Nuclear Instruments and Methods in Physics Research A238 (1985) 127–140
North-Holland, Amsterdam

RELATIONS BETWEEN ELEMENTS OF TRANSFER MATRICES DUE TO THE CONDITION OF SYMPLECTICITY

H. WOLLNIK and M. BERZ

II. Physikalisches Institut der Universität Giessen, D-6300 Giessen, Fed. Rep. Germany

Received 30 January 1985

Relations are derived between elements of transfer matrices. These relations result from the fact that the motion of charged particles from one profile plane to another can be described as a canonical transformation. The derived four first order relations are already known. Similarly, the higher order relations are very useful to check results of numerical ion optical calculations.

1. Introduction

The motion of charged particles in space and time can be described relative to a reference particle in canonically conjugate variables (r_1, r_2, r_3, r_4, r_5, r_6) describing a point in a six-dimensional space as function of an independent variable like, for instance, the time t. Assuming a curvilinear coordinate system, such sets of canonically conjugate variables are

$$\{r_k\} = (r_1, r_2, r_3, r_4, r_5, r_6) = (x, y, \Delta z, p_x, p_y, \Delta p_z), \tag{1a}$$

$$\{r_k\} = (r_1, r_2, r_3, r_4, r_5, r_6) = (x, y, -\Delta s, p_x, p_y, \Delta p), \tag{1b}$$

$$\{r_k\} = (r_1, r_2, r_3, r_4, r_5, r_6) = (x, y, -\Delta t, p_x, p_y, \Delta E), \tag{1c}$$

with $k = 1, 2, 3, 4, 5, 6$. Here s describes the path coordinate, t the time and E the total particle energy while x, y, z describe the position and p_x, p_y, p_z the corresponding components of the particle momentum p. For eq. (1a) the independent variable advantageously is the time t or the path coordinate s, for eq. (1b) it is the time t or the coordinate z along the optic axis and for eq. (1c) it is the path coordinate s or the z coordinate along the optic axis.

Knowing the position r_{k0} at a time t_0 or position z_0 or s_0, respectively, the objective of ion optical calculations [1–3] is to calculate a set of coordinates $\{r_{k1}\}$ describing the position of a particle in phase space at a time t_1 or position z_1 or s_1, respectively. The relation between these two sets of canonically conjugate variables can be expanded in a Taylor's series and written as

$$r_{i1} = \sum_{j=1}^{6} r_{j0}\left\{(r_i|r_j) + \tfrac{1}{2}\sum_{k=1}^{6} r_{k0}\left\{(r_i|r_jr_k) + \tfrac{1}{3}\sum_{l=1}^{6} r_{l0}\left\{(r_i|r_jr_kr_l) + \cdots\right\}\right\}\right\}, \tag{2}$$

with r_i, r_j, r_k, $r_l \cdots$ as defined in eqs. (1). The coefficients

$$(r_i|r_j) = \left(\frac{\partial r_{i1}}{\partial r_{j0}}\right)_{r_{j0}=0}$$

determine the first order properties of the optical system connecting r_{i1} and r_{j0} while the coefficients

$$(r_i|r_jr_k) = \left(\frac{\partial^2 r_{i1}}{\partial r_{jo}\partial r_{k0}}\right)_{r_{j0}=r_{k0}=0}, \qquad (r_i|r_jr_kr_l) = \left(\frac{\partial^3 r_{j1}}{\partial r_{j0}\partial r_{k0}\partial r_{l0}}\right)_{r_{j0}=r_{k0}=r_{l0}=0},$$

0168-9002/85/$03.30 © Elsevier Science Publishers B.V.
(North-Holland Physics Publishing Division)

etc. describe the higher order aberrations. Note that $(r_i|r_jr_k)$ is identical to $(r_i|r_kr_j)$ so that for $j \neq k$ the term $(r_i|r_jr_k)$ describes half of the total aberration coefficient proportional to $r_{j0}r_{k0}$.

2. The condition of symplecticity

Knowing eq. (2) one can compute the 6∗6 Jacobi matrix **A** with 36 elements:

$$w_{ij} = \left(\frac{\partial r_{i1}}{\partial r_{j0}}\right). \tag{3}$$

Here i denotes the row number and j the column number of a specific matrix element where i and j both take values from 1 to 6. Note here that differently than for eq. (2) the partial derivatives of eq. (3) are taken not only at $r_{j0} = 0$.

For the case that both r_i and r_j represent a set of canonically conjugate variables, it has been shown, for instance in ref. [4], that the above Jacobi matrix **A** fulfills the so-called symplectic condition

$$\mathbf{A}^{\mathrm{T}}\mathbf{J}\mathbf{A} = \mathbf{J}, \tag{4}$$

where $\mathbf{A}^{\mathrm{T}}$ is the transpose of **A** and where

$$\mathbf{J} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \end{bmatrix}.$$

Denoting the elements of the matrix $(\mathbf{A}^{\mathrm{T}}\mathbf{J}\mathbf{A})$ by v_{ij}, this condition can also be written as

$$v_{ij} = 0, \qquad \text{except for} \qquad v_{14} = v_{25} = v_{36} = 1, \tag{5a}$$

with $i < j$ because of

$$v_{ij} = -v_{ji}. \tag{5b}$$

Carrying out the matrix multiplication $(\mathbf{A}^{\mathrm{T}}\mathbf{J}\mathbf{A})$ as required by eq. (4) one finds

$$v_{ij} = w_{1i}w_{4j} - w_{4i}w_{1j} + w_{2i}w_{5j} - w_{5i}w_{2j} + w_{3i}w_{6j} - w_{6i}w_{3j}, \tag{6}$$

which also shows that v_{ij} equals $-v_{ji}$. Thus we can restrict ourselves to the case $i < j$ and require only eq. (5a) to be fulfilled.

In order to simplify further calculations, we can rewrite eq. (6) as:

$$v_{ij} = \sum_{\mu=1}^{3} \left(w_{\mu,i}w_{\mu+3,j} - w_{\mu+3,i}w_{\mu,j}\right). \tag{7}$$

The w_{ij} of eqs. (3), (6) now can be obtained by differentiating in eq. (2) the Taylor's series of partial derivatives with respect to r_{j0}:

$$w_{ij} = (r_i|r_j) + \sum_{k=1}^{6} r_{k0}\left\{(r_i|r_jr_k) + \tfrac{1}{2}\sum_{l=1}^{6} r_{l0}\left\{(r_i|r_jr_kr_l) + \cdots\right\}\right\}. \tag{8}$$

Introducing these w_{ij} into eq. (7) one finds after a straightforward substitution

$$v_{ij} = \sum_{\mu=1}^{3}\left\{F_{1,\mu} + \sum_{k=1}^{6} r_{k0}\left\{F_{2,\mu} + \sum_{l=1}^{6} r_{l0}\left\{F_{3,\mu} + \sum_{m=1}^{6} r_{m0}\left\{F_{4,\mu} + \cdots\right\}\right\}\right\}\right\}. \tag{9}$$

where the $F_{\zeta,\mu}$ are given as:

$$F_{\zeta,\mu}=\frac{1}{(\zeta-1)}\Bigg[\binom{\zeta-1}{0}D_\mu(r_i;\ r_k r_l r_m\cdots r_j)$$
$$+\binom{\zeta-1}{1}D_\mu(r_i r_k;\ r_l r_m\cdots r_j)+\cdots+\binom{\zeta-1}{\zeta-1}D_\mu(r_i r_k r_l r_m\cdots;\ r_j)\Bigg],$$
$$D_\mu(M_m;\ M_n)=(r_\mu\,|\,M_m)(r_{\mu+3}\,|\,M_n)-(r_{\mu+3}\,|\,M_m)(r_\mu\,|\,M_n), \tag{10}$$

with $1\le i<j\le 6$ and $k,\ l,\ m,\ n,\ \cdots\in\{1,\ldots,6\}$ and M_m and M_n being monomials of the phase space coordinates $r_1,\ r_2,\ r_3,\ r_4,\ r_5,\ r_6$:

$$M_m=\prod_{\zeta=1}^{6} r_\zeta^{m_\zeta}=r_1^{m_1}r_2^{m_2}r_3^{m_3}r_4^{m_4}r_5^{m_5}r_6^{m_6},$$
$$M_n=\prod_{\zeta=1}^{6} r_\zeta^{n_\zeta}=r_1^{n_1}r_2^{n_2}r_3^{n_3}r_4^{n_4}r_5^{n_5}r_6^{n_6},$$

where n_ζ as well as m_ζ are nonnegative integers 0, 1, 2, 3, 4, ... Note here that

$$\bar{n}=n_1+n_2+n_3+n_4+n_5+n_6\le\zeta,$$
$$\bar{m}=m_1+m_2+m_3+m_4+m_5+m_6\le\zeta,$$

and that very generally

$$D_\mu(M_m;\ M_n)=-D_\mu(M_n;\ M_m).$$

Because of $v_{ij}=-v_{ji}$ and $1\le i<j\le 6$, eq. (5a) represents 15 relations. Comparing the coefficients for all powers of $r_{k0},\ r_{l0},\ r_{m0},\ \ldots$ on both sides of eq. (9) one finds:

$$\sum_{\mu=1}^{3}F_{1,\mu}=\sum_{\mu=1}^{3}D_\mu(r_i;\ r_j)=\delta_{i,j}, \tag{11a}$$
$$\sum_{\mu=1}^{3}F_{2,\mu}=\sum_{\mu=1}^{3}D_\mu(r_i;\ r_k r_j)+D_\mu(r_i r_k;\ r_j)=0, \tag{11b}$$
$$\sum_{\mu=1}^{3}F_{3,\mu}=\sum_{\mu=1}^{3}D_\mu(r_i;\ r_k r_l r_j)+2D_\mu(r_i r_k;\ r_l r_j)+D_\mu(r_i r_k r_l;\ r_j)=0 \tag{11c}$$
$$\sum_{\mu=1}^{3}F_{4,\mu}=\cdots=0, \tag{11d}$$

etc. with $\delta_{i,j}=1$ for $(i,\ j)=(1,4),\ (2,5),\ (3,6)$ and $\delta_{i,j}=0$ for all other cases. Thus
a) eq. (11a) represents 15 relations between the 1st order coefficients of eq. (2),
b) eq. (11b) represents 6 * 15 relations between the 1st and 2nd order coefficients,
c) eq. (11c) represents 21 * 15 relations between the 1st, 2nd and 3rd order coefficients,
d) eq. (11d) represents 56 * 15 relations between the 1st, 2nd, 3rd and 4th order coefficients, etc.

3. General consequences of the choice of the set of coordinates of eq. (1c)

For ion optical problems we like to choose the set of canonically conjugate variables of eq. (1c) with the independent variable being z, i.e., the coordinate along the optic axis, so that

$$r_1=x,\quad r_4=p_x, \tag{12a}$$
$$r_2=y,\quad r_5=p_y, \tag{12b}$$
$$r_3=\Delta t,\quad r_6=\Delta E. \tag{12c}$$

Here, as in any set of coordinates, some coefficients of eq. (2) always vanish so that finally the exuberantly large number of relations due to eqs. (11) is reduced considerably.

3.1. Time invariant fields

Since x_1, y_1, p_{x1}, p_{y1}, ΔE_1 all do not depend on time explicitly, all partial derivatives of these coordinates vanish. This causes all coefficients of eq. (2) to vanish which contain a $r_3^{n_3}$ with $n_3 \neq 0$ in the right hand side of the bracket, i.e.:

$$(r_i \mid \cdots r_3^{n_3} \cdots) = 0, \tag{13a}$$

or $(r_i \mid M_n[n_3 \neq 0]) = 0$ except $(r_3 \mid r_3) = 1$.

3.2. Energy constancy

For most optical systems the particle energy stays unchanged or varies by the same factor for all particles of a beam. This causes all coefficients of eq. (2) to vanish which contain a "$\Delta E = r_6$" in the left side of the bracket, i.e.:

$$(r_6 \mid \cdots) = 0, \tag{13b}$$

or $(r_6 \mid M_n) = 0$ except $(r_6 \mid r_6) = \eta$, with $\eta = 1$ in the event that no accelerations had occurred at all.

3.3. Midplane symmetry

For most ion optical systems one postulates a plane of symmetry which in case of a sector magnet would be the midplane ($y = 0$) in the magnet air gap. In this case all electromagnetic forces must be symmetric with respect to the plane $y = 0$. Thus one finds for $m_2 + m_5 =$ odd

$$(r_1 \mid \cdots r_2^{m_2} r_5^{m_5} \cdots) = (r_3 \mid \cdots r_2^{m_2} r_5^{m_5} \cdots) = (r_4 \mid \cdots r_2^{m_2} r_5^{m_5} \cdots) = (r_6 \mid \cdots r_2^{m_2} r_5^{m_5} \cdots) = 0, \tag{13c}$$

and for $m_2 + m_5 =$ even

$$(r_2 \mid \cdots r_2^{m_2} r_5^{m_5} \cdots) = (r_5 \mid \cdots r_2^{m_2} r_5^{m_5} \cdots) = 0. \tag{13d}$$

The eqs. (13) can also be written as

$$(r_1 \mid M_n^{\text{odd}}) = (r_3 \mid M_n^{\text{odd}}) = (r_4 \mid M_n^{\text{odd}}) = (r_6 \mid M_n^{\text{odd}}) = 0,$$

and

$$(r_2 \mid M_n^{\text{even}}) = (r_5 \mid M_n^{\text{even}}) = 0.$$

Here M_n^{odd} stands for a M_n with $n_2 + n_5 =$ odd and M_n^{even} stands for a M_n with $n_2 + n_5 =$ even.

4. Consequences of eqs. (13) for the D_μ of eqs. (10)

To get a better overview over the different cases let us discuss the total disappearance of D_1, D_2 and D_3 separately, i.e. the cases in which both terms of each D_μ vanish simultaneously.

4.1. The total disappearance of D_1

From eqs. (13c), (13a) one finds $D_1(M_m;\ M_n) = 0$ for

$$M_m = M_m^{\text{odd}} \quad \text{or} \quad M_n = M_n^{\text{odd}}, \tag{14a}$$

or

$$m_3 \neq 0 \qquad \text{or} \qquad n_3 \neq 0. \tag{14b}$$

Note here that the cases of eqs. (14) are the only ones in which D_1 vanishes for arbitrary optical systems.

4.2. The total disappearance of D_2

From eqs. (13a), (13c), (13d) one finds $D_2(M_m;\ M_n) = 0$ for

$$M_m = M_m^{\text{even}} \qquad \text{or} \qquad M_n = M_n^{\text{even}}, \tag{15a}$$

or

$$m_3 \neq 0 \qquad \text{or} \qquad n_3 \neq 0. \tag{15b}$$

Note here also that the cases of eqs. (15) are the only ones in which D_2 vanishes for arbitrary optical systems.

4.3. The total disappearance of D_3

From eq. (13b) one finds $D_3(M_m;\ M_n) = 0$, except for

$$M_m = r_6 \qquad \text{or} \qquad M_n = r_6, \tag{16}$$

which is the same (see the definition of M_m and M_n under eq. (10)) as stating

$$m_1 = m_2 = m_3 = m_4 = m_5 = 0 \qquad \text{and} \qquad m_6 = 1,$$

or

$$n_1 = n_2 = n_3 = n_4 = n_5 = 0 \qquad \text{and} \qquad n_6 = 1.$$

In case one of the conditions of eq. (16) is fulfilled one finds for all nonvanishing D_3:

$$D_3(M_m;\ r_6) = (r3 \mid M_m), \tag{17a}$$

$$D_3(r_6;\ M_n) = -(r_3 \mid M_n). \tag{17b}$$

According to eq. (13c), however, even the D_3 of eqs. (17) vanish if

$$M_m = M_m^{\text{odd}} \qquad \text{or} \qquad M_n = M_n^{\text{odd}}. \tag{18}$$

The restrictions imposed on the D_1, D_2 and D_3 by eqs. (14)–(18) shall now be used to simplify the relations of eqs. (11).

5. Relations between the first order coefficients of eq. (2)

For $\zeta = 1$ in eq. (9) we find $\bar{n} = \bar{m} = 1$, i.e. five m_i and n_i vanish while one m_i and one n_i equal 1. In this case eq. (9) transforms to eq. (11a).

5.1. The case $j \in \{1,4\}$

The condition $i < j$ here implies $j = 4$ and hence $i \in \{1,2,3\}$. Thus the eqs. (15), (16) yield $D_2 = D_3 = 0$ and we find with eqs. (14) only one nontrivial relation:

$$D_1(r_1;\ r_4) = \delta_{1,4} = 1. \tag{19a}$$

5.2. The case $j \in \{2,5\}$

The condition $i<j$ here implies $j=2$, $i=1$ or $j=5$, $i \in \{1,2,3,4\}$. Thus the eqs. (14), (16) yield $D_1 = D_3 = 0$ and we find with eqs. (15) only one nontrivial relation:

$$D_2(r_2;\ r_5) = \delta_{2,5} = 1. \tag{19b}$$

5.3. The case $j = 3$

From eqs. (14)–(16) we infer that $D_1 = D_2 = D_3 = 0$ for arbitrary $i<j$ resulting in only trivial relations.

5.4. The case $j = 6$

Since in this case r_j is even, the eqs. (15) yield $D_2 = 0$. Furthermore, eq. (17a) implies $D_3(r_i;\ r_6) = (r_3 | r_i)$. For $i = 3$ we thus find the identity $(r_3 | r_3) = 1$ and for $i \neq 3$ with eqs. (14):

$$D_1(r_i;\ r_6) + (r_3 | r_i) = 0, \tag{19c}$$

which yields two relations with $i \in \{1,4\}$.

5.5. The resulting relations between coefficients of first order

Using the expressions of eqs. (12) as abbreviations, the eqs. (19a), (19b) read explicitly

$$(x|x)(p_x|p_x) - (p_x|x)(x|p_x) = 1, \tag{20a}$$

$$(y|y)(p_y|p_y) - (y|p_y)(y|p_y) = 1, \tag{20b}$$

which are the well known results of Liouville's theorem, stating that the determinants of the first order transfer matrices for the x- and y-directions both equal one [5]. The eq. (19c) reads explicitly

$$(x|x)(p_x|\Delta E) - (p_x|x)(x|\Delta E) = (\Delta t|x), \tag{20c}$$

$$(x|p_x)(p_x|\Delta E) - (p_x|p_x)(x|\Delta E) = (\Delta t|p_x), \tag{20d}$$

relations which are well known from ref. [6]. The eqs. (20c), (20d) determine longitudinal deviations $(\Delta t|\cdots)$ as functions of lateral deviations. Note here that there are three longitudinal first order terms $[(\Delta t|x), (\Delta t|p_x), (\Delta t|\Delta E)]$, only one of which must be determined truly independently.

6. Relations between first and second order coefficients of eq. (2)

For $\zeta = 2$ in eq. (9) we find either $\bar{n} = 1$, $\bar{m} = 2$ or $\bar{n} = 2$, $\bar{m} = 1$. In this case eq. (9) transforms to eq. (11b).

6.1. The case $j \in \{1,4\}$

The condition $i<j$ implies $j=4$ and $i \in \{1,2,3\}$. Thus the eqs. (16) yield $D_3 = 0$. For $i = 3$ or $k = 3$ we find only trivial relations. For $i \neq 3$ and $k \neq 3$, however, we find from eqs. (14), (15) in the case of r_i = even, i.e. for $i = 1$ and

$$r_k = \text{even:} \qquad D_1(r_1;\ r_k r_4) + D_1(r_1 r_k;\ r_4) = 0, \tag{21a}$$

which yields three relations with $k \in \{1,4,6\}$ while we find for

r_k = odd: only trivial relations.

In the case of r_i = odd, i.e. for $i = 2$ we find for

$$\begin{aligned} r_k &= \text{even:} && \text{only trivial relations,} \\ r_k &= \text{odd:} && D_1(r_2 r_k;\ r_4) + D_2(r_2;\ r_k r_4) = 0, \end{aligned} \tag{21b}$$

which yields two relations with $k \in \{2,5\}$.

6.2. The case $j \in \{2,5\}$

The condition $i < j$ implies $j = 2$, $i = 1$ or $j = 5$, $i \in \{1,2,3,4\}$. Thus the eqs. (16) yield $D_3 = 0$. For $i = 3$ or $k = 3$ we find only trivial relations. For $i \neq 3$ and $k \neq 3$, however, we find from eqs. (14), (15) in the case of r_i = even, i.e. $i \in \{1,4\}$ and

$$\begin{aligned} r_k &= \text{even:} && \text{only trivial relations,} \\ r_k &= \text{odd:} && D_1(r_i;\ r_k r_j) + D_2(r_i r_k;\ r_j) = 0, \end{aligned} \tag{21c}$$

which yields six relations with $k \in \{2,5\}$. In the case of r_i = odd, i.e. for $i = 2$, we find furthermore for

$$r_k = \text{even:} \quad D_2(r_2;\ r_k r_5) + D_2(r_2 r_k;\ r_5) = 0, \tag{21d}$$

which yields three relations with $k \in \{1,4,6\}$ while we find for

r_k = odd: only trivial relations.

6.3. The case $j = 5$

From eqs. (14)–(16) we infer also here that $D_1 = D_2 = D_3 = 0$ for arbitrary $i < j$ resulting in only trivial relations.

6.4. The case $j = 6$

Eq. (17a) here implies $D_3(r_i r_k;\ r_j) = (r_3 \mid r_i r_k)$ and $D_3(r_i;\ r_k r_j) = 0$. For $i = 3$ or $k = 3$ we find only trivial relations. For $i \neq 3$ and $k \neq 3$, however, we find from eqs. (14), (15), (17) in the case of r_i = even, i.e. for $i \in \{1,4\}$ and

$$r_k = \text{even:} \quad D_1(r_i;\ r_k r_6) + D_1(r_i r_k;\ r_6) = -(r_3 \mid r_i r_k), \tag{21e}$$

which yields six relations with $k \in \{1,4,6\}$ while we find for

r_k = odd: only trivial relations.

In the case of r_i = odd, i.e. for $i \in \{2,5\}$ we find furthermore for

$$\begin{aligned} r_k &= \text{even:} && \text{only trivial relations} \\ r_k &= \text{odd:} && D_2(r_i;\ r_k r_6) + D_1(r_i r_k;\ r_6) = -(r_3 \mid r_i r_k), \end{aligned} \tag{21f}$$

which yields four relations with $k \in \{2,5\}$.

6.5 The resulting equations

Using again the abbreviations of eqs. (12) the independent relations of Eqs. (21a)–(21d) read explicitly

$$(x \mid x)(p_x \mid x p_x) - (p_x \mid x)(x \mid x p_x) + (x \mid xx)(p_x \mid p_x) - (p_x \mid xx)(x \mid p_x) = 0, \tag{22a}$$

$$(x \mid x)(p_x \mid p_x p_x) - (p_x \mid x)(x \mid p_x p_x) + (x \mid x p_x)(p_x \mid p_x) - (p_x \mid x p_x)(x \mid p_x) = 0, \tag{22b}$$

$$(x|x)(p_x|\Delta E p_x)-(p_x|x)(x|\Delta E p_x)+(x|x\Delta E)(p_x|p_x)-(p_x|x\Delta E)(x|p_x)=0, \tag{22c}$$

$$(x|p_x)(p_x|yy)-(p_x|p_x)(x|yy)+(y|p_x y)(p_y|y)-(p_y|p_x y)(y|y)=0, \tag{22d}$$

$$(x|p_x)(p_x|p_y y)-(p_x|p_x)(x|p_y y)+(y|p_x p_y)(p_y|y)-(p_y|p_x p_y)(y|y)=0, \tag{22e}$$

$$(x|x)(p_x|yy)-(p_x|x)(x|yy)+(y|xy)(p_y|y)-(p_y|xy)(y|y)=0, \tag{22f}$$

$$(x|x)(p_x|p_y y)-(p_x|x)(x|p_y y)+(y|xp_y)(p_y|y)-(p_y|xp_y)(y|y)=0, \tag{22g}$$

$$(x|x)(p_x|yp_y)-(p_x|x)(x|yp_y)+(y|xy)(p_y|p_y)-(p_y|xy)(y|p_y)=0, \tag{22h}$$

$$(x|x)(p_x|p_y p_y)-(p_x|x)(x|p_y p_y)+(y|xp_y)(p_y|p_y)-(p_y|xp_y)(y|p_y)=0, \tag{22i}$$

$$(x|p_x)(p_x|yp_y)-(p_x|p_x)(x|yp_y)+(y|p_x y)(p_y|p_y)-(p_y|p_x y)(y|p_y)=0, \tag{22j}$$

$$(x|p_x)(p_x|p_y p_y)-(p_x|p_x)(x|p_y p_y)+(y|p_x p_y)(p_y|p_y)-(p_y|p_x p_y)(y|p_y)=0, \tag{22k}$$

$$(y|y)(p_y|xp_y)-(p_y|y)(y|xp_y)+(y|yx)(p_y|p_y)-(p_y|yx)(y|p_y)=0, \tag{22l}$$

$$(y|y)(p_y|p_x p_y)-(p_y|y)(y|p_x p_y)+(y|yp_x)(p_y|p_y)-(p_y|yp_x)(y|p_y)=0, \tag{22m}$$

$$(y|y)(p_y|\Delta E p_y)-(p_y|y)(y|\Delta E p_y)+(y|y\Delta E)(p_y|p_y)-(p_y|y\Delta E)(y|p_y)=0, \tag{22n}$$

where eqs. (22l), (22m) carry no new information and are mere combinations of eqs. (22g), (22h) and eqs. (22e), (22j), respectively. Explicitly the relations of eqs. (21e)–(21f) read

$$(x|x)(p_x|x\Delta E)-(p_x|x)(x|x\Delta E)+(x|xx)(p_x|\Delta E)-(p_x|xx)(x|\Delta E) = (\Delta t|xx), \tag{23a}$$

$$(x|x)(p_x|p_x\Delta E)-(p_x|x)(x|p_x\Delta E)+(x|xp_x)(p_x|\Delta E)-(p_x|xp_x)(x|\Delta E) = (\Delta t|xp_x), \tag{23b}$$

$$(x|x)(p_x|\Delta E\Delta E)-(p_x|x)(x|\Delta E\Delta E)+(x|x\Delta E)(p_x|\Delta E)-(p_x|x\Delta E)(x|\Delta E) = (\Delta t|x\Delta E), \tag{23c}$$

$$(x|p_x)(p_x|x\Delta E)-(p_x|p_x)(x|x\Delta E)+(x|p_x x)(p_x|\Delta E)-(p_x|p_x x)(x|\Delta E) = (\Delta t|p_x x), \tag{23d}$$

$$(x|p_x)(p_x|p_x\Delta E)-(p_x|p_x)(x|p_x\Delta E)+(x|p_x p_x)(p_x|\Delta E)-(p_x|p_x p_x)(x|\Delta E) = (\Delta t|p_x p_x), \tag{23e}$$

$$(x|p_x)(p_x|\Delta E\Delta E)-(p_x|p_x)(x|\Delta E\Delta E)+(x|p_x\Delta E)(p_x|\Delta E)-(p_x|p_x\Delta E)(x|\Delta E) = (\Delta t|p_x\Delta E), \tag{23f}$$

$$(y|y)(p_y|y\Delta E)-(p_y|y)(y|y\Delta E)+(x|yy)(p_x|\Delta E)-(p_x|yy)(x|\Delta E) = (\Delta t|yy), \tag{23g}$$

$$(y|y)(p_y|p_y\Delta E)-(p_y|y)(y|p_y\Delta E)+(x|yp_y)(p_x|\Delta E)-(p_x|yp_y)(x|\Delta E) = (\Delta t|yp_y), \tag{23h}$$

$$(y|p_y)(p_y|y\Delta E)-(p_y|p_y)(y|y\Delta E)+(x|p_y y)(p_x|\Delta E)-(p_x|p_y y)(x|\Delta E) = (\Delta t|p_y y), \tag{23i}$$

$$(y|p_y)(p_y|p_y\Delta E)-(p_y|p_y)(y|p_y\Delta E)+(x|p_y p_y)(p_x|\Delta E)-(p_x|p_y p_y)(x|\Delta E) = (\Delta t|p_y p_y), \tag{23j}$$

where the eqs. (23d), (23i) carry no new information and are mere combinations of eqs. (23b), (22c) and eqs. (23h), (22n), respectively.

Knowing all first order matrix elements the eqs. (22), (23) describe the following relations between the matrix elements of second order:

a) the eqs. (22a)–(22b) represent 2 independent relations between the 6 geometric matrix elements

$$(x|xx), (x|xp_x), (x|p_xp_x), (p_x|xx), (p_x|xp_x), (p_x|p_xp_x)$$

leaving 4 matrix elements to be determined independently;

b) eq. (22c) represents 1 independent relation between the 6 chromatic x-matrix elements

$$(x|x\Delta E), (x|p_x\Delta E), (x|\Delta E\Delta E), (p_x|x\Delta E), (p_x|p_x\Delta E), (p_x|\Delta E\Delta E)$$

leaving 5 matrix elements to be determined independently;

c) eq. (22n) represents 1 independent relation between the 4 chromatic y-matrix elements

$$(y|y\Delta E), (y|p_y\Delta E), (p_y|y\Delta E), (p_y|p_y\Delta E)$$

leaving 3 matrix elements to be determined independently;

d) eqs. (22d)–(22m) represent 8 independent relations which describe all 8 geometric y-matrix elements (note here that eq. (20b) holds)

$$(y|yx), (y|yp_x), (y|p_yx), (y|p_yp_x), (p_y|yx), (p_y|yp_x), (p_y|p_yx), (p_y|p_yp_x)$$

as functions of the 6 geometric x-matrix elements

$$(x|yy), (x|yp_y), (x|p_yp_y), (p_x|yy), (p_x|yp_y), (p_x|p_yp_y)$$

so that none of these y-matrix elements must be determined independently.

e) eqs. (23a)–(23j) represent 7 independent relations which describe all 8 longitudinal matrix elements

$$(\Delta t|xx), (\Delta t|xp_x), (\Delta t|p_xp_x), (\Delta t|x\Delta E), (\Delta t|p_x\Delta E), (\Delta t|yy), (\Delta t|yp_y), (\Delta t|p_yp_y)$$

as functions of the x-matrix elements leaving $(\Delta t|\Delta E\Delta E)$ as the only longitudinal matrix element to be determined independently.

7. Relations between first, second and third order coefficients of eq. (2)

For $\zeta = 3$ in eq. (9) we find either $\bar{n} = 1$, $\bar{m} = 3$ or $\bar{n} = \bar{m} = 2$ or $\bar{n} = 3$, $\bar{m} = 1$. In this case eq. (9) transforms to eq. (11c).

7.1. The case $j \in \{1,4\}$

The condition $i < j$ again implies $j = 4$ and $i \in \{1,2,3\}$. Thus the eqs. (16) yield $D_3 = 0$. For $i = 3$ or $k = 3$ or $l = 3$ we find only trivial relations. For $i \neq 3$, $k \neq 3$ and $l \neq 3$, however, we find from eqs. (14), (15) in the case of r_i = even, i.e., for $i = 1$ and

r_k = even, r_l = even:

$$D_1(r_1;\ r_kr_lr_4) + 2D_1(r_1r_k;\ r_lr_4) + D_1(r_1r_kr_l;\ r_4) = 0, \tag{24a}$$

which yields 9 relations with $k \in \{1,4,6\}$ and $l \in \{1,4,6\}$ while we find for:

r_k = even, r_l = odd: only trivial relations,

r_k = odd, r_l = even: only trivial relations,

r_k = odd, r_l = odd:

$$D_1(r_1;\ r_kr_lr_4) + 2D_2(r_1r_k;\ r_lr_4) + D_1(r_1r_kr_l;\ r_4) = 0, \tag{24b}$$

which yields 4 relations with $k \in \{2,5\}$ and $l \in \{2,5\}$.

In the case of r_i = odd, i.e. for $i = 2$ we find for

r_k = even, r_l = even: only trivial relations,
r_k = even, r_l = odd:

$$D_2(r_2;\ r_k r_l r_4) + 2D_2(r_2 r_k;\ r_l r_4) + D_1(r_2 r_k r_l;\ r_4) = 0, \tag{24c}$$

which yields 6 relations with $k \in \{1,4,6\}$ and $l \in \{2,5\}$ while we find for

r_k = even, r_l = even: only trivial relations,
r_k = odd, r_l = even:

$$D_2(r_2;\ r_k r_l r_4) + 2D_1(r_2 r_k;\ r_l r_4) + D_1(r_2 r_k r_l;\ r_4) = 0, \tag{24d}$$

which yields 6 relations with $k \in \{2,5\}$ and $l \in \{1,4,6\}$.

r_k = odd, r_l = odd: only trivial relations.

7.2. The case $j \in \{2,5\}$

For $i = 3$, $k = 3$ or $l = 3$ we again find only trivial relations. For $i \neq 3$, $k \neq 3$ and $l \neq 3$, however, we obtain from eqs. (14), (15) in the case of r_i = even, i.e. for $i \in \{1,4\}$ and

r_k = even, r_l = even: only trivial relations,
r_k = even, r_l = odd:

$$D_1(r_i;\ r_k r_l r_j) + 2D_1(r_i r_k;\ r_l r_j) + D_2(r_i r_k r_l;\ r_j) = 0, \tag{24e}$$

which yields 18 relations with $k \in \{1,4,6\}$ and $l \in \{2,5\}$ while we find for

r_k = odd, r_l = even:

$$D_1(r_1;\ r_k r_l r_j) + 2D_2(r_i r_k;\ r_l r_j) + D_2(r_i r_k r_l;\ r_j) = 0, \tag{24f}$$

which yields 18 relations with $k \in \{2,5\}$ and $l \in \{1,4,6\}$ and for

r_k = odd, r_l = odd: only trivial relations.

In the case of r_i = odd, i.e. $i = 2$ and $j = 5$ we furthermore find for

r_k = even, r_l = even:

$$D_2(r_2;\ r_k r_l r_5) + 2D_2(r_2 r_k;\ r_l r_5) + D_2(r_2 r_k r_l;\ r_5) = 0, \tag{24g}$$

which yields 9 relations with $k \in \{1,4,6\}$ and $l \in \{1,4,6\}$ while we find for

r_k = even, r_l = odd: only trivial relations,
r_k = odd, r_l = even: only trivial relations,
r_k = odd, r_l = odd:

$$D_2(r_2;\ r_k r_l r_5) + 2D_1(r_2 r_k;\ r_l r_5) + D_2(r_2 r_k r_l;\ r_5) = 0, \tag{24h}$$

which yields 4 relations with $k \in \{2,5\}$ and $l \in \{2,5\}$.

7.3. The case $j = 3$

From eqs. (14)–(16) we infer also here that $D_1 = D_2 = D_3 = 0$ for arbitrary $i < j$ resulting in only trivial relations.

7.4. The case $j = 6$

Eq. (17a) here implies $D_3(r_i r_k r_l;\ r_6) = (r_3 | r_i r_k r_l)$ and $D_3(r_i;\ r_k r_l r_6) = D_3(r_i r_k;\ r_l r_6) = 0$. For $i = 3$ or $k = 3$ or $l = 3$ there are only trivial relations. For $i \neq 3$ and $k \neq 3$ and $l \neq 3$ on the other hand we find in the case of r_i = even; i.e. for $i \in \{1,4\}$ and

r_k = even, r_l = even:

$$D_1(r_i;\ r_k r_l r_6) + 2D_1(r_i r_k;\ r_l r_6) + D_1(r_i r_k r_l;\ r_6) = -(r_3 | r_i r_k r_l), \tag{24i}$$

which yields 18 relations with $k \in \{1,4,6\}$ and $l \in \{1,4,6\}$ while we find for

r_k = even, r_l = odd: only trivial relations,
r_k = odd, r_l = even: only trivial relations,
r_k = odd, r_l = odd:

$$D_1(r_i;\ r_k r_l r_6) + 2D_2(r_i r_k;\ r_l r_6) + D_1(r_i r_k r_l;\ r_6) = -(r_3 | r_i r_k r_l), \tag{24j}$$

which yields 8 relations with $k \in \{2,5\}$ and $l \in \{2,5\}$.

In the case of r_i = odd, i.e. $i \in \{2,5\}$ we furthermore find for

r_k = even, r_l = even: only trivial relations,
r_k = even, r_l = odd:

$$D_2(r_i;\ r_k r_l r_6) + 2D_2(r_i r_k;\ r_l r_6) + D_1(r_i r_k r_l;\ r_6) = -(r_3 | r_i r_k r_l), \tag{24k}$$

which yields 12 relations with $k \in \{1,4,6\}$ and $l \in \{2,5\}$ while we find for

r_k = odd, r_l = even:

$$D_2(r_i;\ r_k r_l r_6) + 2D_1(r_i r_k;\ r_l r_6) + D_1(r_i r_k r_l;\ r_6) = -(r_3 | r_i r_k r_l), \tag{24l}$$

which yields 12 relations again with $k \in \{2,5\}$ and $l \in \{1,4,6\}$.

r_k = odd, r_l = odd: only trivial relations.

Analogously to sects 5.5 and 6.5 the eqs. (24) could now also be written explicitly. Knowing all coefficients of first and second order the eqs. (24) represent

a) 4 relations between the 8 geometric matrix elements

$$(r_i | xxx),\ (r_i | xxp_x),\ (r_i | xp_x p_x),\ (r_i | p_x p_x p_x),$$

with $r_i \in \{x, p_x\}$;

b) 7 relations between the 12 chromatic matrix elements

$$(r_i | xx\Delta E),\ (r_i | xp_x\Delta E),\ (r_i | p_x p_x\Delta E),\ (r_i | x\Delta E\Delta E),\ (r_i | p_x\Delta E\Delta E),\ (r_i | \Delta E\Delta E\Delta E),$$

with $r_i \in \{x, p_x\}$;

c) 3 relations between the 6 chromatic matrix elements

$$(r_i | yx\Delta E),\ (r_i | yp_x\Delta E),\ (r_i | y\Delta E\Delta E),\ (r_i | p_y x\Delta E),\ (r_i | p_y p_x\Delta E),\ (r_i | p_y\Delta E\Delta E),$$

with $r_i \in \{y, p_y\}$;

d) 24 relations between the 24 geometric matrix elements

$$(r_i | yxx),\ (r_i | yxp_x),\ (r_i | yp_x p_x),\ (r_i | yyy),\ (r_i | yyp_y),\ (r_i | yp_y p_y),$$
$$(r_i | p_y xx),\ (r_i | p_y xp_x),\ (r_i | p_y p_x p_x),\ (r_i | p_y yy),\ (r_i | p_y yp_y),\ (r_i | p_y p_y p_y),$$

with $r_i \in \{y, p_y\}$;

e) 50 relations between the 32 longitudinal matrix elements $(\Delta t \mid \cdots)$ leaving only $(\Delta t \mid \Delta E \Delta E \Delta E)$ to be determined independently.

8. The transformation of coordinates

Even though the sets of canonically conjugate variables of eq. (1) are used advantageously to describe the motion of charged particles, they have not been used traditionally but rather the set

$$\bar{q}_1 = x, \qquad \bar{q}_4 = x', \tag{25a}$$
$$\bar{q}_2 = y, \qquad \bar{q}_5 = y', \tag{25b}$$
$$\bar{q}_3 = \Delta t, \qquad \bar{q}_6 = \Delta K/K_0 = k, \tag{25c}$$

with $\Delta t = t - t_0$ and $\Delta K = K - K_0$. Here K_0 is the energy of the reference particle, t_0 is the time it takes this reference particle to traverse the optical system and $x' = \mathrm{d}x/\mathrm{d}t$, $y' = \mathrm{d}y/\mathrm{d}t$ are the slopes in the x- and y-directions, respectively.

In some cases also a slightly different set [7] is used

$$q_1 = x, \qquad q_4 = a = (p/p_0)x'/\sqrt{1 + x'^2 + y'^2}, \tag{26a}$$
$$q_2 = y, \qquad q_5 = b = (p/p_0)y'/\sqrt{1 + x'^2 + y'^2}, \tag{26b}$$
$$q_3 = \Delta t, \qquad q_6 = \Delta E/E = \delta_E, \tag{26c}$$

with eqs. (1c) and (26) we obtain the transformation from this set of coordinates to the canonical coordinates r_i:

$$r_1 = A_1^{qr}(q_1, \dots, q_6) = q_1, \tag{27a}$$
$$r_2 = A_2^{qr}(q_1, \dots, q_6) = q_2, \tag{27b}$$
$$r_3 = A_3^{qr}(q_1, \dots, q_6) = q_3, \tag{27c}$$
$$r_4 = A_4^{qr}(q_1, \dots, q_6) = q_4 p_0, \tag{27d}$$
$$r_5 = A_5^{qr}(q_1, \dots, q_6) = q_5 p_0, \tag{27e}$$
$$r_6 = A_6^{qr}(q_1, \dots, q_6) = q_6 E_0. \tag{27f}$$

Here p_0 stands for the momentum of the reference particle and E_0 for its energy. Similarly one obtains the inverse transformation $q_i = A_i^{qr}$. Assume now that the transformation between profile planes at z_0 and z_1 is given both in symplectic coordinates and those of eqs. (26):

$$k_{r_1} = T_r k_{r0}, \tag{28a}$$
$$k_{q_1} = T_q k_{q0}, \tag{28b}$$

Then we can infer from eqs. (28)

$$T_r = A^{qr} \cdot T_q \cdot A^{rq}, \tag{29a}$$
$$T_q = A^{rq} \cdot T_r \cdot A^{qr}. \tag{29b}$$

The eqs. (29) allow to express the partial derivatives of T_r in terms of partials of A^{qr}, T_q and A^{rq}.

The nonvanishing partials of A^{qr} and A^{rq} are easily found from eqs. (26) as

$$\frac{\partial r_1}{\partial q_1}=\frac{\partial r_2}{\partial q_2}=\frac{\partial r_3}{\partial q_3}=1, \qquad \frac{\partial q_1}{\partial r_1}=\frac{\partial q_2}{\partial r_2}=\frac{\partial q_3}{\partial r_3}=1, \tag{30a}$$

$$\frac{\partial r_4}{\partial q_4}=\frac{\partial r_5}{\partial q_5}=p_0, \qquad \frac{\partial q_4}{\partial r_4}=\frac{\partial q_5}{\partial r_5}=\frac{1}{p_0}, \tag{30b}$$

$$\frac{\partial r_6}{\partial q_6}=E_0, \qquad \frac{\partial q_6}{\partial r_6}=\frac{1}{E_0}. \tag{30c}$$

Consider now the partial derivatives

$$\left(\frac{\partial r_i}{\partial r_k}\right)_{r_{k0}=0}=(r_i|r_k),$$

used in eq. (2). Applying the chain rule twice we obtain

$$(r_i|r_k)=\left(\frac{\partial r_i}{\partial r_k}\right)_{r_{k0}=0}=\sum_{\kappa=1}^{6}\left(\frac{\partial r_{i1}}{\partial q_{\kappa 1}}\right)_{q_{\kappa 1}=0}\left(\frac{\partial q_{\kappa 1}}{\partial r_{k0}}\right)_{r_{k0}=0}$$

$$=\sum_{\kappa=1}^{6}\sum_{\lambda=1}^{6}\left(\frac{\partial r_{i1}}{\partial q_{\kappa 1}}\right)_{q_{\kappa 1}=0}\left(\frac{\partial q_{\kappa 1}}{\partial q_{\lambda 0}}\right)_{q_{\lambda 0}=0}\left(\frac{\partial q_{\lambda 0}}{\partial r_{k0}}\right)_{r_{k0}=0}. \tag{31}$$

The fact that for $\kappa \neq i$ and $\lambda \neq$ k the partial derivatives of A^{qr} and A^{rq} vanish (see eqs. (28), (30)) simplifies eq. (31) to

$$(r_i|r_k)=\left(\frac{\partial r_i}{\partial r_k}\right)_{r_{k0}=0}=\left(\frac{\partial r_{i1}}{\partial q_{i1}}\right)_{q_{i1}=0}(q_i|q_k)\left(\frac{\partial q_{k0}}{\partial r_{k0}}\right)_{r_{k0}=0}. \tag{32a}$$

In a very similar way one obtains the transformation rule for partial derivatives of second and third order:

$$(r_i|r_k r_l)=\left(\frac{\partial^2 r_{i1}}{\partial r_{k0}\partial r_{l0}}\right)_{r_{k0}=r_{l0}=0}=\left(\frac{\partial r_{i1}}{\partial q_{i1}}\right)_{q_{i1}=0}(q_i|q_k q_l)\left(\frac{\partial q_{k0}}{\partial r_{k0}}\right)_{r_{k0}=0}\left(\frac{\partial q_{l0}}{\partial r_{l0}}\right)_{r_{l0}=0}, \tag{32b}$$

$$(r_i|r_k r_l r_m)=\left(\frac{\partial^3 r_{i1}}{\partial r_{k0}\partial r_{l0}\partial r_{m0}}\right)_{r_{k0}=r_{l0}=r_{m0}=0}$$

$$=\left(\frac{\partial r_{i1}}{\partial q_{i1}}\right)_{q_{i1}=0}(q_i|q_k q_l q_m)\left(\frac{\partial q_{k0}}{\partial r_{k0}}\right)_{r_{k0}=0}\left(\frac{\partial q_{l0}}{\partial r_{l0}}\right)_{r_{l0}=0}\left(\frac{\partial q_{m0}}{\partial r_{m0}}\right)_{r_{m0}=0}, \tag{32c}$$

as well as analogous expressions for higher order terms.

With eqs. (32) the relations due to symplecticity can be derived also in the noncanonical coordinates of eq. (26). For this purpose, one only must express all matrix elements $(r_i|r_k)$, $(r_i|r_k r_l)$, $(r_i|r_k r_l r_m)$, $\cdots$ in eqs. (20), (22) etc. in terms of the matrix elements in the noncanonical coordinates $(q_i|q_k)$, $(q_i|q_k q_l)$, $\cdots$ using eqs. (32).

The same pattern as used for the coordinates defined in eqs. (26) can be used for any other set of coordinates, especially those of eqs. (25). However, since in most cases the transformations A^{qr} and A^{rq} are nonlinear, higher order partial derivatives remain in eqs. (30). This usually leads to more complex relations than those in eqs. (32).

For discussions we thank A.J. Dragt, E. Forest and C. Iselin. For financial support we are grateful to the Bundesminister fuer Forschung und Technologie.

References

[1] K.L. Brown, R. Belbeoch and P. Bounin, Rev. Sci. Instr. 35 (1964) 481.
[2] H. Wollnik, Nucl. Instr. and Meth. 52 (1967) 250.
[3] H. Matsuda and H. Wollnik, Nucl. Instr. and Meth. 77 (1970) 283, 103 (1972) 117.
[4] W. Thirring, Lehrbuch der Mathematischen Physik. Vol. 1, Klassische Dynamische Systeme, Wien (1977) and A.J. Dragt, Lectures on Nonlinear Orbit Dynamics, AIP Conf. Proc., vol. 87 (1982).
[5] A.P. Banford, The Transport of Charged Particle Beams (E. and F.N. Spon Ltd., London, 1966).
[6] H. Wollnik and T. Matsuo, Int. J. Mass Spectr. and Ion Phys. 37 (1981) 209.
[7] H. Wollnik, Optics of Charged Particle Beams, in preparation.