Laboratori Nazionali di Frascati

LNF-66/58

M. Bassetti, R. M. Buonanni and M. Placidi : BEAM OPTICS COMPUTATION FOR PARTICLE TRANSPORT SYSTEMS BY MEANS OF AN IMPROVED NEWTON-RAPHSON METHOD.

Estratto da : Nuclear Instr. and Meth. 45, 93 (1966)

NUCLEAR INSTRUMENTS AND METHODS 45 (1966) 93–101; © NORTH-HOLLAND PUBLISHING CO.

BEAM OPTICS COMPUTATION FOR PARTICLE TRANSPORT SYSTEMS BY MEANS OF AN IMPROVED NEWTON-RAPHSON METHOD

M. BASSETTI, R. M. BUONANNI and M. PLACIDI

Laboratori Nazionali di Frascati del CNEN, Roma, Italy

Received 5 April 1966

Computation methods for beam transport systems design are exposed. The IBM 7040 code, by which the beam injection system for the Adone storage ring was designed, is also described. The name of the program is "SYTRAN".

1. Introduction

We will consider a beam transport system, consisting of a succession of bending magnets and quadrupole lenses separated by "drift spaces".

The problem consists in the determination of the sets of parameters fulfilling the channel requirements.

For a given system, the same properties can be obtained through several sets of parameters without changing the essential construction: within the several solutions, the one providing the best compromise between practical requirements is chosen. Each solution is obtained by means of a computer program giving the indications necessary for further selection.

The beam can be represented by points of surfaces limited by closed curves in the horizontal and vertical phase planes. It is convenient to assume an elliptical shape for such curves because in this case the transformation laws along the channel are simple, and moreover an elliptical envelope is a good picture of reality. We will describe only what happens in one phase plane with the assumption that the same is true for the other plane. We will finally assume the particle motions in the two planes to be uncoupled and will use the "rectangular model" in the mathematical treatment of bending magnets and quadrupole lenses.

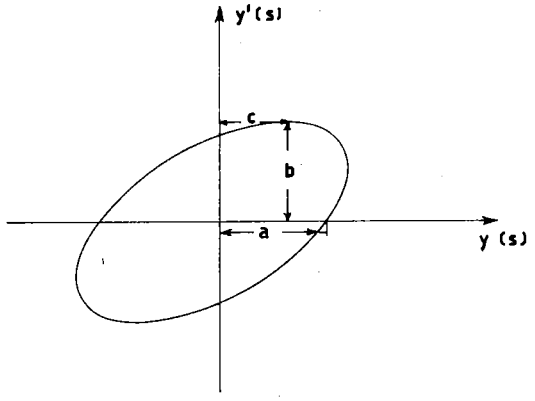


Fig. 1. Beam envelope ellipse in phase plane yy'.

2. Matrix notation and phase plane representation

Let s be the curvilinear coordinate along the optical axis of the system, and $y(s)$ and $y'(s) = \mathrm{d}y(s)/\mathrm{d}s$ the displacement and the slope of a particle with respect to the s-axis. The representative ellipse of the beam equation is then:

$$W = \gamma y^2 + 2\alpha y y' + \beta y'^2 \qquad (\beta\gamma - \alpha^2 = 1). \tag{1}$$

For our purposes it is useful to describe the centered ellipse (1) by means of the three following parameters[1,2]):

$$\begin{aligned} R &= a/b \quad &&\text{"axis ratio"},\\ X &= c/b \quad &&\text{"skew ratio"},\\ W &= ab \quad &&\text{"emittance"}, \end{aligned} \tag{2}$$

with a, b and c defined in fig. 1.

R and X are related to α, β and γ of eq. (1) through the relations:

$$R = 1/\gamma, \qquad X = -\alpha/\gamma.$$

According to Liouville's theorem, W does not change along the channel.

From eqs. (1) and (2) it follows that, for a generic value of s, the maximum values for $y(s)$ and $y'(s)$ are, for a centered ellipse:

$$\begin{aligned} y_{\max}^{(s)} &= (\beta W)^{\frac{1}{2}} = \{W(R + X^2/R)\}^{\frac{1}{2}},\\ y'^{(s)}_{\max} &= (\gamma W)^{\frac{1}{2}} = (W/R)^{\frac{1}{2}}. \end{aligned} \tag{3}$$

To evaluate the effects of the transport system elements upon the phase plane ellipses, the matrix method was used in the computation.

Let us assume for the generical element a 3×3 transfer matrix

$$\mathbf{M} = \begin{vmatrix} a_{11}(s) & a_{12}(s) & a_{13}(s) \\ a_{21}(s) & a_{22}(s) & a_{23}(s) \\ 0 & 0 & 1 \end{vmatrix}. \tag{4}$$

The matrix (4) must be applied to the particle representing vector, whose components are y, y', $\Delta p/p_0$, and whose values at the input and at the output of the element are related by the expression:

$$\begin{vmatrix} y \\ y' \\ \Delta p/p_0 \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} y_0 \\ y'_0 \\ \Delta p/p_0 \end{vmatrix} . \quad (4a)$$

$\Delta p/p_0$ is the fractional momentum difference of the single particle with respect to the central momentum of the channel.

The elements $a_{ik}(s)$ for $(i, k = 1, 2)$ modify the ellipse shape without changing the coordinates of the center; the elements a_{13}, a_{23} (which differ from zero only for the bending magnets) yield a rigid displacement of the ellipse.

We shall use for the matrices of the drift spaces, the bending and the quadrupole lenses, the well known expression (for the rectangular model) that can be found in the literature[3-5]).

By means of some purely algebraic calculations it can be shown that, whenever the particle parameters are transformed according to eq. (4a), the ellipse parameters undergo the following transformations:

$$R(s) = \{(a_{22}^2/R_0)+2a_{21}a_{22}(X_0/R_0)+ +a_{21}^2[R_0+(X_0^2/R_0)]\}^{-1}$$

$$X(s) = R(s)\{(a_{12}a_{22}/R_0)+(1+2a_{12}a_{21})(X_0/R_0)+ +a_{11}a_{21}[R_0+(X_0^2/R_0)]\} \quad (5)$$

and the ellipse center is displaced by a vector Δc whose components are:

$$\Delta y = a_{13}\Delta p/p_0$$
$$\Delta y' = a_{23}\Delta p/p_0, \quad (6)$$

so that the relations (3) are, more generally:

$$y_{\max}^{*(s)} = y_{\max}^{(s)}+|\Delta y|$$
$$y_{\max}^{\prime *(s)} = y_{\max}^{\prime(s)}+|\Delta y'|. \quad (3a)$$

3. Calculation method and use of the program

In a transport system design the requirements the channel must meet at its end or at some intermediate section are:

1. Values of ellipse parameters;
2. Optical properties such as a focus, an image point or first order achromatism;
3. Minimum beam envelope;
4. Maximum resolution in a given section.

From a mathematical point of view, these points yield a set of transcendental equations, for the resolution of which a number of parameters, not less than the number of equations, is needed. The parameters we have found the most convenient are the quadrupole lenses strength and the bending magnets field index, for they have both a considerable influence on the optical features of the channel and an extended range of variability. The solution of the above mentioned non-linear equations system is found by means of a modified Newton-Raphson method which has the following properties:

Let P_0 be a real solution of a certain equation system.

Starting from a point P so that $|P-P_0| < \delta$, where δ is fixed by well defined theorems[6]), the Newton-Raphson method converges to the solution. If $|P-P_0| > \delta$ the method does not converge.

Our modification allows a wider range for δ values so that if the improved method does not reach a real solution one can be sure that the initial point P has been chosen in the neighborhood of complex solution. In such a case the computation goes on until the curve defined by the set of the calculated points crosses the surface defined by the vanishing Jacobian, unless the maximum allowed number of iterations is reached.

The modification we have introduced to the Newton-Raphson method is now explained. Let

$$f_i = 0, \qquad (i = 1, \ldots, J),$$

be the system of equations to be solved, P_j the generic non solution point obtained at the j^{th} iteration, and ΔP_j the increment computated with non-modified N.-R. method, by which it should be

$$P_{j+1} = P_j+\Delta P_j.$$

We introduce at every iteration, a check function F_j defined by

$$F_j = \sum_i [f_i(P_j)]^2$$

and a reduction factor α, less or equal to 1, so that, taking

$$P_{j+1} = P_j+\alpha\Delta P_j \quad (7)$$

it follows

$$F_{j+1} < F_j. \quad (8)$$

In the relation (7) the largest value of $\alpha(\leqq 1)$, compatible with relation (8), is chosen.

The quantity F_j as $j \to \infty$ vanishes in the case of a real solution, otherwise it tends to a non zero value.

To get a solution in the neighborhood of the initial value another small limiting check number h_0 must be

introduced on the absolute value $|\Delta P_j|$: as a consequence the calculation will follow formula (7) until both

$$F_{j+1} < F_j, \tag{8a}$$

$$|\Delta P_j| < h_0, \tag{8b}$$

are verified.

However we are interested only in real solutions and consequently the initial set of the n coordinates of the point P must be chosen following a set of experimental rules we think worthwhile to report:

a. whichever the arrangement of the succession of the quadrupoles, that is doublets or triplets, their strengths should be of alternate sign;

b. the strengths should be chosen so that the focal lengths be of the same order of the distances between the lenses;

c. to match four ellipses (two for each phase plane) at the beginning and at the end of the channel length L, it is convenient to start from a lens system having, over the distance in question, an almost unitary transfer matrix; in this case the quadrupole initial strengths can be easily evaluated in thin lens approximation.

An unitary system, such as the one designed, consists of $2p$ thin lenses ($p \geqq 2$). The effect of a lens doublet and of the corresponding drift spaces is a rotation of each particle by an angle μ in the phase plane; μ verifies the conditions

$$\begin{cases} \mu < \pi \\ p\mu = q\pi, \qquad q = \text{integer} > 0. \end{cases}$$

For $p = 2$ only one value of q is allowed ($q = 1$).

For $p = 3$ two values of q are possible ($q = 1, 2$).

More generally the number of possible q values is $(p-1)$.

The usefulness of a unitary system is due to the fact that it transfers, without changing them (neglecting 2nd order effects), the parameters of the particles and of the ellipses, except for a possible reflection with respect to the origin in the phase plane, over an arbitary space L; the only limitation on the length L arises from the physical dimensions of the lenses.

On the basis of the above mentioned criteria, it is generally possible to match any four ellipses, increasing the q value of the initial unitary system. In the case one has to change a number of elements greater than 4 (number of equations) only four parameters can be varied at time. It is possible to do so because, as stated above, our program calculates and stores the values of the variables which have decreased the initial quantity $\sum f_i^2$ also when a real solution is lacking. Coming back to the matching, if we have six quadrupoles numbered 1-2-3-4-5-6 we will change previously the lenses 1-2-3-4 and afterwards, for example the lenses 2-4-5-6.

Another way is to bind the values for the strengths of certain lenses: this method is suitable mainly when the lens number is a multiple of four and the geometric structure is a periodic one (fig. 2).

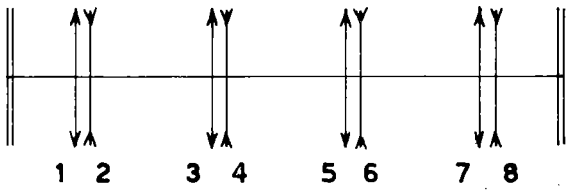


Fig. 2. Scheme of a magnetic channel with eight quadrupole lenses whose strengths $G(m)$ have the following bounds: $G(1) = G(5)$; $G(2) = G(6)$; $G(3) = G(7)$; $G(4) = G(8)$.

The existence of a real solution for a matching problem is a priori the more probable the greater the q value, but it is not convenient to increase this value beyond the minimum indispensable to get a solution, because the field strengths in the focusing elements and the 2nd order chromatic properties of the transport system increase with this parameter.

4. Program description

The channel optical requirements yield a set of transcendantal equations:

$$f_i = 0, \qquad i = 1, 2, \ldots, J \tag{9}$$

and it is in practice important to know the values of the variable parameters for which the functions f_i are smaller than a prefixed value.

We assume the elaboration is finished when for k, being an arbitrarily small assigned value, the following relation is verified

$$\sum_{i=1}^{J} f_i^2 \leqq k_0. \tag{10}$$

Let X_{ij} and Z_{ij} be the horizontal and vertical matrix elements for the part of the system between the initial section and the intermediate one under consideration respectively.

In our program the conditions one may impose (not necessarily all of them) have the following order:

I) $\qquad X_{11} = 0,$

in every system section. Condition I) has the meaning of a focus:

II) $\qquad X_{13} = 0,$

III) $\qquad X_{23} = 0.$

The conditions II) and III) represent first order achromatism conditions for sections of the channel in which are bending magnets [expressions (6)];

IV) $$V-I=0,$$

where V is intended to be the maximum $y^*_{max}(s)$ value, as defined by eq. (3a), in both phase planes, and I is a prefixed value, whose choice will be considered later on. Relation IV) represents the condition for beam envelope minimization.

V) $$2X_{12}/X_{13}-D=0,$$

where D is a number proportional to the desired momentum resolution at a fixed beam section.

The conditions I) and V) together realize the required momentum resolution.

Specific values, O_{ij} and V_{ij} for the total transfer matrix elements TH_{ij}, TV_{ij}, $(i, j = 1, 2)$ can be obtained through the following conditions:

VI) $$TH_{11}-O_{11}=0,$$

VII) $$TH_{12}-O_{12}=0,$$

VIII) $$TH_{21}-O_{21}=0,$$

IX) $$TV_{11}-V_{11}=0,$$

X) $$TV_{12}-V_{12}=0,$$

XI) $$TV_{21}-V_{21}=0.$$

Specific values, ROF, XOF, RVF, XVF, for the final parameters of both phase plane ellipses are reached through the following conditions:

XII) $$ROI-ROF=0,$$

XIII) $$XOI-XOF=0,$$

XIV) $$RVI-RVF=0,$$

XV) $$XVI-XVF=0,$$

where ROI, XOI, RVI and XVI are the R and X values in the two final planes, evaluated by the computer.

Specific values O_{ij} for the chromatic elements of the total transfer matrix can be reached through:

XVI) $$TH_{13}-O_{13}=0,$$

XVII) $$TH_{23}-O_{23}=0.$$

Conditions for first order archromatism can be obtained at the same time in two different beam intermediate sections by means of conditions II) and III), together with the following:

XVIII) $$X_{13}=0,$$

XIX) $$X_{23}=0.$$

By means of simple modifications to the program, the choice of different equations is possible.

We shall now examine the program operations in their general lines.

First the following input data are read:

PIPPOT = k_0 [from relation (10)]

PULSO = beam nominal energy (MeV)

E = emittance (m·rad)*

X, XPR, Z, ZPR = ellipse center coordinates in the initial planes

RO, XO, RV, XV = R and X values in both initial phase planes

ROF, XOF, RVF, XVF = R and X values required in both output phase planes

N = total number of transport system elements (drift spaces included)

NUMIT = upper limit for the number of iterations

NONT = upper limit for the number of partial iterations

N6 = sequence number for the section in which a focus must be realized;

N8, N9 = sequence number for two different beam sections in which first order achromatism is to be obtained (ellipse centers coincident with phase planes origins);

O11, O12, O13, O21, O23, V11, V12, V23 = total horizontal and vertical matrix element values that can be asked for, if final ellips parameters are not assigned;

MM(I), I = 1,..., N = indices which describe the succession of the elements in the system with the following conventions:
MM(I) = 1 for a drift space,
MM(I) = 2 for a quadrupole lens,
MM(I) = 3 for a bending magnet.

According to the value of MM(I) the program reads the following data:

For drift spaces:

EL(I), length (meter)

For 4-pole lenses:

ELM(I), equivalent magnetic length in the rectangular model (meter);

G(I) = k·ELM(I) (in thin lens approximation) (m^{-1}), where k is the elastic constant of the equation describing the particles motion;

G(I) is, in a first approximation, the value of the

* In our computations we can assume for E the same value in horizontal and vertical planes, but this is not in general verified.

element a_{21} of the quadrupole transfer matrix used as a variable.

For bending magnets:

EN(I), field index

R(I), radius of curvature (meter)

ALF(I), deflection (deg)

AC1(I) and AC2(I), trigonometric tangents of input and output edge angles.

The variables are automatically chosen by the program in the following manner.

Let m_1 and m_2 be respectively the number of the quadrupoles and the number of the magnets. The first m_1 variables are the G(I) of the lenses and the next m_2 variables are the field indexes of the magnets.

By means of simple modifications in the program it is possible to choose other variables (for example the lengths of the drift spaces or the lengths of the quadrupoles).

Now instructions corresponding to a set of partial computation in which the total problem is subdivided are given.

For each partial computation the input is:

J number of imposed equations;

N1 initial section of the examined part of the channel;

N2 final section of the examined part of the channel;

NS number of the trials to be done

(MUT(I), I = 1,..., J), LIST

where MUT(I)... are the labels of the equations to be solved and LIST assumes the values 1 if the envelope must be calculated and printed and 2 on the contrary.

KUT(I, 1), I = 1,..., J

KUT(I, 2), I = 1,..., J

.....

KUT(I, NS), I = 1,..., J

labels of the variables to be used.

The program attempts to find a real solution with the first set of variables; if it goes on, it skips the subsequent sets and comes back to the subsequent partial calculation, otherwise it repeats the trial with the next set of variables taking for the initial values the last ones found in the former trial.

When all the conditions are fulfilled one can attempt to solve the problem of minimizing the beam envelope.

To this aim in our program we adopt a method which although does not calculate exactly the minimum envelope nevertheless avoids the use of the Lagrange's multipliers.

The method used is the following:

A further equation IV) and a further variable are adopted; this equation fixes the maximum value of the envelope along the channel. The initial value that may be assigned is the one the channel has with the last values. Successively one tries to decrease this value changing the variables.

If a real solution is found, one goes on with the attempt until a non real solution is found and then takes the last real one.

We shall now give a partial input example:

Suppose we have choosen

J = 4 N1 = 1

N2 = N NS = 2

MUT(1) = 12 MUT(2) = 13

MUT(3) = 14 MUT(4) = 15

LIST = 2

KUT(1) = 3 KUT(2) = 2 KUT(3) = 5 KUT(4) = 4

KUT(1) = 1 KUT(2) = 3 KUT(3) = 2 KUT(4) = 4

In this case the choice of equations represents an ellipse matching problem to be solved by the variables 3-2-5-4 in the first try, and by the variables 1-3-2-4 in a successive try; the imposed LIST value means that the print out of the channel envelope is wanted.

If in a next intermediate computation we assume

J = 6 N1 = 1 N2 = N NS = 1

MUT(1) = 12 MUT(2) = 13 MUT(3) = 14

MUT(4) = 15 MUT(5) = 2 MUT(6) = 3

LIST = 2

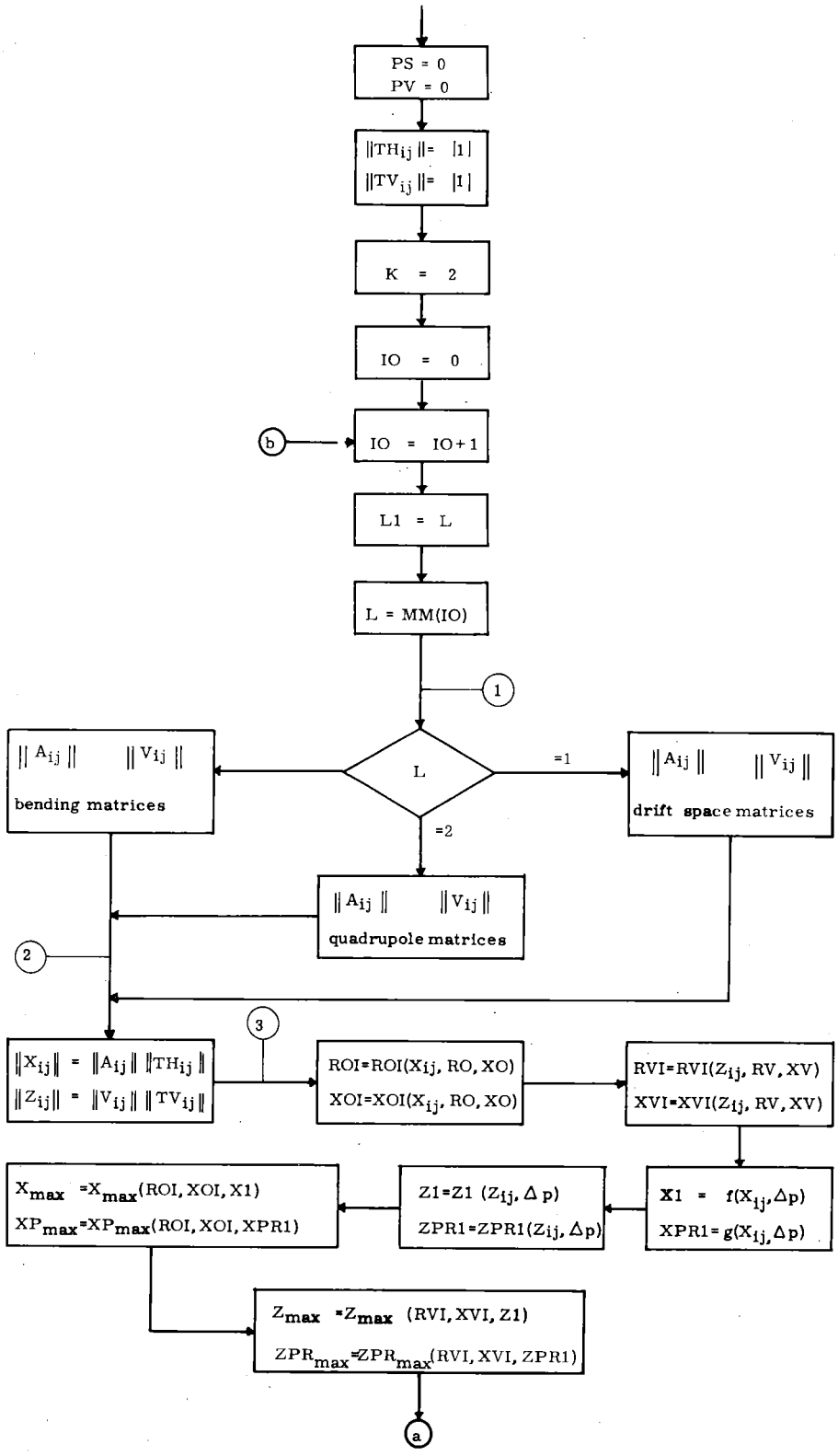
KUT(1) = 3 KUT(2) = 2 KUT(3) = 3

KUT(4) = 4 KUT(5) = 6 KUT(6) = 7

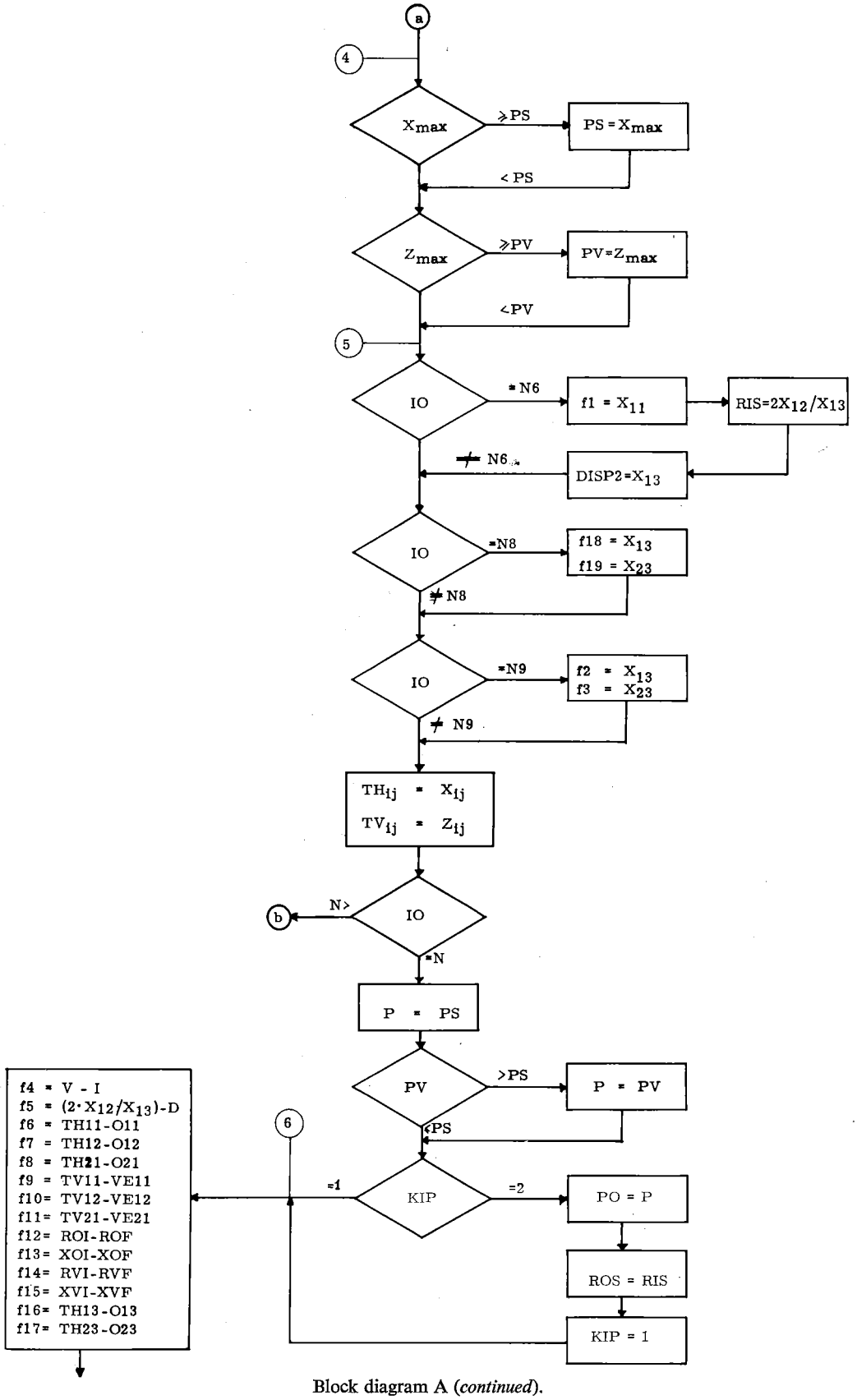
this means that the program tries to solve also the eqs. II) and III) (1^{st} order achromatism at the $n8^{th}$ section) using also variables 6 and 7.

When the input data are read and printed, the program executes the following instructions: (block diagram A).

- Starting from the initial section, it calculates the matrices of all the channel elements, employing initial values parameters (point 1 of block diagram A)
- It executes matrix multiplication, obtaining the two transfer matrices from the initial to a generical section (point 2 of block diagram A).
- For each of them it calculates the coordinates of the center of the ellipse and the associated R and X values, by applying eqs. (4) and (5) to the initial values (point 3 of block diagram A).
- It compares the actual beam envelope values with previous ones and memorizes the largest (point 4 of block diagram A).
- It verifies if the parameters relative to the section must satisfy some equation and, if so, it memorizes their values (point 5 of block diagram A).
- At the final section the assigned equations memorization is complete (point 6 of block diagram A).
- It calculates the quantity (point 11 of block diagram B):



Block diagram A.



Block diagram A (*continued*).

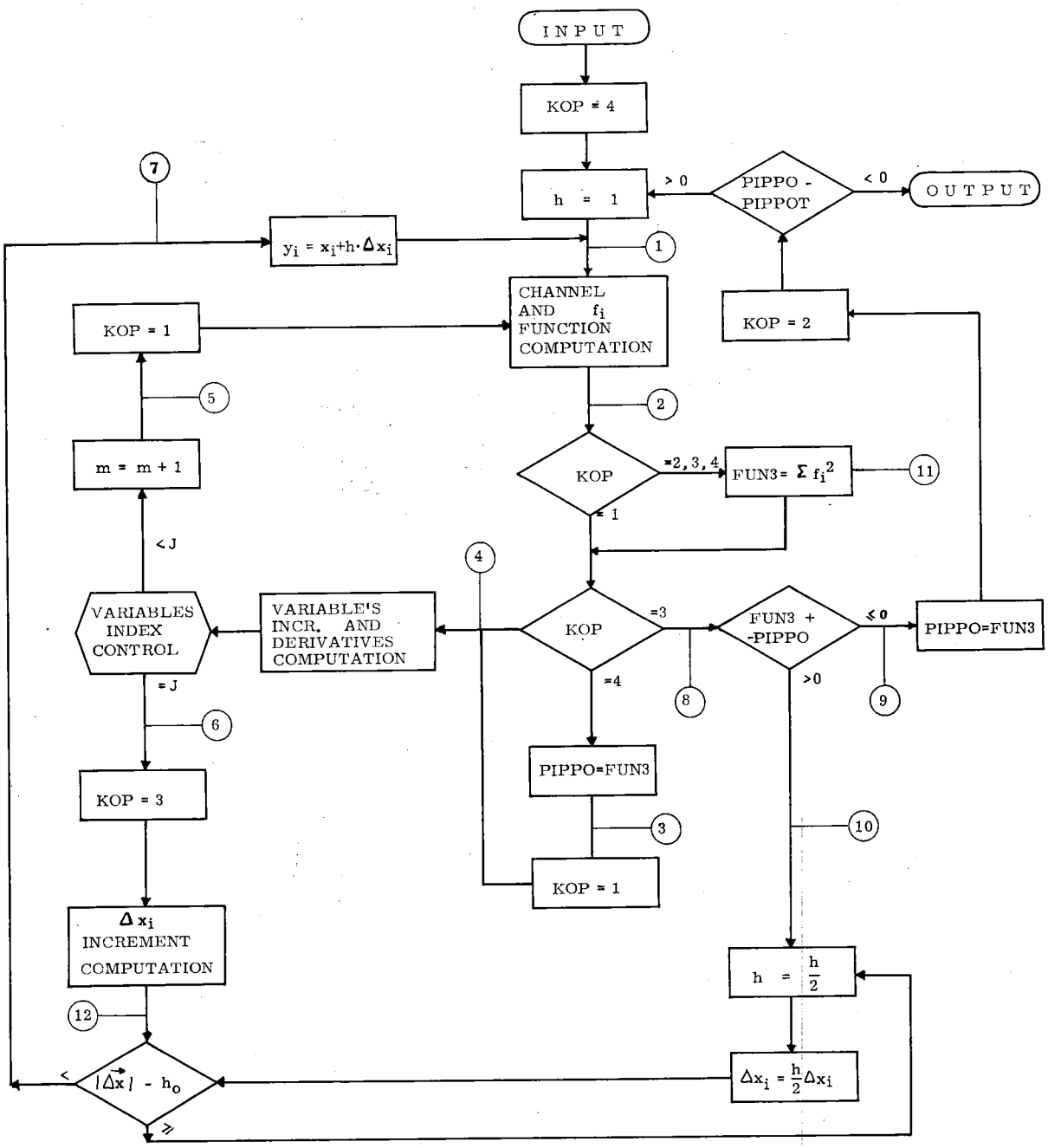
$$\text{FUN3} = \sum f_i^2.$$

- It takes initially (point 3 of block diagram B):

$$\text{PIPPO} = \text{FUN3}$$

and prints its value.

- It calculates the Jacobian of the equations system. In order to perform this calculation, it slightly increases each of the parameters in turn (point 5) and calculates again the functions (path 1-2-4-5-1 of the block diagram).
At the end of the channel it calculates the incremental ratios of the J functions relative to that parameter and treats them as derivatives.
- At this point the program has memorized the values (calculated for the initial values of parameters) of the functions and of their first derivatives and can apply



Block diagram B.

Newton-Raphson's method, to solve the equations.
- It calculates (point 6) the increments to be given to the parameters if they satisfy eq. (8b) it computes the new values for these (point 7). Otherwise it goes to point (10).
- It comes back to point (1) substituting new values for the previous ones.
- After point (8) it compares the calculated FUN3 value with the previous one (memorized as PIPPO).
- If FUN3 is less than PIPPO (point 9), it puts PIPPO = FUN3, prints its value and performs a new iteration with new values for the variables; if FUN3 is greater or equal to PIPPO (point 10), it follows again the path (7-1-2-11-8-10-7) with half-values for the increments and examining, at the end, if FUN3 $\gtrless$ PIPPO. If FUN3 $\geqq$ PIPPO, path (7-1-2-11-8-10-7) is followed again (always with half of the preceding values of increments) and, if after a number of loops the same result is found, partial elaboration is stopped and the program marks that no real solution is in the neighbourhood of the choosen initial set and goes on varying a new set of variables, or solving another set of equations. If FUN3 < PIPPO a new iteration takes place.

After a number of iterations two cases are possible:

a) FUN3 becomes less than k_0. This means that a real solution (from a practical point of view) has been found.

If the IV) [or V)] equation is among the satisfied equations the program tries to decrease the I (or D) value (eqs. IV or V) otherwise it goes on varying another set of equations.

b) When FUN3 does not decrease below the previous value, the computation goes on in the manner described.

The points of the block diagram, where our method differs from the usual Newton-Raphson one are:

1) the computation of the quantity FUN3 = $\sum f_i^2$ and the comparison of the actual to the precedent value (points 11, 8, 9);

2) the computation of the quantity $|\Delta \boldsymbol{x}|$ and the comparison with h_0 (point 12);

3) the multiplication of the increments computed by the Newton-Raphson method by a factor less than one if the comparison between two subsequent FUN3 values or $|\Delta \boldsymbol{x}|$ and h_0 is unsatisfactory.

Besides the general block diagram, we refered to up to now, the one relating the matrix computations more in detail is presented. The running time for a four ellipse matching over a length of 20 m, the initial and final values of R and X being assigned, was about 5 min.

References

1) H. G. Hereward, The properties of particle beams in optical matching systems, in terms of phase-plane ellipse-shapes, PS/Int. TH 59-5 (1959).

2) K. G. Steffen, H. Hultschig and W. Kein, Use of generalized amplitude and phase functions in designing beam transport systems, DESY A 2.70 (1960).

3) E. D. Courant and H. S. Snyder, Ann. Phys. **3** (1958) 1.

4) S. Penner, Rev. Sci. Instr. **32** (1961) 150.

5) P. G. Sona, Ottica degli analizzatori magnetici nell'approssimazione di Gauss. Lab. Naz. di Frascati. Int. Rep. no. 3 (1958).

6) T. L. Saaty and J. Bram, *Nonlinear mathematics* (McGraw-Hill, 1964).