First-Order σ Matrix (Phase Ellipse) Formalism for TRANSPORT

In accelerators and beam transport systems, the behavior of an individual particle is often of less concern than is the behavior of a bundle of particles (the BEAM) of which an individual particle is a member. An extension of the matrix algebra of Eq. (3) provides a convenient means for defining and manipulating this BEAM. TRANSPORT assumes that the bundle of rays constituting a BEAM may correctly be represented in coordinate phase-space by an ellipsoid whose coordinates are the position, angle, and momentum coordinates of the arbitrary rays in the beam about an assumed central trajectory. TRANSPORT is a matrix calculation that truncates the problem to either first- or second-order in a Taylors expansion about the central trajectory. Particles in a BEAM are assumed to lie within the boundaries of the ellipsoid with each point within the ellipsoid representing a possible ray. The sum total of all phase points, the phase space volume, is commonly referred to as the "phase space" occupied by the BEAM. The validity and interpretation of this phase ellipse formalism must be ascertained for each system being designed. However, in general, for charged particle beams in, or emanating, from accelerators, the first-order phase ellipse formalism of TRANSPORT is a reasonable representation of physical reality; but for other applications, such as charged particle spectrometers, caution is in order in its use and interpretation.

The equation of an n-dimensional ellipsoid may be written in matrix form as follows:

$$X(0)^T \ \sigma(0)^{-1} \ X(0) = 1 \tag{27}$$

where $X(0)^T$ is the transpose of the coordinate vector $X(0)$, and $\sigma(0)$ is a real, positive definite, symmetric matrix.

The volume of the n-dimensional ellipsoid defined by sigma is $\frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)}(\det\sigma)^{1/2}$, the area of the projection in one plane is $A = \pi(\det\sigma)^{1/2}$. This is the "phase space" occupied by the beam.

As a particle passes through a system of magnets, it undergoes the matrix transformation of Eq. (3). Combining this transformation with the equation of the initial ellipsoid, and using the identity $RR^{-1} = I$ (the unity matrix), it follows that:

$$X(0)^T \; (R^T R^{T^{-1}}) \;\; \sigma(0)^{-1} \; (R^{-1}R) \;\; X(0) = 1$$

from which:

$$(RX(0))^T \; (R\sigma(0) \; R^T)^{-1} \; (RX(0)) = 1 \tag{28}$$

The equation of the ellipsoid representing the "BEAM" at the end of the system is thus:

$$X(1)^T \;\; \sigma(1)^{-1} \;\; X(1) = 1 \tag{29}$$

where the equation for the sigma matrix at the end may be related to that at the beginning by:

$$\sigma(1) = R\,\sigma(0)\;R^T \tag{30}$$

In addition to calculating the product matrix R, TRANSPORT also computes the sigma "BEAM" matrix at the end of each physical element via Eq. (30).

All of the important physical parameters of the BEAM ellipsoid may be expressed as functions of the matrix elements of the sigma matrix at the location in question. In particular the square roots of the diagonal elements ($\sqrt{\sigma_{ii}}$'s) are the projection of the ellipse upon the coordinate axes and thus represent the maximum extent of the BEAM in the various coordinate directions. The correlation between components (the orientation of the ellipse) is determined by the off-diagonal terms (the σ_{ij}'s). An illustration of this is given below for a 2-dimensional ellipse.

Description of the Sigma BEAM Matrix

Consider a 2-dimensional (x, θ) plane projection of the general 6-dimensional ellipsoid. Let

$$\sigma = \begin{bmatrix} \sigma_{11} & \sigma_{21} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$$

be a real, positive definite, symmetric matrix; the inverse of which is

$$\sigma^{-1} = \frac{1}{\epsilon^2} \begin{bmatrix} \sigma_{22} & -\sigma_{21} \\ -\sigma_{21} & \sigma_{11} \end{bmatrix}$$

where ϵ^2 is the determinant of σ.

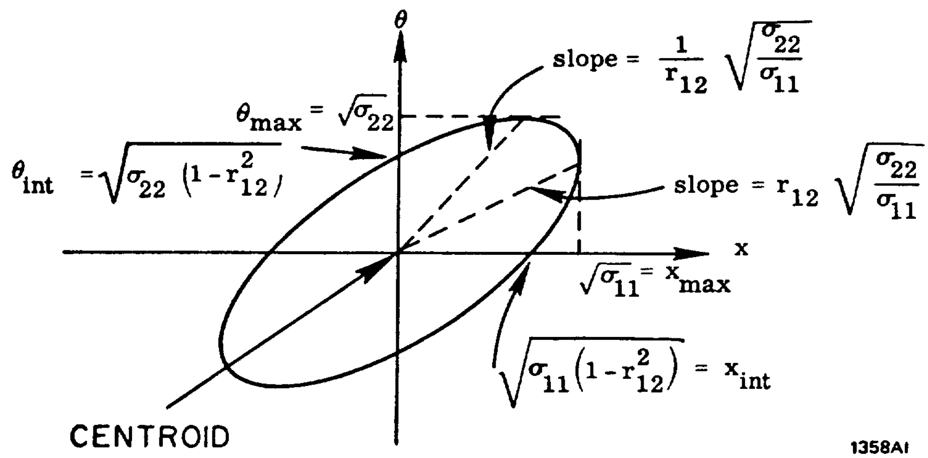
The 2-dimensional coordinate vector (column matrix) and its transpose are:

$$X = \begin{pmatrix} x \\ \theta \end{pmatrix} \quad \text{and} \quad X^T = (x \ \theta)$$

The expansion of the matrix equation $X^T \sigma^{-1} X = 1$ is the equation of the ellipse

$$\sigma_{22} x^2 - 2\sigma_{21} x\theta + \sigma_{11} \theta^2 = \epsilon^2 = \det \sigma \qquad (31)$$

The (x, θ) plane BEAM ellipse represented by Eq. (31) is shown in the following figure along with the physical meaning of the sigma matrix elements.



The area of the ellipse is given by:

$$A = \pi(\det \sigma)^{1/2} = \pi x_{max} \ \theta_{int} = \pi x_{int} \ \theta_{max} \tag{32}$$

A Two Dimensional BEAM Phase Ellipse

The correlation between x and Θ (the orientation of the ellipse) depends upon the off-diagonal term σ_{21}. This correlation is defined as

$$r_{21} = r_{12} = \frac{\sigma_{21}}{\sqrt{\sigma_{11}\ \sigma_{22}}}$$

So defined r always falls in the range

$$-1 \leqslant r \leqslant 1$$

The correlation, r, measures the tilt of the ellipse and the intersection of the ellipse with the coordinate axes.

Since the det R = 1 for all static magnetic beam transport elements, it follows that the determinant of $\sigma(1)$ and $\sigma(0)$ are identical under the transformation of Eq. (30). Hence the "phase space" area is an invariant under the transformation of Eq. (30). This is a statement of Liouville's Theorem for the magnetostatic fields employed and results from the fact that the det R = 1.

It is perhaps worthwhile noting that this 2-dimensional representation of the BEAM matrix has a one to one correspondence with the Courant-Snyder treatment of the theory of the Alternating Gradient Synchrotron* as follows:

$$\begin{bmatrix} \sigma_{11} & \sigma_{21} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} = \epsilon \begin{bmatrix} \beta & -\alpha \\ -\alpha & \gamma \end{bmatrix} \qquad (33)$$

The Phase Ellipse Beam Matrix used by TRANSPORT

For static magnetic systems possessing midplane symmetry, the (x,θ) plane and (y,φ) plane trajectories are decoupled in first-order, i.e. there is no mixing of phase space between the two planes. However for mathematical simplicity and to allow for the possibility of more general systems, the sigma BEAM matrix used in a TRANSPORT calculation has the following general 6-dimensional construction.

	x	θ	y	φ	ℓ	δ
x	σ(11)					
θ	σ(21)	σ(22)				
y	σ(31)	σ(32)	σ(33)			
φ	σ(41)	σ(42)	σ(43)	σ(44)		
ℓ	σ(51)	σ(52)	σ(53)	σ(54)	σ(55)	
δ	σ(61)	σ(62)	σ(63)	σ(64)	σ(65)	σ(66)

* E. D. Courant and H. S. Snyder, "Theory of the Alternating Gradient Synchrotron", Annals of Physics 3, pp 1-48 (1958).

The matrix is symmetric so that only a triangle of elements is needed.

In the printed output this matrix has a somewhat different format for ease of interpretation:

			x	θ	y	φ	ℓ
x	$\sqrt{\sigma(11)}$	CM					
θ	$\sqrt{\sigma(22)}$	MR	r(21)				
y	$\sqrt{\sigma(33)}$	CM	r(31)	r(32)			
φ	$\sqrt{\sigma(44)}$	MR	r(41)	r(42)	r(43)		
ℓ	$\sqrt{\sigma(55)}$	CM	r(51)	r(52)	r(53)	r(54)	
δ	$\sqrt{\sigma(66)}$	PC	r(61)	r(62)	r(63)	r(64)	r(65)

where: $$r(ij) = \frac{\sigma(ij)}{\left[\sigma(ii)\ \ \sigma(jj)\right]^{1/2}} \qquad (34)$$

As a result of the fact that the matrix is positive definite, the r(ij)'s satisfy the relation

$$|r(ij)| \leqslant 1 \qquad (35)$$

The physical meaning of the $\sqrt{\sigma(ii)}$'s is as follows:

$\sqrt{\sigma(11)} = x_{max} =$ The maximum (half)-width of the beam envelope in the x plane at the point of the printout.

$\sqrt{\sigma(22)} = \theta_{max} =$ The maximum (half)-angular divergence of the beam envelope in the x plane.

$\sqrt{\sigma(33)} = y_{max} =$ The maximum (half)-height of the beam envelope.

$\sqrt{\sigma(44)} = \varphi_{max} =$ The maximum (half)-angular divergence of the beam envelope in the y plane.

$\sqrt{\sigma(55)} = \ell_{max} =$ 1/2 the longitudinal extent of the bunch of particles.

$\sqrt{\sigma(66)} = \delta =$ The half-width 1/2 ($\Delta P/P$) of the momentum interval being transmitted by the system.

The units appearing next to the $\sqrt{\sigma(ii)}$'s in the TRANSPORT printout sheet are the units chosen for coordinates x, θ, y, φ, ℓ and $\delta = \Delta P/P$ respectively.

To the immediate left of the listing of the beam envelope size in a TRANSPORT printout, there appears a column of numbers whose values will normally be zero. These numbers are the coordinates of the centroid of the beam phase ellipse (with respect to the initially assumed central trajectory of the system). They may become non-zero under one of three circumstances:

1) When the misalignment (Type Code 8.) is used.

2) When a Beam Centroid shift (Type Code 7.) is used.

or 3) When a 2nd-order calculation (Type Code 17.) is used.

Physical Interpretation of Various Projections of the 2-dimensional BEAM Ellipse

Consider again Eq. (30) $\sigma(1) = R\,\sigma(0)\,R^T$ and expand it in it's most general form for the 2-dimensional (x,θ) plane case.

$$\sigma(1) = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \begin{pmatrix} \sigma_{11}(0) & \sigma_{21}(0) \\ \sigma_{21}(0) & \sigma_{22}(0) \end{pmatrix} \begin{pmatrix} R_{11} & R_{21} \\ R_{12} & R_{22} \end{pmatrix} = \begin{pmatrix} \sigma_{11}(1) & \sigma_{21}(1) \\ \sigma_{21}(1) & \sigma_{22}(1) \end{pmatrix}$$

the result is:

$$\sigma(1) = \left[\begin{array}{c|c} R_{11}^2\sigma_{11}(0)+2R_{11}R_{12}\sigma_{21}(0)+R_{12}^2\sigma_{22}(0) & R_{11}R_{21}\sigma_{11}(0)+(R_{11}R_{22}+R_{12}R_{21})\sigma_{21}(0)+R_{12}R_{22}\sigma_{22}(0) \\ \hline \swarrow & R_{21}^2\sigma_{11}(0)+2R_{21}R_{22}\sigma_{21}(0)+R_{22}^2\sigma_{22}(0) \end{array}\right] \quad (36)$$

In the special case when the initial ellipse is erect i.e., $\sigma_{21}(0)=0$, $\sigma(1)$ reduces to:

$$\sigma(1) = \left[\begin{array}{c|c} R_{11}^2\sigma_{11}(0) + R_{12}^2\sigma_{22}(0) & R_{11}R_{21}\sigma_{11}(0) + R_{12}R_{22}\sigma_{22}(0) \\ \hline \swarrow & R_{21}^2\sigma_{11}(0) + R_{22}^2\sigma_{22}(0) \end{array}\right] \quad (37)$$

Similar results are, of course, obtained for the (y,φ) plane.

If an arbitrary beam transport system is reduced to the most elementary first-order form of representing it as an initial drift distance, followed by a lens action between two principal planes, and a final drift distance; then we observe that for the 2-dimensional representation there are only two basic phase ellipse transformations of interest.

(1) An arbitrary DRIFT distance and

(2) A LENS action

Each of these elementary cases are illustrated on Fig. 8 for both a parallelogram as well as ellipse phase space transformations. Note that a DRIFT followed by a LENS action is not necessarily equal to a LENS action followed by a DRIFT; i.e., the matrices do not necessarily commute.

The phase ellipse transformations for a DRIFT and for a LENS action (between principal planes) as shown in Fig. (8) may be readily calculated using the results of Eq. (37).

The 2-dimensional R matrix representing a drift of distance L is:

$$R(\text{Drift}) = \begin{pmatrix} 1 & L \\ 0 & 1 \end{pmatrix} \tag{38}$$

Substituting into Eq. (37) we find

$$\sigma(1) = \left[\begin{array}{c|c} \sigma_{11}(0) + L^2\sigma_{22}(0) & L\sigma_{22}(0) \\ \hline L\sigma_{22}(0) & \sigma_{22}(0) \end{array}\right] = \left[\begin{array}{c|c} \sigma_{11}(1) & \sigma_{21}(1) \\ \hline \sigma_{21}(1) & \sigma_{22}(1) \end{array}\right] \tag{39}$$

Attaching the physical meaning to the matrix elements yields the following interpretations:

$$\sigma_{11}(1) = \sigma_{11}(0) + L^2\sigma_{22}(0)$$

or

$$(x_1^2)_{max} = (x_0^2)_{max} + L^2\,(\theta_0^2)_{max} \tag{40}$$

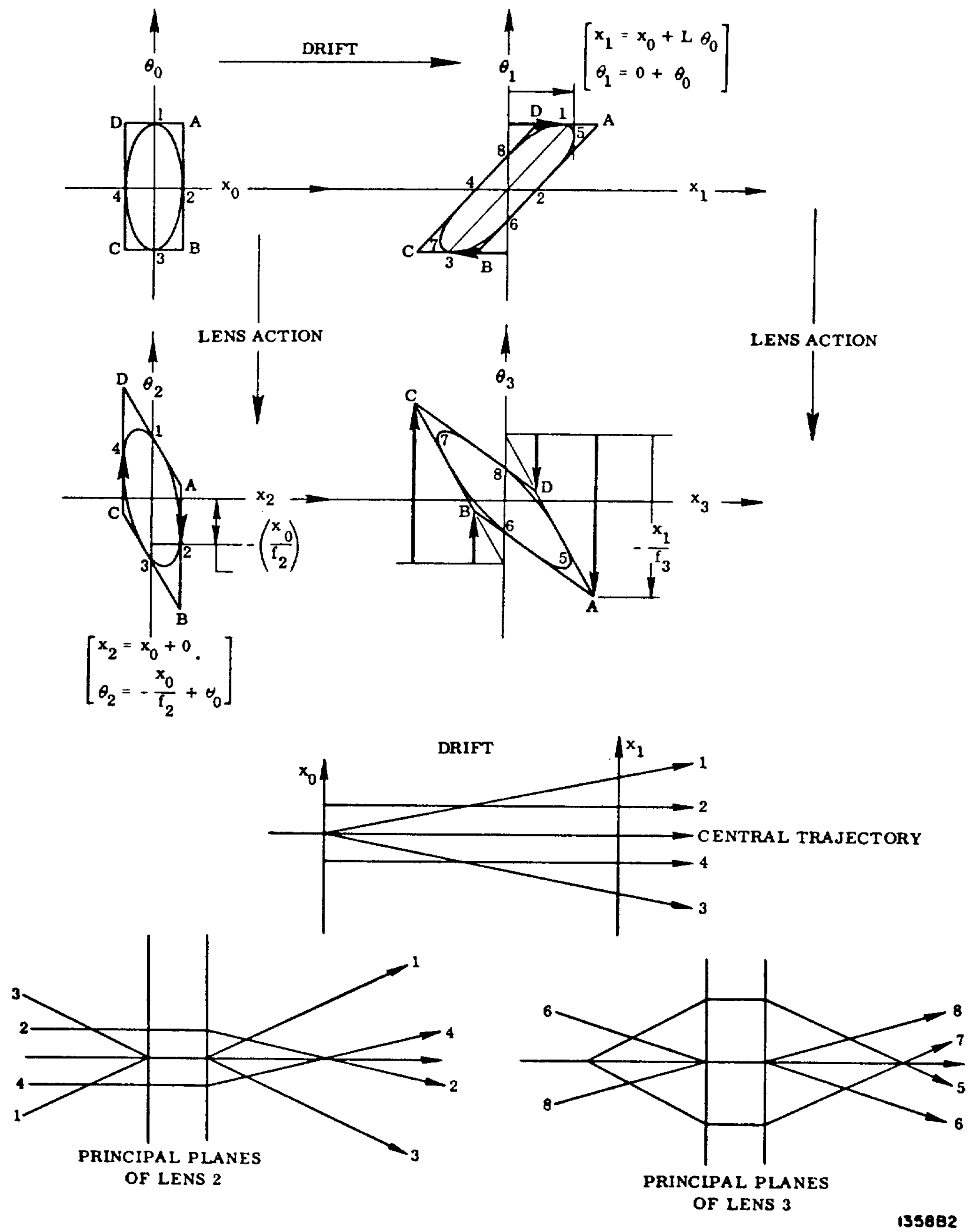

DRIFT
LENS ACTION
LENS ACTION
$x_1 = x_0 + L\,\theta_0$
$\theta_1 = 0 + \theta_0$
$x_2 = x_0 + 0.$
$\theta_2 = -\frac{x_0}{f_2} + \theta_0$
$-\left(\frac{x_0}{f_2}\right)$
$-\frac{x_1}{f_3}$
DRIFT
CENTRAL TRAJECTORY
PRINCIPAL PLANES OF LENS 2
PRINCIPAL PLANES OF LENS 3
135882

Fig. 8

similarly

$$\sigma_{22}(1) = \sigma_{22}(0)$$

or

$$(\theta_1^2)_{max} = (\theta_0^2)_{max} \tag{41}$$

Note that this transformation assumes that the initial phase ellipse is erect, i.e., $\sigma_{21}(0) = 0$.

The 2-dimensional R matrix for a lens actions (between principal planes) is

$$R(\text{Lens}) = \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \tag{42}$$

Substitution into Eq. (37) yields

$$\sigma(2) = \left[\begin{array}{c|c} \sigma_{11}(0) & -\frac{\sigma_{11}(0)}{f} \\ \hline -\frac{\sigma_{11}(0)}{f} & \frac{\sigma_{11}(0)}{f^2} + \sigma_{22}(0) \end{array}\right] = \left[\begin{array}{c|c} \sigma_{11}(2) & \sigma_{21}(2) \\ \hline \sigma_{21}(2) & \sigma_{22}(2) \end{array}\right] \tag{43}$$

Again attaching physical meaning to the matrix elements we have:

$$\sigma_{11}(2) = \sigma_{11}(0)$$

or

$$(x_2^2)_{max} = (x_0^2)_{max} \tag{44}$$

and

$$\sigma_{22}(2) = \frac{\sigma_{11}(0)}{f^2} + \sigma_{22}(0)$$

or

$$(\theta_2^2)_{max} = \frac{1}{f^2} (x_0^2)_{max} + (\theta_0^2)_{max} \tag{45}$$

Note the change in sign of the σ_{21} elements for the Drift and the Lens actions indicating the different sense of orientation of the resulting ellipses as illustrated in Fig. 8.

The Upright Ellipse:

A case of particular interest in any 2-dimensional phase ellipse projection (e.g., the (x,θ) plane or the (y,φ) plane) is when the off-diagonal correlation matrix elements are equal to zero; i.e., an erect ellipse. In a field-free region this corresponds to a so-called "waist" in the BEAM as illustrated in Fig. 9.

It is important to understand correctly the meaning of a waist: For an existing beam, it is the location of the minimum beam size in a given region of the system. Although the waist is the minimum beam size in any given beam, the minimum beam size achievable at a fixed target position by varying the focal strength of the preceding lens system is not the same as the above defined waist. See Fig. 10. In a field-free region, the minimum beam spot size achievable at a fixed target position will occur when the preceding lens system is adjusted such that a waist precedes the target position. Only in the limit of zero phase space area do these quantities occur at the same location. A useful criterion that determines the physical proximity of these quantities is the following: If the system has been adjusted for the smallest spot size at a fixed position and if the size of the beam at the principal planes of the optical system is large compared to its size at the waist, or at the minimum spot size, then the location of these quantities, the waist and the minimum, will closely coincide; if, on the other hand, the size of the beam does not change substantially throughout the system, then the locations of a waist and the minimum beam size may (and usually will) differ substantially. The numerical proximity of these two quantities will be discussed in greater detail later in the report.

In a field free region (i.e., a Drift), the distance to a waist from any location may be readily calculated if the σ matrix at the location is known. Using Eq. (36) and the R matrix for a Drift (Eq. 38) we have for the (x,θ) plane:

$$\sigma_{21}(1) = \sigma_{21}(0) + L\,\sigma_{22}(0) = 0 \text{ (specifying that } \sigma(1) \text{ shall be at a waist)}$$

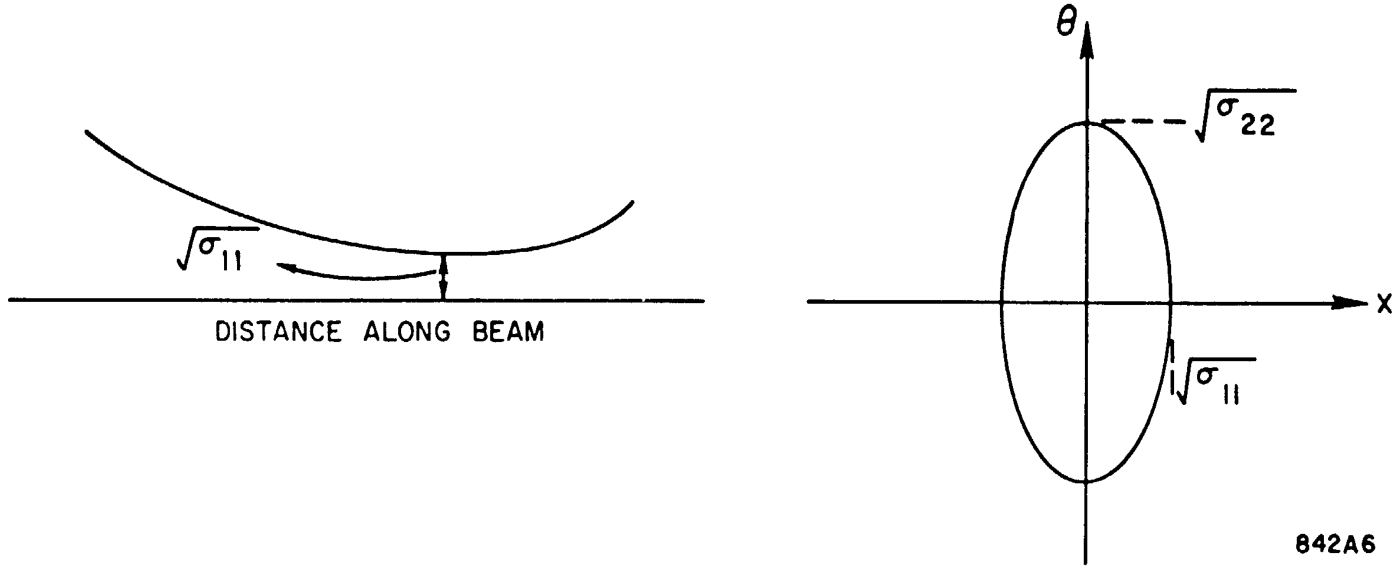


FIG. 9--BEAM WAIST.

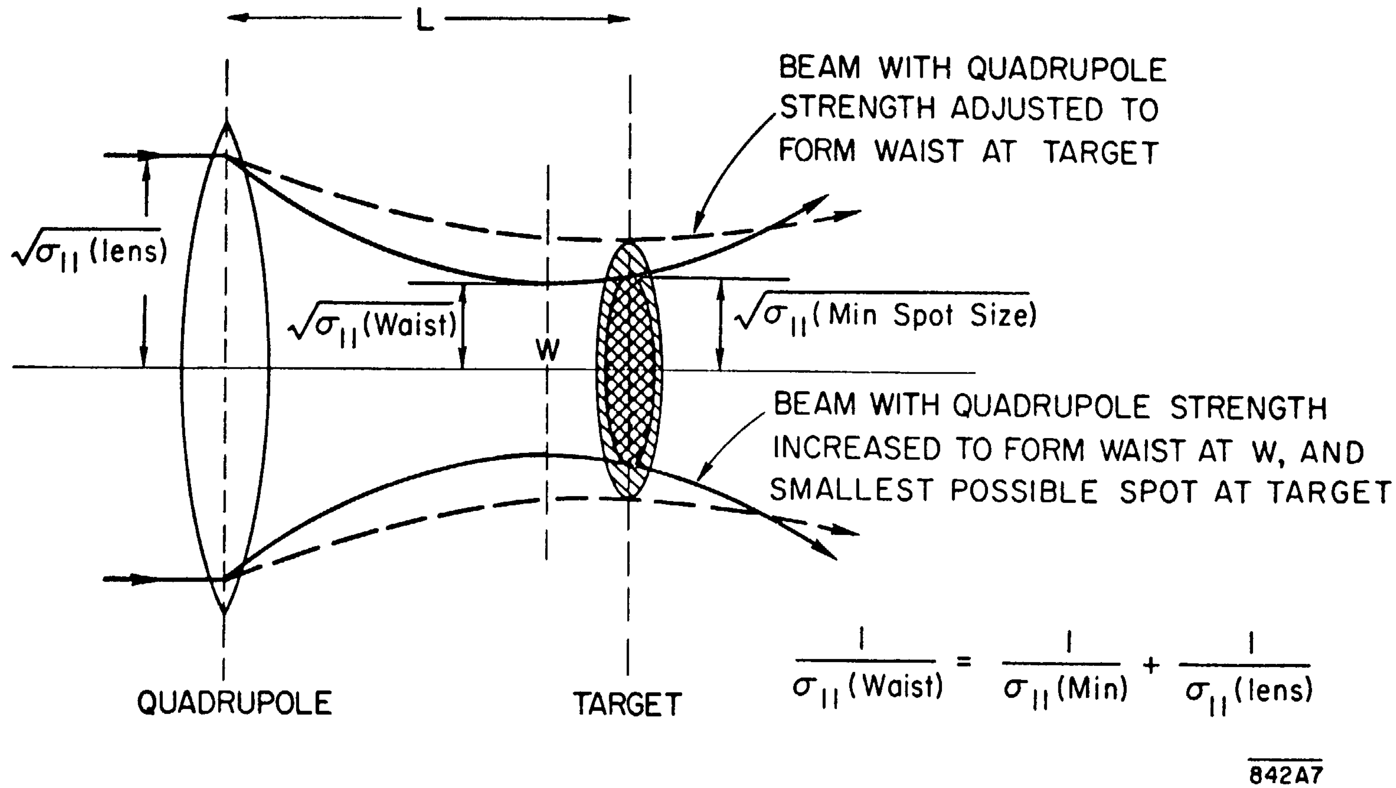

L
BEAM WITH QUADRUPOLE
STRENGTH ADJUSTED TO
FORM WAIST AT TARGET
$\sqrt{\sigma_{11}}$ (lens)
$\sqrt{\sigma_{11}}$(Waist)
W
$\sqrt{\sigma_{11}}$(Min Spot Size)
BEAM WITH QUADRUPOLE STRENGTH
INCREASED TO FORM WAIST AT W, AND
SMALLEST POSSIBLE SPOT AT TARGET
QUADRUPOLE
TARGET
$\frac{1}{\sigma_{11}(\text{Waist})} = \frac{1}{\sigma_{11}(\text{Min})} + \frac{1}{\sigma_{11}(\text{lens})}$
842A7

Fig. 10

or the distance to the waist is:

$$L = -\frac{\sigma_{21}(0)}{\sigma_{22}(0)} = -r_{21}\sqrt{\frac{\sigma_{11}}{\sigma_{22}}} \tag{46}$$

Similarly for the (y,φ) plane the distance to a waist is:

$$L = -\frac{\sigma_{43}(0)}{\sigma_{44}(0)} = -r_{43}\sqrt{\frac{\sigma_{33}}{\sigma_{44}}} \tag{47}$$

Relationship between a Waist and a Parallel-to-Point Image

A parallel-to-point image in the (x,θ) plane occurs when $R_{11} = 0$. The R matrix corresponding to this is

$$R = \left(\begin{array}{c|c} 0 & R_{12} \\ \hline R_{21} & R_{22} \end{array}\right) = \left(\begin{array}{c|c} 0 & s_x \\ \hline c'_x & s'_x \end{array}\right) = \left(\begin{array}{c|c} 0 & f_x \\ \hline -\frac{1}{f_x} & s'_x \end{array}\right) \tag{48}$$

Since $|R| = 1$, $R_{12}R_{21} = -1$ for this situation.

If we assume an erect ellipse $\sigma(0)$ as the beginning of the system, the final beam matrix $\sigma(1)$ is given by substitution of Eq. (48) into Eq. (37) as follows:

$$\sigma(1) = \left[\begin{array}{c|c} R_{12}^2\sigma_{22}(0) & R_{12}R_{22}\sigma_{22}(0) \\ \hline R_{12}R_{22}\sigma_{22}(0) & R_{21}^2\sigma_{11}(0) + R_{22}^2\sigma_{22}(0) \end{array}\right] \tag{49}$$

for parallel-to-point imaging.

Several conclusions may be extracted from this result: The first observation is that a waist and a parallel-to-point image will coincide if $R_{11}=R_{22}=0$. This is equivalent to requiring that the object and image distances (measured to the principal planes) are both equal to the focal length f of the system.

The distance to a waist in this example is:

$$L = -\frac{\sigma_{21}(1)}{\sigma_{22}(1)} = -\frac{R_{22}R_{12}\sigma_{22}(0)}{R_{21}^2\sigma_{11}(0)+R_{22}^2\sigma_{22}(0)} = -\frac{s_x s_x'\sigma_{22}(0)}{(c_x')^2\sigma_{11}(0)+(s_x')^2\sigma_{22}(0)} \quad (50)$$

If $s_x s_x' = R_{12}R_{22} = 0$, a waist and a parallel-to-point image coincide. If $s_x s_x' = R_{12}R_{22} > 0$, the waist precedes the image; and if $s_x s_x' = R_{12}R_{22} < 0$, the waist follows the image; unless $\sigma_{22}(0) = 0$ (zero phase space area) in which case a waist and an image always coincide.

The size of the beam at the image is:

$$(x_1^2)_{max} = \sigma_{11}(1) = R_{12}^2\sigma_{22}(0) = f^2(\theta_0^2)_{max} \quad (51)$$

independent of the source size x_o and of the object distance.

The size of the beam at the waist is:

$$(\text{size at waist})^2 = \frac{|\sigma(0)|}{\sigma_{22}(1)} = \frac{\sigma_{11}(0)\sigma_{22}(0)}{R_{21}^2\sigma_{11}(0) + R_{22}^2\sigma_{22}(0)} \quad (52)$$

If $R_{22} = s_x' = 0$, the two sizes are equal as expected, otherwise the size at the waist is always smaller.

Relationship between a Waist and a Point-to-Point Image

A point-to-point first-order image in the (x,θ) plane occurs when $R_{12} = s_x = 0$. The R matrix representing this case is:

$$R = \left(\begin{array}{c|c} R_{11} & 0 \\ \hline R_{21} & R_{22} \end{array}\right) = \left(\begin{array}{c|c} c_x & 0 \\ \hline c_x' & s_x' \end{array}\right) = \left(\begin{array}{c|c} M & 0 \\ \hline -\frac{1}{f_x} & \frac{1}{M} \end{array}\right) \quad (53)$$

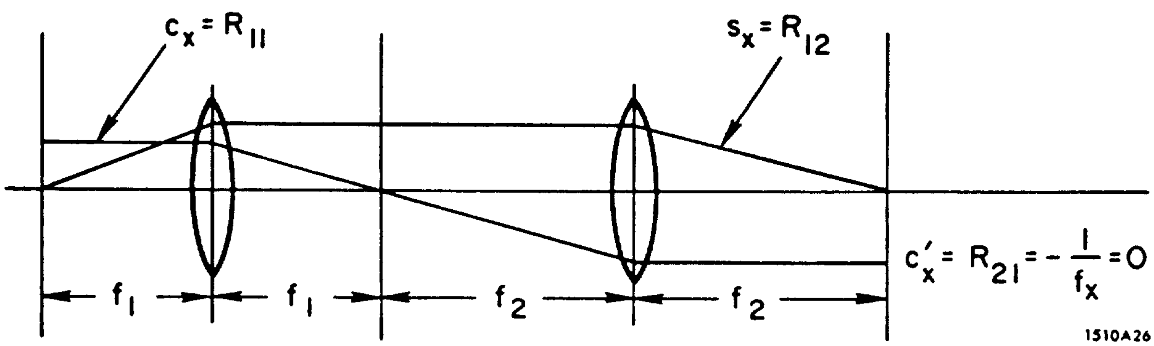
where $|R| = 1 = R_{11}R_{22}$, and M is the magnification.

If we again assume an erect ellipse $\sigma(0)$ as the beginning, the final beam matrix $\sigma(1)$ is given by Eq. (37) as:

$$\sigma(1) = \left[\begin{array}{c|c} R_{11}^2\sigma_{11}(0) & R_{11}R_{21}\sigma_{11}(0) \\ \hline R_{11}R_{21}\sigma_{11}(0) & R_{21}^2\sigma_{11}(0) + R_{22}^2\sigma_{22}(0) \end{array}\right] \qquad (54)$$

for point-to-point imaging.

Our first observation is that except for a zero source size, an image and a waist will coincide only if $R_{12} = R_{21} = 0$. Clearly this is not possible with a single lens; at least two lenses are needed. Such an optical situation is as follows:



The distance to a waist is

$$L = -\frac{\sigma_{21}(1)}{\sigma_{22}(1)} = -\frac{R_{11}R_{21}\sigma_{11}(0)}{R_{21}^2\sigma_{11}(0) + R_{22}^2\sigma_{22}(0)} \qquad (55)$$

So if $R_{11}R_{21} = c_x c_x' = 0$, a waist and a point-to-point image coincide. If $c_x c_x' > 0$, the waist precedes the image and if $c_x c_x' < 0$, the waist follows the image.

The size of the beam at the image is

$$(x_1^2)_{max} = \sigma_{11}(1) = R_{11}^2\sigma_{11}(0) = (Mx_o)^2 \qquad (56)$$

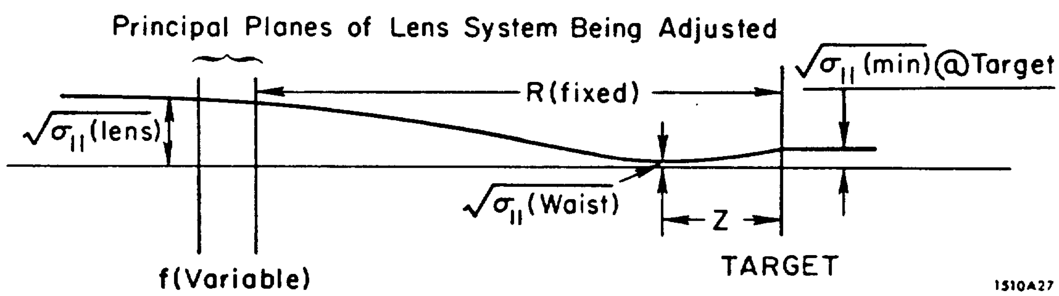
and the size of the beam at the waist is:

$$(\text{size at waist})^2 = \frac{|\sigma(0)|}{\sigma_{22}(1)} = \frac{\sigma_{11}(0)\sigma_{22}(0)}{R_{21}^2\sigma_{11}(0) + R_{22}^2\sigma_{22}(0)} \qquad (57)$$

Thus if $R_{21} = 0$, the two sizes are equal since $|R| = R_{11}R_{22} = 1$. Otherwise the size at the waist is smaller than the image size.

Relationship between a Waist and the Smallest Spot Size Achievable at a Fixed Target Position

Consider the following general situation:



Assume that the size of the beam $\sqrt{\sigma_{11}(\text{lens})}$ at the principal planes of the lens system being adjusted is held constant (i.e., that no other preceding optical elements of the system are being varied); and that the remainder of the system may be represented by a general matrix R which is also held constant. The focal length f is then varied until a minimum spot size $\sqrt{\sigma_{11}(\text{min})}$ is achieved at the target location. The sigma beam matrix at the target position then has the following unique form independent of the orientation of the initial beam ellipse at the lens.

$$\sigma(\text{at target) for a minimum spot size at target} = \begin{bmatrix} \dfrac{R_{12}^2|\sigma|}{\sigma_{11}(\text{lens})} & \dfrac{R_{12}R_{22}|\sigma|}{\sigma_{11}(\text{lens})} \\ \swarrow & \dfrac{\sigma_{11}(\text{lens})}{R_{12}^2} + \dfrac{R_{22}^2|\sigma|}{\sigma_{11}(\text{lens})} \end{bmatrix} = \begin{bmatrix} \sigma_{11}(\text{min}) & \sigma_{21}(\text{min}) \\ \swarrow & \sigma_{22}(\text{min}) \end{bmatrix} \tag{58}$$

thus

$$\sigma_{11}(\text{min}) = \frac{R_{12}^2|\sigma|}{\sigma_{11}(\text{lens})}$$

or

$$x_{\text{min}} = \frac{R_{12}|\sigma|^{1/2}}{x(\text{lens})} \tag{59}$$

If the position of the waist and the minimum beam size both fall within the same field-free region, then the distance to the waist from the target is:

$$Z = -\frac{\sigma_{21}(\text{min})}{\sigma_{22}(\text{min})} = -\frac{R_{12}R_{22}}{R_{22}^2 + \dfrac{\sigma_{11}(\text{lens})}{\sigma_{11}(\text{min})}} = -R_{12}R_{22}\frac{\sigma_{11}(\text{waist})}{\sigma_{11}(\text{lens})} \tag{60}$$

So if $s_x s_x' = R_{12}R_{22} = 0$, the waist and the minimum spot size coincide. If $R_{12}R_{22} > 0$, the waist precedes the target; and if $R_{12}R_{22} < 0$, the waist occurs after the target position.

If the waist and the target positions fall within the same field-free region, the following simple relationship exists between the beam size at the lens $\sqrt{\sigma_{11}(\text{lens})}$, at the waist $\sqrt{\sigma_{11}(\text{waist})}$, and at the target $\sqrt{\sigma_{11}(\text{min})}$.

$$\frac{1}{\sigma_{11}(\text{waist})} = \frac{1}{\sigma_{11}(\text{min})} + \frac{R_{22}^2}{\sigma_{11}(\text{lens})} \tag{61}$$

If now the lens system is readjusted to form a waist at the target position as shown by the dotted lines in Fig. 10, the relative size of this waist and the minimum spot size achieved by the previous lens setting is:

$$\frac{\sigma_{11}(\text{min at target})}{\sigma_{11}(\text{waist at target})} = 1 - \frac{R_{22}^2\sigma_{11}(\text{waist at target})}{\sigma_{11}(\text{lens})} \tag{62}$$

Again we observe that the two quantities approach each other if the size of the beam at the lens is large compared to the beam size at the target.

There are several cases of special interest that may be derived from the above equations:

1) If $R_{22} = 0$ at the target position, then a minimum spot size at the target is also a waist. This corresponds to point-to-parallel imaging from the principal planes of the variable lens system to the target position. Beyond the last lens in the field-free region preceding the target, R_{12} = a constant if $R_{22} = 0$; thus we conclude from Eq. (59) that in this field-free region, the minimum spot size achievable at a target is a waist and is independent of the target position. Such a system is a "Zoom" lens.

2) If there are no lenses beyond the variable lens system, i.e., R is an entirely field-free region (a drift), then R is of the form:

$$R = \begin{pmatrix} 1 & L \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}$$

In this situation $R_{12}R_{22} = L$ is a positive quantity, consequently the waist <u>always</u> precedes a minimum spot size at a target. A case of particular interest is when the minimum spot size achievable is equal to the initial beam size at the lens. It then follows from Eq. (60) that $Z = -L/2$, i.e., a waist occurs midway between the lens and the target. From Eq. (61), the ratio of the size of the beam at the lens and at the waist is:

$$\frac{x(\text{lens})}{x(\text{waist})} = \sqrt{\frac{\sigma_{11}(\text{lens})}{\sigma_{11}(\text{waist})}} = \sqrt{2} \tag{63}$$

Combining this result with Eq. (59),

$$L = R_{12} = \frac{x(\text{lens})x(\text{min})}{|\sigma|^{1/2}} = \frac{x^2(\text{lens})}{x(\text{waist})\theta(\text{waist})}$$

or

$$L = \frac{2\ x(\text{waist})}{\theta(\text{waist})} \tag{64}$$

where L is the longest distance a beam can drift without exceeding its initial size at the lens.

<u>Imaging from an Erect Ellipse to an Erect Ellipse</u>

The general sigma matrix for imaging from an erect ellipse to an erect ellipse may be derived by inspection from Eq. (36) by setting $\sigma_{21}(1) = \sigma_{21}(0) = 0$. The result is:

$$\sigma(1) = \begin{bmatrix} R_{11}^2\sigma_{11}(0) + R_{12}^2\sigma_{22}(0) & 0 \\ 0 & R_{21}^2\sigma_{11}(0) + R_{22}^2\sigma_{22}(0) \end{bmatrix} = \begin{bmatrix} \sigma_{11}(1) & 0 \\ 0 & \sigma_{22}(1) \end{bmatrix} \tag{65}$$

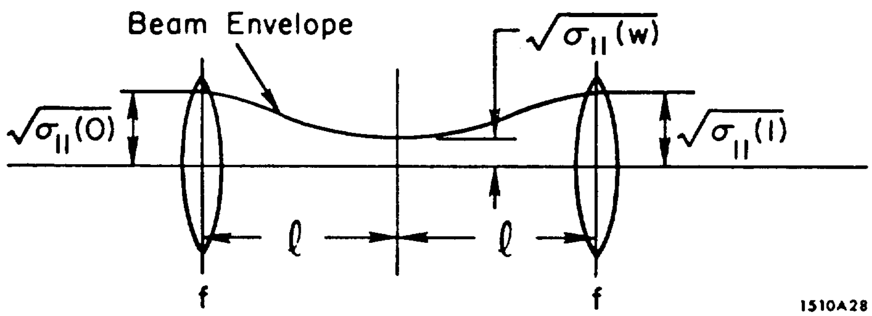
For symmetric magnetic systems $R_{11} = R_{22}$. Using this property and the fact that $|R| = 1$, it follows that $R_{12}R_{21} = (R_{11}^2 - 1)$. So for symmetric magnetic systems Eq. (65) reduces to:

$$\sigma(1) = \begin{bmatrix} \left[-\frac{R_{12}|\sigma|}{R_{21}}\right]^{1/2} & 0 \\ 0 & \left[-\frac{R_{21}|\sigma|}{R_{12}}\right]^{1/2} \end{bmatrix} = \begin{bmatrix} \sigma_{11}(1) & 0 \\ 0 & \sigma_{22}(1) \end{bmatrix} \tag{66}$$

The above equations may be used to calculate the optimum design parameters for periodic beam transport systems.

Example No. 1:

Consider a unit-cell of a periodic focusing array consisting of focusing elements only as indicated below.



The R matrix for the unit-cell, i.e., from the principal planes of the first lens to the principal planes of the second lens, is

$$R = \begin{bmatrix} R_{11} & 2\ell \\ -\frac{1}{f} + \frac{\ell}{2f^2} & R_{11} \end{bmatrix} \tag{67}$$

If now we require that the beam envelope possess symmetry coincident with the lens symmetry i.e., that erect ellipses occur at the principal planes of each lens and a waist midway in between, and furthermore that the beam size at the second lens be kept to a minimum and equal to the beam size at the first lens: then substituting Eq.(67) into Eq. (66) and setting $\sigma_{11}(1)$ to be a minimum yields:

$$\sqrt{\sigma_{11}(1)} = \sqrt{\sigma_{11}(0)}$$

$$f = \ell = \frac{1}{2}\sqrt{\frac{\sigma_{11}(0)}{\sigma_{22}(0)}} = \sqrt{\frac{\sigma_{11}(w)}{\sigma_{22}(w)}} \tag{68}$$

where $\sigma_{11}(0)$ and $\sigma_{22}(0)$ are measured at the principal planes of the first lens.

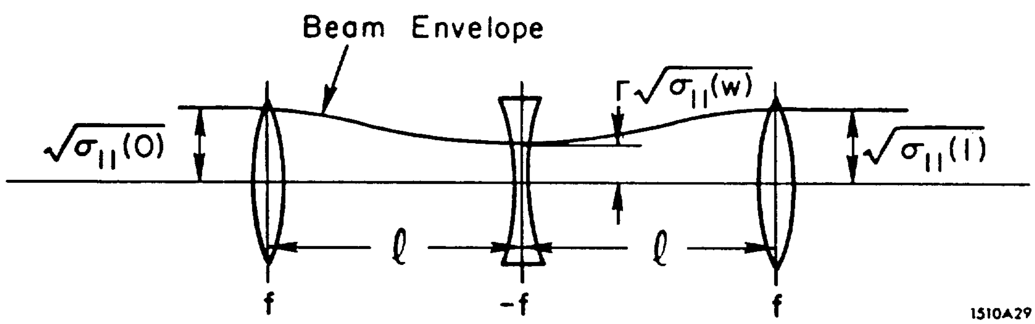
and finally

$$\frac{x(\max)}{x(\min)} = \sqrt{\frac{\sigma_{11}(0)}{\sigma_{11}(w)}} = \sqrt{2} \qquad (69)$$

Note that the ratio of the maximum to the minimum beam size (Eq. 69) is independent of the phase space area and of the lens spacing.

Example No. 2:

If the unit cell is a FODO array as follows:



The R matrix for the unit cell (from the principal planes of the first lens to the principal planes of the third lens) is:

$$R = \left[\begin{array}{c|c} R_{11} & \ell\left(2 + \frac{\ell}{f}\right) \\ \hline -\frac{\ell}{4f^2}\left(2 - \frac{\ell}{f}\right) & R_{11} \end{array}\right] \qquad (70)$$

If we now impose the symmetry requirements that erect ellipses occur at the principal planes of each lens; and that the beam size $\sqrt{\sigma_{11}(1)}$ at lens 3 be a minimum and equal to the beam size $\sqrt{\sigma_{11}(0)}$ at lens 1, then it follows that:

$$\sqrt{\sigma_{11}(1)} = \sqrt{\sigma_{11}(0)}$$

$$\frac{\ell}{f} = \frac{4}{1+\sqrt{5}} \cong 1.236 \qquad (71)$$

$$\frac{\sigma_{11}(0)}{\sigma_{11}(w)} = -\frac{3+\sqrt{5}}{1-\sqrt{5}} \cong 4.23$$

or

$$\frac{x(\max)}{x(\min)} = \sqrt{\frac{\sigma_{11}(0)}{\sigma_{11}(w)}} \cong 2.055 \qquad (72)$$

and finally

$$\ell \cong 0.3003 \sqrt{\frac{\sigma_{11}(0)}{\sigma_{22}(0)}} \qquad (73)$$

where $\sigma_{11}(0)$ and $\sigma_{22}(0)$ are measured at the principal planes of the first lens in the FODO array.

For a FODO quadrupole array where the field strength is held constant for all elements rather than the focal lengths, the results are somewhat different than those above. This case may be readily calculated via TRANSPORT using the above results as initial guesses in the calculation.

Relationship between a First-Order Point-to-Point Image and the Minimum Spot Size Achievable at a Fixed Target Position

This problem is not as easy to explore as were the preceding ones because the question arises "the first-order image of what?" If, however, we restrict the discussion to a thick or thin lens system that does not have intermediate images between the source and the image under consideration, then the following comments are applicable.

The ratio of the minimum beam size to the size of a first-order image at a fixed target position may be calculated using Eq.'s (56) and (59). From Eq. (59) we have:

$$\sigma_{11}(\min) = \frac{L^2|\sigma|}{\sigma_{11}(\text{lens})}$$

and from Eq. (56) the size of a first-order image at the target position is:

$$\sqrt{\sigma_{11}(\text{image})} = |M| \sqrt{\sigma_{11}(\text{object})} = (\frac{L}{p}) \sqrt{\sigma_{11}(\text{object})}$$

where M is the magnification of the first-order image, p is the object distance measured to the principal planes, and L is the distance to the target measured from the principal planes.

The ratio of sizes is

$$\frac{\sigma_{11}(\text{min})}{\sigma_{11}(\text{1st order image})} = \frac{p^2|\sigma|}{\sigma_{11}(\text{object})\sigma_{11}(\text{lens})} \tag{74}$$

Using Eq. (36), we may write

$$\sigma_{11}(\text{lens}) = \sigma_{11}(\text{object}) + 2p\ \sigma_{21}(\text{object}) + p^2\ \sigma_{22}(\text{object}) \tag{75}$$

and since

$$|\sigma| = \sigma_{11}(\text{object})\ \sigma_{22}(\text{object}) - \sigma_{21}{}^2(\text{object})$$

it follows that the first-order image will coincide with the smallest spot size only if the orientation of the initial beam ellipse at the object is such that

$$p\ \sigma_{21}(\text{object}) = -\ \sigma_{11}(\text{object}) \tag{76}$$

or if $\sigma_{11}(\text{object}) = 0$ i.e., for a point source.

For an erect ellipse at the source and the lens adjusted to provide a minimum spot size at the target, it can be shown that the first-order image will always follow the target position (the minimum spot size) by a distance

$$Z = L|M|\ \frac{\sigma_{11}(\text{object})}{\sigma_{11}(\text{lens})} = L|M|\ \frac{x_o^2}{x^2(\text{lens})} \tag{77}$$

where L is the distance to the target position from the principal planes of the lens system, x_o is the source size, and M is the magnification of the first-order image. Again we observe that the ratio of the beam size at the source and the beam size at the "lens" is the criterion determining the proximity of these two quantities.

Orientation of the Major Axes of a Phase Space Ellipse

The matrix equation for a coordinate rotation as shown in Fig. 11 is

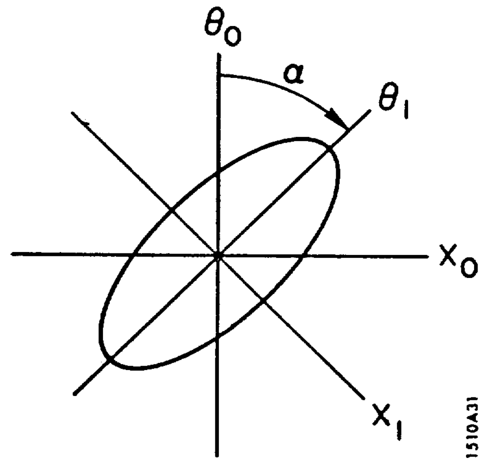


Fig. 11

$$\begin{pmatrix} x_1 \\ \theta_1 \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x_o \\ \theta_o \end{pmatrix} \qquad (78)$$

or

$$X_1 = M\, X_o$$

The equation of an ellipse in either set of coordinates is

$$X^T \sigma^{-1} X = 1 \qquad \text{where} \quad X = \begin{pmatrix} x \\ \theta \end{pmatrix};\ X^T = (x\ \theta)$$

and the transformation from $\sigma(0)$ to $\sigma(1)$ is

$$\sigma(1) = M\, \sigma(0)\, M^T \qquad (79)$$

provided $|M| = 1$, which it does.

If we assume a general ellipse for $\sigma(0)$ and an "erect" ellipse for $\sigma(1)$, i.e.,

$$\sigma(0) = \left[\begin{array}{c|c} \sigma_{11}(0) & \sigma_{21}(0) \\ \hline \sigma_{21}(0) & \sigma_{22}(0) \end{array}\right] \quad \text{and } \sigma(1) = \left[\begin{array}{c|c} \sigma_{11}(1) & 0 \\ \hline 0 & \sigma_{22}(1) \end{array}\right]$$

It follows from Eq. (72) that:

$$\sigma_{21}(1) = 0 = M_{11}M_{21}\sigma_{11}(0) + (M_{11}M_{22} + M_{21}M_{12})\sigma_{12}(0) + M_{12}M_{22}\sigma_{22}(0)$$

from which

$$\text{Tan } 2\alpha = \frac{2\sigma_{21}(0)}{\sigma_{22}(0) - \sigma_{11}(0)} \qquad (80)$$

or using the definition

$$r_{21} = \frac{\sigma_{21}}{\sqrt{\sigma_{11}\ \sigma_{22}}}$$

an alternate form of expressing the ellipse orientation is

$$\operatorname{Tan} 2\alpha = \frac{2r_{21}\sqrt{\sigma_{11}\ \sigma_{22}}}{\sigma_{22} - \sigma_{11}} = \frac{2r_{21}}{\sqrt{\frac{\sigma_{22}}{\sigma_{11}}} - \sqrt{\frac{\sigma_{11}}{\sigma_{22}}}} \tag{81}$$

Clearly α is dependent upon the units chosen for σ_{11} and σ_{22} except in the obvious case of $\alpha = 0$; i.e., an erect ellipse.