L_0 is 0 - there is no linear ramp down. Consequently, the value of a does not matter, as long as it is not too close to 1 that its effect is comparable to the other sections of the magnet. Mathematically, we can express this condition as that the bending angle from the linear ramp down is negligible compared to the bending angle of the other two sections:

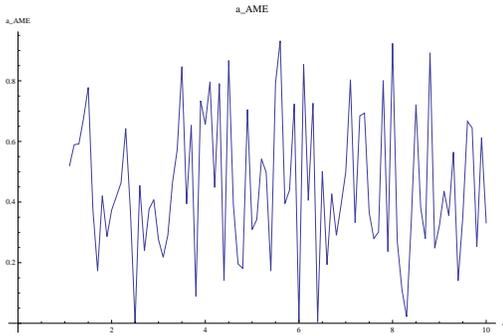
$$\frac{L_0}{\rho_0} \ll \frac{L_1 - L_0}{\rho_1}, \quad (61)$$

and

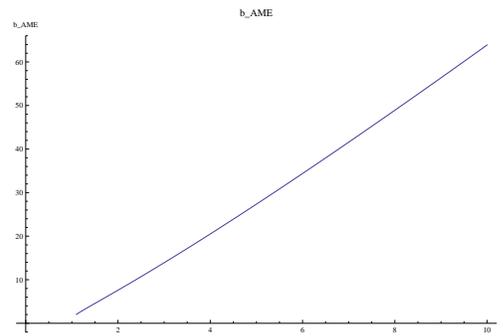
$$\frac{L_0}{\rho_0} \ll \frac{L_2 - L_1}{\rho_2}. \quad (62)$$



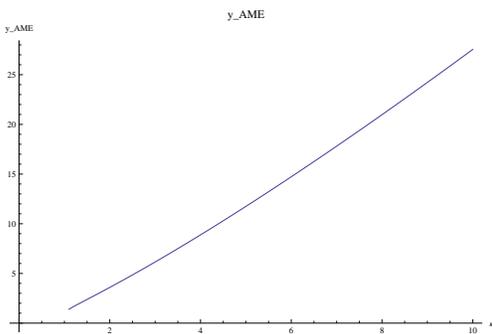
(a) Plot of the emittance reduction factor against κ . The best fit is $-0.2992 + 1.894\kappa - 0.3353\kappa^2 + 0.03198\kappa^3 - 0.001168\kappa^4$.



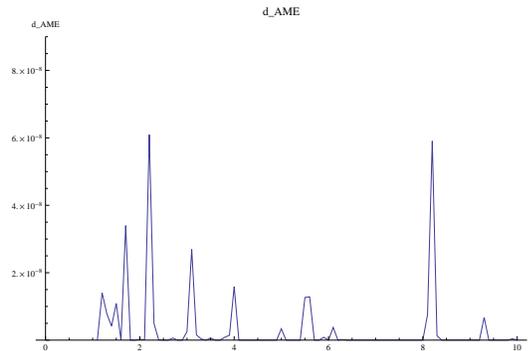
(b) Plot of a_{\min} against κ .



(c) Plot of b_{\min} against κ . The best fit is $-4.092 + 5.461\kappa + 0.2046\kappa^2 - 0.008556\kappa^3 + 0.0001479\kappa^4$.



(d) Plot of y_{\min} against κ . The best fit is $-0.9651 + 2.056\kappa + 0.1210\kappa^2 - 0.005391\kappa^3 + 0.0001228\kappa^4$.



(e) Plot of d_{\min} against κ .

Figure 4: Plots of various useful quantities of the approximation of the optimal AME profile against κ .

Rewriting it in terms of our paramters a, b, y, d , we get:

$$a \ll \frac{1}{d} - 1, \quad (63)$$

and

$$a \ll \frac{y - 1}{bd}. \quad (64)$$

Eyeballing from figure 4, we see that $a \leq 1$, $d \leq 10^{-7}$, $b \leq 70$ and $y \geq 1.4$ (corresponding to the value of $\kappa = 1.1$). The first condition is satisfied without concern, as:

$$1 \ll 10^7, \quad (65)$$

and the second condition is also satisfied, since:

$$1 \ll \frac{1.4 - 1}{70 \times 10^{-7}} = 5.7 \times 10^4. \quad (66)$$

Thus, we do not concern ourselves with the numerical value that a takes in our numerical optimization; the fact that the plot is erratic is but just an artifact of the imperfect minimization methods employed by Mathematica.

As a check of our results, we note that for $\kappa = 2, 4$, and 6 , the emittance reduction factor is $2.42, 3.64$ and 4.41 , in agreement with the values on figure 2 of [2]. For $\kappa = 4$, the optimal AME profile is given by $d \approx 0, a = 0.657$ (irrelevant), $y = 8.82$, and $b = 20.5$, which give to a good approximation the same profile as that obtained in [2], as seen in figure 1.

8 Optimal bending profile $\rho(s)$ for minimum EME

Similarly to sections 6 and 7, where we considered the optimal bending profile that minimized the TME and AME respectively, we consider approximating the optimal profile that gives the minimal effective minimum emittance (EME) by piecewise analytic functions. A glance at the green plot of figure 1, which is the optimal EME obtained by the evolutionary program, suggests that we can perhaps approximate the bending profile as such:

$$\rho(s) = \begin{cases} \infty & 0 \leq s < L_0 \\ \rho_0 & L_0 < s \leq L_1 \\ \rho_0 \left(\frac{r-1}{L-1} \frac{s}{L_1} + \frac{L-r}{L-1} \right) & L_1 < s \leq L_2, \end{cases} \quad (67)$$

where $L \equiv L_2/L_1$ and $r \equiv \rho_1/\rho_0$. We also define for future use $\lambda \equiv L_0/L_1$. We also assume that $\rho_1 > \rho_0$, which means that $L \in [1, \infty), r \in [1, \infty)$ and $\lambda \in [0, 1]$. Physically, this represents a drift space, followed immediately by a flat section followed by a linear ramp up.

We can proceed to the find the projected dispersion generating vector, as per normal, but we note here a simple way of obtaining the expressions for A and c for this set-up, from that of the

(optimal) AME profile. To be precise, assume that we have obtained an expression for A and c for the optimal AME profile:

$$\rho(s) = \begin{cases} \rho_0 & 0 < s \leq s_1 \\ \rho_0 \left(\frac{\tilde{r}-1}{S-1} \frac{s}{s_0} + \frac{S-\tilde{r}}{S-1} \right) & s_1 < s \leq s_2, \end{cases} \quad (68)$$

where \tilde{r} and $S \equiv s_1/s_0$ are defined in the usual way. Let us call the A obtained A_1 , and the c obtained c_1 .

Now, prepend a drift space of length s_0 to the AME profile, so that the beginning of the drift space is at point $-s_0$. Define $\sigma \equiv s_0/s_1$. Then, physically, this set-up is identical to the profile we are considering for the EME, if we make the identification that $L = (s + \sigma)/(1 + \sigma)$, $\tilde{r} = r$, and $\lambda = \sigma/(1 + \sigma)$. Notice also that the projected dispersion generating vector in the drift space is the 0 vector; thus, A_1 and c_1 represent the A and c obtained for the EME profile except that the reference point is at the end of the drift space. To find out A and c with the reference point at the beginning of the drift space (i.e. at the beginning of the entire magnet, and let us call the pair A_0 and c_0), we need only apply the results obtained in section 3 to shift the reference points.

Since the projected dispersion generating vector is the 0 vector, the zeta vector in equation 38 is just $\langle\langle \hat{\xi}_1 \rangle\rangle$, and therefore from equation 37,

$$A_0 = M(s_0|s_1)^{-1} A_1 (M(s_0|s_1)^{-1})^T, \quad (69)$$

where $M(s_0|s_1)$ is the transfer matrix from s_0 to s_1 for a drift space. One then obtains:

$$|A_0| = |A_1|. \quad (70)$$

As for c_0 , one can either go through the same procedure, expressing c_1 in terms of c_0 , but it is easier to argue that since the TME is invariant, independent of the reference point, and is equal to $2\sqrt{|A|}\sqrt{1-c}$, c must therefore be the same:

$$c_0 = c_1. \quad (71)$$

Thus, if we have the expressions for A_1 and c_1 for the optimal AME profile, no further integration is necessary, and all that is required is multiplication by the transfer matrix of the drift space.

In any case, one computes $A(L, r, \lambda)$ and $c(L, r, \lambda)$, and replaces λ in favor of $\kappa \equiv B_{\max}/B_{\text{ref}} = \rho_{\text{ref}}/\rho_0$:

$$\lambda = \frac{-(-1+r)(L-\kappa) + (-1+L)\kappa \ln(r)}{(-1+r)\kappa}. \quad (72)$$

The expression for the EME expression is

$$\langle\langle \mathcal{H} \rangle\rangle_{\text{EME}} = 2\sqrt{|A|} \sqrt{\frac{[1 + (q+3)qc/2](1 + [(1+\tau)q+3]qc/2)}{1+qc}}, \quad (73)$$

where the q parameter is determined by the roots of the cubic equation:

$$(1 + \tau)q^3 + 2(2 + \tau)q^2 + [3 + (2 + \tau)/c]q + 2/c = 0, \quad (74)$$

with $\tau \equiv J_x/J_E$. For the purposes of this paper, we use the value $\tau = 1/2$, which reduces the cubic equation to:

$$\frac{3}{2}q^3 + 5q^2 + \left(3 + \frac{5}{2c}\right)q + \frac{2}{c} = 0. \quad (75)$$

We can solve for c , and input it into right-most term in the expression for the EME, equation 73, to obtain:

$$\sqrt{\frac{-q^2}{48 + 80q + 24q^2}}. \quad (76)$$

We plot $-q^2/(48 + 80q + 24q^2)$ as a function of q , as shown in figure 5.

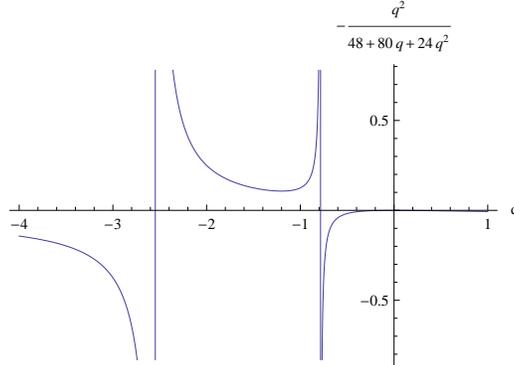


Figure 5: Graph of $-q^2/(48 + 80q + 24q^2)$ against q .

We see that the function is only positive for $(-5 - \sqrt{7})/3 \leq q \leq (-5 + \sqrt{7})/3$, and zero at $q = 0$. The point $q = 0$ obviously cannot be a solution to the cubic equation 75, thus, in choosing which root to pick from the solution of the cubic equation, we pick the one that lies within $1/3(-5 - \sqrt{7}) \leq q \leq 1/3(-5 + \sqrt{7})$.

A slight complication arises when all three roots lie within this range, which occurs for c around and above 0.995. However, we notice that for $0 \leq c \leq 0.995$ (in equation 75), only one of the roots of q is always within the range. Since $0 \leq c \leq 1$ (refer to Appendix C for the proof), we pick that particular root as the solution to the cubic equation.

This root of q is then:

$$q = -\frac{10}{9} - \frac{45c - 46c^2}{9c \left(189c^2 - 190c^3 + 9\sqrt{3}\sqrt{375c^3 - 1003c^4 + 880c^5 - 252c^6}\right)^{1/3}} + \frac{\left(189c^2 - 190c^3 + 9\sqrt{3}\sqrt{375c^3 - 1003c^4 + 880c^5 - 252c^6}\right)^{1/3}}{9c}. \quad (77)$$

We plug this into equation 73, normalize the expression with respect to the EME of the ideal dipole, and then minimize with respect to L and r , for fixed κ , subject to the constraint on the domain $0 \leq \lambda(L, r, \kappa) \leq 1$.

The results for $\kappa = 2, 4, 6, 8$ are surprising. The emittance reduction factors and the corresponding minimum values are:

$$\begin{aligned} \kappa = 2, (L, r, \lambda) &\rightarrow (3.41, 8.19, 2.19 \times 10^{-7}), \text{Factor} = 2.30 \\ \kappa = 4, (L, r, \lambda) &\rightarrow (8.21, 22.3, 4.94 \times 10^{-5}), \text{Factor} = 3.35 \\ \kappa = 6, (L, r, \lambda) &\rightarrow (13.4, 37.4, 1.83 \times 10^{-6}), \text{Factor} = 4.00 \\ \kappa = 8, (L, r, \lambda) &\rightarrow (18.9, 53.1, 1.47 \times 10^{-6}), \text{Factor} = 4.47, \end{aligned} \quad (78)$$

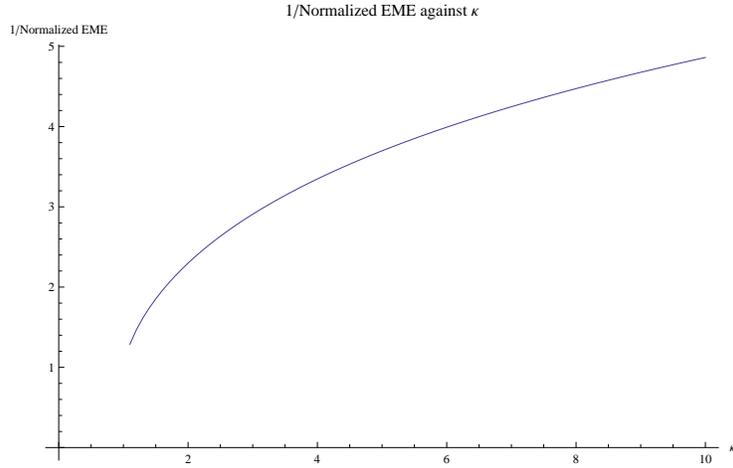
which suggests that the best profile under this assumed bending profile has no drift space, contrary to the profile obtained by the evolutionary program in figure 1! However, the emittance reduction factor is comparable; thus it is inconclusive to say which profile is the optimal one. The *FindMinimum* function in Mathematica might simply not be able to distinguish between the two cases of approximately equal numerical values, or the profile found by Mathematica might be a local (not global) minimum, since the *FindMinimum* function works by the method of steepest descent.

We therefore stress once again that we cannot conclude what the optimal bending profile for the EME is, and perhaps a better piecewise analytic approximation to the green plot in figure 1 might allow the drift space to manifest itself in the optimization.

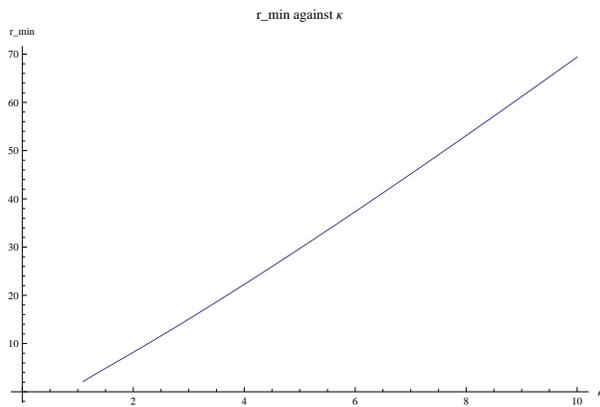
For now, we assume that $\lambda = 0$ for all κ , which greatly simplifies the expressions for the normalized EME. The graphs of the emittance reduction factor together with the optimal r and L are given below in figure 6(a), figure 6(b), and figure 6(c) respectively.

9 Sandwich dipole with reverse bending

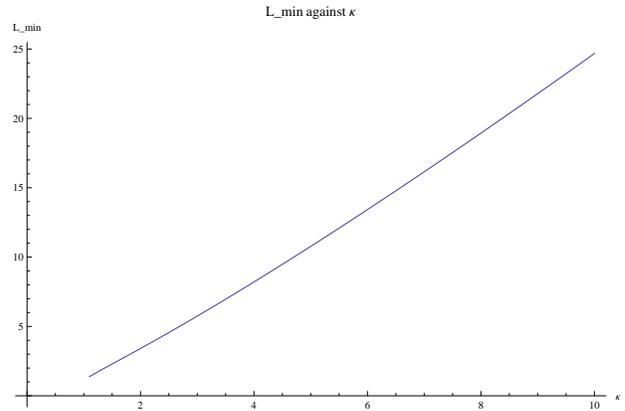
In [1], the author considered a sandwich dipole in which the middle magnet was of constant bending radius ρ_0 and length L_0 , sandwiched between two magnets of constant bending radius ρ_1 and length L_1 , with ρ_0 and ρ_1 greater than 0, and derived the resulting normalized TME for it. In this section we will generalize the set-up to include cases where the magnetic field can be negative, but still keep it symmetric about the center. In this formalism, the set-up is very similar



(a) Plot of $1/\text{Normalized EME}$ against κ . The fourth order polynomial fit to the plot is $-0.1460 + 1.718\kappa - 0.3128\kappa^2 + 0.03013\kappa^3 - 0.001106\kappa^4$.



(b) Plot of r_{\min} against κ . The fourth order polynomial fit to the plot is $-0.8536 + 1.966\kappa + 0.08857\kappa^2 - 0.003727\kappa^3 + 0.00007364\kappa^4$.



(c) Plot of L_{\min} against κ . The fourth order polynomial fit to the plot is $-0.8536 + 1.966\kappa + 0.08857\kappa^2 - 0.003727\kappa^3 + 0.00007364\kappa^4$.

Figure 6: Plots of various useful quantities of the approximation of the optimal EME profile against κ .

to that of the wiggler. Henceforth we shall use the term ‘wiggler’, which should really be read as ‘sandwich dipole with reverse bending’.

The relevant parameters defined for this set-up are as follows: $\mu \equiv \rho_1/\rho_0$ and $\nu \equiv \theta_1/\theta_0$, which are the only 2 parameters needed. If θ is the total bending angle of the wiggler, then $\theta_0 = \theta/(1+2\nu)$ and $\theta_1 = \nu\theta/(1+2\nu)$. Since ρ_0 and ρ_1 can take any real value, the parameter μ and ν can take negative values too. One implication from the definitions is that

$$\text{sgn}(\mu) = \text{sgn}(\nu), \quad (79)$$

which means after obtaining an expression for the normalized TME, $F_{\text{TME}}(\mu, \nu)$, we only consider the first and third quadrants of the \mathbb{R}^2 plane of the parameters μ and ν .

Once again, the expressions obtained are rather complicated, and are given in the appendix B. For now, we contend with two contour plots of the normalized TME of the first and third quadrants in the \mathbb{R}^2 plane. Figure 7 shows the plots.

We note that there is a line of singularities in $F_{\text{TME}}(\mu, \nu)$, corresponding to $\nu = -1/2$. This corresponds to the case when the total bending angle is 0, and can be explained as such: in a wiggler set-up where the total bending angle is 0, i.e. $\nu = -1/2$, there is a finite but non-zero TME, given by the expression in the appendix B. For the case of an ideal, uniform dipole in which the bending angle is 0, the TME is obviously 0, and hence the ratio of a finite number over zero, which is the normalized TME, is necessarily infinite.

As a check, we notice that the algebraic expression for the normalized TME reduces to that of the sandwich dipole in the first quadrant, as it should, given in [1]. Also, if we let $\mu \rightarrow 1$ or $\nu \rightarrow \infty$, the set-up reduces to that of the ideal dipole, and $F_{\text{TME}}(\mu, \nu)$ reduces to a value of 1.

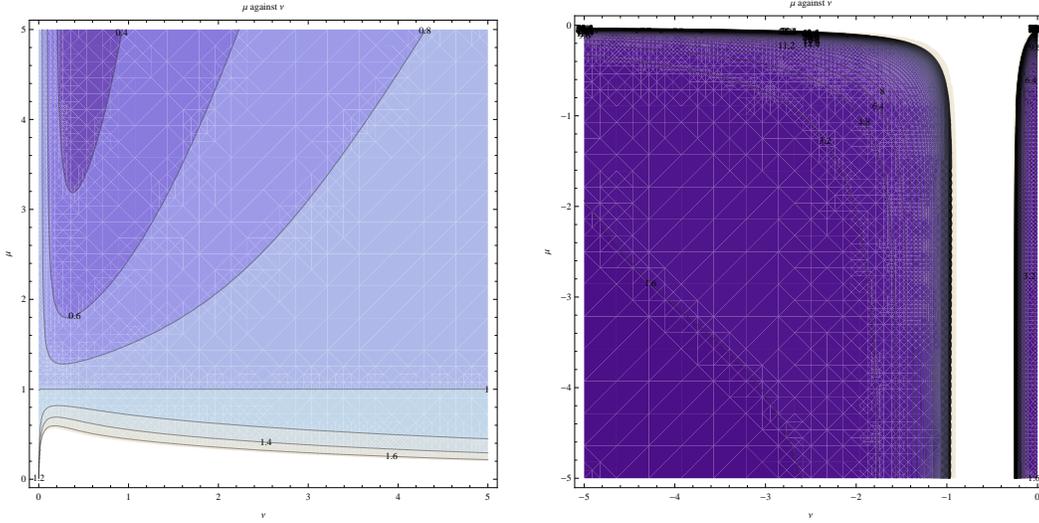
10 Transfer matrix for linearly increasing bending profile

For completeness, we return to the case of approximation to the optimal bending profile for TME, AME and EME, and attempt an exact solution instead of the approximation of the projected dispersion generating vector. This involves solving for the transfer matrix exactly and therefore results in an exact solution for the dispersion generating vector and projected dispersion generating vector (if the expression is integrable). In the worst case scenario, the two functions will be given in terms of quadratures.

The equation to be solved is the differential equation that the dispersion function satisfies:

$$\eta'' + \frac{1}{\rho(s)^2}\eta = \frac{1}{\rho(s)}, \quad (80)$$

where now $\rho(s)$ is $\rho_0(s - s_0)$, a generic equation for a straight line. We also assume that the transfer matrix is from a reference point $s_1 > s_0$ to $s > s_0$, since the only solution for which the transfer is from the point s_0 is the null solution. Note that s_0 cannot refer to the beginning of the magnet, since $\rho = 0$ implies $B = \infty$, a physical impossibility. s_0 is just one of the two variables required to specify a straight line, and should not be confused with the starting point s_1 .



(a) Contour plot of the normalized TME in the first quadrant.

(b) Contour plot of the normalized TME in the third quadrant.

Figure 7: Contour plots of the normalized TME for the wiggler in the first and third quadrant of the \mathbb{R}^2 plane.

From inspection, the particular solution is

$$\eta_p(s) = \rho_0(s - s_0). \quad (81)$$

For the general solution, we plug in an ansatz to the homogenous equation:

$$\eta(s) = (s - s_0)^\alpha, \quad (82)$$

which immediately gives the characteristic equation

$$\alpha(\alpha - 1) + \frac{1}{\rho_0^2}. \quad (83)$$

The solution to the dispersion equation is then:

$$\eta(s) = \sqrt{s - s_0} (A_1(s - s_0)^\omega + A_2(s - s_0)^{-\omega}) + \rho_0(s - s_0). \quad (84)$$

with $\omega \equiv \sqrt{1/4 - 1/\rho_0^2}$.

Now set the initial conditions $\eta_1 \equiv \eta(s_1)$ and $\eta'_1 \equiv \eta'(s_1)$. We can eliminate the constants A_1 and A_2 in favor of η_1 and η'_1 . Once this substitution is made, the function that is multiplied with η_1 is the matrix element $m_1(s)$, the function that is multiplied with η'_1 is the matrix element $m_2(s)$ and the function that does not have a factor or either is the dispersion generating function $\xi(s)$.

The matrix elements are given by:

$$m_1(s) = \frac{(s-s_0)^{\frac{1}{2}-\omega}(-s_0+s_1)^{-\frac{1}{2}-\omega} \left((s-s_0)^{2\omega}(-1+2\omega) + (-s_0+s_1)^{2\omega}(1+2\omega) \right)}{4\omega} \quad (85)$$

$$m_2(s) = \frac{(s-s_0)^{\frac{1}{2}-\omega}(-s_0+s_1)^{\frac{1}{2}-\omega} \left((s-s_0)^{2\omega} - (-s_0+s_1)^{2\omega} \right)}{2\omega} \quad (86)$$

$$m'_1(s) = \frac{(s-s_0)^{-\frac{1}{2}-\omega}(-s_0+s_1)^{-\frac{1}{2}-\omega} \left((s-s_0)^{2\omega} - (-s_0+s_1)^{2\omega} \right) (-1+4\omega^2)}{8\omega} \quad (87)$$

$$m'_2(s) = \frac{(s-s_0)^{-\frac{1}{2}-\omega}(-s_0+s_1)^{\frac{1}{2}-\omega} \left((-s_0+s_1)^{2\omega}(-1+2\omega) + (s-s_0)^{2\omega}(1+2\omega) \right)}{4\omega}, \quad (88)$$

with the transfer matrix

$$M(s_1|s) = \begin{bmatrix} m_1(s) & m_2(s) \\ m'_1(s) & m'_2(s) \end{bmatrix}. \quad (89)$$

The dispersion generating function is given by

$$\xi_{s_1}(s) = -\frac{\rho_0 \left(4(s_0-s)\omega + (s-s_0)^{\frac{1}{2}-\omega}(-s_0+s_1)^{\frac{1}{2}-\omega} \left((-s_0+s_1)^{2\omega}(-1+2\omega) + (s-s_0)^{2\omega}(1+2\omega) \right) \right)}{4\omega}. \quad (90)$$

Expectedly, the Taylor series of $m_1(s)$, $m_2(s)$ and $\xi(s)$ about s_1 are:

$$m_1(s) = 1 + \frac{(-1+4\omega^2)(s-s_1)^2}{8(s_0-s_1)^2} + O((s-s_1)^4) \quad (91)$$

$$m_2(s) = (s-s_1) + \frac{(-1+4\omega^2)(s-s_1)^3}{24(s_0-s_1)^2} + O((s-s_1)^4) \quad (92)$$

$$\xi(s) = \frac{\rho_0(-1+4\omega^2)(s-s_1)^2}{8(s_0-s_1)} + \frac{\rho_0(-1+4\omega^2)(s-s_1)^3}{24(s_0-s_1)^2} + O((s-s_1)^4), \quad (93)$$

which agree with our analysis before.

11 A variational approach to the optimization question

In this last section we present another take on the optimization of the minimum emittance question. This section is still a work in progress, and the material presented here might be inaccurate or even wholly incorrect. If so, the author apologizes for any errors and wishes that the reader views this section as a speculative approach to the problem.

We start by noticing that the optimization problem involves selecting a bending profile $\rho(s)$ subject to certain constraints that will minimize the following expression:

$$\langle\langle \mathcal{H} \rangle\rangle = \frac{\frac{1}{L} \int \frac{\gamma(s)\eta(s)^2 + 2\alpha(s)\eta(s)\eta'(s) + \beta(s)\eta'(s)^2}{|\rho(s)|^3} ds}{\frac{1}{L} \int \frac{1}{\rho(s)^2} ds}, \quad (94)$$

with $\alpha(s) \equiv -\beta'(s)/2$ and $\gamma(s) \equiv (1 + \alpha(s)^2)/\beta(s)$, where $\eta(s)$ and $\beta(s)$ obey the following differential equations:

$$\eta''(s) + \frac{1}{\rho(s)^2}\eta(s) - \frac{1}{\rho(s)} = 0, \text{ and} \quad (95)$$

$$2\beta(s)\beta''(s) - \beta'(s)^2 + \frac{4\beta(s)^2}{\rho(s)^2} - 4 = 0. \quad (96)$$

The variable $\langle\langle\mathcal{H}\rangle\rangle$ is a functional (as opposed to a function), which maps elements (functions, in particular) $(\rho(s) \in V_\rho, \eta(s) \in V_\eta, \beta(s) \in V_\beta)$ from the space $V_\rho \times V_\eta \times V_\beta$ to a complex number \mathbb{C} . That is,

$$\langle\langle\mathcal{H}\rangle\rangle : V_\rho \times V_\eta \times V_\beta \rightarrow \mathbb{C}, \quad (97)$$

subject to the constraints of the differential equations 95 and 96. Thus we hope to be able to use the method of calculus of variations to obtain the functions $\rho(s)$, $\eta(s)$ and $\beta(s)$ that will extremize the functional $\langle\langle\mathcal{H}\rangle\rangle$.

Our first thought is to use the Euler-Lagrange equations, but we run into complications immediately because the Euler-Lagrange equations are only valid for linear functionals; obviously $\langle\langle\mathcal{H}\rangle\rangle$ is not a linear functional and we cannot use such an approach, since it is a ratio of integrals.

It turns out that the theory of optimization of a product of integrals is already a well-studied one, though it is probably not ubiquitously known to the general scientific population. In particular, areas such as the study of the aerodynamics of ballistics frequently deal with the minimization of the product of powers of several integrals. Papers such as [5], [6], and [7] provide the necessary background to deal with such a scenario. Here we shall give an overview and only state the steps involved in computing the minimization of the product of integrals. Little effort has been made to justify or prove the formulas.

We begin with the functional

$$I = \prod_{k=1}^n I_k^{\alpha_k}, \quad (98)$$

where is the product of n integrals

$$I_k = \int_{x_i}^{x_f} f_k(x, y, y') dx, \quad k = 1, \dots, n. \quad (99)$$

The integrals I_k are assumed to be positive, and α_k are given, positive or negative. We also assume that the initial and final points are given:

$$x_i = k_1, \quad y_i = k_2, \quad x_f = k_3, \quad y_f = k_4. \quad (100)$$

The minimization of equation 98 is equivalent to the minimization of $\ln(I)$ since the logarithm is monotonic in nature. Therefore the minimization of equation 98 is achieved at the minimum of the (linear) functional

$$\ln(I) = \sum_{k=1}^n \alpha_k \ln(I_k). \quad (101)$$

The introduction of the n auxiliary variables z_k defined by

$$z'_k - f_k(x, y, y') = 0 \quad \forall k \quad (102)$$

together with the initial conditions

$$(z_k)_i = 0 \quad \forall k \implies (z_k)_f = I_k \quad \forall k, \quad (103)$$

allows us to rewrite equation 101 as

$$\ln(I) = \sum_{k=1}^n \alpha_k \ln((z_k)_f). \quad (104)$$

This is a problem of the Mayer type; one minimizes equation 104 together with the constraints equation 102, and the boundary conditions. We can include the constraints by use of Lagrange multipliers, thus converting the Mayer type problem to a Bolza problem [8]:

$$\tilde{I} = G((z_k)_f) + \int_{x_i}^{x_f} F(x, y, y', z'_k, \lambda_k) dx, \quad (105)$$

where

$$G((z_k)_f) \equiv \sum_{k=1}^n \alpha_k \ln((z_k)_f), \quad (106)$$

and

$$F \equiv \sum_{k=1}^n \lambda_k (f_k - z'_k). \quad (107)$$

Note that F is formally equals to 0.

We can now employ the Euler-Lagrange equation to 105, to obtain $n + 1$ equations

$$\frac{d}{dx} \frac{\partial F}{\partial z'_k} = \frac{\partial F}{\partial z_k}, \quad \forall k, \quad (108)$$

and one equation

$$\frac{d}{dx} \frac{\partial F}{\partial y'} = \frac{\partial F}{\partial y}. \quad (109)$$

From the above two equations, the Lagrange multipliers are seen to be all constant:

$$\lambda'_k = 0 \quad \forall k, \quad (110)$$

while

$$\frac{d}{dx} \frac{\partial f}{\partial y'} = \frac{\partial f}{\partial y}, \quad (111)$$

where

$$f \equiv \sum_{k=1}^n \lambda_k f_k \quad (112)$$

is called the fundamental function. We note that if $\partial f / \partial x = 0$, then by the Beltrami identity $f - y'(\partial f / \partial y') = \text{Constant}$.

The $2n+1$ differential equations (n from the defining equation of z_k , n from the Euler-Lagrange equation w.r.t. z_k , and 1 from the Euler lagrange equation w.r.t. y), have to satisfy the so-called 'transversality condition':

$$\delta G + \left[\left(F - y' \frac{\partial F}{\partial y'} - \sum_{k=1}^n z'_k \frac{\partial F}{\partial z'_k} \right) \delta x + \frac{\partial F}{\partial y'} \delta y + \sum_{k=1}^n \frac{\partial F}{\partial z'_k} \delta z_k \right]_i^f = 0 \quad (113)$$

and in addition the Weierstrass condition

$$\Delta F - \frac{\partial F}{\partial y'} \Delta y' - \sum_{k=1}^n \frac{\partial F}{\partial z'_k} \Delta z'_k \geq 0, \quad (114)$$

for any set of variations $\Delta y'$ and $\Delta z'_k$ consistent with equation 102.

Since we have $2n + 1$ differential equations, $2n$ of which are first order and one is of second order, and one is of second order, we need $2n + 2 + 2$ conditions to specify one solution. $2n$ of these conditions are required by the $2n$ first order equations, 2 are required by the one second order equation and 2 are required to specify x_i and x_f . Indeed we see that n conditions are given by equation 103, 4 by equation 100, and n by equation 113. Thus a solution, if it exists, is uniquely defined.

We can hope to adapt this method of the minimization of the functional $\langle\langle \mathcal{H} \rangle\rangle$. Now though, each integrand I_k has not only x, y, y' in its argument, but second derivatives and additional terms:

$$I_k = \int_0^L f_k(s, \rho, \eta, \eta', \eta'', \beta, \beta', \beta'') ds. \quad (115)$$

We can employ the same method, and define the fundamental function

$$\begin{aligned} f = & \lambda_1 \left(\frac{\gamma(s)\eta(s)^2 + 2\alpha(s)\eta(s)\eta'(s) + \beta(s)\eta'(s)^2}{|\rho(s)|^3} \right) \\ & + \lambda_2 \left(\frac{1}{\rho(s)^2} \right) \\ & + \lambda_3 \left(\eta''(s) + \frac{1}{\rho(s)^2} \eta(s) - \frac{1}{\rho(s)} \right) \\ & + \lambda_4 \left(\beta(s)\beta''(s) - \beta'(s)^2 + \frac{4\beta(s)^2}{\rho(s)^2} - 4 \right) \end{aligned} \quad (116)$$

where $\lambda_1, \dots, \lambda_4$ represent the lagrange multipliers that are to be determined after a solution has been obtained. The Euler-Lagrange equations for a functional $L(x, y, y', y'')$ that involves derivatives up to second order now take the form

$$\frac{\partial L}{\partial y} - \frac{d}{dx} \frac{\partial L}{\partial y'} + \frac{d^2}{dx^2} \frac{\partial L}{\partial y''} = 0, \quad (117)$$

and presumably the transversality condition and the Weierstrass condition have to be modified accordingly too. Thus one sees that theoretically it is possible to obtain a direct solution for $\rho(s), \eta(s), \beta(s)$ using the method of calculus of variations. Unfortunately the differential equations are of high order (≤ 3) and have complicated expressions, and so the possibility of an analytic expression is almost nil. We admit that the variational method probably gives little or no insight at all in the optimization problem. Thus, we reiterate that we hope the reader views this section merely as another method of looking at the same optimization problem for completeness.

12 Conclusion

In this paper, we have successfully approximated the optimal bending profiles for TME, AME and EME with piecewise analytic function, and derived analytic expressions for A and c for all three cases. Using those expressions, we obtained numerical values for the optimal emittance reduction factor, and for various parameters describing the optimal bending profiles. We also employed the method to the case of a wiggler, and obtained analytic expressions for the normalized (TME) emittance reduction factor. In addition to that, we derived an expression for the transfer matrix of a bending profile that was linear in nature: $\rho(s) = \rho_0(s - s_0)$. Lastly, we considered a variational approach to the minimization of $\langle\langle \mathcal{H} \rangle\rangle$ and concluded that while it might be theoretically possible to obtain a solution from such an approach, it is probably too complicated and intractable to have any practical value to the problem of emittance reduction.

13 Acknowledgments

The author Wen Wei Ho would like to thank his mentor Dr. Chun-xi Wang for providing him with the guidance and assistance necessary to embark on this project and make it a successful one.

Appendix

A Expressions for minimal TME

In this section we present the expressions obtained in section 6.

To simplify the algebra, we define the variables

$$x \equiv -1 + r^2 \quad (118)$$

$$y \equiv -1 + r \quad (119)$$

$$z \equiv \ln(r). \quad (120)$$

Then the normalized TME is given by:

$$F_{\text{TME}}(g, r) = \frac{1}{4\sqrt{2}} \sqrt{\frac{f_1 f_2}{f_3}}, \quad (121)$$

where

$$f_1 \equiv 4g^3 r^2 + 6g^2 x + 6g(x - 2z) + 3(x - 2z(1 + z)), \quad (122)$$

$$\begin{aligned} f_2 \equiv & 16g^6 r^4 + 48g^5 r^2 x + 120g^4 r^2 (x - 2z) + 45((1 - 6r - 23r^2)y^2 + 4r^2(-6 + 4r + 2r^2 - z)z) \\ & + 60g^3 r^2 (3 - 8r + 5r^2 - 2z - 6z^2) + 45g^2((1 + 2r + 17r^2)y^2 - 32r^2 yz + 12r^2 z^2) \\ & + 90(-1 + 4r - 4r^2(3 + z) + 4r^3(7 + 2z) + r^4(-19 + 8z)), \text{ and} \end{aligned} \quad (123)$$

$$f_3 \equiv r^2(-1 + (1 + 2g)r^2)(gr + y)^2(g + z)^6. \quad (124)$$

If we now eliminate g in favor of λ (defined in section 6) via

$$g(r, L) \equiv \frac{r - 1}{L - 1} \text{ and} \quad (125)$$

$$L(\lambda, r) = \frac{\lambda(1 - r + \ln(r))}{1 - r + \lambda \ln(r)}, \quad (126)$$

then the normalized TME becomes:

$$\tilde{F}_{\text{TME}}(\lambda, r) = \sqrt{\frac{f_4 f_5}{f_6}}, \quad (127)$$

with

$$\begin{aligned} f_4 \equiv & ((-1 + r)(4r^4 + 2r^3(-7 + 3\lambda) + 2r^2(8 - 9\lambda + 3\lambda^2) \\ & + 3(-5 + 9\lambda - 5\lambda^2 + \lambda^3) + 3r(1 + \lambda - 3\lambda^2 + \lambda^3)) \\ & - 6(-3 - 2r(-1 + \lambda) + 4\lambda + 2r^4\lambda + 2r^3(-3 + \lambda)\lambda - \lambda^2 + r^2\lambda(5 - 4\lambda + \lambda^2)) \ln(r) \\ & + 6(1 - \lambda + r^2(-3 + 2r)\lambda^2 + r^2\lambda^3) \ln(r)^2 - 4r^2\lambda^3 \ln(r)^3), \end{aligned} \quad (128)$$

$$\begin{aligned}
f_5 \equiv & ((-1+r)^2 (16r^8 - 270r(-1+\lambda)^5\lambda + 45(-1+\lambda)^4\lambda^2 + 16r^7(-7+3\lambda) \\
& + 24r^6(15-16\lambda+5\lambda^2) + 4r^5(-175+369\lambda-285\lambda^2+75\lambda^3) \\
& + r^4(1465-5304\lambda+6930\lambda^2-3840\lambda^3+765\lambda^4) \\
& - 6r^3(1+239\lambda-1120\lambda^2+1780\lambda^3-1185\lambda^4+285\lambda^5) \\
& - 3r^2(401-2496\lambda+6475\lambda^2-8700\lambda^3+6315\lambda^4-2340\lambda^5+345\lambda^6)) \\
& - 6(-1+r) (16r^8\lambda - 45r(-1+\lambda)^5\lambda + 15(-1+\lambda)^4\lambda^2 + 8r^7\lambda(-13+5\lambda) \\
& + 8r^6\lambda(37-35\lambda+10\lambda^2) + 2r^5(20-267\lambda+445\lambda^2-305\lambda^3+75\lambda^4) \\
& + r^4(220-679\lambda+350\lambda^2+560\lambda^3-570\lambda^4+135\lambda^5) \\
& - 5r^2(12-121\lambda+439\lambda^2-672\lambda^3+468\lambda^4-135\lambda^5+9\lambda^6) \\
& - 5r^3(52-215\lambda+259\lambda^2-204\lambda^3+276\lambda^4-237\lambda^5+69\lambda^6)) \ln(r) \\
& + 15 (16r^8\lambda^2 + 32r^7(-3+\lambda)\lambda^2 + 3(-1+\lambda)^4\lambda^2 + 48r^6\lambda^2(5-4\lambda+\lambda^2) \\
& - 12r^3\lambda(2-43\lambda+65\lambda^2-5\lambda^3-31\lambda^4+12\lambda^5) + 4r^5(6-2\lambda-85\lambda^2+119\lambda^3-69\lambda^4+15\lambda^5) \\
& + 2r^2\lambda(40-243\lambda+374\lambda^2-168\lambda^3-30\lambda^4+27\lambda^5) \\
& + r^4(-36+96\lambda-393\lambda^2+688\lambda^3-414\lambda^4+72\lambda^5+3\lambda^6)) \ln(r)^2 \\
& - 20r^2\lambda (16r^5\lambda^2 + 24r^4(-3+\lambda)\lambda^2 + 24r^3\lambda^2(5-4\lambda+\lambda^2) \\
& - 3\lambda(-24+11\lambda+63\lambda^2-71\lambda^3+21\lambda^4) + 6r(-9+3\lambda+\lambda^2+25\lambda^3-28\lambda^4+8\lambda^5) \\
& + r^2(54-90\lambda-37\lambda^2+111\lambda^3-69\lambda^4+15\lambda^5)) \ln(r)^3 \\
& + 60r^2\lambda^2 (-9+32\lambda^2+4r^4\lambda^2+4r^3(-3+\lambda)\lambda^2-32\lambda^3+9\lambda^4+2r^2\lambda^2(5-4\lambda+\lambda^2) \\
& + r(18-38\lambda+26\lambda^2-6\lambda^3)) \ln(r)^4 - 24r^2\lambda^3 (15-35\lambda+(27-6r^2+4r^3)\lambda^2 \\
& + (-7+2r^2)\lambda^3) \ln(r)^5 + 16r^4\lambda^6 \ln(r)^6, \tag{129}
\end{aligned}$$

and

$$\begin{aligned}
f_6 = & 32r^2(1-r+\ln(r))^6 ((-1+r)(-1+r+\lambda) - r\lambda \ln(r))^2 \times \\
& ((-1+r) (-1+2r^2+r(-1+\lambda)+\lambda) - 2r^2\lambda \ln(r)). \tag{130}
\end{aligned}$$

B Expressions for sandwich dipole with reverse bending

In this section, we present the final expressions for the case of the wiggler.

The un-normalized TME for the wiggler is just $2\sqrt{|A|(1-c)}$, and is given by:

$$\mathcal{F}_{\text{TME}}(\mu, \nu, L_1, \rho_0) = \frac{L_1^3}{|\rho_0|^3} \mathcal{L}(\mu, \nu), \tag{131}$$

where

$$\mathcal{L}(\mu, \nu) \equiv 2\sqrt{\frac{f_7(f_8 + f_9)}{f_{10}}}, \quad (132)$$

with

$$f_7 \equiv \mu^2 (2\mu\nu(3 + 6\nu + \nu^4) + |\mu|^3) \quad (133)$$

$$f_8 \equiv 4\mu^2\nu^4(15 + 30\mu\nu + 16\mu^2\nu^2) \quad (134)$$

$$f_9 \equiv 4\mu\nu (3 + 15\nu + 10(3 + \mu)\nu^2 + 45\mu\nu^3 + 18\mu^2\nu^4) |\mu|^3 + |\mu|^6 \quad (135)$$

$$f_{10} \equiv 153(\mu + 2\nu)^2(1 + 2\mu\nu)^6|\mu|^6(2\mu\nu + |\mu|^3). \quad (136)$$

For the case when the total bending angle is 0, $\nu = -1/2$. Putting it in into the equation above reduces it to:

$$\mathcal{F}_{\text{TME}}(\mu, L_1, \rho_0) = \frac{L_1^3}{|\rho_0|^3} \sqrt{\frac{\mu^2(\mu^2(15 - 15\mu + 4\mu^2) + \mu(-24 + 25\mu - 9\mu^2)|\mu|^3 + 4|\mu|^6)}{540(-1 + \mu)^8|\mu|^6}}, \quad (137)$$

which is positive and finite for $\mu < 0$. Note that we cannot consider $\mu > 0$, because of the necessary condition $\text{sgn}(\mu) = \text{sgn}(\nu)$.

The normalized TME for the wiggler is given by:

$$F_{\text{TME}}(\mu, \nu) = \sqrt{\frac{f_7(f_8 + f_9)}{f_{11}}}, \quad (138)$$

with

$$f_{11} \equiv (1 + 2\nu)^6(\mu + 2\nu)^2|\mu|^6(2\mu\nu + |\mu|^3). \quad (139)$$

C Bound for c

In this section, we prove the following statement about c :

Theorem C.1. *c lies between the range 0 to 1. That is, $0 \leq c \leq 1$.*

Before we present the proof, we note that $|A| > 0$. The proof of that fact is as follows. Since

$$\begin{aligned} |A| &\equiv |\langle\langle \vec{\xi} \vec{\xi}^T \rangle\rangle| \\ &= \left(\frac{\int \frac{\xi^2}{|\rho|^3} ds}{\int \frac{1}{\rho^2} ds} \right) \left(\frac{\int \frac{\xi'^2}{|\rho|^3} ds}{\int \frac{1}{\rho^2} ds} \right) - \left(\frac{\int \frac{\xi \xi'}{|\rho|^3} ds}{\int \frac{1}{\rho^2} ds} \right)^2, \end{aligned} \quad (140)$$

by firstly the Cauchy-Schwartz inequality

$$\left| \int f(s)\bar{g}(s)ds \right|^2 \leq \left(\int |f(s)|^2 ds \right) \left(\int |g(s)|^2 ds \right), \quad (141)$$

with $f(s) = \xi(s)/|\rho(s)|^{3/2}$ and $g(s) = \xi'(s)/|\rho(s)|^{3/2}$ (equality only holds if and only if $f(s) = g(s)$), and secondly the fact that

$$\int \frac{1}{\rho^2} ds \geq 0, \quad (142)$$

and thirdly $\xi(s) \neq \xi'(s)$, we see that

$$|A| > 0, \quad (143)$$

since equality is never reached.

Proof. We prove first the lower bound of c . The c parameter is defined as $-\text{Tr}(JAJB)/|A|$, which we can write out explicitly to be:

$$c = \frac{\langle\langle \xi^2 \rangle\rangle \langle\langle \xi' \rangle\rangle^2 + \langle\langle \xi'^2 \rangle\rangle \langle\langle \xi \rangle\rangle^2 - 2\langle\langle \xi \xi' \rangle\rangle \langle\langle \xi \rangle\rangle \langle\langle \xi' \rangle\rangle}{\check{\rho}|A|}. \quad (144)$$

The denominator is obviously greater than 0, and we note that the numerator can be written as:

$$c = \begin{bmatrix} \langle\langle \xi' \rangle\rangle & -\langle\langle \xi \rangle\rangle \end{bmatrix} \begin{bmatrix} \langle\langle \xi^2 \rangle\rangle & \langle\langle \xi \xi' \rangle\rangle \\ \langle\langle \xi \xi' \rangle\rangle & \langle\langle \xi'^2 \rangle\rangle \end{bmatrix} \begin{bmatrix} \langle\langle \xi' \rangle\rangle \\ -\langle\langle \xi \rangle\rangle \end{bmatrix}. \quad (145)$$

Since the matrix in the middle is precisely $|A|$, which we see to be a positive definite matrix ($|A| > 0$ and $A_{22} > 0$), the expression 145 is always greater than 0.

Now we prove the upper bound of c .

We define an operation on the space of real integrable functions on the compact interval of the real line $[0, L]$:

$$(u, v) \equiv \langle\langle uv \rangle\rangle - \frac{\langle\langle u \rangle\rangle \langle\langle v \rangle\rangle}{\check{\rho}}. \quad (146)$$

This expression is symmetric: $(u, v) = (v, u)$, linear in the first argument: $(au, v) = a(u, v)$ and $(u + w, v) = (u, v) + (w, v)$ and it is semi-positive definite: $(u, u) \geq 0$ for all u .

The first two properties are obvious, and the third one is true because

$$\begin{aligned} (u, u) &\equiv \langle\langle u^2 \rangle\rangle - \frac{\langle\langle u \rangle\rangle \langle\langle u \rangle\rangle}{\check{\rho}} \\ &= \frac{\int \frac{u^2}{|\rho|^3} ds}{\int \frac{1}{\rho^2} ds} - \frac{\left(\int \frac{u}{|\rho|^3} ds \right)^2}{\int \frac{1}{|\rho|^3} ds \int \frac{1}{\rho^2} ds}, \end{aligned} \quad (147)$$

and we invoke the Cauchy-Schwartz inequality again with $f(s) = u(s)/|\rho(s)^{3/2}|$ and $g(s) = 1/|\rho(s)^{3/2}|$ to get

$$(u, u) \geq 0. \quad (148)$$

The operation we have defined is a semi-inner product, so named because the inner product is only semi-positive definite and not positive definite. That is, there exists a non-zero u such that $(u, u) = 0$. This is satisfied for $u = 1$: $(1, 1) = 0$.

The Cauchy-Schwartz inequality still holds for a semi-inner product, that is:

$$(u, v)^2 \leq (u, u)(v, v). \quad (149)$$

The proof of this can be found in a standard analysis textbook. Let u and v be vectors in the vector space defined above, and let t be a real scalar.

$$\begin{aligned} 0 &\leq (u + tv, u + tv) \\ &= (u, u) + 2t(u, v) + t^2(v, v). \end{aligned} \quad (150)$$

Since this is a polynomial in t , and it is always greater than or equals to 0, its discriminant must be less than or equals to 0:

$$4(u, v)^2 - 4(u, u)(v, v) \leq 0, \quad (151)$$

the desired result.

Letting $u = \xi$ and $v = \xi'$ and a rearrangement of terms in the inequality on this semi-inner product (making use of the form of c in 144) then yields

$$c \leq 1. \quad (152)$$

□

References

- [1] C.-x.Wang, "Minimum emittance in storage rings with uniform or nonuniform dipoles," *Physical Review Special Topics - Accelerators and Beams*, vol. 12, no. 061001, 2009.
- [2] Y. W. C.-x.Wang and Y. Peng, "Optimal dipole-field profiles for emittance reduction in storage rings," *Physical Review Special Topics - Accelerators and Beams*, vol. 14, no. 034001, 2011.
- [3] M. Sommer, "Internal report no. dci/ni/20/81," tech. rep., 1981.
- [4] M. Sands, *The physics of electron storage rings, An introduction*. University of California, Santa Cruz, 1970.

- [5] A. Miele, "On the minimization of the product of the powers of several integrals," *Journal of Optimization Theory and Applications*, vol. 1, no. 1, 1967.
- [6] A. Miele and H.-y.Huang, "Missile shapes of minimum ballistic factor," *Journal of Optimization Theory and Applications*, vol. 1, no. 2, 1967.
- [7] A. I. Bunimovich and A. V. Dubinskii, "A variational method for a generalized class of functionals and its application to problems of aeromechanics," *Izvestiya Akademii Nauk SSSR, Mekhanika Zhidkosti Gaza*, no. 1, pp. 103–111, 1973.
- [8] M. R. Hestenes, "The problem of bolza in the calculus of variations," tech. rep., University of Chicago, 1941.