FIRST-ORDER R MATRIX FORMALISM FOR TRANSPORT

Section II

Beam transport optics may be reduced to a process of matrix multiplication(1,2). To first-order, this is represented by the matrix equation (using the notation of SLAC-75).

$$x_i(t) = \sum_{j=1}^{6} R_{ij}x_j(0) \qquad (1)$$

where

$$x_1=x \quad x_2=\theta \quad x_3=y \quad x_4=\varphi \quad x_5=\ell \quad \text{and} \quad x_6=\delta$$

The determinant $|R| = 1$. This is a direct consequence of the basic equation of motion for a charged particle in a static magnetic field and is a manifestation of Liouville's theorem of conservation of phase space volume. (See SLAC-75, page 41 for a proof that $|R| = 1$.)

For static magnetic systems possessing midplane symmetry, the six simultaneous linear equations represented by Eq. (1) may be expanded in matrix form as follows:

$$\begin{bmatrix} x(t) \\ \theta(t) \\ y(t) \\ \varphi(t) \\ \ell(t) \\ \delta(t) \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} & 0 & 0 & 0 & R_{16} \\ R_{21} & R_{22} & 0 & 0 & 0 & R_{26} \\ 0 & 0 & R_{33} & R_{34} & 0 & 0 \\ 0 & 0 & R_{43} & R_{44} & 0 & 0 \\ R_{51} & R_{52} & 0 & 0 & 1 & R_{56} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ \theta_0 \\ y_0 \\ \varphi_0 \\ \ell_0 \\ \delta_0 \end{bmatrix} \qquad (2)$$

where the transformation is from an initial position $\tau = 0$ to a final position $\tau = t$ measured along the assumed central reference trajectory.

Thus at any specified position in a system, an arbitrary charged particle is represented by a vector (single column matrix), X, whose components are the positions, angles, and momentum of the particle with respect to a specified reference trajectory.

$$\text{i.e.} \quad X = \begin{bmatrix} x \\ \theta \\ y \\ \varphi \\ \ell \\ \delta \end{bmatrix}$$

where:

x = the radial displacement of the arbitrary ray with respect to the assumed central trajectory.

θ = the angle this ray makes in the radial plane with respect to the assumed central trajectory.

y = the transverse displacement of the ray with respect to the assumed central trajectory.

φ = the transverse angle of the ray with respect to the assumed central trajectory.

ℓ = the path length difference between the arbitrary ray and the central trajectory.

$\delta = \Delta P/P$ is the fractional momentum deviation of the ray from the assumed central trajectory.

The magnetic lens is represented by the square matrix, R, which describes the action of the magnet on the particle coordinates. Thus the passage of a charged particle through the system may be represented by the matrix equation:

$$X(1) = R\ X(0) \tag{3}$$

where X(0) is the initial coordinate vector and X(1) is the final coordinate vector of the particle under consideration; R is the transformation matrix for all such particles traversing the system (one particle differing from another only by its initial coordinate vector X(0)).

The traversing of several magnets and interspersing drift spaces is described by the same basic equation but with R now being the product matrix R = R(n)... R(3)R(2)R(1) of the individual matrices of the system elements. TRANSPORT calculates and tabulates the product matrix R representing the system.

The zero elements $R_{13} = R_{14} = R_{23} = R_{24} = R_{31} = R_{32} = R_{41} = R_{42} = R_{36} = R_{46} = 0$ in the R matrix are a direct consequence of midplane symmetry. If midplane symmetry is destroyed, these elements will in general become non-zero. The zero elements in column five occur because the variables x, θ, y, φ, and δ are independent of the path length difference ℓ. The zero's in row six result from the fact that we have restricted the problem to static magnetic fields, i.e., the scalar momentum is a constant of the motion.

In SLAC report 75 (Ref. 1), a physical significance has been attached to the non-zero matrix elements in the first four rows in terms of their identification with characteristic first-order trajectories. We include figures showing these characteristic functions as a convenient reference.

We now wish to relate the elements appearing in column six and those in row five in terms of simple integrals of the characteristic first-order matrix elements $c_x(t) = R_{11}$ and $s_x(t) = R_{12}$. In order to do this, we make use of the Green's function integral, Eq. (43), Section II of SLAC-75, and of the expression for the differential path length in curvilinear coordinates

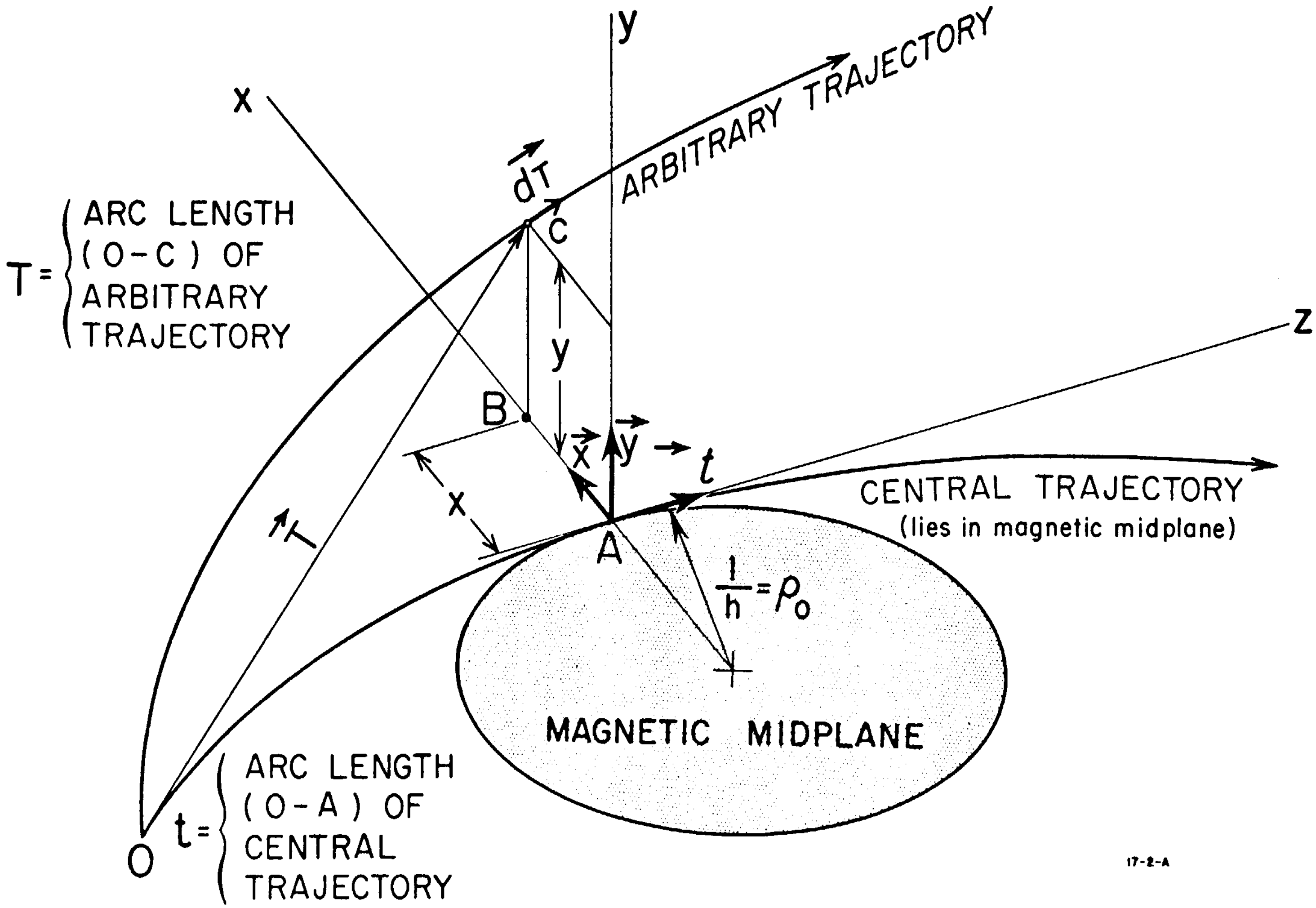

FIG. 1--CURVILINEAR COORDINATE SYSTEM USED IN DERIVATION OF EQUATIONS OF MOTION.

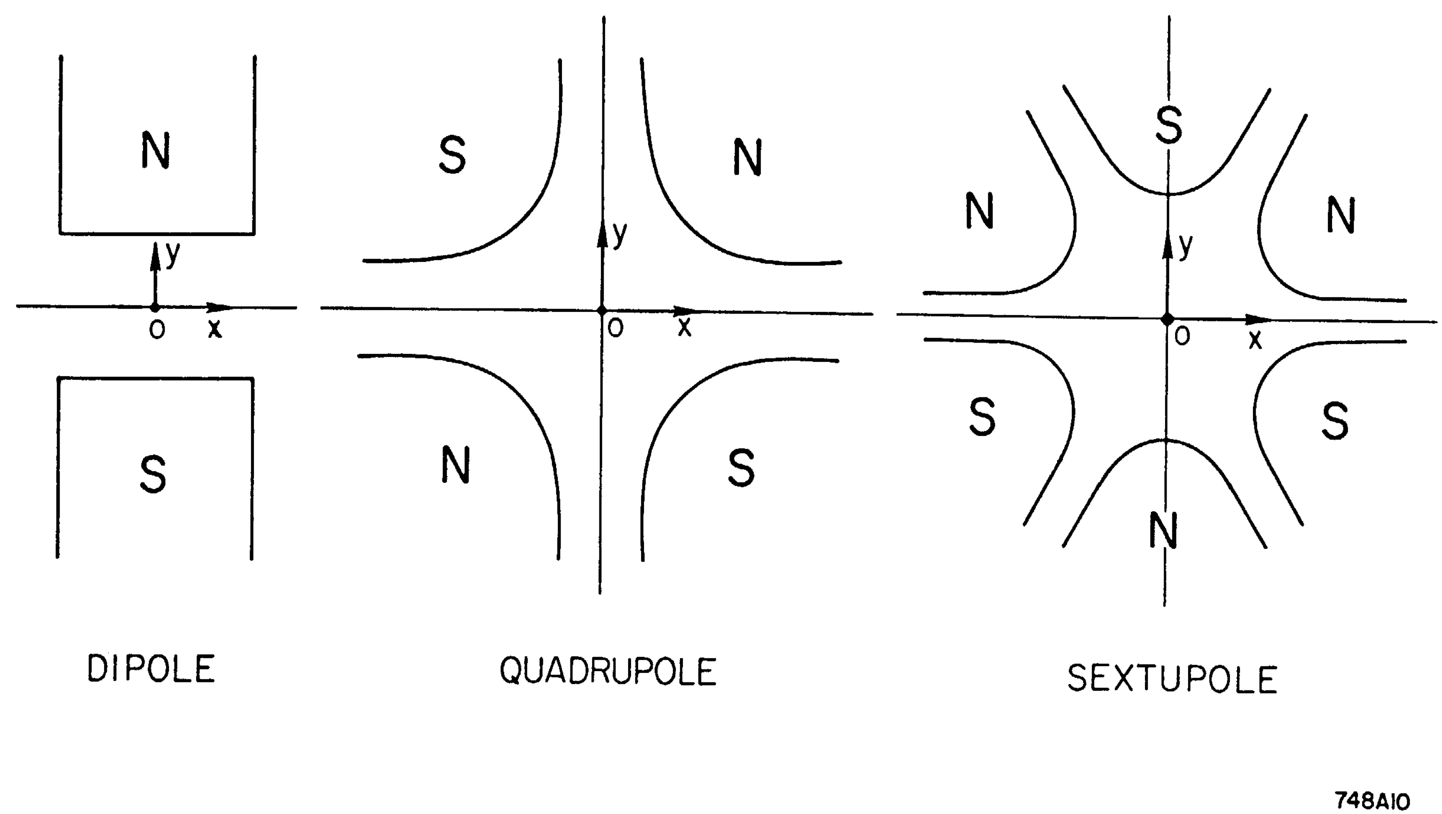

FIG. 2--ILLUSTRATION OF THE MAGNETIC MIDPLANE (X AXIS) FOR DIPOLE, QUADRUPOLE AND SEXTUPOLE ELEMENTS. THE MAGNETIC POLARITIES INDICATE MULTIPOLE ELEMENTS THAT ARE POSITIVE IN RESPECT TO EACH OTHER.

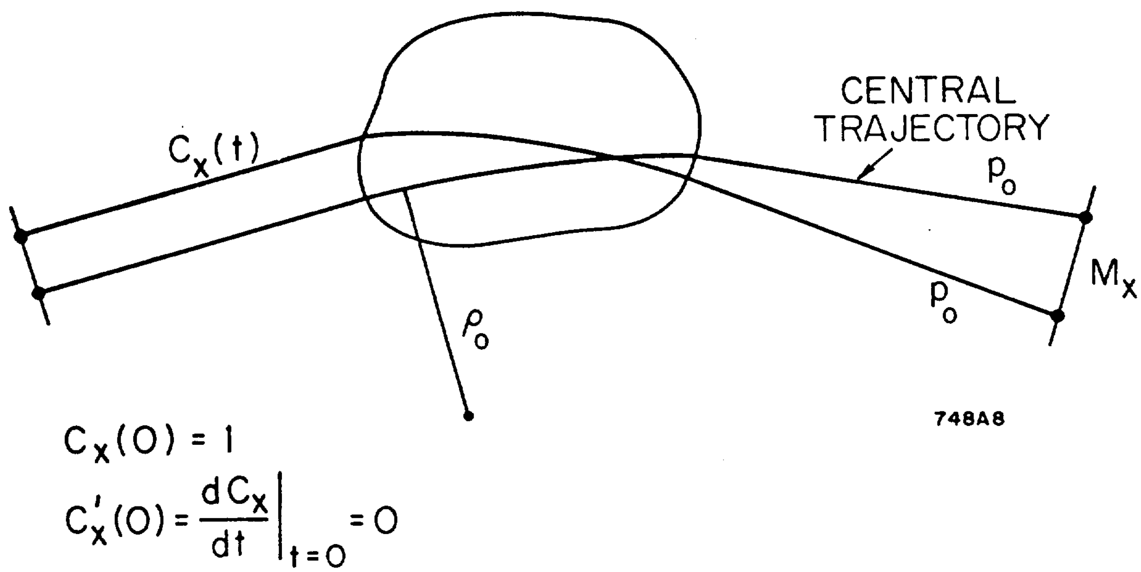

FIG. 3--COSINE-LIKE FUNCTION $c_x(t) = R_{11}$ IN MAGNETIC MIDPLANE. $c_x'(t) = R_{21}$.

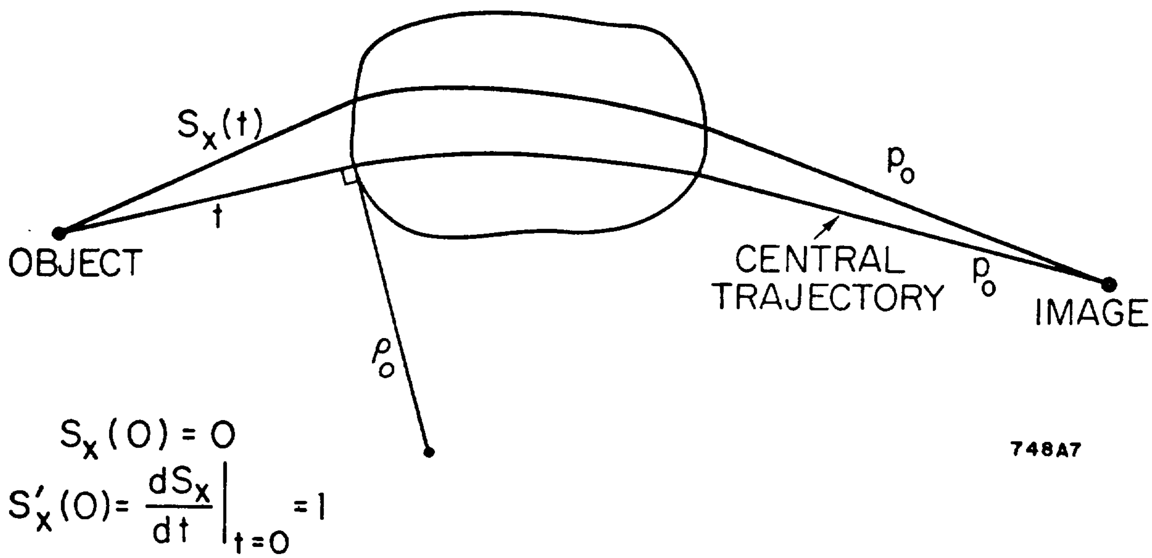

FIG. 4--SINE-LIKE FUNCTION $s_x(t) = R_{12}$ IN MAGNETIC MIDPLANE. $s'_x(t) = R_{22}$.

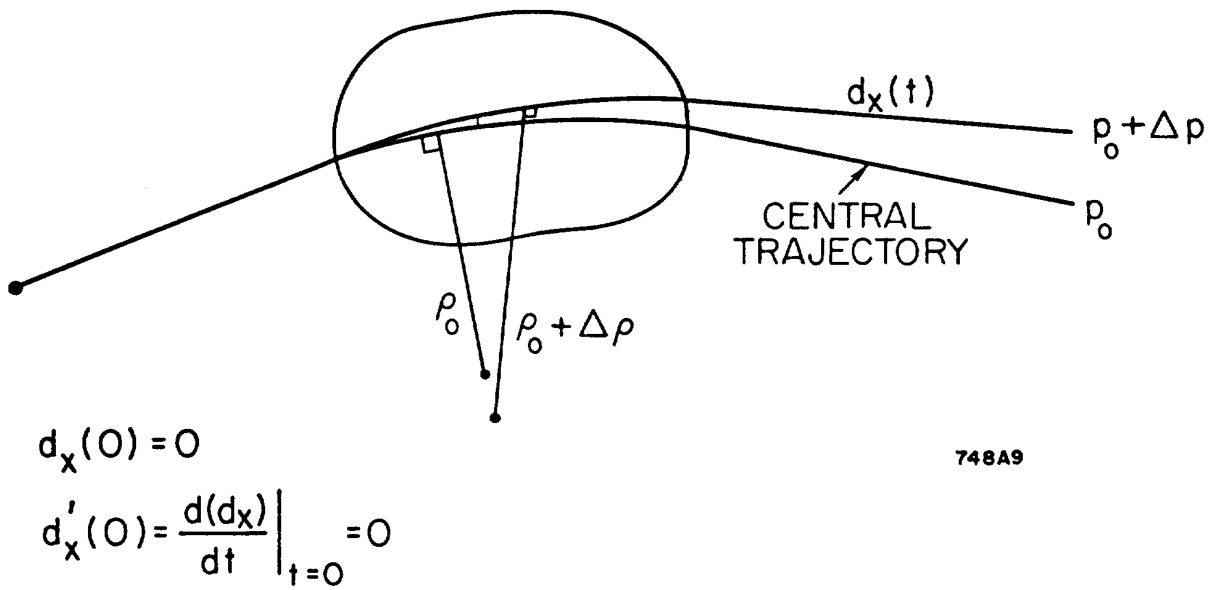

FIG. 5--DISPERSION FUNCTION $d_x(t) = R_{16}$ IN MAGNETIC MIDPLANE. $d'_x(t) = R_{26}$.

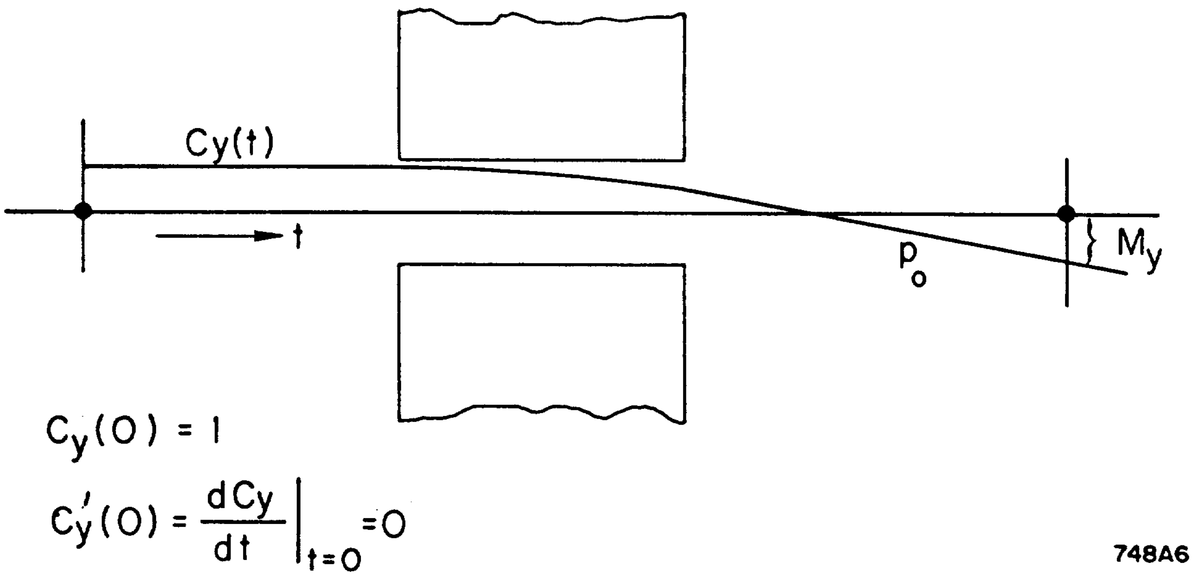

FIG. 6--COSINE-LIKE FUNCTION $c_y(t) = R_{33}$ IN THE NON-BEND (y) PLANE. $c_y'(t) = R_{43}$.

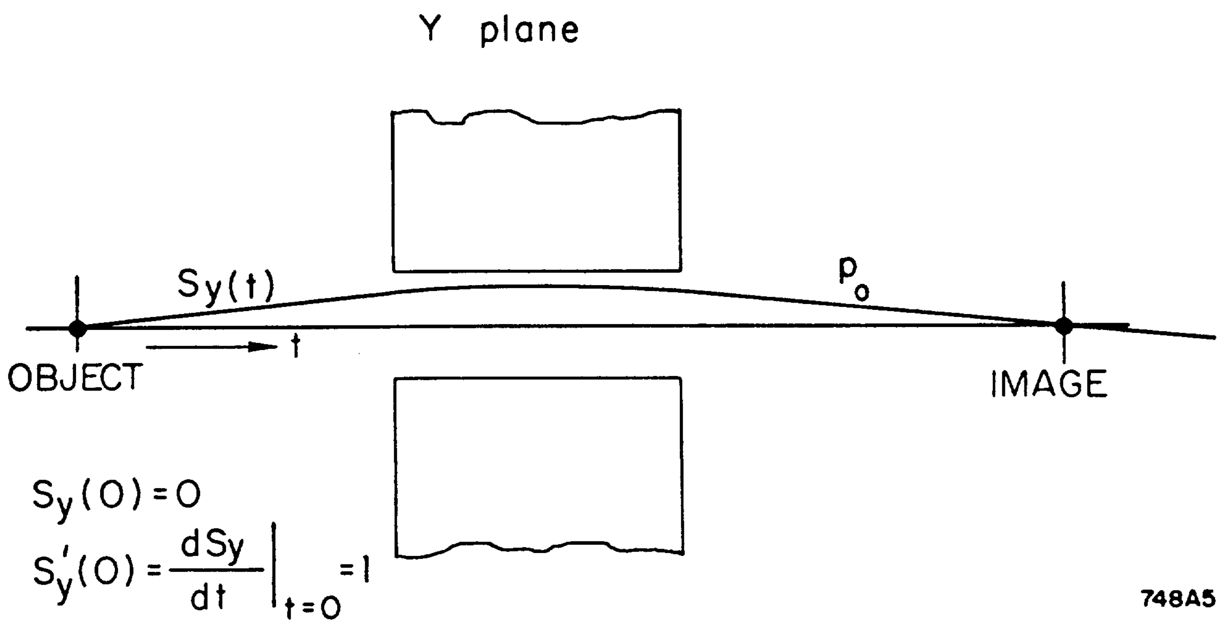

FIG. 7--SINE-LIKE FUNCTION $s_y(t) = R_{34}$ IN NON-BEND (y) PLANE. $s_y'(t) = R_{44}$.

$$dT = \left[(dx)^2 + (dy)^2 + (1+hx)^2 \, (dt)^2 \right]^{1/2} \tag{4}$$

used in the derivation of the equation of motion.

First-Order Dispersion

The spatial dispersion $d_x(t) = R_{16}$ of a system at position t is derived using the Green's function integral, and the driving term $f(\tau) = h(\tau) = \frac{1}{\rho_o(t)}$ for the dispersion (see Table I of SLAC-75). The result is

$$d_x(t) = R_{16} = s_x(t) \int_o^t c_x(\tau) \, h(\tau) d\tau - c_x(t) \int_o^t s_x(\tau) \, h(\tau) d\tau \tag{5}$$

where τ is the variable of integration. Note that $h(\tau)d\tau = d\alpha$ is the differential angle of bend of the central trajectory at any point in the system. Thus first-order dispersion is generated only in regions where the central trajectory is deflected (i.e., in dipole elements.) The angular dispersion is obtained by direct differentiation of $d_x(t)$ with respect to t, the curvilinear distance along the central trajectory.

$$d_x'(t) = R_{26} = s_x'(t) \int_o^t c_x(\tau) \, h(\tau) d\tau - c_x'(t) \int_o^t s_x(\tau) \, h(\tau) d\tau \tag{6}$$

where

$$c_x'(t) = R_{21} \text{ and } s_x'(t) = R_{22}$$

First-Order Path Length

The first-order path length difference is obtained by expanding and integrating Eq. (4) and retaining only the first-order term, i.e.,

$$\ell - \ell_o = (T - t) = \int_o^t x(\tau) \, h(\tau) d\tau + \text{higher order terms}$$

from which

$$\ell = x_o \int_o^t c_x(\tau)\ h(\tau)d\tau + \theta_o \int_o^t s_x(\tau)\ h(\tau)d\tau + \ell_o + \delta \int_o^t d_x(\tau)\ h(\tau)d\tau$$

$$= R_{51}\ x_o + R_{52}\theta_o + \ell_o + R_{56}\delta \qquad (7)$$

Inspection of Eqs. (5), (6), and (7) yields the following useful theorems:

Achromaticity: A system is defined as being achromatic if $R_{16} = R_{26} = 0$, i.e., if $d_x(t) = d_x'(t) = 0$. Therefore it follows from Eq's. (5) and (6) that the necessary and sufficient conditions for achromaticity are that

$$\int_o^t s_x(\tau)\ h(\tau)d\tau = \int_o^t c_x(\tau)\ h(\tau)d\tau = 0 \qquad (8)$$

By comparing Eq. (7) with Eq. (8), we note that if a system is achromatic, all particles of the same momentum will have equal (first-order) path lengths through the system.

Isochronicity: It is somewhat unfortunate that this word has been used in the literature to mean equal path lengths since equal path lengths only imply equal transit times for highly relativistic particles. Nevertheless, from Eq. 7, the necessary and sufficient conditions that the first-order path length of all particles (independent of their initial momenta) will be the same through a system are that $R_{51} = R_{52} = R_{56} = 0$, i.e., if

$$\int_o^t c_x(\tau)\ h(\tau)d\tau = \int_o^t s_x(\tau)\ h(\tau)\ d\tau = \int_o^t d_x(\tau)\ h(\tau)d\tau = 0 \qquad (9)$$

First-Order Imaging

First-order point-to-point imaging in the x plane occurs when $x(t)$ is independent of the initial angle θ_o. This can only be so when

$$s_x(t) = R_{12} = 0. \qquad (10)$$

Similarly first-order point-to-point imaging occurs in the y plane when

$$s_y(t) = R_{34} = 0. \qquad (11)$$

First-order parallel-to-point imaging occurs in the x plane when x(t) is independent of the initial particle position x_o. This will occur only if

$$c_x(t) = R_{11} = 0. \qquad (12)$$

and correspondingly in the y plane, parallel-to-point imaging occurs when

$$c_y(t) = R_{33} = 0. \qquad (13)$$

A parallel ray entering a system exits parallel to the central trajectory if

$$c_x'(t) = R_{21} = 0. \qquad (14)$$

in the x plane; and if

$$c_y'(t) = R_{43} = 0. \qquad (15)$$

in the y plane.

Point-to-parallel imaging occurs in the x plane if

$$s_x'(t) = R_{22} = 0. \qquad (16)$$

and in the y plane if

$$s_y'(t) = R_{44} = 0. \qquad (17)$$

Focal Lengths

A simple ray diagram of a "thick" lens demonstrates that R_{21} and R_{43} have the following physical interpretations

$$c_x'(t) = R_{21} = -\frac{1}{f_x} \text{ and } c_y'(t) = R_{43} = -\frac{1}{f_y} \qquad (18)$$

where f_x and f_y are the system focal lengths in the x and y planes respectively.

Zero Dispersion

For point-to-point imaging, using Eq's. (5) and (10), the necessary and sufficient condition for zero dispersion at an image is

$$d_x(t) = R_{16} = \int_0^t s_x(\tau)\ h(\tau)d\tau = 0 \tag{19}$$

For parallel-to-point imaging, (i.e., $c_x(t) = 0$), the condition for zero dispersion at the image is

$$d_x(t) = R_{16} = \int_0^t c_x(\tau)\ h(\tau)d\tau = 0. \tag{20}$$

Magnification

For monoenergetic point-to-point imaging in the x-plane, the magnification is given by

$$M_x = \frac{x(t)}{x_o} = R_{11} = c_x(t)$$

and in the y plane by

$$M_y = R_{33} = c_y(t) \tag{21}$$

where a negative number means an inverted image.

First-Order Momentum Resolution

For point-to-point imaging the first-order momentum resolving power R_1 (not to be confused with the matrix R) is the ratio of the momentum dispersion to the total image size. Thus if $2x_o$ is the total source size then

$$R_1 = \frac{P}{\Delta P} = \left| \frac{R_{16}}{2x_o R_{11}} \right| = \left| \frac{d_x(t)}{2x_o c_x(t)} \right|$$

For point-to-point imaging $s_x(t) = 0$. Using Eq. (5), the dispersion at an image is

$$d_x(t) = - c_x(t) \int_0^t s_x(\tau)\, h(\tau) d\tau \qquad (22)$$

from which the first-order momentum resolving power R_1 becomes

$$2x_o R_1 = \left| \frac{d_x(t)}{c_x(t)} \right| = \left| \int_0^t s_x(\tau)\, h(\tau) d\tau \right| = \left| R_{52} \right| \qquad (23)$$

Equation (23) for the first-order resolving power of a system may be expressed in a number of useful forms. If we consider a ray (particle) originating at the source with $x_o = 0$ and $\delta = \frac{\Delta P}{P} = 0$ and lying in the midplane. (i.e., a mono-energetic point source), the first-order equation representing the midplane displacement, x, of this trajectory is

$$x(t) = s_x(t)\ \theta_o \qquad (24)$$

We may then rewrite Equation (23) as follows:

$$2x_o R_1 = \int_0^t s_x(\tau)\, h(\tau) d\tau = \frac{1}{\theta_o} \int_0^t x(\tau)\, h(\tau) d\tau = \frac{(\ell - \ell_o)}{\theta_o} = \left| R_{52} \right| \qquad (25)$$

or we may also write it in the form

$$R_1 = \frac{1}{2x_o \theta_o} \int_0^t B \frac{x(t) d\tau}{B \rho_o} = \left(\frac{1}{2x_o \theta_o} \right) \left(\frac{1}{B \rho_o} \right) \int_0 BdA \qquad (26)$$

where $\int BdA$ is the magnetic flux inclosed between the central trajectory and the ray represented by Eq. (24), and $B\rho$ is the magnetic rigidity (momentum) of the central trajectory. Please note, however, that if the ray crosses the central trajectory or the sign of B changes, this changes the sign of the integration.

Some important observations may be made from Eq's. (25) and (26).

1) Resolving particles of different momentum requires that a path length difference must exist between the central trajectory and the trajectory defined by Eq. (22). The greater the path length difference, the greater the resolving power.

2) From Eq. (24), we may define resolving power as the magnetic flux inclosed per unit phase space area per unit momentum ($B\rho$) of the central ray.