Analytical theory of arbitrary-order achromats

Weishi Wan* and Martin Berz
Department of Physics and Astronomy and National Superconducting Cyclotron Laboratory, Michigan State University, East Lansing, Michigan 48824
(Received 11 January 1996)

An analytical theory of arbitrary-order achromats for optical systems with midplane symmetry is presented. It is based on the repeated use of identical cells; but besides mere repetition of cells, mirror symmetry is used to eliminate aberrations. Using mirror imaging of a cell around the x-y and x-z planes, we obtain four kinds of cells: the forward cell (F), the reversed cell (R), the cell in which the direction of bend is switched (S), and the cell where reversion and switching is combined (C). Representing the linear part of the map by a matrix, and the nonlinear part by a single Lie exponent, the symplectic symmetry is accounted for and transfer maps are easily manipulated. It is shown that independent of the choice and arrangement of such cells, for any given order, there is a certain minimum number of constraint conditions that has to be satisfied. It is shown that the minimum number of cells necessary to reach this optimum level is four, and out of the sixty-four possible four-cell symmetry arrangements, four combinations yield such optimal systems. As a proof of principle, the design of a fifth-order achromat is presented. [S1063-651X(96)07709-4]

PACS number(s): 29.27.−a, 41.85.Gy, 41.85.Ja

I. INTRODUCTION

The search for achromats, by which we mean optical systems whose transfer maps are free of any nonlinearities of transverse motion up to a certain order, has generated substantial interest for the past two decades. The ability to transport particles while preserving their phase-space distribution through a certain order has many attractive consequences, and this is why first- and second-order achromats have been widely used in accelerators, storage rings, and beam transport lines. Last but not least, this is also an interesting and challenging problem from a purely theoretical point of view.

Since midplane symmetry has been employed in most of the beam optical systems, canceling half of the transverse aberrations, all achromat theories consider only systems with midplane symmetry. But even under this symmetry, as Table I shows, the number of aberrations that have to be canceled to generate achromats increases drastically with order, even if only the aberrations that are independent due to the symplectic symmetry are considered. In particular, it becomes apparent why even up to second order it is considered unrealistic to attempt to control each aberration with a separate constraint condition. Therefore, the challenge is to achieve achromaticity with as few knobs as possible.

Although the concept of first-order achromats had been widely used in various beam optical systems and accelerators for a long time, it was only until the 1970s that a theory developed by K. Brown enabled the design of realistic second-order achromats in a systematic and elegant way [1].

The theory is based on the following observations. First, any system of n identical cells ($n>1$) with the overall first-order matrix equaling to unity (I) in both transverse planes gives a first-order achromat. When n is not equal to three, it also cancels all second-order geometric aberrations. Second, of all second-order chromatic aberrations, only two are independent. Therefore, they can be corrected by two families of purely second-order elements (in practice usually sextupoles), each responsible for one of them in each transverse plane. These findings make it possible to design a four-cell, second-order achromat with only one dipole, two quadrupoles and two sextupoles per cell.

Because of its simplicity, the second-order achromat concept has been applied to the design of various beam optical systems such as the time-of-flight mass spectrometers both single-pass (TOFI) [2,3] and multipass (ESR) [4,5], the arcs of the Stanford linear collider (SLC), the new facility at SLAC, the final focus test beam [6–9], and the MIT South Hall ring (SHR) [10,11].

Since it is hard to generalize this second-order achromat theory to higher orders, the first third-order achromat theory was developed by Dragt based on normal form theory and Lie algebra [13]. According to the theory, a system of n identical cells is a third-order achromat if the following conditions are met: (1) The tunes of a cell μ_x and μ_y are not full, half, third, or quarter integer resonant, but $n\mu_x$ and $n\mu_y$ are integers. (2) The two chromaticities and five independent third-order aberrations are zero.

Two examples of third-order achromats have been designed. The first design was done by Dragt himself, containing thirty cells ($\mu_x=1/5$ and $\mu_y=1/6$). Each contains ten bends, two quads, two sextupoles, and five octupoles. The

TABLE I. The number of aberrations of orders 2 to 5 for a system with midplane symmetry. The interdependency of aberrations is due to the symplectic symmetry inherent in Hamiltonian dynamics.

Order	2	3	4	5
Aberrations	30	70	140	252
Independent aberrations	18	37	65	110

*Present address: Department of Astrophysical, Planetary and Atmospheric Sciences, University of Colorado at Boulder, Boulder, CO 80309-0391.

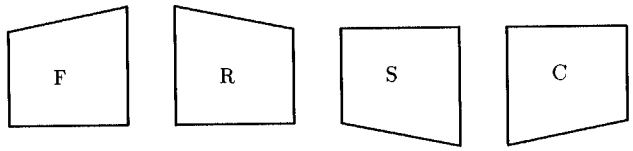

FIG. 1. The geometric relationship among cells F, R, S, and C illustrated by asymmetric boxes.

whole system is a 180° bending arc. The second design was done by Neri [14,15]. It is a seven-cell system with only one bend per cell, while the total bend is also 180°. The tunes of a cell are $\mu_x=1/7$ and $\mu_y=2/7$, which seems to violate the theory because of the third-order resonance $2T_x-T_y=0$. However, achromaticity can still be achieved because the driving terms are canceled by midplane symmetry [14,15], which greatly reduces the number of cells.

Similar to Brown's theory, the Dragt theory cannot be immediately used to find arbitrary-order achromats in a way that the number of cells is independent of the order. The main reason is that for any given order, the tunes of a cell have to be specially chosen such that most, if not all, of the resonances up to one order higher are avoided. Thus the number of cells has to be the smallest possible that makes both tunes of the whole system integers, which depends on the order and usually increases quickly. A second reason is that as the order increases, the difficulty of obtaining analytical formulas increases rapidly because of the inherent complexity of the Baker-Campbell-Hausdorff (BCH) formula.

Our approach for a general achromat theory does not use the normal form method and avoids the resonance concern by introducing mirror symmetry to cancel more aberrations. With these considerations, we are able to study systems with arbitrary numbers of cells and obtain solutions that are independent of the arrangements inside a cell. Because of their simplicity, Lie transformations are used to represent symplectic maps, but instead of an order-by-order factorization, we use a factorization formed by a linear matrix and a single Lie operator, describing the linear and nonlinear parts, respectively. The introduction of mirror symmetry makes it possible for us to obtain four total kinds of cells: the forward cell (F), the reversed cell (R), the switched cell in which the direction of bend is switched (S), and the cell in which reversion and switching is combined (C), as shown schematically in Fig. 1.

In Sec. II, we derive the maps of the cells R, S, and C from the map of the forward cell and then the maps of a four-cell system. Section III contains a classification of the systems with optimal solutions. First it is shown that it is necessary to have at least four cells in a system to achieve an arbitrary-order achromat with a minimum number of constraints, which is followed by the proof of the existence of an optimal solution. It is further shown that four out of the 64 four-cell systems yield the optimal solution while requiring the minimum number of linear constraints (Sec. III D). In Sec. IV, we study the four best systems in detail to find solutions for achromaticity in an order-by-order fashion. First a general solution for arbitrary-order achromats is obtained, although it will be shown that more efficient solutions can be found. Then the optimal solution for achromats up to sixth order is determined. In Sec. V, as a direct application of the theory developed in this paper, the design of a four-cell fifth-order achromat is presented.

II. MAP REPRESENTATIONS

Let us consider a phase space consisting of $2m$ variables $(q_1,\ldots,q_m,p_1,\ldots,p_m)$. Since we do not take into account synchrotron radiation and acceleration, the transverse motion of a beam optical system is described by a symplectic map, which satisfies

$$\mathcal{J}(\vec{M})\cdot\hat{J}\cdot\mathcal{J}(\vec{M})^t=\hat{J},$$

where $\mathcal{J}(\vec{M})$ is the Jacobian matrix of $\vec{M}$ and $\hat{J}$ is an antisymmetric $2m\times 2m$ matrix

$$\hat{J}=\begin{pmatrix} 0 & \hat{I} \\ -\hat{I} & 0 \end{pmatrix}. \tag{1}$$

Under the symplectic symmetry, the transfer map $\vec{M}$ of order n can be represented by a matrix L and a polynomial H of orders three up to $n+1$ through a Lie factorization as

$$\vec{M}=_n(L\vec{I})\circ[\exp(:H:)\vec{I}], \tag{2}$$

where $=_n$ denotes equality up to order n and $\vec{I}$ is the identity map. Up to order n, the inverse of M is

$$\vec{M}^{-1}=_n[\exp(-:H:)\vec{I}]\circ(L^{-1}\vec{I}). \tag{3}$$

For the convenience of further discussion, it is advantageous to define what we want to call a ''standard'' and a ''substandard'' form of the maps:

Definition II.1. For a symplectic map $\vec{M}_S$, the standard form is defined as

$$\vec{M}_S=\exp(:H:)(M_L\vec{I}), \tag{4}$$

where M_L is the linear matrix and H is a polynomial of order three and up. A representation of the form

$$\vec{M}_S=\left[\prod_i \exp(:H_i:)\right](M_L\vec{I}) \tag{5}$$

is called a substandard form.

H is sometimes called the pseudo-Hamiltonian of the map, which is extracted from the map ([12]). Note the difference between Eqs. (2) and (4), where, in the former equation, $\exp(:H:)$ acts on $\vec{I}$ and the resulting map is then composed with the linear map, and, in the latter equation, $\exp(:H:)$ acts on the linear map directly. Apparently, use of the Baker-Campbell-Hausdorff formula in principle allows the transformation of a substandard form into a standard form.

Due to the frequent usage in the sections that follow, two closely related theorems are introduced without proofs.

Theorem II.1. If f is a polynomial of order three or higher on $\mathbb{R}^{2m}$, and g is an arbitrary polynomial, then we have

$$\exp(:f:)g=_n g[\exp(:f:)\vec{I}], \tag{6}$$

where

$$\exp(:f:)\vec{I} = {}_n[\exp(:f:)q_1,\ldots,\exp(:f:)q_n, \times \exp(:f:)q_1,\ldots,\exp(:f:)q_n]. \quad (7)$$

Theorem II.1 is ready to be generalized to the case where g is a polynomial of a map $\vec{M}$. Thus, we have the following theorem:

Theorem II.2. If $\vec{M}$ is a map and g is an arbitrary polynomial on $\mathbb{R}^{2m}$, then we have

$$\exp(:f:)g(\vec{M}) = {}_n g[\exp(:f:)\vec{M}]. \quad (8)$$

Theorem II.2 is probably even more important than Theorem II.1 because, in most cases, the linear map of a system is not unity. Therefore, it is the one which is used directly and very frequently.

From Theorem II.1, we can infer that $\vec{M}$ can be written in the standard form

$$\vec{M} = \exp(:H:)(L\vec{I}). \quad (9)$$

Next let us make some observations about the composition of $\exp(:f:)g(\vec{I})$ and the transfer map $\vec{M}$.

Theorem II.3. If a map $\vec{M}$ on $\mathbb{R}^{2m}$ is symplectic, we have

$$[\exp(:f:)g]\circ(\vec{M}) = {}_n \exp[:f(\vec{M}):]g(\vec{M}). \quad (10)$$

If a map $\vec{M}$ on $\mathbb{R}^{2m}$ is antisymplectic, i.e., $\mathcal{J}(\vec{M})\cdot\hat{J}\cdot\mathcal{J}(\vec{M})^t = -\hat{J}$, we have

$$[\exp(:f:)g]\circ(\vec{M}) = {}_n \exp[-:f(\vec{M}):]g(\vec{M}). \quad (11)$$

Proof: First assume that $\vec{M}$ is a symplectic map, which implies that a Poisson bracket is an invariant under the transformation $\vec{M}$, i.e.,

$$[f(\vec{M}),g(\vec{M})] = ([f,g])(\vec{M}).$$

The rest of the proof is straightforward

$$\begin{aligned}\exp[:f(\vec{M}):]g(\vec{M}) &= {}_n g(\vec{M}) + [f(\vec{M}),g(\vec{M})] + \tfrac{1}{2}[f(\vec{M}),[f(\vec{M}),g(\vec{M})]] + \cdots \\ &= {}_n g(\vec{M}) + ([f,g])(\vec{M}) + \tfrac{1}{2}[f(\vec{M}),([f,g])(\vec{M})] + \cdots \\ &= {}_n g(\vec{M}) + ([f,g])(\vec{M}) + \tfrac{1}{2}([f,[f,g]])\times(\vec{M}) + \cdots \\ &= {}_n[\exp(:f:)g]\circ(\vec{M}).\end{aligned}$$

In the case of $\vec{M}$ being antisymplectic, the proof is basically the same except that the Poisson bracket changes sign under the transformation, i.e.,

$$[f(\vec{M}),g(\vec{M})] = -([f,g])(\vec{M}). \quad (12)$$

Therefore we obtain

$$\begin{aligned}\exp[-:f(\vec{M}):]g(\vec{M}) &= {}_n g(\vec{M}) - [f(\vec{M}),g(\vec{M})] + \tfrac{1}{2}[f(\vec{M}),[f(\vec{M}),g(\vec{M})]] + \cdots \\ &= {}_n g(\vec{M}) + ([f,g])(\vec{M}) + \tfrac{1}{2}[f(\vec{M}), -([f,g])(\vec{M})] + \cdots \\ &= {}_n g(\vec{M}) + ([f,g])(\vec{M}) + \tfrac{1}{2}([f,[f,g]])(\vec{M}) + \cdots \\ &= {}_n[\exp(:f:)g]\circ(\vec{M}),\end{aligned}$$

which concludes the proof.

Let now $\vec{M}^F$ denote the map of the forward cell, where the superscript denotes that it is transversed in the forward direction. In order to obtain the maps $\vec{M}^R$, $\vec{M}^S$, $\vec{M}^C$ of the R, S, and C cells, respectively, Theorem II.3 has to be used repeatedly.

The reversed cell (R) is the one in which the order of the elements is reversed from that of the forward cell. This means that if a particle enters the forward cell at an initial point $(x_i,a_i,y_i,b_i,\delta_i)$ and exits it at a final point $(x_f,a_f,y_f,b_f,\delta_f)$, a particle which enters the reversed cell at $(x_f,-a_f,y_f,-b_f,\delta_f)$ will exit at $(x_i,-a_i,y_i,-b_i,\delta_i)$. This entails that the map of the reversed cell is

$$\vec{M}^R = (R\vec{I})\circ\vec{M}^{-1}\circ(R^{-1}\vec{I}), \quad (13)$$

where

$$R\vec{I} = \begin{pmatrix}1&0&0&0&0\\0&-1&0&0&0\\0&0&1&0&0\\0&0&0&-1&0\\0&0&0&0&1\end{pmatrix}\begin{pmatrix}x\\a\\y\\b\\\delta\end{pmatrix}. \quad (14)$$

Taking into account the fact that R is apparently antisymplectic, we can obtain the standard form of $\vec{M}^R$

$$\begin{aligned}\vec{M}^R &= (R\vec{I})\circ\vec{M}^{-1}\circ(R^{-1}\vec{I}) \\ &= (R\vec{I})\circ[\exp(:-H:)\vec{I}]\circ(L^{-1}\vec{I})\circ(R^{-1}\vec{I}) \\ &= (R\vec{I})\circ[\exp(:-H:)\vec{I}]\circ(L^{-1}R^{-1}\vec{I}) \\ &= (R\vec{I})\circ\{\exp[:H(L^{-1}R^{-1}\vec{I}):] \times (L^{-1}R^{-1}\vec{I})\} \quad \text{(Theorem II.3)} \\ &= \exp[:H(L^{-1}R^{-1}\vec{I}):][(R\vec{I})\circ(L^{-1}R^{-1}\vec{I})] \quad \text{(Theorem II.2)} \\ &= \exp[:H(L^{-1}R^{-1}\vec{I}):](RL^{-1}R^{-1}\vec{I}). \end{aligned} \quad (15)$$

The switched cell (S) is the mirror image of the forward cell about the $y-z$ plane, i.e.,

$$\vec{M}^S = (S\vec{I})\circ\vec{M}\circ(S^{-1}\vec{I}), \quad (16)$$

where

$$S\vec{I}=\begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ a \\ y \\ b \\ \delta \end{pmatrix}. \tag{17}$$

Since the matrix S is apparently symplectic, we have

$$\begin{aligned}\vec{M}^S&=(S\vec{I})\circ\vec{M}\circ(S^{-1}\vec{I})\\&=(S\vec{I})\circ(L\vec{I})\circ[\exp(:H:)\vec{I}]\circ(S^{-1}\vec{I})\\&=(SL\vec{I})\circ\{\exp[:H(S^{-1}\vec{I}):](S^{-1}\vec{I})\}\\&=\exp[:H(S^{-1}\vec{I}):](SLS^{-1}\vec{I}).\end{aligned} \tag{18}$$

The combined cell (C) is switched and reversed cell, whose map is

$$\vec{M}^C=(S\vec{I})\circ(R\vec{I})\circ\vec{M}^{-1}\circ(R^{-1}\vec{I})\circ(S^{-1}\vec{I}). \tag{19}$$

Due to the fact that matrix SR is antisymplectic, similar to the reversed cell, $\vec{M}^C$ can be brought into the standard form

$$\vec{M}^C=\exp[:H(L^{-1}R^{-1}S^{-1}\vec{I}):](SRL^{-1}R^{-1}S^{-1}\vec{I}). \tag{20}$$

In summary, we have the maps of all four-cell types listed below

$$\vec{M}^F=\exp(:H:)(L\vec{I}), \tag{21}$$

$$\vec{M}^R=\exp[:H(L^{-1}R^{-1}\vec{I}):](RL^{-1}R^{-1}\vec{I}), \tag{22}$$

$$\vec{M}^S=\exp[:H(S^{-1}\vec{I}):](SLS^{-1}\vec{I}), \tag{23}$$

$$\vec{M}^C=\exp[:H(L^{-1}R^{-1}S^{-1}\vec{I}):](SRL^{-1}R^{-1}S^{-1}\vec{I}). \tag{24}$$

Now that the maps of the different kinds of symmetry manipulations of the standard cell are known, the map of any given multicell system can be constructed. In particular, it is always possible to determine the substandard form of the map, which can then be turned into the standard form of the map with the BCH formula.

For the purpose of clear nomenclature, let C_i be the ith cell in a k-cell system, i.e., C_i can be F, R, S, or C. Then the map of the total system is denoted by $\vec{M}^{C_1C_2\cdots C_k}$. For example, $\vec{M}^{FRSC}$ represents the map of a four-cell system consisting of the forward cell, followed by the reversed cell, then by the switched cell and ending with the combined cell.

We will illustrate the process of obtaining the substandard forms for the later particularly important four-cell system $FRSC$, and list the results for some other four-cell systems. In the manipulations, we repeatedly make use of Theorem (II.2), Theorem (II.3), and the associativity of ''$\circ$''.

From the definition of system $FRSC$, we can obtain its transfer map from the maps of single cells as

$$\vec{M}^{FRSC}=\vec{M}^C\circ\vec{M}^S\circ\vec{M}^R\circ\vec{M}^F. \tag{25}$$

Note that the order of the maps of the single cells is the reverse of that of the cells, because the initial coordinates of the present cell are the final coordinates of the previous one. Then we bring $\vec{M}^{FRSC}$ into the substandard form via the following steps of iterative character

$$\begin{aligned}\vec{M}^{FRSC}&=\{\exp[:H(L^{-1}R^{-1}S^{-1}\vec{I}):](SRL^{-1}R^{-1}S^{-1}\vec{I})\}\circ\{\exp[:H(S^{-}\vec{I}):](SLS^{-1}\vec{I})\}\circ\{\exp[:H(L^{-1}R^{-1}\vec{I}):]\\&\quad\times(RL^{-1}R^{-1}\vec{I})\}\circ\{\exp[:H(\vec{I}):](L\vec{I})\}\\&=\exp[:H(\vec{I}):]\{[\exp(:H(L^{-1}R^{-1}S^{-1}\vec{I}):)(SRL^{-1}R^{-1}S^{-1}\vec{I})]\circ[\exp(:H(S^{-1}\vec{I}):)(SLS^{-1}\vec{I})]\circ[\exp(:H(L^{-1}R^{-1}\vec{I}):)\\&\quad\times(RL^{-1}R^{-1}\vec{I})]\circ(L\vec{I})\}\quad\text{(Theorem II.2)}\\&=\exp[:H(\vec{I}):]\{[\exp(:H(L^{-1}R^{-1}S^{-1}\vec{I}):)(SRL^{-1}R^{-1}S^{-1}\vec{I})]\circ[\exp(:H(S^{-1}\vec{I}):)\\&\quad\times(SLS^{-1}\vec{I})]\circ[\exp(:H(L^{-1}R^{-1}\cdot L\vec{I}):)(RL^{-1}R^{-1}\cdot L\vec{I})]\}\quad\text{(Theorem II.3)}\cdots\\&=\exp[:H(\vec{I}):]\exp[:H(L^{-1}R^{-1}\cdot L\vec{I}):]\exp[:H(S^{-1}\cdot RL^{-1}R^{-1}\cdot L\vec{I}):]\exp[:H(L^{-1}R^{-1}\cdot LS^{-1}\cdot RL^{-1}R^{-1}\cdot L\vec{I}):]\\&\quad\times(SRL^{-1}R^{-1}\cdot LS^{-1}\cdot RL^{-1}R^{-1}\cdot L\vec{I}).\end{aligned}$$

In a very similar way, the substandard forms of the maps of the systems $\vec{M}^{FRFR}$, $\vec{M}^{FCSR}$, and $\vec{M}^{FCFC}$ are obtained, and altogether we have

$$\begin{aligned}\vec{M}^{FRSC}&=\exp[:H(\vec{I}):]\exp[:H(L^{-1}R^{-1}\cdot L\vec{I}):]\exp[:H(S^{-1}\cdot RL^{-1}R^{-1}\cdot L\vec{I}):]\exp[:H(L^{-1}R^{-1}\cdot LS^{-1}\cdot RL^{-1}R^{-1}\cdot L\vec{I}):]\\&\quad\times(SRL^{-1}R^{-1}\cdot LS^{-1}\cdot RL^{-1}R^{-1}\cdot L\vec{I}),\end{aligned} \tag{26}$$

$$\begin{aligned}\vec{M}^{FRFR}&=\exp[:H(\vec{I}):]\exp[:H(L^{-1}R^{-1}\cdot L\vec{I}):]\exp[:H(RL^{-1}R^{-1}\cdot L\vec{I}):]\exp[:H(L^{-1}R^{-1}\cdot L\cdot RL^{-1}R^{-1}\cdot L\vec{I}):]\\&\quad\times(RL^{-1}R^{-1}\cdot L\cdot RL^{-1}R^{-1}\cdot L\vec{I}),\end{aligned} \tag{27}$$

$$\vec{M}^{FCSR}=\exp[:H(\vec{I}):]\exp[:H(L^{-1}R^{-1}S^{-1}\cdot L\vec{I}):]\exp[:H(RL^{-1}R^{-1}S^{-1}\cdot L\vec{I}):]\exp[:H(L^{-1}R^{-1}\cdot SL\cdot RL^{-1}R^{-1}S^{-1}\cdot L\vec{I}):]$$
$$\times(RL^{-1}R^{-1}\cdot SL\cdot RL^{-1}R^{-1}S^{-1}\cdot L\vec{I}), \tag{28}$$

$$\vec{M}^{FCFC}=\exp[:H(\vec{I}):]\exp[:H(L^{-1}R^{-1}S^{-1}\cdot L\vec{I}):]\exp[:H(SRL^{-1}R^{-1}S^{-1}\cdot L\vec{I}):]$$
$$\times\exp[:H(L^{-1}R^{-1}S^{-1}\cdot L\cdot SRL^{-1}R^{-1}S^{-1}\cdot L\vec{I}):](SRL^{-1}R^{-1}S^{-1}\cdot L\cdot SRL^{-1}R^{-1}S^{-1}\cdot L\vec{I}). \tag{29}$$

As shown in the next section, only the four systems listed here are needed when the solutions of arbitrary-order achromats are determined, because other systems are not as efficient. What will also be shown here is the importance of the substandard form where the optimal four-cell systems are decided. Finally, when solutions of achromats are searched for among the four systems, the standard form of their maps will be obtained from the substandard form using the BCH formula.

III. OPTIMAL FOUR-CELL SYSTEMS

We will now proceed to the analysis of the influence of symmetric arrangements to the cancellation of aberrations in general multicell systems, and then later study the systems found optimal. We begin with a few definitions.

Like in previous achromat theories, we consider only those systems with midplane symmetry. Therefore, the transfer map of the forward cell can be represented in the form of Eq. (9) with its pseudo-Hamiltonian given by

$$H=\sum_{i_x i_a i_y i_b i_\delta} C_{i_x i_a i_y i_b i_\delta} x^{i_x} a^{i_a} y^{i_y} b^{i_b} \delta^{i_\delta}, \tag{30}$$

where $i_x+i_a+i_y+i_b+i_\delta \geqslant 3$, and i_y+i_b is even.

Definition III.1. Let H be the pseudo-Hamiltonian of the system, then we set

$$A(H)=\sum_{i_x i_a i_y i_b i_\delta} C_{i_x i_a i_y i_b i_\delta} x^{i_x} a^{i_a} y^{i_y} b^{i_b} \delta^{i_\delta}$$
$$(i_x+i_a \text{ is odd}, \ i_a+i_b \text{ is even}), \tag{31}$$

$$B(H)=\sum_{i_x i_a i_y i_b i_{gd}} C_{i_x i_a i_y i_b i_\delta} x^{i_x} a^{i_a} y^{i_y} b^{i_b} \delta^{i_\delta}$$
$$(i_x+i_a \text{ is odd}, \ i_a+i_b \text{ is odd}), \tag{32}$$

$$C(H)=\sum_{i_x i_a i_y i_b i_\delta} C_{i_x i_a i_y i_b i_\delta} x^{i_x} a^{i_a} y^{i_y} b^{i_b} \delta^{i_\delta}$$
$$(i_x+i_a \text{ is even}, \ i_a+i_b \text{ is odd}), \tag{33}$$

$$D(H)=\sum_{i_x i_a i_y i_b i_\delta} C_{i_x i_a i_y i_b i_\delta} x^{i_x} a^{i_a} y^{i_y} b^{i_b} \delta^{i_\delta}$$
$$(i_x+i_a \text{ is even}, \ i_a+i_b \text{ is even}). \tag{34}$$

Definition III.2. Define

$$H^F=H(\vec{I})=H(x,a,y,b), \tag{35}$$

$$H^R=H(R\vec{I})=H(x,-a,y,-b), \tag{36}$$

$$H^S=H(S\vec{I})=H(-x,-a,y,b), \tag{37}$$

$$H^C=H(RS\vec{I})=H(-x,a,y,-b). \tag{38}$$

It is easy to show that in decomposed form one can write

$$H^F=A(H)+B(H)+C(H)+D(H), \tag{39}$$

$$H^R=A(H)-B(H)-C(H)+D(H), \tag{40}$$

$$H^S=-A(H)-B(H)+C(H)+D(H), \tag{41}$$

$$H^C=-A(H)+B(H)-C(H)+D(H). \tag{42}$$

A. General properties of *k*-cell systems

Consider a general system of k cells arranged using the above symmetry operations. Using Theorems (II.2) and II.3 repeatedly, its map can be brought to the substandard form in a similar way as in Eq. (26). The result has the form

$$\vec{M}=\exp[:H(\vec{I}):]\exp[:H(M^{(1)}\vec{I}):]\cdots\exp[:H(M^{(k-1)}\vec{I}):]$$
$$\times(M_T\vec{I}), \tag{43}$$

where M_T is the linear matrix of the system and

$$M^{(i)}=\begin{pmatrix} m_{11}^{(i)} & m_{12}^{(i)} & 0 & 0 & m_{15}^{(i)} \\ m_{11}^{(i)} & m_{12}^{(i)} & 0 & 0 & m_{15}^{(i)} \\ 0 & 0 & m_{33}^{(i)} & m_{34}^{(i)} & 0 \\ 0 & 0 & m_{43}^{(i)} & m_{44}^{(i)} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$
$$(i=1,2,\ldots,k-1) \tag{44}$$

is a midplane symmetrical matrix obtained from combinations of the linear matrices of the previous cells and matrices R,R^{-1}, S,S^{-1}, and C,C^{-1} depending on the specific choices of the system. As a result, we have $\det(M^{(i)})=1$.

Using the BCH formula, $\vec{M}$ can be transformed to a single Lie operator acting on a linear map, which is

$$\vec{M} = {}_n \exp[:H(\vec{I}) + H(M^{(1)}\vec{I}) + \cdots + H(M^{(k-1)}\vec{I}) + \text{commutators:}](M_T\vec{I}). \tag{45}$$

In order to achieve achromaticity, it is necessary for the expression between the colons to vanish. Since all commutators are of order four and higher, this entails, in particular, that it is necessary for all the third-order terms in the sum of pseudo-Hamiltonians $\Sigma H(M^{(i)}\vec{I})$ to vanish outright; for higher orders, the terms stemming from the commutators have to balance those from the sum. In the remainder, we shall study several approaches to achieve this.

Next let us show that there is no system that can cancels $D(H(\vec{I}))+D(H(M^{(1)}\vec{I}))+\cdots+D(H(M^{(k-1)}\vec{I}))$ merely by symmetry without making specific choices for the nonlinear settings. First, $D(H(\vec{I}))$ is split into two parts:

Definition III.3. $D^+(H(\vec{I}))$ is defined as the terms in $D(H(\vec{I}))$ with all exponents on x,a,y,b even, which has the form

$$\sum_{n_x,n_a,n_y,n_b,i_\delta} C_{2n_x,2n_a,2n_y,2n_b,i_\delta} x^{2n_x} a^{2n_a} y^{2n_y} b^{2n_b} \delta^{i_\delta}.$$

$D^-(H(\vec{I}))$ is defined as the terms in $D(H(\vec{I}))$ with all exponents on x,a,y,b odd, which has the form

$$\sum_{n_x,n_a,n_y,n_b,i_\delta} C_{2n_x+1,2n_a+1,2n_y+1,2n_b+1,i_\delta} \times x^{2n_x+1} a^{2n_a+1} y^{2n_y+1} b^{2n_b+1} \delta^{i_\delta}.$$

Note that $D=D^+ + D^-$ because of Eq. (34). We then observe

Theorem III.1. For a general k-cell system, it is impossible to cancel any term from $D^+(H(\vec{I})) + D^+(H(M^{(1)}\vec{I})) + \cdots + D^+(H(M^{(k-1)}\vec{I}))$ solely by the symmetrical arrangements of the cells and the choices of special linear matrices without imposing specific relations among the Lie coefficients $C_{i_x i_a i_y i_b i_\delta}$.

Proof: From Eqs. (43) and (44), the sum of $H(M^{(i)}\vec{I})$ is

$$\begin{aligned} &H(\vec{I}) + H(M^{(1)}\vec{I}) + \cdots + H(M^{(k-1)}\vec{I}) \\ &= \sum_{i_x,i_a,i_y,i_b,i_\delta} C_{i_x i_a i_y i_b i_\delta} [x^{i_x} a^{i_a} y^{i_y} b^{i_b} \delta^{i_\delta} + (m_{11}^{(1)} x + m_{12}^{(1)} a \\ &\quad + m_{15}^{(1)} \delta)^{i_x} (m_{21}^{(1)} x + m_{22}^{(1)} a + m_{25}^{(1)} \delta)^{i_a} (m_{33}^{(1)} y \\ &\quad + m_{34}^{(1)} b)^{i_y} (m_{43}^{(1)} y + m_{44}^{(1)} b)^{i_y} \delta^{i_\delta} + \cdots + (m_{11}^{(k-1)} x \\ &\quad + m_{12}^{(k-1)} a + m_{15}^{(k-1)} \delta)^{i_x} (m_{21}^{(k-1)} x + m_{22}^{(k-1)} a \\ &\quad + m_{25}^{(k-1)} \delta)^{i_a} (m_{33}^{(k-1)} y + m_{34}^{(k-1)} b)^{i_y} (m_{43}^{(k-1)} y \\ &\quad + m_{44}^{(k-1)} b)^{i_y} \delta^{i_\delta}], \end{aligned}$$

which entails that the sum of D^+ is

$$\begin{aligned} &D^+(H(\vec{I})) + D^+(H(M^{(1)}\vec{I})) + \cdots + D^+(H(M^{(k-1)}\vec{I})) \\ &= \sum_{n_x,n_a,n_y,n_b,i_\delta} C_{2n_x,2n_a,2n_y,2n_b,i_\delta} [x^{2n_x} a^{2n_a} y^{2n_y} b^{2n_b} \delta^{i_\delta} \\ &\quad + (m_{11}^{(1)} x + m_{12}^{(1)} a + m_{15}^{(1)} \delta)^{2n_x} (m_{21}^{(1)} x + m_{22}^{(1)} a \\ &\quad + m_{25}^{(1)} \delta)^{2n_a} (m_{33}^{(1)} y + m_{34}^{(1)} b)^{2n_y} (m_{43}^{(1)} y + m_{44}^{(1)} b)^{2n_b} \delta^{i_\delta} \\ &\quad + \cdots + (m_{11}^{(k-1)} x + m_{12}^{(k-1)} a + m_{15}^{(k-1)} \delta)^{2n_x} (m_{21}^{(k-1)} x \\ &\quad + m_{22}^{(k-1)} a + m_{25}^{(k-1)} \delta)^{2n_a} (m_{33}^{(k-1)} y \\ &\quad + m_{34}^{(k-1)} b)^{2n_y} (m_{43}^{(k-1)} y + m_{44}^{(k-1)} b)^{2n_b} \delta^{i_\delta}]. \end{aligned}$$

Since by requirement, no specific relationships among the $C_{2n_x,2n_a,2n_y,2n_b,i_\delta}$ are allowed to be imposed, the vanishing of a polynomial associated with $C_{2n_x,2n_a,2n_y,2n_b,i_\delta}$ requires that

$$\begin{aligned} &x^{2n_x} a^{2n_a} y^{2n_y} b^{2n_b} + \sum_{i=1}^{k-1} (m_{11}^{(i)} x + m_{12}^{(i)} a + m_{15}^{(i)} \delta)^{2n_x} (m_{21}^{(i)} x \\ &\quad + m_{22}^{(i)} a + m_{25}^{(i)} \delta)^{2n_a} (m_{33}^{(i)} y + m_{34}^{(i)} b)^{2n_y} (m_{43}^{(i)} y + m_{44}^{(i)} b)^{2n_b} \\ &= 0, \end{aligned}$$

which, because of the even exponents, is impossible to achieve for all points in phase space, regardless of the choice of the $M^{(i)}$.

B. Two- and three-cell systems

The next theorem shows that two- or three-cell systems cannot give optimal solutions for achromats.

Theorem III.2. Two- or three-cell systems cannot cancel $A_3(H(\vec{I}))+A_3(H(M^{(1)}\vec{I}))+\cdots+A_3(H(M^{(k-1)}\vec{I}))$, $B_3(H(\vec{I})) + B_3(H(M^{(1)}\vec{I}))+\cdots+B_3(H(M^{(k-1)}\vec{I}))$ and $C_3(H(\vec{I})) + C_3(H(M^{(1)}\vec{I}))+\cdots+C_3(H(M^{(k-1)}\vec{I}))$, and hence they cannot give optimal systems of achromats.

Proof: Let us first consider two-cell systems. The sum of $H(M^{(i)}\vec{I})$ is

$$\begin{aligned} H(\vec{I}) + H(M^{(1)}\vec{I}) = &\sum_{i_x,i_a,i_y,i_b,i_\delta} C_{i_x i_a i_y i_b i_\delta} [x^{i_x} a^{i_a} y^{i_y} b^{i_b} \delta^{i_\delta} \\ &+ (m_{11}^{(1)} x + m_{12}^{(1)} a + m_{15}^{(1)} \delta)^{i_x} (m_{21}^{(1)} x \\ &+ m_{22}^{(1)} a + m_{25}^{(1)} \delta)^{i_a} (m_{33}^{(1)} y \\ &+ m_{34}^{(1)} b)^{i_y} (m_{43}^{(1)} y + m_{44}^{(1)} b)^{i_y} \delta^{i_\delta}]. \end{aligned}$$

Cancellation of the terms associated with $C_{1,0,0,0,2}$ from $A_3(H(\vec{I}))+A_3(H(M^{(1)}\vec{I}))$ entails that

$$C_{1,0,0,0,2}[x\delta^2 + (m_{11}^{(1)} x + m_{12}^{(1)} a + m_{15}^{(1)} \delta)\delta^2] = 0.$$

Since all coefficients are independent of each other, each term in the above equation has to vanish separately, which gives the solution

$$m_{11}^{(1)}=-1,$$

$$m_{12}^{(1)}=0,$$

$$m_{15}^{(1)}=0;$$

and canceling the terms associated with $C_{0,1,0,0,2}$ from $B_3(H(\vec{I}))+B_3(H(M^{(1)}\vec{I}))$ entails that

$$C_{0,1,0,0,2}[a\,\delta^2+(m_{21}^{(1)}x+m_{22}^{(1)}a+m_{25}^{(1)}\delta)\,\delta^2]=0,$$

which has the solution

$$m_{22}^{(1)}=-1,$$

$$m_{21}^{(1)}=0,$$

$$m_{25}^{(1)}=0.$$

Considering the terms associated with $C_{1,1,0,0,1}$ from $C_3(H(\vec{I}))+C_3(H(M^{(1)}\vec{I}))$, we have

$$C_{1,1,0,0,1}(xa\,\delta+m_{11}^{(1)}m_{22}^{(1)}xa\,\delta)=2C_{1,1,0,0,1}xa\,\delta,$$

which shows that $C_3(H(\vec{I}))+C_3(H(M^{(1)}\vec{I}))$ cannot be canceled. The same conclusion can be reached for three-cell systems in a conceptually similar way that is, however, technically too involved to be exhibited here; for details, we refer to [20].

C. Four-cell systems

In this section we will show that certain four-cell systems can achieve what two- and three-cell systems could not, i.e., they can be brought in symmetrical arrangements to cancel A, B, and C simultaneously.

Theorem III.3. Given a four-cell system, the terms $\Sigma_{i=0}^{3}A_n(H(M^{(i)}\vec{I}))$, $\Sigma_{i=0}^{3}B_n(H(M^{(i)}\vec{I}))$, and $\Sigma_{i=0}^{3}C_n(H(M^{(i)}\vec{I}))$ ($M^{(0)}=\hat{I}$) are canceled for all choices of n, if and only if $H(M^{(1)}\vec{I})$, $H(M^{(2)}\vec{I})$, and $H(M^{(3)}\vec{I})$ equal a permutation of H^R, H^S, and H^C.

Proof: The sufficiency is obvious from Eqs. (39)–(42). While the detailed proof of the necessity is beyond the scope of this paper and the reader is referred to [20], the basic idea behind it is quite simple and shall be illustrated here. The cancellation of the terms has to be achieved through proper choices of the linear matrix as well as the symmetric arrangements. By studying the coefficients of $\Sigma_{i=0}^{3}A_3(H(M^{(i)}\vec{I}))$, $\Sigma_{i=0}^{3}B_3(H(M^{(i)}\vec{I}))$, and $\Sigma_{i=0}^{3}C_3(H(M^{(i)}\vec{I}))$ term by term, one obtains a set of equations for the coefficients of the linear matrix, which, because of the presence of midplane symmetry, however, decouples into a variety of subblocks. Solving the resulting set of equations, which is a very tedious yet rather mechanical process, then shows that there is one and only one solution: the $M^{(i)}$ ($i=2,3,4$) must be a permutation of R, S, and C.

Since it is shown earlier that it is not possible to cancel terms in D^+, the ability to cancel the A, B, and C terms is essentially the optimum of what can be achieved, except for the question of whether or not the D^- terms can also be canceled automatically. However, in practice this is of secondary importance since the number of terms in D^- is much smaller than that in D^+: there is only one term in $D_3^-(H(\vec{I}))$ and $D_4^-(H(\vec{I}))$ as opposed to 15 terms in $D_3(H(\vec{I}))$ and $D_4(H(\vec{I}))$; there are only five terms in $D_5^-(H(\vec{I}))$ and $D_6^-(H(\vec{I}))$, as opposed to 39 terms in $D_5(H(\vec{I}))$ and $D_6(H(\vec{I}))$.

D. The optimal four-cell systems

The previous discussion revealed that the only way to achieve maximum cancellation from symmetry is that the $M^{(i)}$ for $i=2,3,4$ form a permutation of R, S, and C. Since, however, the $M^{(i)}$ contain both contributions from the linear maps as well as from the symmetry operations, it requires some study to determine the systems that satisfy the requirement. The following theorem answers these questions.

Theorem III.4. Among all (64) four-cell systems, there are only four which reach the optimum asserted by Theorem III.3 while imposing the minimum number of constraints on the linear map. They are $FRFR$, $FRSC$, $FCFC$, and $FCSR$.

Proof: While being rather straightforward and mechanical, the entire proof is rather tedious, and so we restrict ourselves here to the first part which leads to the first possible solution of the form $FRFR$.

Suppose the forward cell has a linear matrix L. For the time being, let us restrict ourselves to the x-a-δ block of L, which has the form

$$L_1=\begin{pmatrix} a & b & \eta \\ c & d & \eta' \\ 0 & 0 & 1 \end{pmatrix},$$

$$L_1^{-1}=\begin{pmatrix} d & -b & -d\,\eta+b\,\eta' \\ -c & a & c\,\eta-a\,\eta' \\ 0 & 0 & 1 \end{pmatrix}. \tag{46}$$

(1) Choices on the second cell.

Case 1a FF

Recall that the standard form of a forward cell [Eq. (9)] is

$$\vec{M}^F=\exp(:H:)(L\vec{I}).$$

Therefore, the map of the system FF is

$$\begin{aligned}\vec{M}^{FF}&=\vec{M}^F\circ\vec{M}^F\\&=\{\exp[:H(\vec{I}):](L\vec{I})\}\circ\{\exp[:H(\vec{I}):](L\vec{I})\}\\&=\exp[:H(\vec{I}):](\{\exp[:H(\vec{I}):](L\vec{I})\}\circ(L\vec{I}))\\&=\exp[:H(\vec{I}):]\exp[:H(L\vec{I}):](L\cdot L\vec{I}).\end{aligned}$$

To reach the optimum, L has to be R, S, or C. Since L is symplectic, it can only be S. Hence, the linear matrix is

$$L_1=\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which entails that the conditions of reaching the optimum are

$$\eta=0,$$

$$\eta'=0,$$

$$c=0,$$

$$b=0,$$

$$a=-1.$$

Therefore, five conditions have to be met to reach the optimum.

Case 1b FR: From Eq. (15), we have

$$\begin{aligned}\vec{M}^{FR}&=\vec{M}^{R}\circ\vec{M}^{F}\\&=\{\exp[:H(L^{-1}R^{-1}\vec{I}):]\\&\quad\times(RL^{-1}R^{-1}\vec{I})\}\circ\{\exp[:H(\vec{I}):](L\vec{I})\}\\&=\exp[:H(\vec{I}):]\exp[:H(L^{-1}R^{-1}\cdot L\vec{I}):]\\&\quad\times(RL^{-1}R^{-1}\cdot L\vec{I}).\end{aligned}$$

Specifically, $L_1^{-1}R_1^{-1}\cdot L_1$ can be obtained from Eq. (46), which yields

$$\begin{aligned}L_1^{-1}R_1^{-1}\cdot L_1&=\begin{pmatrix}d&-b&-d\eta+b\eta'\\-c&a&c\eta-a\eta'\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&-1&0\\0&0&1\end{pmatrix}\\&\quad\times\begin{pmatrix}a&b&\eta\\c&d&\eta'\\0&0&1\end{pmatrix}\\&=\begin{pmatrix}ad+bc&2bd&2b\eta'\\-2ac&-(ad+bc)&-2a\eta'\\0&0&1\end{pmatrix}.\end{aligned}\tag{47}$$

Since $L^{-1}R^{-1}\cdot L$ is antisymplectic, it can only be R or C, which leads to the following conditions:

$$bd=0$$

$$ac=0$$

$$b\eta'=0$$

$$a\eta'=0,$$

which are equivalent to

$$\begin{array}{llll}b=0 & & a=0\\c=0 & \text{or} & d=0\\\eta'=0 & & \eta'=0.\end{array}\tag{48}$$

Hence, we obtained two solutions with three conditions, which in turn eliminate the five-condition solution above. For further reference, they are listed below

$$\begin{array}{cc}\text{Solution A:} & \text{Solution B:}\\L_1=\begin{pmatrix}a&0&\eta\\0&d&0\\0&0&1\end{pmatrix}\quad\text{and}\quad & L_1=\begin{pmatrix}0&b&\eta\\c&0&0\\0&0&1\end{pmatrix}.\end{array}\tag{49}$$

(2) Choices for the third cell of the system $FR\times\times$

From Case 1b, we have

$$\vec{M}^{FR}=\exp[:H(\vec{I}):]\exp[:H(L^{-1}R^{-1}L\vec{I}):](RL^{-1}R^{-1}L\vec{I}).$$

Let $\vec{M}^{\times}=\{\exp[:H(M(\times)\vec{I}]:\}[L(\times)\vec{I}]$ be the map of the third cell, which can be F, R, C, or D. Thus, the total map of the three-cell system is

$$\begin{aligned}\vec{M}^{FR\times}&=\vec{M}^{\times}\circ\vec{M}^{FR}\\&=(\exp\{:H[M(\times)\vec{I}]:\}[L(\times)\vec{I}])\circ\{\exp[:H(\vec{I}):]\\&\quad\times\exp[:H(L^{-1}R^{-1}L\vec{I}):](RL^{-1}R^{-1}L\vec{I})\}\\&=\exp[:H(\vec{I}):]\exp[:H(L^{-1}R^{-1}L\vec{I}):]\\&\quad\times[(\exp\{:H[M(\times)\vec{I}]:\}\\&\quad\times[L(\times)\vec{I}])\circ(RL^{-1}R^{-1}L\vec{I})]\\&=\exp[:H(\vec{I}):]\exp[:H(L^{-1}R^{-1}L\vec{I}):]\\&\quad\times\exp\{:H[M(\times)RL^{-1}R^{-1}L\vec{I}]:\}\\&\quad\times[L(\times)RL^{-1}R^{-1}L\vec{I}].\end{aligned}$$

For our convenience, let us define $M^{(2)}(\times)$ as the linear matrix in the pseudo-Hamiltonian for the third cell. For systems $FR\times$, we have

$$M^{(2)}(\times)=M(\times)RL^{-1}R^{-1}L=M(\times)L^{FR},$$

where $L^{FR}=RL^{-1}R^{-1}L$.

Case 2a FRF: Since $M(F)=\hat{I}$, we have $M^{(2)}(F)=RL^{-1}R^{-1}L$. For the two solutions, the x-a-δ block $M^{(2)}(F)$ are listed below. In the case of Solution A, we have

$$M_1^{(2)}(F)=\begin{pmatrix}1&0&0\\0&-1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&-1&0\\0&0&1\end{pmatrix}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix},$$

which does not reach the optimum because it does not satisfy Theorem III.3. In the case of Solution B, we have

$$\begin{aligned}M_1^{(2)}(F)&=\begin{pmatrix}1&0&0\\0&-1&0\\0&0&1\end{pmatrix}\begin{pmatrix}-1&0&0\\0&1&0\\0&0&1\end{pmatrix}\\&=\begin{pmatrix}-1&0&0\\0&-1&0\\0&0&1\end{pmatrix},\end{aligned}$$

which is a possible solution, because it satisfies Theorem III.3 and does not need any more conditions.

(3) Choices for the fourth cells of the systems $FRF\times$

Define $L^{FRF}=LRL^{-1}R^{-1}L$. Similar to case (2), we have

$$M^{(3)}=M(\times)L^{FRF}.$$

Since solution B is the possible solution for this case, the linear matrix of the forward cell is

$$L_1=\begin{pmatrix}0&b&\eta\\c&0&0\\0&0&1\end{pmatrix}.$$

Case 3a FRFF: From $M(F)=\hat{I}$, we have

TABLE II. Optimal four-cell systems and the first-order requirements to achieve their optimum.

Systems	Linear Conditions
FRSC	$(a\|\delta)=0,\ (x\|a)=(a\|x)=0$
FRFR	$(a\|\delta)=0,\ (x\|x)=(a\|a)=0$
FCSR	$(x\|\delta)=0,\ (x\|a)=(a\|x)=0$
FCFC	$(x\|\delta)=0,\ (x\|x)=(a\|a)=0$

$$M_1^{(3)}=LRL^{-1}R^{-1}L.$$

Since

$$R_1L_1^{-1}R_1^{-1}L_1=\begin{pmatrix}-1&0&0\\0&-1&0\\0&0&1\end{pmatrix},$$

we have $M_1^{(3)}=M_1^{(3)}$. Therefore this system is not a solution.

Case 3b FRFR: From $M(R)=L^{-1}R^{-1}$, we have

$$M^{(3)}=L^{-1}R^{-1}LRL^{-1}R^{-1}L.$$

Therefore, the x-a-δ block is

$$M_1^{(3)}(C)=\begin{pmatrix}-1&0&0\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}-1&0&0\\0&-1&0\\0&0&1\end{pmatrix}=\begin{pmatrix}1&0&0\\0&-1&0\\0&0&1\end{pmatrix},$$

which shows that this system is a solution, because it satisfies Theorem III.3.

In total, this concludes the analysis of the path leading to the solution $FRFR$, and we have shown that as needed, it has the form

$$\vec{M}^{FRFR}=\exp(:H^F:)\exp(:H^C:)\exp(:H^S:)\exp(:H^R:)\vec{I}. \tag{50}$$

The next theorem, given without proof, specifies the linear conditions of the y-b block for the optimal systems.

Theorem III.5. For the four systems obtained from the last theorem, the constraints on the y-b block of the linear map are the vanishing of either the diagonal elements or the off-diagonal elements.

To summarize the results of this section, the four optimal systems as well as the linear conditions necessary to satisfy them are listed in Table II.

IV. ORDER-BY-ORDER SOLUTIONS

As the next step, it is necessary to analyze the commutators of lower order terms, in particular, study their interplay with direct sum terms that can never be removed from symmetry alone. The study is simplified by two observations. First of all, the four different maps yielding optimal symmetry cancellation only produce two different substandard forms, namely,

$$\vec{M}^{FRSC}=\vec{M}^{FCFC}=\exp(:H^F:)\exp(:H^R:)\exp(:H^S:)\exp(:H^C:)\vec{I} \tag{51}$$

and

$$\vec{M}^{FCSR}=\vec{M}^{FRFR}=\exp(:H^F:)\exp(:H^C:)\exp(:H^S:)\exp(:H^R:)\vec{I}. \tag{52}$$

Furthermore, one of these two substandard forms can be obtained from the other through a simple transformation. Thus, it is only necessary to study one particular map to determine the conditions for achromaticity.

Lemma IV.1. By switching A and B, the maps $\vec{M}^{FCSR}$ and $\vec{M}^{FCFC}$ are transformed to $\vec{M}^{FRSC}$ and $\vec{M}^{FRFR}$, respectively.

Proof: Under the transformation, we have

$$H^F=A+B+C+D\to B+A+C+D=H^F,$$

$$H^R=A-B-C+D\to B-A-C+D=H^C,$$

$$H^S=-A-B+C+D\to -B-A+C+D=H^S,$$

$$H^C=-A+B-C+D\to -B+A-C+D=H^R,$$

which entails that the transformation of the maps is

$$\vec{M}^{FRSC,FCFC}\to\vec{M}^{FCSR,FRFR},$$

which concludes the proof.

We are now ready for a discussion of various strategies to design achromats. The first method is described by the following theorem, which allows the design of achromats of an arbitrary order, although not in the most economical way.

Theorem IV.1. For a given order n, the optimal systems $FRFR$, $FRSC$, $FCSR$, and $FCFC$ are achromats if $A=_n 0$ or $B=_n 0$, and $D=_{n+1}0$.

So in this case, there are no constraints at all on C terms, and no constraints on A and B terms of the highest order either.

Proof: Let us first consider systems $FRSC$ and $FCFC$. From the proof of Theorem (III.4), they have the same maps, which are

$$\vec{M}=\exp(:H^F:)\exp(:H^R:)\exp(:H^S:)\exp(:H^C:)\vec{I}.$$

Define $H_n=\sum_{i=3}^{n}f_i$ and $f_i=A_i+B_i+C_i+D_i$, where f_i is the sum of the ith order terms in H.

From now we proceed by induction over the order, beginning with the second order. Using the BCH formula, $\vec{M}$ can be transformed to the ''standard'' form, which is

$$\begin{aligned}\vec{M}&=_2\exp(:H_3^F:)\exp(:H_3^R:)\exp(:H_3^S:)\exp(:H_3^C:)\vec{I}\\&=_2\exp(:f_3^F:)\exp(:f_3^R:)\exp(:f_3^S:)\exp(:f_3^C:)\vec{I}\\&=_2\exp(:f_3^F+f_3^R+f_3^S+f_3^C:)\vec{I}\\&=_2\exp(:4D_3:)\vec{I}.\end{aligned}$$

Therefore the second-order solution is $D=_3 0$, as advertised.

To perform the induction step to show the condition to nth order, let us assume that $A=_{n-1}0$ or $B=_{n-1}0$, and $D=_n 0$ are solutions for an $(n-1)$st-order achromat. Since the proof of case $B=_{n-1}0$ is the same as that of $A=_{n-1}0$, only the first case is shown.

In this case, we have

$$H^F_{n-1}=\sum_{i=3}^{n-1}(B_i+C_i),$$

$$H^R_{n-1}=\sum_{i=3}^{n-1}(-B_i-C_i),$$

$$H^S_{n-1}=\sum_{i=3}^{n-1}(-B_i+C_i),$$

$$H^C_{n-1}=\sum_{i=3}^{n-1}(B_i-C_i),$$

which entails that

$$H^F_{n-1}+H^R_{n-1}=H^S_{n-1}+H^C_{n-1}=0$$

and

$$[H^F_{n-1},H^R_{n-1}]=[H^S_{n-1},H^C_{n-1}]=0.$$

Therefore $\vec{M}$ can be transformed to

$$\begin{aligned}\vec{M}&=_n \exp(:H^F_{n+1}:)\exp(:H^R_{n+1}:)\exp(:H^S_{n+1}:)\exp(:H^C_{n+1}:)\vec{I}=_n \exp(:H^F_{n-1}+f^F_n+f^F_{n+1}:)\exp(:H^R_{n-1}+f^R_n\\
&+f^R_{n+1}:)\exp(:H^S_{n-1}+f^S_n+f^S_{n+1}:)\exp(:H^C_{n-1}+f^C_n+f^C_{n+1}:)\vec{I}=_n \exp\{:f^F_n+f^F_{n+1}+f^R_n+f^R_{n+1}+\tfrac{1}{2}([H^F_{n-1},f^R_n]\\
&+[f^F_n,H^R_{n-1}]):\}\exp\{:f^S_n+f^S_{n+1}+f^C_n+f^C_{n+1}+\tfrac{1}{2}([H^S_{n-1},f^C_n]+[f^S_n,H^C_{n-1}]):\}\vec{I}=_n \exp\{:f^F_n+f^F_{n+1}+f^R_n+f^R_{n+1}\\
&+\tfrac{1}{2}([f^F_3,f^R_n]+[f^F_n,f^R_3]):\}\exp\{:f^S_n+f^S_{n+1}+f^C_n+f^C_{n+1}+\tfrac{1}{2}([f^S_3,f^C_n]+[f^S_n,f^C_3]):\}\vec{I}=_n \exp\{:4D_{n+1}+\tfrac{1}{2}([f^F_3,f^R_n]+[f^F_n,f^R_3]\\
&+[f^S_3,f^C_n]+[f^S_n,f^C_3]):\}\vec{I}=_n \exp\{:4D_{n+1}+\tfrac{1}{2}(-[f^R_3,f^R_n]+[f^F_n,f^R_3]-[f^C_3,f^C_n]+[f^S_n,f^C_3]):\}\vec{I}=_n \exp\{:4D_{n+1}+\tfrac{1}{2}([f^R_n\\
&+f^F_n,f^R_3]+[f^C_n+f^S_n,f^C_3]):\}\vec{I}=_n \exp(:4D_{n+1}+[A_n,-B_3-C_3]+[-A_n,B_3-C_3]:)\vec{I}=_n \exp(:4D_{n+1}+2[B_3,A_n]:)\vec{I},\end{aligned}$$

which shows that $A_n=0$ and $D_{n+1}=0$ are a solution for the nth order achromaticity. According to Lemma IV.1, the systems $FRFR$ and $FCSR$ have the same solutions, which concludes the proof. Note that these two solutions are equivalent in terms of the number of conditions, because the number of monomials in A is the same as that in B for a given order n ([20]).

In order to study other strategies to obtain achromats, we remind ourselves that since D cannot be canceled by symmetry, the number of nonlinear conditions cannot be smaller than the number of terms in D. So the best solution under this theory will be that the number of nonlinear conditions equals that of the terms in D.

The next theorem shows that such best solutions exist for up to the fourth order. Furthermore, computational results strongly suggest that indeed the method is also applicable for fifth- and sixth-order achromats, although the tediousness of the algebra has so far prevented us from performing a detailed proof.

Before studying the theorem itself, a few observations are necessary. Considering the commutator of two general terms from the Hamiltonian, we have

$$\begin{aligned}&[C_{i_x i_a i_y i_b i_\delta}x^{i_x}a^{i_a}y^{i_y}b^{i_b}\delta^{i_\delta},C_{i'_x i'_a i'_y i'_b i'_\delta}x^{i'_x}a^{i'_a}y^{i'_y}b^{i'_b}\delta^{i'_\delta}]\\
&\quad=C_{i_x i_a i_y i_b i_\delta}C_{i'_x i'_a i'_y i'_b i'_\delta}[(i_x i'_a\\
&\quad\quad-i_a i'_x)x^{i_x+i'_x-1}a^{i_a+i'_a-1}y^{i_y+i'_y}b^{i_b+i'_b}\delta^{i_\delta+i'_\delta}\\
&\quad\quad+(i_y i'_b-i_b i'_y)x^{i_x+i'_x}a^{i_a+i'_a}y^{i_y+i'_y-1}b^{i_b+i'_b-1}\delta^{i_\delta+i'_\delta}].\end{aligned}$$

If the two terms are from the same part of the Hamiltonian, we have

$$(i_x+i'_x-1)+(i_a+i'_a-1)=(i_x+i_a)+(i'_x+i'_a)-2=\text{even},$$

$$(i_a+i'_a-1)+(i_b+i'_b)=(i_a+i_b)+(i'_a+i'_b)-1=\text{odd},$$

which entails that the commutator gives terms in C.

Similarly, it can be shown that a commutator between terms from any part and those from C gives terms from C; and parts A, B, and D are cyclically connected in that $[A,B]$ gives terms from D, $[A,D]$ gives terms from B and $[B,D]$ gives terms from A.

Theorem IV.2. For the optimal four-cell systems, achromats up to the fourth order can be obtained by canceling D in the total map.

Proof: (1) The second order: From the proof of Theorem IV.1, the map of $FRSC$ is

$$\begin{aligned}\vec{M}&=_2 \exp(:H^F_3:)\exp(:H^R_3;)\exp(:H^S_3:)\exp(:H^C_3:)\vec{I}\\
&=_2 \exp(:4D_3:)\vec{I},\end{aligned}$$

which shows that the second-order solution is $D_3=0$.

(2) The third order:

Also from the proof of Theorem IV.1, the map is

$$\begin{aligned}\vec{M}&=_3 \exp(:H^F_4:)\exp(:H^R_4:)\exp(:H^S_4:)\exp(:H^C_4:)\vec{I}\\
&=_3 \exp(:4D_4+2[B_3,A_3]:)\vec{I}.\end{aligned}$$

Since $[B_3,A_3]$ belongs to D_4, a third-order achromat can be achieved by zeroing $4D_4+2[B_3,A_3]$ instead of canceling D_4 and A_4 (or B_4) separately. Therefore, the best third-order solution is

$$D_4=-\tfrac{1}{2}[B_3,A_3]. \tag{53}$$

(3) The fourth order:
Using the BCH formula, $\vec{M}$ can be transformed to

$$\begin{aligned}
\vec{M} &=_4 \exp(:H_5^F:)\exp(:H_5^R:)\exp(:H_5^S:)\exp(:H_5^C:)\vec{I}\\
&=_4 \exp(:f_3^F+f_4^F+f_5^F:)\exp(:f_3^R+f_4^R+f_5^R:)\exp(:f_3^S+f_4^S+f_5^S:)\exp(:f_3^C+f_4^C+f_5^C:)\vec{I}\\
&=_4 \exp(:4D_5:)\exp(:f_3^F+f_4^F:)\exp(:f_3^R+f_4^R:)\exp(:f_3^S+f_4^S:)\exp(:f_3^C+f_4^C:)\vec{I}\\
&=_4 \exp(:4D_5:)\exp(:f_3^F+f_4^F+f_3^R+f_4^R+\tfrac{1}{2}([f_3^F,f_3^R]+[f_3^F,f_4^R]+[f_4^F,f_3^R])+\tfrac{1}{12}\{(f_3^F,[f_3^F,f_3^R])+(f_3^R,[f_3^R,f_3^F])\}:)\\
&\quad\times\exp(:f_3^S+f_4^S+f_3^C+f_4^C+\tfrac{1}{2}\,([f_3^S,f_3^C]+[f_3^S,f_4^C]+[f_4^S,f_3^C])+\tfrac{1}{12}\{(f_3^S,[f_3^S,f_3^C])+(f_3^C,[f_3^C,f_3^S])\}:)\vec{I}\\
&=_4 \exp(:4D_5:)\exp(:f_4^F+f_4^R+f_4^S+f_4^C+\tfrac{1}{2}([f_3^F,f_3^R]+[f_3^F,f_4^R]+[f_4^F,f_3^R]+[f_3^S,f_3^C]+[f_3^S,f_4^C]+[f_4^S,f_3^C])\\
&\quad+\tfrac{1}{12}\{(f_3^F,[f_3^F,f_3^R])+(f_3^R,[f_3^R,f_3^F])+(f_3^S,[f_3^S,f_3^C])+(f_3^C,[f_3^C,f_3^S])\}+\tfrac{1}{2}\{f_3^F+f_4^F+f_3^R+f_4^R+\tfrac{1}{2}\,([f_3^F,f_3^R]+[f_3^F,f_4^R]\\
&\quad+[f_4^F,f_3^R]),f_3^S+f_4^S+f_3^C+f_4^C+\tfrac{1}{2}([f_3^S,f_3^C]+[f_3^S,f_4^C]+[f_4^S,f_3^C])\}\\
&\quad+\tfrac{1}{12}\{(f_3^F+f_3^R,[f_3^F+f_3^R,f_3^S+f_3^C])+(f_3^S+f_3^C,[f_3^S+f_3^C,f_3^F+f_3^R])\}:)\vec{I}\\
&=_4 \exp(:4D_5:)\exp(:4D_4+2[B_3,A_3]+\tfrac{1}{2}([f_3^F,f_4^R]+[f_4^F,f_3^R]+[f_3^S,f_4^C]+[f_4^S,f_3^C]+[f_3^F+f_3^R,f_4^S+f_4^C]\\
&\quad+[f_4^F+f_4^R,f_3^S+f_3^C])+\tfrac{1}{4}\{(f_3^F+f_3^R,[f_3^F,f_3^R])+([f_3^F,f_3^R],f_3^F+f_3^R)\}+\tfrac{1}{12}\{(f_3^F,[f_3^F,f_3^R])+(f_3^R,[f_3^R,f_3^F])+(f_3^S,[f_3^S,f_3^C])\\
&\quad+(f_3^C,[f_3^C,f_3^S])+(f_3^F+f_3^R,[f_3^F+f_3^R,f_3^S+f_3^C])+(f_3^S+f_3^C,[f_3^S+f_3^C,f_3^F+f_3^R])\}:)\vec{I}\\
&=_4 \exp(:4D_5:)\exp(:\hat{H}:)\vec{I},
\end{aligned}$$

where

$$\begin{aligned}
\hat{H} &= \tfrac{1}{2}([f_3^F,f_4^R]+[f_3^S,f_4^C]+[f_3^F+f_3^R,f_4^S+f_4^C]+[f_4^F,f_3^R]+[f_4^S,f_3^C]+[f_4^F+f_4^R,f_3^S+f_3^C])\tfrac{1}{4}\{(f_3^F+f_3^R,[f_3^F,f_3^R])\\
&\quad+([f_3^F,f_3^R],f_3^F+f_3^R)\}+\tfrac{1}{12}\{(f_3^F,[f_3^F,f_3^R])+(f_3^R,[f_3^R,f_3^F])+(f_3^S,[f_3^S,f_3^C])+(f_3^C,[f_3^C,f_3^S])\\
&\quad+(f_3^F+f_3^R,[f_3^F+f_3^R,f_3^S+f_3^C])+(f_3^S+f_3^C,[f_3^S+f_3^C,f_3^F+f_3^R])\}\\
&=\tfrac{1}{2}([A_3+B_3+C_3,A_4-B_4-C_4+D_4]+[-A_3-B_3+C_3,-A_4+B_4-C_4+D_4]+4[A_3,-A_4+D_4]+[A_4+B_4+C_4\\
&\quad+D_4,A_3-B_3-C_3]+[-A_4-B_4+C_4+D_4,-A_3+B_3-C_3]+4[A_4+D_4,-A_3])\tfrac{1}{4}\{(2A_3,[-A_3-B_3+C_3,-A_3+B_3\\
&\quad-C_3])+(2A_3,[A_3+B_3+C_3,A_3-B_3-C_3])\}+\tfrac{1}{12}\{(f_3^F-f_3^R,[f_3^F,f_3^R])+(f_3^S-f_3^C,[f_3^S,f_3^C])\}\\
&=-[A_3,B_4+C_4]+[B_3+C_3,A_4+D_4]+[A_3,-B_4+C_4]+[-B_3+C_3,-A_4+D_4]\\
&\quad+4[A_3,D_4]+\tfrac{1}{4}\{(2A_3,[-A_3-B_3+C_3,-A_3+B_3-C_3])+(2A_3,[A_3+B_3+C_3,A_3-B_3-C_3])\}\\
&\quad+\tfrac{1}{12}\{(2[B_3+C_3],[A_3+B_3+C_3,A_3-B_3-C_3])+(2[-B_3+C_3],[-A_3-B_3+C_3,-A_3+B_3-C_3])\}\\
&=2[B_4,A_3]+2[B_3,A_4]+2[C_3,D_4]+4[A_3,D_4]+(A_3,[-A_3,B_3-C_3])-(A_3,[A_3,B_3+C_3])\\
&\quad+\tfrac{1}{3}\{-(B_3+C_3,[A_3,B_3+C_3])+(-B_3+C_3,[-A_3,B_3-C_3])\}\\
&=2[B_4,A_3]+2[B_3,A_4]-2(C_3,-\tfrac{1}{2}[B_3,A_3])+4(A_3,-\tfrac{1}{2}[B_3,A_3])-2(A_3,[A_3,B_3])-\tfrac{2}{3}\{(C_3,[A_3,B_3])+(B_3,[A_3,C_3])\}\\
&=2[B_4,A_3]+2[B_3,A_4]+\tfrac{1}{3}(C_3,[A_3,B_3])+\tfrac{2}{3}(B_3,[C_3,A_3]).
\end{aligned}$$

Altogether, the map is

$$\vec{M}={}_4\exp(:4D_5:)\exp(:\hat{H}:)\vec{I}={}_4\exp\{:4D_5+2[B_4,A_3]+2[B_3,A_4]+\tfrac{1}{3}(C_3,[A_3,B_3])+\tfrac{2}{3}(B_3,[C_3,A_3]):\}\vec{I}.$$

Since $2[B_4,A_3]+2[B_3,A_4]+1/3(C_3,[A_3,B_3])+2/3(B_3,[C_3,A_3])$ belongs to D_5, the best solution for the fourth order is

$$D_5=-\{2[B_4,A_3]+2[B_3,A_4]+\tfrac{1}{3}(C_3,[A_3,B_3])+\tfrac{2}{3}(B_3,[C_3,A_3])\}. \quad (54)$$

Currently it does not appear possible to streamline the treatment of the commutators arising from the BCH formula enough to allow an arbitrary-order analysis of the above phenomenon. But as indicated before, computational results confirm that the method also works for orders five and six. We conclude the analysis at this point, but make the

Conjecture IV.1. For the optimal systems, achromats up to an arbitrary order can be obtained by canceling D in the total map.

When a nth-order achromat is reached, the pseudo-Hamiltonian H becomes

$$H=\sum_{i=1}^{n+1} C_{0000i}\delta^i. \quad (55)$$

Therefore the transfer map contains only time-of-flight terms which depend solely on δ. This is also true for the first order due to symplecticity. As a result, to obtain an achromat with no time-of-flight aberrations requires only one more constraint for each order, i.e., the cancellation of $(t|\delta^i)$. A third-order system of this kind has been designed ([16–18]).

V. APPLICATION: A FIFTH-ORDER ACHROMAT

In this section, we present a proof-of-principle design of a fifth-order achromat. While the feasibility of this design cannot formally confirm the, at times, very tedious algebra, for some of which we had to refer to external references because of space limitations, it appears comforting to see that in practice everything comes together as expected.

For the actual design of the device, differential algebraic (DA) techniques were used because of the ease of obtaining the required Lie factorizations to an arbitrary order. Since the Lie factorizations are always obtained from the map, explicit use of the BCH formula is not necessary, which greatly reduces the complexity of the process.

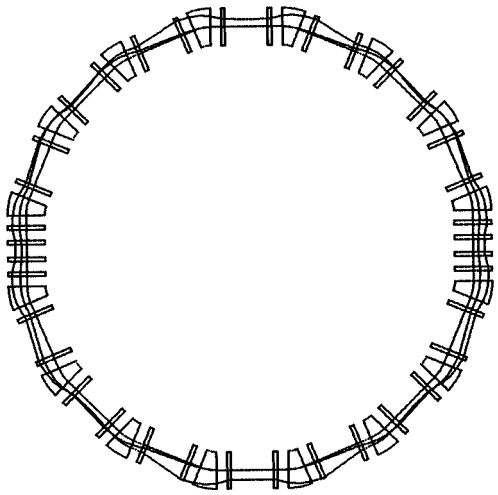

FIG. 2. The FRFR fifth-order achromat: The layout, beam envelope and dispersive ray. The phase advances per cell are $\mu_x=\mu_y=\pi/2$. The circumference is 266.64 m; the emittance is 30π mm mrad; and the dispersion is 0.3%.

Throughout the design processes, the code COSY INFINITY ([19]) is used, which contains all the tools important to beam optical design, including map computation, extraction of Lie coefficients, fitting, tracking, and resolution calculation. First, the DA map of the desired order is computed. Second, relevant Lie coefficients are extracted from the Lie exponent obtained from the map. Third, fitting routines are used to cancel the Lie coefficients which cannot be canceled by symmetry. In our case, we are greatly benefitted from using the package LMDIF by Jorge More from Argonne National Laboratory. The first-order layout of the achromat should avoid large fluctuations in the β functions in order to limit nonlinear aberrations; furthermore, there should be enough room for the insertion of correction multipoles. Another consideration is that, if possible, the number of first-order conditions should be further reduced through symmetry arrangements inside a cell.

The result of these considerations is a ring shown in Fig. 2, which consists of sixteen *FODO* cells plus two dispersion correction sections, each of which includes two quadrupoles. The left half is the forward cell (F) and the right half is the reversed cell (R). Achromaticity is achieved after two turns. The forward cell itself consists of two parts, one of which is the reverse of the other. This guarantees that $(x|x)=(a|a)$ and $(y|y)=(b|b)$. All four *FODO* cells within one part of a

TABLE III. The FRFR fifth-order achromat: The field strengths of the quads and the sextupoles. Only half of them are shown due to mirror symmetry. Numbers in brackets represent powers of 10.

Strengths of the Multipoles (Aperture 10 cm)			
Quadrupoles		Sextupoles	
Gradient (kG/cm)	Field (kG)	Gradient (kG/cm^2)	Field (kG)
−0.162 869	−0.814 344	−0.718 659[−03]	−0.179 665[−01]
0.134 119	0.670 597	0.364 420[−03]	0.911 050[−02]
−0.131 803	−0.659 013		

TABLE IV. The FRFR fifth-order achromat: The field strengths of the octupoles and the decapoles. Note that the multipoles are extremely weak as a result of good linear behavior. Numbers in brackets represent powers of 10.

Strengths of the Multipoles (Aperture 10 cm)			
Octupoles		Decapoles	
Gradient (kG/cm^3)	Field (kG)	Gradient (kG/cm^4)	Field (kG)
−0.996 975[−06]	−0.124 622[−03]	−0.391 808[−06]	−0.244 880[−03]
−0.246 999[−05]	−0.308 749[−03]	0.239 260[−06]	0.149 538[−03]
0.204 723[−05]	0.255 903[−03]	−0.346 336[−07]	−0.216 460[−04]
−0.135 901[−05]	−0.169 876[−03]	−0.413 315[−07]	−0.258 322[−04]
0.951 498[−06]	0.118 937[−03]	0.100 518[−06]	0.628 240[−04]
−0.228 548[−04]	−0.285 685[−02]	−0.501 265[−07]	−0.313 291[−04]
0.177 119[−04]	0.221 399[−02]	−0.953 086[−07]	−0.595 678[−04]
−0.158 309[−04]	−0.197 886[−02]	0.511 256[−06]	0.319 535[−03]
0.420 261[−05]	0.525 326[−03]	−0.305 803[−07]	−0.191 127[−04]
0.871 498[−07]	0.108 937[−04]	−0.775 351[−07]	−0.484 594[−04]
0.377 365[−06]	0.471 706[−04]	0.506 782[−08]	0.316 738[−05]
0.533 332[−05]	0.666 665[−03]	0.153 783[−07]	0.961 144[−05]
0.321 821[−05]	0.402 276[−03]	−0.152 854[−07]	−0.955 335[−05]
0.191 867[−05]	0.239 833[−03]	0.159 598[−06]	0.997 489[−04]
−0.130 343[−05]	−0.162 929[−03]	−0.317 045[−06]	−0.198 153[−03]

cell are identical except that the last one has an extra quadrupole for dispersion correction. Hence, there are three knobs for the first-order design which can zero out $(x|x)$, $(a|a)$, $(y|y)$, $(b|b)$, $(x|\delta)$, and $(a|\delta)$ at the same time. Figure 2 shows that the beam travels around the ring in a very uniform manner, avoiding large ray excursions and β functions. As described in [20] (Sec. 5.2), second-order achromaticity is achieved by symmetrically placing and exciting two pairs of sextupoles in each half.

After the investment in a careful first- and second-order layout, the necessary third- , fourth- , and fifth-order corrections actually turn out to be conceptually straightforward, even though they are computationally more demanding. In the whole process of nonlinear optimization, only two as-

TABLE V. The FRFR fifth-order achromat: The field strengths of the duodecapoles. Note that the multipoles are extremely weak as a result of good linear behavior. Numbers in brackets represent powers of 10.

Strengths of the Duodecapoles (Aperture 10 cm)			
Gradient (kG/cm^5)	Field (kG)	Gradient (kG/cm^5)	Field (kG)
0.260 526[−06]	0.162 829[−03]	0.143 366[−06]	0.896 036[−04]
−0.141 949[−06]	−0.887 180[−04]	0.111 585[−06]	0.697 405[−04]
−0.602 391[−07]	−0.376 494[−04]	−0.392 296[−06]	−0.245 185[−03]
0.115 200[−06]	0.720 003[−04]	0.426 602[−06]	0.266 626[−03]
−0.129 574[−06]	−0.809 839[−04]	−0.251 765[−06]	−0.157 353[−03]
0.167 172[−06]	0.104 483[−03]	0.101 758[−06]	0.635 989[−04]
−0.146 698[−06]	−0.916 861[−04]	−0.812 971[−07]	−0.508 107[−04]
0.109 038[−07]	0.681 489[−05]	0.113 277[−06]	0.707 979[−04]
−0.897 166[−07]	−0.560 728[−04]	−0.423 092[−07]	−0.264 433[−04]
0.905 100[−07]	0.565 687[−04]	−0.733 480[−07]	−0.458 425[−04]
0.422 171[−07]	0.263 857[−04]	0.173 217[−07]	0.108 261[−04]
−0.119 032[−06]	−0.743 948[−04]	0.970 192[−07]	0.606 370[−04]
0.812 032[−07]	0.507 520[−04]	0.745 327[−07]	0.465 829[−04]
−0.859 254[−07]	−0.537 034[−04]	−0.158 631[−06]	−0.991 446[−04]
0.143 652[−06]	0.897 825[−04]	0.230 450[−06]	0.144 031[−03]
−0.192 421[−06]	−0.120 263[−03]	−0.172 798[−06]	−0.107 999[−03]
0.231 122[−06]	0.144 451[−03]	0.923 330[−07]	0.577 081[−04]
−0.729 862[−07]	−0.456 164[−04]	0.126 337[−06]	0.789 607[−04]
−0.102 382[−06]	−0.639 889[−04]	−0.256 941[−06]	−0.160 588[−03]
−0.913 997[−07]	−0.571 248[−04]		

pects seemed to be worth considering. First, the required multipole strengths strongly depend on the average distance between multiples of the same order. In order to keep their strength limited, it is important to have the dimension of the total size of the ring and the dispersive region sufficiently large, and distribute roughly uniformly multipoles of the same order. Second, all the decapoles have to be placed in regions with sufficient dispersion, because all the fourth-order aberrations that remain after third-order corrections are chromatic aberrations. The combination of these considerations results in assuringly weak multipole strengths for third-, fourth-, and fifth-order corrections. (Tables III–V).

VI. CONCLUSION

An analytical theory of arbitrary-order achromats based on mirror symmetries is presented. It is shown that two- and three-cell systems are not optimal to form achromats, they require more conditions than some of the four-cell systems. On the other hand, systems with five or more cells cannot give solutions that are distinctively better than those of four-cell systems. Therefore, four-cell systems seem to be the best choices for arbitrary-order achromats. Four four-cell systems are found optimal for solutions because they require the smallest number of linear conditions.

A general solution for four-cell arbitrary-order achromats based on the optimal systems is presented. Furthermore, a particularly efficient specific solution, which is close to the best solution that can be obtained by the mere use of symmetry, is proved analytically up to the fourth order, and computational results suggest that it is valid up to the sixth order.

ACKNOWLEDGMENTS

We would like to thank Dr. G. Hoffstätter for several useful discussions. This work was supported by the U.S. National Science Foundation under PHY 89-13815, the Alfred P. Sloan Foundation, and the U.S. Department of Energy, Grant No. ER40931.

[1] K. L. Brown, IEEE Trans. Nucl. Sci. **NS-26** (3), 3490 (1979).
[2] J. M. Wouters, D. J. Vieira, H. Wollnik, H. A. Enge, S. Kowalski, and K. L. Brown, Nucl. Instrum. Methods A **240**, 77 (1985).
[3] J. M. Wouters, H. Wollnik, and D. J. Vieira, Nucl. Instrum. Methods A **258**, 331 (1987).
[4] H. Wollnik, Nucl. Instrum. Methods B **26**, 267 (1987).
[5] H. Wollnik, Nucl. Instrum. Methods A **258**, 289 (1987).
[6] K. L. Brown and R. V. Servranckx, IEEE Tran. Nucl. Sci. **NS-32** (5), 2288 (1985).
[7] G. E. Fischer, K. L. Brown, and F. Bulos, (unpublished).
[8] J. J. Murray, K. L. Brown, and T. Fieguth (unpublished).
[9] B. Schwarzschild, Phys. Today **47** (7), 22 (1994).
[10] J. B. Flanz (unpublished).
[11] J. B. Flanz, K. D. Jacobs, R. D. Biron, E. Ihloff, S. Kowalski, Z. Radouch, T. Russ, A. Saab, W. W. Sapp, C. Williamson, A. Zolfaghari, and J. Zumbro (unpublished).
[12] M. Berz, Nucl. Instrum. Methods A **298**, 426 (1990).
[13] A. J. Dragt, Nucl. Instrum. Methods A **258**, 339 (1987).
[14] F. Neri, Michigan State University Cyclotron Laboratory, Technical Report No. MSUCL-767, 1991 (unpublished).
[15] F. Neri (private communication).
[16] W. Wan, E. Goldmann, and M. Berz, *Proceedings of the International Workshop on Nonlinear Problems in Accelerator Physics, Berlin, 1992*, edited by M. Berz, S. Martin, and K. Ziegler, IOP Conf. Series No. 131 (Adam-Hilger, Bristol, England, 1993), p. 201.
[17] W. Wan, E. Goldmann, and M. Berz, in *Computational Accelerator Physics*, edited by Robert Ryne 1993, AIP Conf. Proc. No. 297 (AIP, New York, 1994), p. 143.
[18] W. Wan and M. Berz (unpublished).
[19] M. Berz, COSY INFINITY *Reference Manual Version 6*, Michigan State University Cyclotron Laboratory, Technical Report No. MSUCL-869 (1993) (unpublished).
[20] W. Wan, Ph.D. thesis, Michigan State University, Technical Report No. MSUCL-976, 1995 (unpublished).