Explicit Thin-Lens Solution for an Arbitrary Four by Four Uncoupled Beam Transfer Matrix

V.Balandin*
Deutsches Elektronen-Synchrotron DESY, Notkestrasse 85, 22607 Hamburg, Germany

S.Orlov
Faculty of Computational Mathematics and Cybernetics,
M.V.Lomonosov Moscow State University, 119991 Moscow, Russia
(Dated: October 28, 2011)

In the design of beam transport lines one often meets the problem of constructing a quadrupole lens system that will produce desired transfer matrices in both the horizontal and vertical planes. Nowadays this problem is typically approached with the help of computer routines, but searching for the numerical solution one has to remember that it is not proven yet that an arbitrary four by four uncoupled beam transfer matrix can be represented by using a finite number of drifts and quadrupoles (representation problem) and the answer to this questions is not known not only for more or less realistic quadrupole field models but also for the both most commonly used approximations of quadrupole focusing, namely thick and thin quadrupole lenses. In this paper we make a step forward in resolving the representation problem and, by giving an explicit solution, we prove that an arbitrary four by four uncoupled beam transfer matrix actually can be obtained as a product of a finite number of thin-lenses and drifts.

I. INTRODUCTION

In the design of beam transfer lines one often encounters the problem of finding a combination of quadrupole lenses and field free spaces (drifts) that will produce particular transfer matrices in both the horizontal and vertical planes. Nowadays this problem is typically approached with the help of computer routines which minimize the deviations from the desired matrices as function of the quadrupole strengths, lengths and distances between them. Although very sophisticated software became available for these purposes during the last decades, there is an important theoretical question which has not been answered yet and whose answer could affect the strategy and efficiency of numerical computations. Searching for a numerical solution one has to remember that it is not proven yet that an arbitrary four by four uncoupled beam transfer matrix can be represented by using a finite number of drifts and quadrupoles (representation problem) and the answer to this questions is not known not only for more or less realistic quadrupole field models but also for the both most commonly used approximations of quadrupole focusing, namely thick and thin quadrupole lenses.

In this paper we make a step forward in resolving the representation problem and prove that an arbitrary four by four uncoupled beam transfer matrix actually can be obtained as a product of a finite number of thin-lenses and drifts. Even though our proof uses more thin-lenses than probably needed, we believe that the solution provided is not only of theoretical interest, but could also find some practical applications because it uses explicit analytical formulas connecting thin-lens parameters with the elements of the input beam transfer matrix.

Though the thin-lens kick is the simplest model of the quadrupole focusing, its role in accelerator physics can hardly be overestimated. The thin-lens quadrupole approximation reveals the analogy between light optics and charged particle optics and, if one takes into account difficulties of analytical manipulations with the next by complexity thick-lens quadrupole model [1, 2], is an indispensable tool for understanding principles and limitations of the already available optics modules and for development of the new optics solutions (see, as good examples, papers [3–7]).

The paper by itself is organized as follows. In section II we introduce all needed notations and give the lower bound on the number of drifts and lenses which are required for a solution of the representation problem by providing an example of a matrix which can not be obtained using five thin-lenses and five variable drift spaces. This example and other of our attempts (though omitted in this paper) to find thin-lens solutions for particular beam transfer matrices lead us to the conjecture that in order to represent an arbitrary four by four uncoupled beam transfer matrix one needs at least six thin-lenses if the distances between them can be varied or at least seven thin-lenses if this variation is not allowed.

In section III we prove that an arbitrary four by four uncoupled beam transfer matrix can be obtained as a product of a finite number of thin-lenses and drifts by giving an explicit solution of the thin-lens representation problem which uses equally spaced thin-lenses. The core idea of our approach is the reduction of the initial 2D problem to two independent 1D problems, and the symmetry of the equally spaced thin-lens system allows to make such a reduction while minimizing the number of nonessential technical details. The solution obtained uti-

*Electronic address: vladimir.balandin@desy.de

lizes thirteen lenses if the spacing between them is fixed beforehand and twelve lenses if this distance can be used as an additional parameter. Thus it uses six more lenses than the minimal number stated in our conjecture, but the setting of these six lenses depends only on the distance between lenses and therefore does not depend (at least directly) on the particular input beam transfer matrix.

And, finally, in section IV we generalize our results of the previous section to the case of arbitrarily spaced thin-lenses. As an important by-product of this section we mention the representation of the matrix of the drift-lens system as a product of elementary P matrices (see formulas (66)-(71) below). This representation could be a useful tool for the analytical study of the properties of thin-lens multiplets and it also shows the difference between the role of the variable drift spaces and the role of the variable lens strengths when they are used as fitting parameters. Besides that thin-lens blocks with decoupled transverse actions introduced in this paper are another point of general interest, although the idea of decoupling by itself is not new in the field of accelerator physics (see, for example, [8, 9]).

II. STATEMENT OF THE PROBLEM AND PRELIMINARY CONSIDERATIONS

Let M be an arbitrary four by four uncoupled beam transfer matrix and let the two by two symplectic matrices M_x and M_y be its horizontal and vertical focusing blocks respectively. Let us denote by $Q(g)$ the transfer matrix of the one dimensional thin lens of strength g and by $D(l)$ the transfer matrix of the one dimensional drift space of length l

$$Q(g) \,=\, \begin{pmatrix} 1 & 0 \\ g & 1 \end{pmatrix}, \qquad D(l) \,=\, \begin{pmatrix} 1 & l \\ 0 & 1 \end{pmatrix}. \tag{1}$$

The problem of representation of the matrix M by a thin-lens system can then be written as

$$D(l_n)\, Q(\pm g_n) \cdot \ldots \cdot D(l_1)\, Q(\pm g_1) \,=\, M_{x,y}, \tag{2}$$

where (here and later on) one has to take the upper sign in the combinations $\pm$ and $\mp$ together with the index x and the lower sign together with the index y.

Note that the drift-lens system presented on the left hand side of the equations (2) consists of equal numbers of drifts and lenses and the first element which the beam sees during its passage is a thin-lens. Alternatively one can consider equations

$$Q(\pm g_n)\, D(l_n) \cdot \ldots \cdot Q(\pm g_1)\, D(l_1) \,=\, M_{x,y}, \tag{3}$$

where the first element is a drift space, or one can use the drift-lens system with a non equal number of drifts and lenses which starts and ends with a drift (or a lens), but for the moment this is not important.

There are many unanswered questions related to the equations (2), the most interesting for us in this paper is the following: given a matrix M, does there exist a number n such that these equations have a solution? If the answer to this question is positive, could the number n be chosen independently from the input matrix M and, if it is also possible, what is the minimal n required?

From a mathematical point of view the equations (2) are a system of eight polynomial equations in $2n$ unknowns and for any polynomial system considered over an algebraically closed field of complex numbers there is an algorithmic way to answer the question if this system has infinitely many solutions or has a finite number of solutions, or has no solutions at all. This can be done by transforming the original system to a special form called a Gröbner basis and, very loosely speaking, is an analogue of the Gaussian elimination process in linear algebra [10]. The Gröbner basis can be computed in finitely many steps and, moreover, nowadays its calculation can be done with the help of symbolic manipulation programs like Mathematica and Maple.

Unfortunately, we are interested in the real solutions of the equations (2) constrained additionally by the requirements for the drift lengths to be nonnegative and therefore we can not use all benefits provided by the Gröbner basis theory. Nevertheless, the use of the Gröbner basis approach, although it did not help us to solve the problem in general, it was very useful in providing examples of particular matrices which can not be obtained using a certain number of thin-lenses and drift spaces. For example, using the Gröbner basis technique, it is possible to prove that the matrix M with

$$M_x \,=\, M_y \,=\, \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \tag{4}$$

can not be represented by five thin lenses and five variable drift spaces starting either from a lens like in the equations (2) or from a drift like in the equations (3).

This and many other of our attempts to study the representation problem for particular beam transfer matrices lead us to the conjecture that in order to be able to represent an arbitrary four by four uncoupled beam transfer matrix one needs at least six thin-lenses if the distances between them can be varied or at least seven thin-lenses with nonzero drift spaces between them if this variation is not allowed.

To finish this section, let us note that in the above discussions we made no use of the fact that we are interested not in the general system of polynomial equations, but only in the polynomial system produced by a product of matrices with simple inversion properties

$$Q^{-1}(g) \,=\, Q(-g), \qquad D^{-1}(l) \,=\, D(-l). \tag{5}$$

Choosing some $k = 1, \ldots, n-1$ and using (5) one can rewrite the system (2) in the equivalent form

$$D(l_k)\, Q(\pm g_k) \cdot \ldots \cdot D(l_1)\, Q(\pm g_1) \,=$$

$$Q(\mp g_{k+1})\, D(-l_{k+1}) \cdot \ldots \cdot Q(\mp g_n)\, D(-l_n)\, M_{x,y}. \quad (6)$$

This trick can be used for the elimination of a part of the unknowns from the original system by solving equations (6) with respect to the variables $g_1, \ldots, g_k, l_1, \ldots, l_k$ or one may even think to construct an iterative solution method which could be considered as matrix version of the method of successive elimination of unknowns [6, 11]. Unfortunately, however this approach did not give us any additional noticeable simplifications in the solution of the representation problem.

III. SOLUTION OF 2D PROBLEM USING EQUALLY SPACED THIN LENSES

In this section we will give an explicit solution of the thin-lens representation problem which uses equally spaced thin-lenses. Instead of equations (2) or (3) we will consider the system

$$B(m_n,\, \pm g_n,\, p_n) \cdot \ldots \cdot B(m_1,\, \pm g_1,\, p_1) \,=\, M_{x,y}, \quad (7)$$

where as an elementary building block we take a thin-lens sandwiched between two drift spaces

$$B(m,\, \pm g,\, p) \,=\, D(p)\, Q(\pm g)\, D(m). \quad (8)$$

If the block length $l = m + p > 0$, then one can represent the block transfer matrix in the form

$$B(m,\, \pm g,\, p) \,=\, S^{-1}(m,\, p)\, P(2 \pm lg)\, S(m,\, p), \quad (9)$$

where

$$S(m,\, p) \,=\, \frac{1}{\sqrt{l}} \begin{pmatrix} 1 & m \\ -1 & p \end{pmatrix}. \quad (10)$$

The definition and the properties of the matrix P (and other elementary matrices used in this paper) can be found in Appendix A.

Let us assume that in the system (7) all m_k and all p_k are equal to each other, i.e. that

$$m_1 = \ldots = m_n = m, \quad p_1 = \ldots = p_n = p, \quad (11)$$

and let $l = m + p > 0$. The principle simplification that occurs in this case is that after the substitution of the representation (9) into equations (7) the matrices $S(m,p)$ and $S^{-1}(m,p)$ cancel each other and we obtain

$$P(2 \pm lg_n) \cdot \ldots \cdot P(2 \pm l\, g_1) \,=\, \hat{M}_{x,y}, \quad (12)$$

where

$$\hat{M}_{x,y} \,=\, S(m,\, p)\, M_{x,y}\, S^{-1}(m,\, p). \quad (13)$$

One sees that while the original system (7) is formed by the product of $2n + 1$ interleaved thin-lens and drift matrices (with neighboring drifts lumped together), the system (12) includes only $n + 2$ matrices depending on unknowns and n of them are P matrices.

Nevertheless, the system (12) is still too complicated to find easily its solutions (or even to prove their existence) for an arbitrary matrix M and with the number of lenses n equal to six or seven as required by our conjecture. Instead we will provide an explicit solution which utilizes thirteen lenses if the parameters m and p are fixed and are independent from the input matrix M, and twelve lenses if m and p can be varied. The main idea of our solution is the reduction of the 2D problem (12) to two independent or, more exactly, almost independent 1D problems by constructing thin-lens blocks which can act in the horizontal and vertical planes similar to a single P matrix, but whose actions for both planes can be chosen independently. At first we will consider a solution of the 1D problem in terms of P matrices. As the next step we will introduce a four-lens block with decoupled transverse actions and then will give an explicit solution of the complete 2D problem. Besides that we will discuss the recipe for constructing lens blocks with decoupled transverse actions with more than four lenses.

Before giving the technical details let us consider one more example obtained with the help of the Gröbner basis technique. Let us assume that m and p are fixed and let the matrix M be such that the matrix $\hat{M}$ is equal to the symplectic unit matrix

$$\hat{M}_x \,=\, \hat{M}_y \,=\, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (14)$$

Then this matrix M can not be represented by less than seven thin-lenses and with seven lenses there are many solutions which geometrically can be viewed as six distinct parallel straight lines in the seven dimensional space of lens strengths.

A. 1D problem in terms of P matrices

According to our plan we will prove in this subsection that every real symplectic 2×2 matrix $M = (m_{ij})$ can be represented as a product of at most four P matrices. First, we will consider the case of three P matrices and will find that three P matrices are insufficient for the representation of an arbitrary 2×2 symplectic matrix. Next we will switch to the case of four P matrices and will show that with four P matrices a solution can always be found, but it is always nonunique.

Let us start with the case of three P matrices, i.e. from the equation

$$P(z_3)\, P(z_2)\, P(z_1) \,=\, M. \quad (15)$$

This matrix equation is, in fact, the system of the four equations for the four matrix elements and, as it is well known, due to symplecticity of the matrices on both sides of (15) these four equations should be equivalent to some system consisting of three equations only. It is clear that such a system can not be obtained simply by throwing away one randomly chosen equation from the system (15) and the first benefit of the special form of the matrix P is that the system of the four third order polynomial equations (15) is equivalent to the system

$$\begin{cases} z_1 = -m_{11} - m_{21} \cdot z_3 \\ z_2 = -m_{22} \\ m_{22} \cdot z_3 = -1 - m_{12} \end{cases} \quad (16)$$

which is linear in the unknowns z_1, z_2 and z_3. Moreover, this system already has a triangular form and its solvability depends only on the solvability of the third equation with respect to the variable z_3.

Elementary analysis shows that there are three possibilities for the solutions of the system (16). If $m_{22} \neq 0$, then there exists a unique solution

$$z_1 = \frac{m_{21} - 1}{m_{22}}, \quad z_2 = -m_{22}, \quad z_3 = -\frac{m_{12} + 1}{m_{22}}. \quad (17)$$

If $m_{22} = 0$ and $m_{21} = 1$ (i.e if $M = -P(-m_{11})$), then there exists a one-parameter family of solutions

$$z_1 + z_3 = -m_{11}, \quad z_2 = 0. \quad (18)$$

And, finally, if $m_{22} = 0$ and $m_{21} \neq 1$, then there is no solution at all.

Very loosely speaking, the condition $m_{22} = 0$ defines the two dimensional surface of singularities in the three dimensional space of 2×2 real symplectic matrices. This surface, in the next turn, contains the one dimensional curve selected by the additional relation $m_{21} = 1$. If the matrix M (represented as a point in our three dimensional space) lies outside of the surface of singularities, then a solution for such a matrix exists and is unique. If the point representing the matrix M belongs to the surface of singularities, then we either have many solutions or none depending on whether this point lies on the above defined one dimensional curve or not.

Let us now turn our attention to the equation

$$P(z_4)\,P(z_3)\,P(z_2)\,P(z_1) \,=\, M, \quad (19)$$

which includes four P matrices. The equivalent to this equation system is given below

$$\begin{cases} z_1 = m_{21} - (m_{11} + m_{21} \cdot z_4) \cdot z_3 \\ z_2 = -m_{12} - m_{22} \cdot z_4 \\ (m_{12} + m_{22} \cdot z_4) \cdot z_3 = m_{22} - 1 \end{cases} \quad (20)$$

and the easiest way to obtain it is to substitute into the system (16) the elements of the matrix $P^{-1}(z_4) \cdot M$ instead of the m_{ij}.

The system (20) is not linear anymore, but still has a triangular form and its solvability depends again only on the solvability of the third equation with respect to the variables z_3 and z_4. Because the matrix M is nondegenerated its elements m_{12} and m_{22} can not be equal to zero simultaneously and therefore the expression $m_{12}+m_{22}\cdot z_4$ considered as a function of z_4 can not be equal to zero in more than one point. It means that the last equation in (20) always has solutions and a good way to understand their complete structure is to consider this equation as the equation of a curve on the plane (z_3, z_4). If $m_{22}\cdot(m_{22}-1) \neq 0$ this curve is a hyperbola with two separate branches, if $m_{22} = 1$ it is a degenerate hyperbola consisting of two intersecting lines $z_3 = 0$ and $z_4 = -m_{12}$, and, finally, if $m_{22} = 0$ we have a single straight line $z_3 = -m_{12}^{-1}$. So we see that with the help of the four P matrices a solution of our problem can always be found and is always nonunique.

B. Four-lens block with decoupled transverse actions

Let us denote by $W^{x,y}$ the following combination of four P matrices

$$W^{x,y} = P(2 \pm lg_4)P(2 \pm lg_3)P(2 \pm lg_2)P(2 \pm lg_1), \quad (21)$$

which in the original variables (7) includes four thin-lenses (four-lens block).

If one chooses $\delta = \pm 1$ and if one takes

$$g_2 = \frac{\delta\sqrt{3}}{l}, \qquad g_3 = -\frac{\delta\sqrt{3}}{l}, \quad (22)$$

then the block matrix can be written as

$$W^{x,y} \,=\, -\Lambda^{-1}\left(\sqrt{u^{x,y}}\right)\, P(w^{x,y})\,\Lambda\left(\sqrt{u^{x,y}}\right), \quad (23)$$

where Λ is a diagonal scaling matrix with

$$u^{x,y} \,=\, 2 \mp \delta\sqrt{3}, \qquad u^x \cdot u^y \,=\, 1 \quad (24)$$

and

$$w^x \,=\, 7 \,+\, u^y \cdot lg_1 \,+\, u^x \cdot lg_4, \quad (25)$$

$$w^y \,=\, 7 \,-\, u^x \cdot lg_1 \,-\, u^y \cdot lg_4. \quad (26)$$

Since for any given value of w^x and w^y the equations (25) and (26) can be solved with respect to the variables g_1 and g_4

$$g_1 = \,\frac{\delta}{l} \cdot \frac{u^y \cdot w^x \,+\, u^x \cdot w^y \,-\, 28}{8\sqrt{3}}, \quad (27)$$

$$g_4 = -\frac{\delta}{l} \cdot \frac{u^x \cdot w^x \,+\, u^y \cdot w^y \,-\, 28}{8\sqrt{3}}, \quad (28)$$

the formula (23) gives the result which we were looking for. Both matrices W^x and W^y are similar to a single P matrix (with an inessential minus sign) and both parameters w^x and w^y can be chosen independently by an appropriate setting of the first and the last lenses in the block.

C. Reduction of 2D problem to two independent or almost independent 1D problems

Since with four P matrices we always can solve the 1D problem, let us first consider a combination of four blocks of the type (23). Using (A17) one can show that the total matrix of this sixteen lens system can be written as follows

$$W_4^{x,y}\, W_3^{x,y}\, W_2^{x,y}\, W_1^{x,y} \;=\; \Lambda\left(a^{x,y}\right) \cdot$$

$$P\left(\hat{w}_4^{x,y}\right) P\left(\hat{w}_3^{x,y}\right) P\left(\hat{w}_2^{x,y}\right) P\left(\hat{w}_1^{x,y}\right) \Lambda\left(a^{x,y}\right), \qquad (29)$$

where

$$a^{x,y} \;=\; \sqrt{\frac{u_1^{x,y} u_3^{x,y}}{u_2^{x,y} u_4^{x,y}}} \qquad (30)$$

and

$$\hat{w}_1^{x,y} = \frac{u_2^{x,y} u_4^{x,y}}{u_3^{x,y}} \cdot w_1^{x,y}, \quad \hat{w}_2^{x,y} = \frac{u_3^{x,y}}{u_1^{x,y} u_4^{x,y}} \cdot w_2^{x,y}, \quad (31)$$

$$\hat{w}_3^{x,y} = \frac{u_1^{x,y} u_4^{x,y}}{u_2^{x,y}} \cdot w_3^{x,y}, \quad \hat{w}_4^{x,y} = \frac{u_2^{x,y}}{u_1^{x,y} u_3^{x,y}} \cdot w_4^{x,y}. \quad (32)$$

Plugging this representation into equations (12) we obtain

$$P\left(\hat{w}_4^{x,y}\right) P\left(\hat{w}_3^{x,y}\right) P\left(\hat{w}_2^{x,y}\right) P\left(\hat{w}_1^{x,y}\right) =$$

$$\Lambda^{-1}\left(a^{x,y}\right) \hat{M}_{x,y} \Lambda^{-1}\left(a^{x,y}\right). \qquad (33)$$

Let us choose arbitrary nonnegative m and p with $l = m + p > 0$ and select for each four-lens block its own $\delta = \pm 1$. This, in accordance with the formulas (22), gives us the setting of the eight lenses in our system and this completely determines the matrix on the right hand side of the equations (33). As the last step we take $\hat{w}_k^x$ and $\hat{w}_k^y$ as some solutions of two independent 1D problems of the type (19) and define the strengths of the remaining eight lenses using the formulas (31)-(32) and (27)-(28).

One sees that using four blocks with decoupled transverse actions the complete 2D problem can always be reduced to two easily solvable independent 1D problems. But do we really need four blocks for making such a reduction? The answer is no and the reason for this is as follows. We know that for most of the 2×2 symplectic matrices the 1D problem can be solved with three P matrices, which means that for most of the 4×4 uncoupled beam transfer matrices the 2D problem can also be solved with three blocks. The problem is what to do with the rest? Happily it turns out that by appropriate choice of the parameters m and p one can always move the input matrix M away from the region of unsolvability and, if the variation of m and p is not allowed, this can be done by using only one additional thin-lens. Thus we arrive at the solution announced in the introduction, namely thirteen lenses if the spacing between them is fixed and twelve lenses if this distance can be used as an additional parameter. Below we will consider in detail the case of twelve lenses (three blocks) with variable spacing and the check that the use of an additional lens for the fixed spacing also works we leave as an exercise for the interested reader.

In analogy with (29) the combination of three blocks can be written as

$$W_3^{x,y}\, W_2^{x,y}\, W_1^{x,y} \;=$$

$$-\,\Lambda^{-1}\left(a^{x,y}\right) P\left(\hat{w}_3^{x,y}\right) P\left(\hat{w}_2^{x,y}\right) P\left(\hat{w}_1^{x,y}\right) \Lambda\left(a^{x,y}\right) \quad (34)$$

where

$$a^{x,y} \;=\; \sqrt{\frac{u_1^{x,y} u_3^{x,y}}{u_2^{x,y}}} \qquad (35)$$

and

$$\hat{w}_1^{x,y} = \frac{u_2^{x,y}}{u_3^{x,y}} \cdot w_1^{x,y}, \qquad (36)$$

$$\hat{w}_2^{x,y} = \frac{u_3^{x,y}}{u_1^{x,y}} \cdot w_2^{x,y}, \qquad (37)$$

$$\hat{w}_3^{x,y} = \frac{u_1^{x,y}}{u_2^{x,y}} \cdot w_3^{x,y}. \qquad (38)$$

Plugging again this representation into system (12) we obtain the equations

$$P\left(\hat{w}_3^{x,y}\right) P\left(\hat{w}_2^{x,y}\right) P\left(\hat{w}_1^{x,y}\right) =$$

$$-\,\Lambda\left(a^{x,y}\right) \hat{M}_{x,y} \Lambda^{-1}\left(a^{x,y}\right). \qquad (39)$$

We know that the sufficient condition for these equations to be solvable with respect to the unknowns $\hat{w}_k^{x,y}$ is that the horizontal and vertical parts of the matrix on the right hand side both have nonvanishing r_{22} elements. The direct calculation gives us

$$r_{22}^{x,y} \;=\; \frac{m_{12}^{x,y} - m\, m_{11}^{x,y} - p\, m_{22}^{x,y} + mp\, m_{21}^{x,y}}{m+p}, \qquad (40)$$

where $m_{ij}^{x,y}$ are the elements of the input matrix M.

Looking for a solution one can proceed further in the same manner as in the four block case with only one difference. At the first step one has to take not arbitrary nonnegative m and p, but such m and p that both r_{22}^x and r_{22}^y are nonzero, which due to symplecticity of the matrices M_x and M_y is always possible.

D. Recipe of construction of lens blocks with decoupled transverse actions

In this subsection we give the recipe for the construction of lens blocks with decoupled transverse actions. As we will see, this recipe works not only for the four-lens combination considered above, but is also applicable to blocks with a larger number of lenses.

Let us consider $q-$lens block with $q \geq 4$

$$W^{x,y} = P(2 \pm lg_q) \cdot \ldots \cdot P(2 \pm lg_1), \tag{41}$$

and let us assume that the product of the $(q-2)$ inner matrices in our block takes the form

$$\begin{pmatrix} 0 & u^{x,y} \\ -\left(u^{x,y}\right)^{-1} & e^{x,y} \end{pmatrix}. \tag{42}$$

Then, as one can show by direct multiplication, the total block matrix becomes

$$W^{x,y} \,=\, -\mathrm{sign}(u^{x,y}) \cdot$$

$$\Lambda^{-1}\left(\sqrt{|u^{x,y}|}\right) P(w^{x,y}) \Lambda\left(\sqrt{|u^{x,y}|}\right), \tag{43}$$

where

$$w^{x,y} = \frac{2 \pm lg_1}{|u^{x,y}|} + |u^{x,y}|\,\left(2 \pm lg_q\right) + \mathrm{sign}(u^{x,y})\, e^{x,y}. \tag{44}$$

If for arbitrary given values of w^x and w^y the equations (44) can be solved with respect to the variables g_1 and g_q, then it will be exactly what we need, and the necessary and sufficient condition for such solvability is

$$|u^x| \;\neq\; |u^y|\,. \tag{45}$$

So, in order to construct the q-lens block with the decoupled transverse actions, one has to solve two equations making the r_{11} elements of the x and y parts of the product of the $(q-2)$ inner matrices equal to zero and one has to satisfy one additional inequality constraint (45).

The solution for the four-lens block was already given above and is unique up to a sign change ($\delta = \pm 1$). Let us now consider more complicated (but still analytically solvable) case of five lenses. In this situation all possible solutions which bring the product of the three inner P matrices

$$P(2 \pm lg_4)\, P(2 \pm lg_3)\, P(2 \pm lg_2), \tag{46}$$

to the form (42) can be expressed as a function of parameters l and g_3 as follows:

$$g_2 \;=\; \frac{1}{l} \cdot \frac{lg_3 + \delta \sqrt{((lg_3)^2 - 2) \cdot ((2lg_3)^2 - 9)}}{(lg_3)^2 - 3}, \tag{47}$$

$$g_4 \;=\; \frac{1}{l} \cdot \frac{lg_3 - \delta \sqrt{((lg_3)^2 - 2) \cdot ((2lg_3)^2 - 9)}}{(lg_3)^2 - 3}, \tag{48}$$

$\delta = \pm 1$, and $l > 0$ and g_3 are such that

$$lg_3 \,\in\, \left(-\infty,\, -\sqrt{3}\right) \cup \left(-\sqrt{3},\, -1.5\right] \cup$$

$$\left[-\sqrt{2},\, \sqrt{2}\right] \cup \left[1.5,\, \sqrt{3}\right) \cup \left(\sqrt{3},\, +\infty\right). \tag{49}$$

To complete the block construction we have to select from all these solutions a subset on which the functions

$$u^{x,y} \,=\, 1 \,-\, (lg_2 \mp 2) \cdot (lg_3 \,+\, lg_4) \tag{50}$$

satisfy the inequality (45). As one can check, this can be achieved simply by removing from the set (49) the endpoints of the given set intervals, i.e. by removing the points ± 1.5 and $\pm\sqrt{2}$. So we see that there are many solutions which allow us to construct from five lenses the block with decoupled transverse actions and for selecting one of them some additional optimization criteria could be involved.

Note that in the blocks constructed according to our recipe the setting of the internal lenses does not depend on the setting of the first and the last lenses and depends only on the geometrical block parameters (distances between the lenses), which will be seen more clearly in the following section where we will consider the case of arbitrarily spaced thin lenses.

Note also that the horizontal and the vertical matrices between the first and the last lenses, when calculated using the original representation (7), both have r_{12} elements equal to zero (i.e. the phase advances between the first and the last lenses in the block are always multiples of 180°), but this alone without the inequality (45) satisfied does not give us the block with the decoupled transverse actions.

IV. GENERALIZATION TO THE CASE OF ARBITRARILY SPACED THIN LENSES

When the distances between the lenses are not equal to each other, we immediately lose the advantage of the cancellation of S matrices between the P matrices after substitution of the representation (9) into equations (7). Nevertheless, as we will show below, this case can also be treated with the tools developed in the previous section.

Let us denote by d_{k_1,k_2} the distance between the lenses with the indices k_1 and k_2 ($k_1 \leq k_2$). We start from the observation that for $k = 2, \ldots, n$ the following identity holds

$$S(m_k,\, p_k)\, S^{-1}(m_{k-1},\, p_{k-1}) \,=$$

$$L\left(\frac{l_k}{d_{k-1,k}} - 1\right) \Lambda\left(\frac{d_{k-1,k}}{\sqrt{l_{k-1}\, l_k}}\right) U\left(1 - \frac{l_{k-1}}{d_{k-1,k}}\right), \tag{51}$$

which can be shown by direct multiplication and which requires that all l_k and $d_{k-1,k}$ are positive.

Let us now substitute the representation (9) into the equations (7) and then plug in in the corresponding places the right hand side of the identity (51). After that the property (A19) allows us to eliminate from the result all L and U matrices while shifting their arguments to the arguments of the neighboring P matrices, and leaving us with a product consisting of alternating P and Λ matrices. Although the Λ matrices can not be eliminated completely, they can be moved either on the left or on the right hand side of all P matrices with the help of the property (A17). As the last step we transfer all matrices from the left and right sides of the obtained solid block of the P matrices to the right hand side of our equation, hide them in the matrix $\tilde{M}_{x,y}$ and end up with the equations

$$P(\tilde{v}_n^{x,y}) \cdot \ldots \cdot P(\tilde{v}_1^{x,y}) \;=\; \tilde{M}_{x,y}, \tag{52}$$

which already have the desired form. The detailed structure of the arguments $\tilde{v}_k^{x,y}$ and of the matrix $\tilde{M}_{x,y}$ depends on the particular ways how the individual Λ matrices were moved (to the left or to the right sides) and is given below for the case when during transformations all Λ matrices were moved to the left hand side of the P matrix block. Nevertheless, the expressions given below are general in the sense that they contain an arbitrary positive parameter c_1, and with the proper choice of this parameter one can account for all possible ways of movement of the individual Λ matrices.

$$\tilde{M}_{x,y} = \Lambda(c_n)\, S(m_n, p_n)\, M_{x,y} S^{-1}(m_1, p_1) \Lambda(c_1), \tag{53}$$

$$\tilde{v}_k^{x,y} = c_k^2 l_k \left(\frac{d_{k-1,k+1}}{d_{k-1,k}\, d_{k,k+1}} \pm g_k \right), \quad k = 1, \ldots, n, \tag{54}$$

$$c_k \;=\; \frac{d_{k-1,k}}{\sqrt{l_{k-1}\, l_k}} \cdot \frac{1}{c_{k-1}}, \qquad k = 2, \ldots, n, \tag{55}$$

c_1 is an arbitrary positive parameter and, because we do not have lenses with indices 0 and $n+1$, we use the conventions that

$$d_{0,1} \;=\; l_1, \quad d_{0,2} \;=\; d_{0,1} + d_{1,2}, \tag{56}$$

$$d_{n,n+1} \;=\; l_n, \quad d_{n-1,n+1} \;=\; d_{n-1,n} + d_{n,n+1}. \tag{57}$$

Note that for $c_1 = 1$ and with the condition (11) satisfied the representation (52) turns into the representation (12) as one can expect. Note also that by proper choice of c_1 one can for even n satisfy $c_n = c_1$ and for odd n satisfy $c_n = c_1^{-1}$.

Now in order to continue we need a lens block with the decoupled transverse actions and, as it is not difficult to check, the recipe given in the previous section is applicable without any changes. For the construction of the q-lens block we still need to bring the product of the $(q-2)$ inner matrices to the form (42) while also satisfying the inequality constraint (45). For the four-lens case

$$W^{x,y} \;=\; P(\tilde{v}_4^{x,y})\, P(\tilde{v}_3^{x,y})\, P(\tilde{v}_2^{x,y})\, P(\tilde{v}_1^{x,y}) \tag{58}$$

the two equations making the r_{11} elements of the x and y parts of the product of the two inner matrices equal to zero are

$$\tilde{v}_2^{x,y} \cdot \tilde{v}_3^{x,y} = 1, \tag{59}$$

and have a solution

$$g_2 = \;\frac{\delta}{d_{1,2}} \cdot \sqrt{\frac{d_{1,4}}{d_{2,3}} \cdot \frac{d_{1,3}}{d_{2,4}}}, \tag{60}$$

$$g_3 = -\frac{\delta}{d_{3,4}} \cdot \sqrt{\frac{d_{1,4}}{d_{2,3}} \cdot \frac{d_{2,4}}{d_{1,3}}}, \tag{61}$$

which again is unique up to a sign change ($\delta = \pm 1$). The values $u^{x,y}$ for this solution are

$$u^{x,y} \;=\; \tilde{v}_3^{x,y} \;=\; \frac{c_3^2\, l_3}{d_{3,4}} \cdot \left(\frac{d_{2,4}}{d_{2,3}} \mp \delta \cdot \sqrt{\frac{d_{1,4}}{d_{2,3}} \cdot \frac{d_{2,4}}{d_{1,3}}} \right). \tag{62}$$

Both of them are positive and clearly satisfy the inequality (45). With this choice for g_2 and g_3 the total block matrix takes the form

$$W^{x,y} \;=\; -\Lambda^{-1}\left(\sqrt{u^{x,y}}\right)\, P(w^{x,y})\, \Lambda\left(\sqrt{u^{x,y}}\right) \tag{63}$$

where

$$w^{x,y} \;=\; (u^{x,y})^{-1} \cdot \tilde{v}_1^{x,y} \;+\; u^{x,y} \cdot \tilde{v}_4^{x,y} \;-\; 1, \tag{64}$$

and all results of the previous section concerning the reduction of the 2D problem to two 1D problems become applicable with some minor changes connected with the difference in the matrices $\hat{M}_{x,y}$ and $\tilde{M}_{x,y}$ defined by the relations (13) and (53) respectively. Note that if, when placed in the beam line, the actual decoupling block starts from the lens with the index k, one has simply to add $k-1$ to the indices $1, 2, 3$ and 4 in all above formulas.

A. Removing of superfluous parameters

The equations (7) contain $2n$ parameters which specify the drift lengths $(m_1, p_1, \ldots, m_n, p_n)$ while only $n+1$ parameters, namely $m_1, d_{1,2}, \ldots, d_{n-1,n}, p_n$ have a clear physical meaning and are independent. Let us have a closer look at the formulas (52)-(57) and count how many superfluous parameters are still left in them and then show ways how to remove them.

The superfluous parameters p_1 and m_n are clearly present, either directly as the arguments of S matrices or through the lengths of the first and the last building blocks l_1 and l_n. And actually that is all. The presence of the other superfluous parameters through the values $l_2, \ldots, l_{n-1}$ is completely imaginary. To show this let us note that these values can enter the main formulas (52)-(54) only through the values c_1 and c_n and through the combinations $c_1^2 l_1, \ldots, c_n^2 l_n$. So if we choose c_1 to be independent from $l_2, \ldots, l_{n-1}$, then these parameters can enter in none of the combinations $c_k^2 l_k$ due to the recursion relation

$$c_k^2 \, l_k \;=\; d_{k-1,k}^2 \cdot \frac{1}{c_{k-1}^2 \, l_{k-1}} \qquad k = 2, \ldots, n, \tag{65}$$

which follows from the recursion relation (55), and likewise they can not enter the value c_n because one can write that $c_n = \sqrt{c_n^2 l_n / l_n}$.

Thus there are only two superfluous parameters, p_1 and m_n, present in our formulas, either directly or through the values l_1 and l_n. Do we need to remove them? In general not, because it is clear that none of the physically meaningful answers will depend on them and, in this sense, their absence in the final results (like in formulas (60) and (61)) could work as some indirect indicator of the correctness of the calculations. But from another point of view, it seems better not to have any superfluous parameters from which one can expect nothing except some possible additional complications.

The simplest way to remove the parameters p_1 and m_n from the formulas (52)-(54) is to make them functions of the physically meaningful parameters. For example, one can take $p_1 = 0.5 \cdot d_{1,2}$ and $m_n = 0.5 \cdot d_{n-1,n}$. However the way which we prefer is the modification of the formulas (52)-(54) in such a way that the superfluous parameters will disappear automatically. In doing so let us first present the final result and then make some remarks on how it can be obtained.

$$P(v_n^{x,y}) \cdot \ldots \cdot P(v_1^{x,y}) \;=\; \check{M}_{x,y}, \tag{66}$$

$$\check{M}_{x,y} \;=\; J \Lambda^{-1}(b_n) U(-p_n) \, M_{x,y} \, U(-m_1) \Lambda(b_1), \tag{67}$$

$$v_1^{x,y} = b_1^2 \left(\frac{1}{d_{1,2}} \pm g_1 \right), \tag{68}$$

$$v_k^{x,y} = b_k^2 \left(\frac{d_{k-1,k+1}}{d_{k-1,k} \, d_{k,k+1}} \pm g_k \right), \;\; k = 2, \ldots, n-1, \tag{69}$$

$$v_n^{x,y} = b_n^2 \left(\frac{1}{d_{n-1,n}} \pm g_n \right), \tag{70}$$

$$b_1 > 0, \quad b_k \;=\; d_{k-1,k} \cdot \frac{1}{b_{k-1}}, \quad k = 2, \ldots, n. \tag{71}$$

In order to obtain these formulas from (52)-(57) let us first introduce the parameters $b_k = c_k \sqrt{l_k}$ and then assume that c_1 is chosen in such a way that b_1 does not depend on any superfluous parameter (for example, one simply can take $c_1 = 1/\sqrt{l_1}$). After this one sees that the parameters l_1 and l_n enter the left hand side of the equation (52) only through the matrices $P(\tilde{v}_1^{x,y})$ and $P(\tilde{v}_n^{x,y})$. Due to the property (A19) these matrices can be decomposed into the following products

$$P(\tilde{v}_1^{x,y}) \;=\; P(v_1^{x,y}) \, L(c_1^2) \;=\; P(v_1^{x,y}) \, L(b_1^2 \, / \, l_1), \tag{72}$$

$$P(\tilde{v}_n^{x,y}) = U(-c_n^2) \, P(v_n^{x,y}) = U(-b_n^2 \, / \, l_n) \, P(v_n^{x,y}). \tag{73}$$

As the last step, one has to substitute these decompositions back into the equation (52), transfer U and L to the right hand side and, after some straightforward manipulations, arrive at the final result described in the above formulas (66)-(71).

Note that the whole story about the presence of the superfluous parameters is the result of our desire to have the expressions for the problem description (expressions (52)-(57)) which reduces to the highly symmetric expressions (12) and (13) in the limit of equal distances between thin lenses. If one does not require that, then, as we will outline below, it is possible to arrive at the representation (66)-(71) without using the identity (9).

According to (A20) and (A21) the matrix of the building block can be written as

$$B(m, \, \pm g, \, p) \;=\; P(-p) \, P(\pm g) \, P(-m) \, J. \tag{74}$$

Substituting this representation in the original equations (7) and using that due to (A8)

$$P(-m_k) \, J \, P(-p_{k-1}) \;=\; -P(-d_{k-1,k}) \tag{75}$$

we obtain

$$P(\pm g_n) P(-d_{n-1,n}) \cdot \ldots \cdot P(-d_{1,2}) P(\pm g_1) \Lambda(b_1) =$$

$$(-1)^{n-1} J \, U(-p_n) \, M_{x,y} \, U(-m_1) \, \Lambda(b_1), \tag{76}$$

where we have already introduced an arbitrary positive parameter b_1. Now, assuming that all distances between lenses are positive and using (A16), we can replace for each $k = 2, \ldots, n$ the matrix $P(-d_{k-1,k})$ by the matrix $-\Lambda(d_{k-1,k})$ with simultaneous adding to the arguments of the two neighboring P matrices the value $d_{k-1,k}^{-1}$. After these manipulations we arrive at the expression

$$P(d_{n-1,n}^{-1} \pm g_n) \, \Lambda(d_{n-1,n}) \cdot$$

$$P(d_{n-1,n}^{-1} + d_{n-2,n-1}^{-1} \pm g_{n-1}) \, \Lambda(d_{n-2,n-1}) \cdot \ldots$$

$$\ldots \cdot P(d_{2,3}^{-1} + d_{1,2}^{-1} \pm g_2) \Lambda(d_{1,2}) P(d_{1,2}^{-1} \pm g_1) \Lambda(b_1) =$$

$$J \, U(-p_n) \, M_{x,y} \, U(-m_1) \, \Lambda(b_1), \tag{77}$$

and the last step, which is still necessary in order to obtain formulas (66)-(71), is to move all Λ matrices to the left in the left hand side of the equation (77) using the identity (A17) with a subsequent transfer of the matrix $\Lambda(b_n^{-1})$ from the left to the right hand side of the obtained equality.

Acknowledgments

The authors are thankful to Winfried Decking, Nina Golubeva and Helmut Mais for support and their interest in this work. The careful reading of the manuscript by Helmut Mais is gratefully acknowledged.

Appendix A: Elementary matrices and their properties

The elementary symplectic P matrix which is defined as follows

$$P(a) \,=\, \begin{pmatrix} a & 1 \\ -1 & 0 \end{pmatrix} \tag{A1}$$

and which we use extensively throughout this paper was found empirically by the usual trial and error method during attempts to reduce the problem of analytical study of thin-lens multiplets to some "more manageable" form. As we will see below, this matrix possesses many interesting properties not only by itself, but also in combination with the other elementary matrices. Although not widely known in the scientific community, it was no surprise, as we found later, that it was successfully used in some special area of abstract algebra [12].

In order to give an expression for the product of n elementary P matrices let us first define a sequence of polynomials κ_n in the variables $z_1, \ldots, z_n$ recursively by the following equations

$$\kappa_{-1} \,=\, 0, \quad \kappa_0 \,=\, 1, \tag{A2}$$

$$\kappa_n(z_1, \ldots, z_n) \,=\, z_n \cdot \kappa_{n-1}(z_1, \ldots, z_{n-1}) - \kappa_{n-2}(z_1, \ldots, z_{n-2}), \quad n \,\geq\, 1. \tag{A3}$$

With these notations we assert that

$$P(a_n) \cdot \ldots \cdot P(a_1) \,=\, \begin{pmatrix} \kappa_n(a_1, \ldots, a_n) & \kappa_{n-1}(a_2, \ldots, a_n) \\ -\kappa_{n-1}(a_1, \ldots, a_{n-1}) & -\kappa_{n-2}(a_2, \ldots, a_{n-1}) \end{pmatrix}, \tag{A4}$$

which is clear for $n = 1$ and in general case can be proven by induction. Because such induction can be made in two different ways, either by adding one more P matrix from the left or from the right side, it is easy to see that the polynomials κ_n can also be defined by (A2) and by the recursion relation

$$\kappa_n(z_1, \ldots, z_n) \,=\, z_1 \cdot \kappa_{n-1}(z_2, \ldots, z_n) - \kappa_{n-2}(z_3, \ldots, z_n), \quad n \,\geq\, 1. \tag{A5}$$

Comparison of (A3) and (A5) implies that

$$\kappa_n(z_1, z_2, \ldots, z_{n-1}, z_n) \equiv \kappa_n(z_n, z_{n-1}, \ldots, z_2, z_1). \tag{A6}$$

According to (A4) we can write down the matrix of the product of any number of elementary P matrices without making any matrix multiplications. In this connection let us note that the problem of deriving some recursion relations which allow to obtain the transfer matrix of an arbitrary multiplet without actual matrix multiplications was also addressed in [1].

It is clear that the matrix $P(0)$ coincides with the 2×2 symplectic unit matrix J, i.e. that

$$P(0) \,=\, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \,=\, J, \tag{A7}$$

and the following relations between the P matrices can be easily verified by direct multiplication:

$$P(a)\, J\, P(b) \,=\, -P(a + b), \tag{A8}$$

$$P^3(\pm 1) \,=\, \mp I, \tag{A9}$$

$$P^{-1}(a) \,=\, J\, P(-a)\, J \,=\, a \cdot I \,-\, P(a), \tag{A10}$$

$$P(a)\, P^{-1}(b) \,=\, -P(a - b)\, J, \tag{A11}$$

$$P(a)\, P^{-1}(b)\, P(c) \,=\, P(a - b + c), \tag{A12}$$

where I is the 2×2 identity matrix.

Let us now introduce three more elementary matrices. The diagonal (scaling) matrix

$$\Lambda(a) \,=\, \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} \tag{A13}$$

and the lower and upper triangular matrices with unit diagonal elements

$$L(a) \,=\, \begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix}, \quad U(a) \,=\, \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}. \tag{A14}$$

Note that although the matrices L and U formally coincide with the matrices of the thin lens and the drift space respectively, we have introduced them in order to distinguish the situations where matrix of lens or drift has physical meaning and where the usage of low or upper triangular matrix is simply the reflection of the mathematical technique used.

We have the following relations between the matrices P, Λ, L and U:

$$P(a)\, P(a^{-1})\, P(a) \,=\, \Lambda(-a), \tag{A15}$$

$$P(a)\, P(b^{-1})\, P(c) \,=\, P(a - b)\, \Lambda(b^{-1})\, P(c - b), \tag{A16}$$

$$\Lambda(a)\, P(b)\, \Lambda(a) \,=\, P(a^2 b), \tag{A17}$$

$$P(a^{-1}) \,=\, L(-a)\,\Lambda(a^{-1})\,U(a), \qquad \text{(A18)}$$

$$U(a)\,P(b)\,L(c) \,=\, P(b+c-a), \qquad \text{(A19)}$$

$$L(a) \,=\, -J\,P(a), \qquad \text{(A20)}$$

$$U(a) \,=\, -P(-a)\,J. \qquad \text{(A21)}$$

Although these relations are elementary, they are basic for all results of this paper.

---

[1] E.Regenstreif, *Phase-space transformation by means of quadrupole multiplets*, CERN 67-6, March 1967.

[2] E.Regenstreif, *Possible and impossible phase-space transformations by means of alternating-gradient doublets and triplets*, CERN 67-8, March 1967.

[3] K.L.Brown, R.V.Servranckx, *Optics modules for circular accelerator design*, NIM A258, 1987.

[4] B.W.Montague, F.Ruggiero, *Apochromatic Focusing for Linear Colliders*, CLIC-Note 37, May 1987.

[5] B.Zotter, *Design considerations for a chromatically corrected final focus for TeV colliders*, CLIC-Note 64, June 1988.

[6] O.Napoly, *Thin lens telescopes for final focus systems*, CERN/LEP-TH/89-69, CLIC Note 102, November 1989.

[7] E.T.d'Amigo, G.Guignard, *Analysis of generic insertions made of two symmetric triplets*, CERN SL/98-014(AP), May 1998.

[8] T.Roser, *Decoupled Beam Line Tuning*, AGS/AD/Tech. Note No. 357, January 1992.

[9] N.J.Walker, J.Irwin, M.Woodley, *Global Tuning Knobs for the SLC Final Focus*, Proceedings of PAC'93, Washington, DC, USA.

[10] D.Cox, J.Little, D.O'Shea, *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Computer Algebra*, Undergraduate Texts in Mathematics. Springer, New York, 1992.

[11] Y-Chiu Chao, J.Irwin, *Solution of a three-thin-lens system with arbitrary transfer properties*, SLAC-PUB-5834, October 1992.

[12] P.M.Cohn, *On the structure of the GL_2 of a ring*, Publ. Math. I.H.E.S., No.30, Paris, p.5-53 (1966).