This is to certify that the

dissertation entitled

"Theory and Applications of Arbitrary-Order Achromats"

presented by

Weishi Wan

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Physics

Martin Berz
Major professor

Date May 10, 1995

Theory and Applications of Arbitrary-Order Achromats

By

Weishi Wan

A DISSERTATION

submitted to
Michigan State University
in partial fulfillment of the requirements
for the Degree of

DOCTOR OF PHILOSOPHY

Department of Physics and Astronomy

1995

ABSTRACT

An Analytical Theory of Arbitrary Order Achromats and Their Applications

By

Weishi Wan

An analytical theory of arbitrary order achromats for optical systems with midplane symmetry is presented. Besides repetition of cells, mirror symmetry is used to eliminate aberrations. Using mirror imaging 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 easily accounted for and maps are easily manipulated.

It is shown that, independent of the choice and arrangement of such cells, there is a certain minimum number of conditions for a given order; for example, this number is five for the first order, four for the second order, fifteen for the third order, fifteen for the fourth order, and thirty-nine for the fifth and sixth orders. It is shown that the minimum number of cells necessary to reach this optimum level is four, and four of the sixty-four possible four-cell symmetry arrangements are optimal systems. Various third-, fourth- and fifth-order achromats are designed and potential applications are discussed.

To my wife Juxiang and my parents.

ACKNOWLEDGMENTS

Throughout the five years of my graduate study at Michigan State University, I have constantly received help from numerous individuals and organizations. I would like to take this chance to express my gratitude to each of them.

First of all, I am grateful to my thesis advisor, Prof. Martin Berz, for exposing me to the interesting topic of high-order achromats, for guiding my research work through the past four years, for suggesting many important revisions of my thesis, for teaching me critical and rigorous thinking, and, most importantly, for his faith in science and his dedication to scientific research which inspired me to continue my quest for a scientific career. My thanks go to Prof. Henry Blosser, Prof. Jerry Nolen, Prof. Wayne Repko, Prof. Peter Schroeder, and Prof. Brad Sherrill, for kindly serving on my advisory committee and for many helpful comments and suggestions about my dissertation.

I would like to thank Prof. Karl L. Brown and Prof. Roger V. Servranckx for various helpful discussions, Dr. Jorge Moré for providing me the optimization package LMDIF, and Dr. Felix Marti for allowing me access to his private literature collection. I would like to thank Don Lawton for drawing the wonderful pictures used for my presentations. I also thank Prof. Julius Kovacs for supporting my first year of study at MSU.

Special thanks goes to my former colleague, Dr. Georg Hoffstätter, for many stimulating discussions and a courageous act when my life was threatened. I would like to thank other members of our group, Dr. Ralf Degenhardt, Kyoko Fuchi, Khodr Shamseddine, and Meng Zhao, for discussions and friendship. I give my thanks to Jing Wang, Dr. Xiaoyu Wu, and Bo Zhang for the exciting soccer games we played together.

This work was supported in part by the U.S. National Science Foundation, Grant No. PHY 89-13815, and the Alfred P. Sloan Foundation. I express my deep appreciation to them for their continuous support of basic scientific research.

I am deeply indebted to my wife Juxiang for her endurance of my long working days, especially on weekends; for her support of my commitment to scientific research at a time when many are leaving the discipline; for her encouragement; and for her pressure on me when I was lazy at times. It is hard to imagine finishing my graduate study without her. I am grateful to my parents for their wholehearted love, support, and inspiration throughout my life. I am extremely sad, however, because my mother is not able to share my happiness as I finish my graduate study.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 The Map Method in Beam Physics

In the past six decades, accelerators and other beam optical systems have gone through tremendous improvements [Wied93]. The energy that can be reached by accelerators rose from a few MeV to 2 TeV [Law32, Dugan91, Finley91]; and the beam spot size in the Stanford Linear Collider has been decreased to 75 nm × 1 μm [Schwar94]. All these improvements come from the detailed knowledge of the motion of charged particles in electromagnetic fields, which is governed by a system of first-order ordinary differential equations (ODEs)

$$\frac{d}{dt}\vec{x} = \vec{f}(\vec{x}, t). \tag{1.1}$$

Since in general the ODEs are nonlinear and complicated, computers become a very useful tool in solving the ODEs numerically. Through the years, there have been mainly two ways to do this: The ray tracing method and the map method. The basic feature of the ray tracing method [Gordon59] is to send many particles through a numerical integrator derived from the ODEs (for example, a Runge-Kutta integrator), and obtain the final positions and angles of all particles. By studying the distribution of the particles in the phase space (at a fixed position around the reference orbit, called Poincare section), the dynamic behavior of the beam is studied. Two examples for the different computer codes are RAYTRACE [Kowal85], which is used for spectrographs, and TEAPOT [Tal87], which is used for repetitive systems, especially synchrotrons.

Although ray tracing is conceptually simple, it is time-consuming and does not readily provide the direct links between the final and initial coordinates. The map method, on the other hand, focuses on finding the analytical relation between the final and initial coordinates through solving the ODEs. Once these relations, functions between the initial and final coordinates are obtained (the transfer map), all information about the particle motion is known. Therefore, major efforts have been devoted to find the map of the ODEs. It is customary and advantageous to use curvilinear coordinates along a reference particle, such that other particles are always close to the origin and the map is origin preserving. Hence, perturbation theory has been the major tool for solving the ODEs.

Since all particles stay close to the origin, the nonlinearity in the map is weak. This is why the map can be approximately represented by a truncated Taylor series. With orders typically reaching ten, an accuracy in the range of 10 digits can be achieved. The coefficients in the Taylor series, except $\partial x_f/\partial x_i$, $\partial a_f/\partial a_i$, etc., are called the aberrations. A intuitive way to obtain the aberrations is to numerically differentiate the data of the ray-tracing output, like in the code MOTER [Thiess72], where certain low-order derivatives can be extracted. Yet this method of obtaining the derivatives is rather cumbersome and limited, due to the loss of accuracy resulting from the numerical differentiation. This is the main reason for the development of the map method.

Furthermore, for our convenience, the arc length of the reference particle, rather than the time, is used as the independent variable. As a result, the phase space variables are x, p_x, y, p_y, t, and E. Therefore, the ODEs are transformed to

$$\frac{d}{ds}\vec{x} = \vec{f}(\vec{x}, s). \tag{1.2}$$

In order to keep the variables small and simultaneously canonical, a new set of variables, x, a, y, b, Δt, and δ_K, are used in computer codes like COSY INFINITY [Berz93] and also in this thesis, where $a = p_x/p_0$, $b = p_y/p_0$, $\Delta t = (t - t_0)v_0\gamma/(1+\gamma)$, and $\delta_K = (E_k - E_{k0})/E_{k0}$. Note that all quantities with subscript 0 are associated with the reference particle. In these coordinates, the map is denoted by $\vec{M} = (m_x, m_a, m_y, m_b, m_t, m_\delta)$, where

$$\begin{pmatrix} x_f \\ a_f \\ y_f \\ b_f \\ \Delta t_f \\ \delta_f \end{pmatrix} = \vec{M} \begin{pmatrix} x_i \\ a_i \\ y_i \\ b_i \\ \Delta t_i \\ \delta_i \end{pmatrix}. \tag{1.3}$$

When m_x is written as a Taylor series, the coefficient of the term $x_i^{i_x} a_i^{i_a} y_i^{i_y} b_i^{i_b} \Delta t_i^{i_t} \delta_i^{i_\delta}$ is represented by $(x|x^{i_x} a^{i_a} y^{i_y} b^{i_b} \Delta t^{i_t} \delta^{i_\delta})$. The same rule applies to the other functions. For example, $(x|x)$ is the coefficient of term x_i in m_x $(\partial x_f/\partial x_i)$, which is the magnification; $(a|x)$ is the coefficient of term x_i in m_a, which is the defocusing power; and $(x|a^2)$ is the coefficient of a_i^2 in m_x, which is the second-order opening aberration.

For circular machines, the picture of aberrations can not describe the key aspects of particle motion conveniently, because they are averaged out over many turns, with only the ones with the same periodicity as the motion being important. Therefore, instead of magnification, focusing power, and aberrations, the concepts such as tunes, betatron functions, and resonances are used to describe the motion in a circular machine. Generally speaking, the tunes (T) are defined as the remainder of the number of periods of the motion in one plane over the number of turns, which is a measure of the entire motion, including the nonlinearities [Berz92a]. To the first order, the tune of the ith component is

$$T_i = \arccos(Tr(L_i))/2\pi, \tag{1.4}$$

where L_i is the linear matrix of one turn for the ith componant and $Tr(L_i)$ is the trace of it. The tunes of the transverse motion are called the betatron tunes and the tune of the longitudinal motion is called the synchrotron tune. In this thesis, only betatron tunes (T_x and T_y) are relevant. When the relation

$$lT_x + mT_y = n \quad (l,\ m,\ \text{and}\ n\ \text{are integers.}) \tag{1.5}$$

holds, the motion is said to be on an $(l+m)$th order resonance. As shown in Section 3.3, there are always aberrations which grow exponentially under this resonance, called the driving terms of the resonance. A motion which is on a resonance with nonzero driving terms is unstable and the particle will eventually hit the wall and get lost. Since there are infinitely many resonances, virtually every particle will be lost after a long time. Practically speaking, particles are only stored in a machine for a certain period, which means that mostly the low-order resonances are important, because they cause the growth of certain lower-order aberrations, which are initially bigger than higher-order aberrations.

After the development of the theory of the alternating-gradient synchrotron [Cour58], large synchrotrons took the center stage of high energy accelerators; the map method has been developed with this. Currently there are dozens of various computer codes in use, ranging from first-order codes like COMFORT [Wood83] and SYNCH [Garren75], to second- and third-order codes based on explicit formulas, like TRANSPORT [Brown73], MAD [Iselin85, Iselin88], DIMAD [Serv85], TRIO [Matsuo76], GIOS [Woll87a], and MARYLIE [Dragt85], to higher-order codes of the same approach, like COSY 5.0

[Berz87a], and finally to arbitrary-order codes that do not rely on explicit formulas, like TLIE [Zeijts92], ZLIB [Yan90], and COSY INFINITY [Berz90, Berz93]. It is worth noting that it is possible to compute transfer maps of an arbitrary order after the emergence of the differential algebraic (DA) techniques [Berz89]. Among the high-order codes (beyond the third-order), COSY 5.0 is a fifth-order one, and the rest are all DA codes. Of all the DA codes, COSY INFINITY is the first and, up to now, probably the most general code. It has been shown through the history of the code development that DA techniques are probably the only practical way to obtain the transfer map of an arbitrary order. Therefore, they should be discussed in more detail.

Without involving too much mathematics (see, for example, [Berz90, Berz92b]), the differential algebraic techniques can be viewed as a way to solve the ODEs to an arbitrary order in one attempt without loosing accuracy. The keys are, first, that the arbitrary-order Taylor expansion of a large class of functions, whose variables are Taylor series with constant parts, can be obtained through a finite number of operations; second, that Taylor expansions of all functions are done simultaneously, instead of the traditional way of obtaining higher-order solutions through lower-order ones.

Since it is only necessary and possible to obtain a finite number of terms from the Taylor series of a function, the truncation order n is always specified when Taylor expansion is done, and terms of higher orders are neglected. Therefore, infinitely often differentiable functions, including $\vec{f}$ of the beam optical systems, can be expanded around any given point, and the expansion up to order n can be obtained by a finite number of additions and multiplications, even when $\vec{f}$ is complicated. Together with addition, multiplication, and differentiation of polynomials, expressed in the space of polynomials up to order n and properly implemented in computers, the ODEs can be solved with a DA-based numerical integrator where the values of the phase space variables are replaced by the DA vectors containing the constants and derivatives of those variables. At the end, the solution contains not only the final values but also all derivatives up to a certain order, which are the aberrations in the transfer map. It is worth noting that in the code COSY INFINITY, the integrator is a modified eighth-order Runge-Kutta integrator with automatic step size control to ensure that the accuracy of the solution is compatible with the specific order.

When the electromagnetic field in a beam optical system does not change longitudinally, there is a much quicker way to solve the ODEs using the so-called flow operator. In this case, $\vec{f}$ in the ODEs does not depend on s, which leads to the result that the transfer map is

$$\vec{M} = \exp((s - s_0)L_{\vec{f}})\vec{I}, \tag{1.6}$$

where $L_{\vec{f}} = \vec{f} \cdot \vec{\nabla}$ is the flow operator, and $\vec{I} = (x_i, a_i, y_i, b_i, \Delta t_i, \delta_i)$. Thus, the map is obtained after only one step. Note that when $\vec{f}$ represents a Hamiltonian system, the flow operator $L_{\vec{f}}$ becomes a Lie operator similar to that discussed in Section 2.1.

The DA techniques offer a power tool to study beam optical systems, including tracking through a high-order map, map manipulations such as composition and inversion, computing generating functions and Lie factorizations, studying parameter dependence of certain quantities [Berz92c], and suppressing resonances through normal forms [Berz92a, Berz92b, Berz93].

1.2 An Overview of Achromats

The search for achromats up to a certain order has generated substantial interest for the past two decades. Here an achromat is defined as a beam optical system

Order	1	2	3	4	5
Aberrations	6	30	70	140	252
Independent aberrations	6	18	37	65	110

Table 1.1: The number of aberrations of orders 1 to 5 for a system with midplane symmetry. The interdependency of aberrations comes from symplecticity, which is a property of a Hamiltonian system.

whose map of the transverse motion is free of aberrations up to a certain order. The advantage of achromats is that all aberrations of the transverse motion are cancelled, as are all aberrations of the longitudinal motion except $(t|\delta^n)$; hence, an achromat transports charged particles without distortion of the transverse motion. This is why first- and second-order achromats have been so 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, cancelling half of the transverse aberrations, all achromat theories consider only systems with midplane symmetry. Table 1.1 lists the number of aberrations that have to be cancelled for achromats up to the fifth order. It shows that the number of independent aberrations grows so rapidly with increasing order that even up to only the second order it is unrealistic to obtain an achromat by providing each aberration with a knob. Therefore, the challenge is how to achieve an achromat with as few knobs as possible.

Due to their simplicity, first-order achromats have been widely used, especially the partial achromats where only dispersion is corrected. If we exclude the techniques of dispersion matching and the dispersion suppressor (see, for example [Wied93]), there are basically two ways of obtaining a first-order achromat, namely, through mirror symmetry and repetition. The principle of a mirror symmetrical first-order achromat

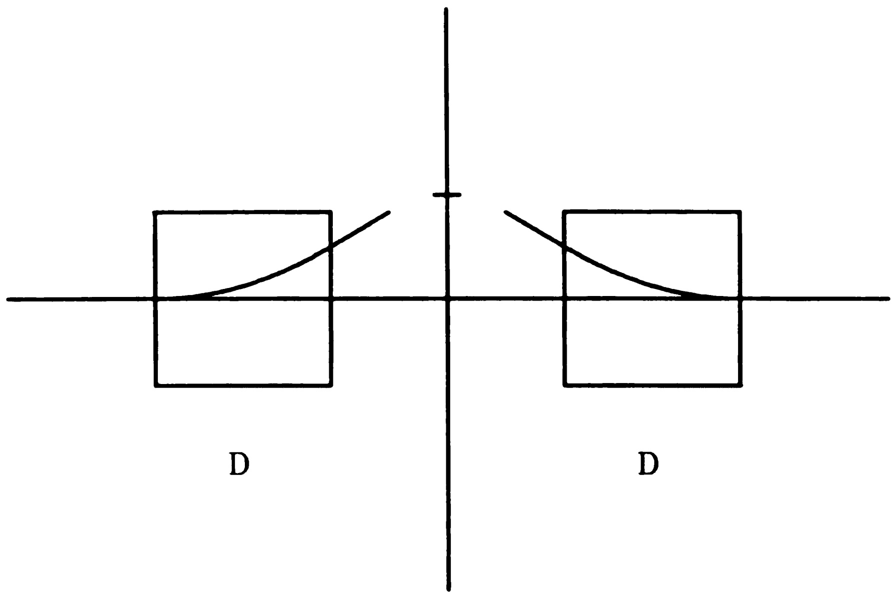

Figure 1.1: The principle of a mirror symmetrical, first-order achromat. The conditions are the cancellation of $(a|\delta)$, $(a|x)$, $(x|a)$, $(b|y)$, and $(y|b)$ in the middle where the vertical line lies.

is illustrated in Fig. 1.1. When $(a|x)$, $(x|a)$, $(b|y)$, $(y|b)$ and $(a|\delta)$ vanish in the middle, mirror symmetry entails that the total first-order map is $\vec{I}$ (here $\vec{I}$ stands for the identity map).

Mirror symmetrical achromats are mainly used as building blocks of synchrotron light sources [Jack87] due to the low equilibrium emittance achieved (see, for example, [Wied93]). As examples, the lattices of the Advanced Photon Source (APS) at Argonne National Laboratory (ANL), and the Advanced Light Source at Lawrence Berkeley Laboratory (LBL) are shown in Figures 1.2 and 1.3, respectively [Murphy92].

As shown in Section 3.2, any repetitive system with integer tunes is a first-order achromat. Since a repetitive first-order achromat (except a 3-cell system) cancels all second-order geometric aberrations, it has been used for bending arcs of various synchrotrons and storage rings, especially as a 180° bending arc of a racetrack lattice [Serv83, Serv83, Litv93, Wu93a, Wu93b]. Fig. 1.4 presents the lattice of the Duke

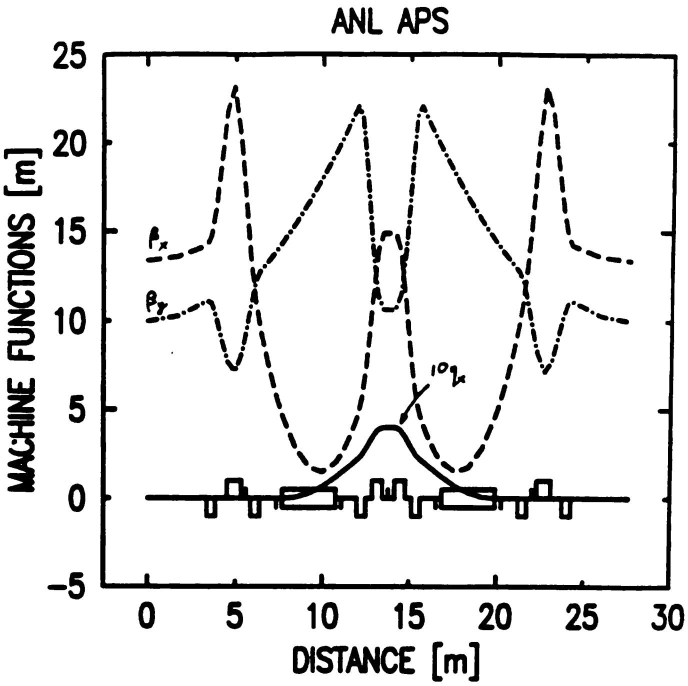

Figure 1.2: The Double Bend Achromat lattice of the APS at ANL.

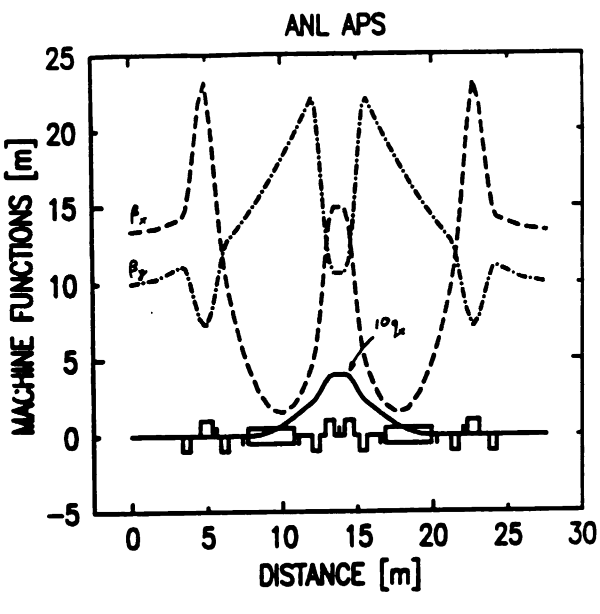

Figure 1.3: The Triple Bend Achromat lattice of the ALS at LBL.

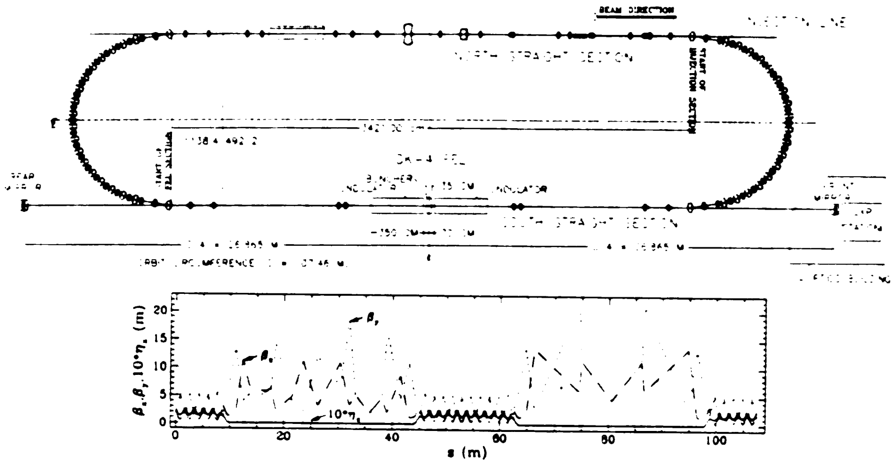

Figure 1.4: The layout and lattice functions of the Duke FEL storage ring, where each arc is made up of 10 FODO cells, with phase advances per cell chosen as $\psi_x = (3/10)2\pi$ and $\psi_y = (1/10)2\pi$.

FEL (Free Electron Laser) storage ring.

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 in the 1970s that a theory developed by K. Brown enabled the design of realistic second-order achromats in a systematic and elegant way [Brown79, Brown82a].

The theory is based on the following observations. First, any system of n identical cells ($n > 1$), with the overall first-order matrix equaling unity ($\vec{I}$) 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 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. Detailed studies on this theory will be shown in Section 3.2.

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) [Wouter85, Wouter87] and multi-pass(ESR) [Woll87b, Woll87c], the Arcs of the Stanford Linear Collider (SLC), the new facility at SLAC, the Final Focus Test Beam [Brown85, Brown87a, Brown87b, Schwar94], and the MIT South Hall Ring (SHR) [Flanz89a, Flanz89b].

Since it is hard to generalize the second-order achromat theory to higher orders, the first third-order achromat theory was developed by Dr. Alex Dragt based on normal form theory and Lie algebra [Dragt87]. 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 cells T_x and T_y are not full, half, third or quarter integer resonant, but nT_x and nT_y are integers. (2) The two chromaticities and five independent third-order aberrations are zero. Details of the theory will be discussed in Section 3.3.

Two examples of third-order achromats have been designed. The first design was done by Dragt himself, containing thirty cells with $T_x = 1/5$ and $T_y = 1/6$. Each contains ten bends, two quads, two sextupoles and five octupoles. The whole system forms a 180° bending arc. The second design was done by Neri [Neri91]. It is a seven-cell system with only one bend per cell, with the total bend also being 180°. The tunes of a cell are $T_x = 1/7$ and $T_y = 2/7$, which seems to violate the theory because of the third-order resonance $2T_x - T_y = 0$. However, the achromaticity can still be achieved because the driving terms are cancelled by midplane symmetry (see Section 3.3). This approach greatly reduces the number of cells.

Similar to Brown's theory, the Dragt theory cannot be immediately used to find arbitrary order achromats in such a way that the number of the cells in a achromat 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 system cells has to be the smallest

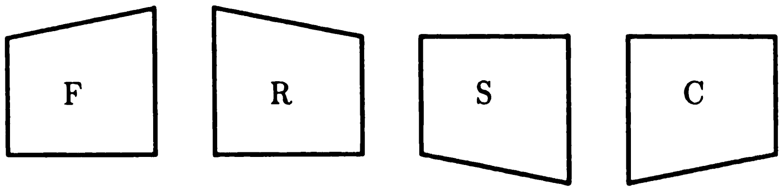

Figure 1.5: The geometric relationship among cells F, R, S, and C illustrated by asymmetric boxes.

number 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 an analytical formula increases rapidly because of the complexity of the Baker-Campbell-Hausdorff 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) (Fig. 1.5).

This thesis is organized as follows: The Lie representations of symplectic maps are discussed in Chapter 2, including the definition of a Lie operator, various methods of Lie factorization, and the Baker-Campbell-Hausdorff formula. As a comparison to our achromat theory, the theories of repetitive achromats up to order three are presented in Chapter 3. In Chapter 4, the analytical theory for arbitrary order achromats is studied with a detailed proof provided for every theorem. Chapter 5 consists of four example designs of achromats of orders three to five. Also presented are studies on various aspects of the nonlinear dynamical behavior of those examples. Finally, a summary concludes the thesis.

Chapter 2

The Lie Representation of Symplectic Maps

In the following chapter the Lie representation of symplectic maps, developed by Dragt and Finn [Dragt76, Dragt81] is outlined, proceeded by the introduction of Lie transformations and symplecticity. Only the results closely related to the achromat theory are presented.

2.1 Lie Transformations and Symplecticity

First, let us review the definition and basic properties of the Poisson bracket, since a Lie transformation is defined based on the Poisson bracket. On a phase space $\mathcal{R}^{2m}$ with variables $(q_1, \cdots, q_m, p_1, \cdots, p_m)$, a Poisson bracket of functions f and g is defined as

$$[f, g] = \sum_{i=1}^{m} (\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}) = \vec{\nabla} f \cdot \hat{J} \cdot \vec{\nabla} g^t, \tag{2.1}$$

where

$$\vec{\nabla} f = (\frac{\partial f}{\partial q_1}, \cdots, \frac{\partial f}{\partial q_m}, \frac{\partial f}{\partial p_1}, \cdots, \frac{\partial f}{\partial p_m})$$

and $\hat{J}$ is an antisymmetric $2m \times 2m$ matrix

$$\hat{J} = \begin{pmatrix} 0 & \hat{I} \\ -\hat{I} & 0 \end{pmatrix}. \tag{2.2}$$

As an example, note that

$$[I_i, I_j] = J_{ij}, \tag{2.3}$$

where $\vec{I} = (q_1, \cdots, q_m, p_1, \cdots, p_m)$ and J_{ij} is the (ij) element of the matrix $\hat{J}$.

It is well known that Poisson brackets have the following properties:

$$[f, g+h] = [f,g] + [f,h], \tag{2.4}$$

$$[f, tg] = t[f,g] \quad \text{where } t \text{ is an arbitrary constant}, \tag{2.5}$$

$$[f, gh] = [f,g]h + g[f,h], \tag{2.6}$$

$$[f,[g,h]] + [g,[h,f]] + [h,[f,g]] = 0. \tag{2.7}$$

With the Poisson bracket defined as the multiplication on $\mathcal{R}^{2m}$, the space of functions on $\mathcal{R}^{2m}$ form a Lie algebra.

For any function f on $\mathcal{R}^{2m}$, a Lie operator $:f:$ acting on another function g on $\mathcal{R}^{2m}$ is defined as

$$:f: g = [f,g]. \tag{2.8}$$

The zero power of $:f:$ is defined as

$$:f:^0 g = g \tag{2.9}$$

and the square of $:f:$ is

$$:f:^2 g = [f,[f,g]]; \tag{2.10}$$

with higher powers being defined in the same way. Similar to the Poisson brackets, Lie operators have the following properties:

$$:f:(g+h) \;=\; :f:g+:f:h, \tag{2.11}$$

$$:f:(tg) \;=\; t:f:g, \quad \text{where } t \text{ is an arbitrary constant,} \tag{2.12}$$

$$:f:(gh) \;=\; (:f:g)h+g(:f:h), \tag{2.13}$$

$$:f::g:h-:g::f:h \;=\; :[f,g]:h. \tag{2.14}$$

Note that, with the multiplication defined as

$$:f:\times:g: =:f::g:-:g::f:, \tag{2.15}$$

the Lie operators on $\mathcal{R}^{2m}$ form another Lie algebra.

A useful relation that $:f:^n$ obeys is the Leibniz rule

$$:f:^n(gh)=\sum_{m=0}^{n}\begin{pmatrix} n \\ m \end{pmatrix}(:f:^m g)(:f:^{n-m} h), \tag{2.16}$$

where

$$\begin{pmatrix} n \\ m \end{pmatrix}=\frac{n!}{m!(n-m)!}.$$

To prove it, the mathematical induction method is used. First, for $n=1$, we have

$$:f:(gh)=(:f:g)h+g(:f:h),$$

which satisfies eq. (2.16). Second, for $n-1$, we assume eq. (2.16) holds. Third, for n, we have

$$\begin{aligned} :f:^n(gh) &= :f:(:f:^{n-1}(gh)) \\ &= :f:\left(\sum_{m=0}^{n-1}\begin{pmatrix} n-1 \\ m \end{pmatrix}(:f:^m g)(:f:^{n-1-m} h)\right) \\ &= \sum_{m=0}^{n-1}\begin{pmatrix} n-1 \\ m \end{pmatrix}(:f:^{m+1} g)(:f:^{n-1-m} h) \end{aligned}$$

$$+\sum_{m=0}^{n-1}\begin{pmatrix} n-1 \\ m \end{pmatrix}(:f:^m g)(:f:^{n-m} h)$$

$$= \sum_{m=1}^{n}\begin{pmatrix} n-1 \\ m-1 \end{pmatrix}(:f:^m g)(:f:^{n-m} h)$$

$$+\sum_{m=0}^{n-1}\begin{pmatrix} n-1 \\ m \end{pmatrix}(:f:^m g)(:f:^{n-m} h)$$

$$= \sum_{m=0}^{n}\begin{pmatrix} n \\ m \end{pmatrix}(:f:^m g)(:f:^{n-m} h),$$

which concludes the proof.

A Lie transformation associated with a function f on $\mathcal{R}^{2m}$ is defined as

$$\exp(:f:) = \sum_{n=0}^{\infty}\frac{1}{n!}:f:^n . \tag{2.17}$$

To avoid the subtle questions connected to the convergence of a Lie transformation, where even the definition of the norm is not clear, we require that f is a polynomial of orders 3 and up, expansion is always truncated at order n, and the functions, on which the Lie transformation acts, are also polynomials. Since no infinite series is involved, the conclusions are completely rigorous. On the other hand, due to the fact that n is an arbitrary natural number, this approach does lose generality. Practically speaking, this is always the case for the DA maps, which makes this treatment fit perfectly to the implementation of the Lie transformation. Therefore, the new definition of a Lie transformation is:

$$\exp(:f:) =_n \sum_{i=0}^{n}\frac{1}{i!}:f:^i . \tag{2.18}$$

Lie transformations have the following properties:

$$\exp(:f:)(g+h) =_n \exp(:f:)g + \exp(:f:)h, \tag{2.19}$$

$$\exp(:f:)tg =_n t\exp(:f:)g, \tag{2.20}$$

$$\exp(:f:)(gh) =_n (\exp(:f:)g)(\exp(:f:)h), \tag{2.21}$$

$$\exp(:f:)[g,h] =_n [\exp(:f:)g, \exp(:f:)h], \tag{2.22}$$

The proof of eq. (2.21) makes use of the Leibniz rule (eq. 2.16):

$$\begin{aligned}
& (\exp(:f:)g)(\exp(:f:)h) \\
=_n & \left(\sum_{l=0}^{n} \frac{1}{l!} :f:^l g\right)\left(\sum_{m=0}^{n} \frac{1}{m!} :f:^m h\right) \\
=_n & \left(\sum_{l=0}^{n}\left(\sum_{m=0}^{l} \frac{1}{m!} :f:^m g \frac{1}{(l-m)!} :f:^{l-m} h\right)\right) \\
=_n & \left(\sum_{l=0}^{n}\frac{1}{l!}\left(\sum_{m=0}^{l} \frac{l!}{m!(l-m)!} :f:^m g :f:^{l-m} h\right)\right) \\
=_n & \left(\sum_{l=0}^{n}\frac{1}{l!}\left(\sum_{m=0}^{l} \binom{l}{m} :f:^m g :f:^{l-m} h\right)\right) \\
=_n & \sum_{l=0}^{n}\frac{1}{l!} :f:^l (gh) \quad \text{(Leibniz rule)} \\
=_n & \exp(:f:)(gh).
\end{aligned} \tag{2.23}$$

Equation (2.22) can be obtained directly from eq. (2.21).

Now let us look at a Lie transformation acting on a polynomial

$$g(q_1, \cdots, q_m, p_1, \cdots, p_m) = \sum_{i_{q_1},\cdots,i_{q_m},i_{p_1},\cdots,i_{p_m}} a_{i_{q_1},\cdots,i_{p_m}} q_1^{i_{q_1}} \cdots q_m^{i_{q_m}} p_1^{i_{p_1}} \cdots p_m^{i_{p_m}}, \tag{2.24}$$

where $i_{q_1} + \cdots + i_{q_n} + i_{p_1} + \cdots + i_{p_n} \leq n$. Using eqs. (2.19) and (2.21) repeatedly, we obtain

$$\begin{aligned}
\exp(:f:)g &=_n \exp(:f:)\left(\sum_{i_{q_1},\cdots,i_{q_m},i_{p_1},\cdots,i_{p_m}} a_{i_{q_1},\cdots,i_{p_m}} q_1^{i_{q_1}} \cdots q_m^{i_{q_m}} p_1^{i_{p_1}} \cdots p_m^{i_{p_m}}\right) \\
&=_n \sum_{i_{q_1},\cdots,i_{q_m},i_{p_1},\cdots,i_{p_m}} a_{i_{q_1},\cdots,i_{p_m}} \exp(:f:)(q_1^{i_{q_1}} \cdots q_m^{i_{q_m}} p_1^{i_{p_1}} \cdots p_m^{i_{p_m}}) \\
&=_n \sum_{i_{q_1},\cdots,i_{q_m},i_{p_1},\cdots,i_{p_m}} a_{i_{q_1},\cdots,i_{p_m}} (\exp(:f:)q_1)^{i_{q_1}} \cdots (\exp(:f:)q_m)^{i_{q_m}}
\end{aligned}$$

$$(\exp(:f:)p_1)^{i_{p_1}} \cdots (\exp(:f:)p_m)^{i_{p_m}}$$
$$=_n \quad g(\exp(:f:)q_1, \cdots, \exp(:f:)q_m, \exp(:f:)p_1, \cdots, \exp(:f:)p_m),$$

which leads to the following important theorem:

Theorem 2.1 *If f is a polynomial of order 3 or higher on $\mathcal{R}^{2m}$, and g is an arbitrary polynomial, then we have*

$$\exp(:f:)g =_n g(\exp(:f:)\vec{I}), \tag{2.25}$$

where

$$\exp(:f:)\vec{I} =_n (\exp(:f:)q_1, \cdots, \exp(:f:)q_n, \exp(:f:)q_1, \cdots, \exp(:f:)q_n). \tag{2.26}$$

Theorem 2.1 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 2.2 *If $\vec{M}$ is a map and g is an arbitrary polynomial on $\mathcal{R}^{2m}$, then we have*

$$\exp(:f:)g(\vec{M}) =_n g(\exp(:f:)\vec{M}). \tag{2.27}$$

Theorem 2.2 is probably even more important than Theorem 2.1 because in most cases, the linear map of a system is unity. Therefore, it is the one which is used directly and very frequently in Chapter 3 and 4.

Another useful concept is symplecticity. A $2m$ dimensional vector function $\vec{M}$ is symplectic when its Jacobian matrix $\text{Jac}(\vec{M})$ satisfies

$$\text{Jac}(\vec{M}) \cdot \hat{J} \cdot \text{Jac}(\vec{M})^t = \hat{J}. \tag{2.28}$$

This kind of function is important in that the transfer map of a Hamiltonian system is always symplectic [Gold80, Arnold89]. Symplectic maps have the following properties:

$$[f(\vec{M}), g(\vec{M})] = ([f,g])(\vec{M}), \tag{2.29}$$

$$\det(\mathrm{Jac}(\vec{M})) = 1, \tag{2.30}$$

$$\mathrm{Jac}(\vec{M} \circ \vec{N}) \cdot \hat{J} \cdot \mathrm{Jac}(\vec{M} \circ \vec{N})^t = \hat{J}, \tag{2.31}$$

$$\mathrm{Jac}(\vec{M}^{-1}) \cdot \hat{J} \cdot \mathrm{Jac}(\vec{M}^{-1})^t = \hat{J}, \tag{2.32}$$

where $\vec{M}$ and $\vec{N}$ are symplectic maps and "o" denotes the composition of two functions. Equations (2.29) and (2.31) can be proven using the chain rule of derivatives. The proof of eq. (2.31) can be found in reference [Dragt81]. Since the Jacobian of $\vec{M}^{-1}$ is the inverse of that of $\vec{M}$, we obtain

$$\begin{aligned}
&\mathrm{Jac}(\vec{M}^{-1}) \cdot \hat{J} \cdot \mathrm{Jac}(\vec{M}^{-1})^t \\
&\quad = \mathrm{Jac}(\vec{M})^{-1} \cdot \hat{J} \cdot (\mathrm{Jac}(\vec{M})^{-1})^t \\
&\quad = \mathrm{Jac}(\vec{M})^{-1} \cdot \hat{J} \cdot (\mathrm{Jac}(\vec{M})^{-1})^t \\
&\quad = (\mathrm{Jac}(\vec{M})^t \cdot -\hat{J} \cdot \mathrm{Jac}(\vec{M}))^{-1} \\
&\quad = (-\hat{J})^{-1} \\
&\quad = \hat{J},
\end{aligned}$$

which shows that $\vec{M}^{-1}$ is also symplectic (eq. 2.32). Equations (2.31) and (2.32) indicate that, with the composition "o" as the multiplication, symplectic maps form a group.

From eq. (2.21), we have

$$[\exp(:f:)I_i, \exp(:f:)I_j] \tag{2.33}$$

$$=_n \ \exp(:f:)[I_i, I_j] \tag{2.34}$$

$$=_n \ \exp(:f:)J_{ij} \tag{2.35}$$

$$=_n \ J_{ij}, \tag{2.36}$$

which leads to the following theorem:

Theorem 2.3 *If a map $\vec{M}$ on $\mathcal{R}^{2m}$ has the form $\vec{M} =_n \exp(:f:)\vec{I}$, it is symplectic.*

In the next section, it will be shown that any symplectic map can be represented by either one or a series of Lie transformations up to an arbitrary order.

Next let us study the composition of $\exp(:f:)g(\vec{I})$ and map $\vec{M}$. Note that this operation plays a key role in the development of achromat theories.

Theorem 2.4 *If a map $\vec{M}$ on $\mathcal{R}^{2m}$ is symplectic, we have*

$$(\exp(:f:)g) \circ (\vec{M}) =_n \exp(:f(\vec{M}):)g(\vec{M}). \tag{2.37}$$

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

$$(\exp(:f:)g) \circ (\vec{M}) =_n \exp(-:f(\vec{M}):)g(\vec{M}). \tag{2.38}$$

Proof:

First assume that $\vec{M}$ is a symplectic map, which implies that a Poisson bracket is an invariant under the transformation $\vec{M}$ (eq. 2.29), i.e.:

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

The rest of the proof is straightforward:

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

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.:

$$[f(\vec{M}), g(\vec{M})] = -([f, g])(\vec{M}). \tag{2.39}$$

Therefore we obtain

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

which concludes the proof.

2.2 Lie Factorizations

In this section, proofs are given for the fact that any symplectic map can be represented, up to a given order, by a product of Lie transformations, which is referred to as a Lie factorization. Various methods of Lie factorization are presented, including Dragt-Finn factorization [Dragt76] and the single-exponent factorization.

Theorem 2.5 *Let $\vec{M}$ be a symplectic map on $\mathcal{R}^{2m}$. Then, to an arbitrary order n, there exists a matrix L and homogeneous polynomials f_3, f_4, $\cdots$, f_{n+1} of orders 3, 4, $\cdots$ and $n+1$, respectively, such that*

$$\vec{M} =_n (L\vec{I}) \circ (\exp(: f_3 :)\vec{I}) \circ (\exp(: f_4 :)\vec{I}) \circ \cdots \circ (\exp(: f_{n+1} :)\vec{I}). \tag{2.40}$$

Note that L is simply a matrix; $L\vec{I}$, on the other hand, is the linear map, which is a set of functions. Similarly, $\exp(:f:)$ is just an operator while $\exp(:f:)\vec{I}$ is the map generated by it.

Proof:

The induction method is used.

(1) The second order:

Define $\vec{M}_1 = (L^{-1}\vec{I}) \circ \vec{M}$ and $\vec{N}_1 = \vec{M}_1 - \vec{I}$. According to eq. (2.31), $\vec{M}_1$ is symplectic. Besides, note that $\vec{M}_1 =_1 \vec{I}$ and $\vec{N}_2$ contain terms of orders two and up. Suppose $\vec{M}_1 =_2 \exp(:f_3:)\vec{I}$. We have

$$\vec{M}_1 \;=_2\; \vec{I} + [f_3, \vec{I}] \;=_2\; -\vec{\nabla} f_3 \cdot \hat{J},$$

which entails that

$$\vec{\nabla} f_3 \;=_2\; -(\vec{M}_1 - \vec{I}) \cdot \hat{J} \;=_2\; -\vec{N}_1 \cdot \hat{J}.$$

On the other hand, the symplecticity of $\vec{M}_1$ gives

$$\begin{aligned} & \mathrm{Jac}(\vec{M}_1) \cdot \hat{J} \cdot \mathrm{Jac}(\vec{M}_1)^t = \hat{J}, \\ \Rightarrow\; & (\hat{I} + \mathrm{Jac}(\vec{N}_1)) \cdot \hat{J} \cdot (\hat{I} + \mathrm{Jac}(\vec{M}_1)^t) = \hat{J}, \\ \Rightarrow\; & \hat{J} + \mathrm{Jac}(\vec{N}_1)\hat{J} + \hat{J}\mathrm{Jac}(\vec{M}_1)^t =_1 \hat{J}, \\ \Rightarrow\; & \mathrm{Jac}(\vec{N}_1)\hat{J} =_1 (\mathrm{Jac}(\vec{N}_1)\hat{J})^t. \end{aligned} \tag{2.41}$$

Here, $\mathrm{Jac}(\vec{N}_1) \cdot \hat{J} \cdot \mathrm{Jac}(\vec{N}_1)^t$ is eliminated because it contains terms of orders two and up. According to the potential theorem [Meyer91], there exists a function f, which satisfies $\vec{\nabla} f = \vec{g}$, if and only if $\vec{g}$ satisfies

$$\frac{\partial g_i}{\partial x_j} = \frac{\partial g_j}{\partial x_i}. \tag{2.42}$$

Therefore, eq. (2.41) shows that f_3 does exist.

(2) The $(n-1)$st order:

Assume that the theorem holds, i.e.:

$$\vec{M} =_{n-1} (L\vec{I}) \circ (\exp(:f_3:)\vec{I}) \circ \cdots \circ (\exp(:f_n:)\vec{I}).$$

(3) The nth order:

Define

$$\vec{M}_{n-1} = (\exp(:f_n:)\vec{I})^{-1} \circ \cdots \circ (\exp(:f_3:)\vec{I})^{-1} \circ (L^{-1}\vec{I}) \circ \vec{M}.$$

Hence $\vec{M}_{n-1} =_{n-1} \vec{I}$. According to Theorem 2.3 and eq. (2.31), $\vec{M}_{n-1}$ is symplectic. Now define $\vec{N}_{n-1} = \vec{M}_{n-1} - \vec{I}$. Note that $\vec{N}_{n-1}$ contains terms of orders n and up. Suppose $\vec{M}_{n-1} =_n \exp(:f_{n+1}:)\vec{I}$. Similar to the second-order case, we have

$$\vec{M}_{n-1} \;=_n\; \vec{I} + [f_{n+1}, \vec{I}] \;=_n\; -\vec{\nabla} f_{n+1} \cdot \hat{J}.$$

Since $\mathrm{Jac}(\vec{N}_{n-1})$ contains terms of orders $(n-1)$ and up, the symplecticity of $\vec{M}_{n-1}$ entails that

$$\mathrm{Jac}(\vec{N}_{n-1})\hat{J} \;=_{n-1}\; (\mathrm{Jac}(\vec{N}_{n-1})\hat{J})^t.$$

Hence f_{n+1} exists. Finally, it has been proven that

$$\vec{M} \;=_n\; (L\vec{I}) \circ (\exp(:f_3:)\vec{I}) \circ (\exp(:f_4:)\vec{I}) \circ \cdots \circ (\exp(:f_{n+1}:)\vec{I}),$$

which concludes the proof.

Using Theorem 2.4 repeatedly, eq. (2.40) can be written into

$$\vec{M} =_n \exp(:f_{n+1}:)\exp(:f_n:)\cdots\exp(:f_3:)(L^{-1}\vec{I}), \tag{2.43}$$

where $\exp(:f_i:)$ $(i = 3, \cdots, n+1)$ are operators acting on functions obtained from previous Lie transformations. Please note the difference between $\exp(:f:)\exp(:g:)\vec{I}$

and $(\exp(:f:)\vec{I}) \circ (\exp(:g:)\vec{I})$. In the former expression, $\exp(:f:)$ acts on the map $\exp(:g:)\vec{I}$; in latter expression, $\exp(:f:)$ and $\exp(:g:)$ act on the unity map $\vec{I}$ separately, and the resulting maps $\exp(:f:)\vec{I}$ and $\exp(:g:)\vec{I}$ are composed afterwards.

It is worth noting that $\vec{M}$ can also be expressed in the reverse order, where

$$\vec{M} =_n (\exp(:f_{n+1}:)\vec{I}) \circ (\exp(:f_n:)\vec{I}) \circ \cdots \circ (\exp(:f_3:)\vec{I}) \circ (L\vec{I}). \tag{2.44}$$

The proof is the same except that $\vec{M}_i$ is defined as

$$\vec{M}_i = \vec{M} \circ (L^{-1}\vec{I}) \circ (\exp(:f_3:)\vec{I})^{-1} \circ \cdots \circ (\exp(:f_{i+1}:)\vec{I})^{-1}.$$

The next theorem shows that it is also possible to represent a symplectic map using a single Lie transformation.

Theorem 2.6 *Let $\vec{M}$ be a symplectic map on $\mathcal{R}^{2m}$. Then to an arbitrary order n, there exists a matrix L and a polynomial H of orders 3 and 4, up to $n+1$, such that*

$$\vec{M} =_n (L\vec{I}) \circ (\exp(:H:)\vec{I}). \tag{2.45}$$

Proof:

First, define $\vec{M}_1 = (L^{-1}) \circ \vec{M}$.

(1) The second order:

From the proof of the last theorem, we know that there exists a function f_3 of order 3 satisfying

$$\vec{M}_1 =_2 \exp(:f_3:)\vec{I}.$$

(2) The third order:

Suppose there is a function f_4 of order 4 such that

$$\begin{aligned}
\vec{M}_1 &=_3 \exp(:f_3+f_4:)\vec{I} \\
&=_3 \vec{I}+[f_3+f_4,\vec{I}]+\frac{1}{2}[f_3+f_4,[f_3+f_4,\vec{I}]] \\
&=_3 \vec{I}+[f_3,\vec{I}]+\frac{1}{2}[f_3,[f_3,\vec{I}]]+[f_4,\vec{I}] \\
&=_3 \exp(:f_3:)\vec{I}+[f_4,\vec{I}] \\
&=_3 \exp(:f_3:)\vec{I}+\vec{\nabla} f_4\cdot\hat{J}. \qquad (2.46)
\end{aligned}$$

The removal of f_4 from $[f_3+f_4,[f_3+f_4,\vec{I}]]$ is due to the fact that f_4 is of order 4; hence $[f_3,[f_4,\vec{I}]]$ and $[f_4,[f_3,\vec{I}]]$ give fourth-order terms.

Now define $\vec{N}_2 = \vec{M}_1 - \exp(:f_3:)\vec{I}$, which contains terms of orders 3 and up. Thus from eq. (2.46) we obtain

$$\vec{\nabla} f_4 =_3 -(\vec{M}-\exp(:f_3:)\vec{I})\cdot\hat{J}$$

Since $\vec{M}_1$ is symplectic, we have

$$\begin{aligned}
& \mathrm{Jac}(\vec{M}_1)\cdot\hat{J}\cdot\mathrm{Jac}(\vec{M}_1)^t =_2 \hat{J} \\
\Rightarrow\ & (\mathrm{Jac}(\vec{N}_1)+\mathrm{Jac}(\exp(:f_3:)\vec{I}))\cdot\hat{J}\cdot(\mathrm{Jac}(\vec{N}_1)^t+\mathrm{Jac}(\exp(:f_3:)\vec{I})^t) =_2 \hat{J}, \\
\Rightarrow\ & \mathrm{Jac}(\vec{N}_1)\cdot\hat{J}\cdot\mathrm{Jac}(\exp(:f_3:)\vec{I})^t+\mathrm{Jac}(\exp(:f_3:)\vec{I})\cdot\hat{J}\cdot\mathrm{Jac}(\vec{N}_1)^t, \\
& +\mathrm{Jac}(\exp(:f_3:)\vec{I}))\cdot\hat{J}\cdot\mathrm{Jac}(\exp(:f_3:)\vec{I})^t =_2 \hat{J}, \\
\Rightarrow\ & \mathrm{Jac}(\vec{N}_1)\cdot\hat{J}\cdot\mathrm{Jac}(\exp(:f_3:)\vec{I})^t+\mathrm{Jac}(\exp(:f_3:)\vec{I})\cdot\hat{J}\cdot\mathrm{Jac}(\vec{N}_1)^t =_2 0, \\
& \quad (\exp(:f_3:)\vec{I} \text{ is symplectic.}) \\
\Rightarrow\ & \mathrm{Jac}(\vec{N}_1)\cdot\hat{J}\cdot\hat{I}^t+\hat{I}\cdot\hat{J}\cdot\mathrm{Jac}(\vec{N}_1)^t =_2 0, \\
& \quad (\mathrm{Jac}(\vec{N}_1) \text{ is of orders 2 and up.}) \\
\Rightarrow\ & \mathrm{Jac}(\vec{N}_1)\hat{J} =_2 (\mathrm{Jac}(\vec{N}_1)\hat{J})^t.
\end{aligned}$$

Therefore f_4 exists. Choosing $H = f_3 + f_4$, we have

$$\vec{M} =_3 (L\vec{I})\circ(\exp(:H:)\vec{I}).$$

(3) The $(n-1)$st order:

Assume that the theorem holds, i.e.

$$\vec{M} =_{n-1} (L\vec{I}) \circ (\exp(:H:)\vec{I}),$$

where $H = f_3 + f_4 + \cdots + f_n$.

(4) The nth order:

Suppose there is a function f_{n+1} of order $n+1$ such that

$$\begin{aligned} \vec{M}_1 &=_n \exp(:f_3 + f_4 + \cdots + f_{n+1}:)\vec{I} \\ &=_n \exp(:f_3 + f_4 + \cdots + f_n:)\vec{I} + \vec{\nabla} f_{n+1} \cdot \hat{J}. \end{aligned} \tag{2.47}$$

Define $\vec{N}_{n-1} = \vec{M}_1 - \exp(:H:)\vec{I}$, which contains terms of orders n and up. Thus from eq. (2.47) we obtain

$$\vec{\nabla} f_4 =_n -(\vec{M} - \exp(:H:)\vec{I}) \cdot \hat{J}$$

Similar to the third-order case, the symplecticity of $\vec{M}_1$ gives

$$\Rightarrow \quad \mathrm{Jac}(\vec{N}_{n-1})\hat{J} =_{n-1} (\mathrm{Jac}(\vec{N}_{n-1})\hat{J})^t,$$

which implies that f_{n+1} exists. Choosing $H = f_3 + f_4 + \cdots + f_{n+1}$, we have

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

which concludes the proof.

Note that the proofs for the above factorization theorems provide algorithms for obtaining the Lie factorization of arbitrary orders. In practice, they are used by Differential Algebraic codes such as COSY INFINITY [Berz93].

2.3 The Baker-Campbell-Hausdorff Formula

In this section we will study an important formula, the Baker-Campbell-Hausdorff formula, which combines two Lie transformations into one. The complete proof of it involves knowledge of Lie algebra that is beyond the scope of this thesis [Dynkin62, Vara84]. Instead, we will present a partial proof to show that it holds to the third order for Lie operators.

Theorem 2.7 *Let A and B be two functions on $\mathcal{R}^{2m}$. The following relation holds:*

$$\begin{aligned}&\exp(:A:)\exp(:B:)\\ &\quad =_3 \exp(:A+B+\frac{1}{2}[A,B]+\frac{1}{12}([A,[A,B]]+[B,[B,A]]):),\end{aligned}$$

where A and B are polynomials of orders 3 and higher, and " $=_3$" means the truncation of the polynomial of Lie operators at the third order.

Proof:

From eq. (2.14), we have

$$:A::B:-:B::A:=:[A,B]:.$$

Using this relation repeatedly, the rest of the proof is straightforward.

The left-hand side can be transformed to

$$\begin{aligned}&\exp(:A:)\exp(:B:)\\ &\quad =_3 (1+:A:+\frac{1}{2}:A:^2+\frac{1}{6}:A:^3)(1+:B:+\frac{1}{2}:B:^2+\frac{1}{6}:B:^3)\\ &\quad =_3 1+:A:+\frac{1}{2}:A:^2+\frac{1}{6}:A:^3+:B:+:A::B:+\frac{1}{2}:A:^2:B:\\ &\qquad +\frac{1}{2}:B:^2+\frac{1}{2}:A::B:^2+\frac{1}{6}:B:^3\\ &\quad =_3 1+:(A+B):+\frac{1}{2}(:A:^2+:A::B:+:B::A:+:B:^2+:[A,B]:)\end{aligned}$$

$$+\frac{1}{6}(:A:^3 + 3:A:^2\!:B: + 3:A::B:^2 + :B:^3)$$

$$=_3 1 + :(A+B): + \frac{1}{2}:(A+B):^2 + \frac{1}{2}:[A,B]:$$

$$+\frac{1}{6}(:A:^3 + :A:^2\!:B: + :A::B::A: + :A::[A,B]: + :B::A:^2$$

$$+ :A::[A,B]: + :[A,B]::A: + :A::B:^2 + :B::A::B:$$

$$+ :[A,B]::B: + :B:^2\!:A: + :[A,B]::B: + :B::[A,B]: + :B:^3)$$

$$=_3 1 + :(A+B): + \frac{1}{2}:(A+B):^2 + \frac{1}{2}:[A,B]: + \frac{1}{6}:(A+B):^3$$

$$+\frac{1}{6}(2:A::[A,B]: + :[A,B]::A: + :B::[A,B]: + 2:[A,B]::B:).$$

The right-hand side can be transformed to

$$\exp(:A + B + \frac{1}{2}[A,B] + \frac{1}{12}([A,[A,B]] + [B,[B,A]]) + \cdots :)$$

$$=_3 1 + :(A + B + \frac{1}{2}[A,B] + \frac{1}{12}([A,[A,B]] + [B,[B,A]])):$$

$$+\frac{1}{2}:(A + B + \frac{1}{2}[A,B]):^2 + \frac{1}{6}:(A+B):^3$$

$$=_3 1 + :(A+B): + \frac{1}{2}:(A+B):^2 + \frac{1}{2}:[A,B]: + \frac{1}{6}:(A+B):^3$$

$$+\frac{1}{4}(:(A+B)::[A,B]: + :[A,B]::(A+B):)$$

$$+\frac{1}{12}(:A::[A,B]: - :[A,B]::A: + :B::[A,B]: - :[A,B]::B:)$$

$$=_3 1 + :(A+B): + \frac{1}{2}:(A+B):^2 + \frac{1}{2}:[A,B]: + \frac{1}{6}:(A+B):^3$$

$$+\frac{1}{6}(2:A::[A,B]: + :[A,B]::A: + :B::[A,B]: + 2:[A,B]::B:).$$

Altogether, the left-hand side equals the right-hand side, which ends the proof.

It is worth noting that only commutators appear in the right-hand side of the B-C-H formula, which is remarkable in that, for any Lie algebra, the manipulation of the corresponding Lie transformation does not require extra operations. Besides, the B-C-H formula links the Lie operators with the Poisson brackets.

A direct result of the B-C-H formula is that we are able to find the inverse of $\exp(:f:)$ easliy. Since $[f,f]=0$, we obtain

$$\exp(:f:)\exp(:-f:) \;=\; \exp(:f+(-f):)=\exp(:0:)=1.$$

Therefore, the inverse of $\exp(:f:)$ is $\exp(:-f:)$. Futhermore, the inverse map of $\vec{M}$, in the form of a single-factor Lie factorization, i.e., $\vec{M}=(L\vec{I})\circ(\exp(:H:)\vec{I})$, is

$$\vec{M}^{-1}=(\exp(:-H:)\vec{I})\circ(L^{-1}\vec{I}), \tag{2.48}$$

which is shown below:

$$\begin{aligned}
\vec{M}^{-1}\circ\vec{M} &= (\exp(:-H:)\vec{I})\circ(L^{-1}\vec{I})\circ(L\vec{I})\circ(\exp(:H:)\vec{I}) \\
&= (\exp(:-H:)\vec{I})\circ(\exp(:H:)\vec{I}) \\
&= \exp(:H:)\exp(:-H:)\vec{I}=\vec{I}; \\
\vec{M}\circ\vec{M}^{-1} &= (L\vec{I})\circ(\exp(:H:)\vec{I})\circ(\exp(:-H:)\vec{I})\circ(L^{-1}\vec{I}) \\
&= (L\vec{I})\circ(\exp(:H:)\exp(:-H:)\vec{I})\circ(L^{-1}\vec{I}) \\
&= (L\vec{I})\circ(L^{-1}\vec{I})=\vec{I}.
\end{aligned}$$

Chapter 3

Repetitive Achromat Theory

In this chapter the theories of repetitive achromats up to order three will be discussed. This will give us an overview of what has been achieved and the remaining difficulties. All achromat theories, including the arbitrary-order theory presented in Chapter 4, deal with systems with midplane symmetry. The first section is devoted to the definition and implications of midplane symmetry.

3.1 Midplane Symmetry

In a system with midplane symmetry, two particles that are symmetrical about the midplane at the beginning stay symmetrical afterwards. In other words, let us consider a beam optical system with transfer map $\vec{M}$. Suppose a particle enters it at $(x_i,a_i,y_i,b_i,t_i,\delta_i)$ and exits at $(x_f,a_f,y_f,b_f,t_f,\delta_f)$. Another particle that enters it at $(x_i,a_i,-y_i,-b_i,t_i,\delta_i)$ will exit at $(x_f,a_f,-y_f,-b_f,t_f,\delta_f)$. Now the matrix P is defined as

$$P\vec{I} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ a \\ y \\ b \\ t \\ \delta \end{pmatrix}. \tag{3.1}$$

Hence, midplane symmetry can be expressed as

$$(P\vec{I})\circ\vec{M}\circ(P^{-1}\vec{I})=\vec{M}. \tag{3.2}$$

Since we are only interested in symplectic maps, $\vec{M}$ can be represented by a linear matrix and a Lie transformation (see Section 2.2):

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

Inserting into eq. (3.2) and using Theorems 2.4 and 2.2, we have

$$\begin{aligned} & (L\vec{I})\circ(\exp(:H:)\vec{I}) = (P\vec{I})\circ(L\vec{I})\circ(\exp(:H:)\vec{I})\circ(P^{-1}\vec{I}) \\ \Rightarrow\ & (L\vec{I})\circ(\exp(:H:)\vec{I}) = \exp(:H(P^{-1}\vec{I}):)(P\cdot L\cdot P^{-1}\vec{I}). \end{aligned}$$

For first-order matrices, midplane symmetry requires that

$$L=P\cdot L\cdot P^{-1}, \tag{3.3}$$

which implies that for higher orders

$$\begin{aligned} & \exp(:H(\vec{I}):)=\exp(:H(P^{-1}\vec{I}):), \\ \Rightarrow\ & H(\vec{I})=H(P^{-1}\vec{I}). \end{aligned} \tag{3.4}$$

Equations (3.3) and (3.4) determine that

$$L=\begin{pmatrix} (x|x) & (x|a) & 0 & 0 & (x|t) & (x|\delta) \\ (a|x) & (a|a) & 0 & 0 & (a|t) & (a|\delta) \\ 0 & 0 & (y|y) & (y|b) & 0 & 0 \\ 0 & 0 & (b|y) & (b|b) & 0 & 0 \\ (t|x) & (t|a) & 0 & 0 & (t|t) & (t|\delta) \\ (\delta|x) & (\delta|a) & 0 & 0 & (\delta|t) & (\delta|\delta) \end{pmatrix} \tag{3.5}$$

and

$$H=\sum_{i_x i_a i_y i_b i_t i_\delta} C_{i_x i_a i_y i_b i_t i_\delta}\, x^{i_x} a^{i_a} y^{i_y} b^{i_b} t^{i_t} \delta^{i_\delta}, \tag{3.6}$$

where $i_x + i_a + i_y + i_b + i_t + i_\delta \geq 3$ and $i_y + i_b$ is even. Here the independent variable is the arc length of a reference trajectory s and $x_i, a_i, y_i, b_i, t_i, \delta_i$ are canonical variables.

An achromat can be achieved only when acceleration is not present and synchrotron radiation can be neglected. Therefore, the transfer map of such a system is not only symplectic but also time-independent (static) and energy-conserving. These put more constraints on the transfer map, and we now have

$$L = \begin{pmatrix} (x|x) & (x|a) & 0 & 0 & 0 & (x|\delta) \\ (a|x) & (a|a) & 0 & 0 & 0 & (a|\delta) \\ 0 & 0 & (y|y) & (y|b) & 0 & 0 \\ 0 & 0 & (b|y) & (b|b) & 0 & 0 \\ (t|x) & (t|a) & 0 & 0 & (t|t) & (t|\delta) \\ 0 & 0 & 0 & 0 & 0 & (\delta|\delta) \end{pmatrix} \tag{3.7}$$

and

$$H = \sum_{i_x i_a i_y i_b i_\delta} C_{i_x i_a i_y i_b i_\delta} x^{i_x} a^{i_a} y^{i_y} b^{i_b} \delta^{i_\delta}, \tag{3.8}$$

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

It is clear that, from eq. (3.8), we have

$$H = \sum_{i_\delta = 3} C_{0,0,0,0,i_\delta} \delta^{i_\delta} \tag{3.9}$$

when an achromat is reached. Hence, only the terms $(t|\delta^n)$ $(n = 2, 3, \cdots)$ are left in an achromat. Futhermore, symplecticity [Berz85] implies that

$$(t|x) = -(x|x)(a|\delta_k) + (a|x)(x|\delta_k), \tag{3.10}$$

$$(t|a) = -(x|a)(a|\delta_k) + (a|a)(x|\delta_k), \tag{3.11}$$

which means that this is also true for the first order.

3.2 Second-Order Achromat Theory

Since a second-order achromat is always based on a first-order achromat, let us first study how to obtain a repetitive first-order achromat. Consider a system consisting of n identical cells ($n > 1$) and let L_x be the x-matrix of one cell, which has the form

$$L_x = \begin{pmatrix} (x|x) & (x|a) & (x|\delta) \\ (a|x) & (a|a) & (a|\delta) \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} M & \vec{\omega} \\ 0 & 1 \end{pmatrix}. \tag{3.12}$$

Therefore, the total x-matrix L_{Tx} is

$$\begin{aligned} L_{Tx} &= L_x^n \\ &= \begin{pmatrix} M^n & (M^{n-1} + M^{n-2} + \cdots + \hat{I})\vec{\omega} \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} M^n & (M^n - \hat{I})(M - \hat{I})^{-1}\vec{\omega} \\ 0 & 1 \end{pmatrix}. \end{aligned} \tag{3.13}$$

Equation (3.13) shows that the dispersion vanishes when $M^n = \hat{I}$, i.e., the phase advance equals a multiple of 2π. Together with the requirement that $L_y = \hat{I}$, a first-order achromat is reached when the tunes of the whole system are integers.

Brown's second-order achromat theory is built on a repetitive first-order achromat, as described above. It consists of the following two theorems:

Theorem A: If a system contains N identical cells ($N > 1$ and $N \neq 3$), all second-order geometric aberrations vanish when, for a cell, both transverse planes have the same non-integer tunes and the phase advance of the system is a multiple of 2π.

Proof:

Here we adopt K. Brown's original notation where the linear matrix is represented by R and the second-order matrix is T. Therefore, the second-order map is

$$x_{i,1} = \sum_{j} R_{ij} x_{j,0} + \sum_{j,k} T_{ijk} x_{j,0} x_{k,0}, \tag{3.14}$$

where $x_{i,0}$ and $x_{i,1}$ are initial and final coordinates, respectively. According to perturbation theory, high-order solutions can be obtained through the lower-order solution and the inhomogeneous part of the ODEs. In the case of Brown's theory, the ODEs are expanded to the second order, where the the terms in the inhomogeneous part are called the driving terms [Brown82b]. With all driving terms obtained, T_{ijk} can be expressed as a function of R_{ij}, which is

$$T_{ijk} = \int_0^L K_p(s)(R_{ij}(s))^n (R_{ik}(s))^m ds, \quad \text{with } (n+m) = 3, \tag{3.15}$$

and where $K_p(s)$ is the multiple strength at s. For geometric aberrations, R_{ij} should come from the geometric part of R only, which is

$$\begin{pmatrix} \cos\psi(s) + \alpha(s)\sin\psi(s) & \beta(s)\sin\psi(s) \\ -\gamma(s)\sin\psi(s) & \cos\psi(s) - \alpha(s)\sin\psi(s) \end{pmatrix}.$$

As a result, T_{ijk} can be written into

$$T_{ijk} = \int_0^L F_p(s)\sin^n(\psi(s))\cos^m(\psi(s)) ds. \tag{3.16}$$

Since $\sin^n(\psi(s))\cos^m(\psi(s))$ $(m+n=3)$ gives only $e^{\pm i\psi(s)}$ and $e^{\pm 3i\psi(s)}$, the conditions for all second-order geometric aberrations to vanish are

$$\int_0^L F_p e^{\pm i\psi} ds = 0 \quad \text{and} \quad \int_0^L F_p e^{\pm 3i\psi} ds = 0.$$

Due to the fact that the system consists of individual cells, the integral conditions above become the following sums

$$\sum_{k=1}^{N} \tilde{F}_k e^{\pm i\psi_k} = 0 \quad \text{and} \quad \sum_{k=1}^{N} \tilde{F}'_k e^{\pm 3i\psi_k} = 0,$$

where

$$
\begin{aligned}
\tilde{F}_k &= \int_{\frac{k}{N}L+s_0-\Delta s}^{\frac{k}{N}L+s_0+\Delta s} F_p(s)e^{\pm i(\psi(s)-\psi_k(s_0))}ds \\
&= \int_{-\Delta s}^{\Delta s} F_p(\frac{k}{N}L+s_0+\tilde{s})e^{\pm i(\psi(\frac{k}{N}L+s_0+\tilde{s})-\psi_k(s_0))}d\tilde{s} \\
&= \int_{-\Delta s}^{\Delta s} F_p(\tilde{s})e^{\pm i(\Delta\psi(\tilde{s}))}d\tilde{s}, \qquad (3.17)
\end{aligned}
$$

$$
\tilde{F}'_k = \int_{-\Delta s}^{\Delta s} F_p(\tilde{s})e^{\pm 3i(\Delta\psi(\tilde{s}))}d\tilde{s}. \qquad (3.18)
$$

Here, N, L, s_0, and Δs are the number of cells, the length of the system, the position of the center of the element considered, and the half-length of the element, respectively. Repetition of the system is used to obtain eq. (3.17) and (3.18). Since $\tilde{F}_k$ and $\tilde{F}'_k$ are independent of k, eq. (3.17) is further reduced to

$$
\sum_{k=1}^{N} e^{\pm i\psi_k} = 0 \quad \text{and} \quad \sum_{k=1}^{N} e^{\pm 3i\psi_k} = 0.
$$

In conclusion, all second-order aberrations vanish when $N \neq 3$, $N\psi_{x,y} = 2m_{x,y}\pi$, and $m_{x,y} \neq 2mN$ ($m = 1, 2, \cdots$) (see Figs. 3.1 and 3.2).

The second theorem deals with the correction of second-order chromatic aberrations left in a system satisfying **Theorem A**.

Theorem B: For a system that satisfies **Theorem A** and $N > 3$, a second-order achromat is achieved when two families of sextupole components are adjusted so as to make one chromatic aberration in each transverse plane vanish. In other words, only two chromatic aberrations are independent.

The proof of this theorem can be found in reference [Carey81]. Another proof using normal form theory will be given in Section 3.3 as part of the third-order achromat theory. A typical four-cell, second-order achromat is shown in Fig. 3.3.

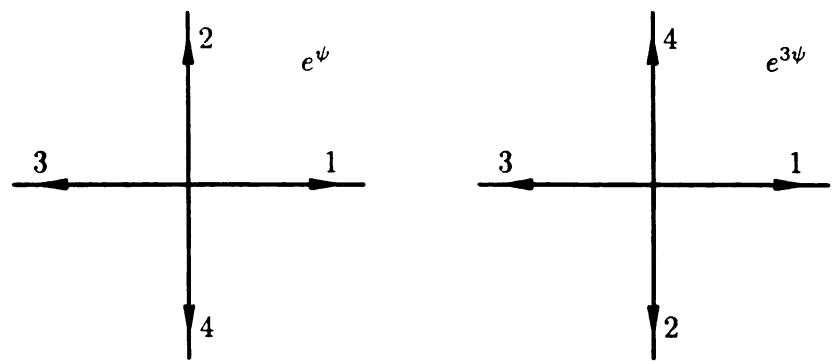

Figure 3.1: Complex plane diagram for second-order geometric aberrations of a four-cell repetitive system with phase advances 2π.

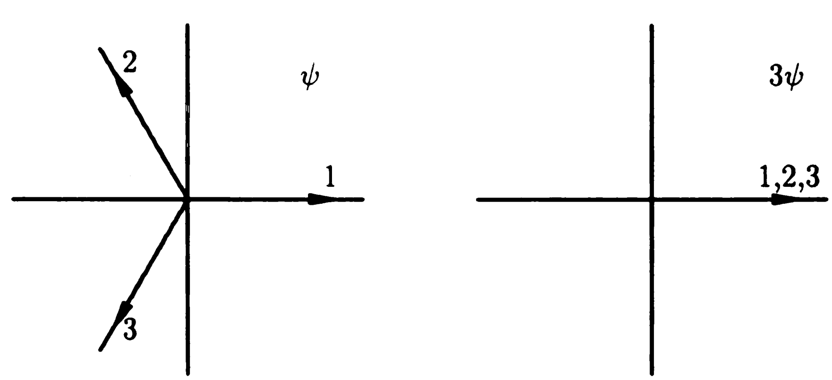

Figure 3.2: Complex plane diagram for second-order geometric aberrations of a three-cell repetitive system with phase advances 2π.

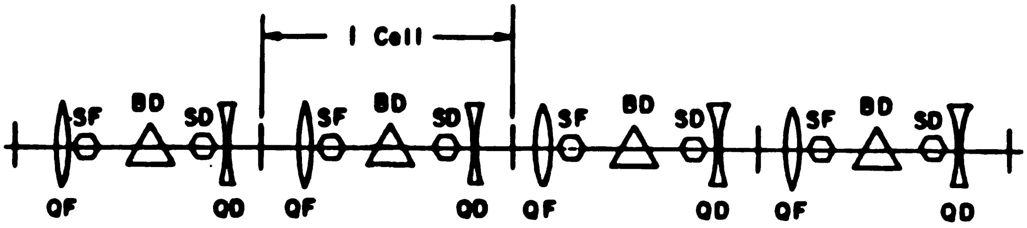

Figure 3.3: K. Brown's four-cell, second-order achromat. Quadrupoles are used to tune the system to phase advance 2π in both transfer planes, and two families of sextupoles, SF and SD, are used to correct chromatic second-order aberrations in the x- and y-planes, respectively. To make the sextupoles weak, they are placed such that β_x is larger at SF and β_y is larger at SD.

3.3 Third-Order Achromat Theory

In the mid 80's, Dragt [Dragt87] developed a third-order achromat theory for repetitive systems based on the normal form theory [Dragt79]. Although the same result can be obtained from other normal form algorithms, the Lie factorization has the advantage of explicitly showing the number of independent aberrations to be corrected. In practice, however, it is difficult to implement the Lie normal form beyond the fifth order; hence DA techniques have to be used, either combined with Lie algebraic techniques [Forest89] or by themselves [Berz92a], to compute the normal form map.

The key idea of this theory is that when an achromat is achieved in the normal form coordinates, it is achieved in any set of coordinates. Since the transfer map in the normal form coordinates is much simpler than that in the original curvilinear coordinates, the conditions for an achromat become much clearer. Consider an n-cell symplectic system with midplane symmetry. From Section 2.2, the transfer map of one cell can be written into

$$\vec{M} =_n (L\vec{I}) \circ (\exp(:f_3:)\vec{I}) \circ (\exp(:f_4:)\vec{I}) \circ \cdots \circ (\exp(:f_{n+1}:)\vec{I}).$$

To ord

$\vec{M}$ =

Since

transf

5 x 5

A

whic

A

whe

Sup

eige

can

To order 3, we have

$$\vec{M} =_3 (L\vec{I}) \circ (\exp(:f_3:)\vec{I}) \circ (\exp(:f_4:)\vec{I}). \tag{3.19}$$

Since t is not of interest, it is discarded. Therefore $\vec{I} = (x, a, y, b, \delta)$. The normal form transformation is done order by order [Forest89]. For the first order, there exists a 5×5 symplectic matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & & & a_{15} \\ a_{21} & a_{22} & & & a_{25} \\ & & a_{33} & a_{34} & \\ & & a_{43} & a_{44} & \\ & & & & 1 \end{pmatrix}. \tag{3.20}$$

which satisfies

$$A \cdot L \cdot A^{-1} = R, \tag{3.21}$$

where

$$R = \begin{pmatrix} e^{i\mu_x} & & & & \\ & e^{-i\mu_x} & & & \\ & & e^{i\mu_y} & & \\ & & & e^{-i\mu_y} & \\ & & & & 1 \end{pmatrix}. \tag{3.22}$$

Suppose $s_{x,y}$ is the eigenvector of $e^{i\mu_{x,y}}$. The fact that L is real entails that $\bar{s}_{x,y}$ is the eigenvector of $e^{-i\mu_{x,y}}$.

Now let us define $\vec{N}_1$ as the transfer map in the eigenvector coordinates, which can be transformed to

$$\begin{aligned} \vec{N}_1 &= (A\vec{I}) \circ \vec{M} \circ (A^{-1}\vec{I}) \\ &= (A\vec{I}) \circ (L\vec{I}) \circ (\exp(:f_3:)\vec{I}) \circ (\exp(:f_4:)\vec{I}) \circ (A^{-1}\vec{I}) \end{aligned}$$

Note

To t

wh

w

$$\begin{aligned}
&= (A\vec{I}) \circ (L\vec{I}) \circ (A^{-1}\vec{I}) \circ (A\vec{I}) \circ (\exp(:f_3:)\vec{I}) \circ (A^{-1}\vec{I}) \\
&\quad \circ (A\vec{I}) \circ (\exp(:f_4:)\vec{I}) \circ (A^{-1}\vec{I}) \\
&= (R\vec{I}) \circ (\exp(:f_3(A^{-1}\vec{I}):)\vec{I}) \circ (\exp(:f_4(A^{-1}\vec{I}):)\vec{I}) \\
&= (R\vec{I}) \circ (\exp(:g_3:)\vec{I}) \circ (\exp(:g_4:)\vec{I}). \qquad (3.23)
\end{aligned}$$

Note that $\vec{I}$ is the unity map in the eigenvector coordinates.

Next we define

$$\vec{N}_2 = (\exp(:G_3:)\vec{I}) \circ \vec{N}_1 \circ (\exp(:-G_3:)\vec{I}). \qquad (3.24)$$

To the second order, we have

$$\begin{aligned}
\vec{N}_2 &=_2 (\exp(:G_3:)\vec{I}) \circ \vec{N}_1 \circ (\exp(:-G_3:)\vec{I}) \\
&=_2 (\exp(:G_3(R\vec{I}):)R\vec{I}) \circ (\exp(:g_3:)\vec{I}) \circ (\exp(:-G_3:)\vec{I}) \\
&=_2 \exp(:-G_3:)\exp(:g_3:)\exp(:G_3(R\vec{I}):)(R\vec{I}) \\
&=_2 \exp(:g_3-(G_3-G_3(R\vec{I})):)(R\vec{I}), \qquad (3.25)
\end{aligned}$$

where use of the B-C-H formula has been made.

Since G_3 and g_3 are polynomials of order 3, in general they have the form

$$G_3 = \sum_{m_x n_x m_y n_y i_\delta} G_{m_x n_x m_y n_y i_\delta} s_x^{m_x} \bar{s}_x^{n_x} s_y^{m_y} \bar{s}_y^{n_y} \delta^{i_\delta}, \qquad (3.26)$$

$$g_3 = \sum_{m_x n_x m_y n_y i_\delta} g_{m_x n_x m_y n_y i_\delta} s_x^{m_x} \bar{s}_x^{n_x} s_y^{m_y} \bar{s}_y^{n_y} \delta^{i_\delta}, \qquad (3.27)$$

where $m_x + n_x + m_y + n_y + i_\delta = 3$ and $m_y + n_y$ is even. Therefore, we have

$$\begin{aligned}
& g_3 - (G_3 - G_3(R\vec{I})) \\
&= \sum_{m_x n_x m_y n_y i_\delta} \left(g_{m_x n_x m_y n_y i_\delta} - \left(1 - e^{i((m_x-n_x)\mu_x+(m_y-n_y)\mu_y)}\right) G_{m_x n_x m_y n_y i_\delta}\right) \\
&\qquad s_x^{m_x} \bar{s}_x^{n_x} s_y^{m_y} \bar{s}_y^{n_y} \delta^{i_\delta} \\
&= \sum_{m_x n_x m_y n_y i_\delta} \left(g_{m_x n_x m_y n_y i_\delta} - \left(1 - e^{i\vec{\mu}\cdot(\vec{m}-\vec{n})}\right) G_{m_x n_x m_y n_y i_\delta}\right) s_x^{m_x} \bar{s}_x^{n_x} s_y^{m_y} \bar{s}_y^{n_y} \delta^{i_\delta}, \qquad (3.28)
\end{aligned}$$

where $\vec{\mu} = (\mu_x, \mu_y)$, $\vec{m} = (m_x, m_y)$ and $\vec{n} = (n_x, n_y)$. As a result, all the terms which satisfy $\vec{\mu} \cdot (\vec{m} - \vec{n}) \neq 2n\pi$ can be removed. This is why normal form transformation simplifies the transfer map.

Futhermore, there are two categories of terms that remain in the normal form map. The first category satisfies $\vec{m} = \vec{n}$, i.e.,

$$m_x = n_x \quad \text{and} \quad m_y = n_y, \tag{3.29}$$

which does not depend on $\vec{\mu}$. Since these terms cannot be removed intrinsically by any normal form transformation, they are the minimum independent conditions for obtaining an achromat. Note that they are also responsible for the amplitude-dependent tune shifts.

The other category of terms remaining in the normal form map consists of the nontrivial solutions to the equation

$$\vec{\mu} \cdot (\vec{m} - \vec{n}) = 2n\pi, \tag{3.30}$$

which are tune-dependent and can be removed by carefully choosing the tunes or by having a certain symmetry. The tunes that give eq. (3.30) non-trivial solutions are called resonances; and the terms associated with the solutions are called the driving terms.

In the case of the second order, $\vec{m}$ and $\vec{n}$ must also satisfy

$$m_x + n_x + m_y + n_y \leq 3. \tag{3.31}$$

Therefore, from eqs. (3.29) and (3.31), the tune-shift terms are $g_{110001} s_x \bar{s}_x \delta$, $g_{001111} s_y \bar{s}_y \delta$ and $g_{000003} \delta^3$, and from eqs. (3.30) and (3.31) the resonance driving terms can also be obtained.

Since the vector (x, a, y, b, δ) is real, from

$$\begin{pmatrix} s_x \\ \bar{s}_x \\ s_y \\ \bar{s}_y \\ \delta \end{pmatrix} = A \begin{pmatrix} x \\ a \\ y \\ b \\ \delta \end{pmatrix} \quad \text{we have} \quad A^{-1} \begin{pmatrix} s_x \\ \bar{s}_x \\ s_y \\ \bar{s}_y \\ \delta \end{pmatrix} = \begin{pmatrix} x \\ a \\ y \\ b \\ \delta \end{pmatrix} \quad \text{is real.}$$

Therefore, from the fact that all coefficients in f_3 are real, we have

$$g_3(s_x, \bar{s}_x, s_y, \bar{s}_y, \delta) = f_3(A^{-1}\vec{I}) \quad \text{is real.}$$

Thus we obtain

$$\begin{aligned} & \sum_{m_x n_x m_y n_y i_\delta} g_{m_x n_x m_y n_y i_\delta} s_x^{m_x} \bar{s}_x^{n_x} s_y^{m_y} \bar{s}_y^{n_y} \delta^{i_\delta} \\ = & \sum_{m_x n_x m_y n_y i_\delta} \bar{g}_{m_x n_x m_y n_y i_\delta} \bar{s}_x^{m_x} s_x^{n_x} \bar{s}_y^{m_y} s_y^{n_y} \delta^{i_\delta} \\ = & \sum_{m_x n_x m_y n_y i_\delta} \bar{g}_{n_x m_x n_y m_y i_\delta} s_x^{m_x} \bar{s}_x^{n_x} s_y^{m_y} \bar{s}_y^{n_y} \delta^{i_\delta}. \end{aligned}$$

Since this relation holds for any point, the coefficient of each term has to be equal separately, which gives

$$g_{m_x n_x m_y n_y i_\delta} = \bar{g}_{n_x m_x n_y m_y i_\delta}. \tag{3.32}$$

When $m_x = n_x$ and $m_y = n_y$, we have

$$g_{m_x m_x m_y m_y i_\delta} = \bar{g}_{m_x m_x m_y m_y i_\delta}, \tag{3.33}$$

which means that the coefficients of the tune-shift terms are real. Therefore, there are only two independent chromatic aberrations left when driving terms do not appear, which proves **Theorem B** from the last section.

Note that we have

$$G_{m_x n_x m_y n_y i_\delta} = \frac{g_{m_x n_x m_y n_y i_\delta}}{1 - e^{i\vec{\mu}\cdot(\vec{m}-\vec{n})}} \tag{3.34}$$

for ter

$\bar{G}_{n,}$

entai

coeff

is re

I

such

the

reso

are

ter

O

for terms that satisfy $\vec{\mu}\cdot(\vec{m}-\vec{n}) \neq 2n\pi$. Thus, the relation

$$
\begin{aligned}
\bar{G}_{n_x m_x n_y m_y i_\delta} &= \frac{\bar{g}_{n_x m_x n_y m_y i_\delta}}{1-e^{-i\vec{\mu}\cdot(\vec{n}-\vec{m})}} \\
&= \frac{g_{m_x n_x m_y n_y i_\delta}}{1-e^{i\vec{\mu}\cdot(\vec{m}-\vec{n})}} \\
&= G_{m_x n_x m_y n_y i_\delta}
\end{aligned}
\tag{3.35}
$$

entails that this part of G_3 is real. Since the rest of G_3 has no effect on $\vec{N}_2$, the coefficients can be chosen as zeroes. Therefore, G_3 is real. It also follows that $G_3(R\vec{I})$ is real.

Regarding the driving terms, one way to eliminate them is to choose the tunes such that no resonances equal to or below order 3 occur; the other way is to choose the tunes such that some resonances are avoided and the driving terms of the present resonances are cancelled by midplane symmetry.

In the case of Brown's second-order achromats, $\mu_x = \mu_y = (q/p)2\pi$, where q and p are natural numbers, $q < p$, and q/p is an irreducible fraction. Therefore, the driving terms are the non-trivial solutions of

$$
\frac{q}{p}(m_x - n_x + m_y - n_y) = n \quad (n \text{ is an integer}). \tag{3.36}
$$

On the other hand, $2p \geq 4$ and $m_x + n_x + m_y + n_y \leq 3$ entail that

$$
m_x - n_x + m_y - n_y = -p,\ 0,\ p. \tag{3.37}
$$

1) <u>$p > 3$</u>:

Equation (3.37) can be further reduced to

$$
m_x - n_x + m_y - n_y = 0. \tag{3.38}
$$

Equation (3.38) has these non-trivial solutions

$$
\left\{ \begin{array}{rcr} m_x - n_x &=& 1 \\ m_y - n_y &=& -1 \end{array} \right. \quad \text{and} \quad \left\{ \begin{array}{rcr} m_x - n_x &=& -1 \\ m_y - n_y &=& 1 \end{array} \right. , \tag{3.39}
$$

which can be transformed to

$$\begin{cases} m_x &= 1 \\ n_x &= 0 \\ m_y &= 0 \\ n_y &= 1 \end{cases} \text{and} \begin{cases} m_x &= 0 \\ n_x &= 1 \\ m_y &= 1 \\ n_y &= 0 \end{cases}. \tag{3.40}$$

These solutions entail that the driving terms are $s_x\bar{s}_y\delta$ and $\bar{s}_x s_y\delta$. Since midplane symmetry cancels them, only tune-shift terms are left in such a system, which requires at least four cells to be a first-order achromat.

2) $\underline{p = 3}$:

From the requirements of midplane symmetry discussed above, eq. (3.37) can be reduced to

$$m_x - n_x + m_y - n_y = -3,\ 3. \tag{3.41}$$

With midplane symmetry taken into account, it has the non-trivial solutions

$$\begin{cases} m_x - n_x &= -3,\ -1 \\ m_y - n_y &= 0,\ -2 \end{cases} \text{and} \begin{cases} m_x - n_x &= 3,\ 1 \\ m_y - n_y &= 0,\ 2 \end{cases}, \tag{3.42}$$

which are equivalent to

$$\begin{cases} m_x &= 0 \\ n_x &= 3 \\ m_y &= 0 \\ n_y &= 0 \end{cases}, \begin{cases} m_x &= 0 \\ n_x &= 1 \\ m_y &= 0 \\ n_y &= 2 \end{cases}, \begin{cases} m_x &= 3 \\ n_x &= 0 \\ m_y &= 0 \\ n_y &= 0 \end{cases}, \text{and} \begin{cases} m_x &= 1 \\ n_x &= 0 \\ m_y &= 2 \\ n_y &= 0 \end{cases}. \tag{3.43}$$

Hence the driving terms are s_x^3, $s_x s_y^2$, $\bar{s}_x^3$ and $\bar{s}_x\bar{s}_y^2$. This shows that not all second-order geometric aberrations vanish in a three-cell system, which agrees with **Theorem A**.

3) $\underline{p = 2}$:

Similar to case 2), we have

$$m_x - n_x + m_y - n_y = -2,\ 2, \tag{3.44}$$

which gives the non-trivial solutions

$$\left\{\begin{array}{lcl} m_x - n_x & = & -2, \quad 0 \\ m_y - n_y & = & 0, \quad -2 \end{array}\right. \text{ and } \left\{\begin{array}{lcl} m_x - n_x & = & 2, \quad 0 \\ m_y - n_y & = & 0, \quad 2 \end{array}\right. . \tag{3.45}$$

Hence we have

$$\left\{\begin{array}{lcl} m_x & = & 0 \\ n_x & = & 2 \\ m_y & = & 0 \\ n_y & = & 0 \end{array}\right., \left\{\begin{array}{lcl} m_x & = & 0 \\ n_x & = & 0 \\ m_y & = & 0 \\ n_y & = & 2 \end{array}\right., \left\{\begin{array}{lcl} m_x & = & 2 \\ n_x & = & 0 \\ m_y & = & 0 \\ n_y & = & 0 \end{array}\right. \text{ and } \left\{\begin{array}{lcl} m_x & = & 0 \\ n_x & = & 0 \\ m_y & = & 2 \\ n_y & = & 0 \end{array}\right., \tag{3.46}$$

which entail that the driving terms are $s_x^2\delta$, $\bar{s}_x^2\delta$, $s_y^2\delta$ and $\bar{s}_y^2\delta$. This shows that all second-order geometric aberrations vanish, yet not all driving terms disappear. From the proof of **Theorem B** given by D. Carey (see eq. (18) in [Carey81]), it is clear that there are two independent second-order chromatic aberrations only when $\mu \neq \pi$. In conclusion, normal form theory gives an alternative proof of K. Brown's theorems.

Now let us move to the third order. From eq. (3.23) and (3.24), we have

$$\vec{N}_2 = (\exp(:G_3:)\vec{I}) \circ \vec{N}_1 \circ (\exp(:-G_3:)\vec{I}) \tag{3.47}$$

$$= (\exp(:G_3:)\vec{I}) \circ (R\vec{I}) \circ (\exp(:g_3:)\vec{I}) \tag{3.48}$$

$$\circ(\exp(:g_4:)\vec{I}) \circ (\exp(:-G_3:)\vec{I}) \tag{3.49}$$

$$=_3 \exp(:-G_3:)\exp(:g_4:)\exp(:g_3:)\exp(:G_3(R\vec{I}):)(R\vec{I})$$

$$=_3 \exp(:g_3 - G_3 + G_3(R\vec{I}):)$$

$$\exp(:g_4 + \frac{1}{2}([g_3, G_3] + [g_3, G_3(R\vec{I})] + [G_3(R\vec{I}), G_3]):)(R\vec{I})$$

$$=_3 \exp(:g_4 + \frac{1}{2}([g_3, G_3] + [g_3, G_3(R\vec{I})] + [G_3(R\vec{I}), G_3]):)$$

$$\exp(:g_3 - G_3 + G_3(R\vec{I}):)(R\vec{I})$$

$$=_3 (R\vec{I}) \circ (\exp(:h_3:)\vec{I}) \circ (\exp(:h_4:)\vec{I}), \tag{3.50}$$

where

$$h_3 = g_3 - G_3 + G_3(R\vec{I}), \tag{3.51}$$

$$h_4 = g_4 + \frac{1}{2}([g_3, G_3] + [g_3, G_3(R\vec{I})] + [G_3(R\vec{I}), G_3]). \tag{3.52}$$

Now define $\vec{N}_3$ as

$$\vec{N}_3 = (\exp(:H_4:)\vec{I}) \circ \vec{N}_2 \circ (\exp(:-H_4:)\vec{I}). \tag{3.53}$$

To the third order, we have

$$\begin{aligned}
\vec{N}_3 &=_3 (\exp(:H_4(R\vec{I}):)R\vec{I}) \circ (\exp(:h_3:)\vec{I}) \circ (\exp(:h_4:)\vec{I}) \circ (\exp(:-H_4:)\vec{I}) \\
&=_3 \exp(:-H_4:)\exp(:h_4:)\exp(:h_3:)\exp(:H_4(R\vec{I}):)(R\vec{I}) \\
&=_3 \exp(:h_3 + h_4 - (H_4 - H_4(R\vec{I})):)(R\vec{I}),
\end{aligned} \tag{3.54}$$

where use of the B-C-H formula has been made.

Similar to the second order, when resonances are avoided and/or driving terms not present, only tune-shift terms are left, which are $h_{22000}(s_x\bar{s}_x)^2$, $h_{11110}(s_x\bar{s}_x)s_y\bar{s}_y$, $h_{00220}(s_y\bar{s}_y)^2$, $h_{11002}s_x\bar{s}_x\delta^2$, $h_{00112}s_y\bar{s}_y\delta^2$, and $h_{00004}\delta^4$. Therefore, the total pseudo-Hamiltonian is

$$\begin{aligned}
h_T &= g_{11001}s_x\bar{s}_x\delta + g_{00111}s_y\bar{s}_y\delta + g_{00003}\delta^3 \\
&\quad + h_{22000}(s_x\bar{s}_x)^2 + h_{11110}s_x\bar{s}_x s_y\bar{s}_y h_{00220}(s_y\bar{s}_y)^2 \\
&\quad + h_{11002}s_x\bar{s}_x\delta^2 + h_{00112}s_y\bar{s}_y\delta^2 + h_{00004}\delta^4.
\end{aligned} \tag{3.55}$$

Since g_3, g_4, G_3 and $G_3(R\vec{I})$ are real, h_4 is also real. Hence, $h_{m_x m_x m_y m_y i_\delta}$ is real. Therefore, 5 third-order knobs are needed to achieve a third-order achromat.

Define real coordinates

$$X, Y = (s_{x,y} + \bar{s}_{x,y})/2, \tag{3.56}$$

$$P_X, P_Y = (s_{x,y} - \bar{s}_{x,y})/2i, \tag{3.57}$$

where X and P_X are the real linear combinations of x and a, and Y and P_Y are the real linear combinations of y and b. The linear matrix in the new coordinates is

$$R = \begin{pmatrix} \cos(\mu_x) & \sin(\mu_x) & & & \\ -\sin(\mu_x) & \cos(\mu_x) & & & \\ & & \cos(\mu_y) & \sin(\mu_y) & \\ & & -\sin(\mu_y) & \cos(\mu_y) & \\ & & & & 1 \end{pmatrix}, \quad (3.58)$$

where the linear motion is simply a rotation of X and P_X, and Y and P_Y. Usually the X-P_X space is called the normal form space. Futhermore, the pseudo-Hamiltonian has the form

$$\begin{aligned} h_T &= g_{11001}(X^2+P_X^2)\delta + g_{00111}(Y^2+P_Y^2)\delta + g_{00003}\delta^3 \\ &\quad + h_{22000}(X^2+P_X^2)^2 + h_{11110}(X^2+P_X^2)(Y^2+P_Y^2)h_{00220}(Y^2+P_Y^2)^2 \\ &\quad + h_{11002}(X^2+P_X^2)\delta^2 + h_{00112}(Y^2+P_Y^2)\delta^2 + h_{00004}\delta^4. \\ &= w'_x(X^2+P_X^2)\delta + w'_y(Y^2+P_Y^2)\delta + e\delta^3 \\ &\quad + a(X^2+P_X^2)^2 + b(X^2+P_X^2)(Y^2+P_Y^2)c(Y^2+P_Y^2)^2 \\ &\quad + w''_x(X^2+P_X^2)\delta^2 + w''_x(Y^2+P_Y^2)\delta^2 + f\delta^4. \end{aligned} \quad (3.59)$$

Here, w'_x and w'_y are first-order chromaticities; w''_x and w''_y are second-order chromaticities; a,b and c are anharmonicities.

In summary, a third-order achromat is achieved when the total tunes of both transverse planes are integers, all resonances of order 4 and below are avoided or all driving terms of present resonances vanish, and all first- and second-order chromaticities and anharmonicities are corrected.

When a third-order achromat is reached in the normal form coordinates, we have

$$\vec{N}_3 \;=_3\; (R\vec{I}) \circ (\exp(:e\delta^3 + f\delta^4:)\vec{I}). \quad (3.60)$$

Therefore, the total map of the n-cell achromat is

$$\vec{N}_3^n \;=_3\; ((R\vec{I}) \circ (\exp(:e\delta^3 + f\delta^4:)\vec{I}))^n$$

$$=_3 \quad \exp(n : e\delta^3 + f\delta^4 :)\vec{I}. \tag{3.61}$$

Since the normal form transformation

$$\vec{A} = (\exp(: H_4 :)\vec{I}) \circ (\exp(: G_3 :)\vec{I}) \circ (A\vec{I}) \tag{3.62}$$

does not contain t, $\vec{N}_3^n$ and $\vec{A}$ commute. As a result, the total map in the original coordinates is

$$\begin{aligned}
\vec{M}_T &=_3 \quad \vec{A} \circ \vec{N}_3^n \circ \vec{A}^{-1} \\
&=_3 \quad (\exp(: H_4 :)\vec{I}) \circ (\exp(: G_3 :)\vec{I}) \circ (A\vec{I}) \circ \vec{N}_3^n \\
&\qquad \circ(A^{-1}\vec{I}) \circ (\exp(: -G_3 :)\vec{I}) \circ (\exp(: -H_4 :)\vec{I}) \\
&=_3 \quad \exp(: -H_4 :)\exp(: -G_3 :)\exp(n : e\delta^3 + f\delta^4 :)\exp(: G_3 :)\exp(: H_4 :)\vec{I} \\
&=_3 \quad \exp(n : e\delta^3 + f\delta^4 :)\vec{I} \;=\; \vec{N}_3^n.
\end{aligned} \tag{3.63}$$

Indeed, $\vec{M}_T$ is a third-order achromat, and the only aberrations left are $(t|\delta^i)$, where $i = 1, 2, \cdots, n$.

Let us come back to the examples designed by Dragt and Neri. Dragt's example is very straightforward, because all resonances of order 4 and below are avoided. Neri's example is more complicated because some resonances are present and their driving terms are cancelled by midplane symmetry. In his example, the tunes are $\mu_x/2\pi = 1/7$ and $\mu_y/2\pi = 2/7$, which have the driving terms satisfying

$$m_x - n_x + 2(m_y - n_y) = -7, \ 0, \ 7. \tag{3.64}$$

It gives the non-trivial solutions

$$\left\{ \begin{array}{lcrrrr} m_x - n_x & = & -1, & 2, & -2, & 1 \\ m_y - n_y & = & -3, & -1, & 1, & 3 \end{array} \right. \tag{3.65}$$

However, midplane symmetry requires that $m_y + n_y$ be even, which entails that $m_y - n_y$ also be even. Hence, all the terms are cancelled.

As a summary of this chapter, let us look at the perspective of extending Dragt's theory to higher orders. If we follow Neri's choice of resonance, it can be shown that the minimum number of cells for a fourth-order achromat is 11, with $T_x = 1/11$ and $T_y = 2/11$. For fifth-order achromats, Neri's approach fails because the driving terms $\bar{s}_x^4 s_y^2$ and $s_x^4 \bar{s}_y^2$ cannot be cancelled by midplane symmetry. To circumvent this difficulty, different choices of the tunes have to be made. The next choices are $T_x = 3T_y$ or $3T_x = T_y$. The first one gives a smaller minimum number of cells, which is 17, and the tunes are $T_x = 3/17$ and $T_y = 1/17$. These examples show that there is no established pattern for the choices of the tunes or the number of the cells to design an arbitrary-order achromat, and the driving terms for a resonance increase rapidly with the order, as does the minimum number of cells. The theory presented in the next chapter allows us to solve these difficulties.

Chapter 4

Arbitrary-Order Achromat Theory

In the chapter that follows, an analytical theory of arbitrary-order achromats will be developed. In Section 4.1, the maps of cells R, S, and C will be derived from the map of the forward cell, and that of a four-cell system from those of single cells. Section 4.2 contains a classification of the systems with the best solutions. In the first part (Section 4.2.3), it is shown that it is necessary to have at least four cells in a system to achieve an efficient arbitrary-order achromat. Then the proof of the existence of an optimal solution is given. It is further shown that 4 out of 64 four-cell systems give the optimal solution while requiring a minimum number of linear constraints (Section 4.2.4). In Section 4.3, the four best systems are studied in detail to determine the solutions for achromats order by order. First, a general solution for arbitrary-order achromats is obtained, even though it is not the optimal solution we can obtain from this theory, regarding the number of conditions that have to be satisfied. Then the optimal solutions for achromats up to the sixth order are found.

4.1 Map Representations

Let us consider a phase space consisting of $2m$ variables $(q_1, \cdots, q_m, p_1, \cdots, p_m)$. 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. Therefore, its transfer map $\vec{M}$ of order n can be represented by a matrix L and a polynomial H of orders 3 up to $n+1$ through Lie factorization (Theorem 2.6) via

$$\vec{M} =_n (L\vec{I}) \circ (\exp(: H :)\vec{I}). \tag{4.1}$$

Furthermore, its inverse is (eq. 2.48)

$$\vec{M}^{-1} =_n (\exp(- : H :)\vec{I}) \circ (L^{-1}\vec{I}). \tag{4.2}$$

Next, let us define a "standard" and a "sub-standard" form of the maps. The advantage of these forms will become evident later.

Definition 4.1 *For a symplectic map $\vec{M_S}$, the standard form is defined as*

$$\vec{M_S} = \exp(: H :)(M_L\vec{I}), \tag{4.3}$$

where H is the pseudo-Hamiltonian of orders three and up, and M_L is the linear matrix. A representation of the form

$$\vec{M_S} = \left[\prod_i \exp(: H_i :)\right](M_L\vec{I}) \tag{4.4}$$

is called a sub-standard form.

Note again the difference between eqs. (4.1) and (4.3), where, in the former equation, $\exp(: H :)$ acts on $\vec{I}$ and the resulting map is then composed to the linear map, and, in the latter equation, $\exp(: H :)$ acts on the linear map directly. Like the composition "o", Lie operators are are also associative, which implies the associativity of Lie transformations. Apparently, use of the Baker-Campbell-Hausdorff formula allows the transformation of a sub-standard form into a standard form.

From Theorem 2.1, $\vec{M}$ can be written in the standard form

$$\vec{M}^F = \exp(:H:)(L\vec{I}), \tag{4.5}$$

which is called the forward map $\vec{M}^F$.

To obtain the maps of R, S, and C cells, Theorem 2.4 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, δ_i) and exits it at a final point (x_f, a_f, y_f, b_f, δ_f), a particle which enters the reversed cell at (x_f, $-a_f$, y_f, $-b_f$, δ_f) will exit at (x_i, $-a_i$, y_i, $-b_i$, δ_i). This determines that the map of the reversed cell is

$$\vec{M}^R = (R\vec{I}) \circ \vec{M}^{-1} \circ (R^{-1}\vec{I}), \tag{4.6}$$

where

$$R\vec{I} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ a \\ y \\ b \\ \delta \end{pmatrix}. \tag{4.7}$$

Taking into account the fact that R is antisymplectic (see Appendix A), we can get the standard form of $\vec{M}^R$:

$$\begin{aligned} \vec{M}^R &= (R\vec{I}) \circ \vec{M}^{-1} \circ (R^{-1}\vec{I}) \\ &= (R\vec{I}) \circ (\exp(:-H:)\vec{I}) \circ (L^{-1}\vec{I}) \circ (R^{-1}\vec{I}) \\ &= (R\vec{I}) \circ (\exp(:-H:)\vec{I}) \circ (L^{-1}R^{-1}\vec{I}) \end{aligned}$$

$$= (R\vec{I}) \circ (\exp(: H(L^{-1}R^{-1}\vec{I}) :)(L^{-1}R^{-1}\vec{I})) \quad \text{(Theorem 2.4)}$$

$$= \exp(: H(L^{-1}R^{-1}\vec{I}) :)((R\vec{I}) \circ (L^{-1}R^{-1}\vec{I})) \quad \text{(Theorem 2.2)}$$

$$= \exp(: H(L^{-1}R^{-1}\vec{I}) :)(RL^{-1}R^{-1}\vec{I}). \tag{4.8}$$

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

$$\vec{M}^S = (S\vec{I}) \circ \vec{M} \circ (S^{-1}\vec{I}), \tag{4.9}$$

where

$$S\vec{I} = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ a \\ y \\ b \\ \delta \end{pmatrix}. \tag{4.10}$$

Since the matrix S is symplectic (see Appendix A), we have

$$\begin{aligned} \vec{M}^S &= (S\vec{I}) \circ \vec{M} \circ (S^{-1}\vec{I}) \\ &= (S\vec{I}) \circ (L\vec{I}) \circ (\exp(: H :)\vec{I}) \circ (S^{-1}\vec{I}) \\ &= (SL\vec{I}) \circ (\exp(: H(S^{-1}\vec{I}) :)(S^{-1}\vec{I})) \\ &= \exp(: H(S^{-1}\vec{I}) :)(SLS^{-1}\vec{I}). \end{aligned} \tag{4.11}$$

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

$$\vec{M}^C = (S\vec{I}) \circ (R\vec{I}) \circ \vec{M}^{-1} \circ (R^{-1}\vec{I}) \circ (S^{-1}\vec{I}). \tag{4.12}$$

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

$$\vec{M}^C = \exp(: H(L^{-1}R^{-1}S^{-1}\vec{I}) :)(SRL^{-1}R^{-1}S^{-1}\vec{I}). \tag{4.13}$$

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

$$\vec{M}^F = \exp(:H:)(L\vec{I}), \tag{4.14}$$

$$\vec{M}^R = \exp(:H(L^{-1}R^{-1}\vec{I}):)(RL^{-1}R^{-1}\vec{I}), \tag{4.15}$$

$$\vec{M}^S = \exp(:H(S^{-1}\vec{I}):)(SLS^{-1}\vec{I}), \tag{4.16}$$

$$\vec{M}^C = \exp(:H(L^{-1}R^{-1}S^{-1}\vec{I}):)(SRL^{-1}R^{-1}S^{-1}\vec{I}). \tag{4.17}$$

Since we have the maps of the different kinds of cells, we are ready to construct the map of any given multi-cell system. As examples, the total maps of a few four-cell systems are presented here. First, let us define a symbol denoting the map of a multi-cell system.

Definition 4.2 *Let C_i be the ith cell in a k-cell system, i.e., C_i can be F, R, S, or C. 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. Next we will determine the sub-standard form for a variety of four-cell combinations, which will be very useful later. As a proof of principle, we show this process for system FRSC. In the transformation, we repeatedly make use of eq. (2.1), Theorem 2.4, and the associativity of "$\circ$". From the definition of system FRSC, we can obtain its transfer map from the maps of single cells, which is

$$\vec{M}^{FRSC} = \vec{M}^C \circ \vec{M}^S \circ \vec{M}^R \circ \vec{M}^F. \tag{4.18}$$

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. Thus, $\vec{M}^{FRSC}$ is transformed to

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

$$\circ[\exp(:H(S^{-1}\vec{I}):)(SLS^{-1}\vec{I})]$$

$$\circ[\exp(:H(L^{-1}R^{-1}\vec{I}):)(RL^{-1}R^{-1}\vec{I})]\circ[\exp(:H(\vec{I}):)(L\vec{I})]$$

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

$$\circ[\exp(:H(S^{-1}\vec{I}):)(SLS^{-1}\vec{I})]$$

$$\circ[\exp(:H(L^{-1}R^{-1}\vec{I}):)(RL^{-1}R^{-1}\vec{I})]\circ(L\vec{I})\}$$

(Theorem 2.2)

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

$$\circ[\exp(:H(S^{-1}\vec{I}):)(SLS^{-1}\vec{I})]$$

$$\circ[\exp(:H(L^{-1}R^{-1}\cdot L\vec{I}):)(RL^{-1}R^{-1}\cdot L\vec{I})]\}$$

(Theorem 2.4)

$$= \exp(:H(\vec{I}):)\exp(:H(L^{-1}R^{-1}\cdot L\vec{I}):)$$

$$\{[\exp(:H(L^{-1}R^{-1}S^{-1}\vec{I}):)(SRL^{-1}R^{-1}S^{-1}\vec{I})]$$

$$\circ[\exp(:H(S^{-1}\vec{I}):)(SLS^{-1}\vec{I})]\circ(RL^{-1}R^{-1}\cdot L\vec{I})\}$$

(Theorem 2.2)

$$= \exp(:H(\vec{I}):)\exp(:H(L^{-1}R^{-1}\cdot L\vec{I}):)$$

$$\{[\exp(:H(L^{-1}R^{-1}S^{-1}\vec{I}):)(SRL^{-1}R^{-1}S^{-1}\vec{I})]$$

$$\circ[\exp(:H(S^{-1}\cdot RL^{-1}R^{-1}\cdot L\vec{I}):)(SLS^{-1}\cdot RL^{-1}R^{-1}\cdot L\vec{I})]\}$$

(Theorem 2.4)

$$= \exp(:H(\vec{I}):)\exp(:H(L^{-1}R^{-1}\cdot L\vec{I}):)$$

$$\exp(:H(S^{-1}\cdot RL^{-1}R^{-1}\cdot L\vec{I}):)\{[\exp(:H(L^{-1}R^{-1}S^{-1}\vec{I}):)$$

$$(SRL^{-1}R^{-1}S^{-1}\vec{I})]\circ(SLS^{-1}\cdot RL^{-1}R^{-1}\cdot L\vec{I})\}$$

(Theorem 2.2)

$$= \exp(:H(\vec{I}):)\exp(:H(L^{-1}R^{-1}\cdot L\vec{I}):)$$

$$\exp(:H(S^{-1}\cdot RL^{-1}R^{-1}\cdot L\vec{I}):)$$
$$\exp(:H(L^{-1}R^{-1}\cdot LS^{-1}\cdot RL^{-1}R^{-1}\cdot L\vec{I}):)$$
$$(SRL^{-1}R^{-1}\cdot LS^{-1}\cdot RL^{-1}R^{-1}\cdot L\vec{I}),$$

(Theorem 2.4)

which is now in the sub-standard form. In a similar way, the sub-standard forms of the maps of the systems $\vec{M}^{FRFR}$, $\vec{M}^{FCSR}$, and $\vec{M}^{FCFC}$ are obtained. Together with $\vec{M}^{FRSC}$, we have

$$\begin{aligned}
\vec{M}^{FRSC} &= \exp(:H(\vec{I}):)\exp(:H(L^{-1}R^{-1}\cdot L\vec{I}):) \\
&\quad \exp(:H(S^{-1}\cdot RL^{-1}R^{-1}\cdot L\vec{I}):) \\
&\quad \exp(:H(L^{-1}R^{-1}\cdot LS^{-1}\cdot RL^{-1}R^{-1}\cdot L\vec{I}):) \\
&\quad (SRL^{-1}R^{-1}\cdot LS^{-1}\cdot RL^{-1}R^{-1}\cdot L\vec{I}), \qquad (4.19) \\
\vec{M}^{FRFR} &= \exp(:H(\vec{I}):)\exp(:H(L^{-1}R^{-1}\cdot L\vec{I}):) \\
&\quad \exp(:H(RL^{-1}R^{-1}\cdot L\vec{I}):) \\
&\quad \exp(:H(L^{-1}R^{-1}\cdot L\cdot RL^{-1}R^{-1}\cdot L\vec{I}):) \\
&\quad (RL^{-1}R^{-1}\cdot L\cdot RL^{-1}R^{-1}\cdot L\vec{I}), \qquad (4.20) \\
\vec{M}^{FCSR} &= \exp(:H(\vec{I}):)\exp(:H(L^{-1}R^{-1}S^{-1}\cdot L\vec{I}):) \\
&\quad \exp(:H(RL^{-1}R^{-1}S^{-1}\cdot L\vec{I}):) \\
&\quad \exp(:H(L^{-1}R^{-1}\cdot SL\cdot RL^{-1}R^{-1}S^{-1}\cdot L\vec{I}):) \\
&\quad (RL^{-1}R^{-1}\cdot SL\cdot RL^{-1}R^{-1}S^{-1}\cdot L\vec{I}), \qquad (4.21) \\
\vec{M}^{FCFC} &= \exp(:H(\vec{I}):)\exp(:H(L^{-1}R^{-1}S^{-1}\cdot L\vec{I}):) \\
&\quad \exp(:H(SRL^{-1}R^{-1}S^{-1}\cdot L\vec{I}):) \\
&\quad \exp(:H(L^{-1}R^{-1}S^{-1}\cdot L\cdot SRL^{-1}R^{-1}S^{-1}\cdot L\vec{I}):) \\
&\quad (SRL^{-1}R^{-1}S^{-1}\cdot L\cdot SRL^{-1}R^{-1}S^{-1}\cdot L\vec{I}). \qquad (4.22)
\end{aligned}$$

As shown in the next section, only those 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 sub-standard 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 sub-standard form using the B-C-H formula.

4.2 Optimal Four-Cell Systems

In this section, we first study general multi-cell systems and then all possible four-cell systems using the maps obtained in the last section. The goal is to find the systems that require the fewest conditions in order to be converted to achromats of a given order. A few definitions have to be mentioned before the study can be started. They are the keys to the proofs of the theorems discussed later in this section.

Like the previous 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. (4.5) with its pseudo-Hamiltonian given by eq. (3.8), which we write as

$$H = \sum_{i_x i_a i_y i_b i_\delta} C_{i_x i_a i_y i_b i_\delta} x^{i_x} a^{i_a} y^{i_y} b^{i_b} \delta^{i_\delta}, \tag{4.23}$$

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

Definition 4.3 *Define the polynomials $A(H)$, $B(H)$, $C(H)$, and $D(H)$ as the parts of H which satisfy*

$$A(H) = \sum_{i_x i_a i_y i_b i_\delta} C_{i_x i_a i_y i_b i_\delta} x^{i_x} a^{i_a} y^{i_y} b^{i_b} \delta^{i_\delta} \quad (i_x + i_a \text{ is odd}, \; i_a + i_b \text{ is even}), \tag{4.24}$$

$$B(H) = \sum_{i_x i_a i_y i_b i_\delta} C_{i_x i_a i_y i_b i_\delta} x^{i_x} a^{i_a} y^{i_y} b^{i_b} \delta^{i_\delta} \quad (i_x + i_a \text{ is odd}, \; i_a + i_b \text{ is odd}), \tag{4.25}$$

$$C(H) = \sum_{i_x i_a i_y i_b i_\delta} C_{i_x i_a i_y i_b i_\delta} x^{i_x} a^{i_a} y^{i_y} b^{i_b} \delta^{i_\delta} \ (i_x + i_a \text{ is even}, \ i_a + i_b \text{ is odd}), \quad (4.26)$$

$$D(H) = \sum_{i_x i_a i_y i_b i_\delta} C_{i_x i_a i_y i_b i_\delta} x^{i_x} a^{i_a} y^{i_y} b^{i_b} \delta^{i_\delta} \ (i_x + i_a \text{ is even}, \ i_a + i_b \text{ is even}). \quad (4.27)$$

Definition 4.4 *Define*

$$H^F = H(\vec{I}) = H(x, a, y, b), \quad (4.28)$$

$$H^R = H(R\vec{I}) = H(x, -a, y, -b), \quad (4.29)$$

$$H^S = H(S\vec{I}) = H(-x, -a, y, b), \quad (4.30)$$

$$H^C = H(RS\vec{I}) = H(-x, a, y, -b). \quad (4.31)$$

It is easy to show that

$$H^F = A(H) + B(H) + C(H) + D(H), \quad (4.32)$$

$$H^R = A(H) - B(H) - C(H) + D(H), \quad (4.33)$$

$$H^S = -A(H) - B(H) + C(H) + D(H), \quad (4.34)$$

$$H^C = -A(H) + B(H) - C(H) + D(H). \quad (4.35)$$

4.2.1 General Properties of k-Cell Systems

Consider a general system of k cells arranged using the above symmetry operations. Using Theorems (2.2) and 2.4 repeatedly, its map can be brought to the sub-standard form in a similar way as in eq. (4.19). The result has the form

$$\vec{M} = \exp(: H(\vec{I}) :) \exp(: H(M^{(1)}\vec{I}) :) \cdots \exp(: H(M^{(k-1)}\vec{I}) :)(M_T\vec{I}), \quad (4.36)$$

where M_T is the linear matrix of the system and

$$M^{(i)} = \begin{pmatrix} m_{11}^{(i)} & m_{12}^{(i)} & 0 & 0 & m_{15}^{(i)} \\ m_{11}^{(i)} & m_{12}^{(i)} & 0 & 0 & m_{15}^{(i)} \\ 0 & 0 & m_{33}^{(i)} & m_{34}^{(i)} & 0 \\ 0 & 0 & m_{43}^{(i)} & m_{44}^{(i)} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (i = 1, 2, \cdots, k-1) \quad (4.37)$$

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 B-C-H formula, $\vec{M}$ can be transformed to a single Lie operator acting on a linear map, which is

$$\vec{M} =_n \exp(: H(\vec{I}) + H(M^{(1)}\vec{I}) + \cdots + H(M^{(k-1)}\vec{I}) + \text{commutators} :)(M_T\vec{I}), \quad (4.38)$$

where the commutators are polynomials of order 4 and higher, and M_T is the linear matrix of the whole system.

For the time being, we restrict ourself to $\sum H(M^{(i)}\vec{I})$, hoping that systems which cancel the most number of terms in $\sum H(M^{(i)}\vec{I})$, called the optimal systems, do the same to the total pseudo-Hamiltonian. In Section 4.3, this will be shown to be true at least up to the sixth order. Due to the fact that all third-order terms in the total pseudo-Hamiltonian are contained in $H_3(\vec{I}) + H_3(M^{(1)}\vec{I}) + \cdots + H_3(M^{(k-1)}\vec{I})$, those not cancelled by symmetry must be cancelled by adjusting the second-order elements. Therefore, $\sum H_3(M^{(i)}\vec{I})$ is used to find the necessary conditions for the optimal systems that cancel the most number of terms in $\sum H_3(M^{(i)}\vec{I})$ (Theorems 4.2 and 4.3). Although the restriction of $\sum H_3(M^{(i)}\vec{I})$ makes it more difficult to prove Theorem 4.3, it makes the logical structure of the proof more straightforward.

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

Definition 4.5 *$D^+(H(\vec{I}))$ 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 $\sum_{n_x,n_a,n_y,n_b,i_\delta} C_{2n_x+1,2n_a+1,2n_y+1,2n_b+1,i_\delta} x^{2n_x+1} a^{2n_a+1} \; y^{2n_y+1} b^{2n_b+1} \delta^{i_\delta}$. *Here* n_x, n_a, n_y, n_b *are non-negative integers.*

Note that $D = D^+ + D^-$ because of eq. (4.27).

Theorem 4.1 *For a given 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 under the assumption that no relations are assumed among the Lie coefficients* $C_{i_x i_a i_y i_b i_\delta}$.

Proof:

From eq. (4.36) and (4.37), the sum of $H(M^{(i)}\vec{I})$ is

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

which entails that the sum of D^+ is

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

Since there are no connections demanded among $C_{2n_x,2n_a,2n_y,2n_b,i_\delta}$, the vanishing of a polynomial associated with $C_{2n_x,2n_a,2n_y,2n_b,i_\delta}$ requires that

$$x^{2n_x}a^{2n_a}y^{2n_y}b^{2n_b}+\sum_{i=1}^{k-1}(m_{11}^{(i)}x+m_{12}^{(i)}a+m_{15}^{(i)}\delta)^{2n_x}(m_{21}^{(i)}x+m_{22}^{(i)}a+m_{25}^{(i)}\delta)^{2n_a}$$

$$(m_{33}^{(i)}y+m_{34}^{(i)}b)^{2n_y}(m_{43}^{(i)}y+m_{44}^{(i)}b)^{2n_b}=0$$

for any point in the phase space. Due to the fact that all quantities that appear in this polynomial are real numbers, they can not be cancelled regardless of the choices for $M^{(i)}$. Therefore, these terms have to be cancelled separately when the achromats are designed.

4.2.2 Two- and Three-Cell Systems

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

Theorem 4.2 *Two- or Three-cell systems can not 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* $C_3(H(\vec{I}))+C_3(H(M^{(1)}\vec{I}))+\cdots+C_3(H(M^{(k-1)}\vec{I}))$.

Proof.

(1) <u>Two-cell systems:</u>

The sum of $H(M^{(i)}\vec{I})$ is

$$H(\vec{I})+H(M^{(1)}\vec{I})=\sum_{i_x,i_a,i_y,i_b,i_\delta}C_{i_xi_ai_yi_bi_\delta}\Big(x^{i_x}a^{i_a}y^{i_y}b^{i_b}\delta^{i_\delta}$$

$$+(m_{11}^{(1)}x+m_{12}^{(1)}a+m_{15}^{(1)}\delta)^{i_x}(m_{21}^{(1)}x+m_{22}^{(1)}a+m_{25}^{(1)}\delta)^{i_a}$$

$$(m_{33}^{(1)}y+m_{34}^{(1)}b)^{i_y}(m_{43}^{(1)}y+m_{44}^{(1)}b)^{i_y}\delta^{i_\delta}\Big).$$

Cancelling 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

$$\begin{cases} m_{11}^{(1)} &= -1, \\ m_{12}^{(1)} &= 0, \\ m_{15}^{(1)} &= 0. \end{cases}$$

and cancelling 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

$$\begin{cases} m_{22}^{(1)} &= -1, \\ m_{21}^{(1)} &= 0, \\ m_{25}^{(1)} &= 0; \end{cases}$$

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 cancelled.

(2) Three-cell systems:

Similarly, the sum of $H(M^{(i)}\vec{I})$ is

$$H(\vec{I}) + H(M^{(1)}\vec{I}) + H(M^{(2)}\vec{I}) = \sum_{i_x,i_a,i_y,i_b,i_\delta} C_{i_x i_a i_y i_b i_\delta} \left(x^{i_x}a^{i_a}y^{i_y}b^{i_b}\delta^{i_\delta}\right.$$

$$+(m_{11}^{(1)}x + m_{12}^{(1)}a + m_{15}^{(1)}\delta)^{i_x}(m_{21}^{(1)}x + m_{22}^{(1)}a + m_{25}^{(1)}\delta)^{i_a}$$
$$(m_{33}^{(1)}y + m_{34}^{(1)}b)^{i_y}(m_{43}^{(1)}y + m_{44}^{(1)}b)^{i_y}\delta^{i_\delta}$$
$$+(m_{11}^{(2)}x + m_{12}^{(2)}a + m_{15}^{(2)}\delta)^{i_x}(m_{21}^{(2)}x + m_{22}^{(2)}a + m_{25}^{(2)}\delta)^{i_a}$$
$$(m_{33}^{(2)}y + m_{34}^{(2)}b)^{i_y}(m_{43}^{(2)}y + m_{44}^{(2)}b)^{i_y}\delta^{i_\delta}\Big).$$

Cancellation of the terms with $C_{1,0,0,0,2}$ requires that

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

which entails that

$$\left\{\begin{array}{l} 1 + m_{11}^{(1)} + m_{11}^{(2)} = 0 \\ m_{12}^{(1)} + m_{12}^{(2)} = 0 \\ m_{15}^{(1)} + m_{15}^{(2)} = 0 \end{array}\right., \quad \Rightarrow \left\{\begin{array}{l} m_{11}^{(2)} = -(1 + m_{11}^{(1)}) \\ m_{12}^{(1)} = -m_{12}^{(2)} \\ m_{15}^{(1)} = -m_{15}^{(2)} \end{array}\right. . \tag{4.39}$$

Now, let us look at those terms with $C_{3,0,0,0,0}$, which have the form

$$\begin{aligned} & x^3 + (m_{11}^{(1)}x + m_{12}^{(1)}a + m_{15}^{(1)}\delta)^3 + (m_{11}^{(2)}x + m_{12}^{(2)}a + m_{15}^{(2)}\delta)^3 \\ & \quad = x^3 + (m_{11}^{(1)}x + m_{12}^{(1)}a)^3 + 3(m_{11}^{(1)}x + m_{12}^{(1)}a)^2 m_{15}^{(1)}\delta \\ & \quad +3(m_{11}^{(1)}x + m_{12}^{(1)}a)(m_{15}^{(1)})^2\delta^2 + (m_{15}^{(1)})^3\delta^3 \\ & \quad +(m_{11}^{(2)}x + m_{12}^{(2)}a)^3 + 3(m_{11}^{(2)}x + m_{12}^{(2)}a)^2 m_{15}^{(2)}\delta \\ & \quad +3(m_{11}^{(2)}x + m_{12}^{(2)}a)(m_{15}^{(2)})^2\delta^2 + (m_{15}^{(2)})^3\delta^3 \end{aligned} \tag{4.40}$$

$$\begin{aligned} & \quad = (1 + (m_{11}^{(1)})^3 + (m_{11}^{(2)})^3)x^3 + 3((m_{11}^{(1)})^2 m_{12}^{(1)} + (m_{11}^{(2)})^2 m_{12}^{(2)})x^2 a \\ & \quad +3(m_{11}^{(1)}(m_{12}^{(1)})^2 + m_{11}^{(2)}(m_{12}^{(2)})^2)xa^2 + ((m_{12}^{(1)})^3 + (m_{12}^{(2)})^3)a^3 \\ & \quad +3((m_{11}^{(1)})^2 m_{15}^{(1)} + (m_{11}^{(2)})^2 m_{15}^{(2)})x^2\delta + 6(m_{11}^{(1)}m_{12}^{(1)}m_{15}^{(1)} + m_{11}^{(2)}m_{12}^{(2)}m_{15}^{(2)})xa\delta \\ & \quad +3((m_{12}^{(1)})^2 m_{15}^{(1)} + (m_{12}^{(2)})^2 m_{15}^{(2)})a^2\delta + 3(m_{11}^{(1)}(m_{15}^{(1)})^2 + m_{11}^{(2)}(m_{15}^{(2)})^2)x\delta^2 \\ & \quad +3(m_{12}^{(1)}(m_{15}^{(1)})^2 + m_{12}^{(2)}(m_{15}^{(2)})^2)a\delta^2 + ((m_{15}^{(1)})^3 + (m_{15}^{(2)})^3)\delta^3. \end{aligned} \tag{4.41}$$

Similar to the case of the two-cell system, each term in eq. (4.41) has to vanish separately.

Inserting $m_{11}^{(2)} = -(1+m_{11}^{(1)})$ from eq. (4.39) into $(1+(m_{11}^{(1)})^3 + (m_{11}^{(2)})^3) = 0$ from eq. (4.41), we have

$$\begin{aligned} & 1 + (m_{11}^{(1)})^3 - (1 + m_{11}^{(1)})^3 = 0, \\ \Rightarrow \; & m_{11}^{(1)}(m_{11}^{(1)} + 1) = 0, \\ \Rightarrow \; & m_{11}^{(1)} = 0 \quad \text{or} \quad m_{11}^{(1)} = -1. \end{aligned}$$

Since $m_{11}^{(2)} = -(1 + m_{11}^{(1)})$, the solutions are

$$\begin{cases} m_{11}^{(1)} = 0 \\ m_{11}^{(2)} = -1 \end{cases} \quad \text{or} \quad \begin{cases} m_{11}^{(1)} = -1 \\ m_{11}^{(2)} = 0 \end{cases}.$$

Inserting either of the two above solutions into $(m_{11}^{(1)})^2 m_{12}^{(1)} + (m_{11}^{(2)})^2 m_{12}^{(2)} = 0$ (eq. 4.41), and combining this with eq. (4.39), we have $m_{12}^{(1)} = m_{12}^{(2)} = 0$. Similarly, from $(m_{11}^{(1)})^2 m_{15}^{(1)} + (m_{11}^{(2)})^2 m_{15}^{(2)} = 0$, we obtain $m_{15}^{(1)} = m_{15}^{(2)} = 0$.

In summary, the solutions cancelling the terms with $C_{1,0,0,0,2}$ and $C_{3,0,0,0,0}$ are

$$\begin{cases} m_{11}^{(1)} = 0 \\ m_{11}^{(2)} = -1 \\ m_{12}^{(1)} = m_{12}^{(2)} = 0 \\ m_{15}^{(1)} = m_{15}^{(2)} = 0 \end{cases} \quad \text{or} \quad \begin{cases} m_{11}^{(1)} = -1 \\ m_{11}^{(2)} = 0 \\ m_{12}^{(1)} = m_{12}^{(2)} = 0 \\ m_{15}^{(1)} = m_{15}^{(2)} = 0 \end{cases}.$$

Similarly, the solutions which cancel the terms with $C_{0,1,0,0,2}$ and $C_{0,3,0,0,0}$ are

$$\begin{cases} m_{22}^{(1)} = 0 \\ m_{22}^{(2)} = -1 \\ m_{21}^{(1)} = m_{21}^{(2)} = 0 \\ m_{25}^{(1)} = m_{25}^{(2)} = 0 \end{cases} \quad \text{or} \quad \begin{cases} m_{22}^{(1)} = -1 \\ m_{22}^{(2)} = 0 \\ m_{21}^{(1)} = m_{21}^{(2)} = 0 \\ m_{25}^{(1)} = m_{25}^{(2)} = 0 \end{cases}.$$

For all the combinations, there is at least one matrix whose determinant equals zero. So there is no solution that cancels $A_3(H(\vec{I})) + A_3(H(M^{(1)}\vec{I})) + A_3(H(M^{(2)}\vec{I}))$ and $B_3(H(\vec{I})) + B_3(H(M^{(1)}\vec{I})) + B_3(H(M^{(2)}\vec{I}))$ simultaneously, which concludes the proof.

4.2.3 Four-Cell Systems

In this section we will show that certain four-cell systems do cancel A, B and C at the same time.

Theorem 4.3 *Given a four-cell system, the terms* $\sum_{i=0}^{3} A_n(H(M^{(i)}\vec{I}))$, $\sum_{i=0}^{3} B_n(H(M^{(i)}\vec{I}))$, *and* $\sum_{i=0}^{3} C_n(H(M^{(i)}\vec{I}))$ $(M^{(0)} = \hat{I})$ *are cancelled 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:

(1) The sufficiency is obvious from eqs. (4.32)-(4.35).

(2) The necessity:

Since $\sum_{i=0}^{3} A_3(H(M^{(i)}\vec{I}))$, $\sum_{i=0}^{3} B_3(H(M^{(i)}\vec{I}))$, and $\sum_{i=0}^{3} C_3(H(M^{(i)}\vec{I}))$ can only be cancelled by symmetry and the first-order arrangements, the necessary conditions for the vanishing of A_3, B_3, and C_3, and hence for the vanishing of A_n, B_n, and C_n, are also from symmetry and the linear map only. Therefore, A_3, B_3, and C_3 are selected to determine the necessary conditions.

To prove the claim, the groups of solutions for the smallest decoupled sets of equations are found first; the connections between two of them are found from the equations containing the variables of the two groups, which form a second-level group. and so on. Eventually, a necessary solution for all the equations is found, from which the conclusion is drawn.

First observe that since all four cells contribute as a sum, cancelling the terms associated with coefficients $C_{1,0,0,0,2}$, $C_{0,1,0,0,2}$, $C_{3,0,0,0,0}$, and $C_{0,3,0,0,0}$ requires that

$$x\delta^2 + \sum_{i=1}^{3}(m_{11}^{(i)}x + m_{12}^{(i)}a + m_{15}^{(i)}\delta)\delta^2 = 0, \tag{4.42}$$

Th

us

(4

we

wh

all

$$a\delta^2 + \sum_{i=1}^{3}(m_{21}^{(i)}x + m_{22}^{(i)}a + m_{25}^{(i)}\delta)\delta^2 = 0, \tag{4.43}$$

$$x^3 + \sum_{i=1}^{3}(m_{11}^{(i)}x + m_{12}^{(i)}a + m_{15}^{(i)}\delta)^3 = 0, \tag{4.44}$$

$$a^3 + \sum_{i=1}^{3}(m_{21}^{(i)}x + m_{22}^{(i)}a + m_{25}^{(i)}\delta)^3 = 0. \tag{4.45}$$

The fact that eqs. (4.42) and (4.44) are decoupled from eqs. (4.43) and (4.45) allows us to solve the two groups separately. Let us first concentrate on eqs. (4.42) and (4.44). Since the coefficient of each term in the equations has to vanish separately, we obtain the necessary equations for the coefficients, which are

$$1 + m_{11}^{(1)} + m_{11}^{(2)} + m_{11}^{(3)} = 0, \tag{1}$$

$$m_{12}^{(1)} + m_{12}^{(2)} + m_{12}^{(3)} = 0, \tag{2}$$

$$m_{15}^{(1)} + m_{15}^{(2)} + m_{15}^{(3)} = 0, \tag{3}$$

$$1 + (m_{11}^{(1)})^3 + (m_{11}^{(2)})^3 + (m_{11}^{(3)})^3 = 0, \tag{4}$$

$$(m_{11}^{(1)})^2 m_{12}^{(1)} + (m_{11}^{(2)})^2 m_{12}^{(2)} + (m_{11}^{(3)})^2 m_{12}^{(3)} = 0, \tag{5}$$

$$m_{11}^{(1)}(m_{12}^{(1)})^2 + m_{11}^{(2)}(m_{12}^{(2)})^2 + m_{11}^{(3)}(m_{12}^{(3)})^2 = 0, \tag{6}$$

$$(m_{12}^{(1)})^3 + (m_{12}^{(2)})^3 + (m_{12}^{(3)})^3 = 0, \tag{7}$$

$$(m_{11}^{(1)})^2 m_{15}^{(1)} + (m_{11}^{(2)})^2 m_{15}^{(2)} + (m_{11}^{(3)})^2 m_{15}^{(3)} = 0, \tag{8}$$

$$m_{11}^{(1)} m_{12}^{(1)} m_{15}^{(1)} + m_{11}^{(2)} m_{12}^{(2)} m_{15}^{(2)} + m_{11}^{(3)} m_{12}^{(3)} m_{15}^{(3)} = 0, \tag{9}$$

$$(m_{12}^{(1)})^2 m_{15}^{(1)} + (m_{12}^{(2)})^2 m_{15}^{(2)} + (m_{12}^{(3)})^2 m_{15}^{(3)} = 0, \tag{10}$$

$$m_{11}^{(1)}(m_{15}^{(1)})^2 + m_{11}^{(2)}(m_{15}^{(2)})^2 + m_{11}^{(3)}(m_{15}^{(3)})^2 = 0, \tag{11}$$

$$m_{12}^{(1)}(m_{15}^{(1)})^2 + m_{12}^{(2)}(m_{15}^{(2)})^2 + m_{12}^{(3)}(m_{15}^{(3)})^2 = 0, \tag{12}$$

$$(m_{15}^{(1)})^3 + (m_{15}^{(2)})^3 + (m_{15}^{(3)})^3 = 0, \tag{13}$$

where eqs. (1)-(3) come from eq. (4.42), and eqs. (4)-(13) from eq. (4.44). Note that all these equations are invariant under a permutation of the upper indices, as they should be.

Since eqs. (1) and (4) are decoupled from the other ones and only contain m_{11}, let us first solve them. Equation (1) can be used to solve for $m_{11}^{(3)}$ to obtain

$$m_{11}^{(3)} = -(1 + m_{11}^{(1)} + m_{11}^{(2)}). \qquad (*)$$

Inserting this into eq. (4), we have $1 + (m_{11}^{(1)})^3 + (m_{11}^{(2)})^3 - (1 + m_{11}^{(1)} + m_{11}^{(2)})^3 = 0$, which can be reduced to

$$(m_{11}^{(1)} + m_{11}^{(2)})(m_{11}^{(1)} + 1)(m_{11}^{(2)} + 1) = 0.$$

Using eq. (*), we obtain three solutions for m_{11}, which are

$$\begin{cases} m_{11}^{(2)} = -m_{11}^{(3)} \\ m_{11}^{(1)} = -1 \end{cases}, \quad \begin{cases} m_{11}^{(1)} = -m_{11}^{(3)} \\ m_{11}^{(2)} = -1 \end{cases}, \quad \begin{cases} m_{11}^{(1)} = -m_{11}^{(2)} \\ m_{11}^{(3)} = -1 \end{cases}, \tag{4.46}$$

As is to be expected from the permutational symmetry, the three solutions in eq. (4.46) are permutations of each other. Without the loss of generality, we select one of them, which induces one particular choice of the permutation, and all the other solutions can be obtained at the end. Therefore, the representative solution of eqs. (1) and (4) is

$$\begin{cases} m_{11}^{(1)} = -m_{11}^{(2)} \\ m_{11}^{(3)} = -1 \end{cases}. \tag{4.47}$$

From now on, solutions of all other variable are expressed with respect to this representative solution.

Now let us turn to eq. (3). from which $m_{15}^{(3)}$ can be solved. We have

$$m_{15}^{(3)} = -(m_{15}^{(1)} + m_{15}^{(2)}). \qquad (*')$$

Inserting this into eq. (13), we have $(m_{15}^{(1)})^3 + (m_{15}^{(2)})^3 - (m_{15}^{(1)} + m_{15}^{(2)})^3 = 0$, which can be reduced to

$$m_{15}^{(1)} m_{15}^{(2)} (m_{15}^{(1)} + m_{15}^{(2)}) = 0.$$

From eq. $(*')$, the solutions of m_{15} are

$$\begin{cases} m_{15}^{(2)} = -m_{15}^{(3)} \\ m_{15}^{(1)} = 0 \end{cases}, \quad \begin{cases} m_{15}^{(1)} = -m_{15}^{(3)} \\ m_{15}^{(2)} = 0 \end{cases}, \quad \text{or} \quad \begin{cases} m_{15}^{(1)} = -m_{15}^{(2)} \\ m_{15}^{(3)} = 0 \end{cases}. \tag{4.48}$$

Note that the other three solutions are exactly the same as those shown in eq. (4.48), respectively.

Combining eqs. (4.47) and (4.48), we obtain the the solutions for m_{11} and m_{15} satisfying eqs. (1), (3), (4), and (13), which are

$$\begin{cases} m_{15}^{(2)} = -m_{15}^{(3)}, m_{15}^{(1)} = 0 \\ m_{11}^{(1)} = -m_{11}^{(2)}, m_{11}^{(3)} = -1 \end{cases}, \quad \begin{cases} m_{15}^{(1)} = -m_{15}^{(3)}, m_{15}^{(2)} = 0 \\ m_{11}^{(1)} = -m_{11}^{(2)}, m_{11}^{(3)} = -1 \end{cases},$$

$$\text{or} \quad \begin{cases} m_{15}^{(1)} = -m_{15}^{(2)}, m_{15}^{(3)} = 0 \\ m_{11}^{(1)} = -m_{11}^{(2)}, m_{11}^{(3)} = -1 \end{cases}. \tag{4.49}$$

Next, let us now decide how these solutions satisfy eqs. (8) and (11). Together with the four equations above, they form a second-level group which contains purely m_{11} and m_{15}.

Inserting the first solution from eq. (4.49) into (8) and (11), we have

$$\begin{cases} (m_{11}^{(2)})^2 m_{15}^{(2)} + (-1)^2(-m_{15}^{(2)}) & = 0 \\ (m_{11}^{(2)})(m_{15}^{(2)})^2 + (-1)(-m_{15}^{(2)})^2 & = 0 \end{cases},$$

which has the solutions $m_{15}^{(2)} = 0$ or $m_{11}^{(2)} = 1$. In the case of $m_{15}^{(2)} = 0$, we have $m_{15}^{(3)} = -m_{15}^{(2)} = 0$, which gives

$$\begin{cases} m_{15}^{(1)} = m_{15}^{(2)} = m_{15}^{(3)} = 0 \\ m_{11}^{(1)} = -m_{11}^{(2)}, m_{11}^{(3)} = -1 \end{cases}.$$

In the case of $m_{11}^{(2)} = 1$, we have $m_{11}^{(3)} = -m_{11}^{(2)} - 1$, which gives

$$\begin{cases} m_{15}^{(2)} = -m_{15}^{(3)}, m_{15}^{(1)} = 0 \\ m_{11}^{(1)} = -1, m_{11}^{(2)} = 1, m_{11}^{(3)} = -1 \end{cases}. \tag{4.50}$$

Similarly, inserting the second solution from eq. (4.49) into (8) and (11) gives the solutions

$$\begin{cases} m_{15}^{(1)} = m_{15}^{(2)} = m_{15}^{(3)} = 0, \\ m_{11}^{(1)} = -m_{11}^{(2)}, m_{11}^{(3)} = -1 \end{cases} \quad \text{or} \quad \begin{cases} m_{15}^{(1)} = -m_{15}^{(3)}, m_{15}^{(2)} = 0 \\ m_{11}^{(1)} = 1, m_{11}^{(2)} = -1, m_{11}^{(3)} = -1 \end{cases}.$$

Since the last solution from eq. (4.49) happens to satisfy eqs. (8) and (11), and the solutions from eqs. (4.50) and (4.51) are special cases of eq. (4.49)'s solution, those solutions satisfying the equations containing only $m_{11}^{(1)}$ and $m_{15}^{(1)}$ are

$$\begin{cases} m_{15}^{(1)} = 0, m_{15}^{(2)} = 0, m_{15}^{(3)} = 0 \\ m_{11}^{(1)} = -m_{11}^{(2)}, m_{11}^{(3)} = -1 \end{cases}, \quad \text{or} \quad \begin{cases} m_{15}^{(1)} = -m_{15}^{(2)}, m_{15}^{(3)} = 0 \\ m_{11}^{(1)} = -m_{11}^{(2)}, m_{11}^{(3)} = -1 \end{cases}. \tag{4.51}$$

Note that the first solution is actually a special case of the second one. But it is kept as a separate solution for the convenience of the proof.

A little inspection shows that the equations containing only m_{12}, which are eqs. (2) and (7), are exactly the same as eqs. (3) and (13) for m_{15}. Therefore, m_{12} has the same solution as m_{15}, which means that the solutions satisfying those equations containing only $m_{11}^{(1)}$ and $m_{12}^{(1)}$ are exactly the same as those of the equations for m_{11} and m_{15}, which are

$$\begin{cases} m_{12}^{(1)} = m_{12}^{(2)} = m_{12}^{(3)} = 0 \\ m_{11}^{(1)} = -m_{11}^{(2)}, m_{11}^{(3)} = -1 \end{cases} \quad \text{or} \quad \begin{cases} m_{12}^{(1)} = -m_{12}^{(2)}, m_{12}^{(3)} = 0 \\ m_{11}^{(1)} = -m_{11}^{(2)}, m_{11}^{(3)} = -1 \end{cases}. \tag{4.52}$$

After checking with eq. (9), (10) and (12), we obtain the solutions for the whole system of equations, which are

$$(1) \begin{cases} m_{12}^{(1)} = m_{12}^{(2)} = m_{12}^{(3)} = 0 \\ m_{15}^{(1)} = m_{15}^{(2)} = m_{15}^{(3)} = 0 \\ m_{11}^{(1)} = -m_{11}^{(2)}, m_{11}^{(3)} = -1 \end{cases}, \quad (2) \begin{cases} m_{12}^{(1)} = m_{12}^{(2)} = m_{12}^{(3)} = 0 \\ m_{15}^{(1)} = -m_{15}^{(2)}, m_{15}^{(3)} = 0 \\ m_{11}^{(1)} = -m_{11}^{(2)}, m_{11}^{(3)} = -1 \end{cases},$$

$$(3) \begin{cases} m_{12}^{(1)} = -m_{12}^{(2)}, m_{12}^{(3)} = 0 \\ m_{15}^{(1)} = m_{15}^{(2)} = m_{15}^{(3)} = 0 \\ m_{11}^{(1)} = -m_{11}^{(2)}, m_{11}^{(3)} = -1 \end{cases}, \quad (4) \begin{cases} m_{12}^{(1)} = -m_{12}^{(2)}, m_{12}^{(3)} = 0 \\ m_{15}^{(1)} = -m_{15}^{(2)}, m_{15}^{(3)} = 0 \\ m_{11}^{(1)} = -m_{11}^{(2)}, m_{11}^{(3)} = -1 \end{cases}.$$

Again, the equations of m_{21}, m_{22} and m_{25}, given by eqs. (4.43) and (4.45), are exactly the same as those of m_{11}, m_{12} and m_{15}. Since the solution for m_{11} has been fixed, all three solutions have to be taken account, which are

$$(1') \begin{cases} m_{21}^{(1)} = m_{21}^{(2)} = m_{21}^{(3)} = 0 \\ m_{25}^{(1)} = m_{25}^{(2)} = m_{25}^{(3)} = 0 \\ m_{22}^{(2)} = -m_{22}^{(3)}, m_{22}^{(1)} = -1 \end{cases}, \quad \begin{cases} m_{21}^{(1)} = m_{21}^{(2)} = m_{21}^{(3)} = 0 \\ m_{25}^{(1)} = m_{25}^{(2)} = m_{25}^{(3)} = 0 \\ m_{22}^{(1)} = -m_{22}^{(3)}, m_{22}^{(2)} = -1 \end{cases},$$

$$\text{or}\quad \begin{cases} m_{21}^{(1)} = m_{21}^{(2)} = m_{21}^{(3)} = 0 \\ m_{25}^{(1)} = m_{25}^{(2)} = m_{25}^{(3)} = 0 \\ m_{22}^{(1)} = -m_{22}^{(2)}, m_{22}^{(3)} = -1 \end{cases} ;$$

$$(2')\quad \begin{cases} m_{21}^{(1)} = m_{21}^{(2)} = m_{21}^{(3)} = 0 \\ m_{25}^{(2)} = -m_{25}^{(3)}, m_{25}^{(1)} = 0 \\ m_{22}^{(2)} = -m_{22}^{(3)}, m_{22}^{(1)} = -1 \end{cases} , \begin{cases} m_{21}^{(1)} = m_{21}^{(2)} = m_{21}^{(3)} = 0 \\ m_{25}^{(1)} = -m_{25}^{(3)}, m_{25}^{(2)} = 0 \\ m_{22}^{(1)} = -m_{22}^{(3)}, m_{22}^{(2)} = -1 \end{cases} ,$$

$$\text{or}\quad \begin{cases} m_{21}^{(1)} = m_{21}^{(2)} = m_{21}^{(3)} = 0 \\ m_{25}^{(1)} = -m_{25}^{(2)}, m_{25}^{(3)} = 0 \\ m_{22}^{(1)} = -m_{22}^{(2)}, m_{22}^{(3)} = -1 \end{cases} ;$$

$$(3')\quad \begin{cases} m_{21}^{(2)} = -m_{21}^{(3)}, m_{21}^{(1)} = 0 \\ m_{25}^{(1)} = m_{25}^{(2)} = m_{25}^{(3)} = 0 \\ m_{22}^{(2)} = -m_{22}^{(3)}, m_{22}^{(1)} = -1 \end{cases} , \begin{cases} m_{21}^{(1)} = -m_{21}^{(3)}, m_{21}^{(2)} = 0 \\ m_{25}^{(1)} = m_{25}^{(2)} = m_{25}^{(3)} = 0 \\ m_{22}^{(1)} = -m_{22}^{(3)}, m_{22}^{(2)} = -1 \end{cases} ,$$

$$\text{or}\quad \begin{cases} m_{21}^{(1)} = -m_{21}^{(2)}, m_{21}^{(3)} = 0 \\ m_{25}^{(1)} = m_{25}^{(2)} = m_{25}^{(3)} = 0 \\ m_{22}^{(1)} = -m_{22}^{(2)}, m_{22}^{(3)} = -1 \end{cases} ;$$

$$(4')\quad \begin{cases} m_{21}^{(2)} = -m_{21}^{(3)}, m_{21}^{(1)} = 0 \\ m_{25}^{(2)} = -m_{25}^{(3)}, m_{25}^{(1)} = 0 \\ m_{22}^{(2)} = -m_{22}^{(3)}, m_{22}^{(1)} = -1 \end{cases} , \begin{cases} m_{21}^{(1)} = -m_{21}^{(3)}, m_{21}^{(2)} = 0 \\ m_{25}^{(1)} = -m_{25}^{(3)}, m_{25}^{(2)} = 0 \\ m_{22}^{(1)} = -m_{22}^{(3)}, m_{22}^{(2)} = -1 \end{cases} ,$$

$$\text{or}\quad \begin{cases} m_{21}^{(1)} = -m_{21}^{(2)}, m_{21}^{(3)} = 0 \\ m_{25}^{(1)} = -m_{25}^{(2)}, m_{25}^{(3)} = 0 \\ m_{22}^{(1)} = -m_{22}^{(2)}, m_{22}^{(3)} = -1 \end{cases} .$$

So far there is an abundance of solutions. By using additional conditions connecting matrix elements, their number will be reduced. Let us consider the terms with $C_{1,2,0,0,0}$, $C_{2,1,0,0,0}$, and $C_{1,1,0,0,1}$. Cancelling them requires that

$$xa^2 + \sum_{i=1}^{3}(m_{11}^{(i)}x + m_{12}^{(i)}a + m_{15}^{(i)}\delta)(m_{21}^{(i)}x + m_{22}^{(i)}a + m_{25}^{(i)}\delta)^2 = 0, \tag{4.53}$$

$$x^2a + \sum_{i=1}^{3}(m_{11}^{(i)}x + m_{12}^{(i)}a + m_{15}^{(i)}\delta)^2(m_{21}^{(i)}x + m_{22}^{(i)}a + m_{25}^{(i)}\delta) = 0, \tag{4.54}$$

$$xa\delta + \sum_{i=1}^{3}(m_{11}^{(i)}x + m_{12}^{(i)}a + m_{15}^{(i)}\delta)(m_{21}^{(i)}x + m_{22}^{(i)}a + m_{25}^{(i)}\delta)\delta = 0. \quad (4.55)$$

Next, all possible combinations of the two sets of solutions are considered, and the final solutions are decided. We will start from the simplest case, and then move to the more complicated ones.

Case $(1)-(1')$:

Since $m_{12}^{(i)} = m_{15}^{(i)} = m_{21}^{(i)} = m_{25}^{(i)} = 0$ $(i = 1,2,3)$, the equations (4.53)-(4.55) are simplified to

$$xa^2 + \sum_{i=1}^{3}(m_{11}^{(i)}x)(m_{22}^{(i)}a)^2 = 0,$$

$$x^2a + \sum_{i=1}^{3}(m_{11}^{(i)}x)^2(m_{22}^{(i)}a) = 0,$$

$$xa\delta + \sum_{i=1}^{3}(m_{11}^{(i)}x)(m_{22}^{(i)}a)\delta = 0,$$

where the coefficients are

$$1 + m_{11}^{(1)}(m_{22}^{(1)})^2 + m_{11}^{(2)}(m_{22}^{(2)})^2 + m_{11}^{(3)}(m_{22}^{(3)})^2 = 0, \quad (4.56)$$

$$1 + (m_{11}^{(1)})^2 m_{22}^{(1)} + (m_{11}^{(2)})^2 m_{22}^{(2)} + (m_{11}^{(3)})^2 m_{22}^{(3)} = 0, \quad (4.57)$$

$$1 + m_{11}^{(1)}m_{22}^{(1)} + m_{11}^{(2)}m_{22}^{(2)} + m_{11}^{(3)}m_{22}^{(3)} = 0. \quad (4.58)$$

Now let us look at the first solution from $(1')$, which gives

$$\begin{cases} m_{22}^{(2)} = -m_{22}^{(3)} \\ m_{22}^{(1)} = -1 \end{cases}.$$

Inserting this into eqs. (4.56)-(4.58), we have

$$1 + (-m_{11}^{(2)})(-1)^2 + m_{11}^{(2)}(m_{22}^{(2)})^2 + (-1)(-m_{22}^{(2)})^2 = 0,$$

$$1 + (-m_{11}^{(2)})^2(-1) + (m_{11}^{(2)})^2 m_{22}^{(2)} + (-1)^2(-m_{22}^{(2)}) = 0,$$

$$1 + (-m_{11}^{(2)})(-1) + m_{11}^{(2)}m_{22}^{(2)} + (-1)(-m_{22}^{(2)}) = 0,$$

whi

Af

S

I

which can be transformed to

$$1 - m_{11}^{(2)} + m_{11}^{(2)}(m_{22}^{(2)})^2 - (m_{22}^{(2)})^2 = 0,$$

$$1 - (m_{11}^{(2)})^2 + (m_{11}^{(2)})^2 m_{22}^{(2)} - m_{22}^{(2)} = 0,$$

$$1 + m_{11}^{(2)} + m_{11}^{(2)} m_{22}^{(2)} + m_{22}^{(2)} = 0.$$

After factorization, the equations are

$$(1 - m_{11}^{(2)})(1 - (m_{22}^{(2)})^2) = 0,$$

$$(1 - (m_{11}^{(2)})^2)(1 - m_{22}^{(2)}) = 0,$$

$$(1 + m_{11}^{(2)})(1 + m_{22}^{(2)}) = 0.$$

Straightforward arithmetic reveals that the solutions are

$$\begin{cases} m_{11}^{(2)} &= -1 \\ m_{22}^{(2)} &= \pm 1 \end{cases}, \quad \text{or} \quad \begin{cases} m_{11}^{(2)} &= \pm 1 \\ m_{22}^{(2)} &= -1 \end{cases}.$$

In conclusion, the final results for the first solution are

$$\begin{cases} m_{11}^{(1)} = 1, m_{11}^{(2)} = -1, m_{11}^{(3)} = -1 \\ m_{22}^{(1)} = -1, m_{22}^{(2)} = \pm 1, m_{22}^{(3)} = \mp 1 \end{cases}, \text{ or } \begin{cases} m_{11}^{(1)} = \mp 1, m_{11}^{(2)} = \pm 1, m_{11}^{(3)} = -1 \\ m_{22}^{(1)} = -1, m_{22}^{(2)} = -1, m_{22}^{(3)} = 1 \end{cases}.$$

Written in the form of matrices, the solutions of this case are

$$M_1^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, M_1^{(2)} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \pm 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, M_1^{(3)} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \mp 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

or

$$M_1^{(1)} = \begin{pmatrix} \mp 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, M_1^{(2)} = \begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, M_1^{(3)} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

which correspond to the cases that $(M_1^{(1)}, M_1^{(2)}, M_1^{(3)}) = (R_1, C_1, S_1)$, (R_1, S_1, C_1), (S_1, R_1, C_1), or (R_1, S_1, C_1). Therefore, the first solution of $(1')$ agrees with the assertion of the theorem. Note that $M_1^{(3)}$ can only be S or C, because $m_{11}^{(3)}$ is fixed to be -1. This is why Case $(1) - (1')$ gives four, rather than six, solutions. It can be shown similarly that the second solution of $(1')$ draws the same conclusion.

In the case of the third solution, we have

$$\begin{cases} m_{22}^{(1)} = -m_{22}^{(2)} \\ m_{22}^{(3)} = -1 \end{cases}.$$

Inserting it into eqs. (4.56)-(4.58) shows that eqs. (4.56) and (4.57) are automatically satisfied, and only eq. (4.58) gives a nontrivial solution, which has the form 2 + $2m_{11}^{(1)}m_{22}^{(1)} = 0$. The solution is $m_{11}^{(1)}m_{22}^{(1)} = -1$. Together with $m_{11}^{(3)} = -1$, we have

$$\begin{cases} m_{11}^{(1)}m_{22}^{(1)} = -1 \\ m_{11}^{(3)} = -1 \end{cases} \tag{4.59}$$

It can be shown that

$$m_{11}^{(1)} = -m_{22}^{(1)} \quad \text{(see Appendix B).} \tag{4.60}$$

Thus, the solutions are

$$\begin{cases} m_{11}^{(1)} = \pm 1, m_{11}^{(2)} = \mp 1, m_{11}^{(3)} = -1 \\ m_{22}^{(2)} = \pm 1, m_{22}^{(1)} = \mp 1, m_{22}^{(3)} = -1 \end{cases},$$

or in the matrix form

$$M_1^{(1)} = \begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & \mp 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, M_1^{(2)} = \begin{pmatrix} \mp 1 & 0 & 0 \\ 0 & \pm 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, M_1^{(3)} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0, \\ 0 & 0 & 1 \end{pmatrix}.$$

which correspond to the case that $(M_1^{(1)}, M_1^{(2)}, M_1^{(3)}) = (R_1, C_1, S_1)$ or (C_1, R_1, S_1). Therefore, Case (1) – (1′) agrees with the assertion of the theorem.

Case (1) – (2′):

Here $m_{12}^{(i)} = m_{15}^{(i)} = m_{21}^{(i)} = 0$ $(i = 1, 2, 3)$; thus the equations (4.53)-(4.55) become

$$xa^2 + \sum_{i=1}^{3}(m_{11}^{(i)}x)(m_{22}^{(i)}a + m_{25}^{(i)}\delta)^2 = 0, \tag{4.61}$$

$$x^2a + \sum_{i=1}^{3}(m_{11}^{(i)}x)^2(m_{22}^{(i)}a + m_{25}^{(i)}\delta) = 0, \tag{4.62}$$

$$xa\delta + \sum_{i=1}^{3}(m_{11}^{(i)}x)(m_{22}^{(i)}a + m_{25}^{(i)}\delta)\delta = 0, \tag{4.63}$$

which are equivalent to

$$1+m_{11}^{(1)}(m_{22}^{(1)})^2+m_{11}^{(2)}(m_{22}^{(2)})^2+m_{11}^{(3)}(m_{22}^{(3)})^2=0,$$

$$1+(m_{11}^{(1)})^2m_{22}^{(1)}+(m_{11}^{(2)})^2m_{22}^{(2)}+(m_{11}^{(3)})^2m_{22}^{(3)}=0,$$

$$1+m_{11}^{(1)}m_{22}^{(1)}+m_{11}^{(2)}m_{22}^{(2)}+m_{11}^{(3)}m_{22}^{(3)}=0,$$

$$m_{11}^{(1)}(m_{25}^{(1)})^2+m_{11}^{(2)}(m_{25}^{(2)})^2+m_{11}^{(3)}(m_{25}^{(3)})^2=0,$$

$$(m_{11}^{(1)})^2m_{25}^{(1)}+(m_{11}^{(2)})^2m_{25}^{(2)}+(m_{11}^{(3)})^2m_{25}^{(3)}=0,$$

$$m_{11}^{(1)}m_{25}^{(1)}+m_{11}^{(2)}m_{25}^{(2)}+m_{11}^{(3)}m_{25}^{(3)}=0,$$

$$m_{11}^{(1)}m_{22}^{(1)}m_{25}^{(1)}+m_{11}^{(2)}m_{22}^{(2)}m_{25}^{(2)}+m_{11}^{(3)}m_{22}^{(3)}m_{25}^{(3)}=0.$$

From the first three equations, we obtain the same solution as in Case $(1)-(1')$. First let us look at the case of $(M_1^{(1)},M_1^{(2)},M_1^{(3)})=(R_1,S_1,C_1)$. Inserting this into the last three equations, we have

$$m_{25}^{(1)}+m_{25}^{(2)}+(m^{(3)})_{25} = 0, \qquad (1'')$$

$$m_{25}^{(1)}-m_{25}^{(2)}-(m^{(3)})_{25} = 0, \qquad (2'')$$

$$-m_{25}^{(1)}+m_{25}^{(2)}-(m^{(3)})_{25} = 0. \qquad (3'')$$

From $(2'')$ and $(3'')$, we have $m_{25}^{(3)}=0$, and from $(1'')$ and $(3'')$, we have $m_{25}^{(1)}=0$, Therefore, the solution is $m_{25}^{(1)}=m_{25}^{(2)}=m_{25}^{(3)}=0$, which is means that this case is reduced to Case $(1)-(1')$. The other three solutions from Case $(1)-(1')$ give the same results, which is a consequence of the permutational symmetry.

Note that Cases $(1)-(3')$ and $(1)-(4')$ are the same as Case $(1)-(2')$.

Case $(2)-(2')$:

Here, $m_{12}^{(i)}=m_{21}^{(i)}=0$ $(i=1,2,3)$; thus the equations (4.53)-(4.55) become

$$xa^2+\sum_{i=1}^{3}(m_{11}^{(1)}x+m_{15}^{(i)}\delta)(m_{22}^{(i)}a+m_{25}^{(i)}\delta)^2=0,$$

$$x^2 a + \sum_{i=1}^{3} (m_{11}^{(1)} x + m_{15}^{(i)} \delta)^2 (m_{22}^{(i)} a + m_{25}^{(i)} \delta) = 0,$$

$$x a \delta + \sum_{i=1}^{3} (m_{11}^{(1)} x + m_{15}^{(i)} \delta)(m_{22}^{(i)} a + m_{25}^{(i)} \delta)\delta = 0.$$

Cancelling the terms $a^2\delta$, $a\delta^2$, and $a\delta$ requires that

$$m_{15}^{(1)}(m_{22}^{(1)})^2 + m_{15}^{(2)}(m_{22}^{(2)})^2 + m_{15}^{(3)}(m_{22}^{(3)})^2 = 0,$$

$$(m_{15}^{(1)})^2 m_{22}^{(1)} + (m_{15}^{(2)})^2 m_{22}^{(2)} + (m_{15}^{(3)})^2 m_{22}^{(3)} = 0,$$

$$m_{15}^{(1)} m_{22}^{(1)} + m_{15}^{(2)} m_{22}^{(2)} + m_{15}^{(3)} m_{22}^{(3)} = 0.$$

Since the diagonal elements are the same as those obtained in Case (1) – (1′), we obtain $m_{15}^{(1)} = m_{15}^{(2)} = m_{15}^{(3)} = 0$. Therefore this case is simplified to Case (1) – (2′), which gives the same solution as Case (1) – (1′). Note that Case (2) – (3′) and Case (2) – (4′) are the same as Case (2) – (2′).

Case (3) – (3′):

Here, $m_{15}^{(i)} = m_{25}^{(i)} = 0$ $(i = 1, 2, 3)$; thus the equations (4.53)-(4.55) become

$$x a^2 + \sum_{i=1}^{3} (m_{11}^{(1)} x + m_{12}^{(i)} a)(m_{21}^{(i)} x + m_{22}^{(i)} a)^2 = 0, \tag{4.64}$$

$$x^2 a + \sum_{i=1}^{3} (m_{11}^{(1)} x + m_{12}^{(i)} a)^2 (m_{21}^{(i)} x + m_{22}^{(i)} a) = 0, \tag{4.65}$$

$$x a \delta + \sum_{i=1}^{3} (m_{11}^{(1)} x + m_{12}^{(i)} a)(m_{21}^{(i)} x + m_{22}^{(i)} a)\delta = 0. \tag{4.66}$$

Cancelling the terms a^3 and $a^2\delta$ requires that

$$m_{12}^{(1)}(m_{22}^{(1)})^2 + m_{12}^{(2)}(m_{22}^{(2)})^2 + m_{12}^{(3)}(m_{22}^{(3)})^2 = 0,$$

$$(m_{12}^{(1)})^2 m_{22}^{(1)} + (m_{12}^{(2)})^2 m_{22}^{(2)} + (m_{12}^{(3)})^2 m_{22}^{(3)} = 0,$$

$$m_{12}^{(1)} m_{22}^{(1)} + m_{12}^{(2)} m_{22}^{(2)} + m_{12}^{(3)} m_{22}^{(3)} = 0.$$

Inserting this in $m_{12}^{(1)} = -m_{12}^{(2)}$ and $m_{12}^{(3)} = 0$, we have

$$m_{12}^{(1)}((m_{22}^{(1)})^2 - (m_{22}^{(2)})^2) = 0,$$

$$(m_{12}^{(1)})^2(m_{22}^{(1)} + m_{22}^{(2)}) = 0,$$

$$m_{12}^{(1)}(m_{22}^{(1)} - m_{22}^{(2)}) = 0,$$

which have the solutions

$$m_{12}^{(1)} = 0 \quad \text{or} \quad m_{22}^{(1)} = m_{22}^{(2)} = 0.$$

Let us consider these two solutions separately.

(a) $m_{22}^{(1)} = m_{22}^{(2)} = 0$

From solution (3′), we have $m_{22}^{(3)} = -1$. Together with $m_{12}^{(3)} = 0$, the equations (4.64) and (4.66) become

$$xa^2 + (m_{11}^{(1)}x + m_{12}^{(1)}a)(m_{21}^{(1)}x)^2 + (m_{11}^{(2)}x + m_{12}^{(2)}a)(m_{21}^{(2)}x)^2$$
$$+(m_{11}^{(3)}x)(m_{21}^{(3)}x - a)^2 = 0,$$

$$x^2a + (m_{11}^{(1)}x + m_{12}^{(1)}a)^2(m_{21}^{(1)}x) + (m_{11}^{(2)}x + m_{12}^{(2)}a)^2(m_{21}^{(2)}x)$$
$$+(m_{11}^{(3)}x)^2(m_{21}^{(3)}x - a) = 0,$$

$$xa + (m_{11}^{(1)}x + m_{12}^{(1)}a)(m_{21}^{(1)}x) + (m_{11}^{(2)}x + m_{12}^{(2)}a)(m_{21}^{(2)}x)$$
$$+(m_{11}^{(3)}x)(m_{21}^{(3)}x - a) = 0.$$

Cancelling the terms xa^2, x^2a and xa requires that

$$1 + m_{11}^{(3)} = 0, \quad 1 - (m_{11}^{(3)})^2 = 0, \quad 1 - m_{11}^{(3)}x = 0,$$

which have no solution.

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

From solution (3), we have $m_{12}^{(2)} = 0$. Hence, this case is reduced to Case $(1)-(3')$.

Similarly, Cases (3) – (4′) and (4) – (4′) can be simplified to Case (1) – (4′) and Case (2) – (4′) respectively, which concludes the x-motion.

Now let us study the y-b sub-matrices. Since it is proven that the x-a sub-matrices are permutations of R_1, S_1, and C_1, and permutational symmetry holds, one solution out of six can be chosen without the loss of generality. Here we choose $(M_1^{(1)}, M_1^{(2)}, M_1^{(3)}) = (R_1, S_1, C_1)$.

Cancelling the terms with $C_{1,0,2,0,0}$ and $C_{0,1,2,0,0}$ requires that

$$y^2 + (m_{33}^{(1)}y + m_{34}^{(1)}b)^2 - (m_{33}^{(2)}y + m_{34}^{(2)}b)^2 - (m_{33}^{(3)}y + m_{34}^{(3)}b)^2 = 0, \qquad (1)$$

$$y^2 - (m_{33}^{(1)}y + m_{34}^{(1)}b)^2 - (m_{33}^{(2)}y + m_{34}^{(2)}b)^2 + (m_{33}^{(3)}y + m_{34}^{(3)}b)^2 = 0. \qquad (2)$$

From (1) + (2), we obtain $y^2 - (m_{33}^{(2)}y + m_{34}^{(2)}b)^2 = 0$, which has the solutions

$$\left\{ \begin{array}{lcl} m_{33}^{(2)} & = & 1 \\ m_{34}^{(2)} & = & 0 \end{array} \right. \quad \text{or} \quad \left\{ \begin{array}{lcl} m_{33}^{(2)} & = & -1 \\ m_{34}^{(2)} & = & 0 \end{array} \right. .$$

From (1) – (2), we obtain $(m_{33}^{(1)}y + m_{34}^{(1)}b)^2 - (m_{33}^{(3)}y + m_{34}^{(3)}b)^2 = 0$, which has the solutions

$$\left\{ \begin{array}{l} m_{33}^{(1)} = m_{33}^{(3)} \\ m_{34}^{(1)} = m_{34}^{(3)} \end{array} \right. \quad \text{or} \quad \left\{ \begin{array}{l} m_{33}^{(1)} = -m_{33}^{(3)} \\ m_{34}^{(1)} = -m_{34}^{(3)} \end{array} \right. .$$

Altogether, we have the solutions

$$\left\{ \begin{array}{l} m_{33}^{(1)} = \pm m_{33}^{(3)}, m_{33}^{(2)} = 1 \\ m_{34}^{(1)} = \pm m_{34}^{(3)}, m_{34}^{(2)} = 0 \end{array} \right. \quad \text{or} \quad \left\{ \begin{array}{l} m_{33}^{(1)} = \pm m_{33}^{(3)}, m_{33}^{(2)} = -1 \\ m_{34}^{(1)} = \pm m_{34}^{(3)}, m_{34}^{(2)} = 0 \end{array} \right. . \qquad (4.67)$$

Cancelling the terms with $C_{1,0,0,2,0}$ and $C_{0,1,0,2,0}$ requires that

$$b^2 + (m_{43}^{(1)}y + m_{44}^{(1)}b)^2 - (m_{43}^{(2)}y + m_{44}^{(2)}b)^2 - (m_{43}^{(3)}y + m_{44}^{(3)}b)^2 = 0,$$

$$b^2 - (m_{43}^{(1)}y + m_{44}^{(1)}b)^2 - (m_{43}^{(2)}y + m_{44}^{(2)}b)^2 + (m_{43}^{(3)}y + m_{44}^{(3)}b)^2 = 0,$$

which are exactly the same as those for m_{33} and m_{34}. Therefore, the solutions of m_{43} and m_{44} are

$$\left\{ \begin{array}{l} m_{43}^{(1)} = \pm m_{43}^{(3)}, m_{43}^{(2)} = 0 \\ m_{44}^{(1)} = \pm m_{44}^{(3)}, m_{44}^{(2)} = 1 \end{array} \right. \quad \text{or} \quad \left\{ \begin{array}{l} m_{43}^{(1)} = \pm m_{43}^{(3)}, m_{43}^{(2)} = 0 \\ m_{44}^{(1)} = \pm m_{44}^{(3)}, m_{44}^{(2)} = -1 \end{array} \right. . \qquad (4.68)$$

Cancelling the terms with $C_{1,0,1,1,0}$ and $C_{0,1,1,1,0}$ requires that

$$yb + (m_{33}^{(1)}y + m_{34}^{(1)}b)(m_{43}^{(1)}y + m_{44}^{(1)}b)$$
$$-(m_{33}^{(2)}y)(m_{44}^{(2)}b) - (m_{33}^{(3)}y + m_{34}^{(3)}b)(m_{43}^{(3)}y + m_{44}^{(3)}b) = 0, \qquad (5)$$
$$yb - (m_{33}^{(1)}y + m_{34}^{(1)}b)(m_{43}^{(1)}y + m_{44}^{(1)}b)$$
$$-(m_{33}^{(2)}y)(m_{44}^{(2)}b) + (m_{33}^{(3)}y + m_{34}^{(3)}b)(m_{43}^{(3)}y + m_{44}^{(3)}b) = 0. \qquad (6)$$

From (5) + (6), we obtain $yb - (m_{33}^{(2)}y)(m_{44}^{(2)}b) = 0$, which has the solutions

$$m_{33}^{(2)} = m_{44}^{(2)} = 1 \quad \text{or} \quad m_{33}^{(2)} = m_{44}^{(2)} = -1.$$

From (5) − (6), we have

$$(m_{33}^{(1)}y + m_{34}^{(1)}b)(m_{43}^{(1)}y + m_{44}^{(1)}b) = (m_{33}^{(3)}y + m_{34}^{(3)}b)(m_{43}^{(3)}y + m_{44}^{(3)}b),$$

which entails that

$$m_{33}^{(1)}m_{43}^{(1)} = m_{33}^{(3)}m_{43}^{(3)},$$
$$m_{33}^{(1)}m_{44}^{(1)} + m_{34}^{(1)}m_{43}^{(1)} = m_{33}^{(3)}m_{44}^{(3)} + m_{34}^{(3)}m_{43}^{(3)},$$
$$m_{34}^{(1)}m_{44}^{(1)} = m_{34}^{(3)}m_{44}^{(3)}.$$

Inserting eq. (4.67) into the three equations above, we obtain $m_{43}^{(1)} = \pm m_{43}^{(3)}$ and $m_{44}^{(1)} = \pm m_{44}^{(3)}$, which shows that the total solutions are

$$\begin{cases} m_{33}^{(1)} = \pm m_{33}^{(3)}, m_{33}^{(2)} = 1 \\ m_{34}^{(1)} = \pm m_{34}^{(3)}, m_{34}^{(2)} = 0 \\ m_{43}^{(1)} = \pm m_{43}^{(3)}, m_{43}^{(2)} = 0 \\ m_{44}^{(1)} = \pm m_{44}^{(3)}, m_{44}^{(2)} = 1 \end{cases} \quad \text{or} \quad \begin{cases} m_{33}^{(1)} = \pm m_{33}^{(3)}, m_{33}^{(2)} = -1 \\ m_{34}^{(1)} = \pm m_{34}^{(3)}, m_{34}^{(2)} = 0 \\ m_{43}^{(1)} = \pm m_{43}^{(3)}, m_{43}^{(2)} = 0 \\ m_{44}^{(1)} = \pm m_{44}^{(3)}, m_{44}^{(2)} = -1 \end{cases}.$$

Cancelling the terms with $C_{0,0,1,1,1}$ requires that

$$yb + (m_{33}^{(1)}y + m_{34}^{(1)}b)(m_{43}^{(1)}y + m_{44}^{(1)}b)$$
$$+(m_{33}^{(2)}y)(m_{44}^{(2)}b) + (m_{33}^{(3)}y + m_{34}^{(3)}b)(m_{43}^{(3)}y + m_{44}^{(3)}b) = 0 \qquad (7)$$

From

whi

Du

wh

Ap

M

$\sum$

tio

From (7) − (6), we obtain

$$(m_{33}^{(2)}y)(m_{44}^{(2)}b) + (m_{33}^{(3)}y + m_{34}^{(3)}b)(m_{43}^{(3)}y + m_{44}^{(3)}b) = 0,$$

which entails that

$$m_{33}^{(3)}m_{43}^{(3)} = 0,$$

$$1 + m_{33}^{(3)}m_{44}^{(3)} + m_{34}^{(3)}m_{43}^{(3)} = 0,$$

$$m_{34}^{(3)}m_{44}^{(3)} = 0.$$

Due to the fact that $\det(M^{(3)}) \neq 0$, the solutions are

$$a) \begin{cases} m_{33}^{(3)} = 0 \\ m_{44}^{(3)} = 0 \\ m_{34}^{(3)}m_{43}^{(3)} = -1 \end{cases} \quad \text{or} \quad b) \begin{cases} m_{34}^{(3)} = 0 \\ m_{43}^{(3)} = 0 \\ m_{33}^{(3)}m_{44}^{(3)} = -1 \end{cases} .$$

Case a): $m_{34}^{(3)}m_{43}^{(3)} = -1$ implies that $\det(M^{(3)}) = -\det(M_1^{(3)}) \cdot m_{34}^{(3)}m_{43}^{(3)} = -1$, which is impossible.

Case b): Similar to the x-a-δ sub-matrix, it can be shown that $m_{33}^{(3)} = -m_{44}^{(3)}$ (see Appendix B). Taking into account midplane symmetry, this means that $M_2^{(1)} = R_2$, $M_2^{(2)} = S_2$, and $M_2^{(3)} = C_2$.

In summary, $M^{(1)} = R$, $M^{(2)} = S$, and $M^{(3)} = C$, which concludes the proof.

Altogether, we have proven that there is only one way to cancel $\sum A_3(H(M^{(i)}\vec{I}))$, $\sum B_3(H(M^{(i)}\vec{I}))$, and $\sum C_3(H(M^{(i)}\vec{I}))$, which is that $M^{(i)}$ $(i = 1,2,3)$ is a permutation of R, S, and C.

Since this is the only solution, it is the best a four-cell system can do. If we consider a system with more cells, there is a possibility that we can find solutions cancelling $\sum D^-(H(M^{(i)}\vec{I}))$ as well. Due to the fact that $D^-(H(\vec{I}))$ is only a small part of $D(H(\vec{I}))$, a solution which cancels $\sum D^-(H(M^{(i)}\vec{I}))$ will make only limited improvements compared to the current solution. To illustrate this, there is only 1 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 5 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 $D_6(H(\vec{I}))$.

4.2.4 The Optimal Four-Cell Systems

Now the question is: What kinds of systems have the properties stated in Theorem 4.3, under what conditions, and which of them requires the least number of conditions. The following theorem answers these questions.

Theorem 4.4 *Among all (sixty-four) four-cell systems, there are only four which reach the optimum asserted by Theorem 4.3 while imposing the minimum number of constraints on the linear map. They are* FRFR, FRSC, FCFC, *and* FCSR.

Proof:

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

$$L_1 = \begin{pmatrix} a & b & \eta \\ c & d & \eta' \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad L_1^{-1} = \begin{pmatrix} d & -b & -d\eta + b\eta' \\ -c & a & c\eta - a\eta' \\ 0 & 0 & 1 \end{pmatrix}. \tag{4.69}$$

(1) Choices on the second cell.

Case 1a FF:

Recall that the standard form of a forward cell (eq. 4.5) is

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

$\vec{M}$

To r

S. H

whi

S

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

$$\begin{cases} \eta &= 0, \\ \eta' &= 0, \\ c &= 0, \\ b &= 0, \\ a &= -1. \end{cases}$$

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

<u>Case 1b</u> FR:

From eq. (4.8), we have

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

Specifically, $L_1^{-1}R_1^{-1} \cdot L_1$ can be obtained from eq. (4.69), which yields

$$L_1^{-1}R_1^{-1} \cdot L_1 = \begin{pmatrix} d & -b & -d\eta + b\eta' \\ -c & a & c\eta - a\eta' \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & b & \eta \\ c & d & \eta' \\ 0 & 0 & 1 \end{pmatrix}$$

Since

cond

whic

Her

five

$$= \begin{pmatrix} ad+bc & 2bd & 2b\eta' \\ -2ac & -(ad+bc) & -2a\eta' \\ 0 & 0 & 1 \end{pmatrix}. \tag{4.70}$$

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

$$\begin{cases} bd &= 0 \\ ac &= 0 \\ b\eta' &= 0 \\ a\eta' &= 0 \end{cases},$$

which are equivalent to

$$\begin{cases} b &= 0 \\ c &= 0 \\ \eta' &= 0 \end{cases} \quad \text{or} \quad \begin{cases} a &= 0 \\ d &= 0 \\ \eta' &= 0 \end{cases}. \tag{4.71}$$

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

$$\text{Solution A}: \qquad\qquad \text{Solution B}:$$

$$L_1 = \begin{pmatrix} a & 0 & \eta \\ 0 & d & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad L_1 = \begin{pmatrix} 0 & b & \eta \\ c & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{4.72}$$

<u>Case 1c</u> FS:

Similar to FR, the map of FS is

$$\begin{aligned} \vec{M}^{FS} &= \vec{M}^S \circ \vec{M}^F \\ &= (\exp(:H(S^{-1}\vec{I}):)(SL^{-1}S^{-1}\vec{I})) \circ (\exp(:H(\vec{I}):)(L\vec{I})) \\ &= \exp(:H(\vec{I}):)\exp(:H(S^{-1}\cdot L\vec{I}):)(SL^{-1}S^{-1}\cdot L\vec{I}). \end{aligned}$$

Sinc

whi

whi

For

the

the

Since S is symplectic, $S^{-1} \cdot L$ is also symplectic. Therefore, $S^{-1} \cdot L$ can only be S, which means that $L = \hat{I}$. Like Case 1a, the conditions are

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

which is not an optimal solution.

Case 1d FC:

The transfer map of the system FC is

$$\begin{aligned} \vec{M}^{FC} &= \vec{M}^{C} \circ \vec{M}^{F} \\ &= (\exp(: H(L^{-1}R^{-1}S^{-1}\vec{I}) :)(RSL^{-1}R^{-1}S^{-1}\vec{I})) \circ (\exp(: H(\vec{I}) :)(L\vec{I})) \\ &= \exp(: H(\vec{I}) :) \exp(: H(L^{-1}R^{-1}S^{-1} \cdot L\vec{I}) :)(RSL^{-1}R^{-1}S^{-1} \cdot L\vec{I}). \end{aligned}$$

For this case, $L^{-1}R^{-1}S^{-1} \cdot L$ can be R or C, because it is antisymplectic. Since it has the form

$$\begin{aligned} L_1^{-1}R_1^{-1}S_1^{-1} \cdot L_1 &= \begin{pmatrix} d & -b & -d\eta + b\eta' \\ -c & a & c\eta - a\eta' \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & b & \eta \\ c & d & \eta' \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -(ad+bc) & -2bd & -2d\eta \\ 2ac & ad+bc & 2c\eta \\ 0 & 0 & 1 \end{pmatrix}, \end{aligned} \tag{4.73}$$

the solutions for an optimal system are

$$\begin{cases} b &= 0 \\ c &= 0 \\ \eta &= 0 \end{cases} \quad \text{or} \quad \begin{cases} a &= 0 \\ d &= 0 \\ \eta &= 0 \end{cases}.$$

The fact that these solutions require only three conditions makes them another set of candidates, which are

Solution A : Solution B :

$$L_1 = \begin{pmatrix} a & 0 & 0 \\ 0 & d & \eta' \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad L_1 = \begin{pmatrix} 0 & b & 0 \\ c & 0 & \eta' \\ 0 & 0 & 1 \end{pmatrix}. \tag{4.74}$$

In conclusion, in order to impose minimum conditions, the second cell can only be R or C.

(2) Choices for the third cell of the system FR××

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

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

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

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

whe

(

S

r-a-

I

whic

I

whic

any

(

F

A

I

whic

a five

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

<u>Case 2a</u> FRF:

Since $M(F) = \hat{I}$, we have $M^{(2)}(F) = RL^{-1}R^{-1} \cdot L$. For the two solutions, the x-a-δ block $M^{(2)}(F)$ are listed below.

In 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 agree with Theorem 4.3.

In case of Solution B, 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 is a possible solution, because it agrees with Theorem 4.3 and does not need any more conditions.

<u>Case 2b</u> FRR:

From $M(R) = L^{-1}R^{-1}$, we obtain

$$M^{(2)}(R) = L^{-1}R^{-1} \cdot RL^{-1}R^{-1} \cdot L = L^{-1} \cdot L^{-1}R^{-1} \cdot L.$$

In case of Solution A, we have

$$M_1^{(2)}(R) = \begin{pmatrix} d & 0 & -d\eta \\ 0 & a & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} d & 0 & -d\eta \\ 0 & -a & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which does not reach the optimum unless $d = 1$ and $\eta = 0$. This case corresponds to a five-condition solution, which is eliminated.

In case of Solution B, we have

$$M_1^{(2)}(R) = \begin{pmatrix} 0 & -b & 0 \\ -c & 0 & c\eta \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -b & 0 \\ c & 0 & c\eta \\ 0 & 0 & 1 \end{pmatrix},$$

which does not reach the optimum because it does not agree with Theorem 4.3.

<u>Case 2c</u> FRS:

From $M(S) = S^{-1}$, we obtain

$$M^{(2)}(S) = S^{-1} \cdot RL^{-1}R^{-1} \cdot L.$$

In case of Solution A, we have

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

which is a possible solution, because it agrees with Theorem 4.3.

In case of Solution B, we have

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

which does not reach the optimum, because it does not agree with Theorem 4.3.

<u>Case 2d</u> FRC:

From $M(C) = L^{-1}R^{-1}S^{-1}$, we obtain

$$M^{(2)}(C) = L^{-1}R^{-1}S^{-1} \cdot RL^{-1}R^{-1} \cdot L.$$

In case of Solution A, we have

$$M_1^{(2)}(C) = \begin{pmatrix} d & 0 & -d\eta \\ 0 & a & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} -d & 0 & -d\eta \\ 0 & a & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which does not reach the optimum unless $d = 1$ and $\eta = 0$. This case corresponds to a five-condition solution, which is eliminated.

In case of Solution A, we have

$$M_1^{(2)}(C) = \begin{pmatrix} 0 & -b & 0 \\ -c & 0 & c\eta \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & b & 0 \\ -c & 0 & c\eta \\ 0 & 0 & 1 \end{pmatrix},$$

which does not reach the optimum, because it does not agree with Theorem 4.3.

(3) Choices for the fourth cells of the systems FRF× and FRS×.

Case 3a FRF×:

Define $L^{FRF} = L \cdot RL^{-1}R^{-1} \cdot L$. 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

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

FRFF:

From $M(F) = \hat{I}$, we have

$$M^{(3)} = L \cdot RL^{-1}R^{-1} \cdot L.$$

Since

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

this system is not a solution (see Case 2b).

FRFR:

From $M(R) = L^{-1}R^{-1}$, we have

$$M^{(3)} = L^{-1}R^{-1} \cdot L \cdot RL^{-1}R^{-1} \cdot L.$$

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 agrees with Theorem 4.3.

In summary, the total map of system FRFR is

$$\vec{M}^{FRFR} = \exp(:H^F:)\exp(:H^C:)\exp(:H^S:)\exp(:H^R:)\vec{I}. \tag{4.75}$$

FRFS:

From $M(S) = S^{-1}$, we have

$$M^{(3)} = S^{-1} \cdot L \cdot RL^{-1}R^{-1} \cdot L.$$

Since

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

which does not agree with Theorem 4.3, this system is not a solution.

FRFC:

From $M(C) = L^{-1}R^{-1}S^{-1}$, we have

$$M^{(3)} = L^{-1}R^{-1}S^{-1} \cdot L \cdot RL^{-1}R^{-1} \cdot L.$$

Therefore, the x-a-b block is

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

$$= \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 2c\eta \\ 0 & 0 & 1 \end{pmatrix},$$

which is not a solution, because it does not agree with Theorem 4.3.

Case 3b FRS×:

Define $L^{FRS} = SLS^{-1} \cdot RL^{-1}R^{-1} \cdot L$. We have

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

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

$$L_1 = \begin{pmatrix} a & 0 & \eta \\ 0 & d & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

FRSF:

From $M(F) = \hat{I}$, we have

$$M^{(3)} = SLS^{-1} \cdot RL^{-1}R^{-1} \cdot L.$$

Since

$$S_1^{-1} \cdot R_1 L_1^{-1} R_1^{-1} \cdot L_1 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which does not agree with Theorem 4.3, this system is not a solution.

FRSR:

From $M(R) = L^{-1}R^{-1}$, we have

$$M^{(3)} = L^{-1}R^{-1} \cdot SLS^{-1} \cdot RL^{-1}R^{-1} \cdot L.$$

Therefore, the x-a-δ block is

$$M_1^{(3)}(R) = \begin{pmatrix} d & 0 & -d\eta \\ 0 & a & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 & \eta \\ 0 & d & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & -2d\eta \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which is not a solution, because it does not agree with Theorem 4.3.

FRSS:

From $M(S) = S^{-1}$, we have

$$M^{(3)} = S^{-1} \cdot SLS^{-1} \cdot RL^{-1}R^{-1} \cdot L = LS^{-1} \cdot RL^{-1}R^{-1} \cdot L,$$

which is the same as FRSF.

FRSC:

From $M(C) = L^{-1}R^{-1}S^{-1}$, we have

$$M^{(3)} = L^{-1}R^{-1}S^{-1} \cdot SLS^{-1} \cdot RL^{-1}R^{-1} \cdot L = L^{-1}R^{-1} \cdot LS^{-1} \cdot RL^{-1}R^{-1} \cdot L.$$

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 agrees with Theorem 4.3.

In summary, the total map of system FRSC is

$$\vec{M}^{FRSC} = \exp(:H^F:)\exp(:H^R:)\exp(:H^S:)\exp(:H^C:)\vec{I}, \tag{4.76}$$

which satisfies the requirements.

(4) Choices for the third cell of system FC××.

Define $L^{FC} = SRL^{-1}R^{-1}S^{-1}\cdot L$. For systems FC×, we have

$$M^{(2)}(\times) = M(\times)\cdot L^{FC}.$$

<u>Case 4a</u> FCF:

From $M(F) = \hat{I}$, we have

$$M^{(2)}(F) = SRL^{-1}R^{-1}S^{-1}\cdot L.$$

In 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 agree with Theorem 4.3.

In case of Solution B, 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 is a possible solution, because it agrees with Theorem 4.3.

<u>Case 4b</u> FCR:

From $M(R) = L^{-1}R^{-1}$, we have

$$M^{(2)}(R) = L^{-1}R^{-1}\cdot SRL^{-1}R^{-1}S^{-1}\cdot L.$$

Since

$$R_1^{-1}\cdot S_1R_1L_1^{-1}R_1^{-1}S_1^{-1}\cdot L_1=\begin{pmatrix}1&0&0\\0&-1&0\\0&0&1\end{pmatrix}\quad\text{or}\quad\begin{pmatrix}-1&0&0\\0&1&0\\0&0&1\end{pmatrix},$$

which does not agree with Theorem 4.3, this is not a solution.

Case 4c FCS:

From $M(S)=S^{-1}$, we have

$$M^{(2)}(S)=S^{-1}\cdot SRL^{-1}R^{-1}S^{-1}\cdot L=RL^{-1}R^{-1}S^{-1}L.$$

In case of Solution A, we have

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

which is a possible solution, because it agrees with Theorem 4.3.

In case of Solution B, we have

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

which does not reach the optimum, because it does not agree with Theorem 4.3.

Case 4d FCC:

From $M(C)=L^{-1}R^{-1}S^{-1}$, we have

$$M^{(2)}(C)=L^{-1}R^{-1}S^{-1}\cdot SRL^{-1}R^{-1}S^{-1}\cdot L=L^{-1}\cdot L^{-1}R^{-1}S^{-1}\cdot L.$$

Since

$$L_1^{-1}R_1^{-1}S_1^{-1}\cdot L_1=\begin{pmatrix}-1&0&0\\0&1&0\\0&0&1\end{pmatrix}\quad\text{or}\quad\begin{pmatrix}1&0&0\\0&-1&0\\0&0&1\end{pmatrix},$$

which does not agree with Theorem 4.3, this is not a solution.

(5) Choices for the fourth cell of the systems FCF$\times$ and FCS$\times$.

<u>Case 5a</u> FCF$\times$:

Define $L^{FCF} = L \cdot SRL^{-1}R^{-1}S^{-1} \cdot L$. We have

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

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

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

FCFF:

From $M(F) = \hat{I}$, we have

$$M^{(3)}(F) = L \cdot SRL^{-1}R^{-1}S^{-1} \cdot L.$$

Since

$$S_1R_1L_1^{-1}R_1^{-1}S_1^{-1} \cdot L_1 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which does not agree with Theorem 4.3, this system is not a solution.

FCFR:

From $M(R) = L^{-1}R^{-1}$, we have

$$M^{(3)}(R) = L^{-1}R^{-1} \cdot L \cdot SRL^{-1}R^{-1}S^{-1} \cdot L.$$

Therefore, the x-a-δ block is

$$M_1^{(3)}(R) = \begin{pmatrix} 0 & -b & b\eta' \\ -c & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & b & 0 \\ c & 0 & \eta' \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 2b\eta' \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which is not a solution, because it does not agree with Theorem 4.3.

FCFS:

From $M(S) = R^{-1}$, we have

$$M^{(3)}(S) = S^{-1} \cdot L \cdot SRL^{-1}R^{-1}S^{-1} \cdot L.$$

Since

$$S_1 R_1 L_1^{-1} R_1^{-1} S_1^{-1} \cdot L_1 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which does not agree with Theorem 4.3, this system is not a solution.

FCFC:

From $M(C) = L^{-1}R^{-1}S^{-1}$, we have

$$M^{(3)}(C) = L^{-1}R^{-1}S^{-1} \cdot L \cdot SRL^{-1}R^{-1}S^{-1} \cdot L.$$

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 agrees with Theorem 4.3.

In summary, the total map of system FCFC is

$$\vec{M}^{FCFC} = \exp(:H^F:)\exp(:H^R:)\exp(:H^S:)\exp(:H^C:)\vec{I}, \tag{4.77}$$

which satisfies the requirements.

<u>Case 5b</u> FCS×:

Define $L^{FCS} = SLS^{-1} \cdot SRL^{-1}R^{-1}S^{-1} \cdot L$. We have

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

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

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

FCSF:

From $M(F) = \hat{I}$, we have

$$M^{(3)}(F) = SLS^{-1} \cdot SRL^{-1}R^{-1}S^{-1} \cdot L = SL \cdot RL^{-1}R^{-1}S^{-1} \cdot L.$$

Since

$$R_1L_1^{-1}R_1^{-1}S_1^{-1}L_1 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which does not agree with Theorem 4.3, this system is not a solution.

FCSR:

From $M(R) = L^{-1}R^{-1}$, we have

$$M^{(3)}(R) = L^{-1}R^{-1} \cdot SLS^{-1} \cdot SRL^{-1}R^{-1}S^{-1} \cdot L = L^{-1}R^{-1} \cdot SL \cdot RL^{-1}R^{-1}S^{-1} \cdot L. \quad (4.78)$$

Therefore, the x-a-δ block is

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

which shows that this system is a solution, because it agrees with Theorem 4.3.

In summary, the total map of system FCSR is

$$\vec{M}^{FCSR} = \exp(:H^F:)\exp(:H^C:)\exp(:H^S:)\exp(:H^R:)\vec{I}, \quad (4.79)$$

which satisfies the requirements.

FCSS:

From $M(S) = S^{-1}$, we have

$$M^{(3)}(S) = S^{-1} \cdot SLS^{-1} \cdot SRL^{-1}R^{-1}S^{-1} \cdot L = L \cdot RL^{-1}R^{-1}S^{-1} \cdot L,$$

which is the same as FCSF.

FCSC:

From $M(C) = L^{-1}R^{-1}S^{-1}$, we have

$$M^{(3)}(C) = L^{-1}R^{-1}S^{-1} \cdot SLS^{-1} \cdot SRL^{-1}R^{-1}S^{-1} \cdot L = L^{-1}R^{-1}L \cdot RL^{-1}R^{-1}S^{-1} \cdot L. \quad (4.80)$$

Therefore, the x-a-δ block is

$$\begin{aligned} M_1^{(3)}(C) &= \begin{pmatrix} d & 0 & 0 \\ 0 & a & -a\eta' \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 & 0 \\ 0 & d & \eta' \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & -2a\eta' \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

which is not a solution, because it does not agree with Theorem 4.3, and this concludes the proof.

The next theorem gives the linear conditions of the y-b block for the optimal systems.

Theorem 4.5 *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.*

Proof:

Let $L_2 = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$ be the y-b block of L.

(1) FRFR:

From eq. (4.20), the total map of system FRFR is

$$\begin{aligned}\vec{M}^{FRFR} &= \exp(:H(\vec{I}):)\exp(:H(L^{-1}R^{-1}\cdot L\vec{I}):) \\ &\quad \exp(:H(RL^{-1}R^{-1}\cdot L\vec{I}):) \\ &\quad \exp(:H(L^{-1}R^{-1}\cdot L\cdot RL^{-1}R^{-1}\cdot L\vec{I}):)\vec{I}.\end{aligned}$$

From the second cell, we have

$$L_2^{-1}R_2^{-1}\cdot L_2 = \begin{pmatrix} h & -f \\ -g & e \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} eh+fg & 2fh \\ -2eg & -(eh+fg) \end{pmatrix}, \quad (4.81)$$

which gives the following possible solutions.

$$\text{Solution A}: \begin{cases} e = 0 \\ h = 0 \end{cases} \quad \text{and} \quad \text{Solution B}: \begin{cases} f = 0 \\ g = 0 \end{cases}.$$

In the case of solution A, we have

$$L_2^{-1}R_2^{-1}L_2 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$R_2L_2^{-1}R_2^{-1}\cdot L_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$L_2^{-1}R_2^{-1}\cdot L_2\cdot R_2L_2^{-1}R_2^{-1}\cdot L_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

which is a solution

In the case of solution B, we have

$$L_2^{-1}R_2^{-1}\cdot L_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$R_2L_2^{-1}R_2^{-1}\cdot L_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$L_2^{-1}R_2^{-1}\cdot L_2\cdot R_2L_2^{-1}R_2^{-1}\cdot L_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

Systems	Linear Conditions
F R S C	$(a\|\delta)=0,\ (x\|a)=(a\|x)=0$
F R F R	$(a\|\delta)=0,\ (x\|x)=(a\|a)=0$
F C S R	$(x\|\delta)=0,\ (x\|a)=(a\|x)=0$
F C F C	$(x\|\delta)=0,\ (x\|x)=(a\|a)=0$

Figure 4.1: Optimal four-cell systems and the first-order requirements to achieve their optimum.

which is also a solution.

Because of midplane symmetry,

$$\vec{M}^{FRFR} = \exp(:H^F:)\exp(:H^C:)\exp(:H^S:)\exp(:H^R:)\vec{I} \tag{4.82}$$

in both cases.

(2) FRSC, FCFC, FCSR

Since $S_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, we have $S_2 = F_2$ *and* $C_2 = R_2$. This implies that FRSC, FCFC, and FCSR have the same matrix M_2^i $(i = 1,2,3)$ as that of FRFR. So, the requirements on L_2 should also be the same, which concludes the proof.

All four optimal systems are listed in Table 4.1.

4.3 Order-by-Order Solutions

Next we will study the nonlinear conditions for higher-order achromats of the optimal systems. Of the four systems, there are only two different maps, which are

$$\vec{M}^{FRSC} = \vec{M}^{FCFC} = \exp(:H^F:)\exp(:H^R:)\exp(:H^S:)\exp(:H^C:)\vec{I} \tag{4.83}$$

and

$$\vec{M}^{FCSR} = \vec{M}^{FRFR} = \exp(:H^F:)\exp(:H^C:)\exp(:H^S:)\exp(:H^R:)\vec{I}. \tag{4.84}$$

The lemma below shows that one can be obtained from the other through a simple transformation. Thus, we need to study only one map to find out the conditions for achromats.

Lemma 4.1 *By switching A and B, the maps $\vec{M}^{FRSC}$ and $\vec{M}^{FCFC}$ are transformed to $\vec{M}^{FRSC}$ and $\vec{M}^{FRFR}$, respectively.*

Proof:

Under the transformation, we have

$$
\begin{aligned}
H^F &= A+B+C+D \quad\rightarrow\quad B+A+C+D = H^F, \\
H^R &= A-B-C+D \quad\rightarrow\quad B-A-C+D = H^C, \\
H^S &= -A-B+C+D \quad\rightarrow\quad -B-A+C+D = H^S, \\
H^C &= -A+B-C+D \quad\rightarrow\quad -B+A-C+D = H^R,
\end{aligned}
$$

which entails that the transformation of the maps is

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

which concludes the proof.

The next theorem gives a general solution for arbitrary-order achromats. It is not necessarily the best possible, yet it holds for an arbitrary order.

Theorem 4.6 *For a given order n, the optimal systems* FRFR, FRSC, FCSR, *and* FCFC *are achromats if $A =_n 0$ or $B =_n 0$, and $D =_{n+1} 0$.*

Proof:

The mathematical induction method is used.

Let us first consider systems FRSC and FCFC. From the proof of Theorem 4.4, 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.

(1) The second order:

Using the B-C-H 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 $D =_3 0$.

(2) The third order:

Similar to the second-order case, $\vec{M}$ is simplified to

$$\begin{aligned}
\vec{M} &=_3 \exp(:H_4^F:)\exp(:H_4^R:)\exp(:H_4^S:)\exp(:H_4^C:)\vec{I} \\
&=_3 \exp(:f_3^F + f_4^F:)\exp(:f_3^R + f_4^R:) \\
&\quad \exp(:f_3^S + f_4^S:)\exp(:f_3^C + f_4^C:)\vec{I} \\
&=_3 \exp(:f_4^F + f_4^R + f_4^S + f_4^C:)\exp(:f_3^F:) \\
&\quad \exp(:f_3^R:)\exp(:f_3^S:)\exp(:f_3^C:)\vec{I} \\
&=_3 \exp(:4D_4:)\exp(:f_3^F + f_3^R + \frac{1}{2}[f_3^F, f_3^R]:) \\
&\quad \exp(:f_3^S + f_3^C + \frac{1}{2}[f_3^S, f_3^C]:)\vec{I} \\
&=_3 \exp(:4D_4 + \frac{1}{2}([f_3^F, f_3^R] + [f_3^S, f_3^C] + [f_3^F + f_3^R, f_3^S + f_3^C]):)\vec{I},
\end{aligned}$$

where

$$
\begin{aligned}
&[f_3^F, f_3^R] + [f_3^S, f_3^C] + [f_3^F + f_3^R, f_3^S + f_3^C] \\
&\quad = [A_3 + B_3 + C_3 + D_3, A_3 - B_3 - C_3 + D_3] \\
&\qquad + [-A_3 - B_3 + C_3 + D_3, -A_3 + B_3 - C_3 + D_3] \\
&\qquad + 4[A_3 + D_3, -A_3 + D_3] \\
&\quad = 2[B_3 + C_3, A_3 + D_3] + 2[-B_3 + C_3, -A_3 + D_3] + 8[A_3, D_3] \\
&\quad = 4[B_3, A_3] + 4[C_3, D_3] + 8[A_3, D_3] \\
&\quad = 4[B_3, A_3].
\end{aligned}
$$

Altogether, the map has the form

$$\vec{M} =_3 \exp(: 4D_4 + 2[B_3, A_3] :)\vec{I},$$

which entails that $A =_3 0$ or $B =_3 0$, and $D =_4 0$ are solutions for a third-order achromat.

(3) The nth order:

Let us assume that $A =_{n-1} 0$ or $B =_{n-1} 0$, and $D =_n 0$ are solutions for an $(n-1)$th-order achromat.

a) $A =_{n-1} 0$, $D =_n 0$.

In this case, we have

$$
\begin{aligned}
H_{n-1}^F &= \sum_{i=3}^{n-1} (B_i + C_i), \\
H_{n-1}^R &= \sum_{i=3}^{n-1} (-B_i - C_i), \\
H_{n-1}^S &= \sum_{i=3}^{n-1} (-B_i + C_i),
\end{aligned}
$$

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

which entails that

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

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

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

which shows that $A_n = 0$ and $D_{n+1} = 0$ are a solution for the nth order.

In conclusion, $A =_n 0$ and $D =_{n+1} 0$ are a solution for an nth-order achromat.

b) $B =_{n-1} 0$, $D =_n 0$.

In this case, we have

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

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

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

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

which implies that

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

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

$$\begin{aligned}
\vec{M} \;&=_n\; \exp(:H^F_{n-1} + f^F_n + f^F_{n+1}:)\exp(:H^R_{n-1} + f^R_n + f^R_{n+1}:) \\
&\quad \exp(:H^S_{n-1} + f^S_n + f^S_{n+1}:)\exp(:H^C_{n-1} + f^C_n + f^C_{n+1}:)\vec{I} \\
&=_n\; \exp(:4D_{n+1}:)\exp(:H^F_{n-1} + f^F_n:) \\
&\quad \exp(:f^R_n + f^S_n + \frac{1}{2}([H^R_{n-1}, f^S_n] + [f^R_n, H^S_{n-1}]):)\exp(:H^C_{n-1} + f^C_n:)\vec{I} \\
&=_n\; \exp(:4D_{n+1}:)\exp(:H^F_{n-1} + f^F_n:)\exp(:H^C_{n-1} + f^R_n + f^S_n + f^C_n \\
&\quad +\frac{1}{2}([H^R_{n-1}, f^S_n] + [f^R_n, H^S_{n-1}] + [f^R_n + f^S_n, H^C_{n-1}]):)\vec{I} \\
&=_n\; \exp(:4D_{n+1} + \frac{1}{2}([H^F_{n-1}, f^R_n + f^S_n + f^C_n] + [f^F_n, H^C_{n-1}] \\
&\quad +[H^R_{n-1}, f^S_n] + [f^R_n, H^S_{n-1}] + [f^R_n + f^S_n, H^C_{n-1}]):)\vec{I} \\
&=_n\; \exp(:4D_{n+1} + \frac{1}{2}([f^F_3, f^R_n + f^S_n + f^C_n] + [f^F_n, f^C_3] \\
&\quad +[f^R_3, f^S_n] + [f^R_n, f^S_3] + [f^R_n + f^S_n, f^C_3]):)\vec{I} \\
&=_n\; \exp(:4D_{n+1} + \frac{1}{2}([f^R_n + f^S_n, f^S_3 + f^C_3]):)\vec{I} \\
&=_n\; \exp(:4D_{n+1} + 2[B_n, A_3]:)\vec{I},
\end{aligned}$$

which shows that $B =_n 0$ and $D =_{n+1} 0$ are also a solution for an nth-order achromat. According to Lemma 4.1, the systems FRFR and FCSR have the same solutions, which concludes the proof.

Considering the number of conditions, the two solutions obtained above are equivalent. This is shown by the following theorem.

Theorem 4.7 *For a given order n, the number of monomials in A is the same as that in B.*

Proof:

A general term from H is $C_{i_x i_a i_y i_b i_\delta} x^{i_x} a^{i_a} y^{i_y} b^{i_b} \delta^{i_\delta}$, where $(i_y + i_b)$ is even.

For terms from A_n, $i_x + i_a$ is odd and $i_a + i_b$ is even. Specifically, when i_x is odd, i_a is even, i_b is even, and i_y is even; when i_x is even, i_a is odd, i_b is odd, and i_y is odd. Therefore A_n can be decomposed to

$$
\begin{aligned}
A_n &= \sum_{\substack{i_b \\ \text{even}}} \sum_{\substack{i_y \\ \text{even}}} \sum_{\substack{i_a \\ \text{even}}} \sum_{\substack{i_x \\ \text{odd}}} C_{i_x i_a i_y i_b i_\delta} x^{i_x} a^{i_a} y^{i_y} b^{i_b} \delta^{i_\delta} \\
&+ \sum_{\substack{i_b \\ \text{odd}}} \sum_{\substack{i_y \\ \text{odd}}} \sum_{\substack{i_a \\ \text{odd}}} \sum_{\substack{i_x \\ \text{even}}} C_{i_x i_a i_y i_b i_\delta} x^{i_x} a^{i_a} y^{i_y} b^{i_b} \delta^{i_\delta} \quad (i_x + i_a + i_y + i_b + i_\delta = n).
\end{aligned}
$$

For terms from B_n, $i_x + i_a$ is odd and $i_a + i_b$ is odd. Specifically, when i_x is odd, i_a is even, i_b is odd, and i_y is odd; when i_x is even, i_a is odd, i_b is even, and i_y is even. Therefore B_n can be decomposed to

$$
B_n = \sum_{\substack{i_b \\ \text{even}}} \sum_{\substack{i_y \\ \text{even}}} \sum_{\substack{i_a \\ \text{odd}}} \sum_{\substack{i_x \\ \text{even}}} C_{i_x i_a i_y i_b i_\delta} x^{i_x} a^{i_a} y^{i_y} b^{i_b} \delta^{i_\delta}
$$

$$+ \sum_{\substack{i_b \\ \text{odd}}} \sum_{\substack{i_y \\ \text{odd}}} \sum_{\substack{i_a \\ \text{even}}} \sum_{\substack{i_x \\ \text{odd}}} C_{i_x i_a i_y i_b i_\delta} x^{i_x} a^{i_a} y^{i_y} b^{i_b} \delta^{i_\delta} \quad (i_x + i_a + i_y + i_b + i_\delta = n).$$

For any given i_y and i_b, $\sum_{\substack{i_a \\ \text{even}}} \sum_{\substack{i_x \\ \text{odd}}} = \sum_{\substack{i_a \\ \text{odd}}} \sum_{\substack{i_x \\ \text{even}}}$, which concludes the proof.

Since D cannot be cancelled by symmetry, we always have to kill it to get an achromat. This means that 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 a best solution exists for up to the fourth order. According to computational results, this solution is also valid for the fifth- and sixth-order achromats. Before studying the theorem itself, a few observations necessary to the proof are shown below.

Considering the commutator of two general terms from the Hamiltonian, we have

$$[C_{i_x i_a i_y i_b i_\delta} x^{i_x} a^{i_a} y^{i_y} b^{i_b} \delta^{i_\delta}, C_{i'_x i'_a i'_y i'_b i'_\delta} x^{i'_x} a^{i'_a} y^{i'_y} b^{i'_b} \delta^{i'_\delta}]$$

$$= C_{i_x i_a i_y i_b i_\delta} C_{i'_x i'_a i'_y i'_b i'_\delta} ((i_x i'_a - i_a i'_x) x^{i_x+i'_x-1} a^{i_a+i'_a-1} y^{i_y+i'_y} b^{i_b+i'_b} \delta^{i_\delta+i'_\delta}$$

$$+ (i_y i'_b - i_b i'_y) x^{i_x+i'_x} a^{i_a+i'_a} y^{i_y+i'_y-1} b^{i_b+i'_b-1} \delta^{i_\delta+i'_\delta}).$$

(1) 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.

(2) If the first term is from A, i.e., $i_x + i_a$ is odd and $i_a + i_b$ is even, and the second term is from B, i.e. $i'_x + i'_a$ is odd and $i'_a + i'_b$ is odd, we have

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

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

which entails that $[A, B]$ gives terms in D.

(3) If the first term is from A, i.e., $i_x + i_a$ is odd and $i_a + i_b$ is even, and the second term is from D, i.e., $i'_x + i'_a$ is even and $i'_a + i'_b$ is even, we have

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

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

which entails that $[A, D]$ gives terms in B.

(4) If the first term is from B, i.e., $i_x + i_a$ is odd and $i_a + i_b$ is odd, and the second term is from D, i.e., $i'_x + i'_a$ is even and $i'_a + i'_b$ is even, we have

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

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

which entails that $[B, D]$ gives terms in A.

(5) If the first term is from A, i.e., $i_x + i_a$ is odd and $i_a + i_b$ is even, and the second term is from C, i.e., $i'_x + i'_a$ is even and $i'_a + i'_b$ is odd, we have

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

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

which entails that $[A, C]$ gives terms in A.

(6) If the first term is from B, i.e., $i_x + i_a$ is odd and $i_a + i_b$ is odd, and the second term is from C, i.e., $i'_x + i'_a$ is even and $i'_a + i'_b$ is odd, we have

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

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

which entails that $[B, C]$ gives terms in B.

(7) If the first term is from D, i.e., $i_x + i_a$ is even and $i_a + i_b$ is even, and the second term is from C', i.e., $i'_x + i'_a$ is even and $i'_a + i'_b$ is odd, we have

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

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

which entails that $[D, C']$ gives terms in D.

Theorem 4.8 *For the optimal systems, achromats up to the fourth order can be obtained by cancelling D in the total map.*

Proof:

(1) Systems FRSC and FCFC:

a) The second order:

From the proof of Theorem 4.6, the map is

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

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

b) The third order:

Also from the proof of Theorem 4.6, the map is

$$\begin{aligned} \vec{M} &=_3 \exp(:H_4^F:)\exp(:H_4^R:)\exp(:H_4^S:)\exp(:H_4^C:)\vec{I} \\ &=_3 \exp(:4D_4 + 2[B_3, A_3]:)\vec{I}. \end{aligned}$$

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 cancelling D_4 and A_4 (or B_4) separately. Therefore, the best third-order solution is

$$D_4 = -\frac{1}{2}[B_3, A_3]. \tag{4.85}$$

c) The fourth order:

Using the B-C-H formula, $\vec{M}$ can be transformed to

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

$$=_4 \quad \exp(:4D_5:)\exp(:\hat{H}:)\vec{I},$$

where

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

$$= 2[B_4, A_3] + 2[B_3, A_4] - 2[C_3, -\frac{1}{2}[B_3, A_3]] + 4[A_3, -\frac{1}{2}[B_3, A_3]]$$
$$-2[A_3, [A_3, B_3]] - \frac{2}{3}([C_3, [A_3, B_3]] + [B_3, [A_3, C_3]])$$
$$= 2[B_4, A_3] + 2[B_3, A_4] + \frac{1}{3}[C_3, [A_3, B_3]] + \frac{2}{3}[B_3, [C_3, A_3]].$$

Altogether, the map is

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

Since $2[B_4, A_3] + 2[B_3, A_4] + \frac{1}{3}[C_3, [A_3, B_3]] + \frac{2}{3}[B_3, [C_3, A_3]]$ belongs to D_5, the best solution for the fourth order is

$$D_5 = -(2[B_4, A_3] + 2[B_3, A_4] + \frac{1}{3}[C_3, [A_3, B_3]] + \frac{2}{3}[B_3, [C_3, A_3]]). \tag{4.86}$$

(2) Systems FCSR and FRFR:

The best solution for a fourth-order achromat can be obtained by switching A and B, which is

$$D_3 = 0, \tag{4.87}$$
$$D_4 = -\frac{1}{2}[A_3, B_3], \tag{4.88}$$
$$D_5 = -(2[A_4, B_3] + 2[A_3, B_4] + \frac{1}{3}[C_3, [B_3, A_3]] + \frac{2}{3}[A_3, [C_3, B_3]]), \tag{4.89}$$

which concludes the proof.

Computer results show that this property holds for the fifth and sixth orders also.

As the conclusion for this chapter, the following conjecture is presented:

Conjecture 4.1 *For the optimal systems, achromats up to an arbitrary order can be obtained by cancelling D in the total map.*

Chapter 5

Applications

When a beam optical system is actually designed, one of the design codes has to be used (see Section 1.1). In the case of high-order achromats, DA techniques are needed, because when the order is as high as 5, the map and Lie factorization can be computed efficiently only through DA techniques. The proofs in Theorems 2.5 and 2.6 actually provide the DA algorithms for computing various Lie factorizations. Since the Lie factorizations are always obtained from the map, the B-C-H formula is not necessary, which greatly simplifies the complexity of the processes. Due to the fact that no second-order polynomials appear in the Lie exponents, this algorithm works for an arbitrary order as well. Throughout the design processes, COSY INFINITY is used, which contains all the tools important to beam optical design. Therefore, everything can be done in one shot, including map computation, Lie coefficients extraction, 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 cancelled by symmetry. In our case, sometimes the package LMDIF from Argonne National Laboratory was used, other times a matrix inverter was used.

In the next sections, several designs of third-, fourth-, and fifth-order achromats are presented. All of them are direct results of the arbitrary-order achromat theory presented in Chapter 4; and the circular systems are designed in such a way that Brown's second-order achromat theory can be applied, where two second-order knobs are saved. One of them, a circular system (see Section 5.2), is not only a third-order achromat, but is also energy-isochronous to the third order, which is an ideal multi-pass time-of-flight mass spectrometer. Another design, a single-pass third-order achromat (see Section 5.1), presents a compact system which can be transformed to a single-pass time-of-flight mass spectrometer. Two more designs push the orders to 4 and 5, which further verified the arbitrary-order achromat theory.

5.1 A Third-Order Achromat - FRSC

5.1.1 The First-Order Layout

Our design is aimed at a system for 200 MeV protons [Wan93a]. The first-order forward cell contains a 120-degree inhomogeneous bending magnet and four quadrupoles (Fig. 5.1). The compact layout shows the potential application of high-order achromats for single-pass time-of-flight spectrometers. The forward cell is symmetric around the midpoint, which entails that $(x|x) = (a|a)$ and $(y|y) = (b|b)$. So we need to fit only 5 conditions, instead of 7, in order to obtain a map of the form shown in Table 5.1. During the process of fitting, the two drifts between quadrupoles and dipole are fixed; the variables are the field index of the dipole, the two field strengths of the quadrupoles, and the drifts before and between the quadrupoles. Table 5.2 shows the parameters chosen as our starting conditions for higher order optimization. The long drifts are spaces where higher-order elements will be placed.

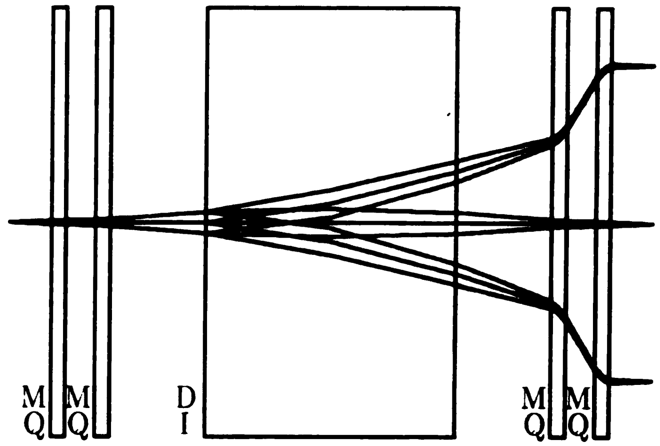

Figure 5.1: The FRSC third-order achromat: The first-order forward cell. Only the bending magnet and the quadrupoles are shown. The phase advances are $\mu_x = \pi$ and $\mu_y = \pi/2$. It also shows that at the end of the cell, the dispersion is not corrected, but dispersive rays are parallel to the on-energy rays that start with the same initial conditions.

5.1.2 The Second- and Third-Order Achromat

To make this system a second-order achromat, ten sextupoles were inserted symmetrically with respect to the dipole (but not symmetrically excited). The sextupoles provide us the knobs to correct second-order aberrations. The required values are rather weak (Table 5.4), which indicates that the first-order layout gives weak higher-

-1.000000	0.0000000E+00	0.0000000E+00	0.0000000E+00	0.2463307E-15	100000
0.0000000E+00	-1.000000	0.0000000E+00	0.0000000E+00	-7.366987	010000
0.0000000E+00	0.0000000E+00	0.1387779E-15	1.000000	0.0000000E+00	001000
0.0000000E+00	0.0000000E+00	-1.000000	0.1526557E-15	0.0000000E+00	000100
0.0000000E+00	0.0000000E+00	0.0000000E+00	0.0000000E+00	1.000000	000010
7.366987	-0.2775558E-15	0.0000000E+00	0.0000000E+00	1.733828	000001

Table 5.1: The FRSC third-order achromat: The COSY INFINITY output of the first-order map of the forward cell. The first five columns are Taylor coefficients of x_f, a_f, y_f, b_f, and t_f as functions of x_i, a_i, y_i, b_i, t_i, and δ_i, and the sixth column contains the powers associated with the coefficients on the same row. From left to right, the sub-columns stand for x_i, a_i, y_i, b_i, t_i, and δ_i, respectively.

Element	Field/Length	Gradient
Drift 1	0.871354(m)	
Quadrupole 1	1.86498(kG)	0.372997(kG/cm)
Drift 2	0.624612(m)	
Quadrupole 2	-2.67177(kG)	-0.534353(kG/cm)
Drift 3	2.0(m)	
Dipole	8.6(kG)	
Field index(n_1)		0.398455

Table 5.2: The FRSC third-order achromat: The field strengths and drift lengths of the first-order layout. n_1 is the first-order derivative of the field of the bending magnet over the bending radius r (dimensionless).

order aberrations, and the newly introduced sextupoles will not produce strong third-order aberrations either. The map of the four-cell system was then computed. As the theory predicted, it is free of all second-order aberrations. Note that this system was designed before the analytical theory was developed. The requirement of ten sextupoles was obtained from a computer theorem-proving program discussed in references [Wan92, Wan93a].

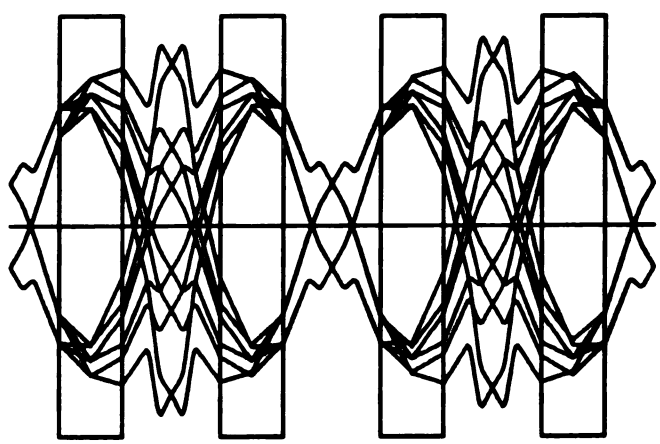

Figure 5.2: The FRSC third-order achromat: The third-order x-z beam envelope where only the bending magnets are shown. The mirror symmetry between cells F and R and that between cells S and C is clearly shown in the beam trajectories.

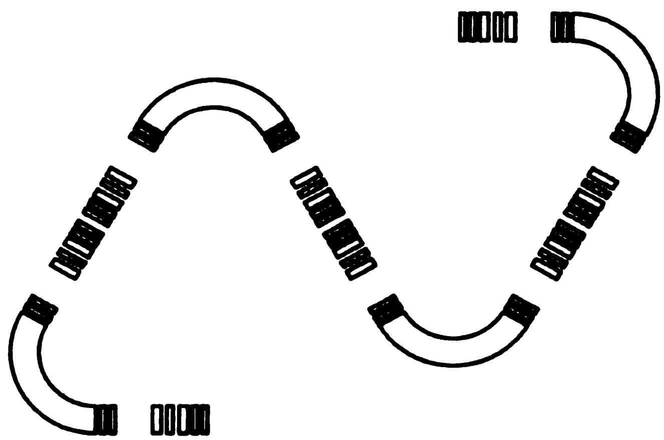

Figure 5.3: The FRSC third-order achromat: The third-order lab layout. The "S"-shaped geometry helps make it a compact system.

The last step is to correct third-order aberrations. All third-order elements are superimposed in the existing dipoles, quadrupoles, and sextupoles because this configuration tends to require weaker octupole fields. Specifically, 14 octupoles are placed inside other multipoles and an octupole component is added to the inhomogeneous dipole field through the curvature of the pole faces. So there are 15 variables for 15 conditions. The results show that very weak octupoles can meet all the conditions (Table 5.4). Figures 5.2 and 5.3 show the beam envelope and the lab coordinate layout of the whole system. Regarding the area it occupies, this is a rather compact system. Table 5.3 presents the third-order map, which shows that the only terms left that are nonzero are the dependents of time-of-flight on energy spread up to the third order. This makes the system an effective time-of-flight spectrograph.

Using an emittance of 1 mm mrad and a momentum spread of 1%, Figure 5.4 shows the eighth-order beam around the final focal point. The sum of the aberrations at the focal point is about 10 μm horizontally and 3 μm vertically, which is quite small.

1.000000	0.0000000E+00	0.0000000E+00	0.0000000E+00	0.0000000E+00	100000
0.0000000E+00	1.000000	0.0000000E+00	0.0000000E+00	0.0000000E+00	010000
0.0000000E+00	0.0000000E+00	1.000000	0.0000000E+00	0.0000000E+00	001000
0.0000000E+00	0.0000000E+00	0.0000000E+00	1.000000	0.0000000E+00	000100
0.0000000E+00	0.0000000E+00	0.0000000E+00	0.0000000E+00	1.000000	000010
0.0000000E+00	0.0000000E+00	0.0000000E+00	0.0000000E+00	6.935312	000001
0.0000000E+00	0.0000000E+00	0.0000000E+00	0.0000000E+00	-21.18904	000002
0.0000000E+00	0.0000000E+00	0.0000000E+00	0.0000000E+00	59.36542	000003

Table 5.3: The FRSC third-order achromat: The COSY output of the third-order map. A third-order achromat is reached and only $(t|\delta^n)$ $(n = 1,2,3)$ is not cancelled. (Any number smaller than 1E-11 is set to zero.)

5.2 A Third-Order Achromat - FRFR

A third-order achromat based on the first-order layout of the Experimental Storage Ring (ESR) at Germany is designed to convert it into a time-of-flight mass spectrometer without using electron cooling, so that the masses of short-lived radioactive nuclei can be directly measured.

The ESR ring contains six dipoles, twenty quadrupoles, and eight sextupoles, as well as RF cavities, beam cooling devices, and the injection-extraction system. Two long, straight sections divide it into two identical parts, each of which is symmetric about its center. (Figure 5.5) [Franz87]. It is much easier to take half rather than a quarter of the ring as the forward cell. Consequently, the other half should be the reversed cell, and an achromat corresponds to two turns of the ring.

Since this is a FRFR system, the first-order matrix of the forward cell has to satisfy $(x|x) = (a|a) = (x|\delta) = 0$ (Table 4.1). In this design, the quadrupoles are excited symmetrically, which ensures that $(x|x) = (a|a)$ and $(y|y) = (b|b)$ and reduces the number of first-order knobs from 5 to 3. In order to make a system a mass spectrometer, $(t|\delta)$ has be cancelled, too, which adds one more condition and hence the total number of conditions is 4. After the fine tuning of the free knob (the third

Strengths of the Multipoles (Aperture 10 cm)			
Sextupoles		Octupoles	
Gradient (kG/cm^2)	Field (kG)	Gradient (kG/cm^3)	Field (kG)
-0.254023E-03	-0.635057E-02	-0.673151E-03	-0.841438E-01
0.143424E-02	0.358559E-01	0.110241E-02	0.137801
-0.248686E-02	-0.621716E-01	-0.359389E-03	-0.449236E-01
-0.484701E-03	-0.121175E-01	0.180765E-03	0.225956E-01
0.304526E-02	0.761314E-01	-0.506326E-04	-0.632907E-02
0.152721E-02	0.381804E-01	0.854076E-04	0.106759E-01
0.147310E-02	0.368275E-01	-0.794297E-04	-0.992872E-02
-0.270373E-02	-0.675933E-01	-0.592687E-04	-0.740859E-02
-0.644784E-03	-0.161196E-01	0.615846E-04	0.769808E-02
0.187468E-02	0.468670E-01	-0.613805E-04	-0.767256E-02
		0.320186E-03	0.400233E-01
		-0.886455E-03	-0.110807
		0.287362E-02	0.359203
		-0.177754E-02	-0.222193
		-1.12893(n_3)	

Table 5.4: The FRSC third-order achromat: Field strengths of the sextupoles and the octupoles. The field index, n_3, is the third-order derivative of the field of the bending magnet over the bending radius (dimensionless).

quadrupole), the best-behaved first-order solution was found and the horizontal beam envelope is shown in Figure 5.6. The strengths of the quadrupoles are displayed in Table 5.5.

According to the analytical theory, four independent sextupoles are required to obtain a second-order achromat. However, because of the fact that for the first order, cell R is identical to cell F, a simplification is possible based on Brown's theory of second-order achromats[Brown79, Carey81]. As shown in Section 3.2, a second-order achromat can be achieved by placing two families of sextupoles in the dispersive region of every cell and correcting one second-order chromatic aberration in each transverse plane. This is the second reason why the forward cell is made symmetric itself and $(y|y)$ is cancelled instead of $(y|b)$ and $(b|y)$. Even though in principle a second-order

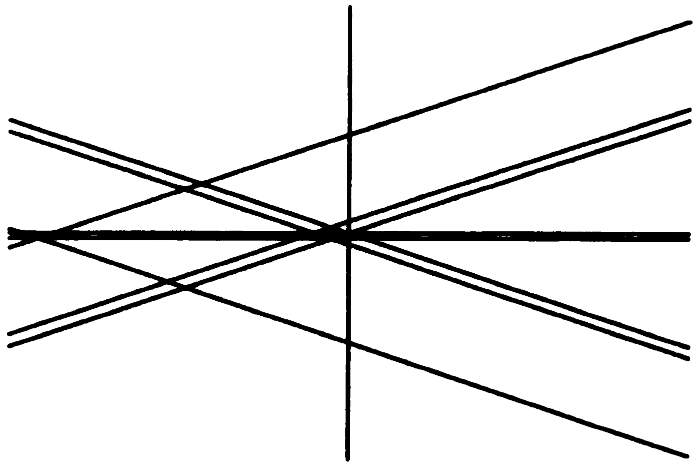

Figure 5.4: The FRSC third-order achromat: Remaining aberrations up to the eighth order (scale: 30 μm × 20 μm). The vertical line represents the final focal plane, where the deviations are around 10 μm.

achromat can be achieved with two sextupoles per cell (half ring), the constraint that the second half be the reverse of the first requires that the sextupoles be split into symmetrically excited pairs. Again, one more pair is inserted to kill the term $(t|\delta^2)$ (Table 5.6). Sixteen octupoles are introduced to correct all the remaining third-order aberrations, including $(t|\delta^3)$. The positions of some of the multipoles are carefully chosen to minimize the required field strengths, which results in a realistic setting (Table 5.8).

Since our goal is to make ESR a multi-pass time-of-flight mass spectrometer, its dynamic aperture has to be studied in detail. The dynamic aperture was obtained by tracking the system of a number of turns using its 8th-order one turn map. For particles of a momentum spread of $\pm 0.25\%$ to survive 100 turns, the horizontal and vertical apertures are roughly 1 π mm mrad.

The mass resolution was determined in a statistical way: First, a large number of particles (1000) with a certain mass deviation inside a certain phase space area are sent through the 8th-order one-turn map n times and the n-turn time-of-flight

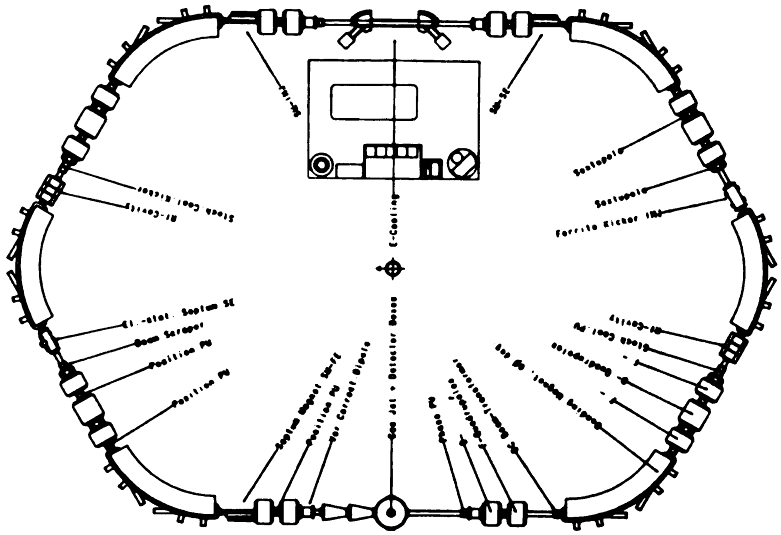

Figure 5.5: The FRFR third-order achromat: The original layout of the Experimental Storage Ring (ESR) at Darmstadt, Germany

of each particle is computed. Second, the random errors of the detector are taken into account, which in this design is chosen as 100 ps maximum, and the predicted mass deviation of each particle is calculated. Finally, the difference between the predicted and initial mass deviation of each particle is obtained and the n-turn mass resolution is determined by calculating the inverse of the average differences. The mass resolution's dependence on the number of turns is presented in Figure 5.8 .

5.3 A Fourth-Order Achromat - FRFR

A fourth-order achromat was also designed using the best solution obtained in Section 4.3. It is a storage ring with mirror symmetry. Since cells S and C are not allowed, the only choice is FRFR. The first half of the ring forms the forward cell (F) and the second half is the reversed cell (R). So, like the last example, the achromat is achieved after two turns. In the process of the first-order design, the four elements of $(x|x)$, $(x|\delta)$, $(y|b)$ and $(b|y)$ were fitted to zero. Since the first-order layout of each cell is symmetric about its own center, $(a|a)$ equals $(x|x)$, which implies that $(a|a)$

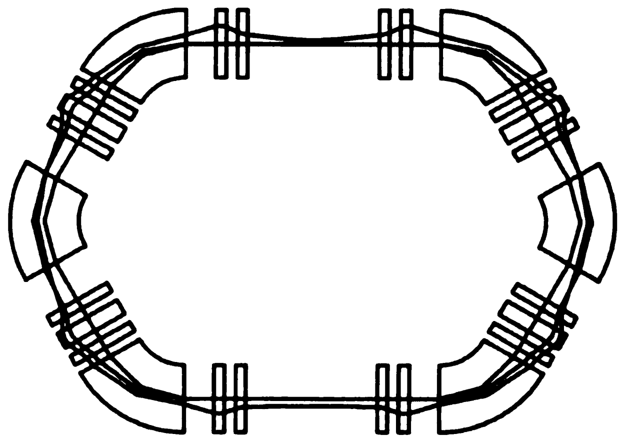

Figure 5.6: The FRFR third-order achromat: The first-order beam envelope of the horizontal (x-z) plane. The curve that does not coincide with the z-axis in the straight section is the β-function. The other curve is the dispersive ray. The circumference is 108.36 m; the emittance is 12.5π mm mrad; and the dispersion is 0.7%.

is also cancelled. Ten quadrupoles are used in each cell, which give five free knobs (Figure 5.9). Four of them are used for cancelling the four above terms with one left for optimizing the first-order solution.

There are four sextupoles in each cell to correct second-order aberrations, all of which have to be placed inside the dispersive region because all second-order aberrations left are chromatic ones. This is also true for the decapoles, but not for the octupoles. By placing the fifteen octupoles and fifteen decapoles more or less equidistant, a fourth-order achromat requiring very weak multipoles is achieved (Tables 5.9 and 5.10).

To study the influence of the remaining aberrations, the final spots of beams starting from a square lattice are plotted and shown in Fig. 5.10. It clearly shows that the achromat is reached at the end of the second turn, because up to the fourth order there are still significant distortions after one turn. It also shows that the higher-order remaining aberrations grow exponentially with the emittance.

Strengths of the Multipoles (Aperture 10 cm)			
Quadrupoles		Sextupoles	
Gradient (kG/cm)	Field (kG)	Gradient (kG/cm^2)	Field (kG)
-0.624520	-3.12260	-0.208441E-01	-0.521102
0.742699	3.71350	0.106122E-01	0.265306
-0.212000	-1.06000	-0.260702E-02	-0.651754E-01
0.389363	1.94682		
-0.323636	-1.61818		

Table 5.5: The FRFR third-order achromat: The field strengths of the quadrupoles and the sextupoles. Since they are excited symmetrically, only half of the multipoles are listed.

```
0.4551914E-14 0.2306961     0.0000000E+00 0.0000000E+00 0.0000000E+00 1000000
-4.334708     0.4163336E-14 0.0000000E+00 0.0000000E+00 0.0000000E+00 0100000
0.0000000E+00 0.0000000E+00-0.1443290E-14 0.5119279     0.0000000E+00 0010000
0.0000000E+00 0.0000000E+00 -1.953400    -0.1817990E-14 0.0000000E+00 0001000
0.0000000E+00 0.0000000E+00 0.0000000E+00 0.0000000E+00  1.000000     0000100
0.0000000E+00 0.0000000E+00 0.0000000E+00 0.0000000E+00 0.1679212E-14 0000010
0.0000000E+00 0.0000000E+00 0.0000000E+00 0.0000000E+00 -4.187160     0000001
------------------------------------------------------------------------------
```

Table 5.6: The FRFR third-order achromat: The first-order map of half the ring. (The zeroes are numbers smaller than 1E-15) The phase advances are $\mu_x = \mu_y = \pi/2$. Note that $(t|\delta)$ also vanishes

5.4 A Fifth-Order Achromat - FRFR

As a proof of principle, a circular fifth-order achromat is designed, which proves Theorem 4.8 and adds a case supporting Conjecture 4.1. The first-order layout should avoid large changes in the beta functions in order to minimize 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 Figure 5.11, which consists of sixteen FODO cells plus two dispersion correction sections, each of which includes two

1.000000	0.0000000E+00	0.0000000E+00	0.0000000E+00	0.0000000E+00	1000000
0.0000000E+00	1.000000	0.0000000E+00	0.0000000E+00	0.0000000E+00	0100000
0.0000000E+00	0.0000000E+00	1.000000	0.0000000E+00	0.0000000E+00	0010000
0.0000000E+00	0.0000000E+00	0.0000000E+00	1.000000	0.0000000E+00	0001000
0.0000000E+00	0.0000000E+00	0.0000000E+00	0.0000000E+00	1.000000	0000100
0.0000000E+00	0.0000000E+00	0.0000000E+00	0.0000000E+00	-16.74864	0000001
0.0000000E+00	0.0000000E+00	0.0000000E+00	0.0000000E+00	22.80066	0000011
0.0000000E+00	0.0000000E+00	0.0000000E+00	0.0000000E+00	-0.2394999	0000002
0.0000000E+00	0.0000000E+00	0.0000000E+00	0.0000000E+00	-23.27966	0000021
0.0000000E+00	0.0000000E+00	0.0000000E+00	0.0000000E+00	-3.467360	0000012
0.0000000E+00	0.0000000E+00	0.0000000E+00	0.0000000E+00	0.4997533	0000003

Table 5.7: The FRFR third-order achromat: The third-order two-turn map. (The zeroes are numbers smaller than 1E-8.) Note that time-of-flight terms, which depend on energy and mass, cannot be cancelled along with $(t|\delta^n)$ due to the fact that the magnetic field only distinguishes magnetic rigidity, which is a function of δ_k and δ_m.

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 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 5.11 shows that the beam travels around the ring in a very uniform manner, avoiding large ray excursions and beta functions. As described in the last example, a second-order achromat 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 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 aspects seem 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

Strengths of the Octupoles (Aperture 10 cm)			
Gradient (kG/cm^3)	Field (kG)	Gradient (kG/cm^3)	Field (kG)
-0.128089E-02	-0.160111	-0.882350E-02	-1.10294
0.180353E-02	0.225441	0.581219E-03	0.726523E-01
0.275862E-02	0.344828	0.394693E-02	0.493366
-0.429976E-02	-0.537470	0.706907E-02	0.883634
0.344373E-02	0.430466	0.453311E-02	0.566639
0.153106E-03	0.191382E-01	-0.114925E-01	-1.43656
-0.263437E-02	-0.329296	0.308289E-02	0.385361
0.904441E-02	1.13055	0.148141E-01	1.85176

Table 5.8: The FRFR third-order achromat: The field strengths of the octupoles.

Strengths of the Multipoles (Aperture 10cm)			
Quadrupoles		Sextupoles	
Gradient (kG/cm)	Field (kG)	Gradient (kG/cm^2)	Field (kG)
0.277788	1.38894	-0.114852E-02	-0.287130E-01
-0.273114	-1.36557	-0.184780E-03	-0.461950E-02
0.881770E-01	0.440885	0.101651E-02	0.254127E-01
-0.282065E-01	-0.141032	-0.941644E-03	-0.235411E-01
-0.361497E-01	-0.180748		

Table 5.9: The FRFR fourth-order achromat: The field strengths of the quads and the sextupoles. Only half of them are shown due to mirror symmetry.

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 weak multipole strengths for third-, fourth-, and fifth-order corrections. (Tables 5.12 and 5.13).

The 1000-turn dynamic aperture for both horizontal and vertical motions are studied using the 11th-order map. For particles of a momentum spread of $\pm 0.5\%$ to survive 1000 turns, the dynamic aperture is at least 100 π mm mrad both horizontally

Strengths of the Multipoles (Aperture 10cm)			
Octupoles		Decapoles	
Gradient (kG/cm^3)	Field (kG)	Gradient (kG/cm^4)	Field (kG)
0.865187E-04	0.108148E-01	-0.184214E-03	-0.115134
-0.730041E-04	-0.912551E-02	0.993688E-04	0.621055E-01
0.118558E-03	0.148198E-01	0.267539E-03	0.167212
-0.101019E-03	-0.126274E-01	0.353274E-03	0.220796
0.474184E-04	0.592730E-02	-0.577517E-03	-0.360948
-0.663923E-05	-0.829904E-03	-0.153650E-03	-0.960314E-01
-0.132363E-04	-0.165454E-02	0.764958E-03	0.478099
0.466092E-04	0.582615E-02	-0.433794E-03	-0.271121
-0.849342E-04	-0.106168E-01	0.405756E-03	0.253597
-0.515629E-05	-0.644536E-03	-0.843814E-03	-0.527384
0.280953E-04	0.351191E-02	0.538596E-03	0.336623
0.125900E-03	0.157375E-01	0.369327E-03	0.230830
-0.423299E-03	-0.529124E-01	-0.128477E-03	-0.802978E-01
0.245393E-03	0.306741E-01	-0.177871E-03	-0.111169
0.796716E-04	0.995896E-02	0.207354E-03	0.129596

Table 5.10: The FRFR fourth-order achromat: The field strengths of the octupoles and the decapoles. Here all the multipoles are very weak.

and vertically. It is much larger than the acceptance of the ring, which is about 30 $\sim$ 40 π mm mrad. As an example, Figure 5.12 shows the horizontal motion of on-energy particles up to 1000 turns. The absence of linear effects, as well as any aberration up to order five, leads to a behavior that is entirely determined by nonlinearities of order six and higher.

The time-of-flight energy resolution of this ring is determined in a statistical manner similar to the example discussed in Section 5.2 except that the one-turn map is of the ninth-order. The dependence of the resolution on the number of turns is presented in Figure 5.13 .

Strengths of the Multipoles (Aperture 10 cm)			
Quadrupoles		Sextupoles	
Gradient (kG/cm)	Field (kG)	Gradient (kG/cm^2)	Field (kG)
-0.162869	-0.814344	-0.718659E-03	-0.179665E-01
0.134119	0.670597	0.364420E-03	0.911050E-02
-0.131803	-0.659013		

Table 5.11: The FRFR fifth-order achromat: The field strengths of the quads and the sextupoles. Only half of them are shown due to mirror symmetry.

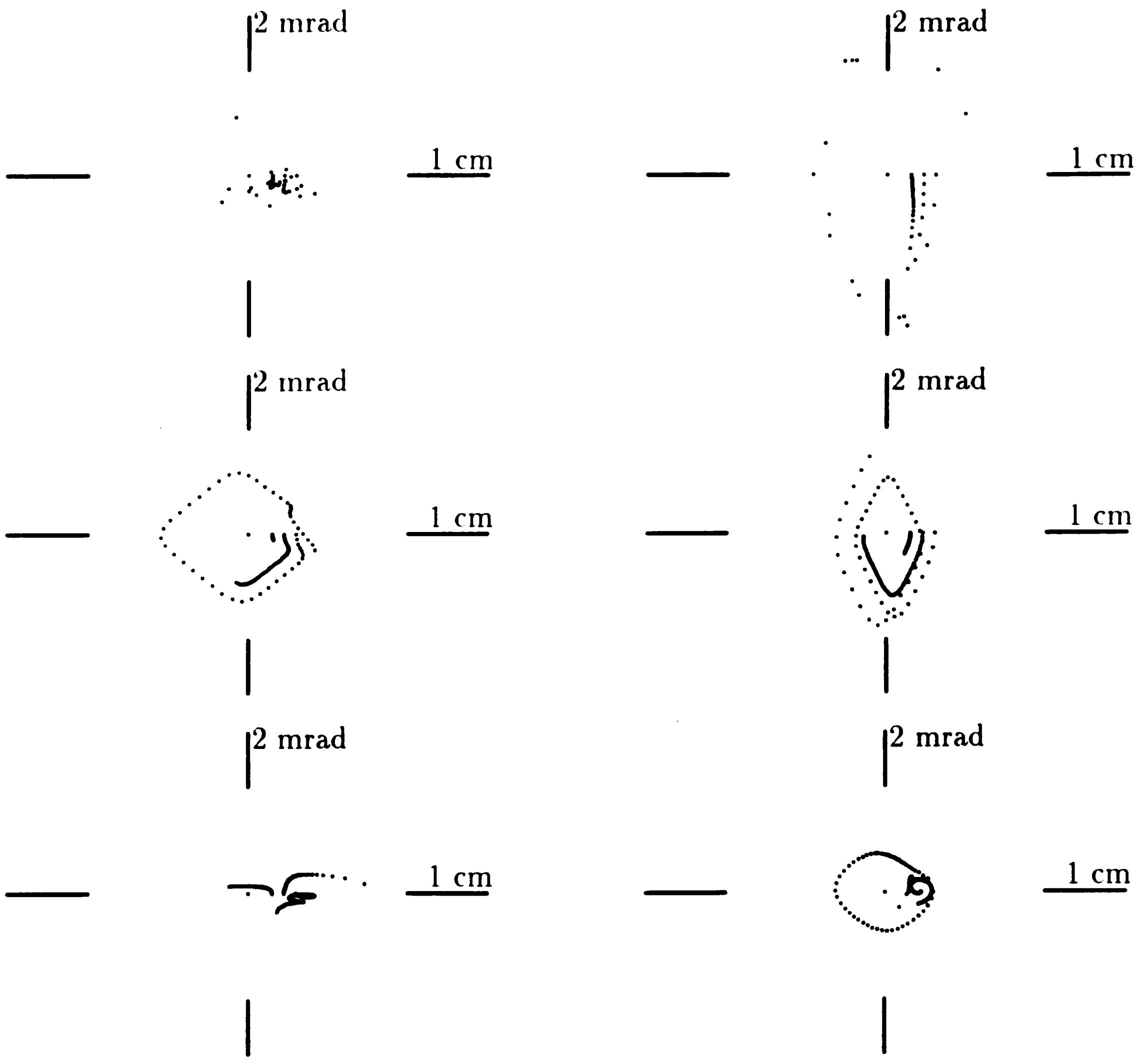

Figure 5.7: The FRFR third-order achromat: The 8th-order 1000-turn tracking picture. The left and right columns display those of x- and y-motion, respectively; the top, middle, and bottom rows show those of $\delta = -0.1\%$, 0, and 0.1%, respectively.

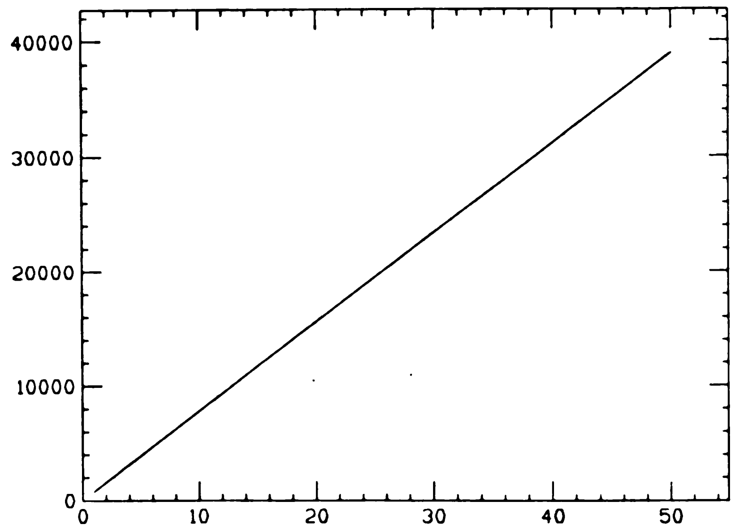

Figure 5.8: The FRFR third-order achromat: The multi-turn mass resolution as a function of the number of turns. Due to the small emittance, the resolution increases almost linearly with the number of turns.

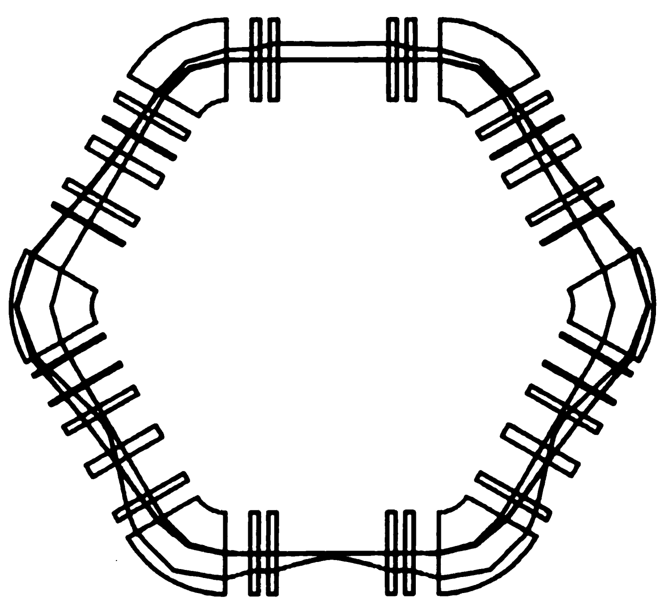

Figure 5.9: The FRFR fourth-order achromat: The layout, beam envelope and dispersive ray. The phase advances are $\mu_x = \mu_y = \pi/2$. The circumference is 147.35 m; the emittance is 20π mm mrad; and the dispersion is 0.6%.

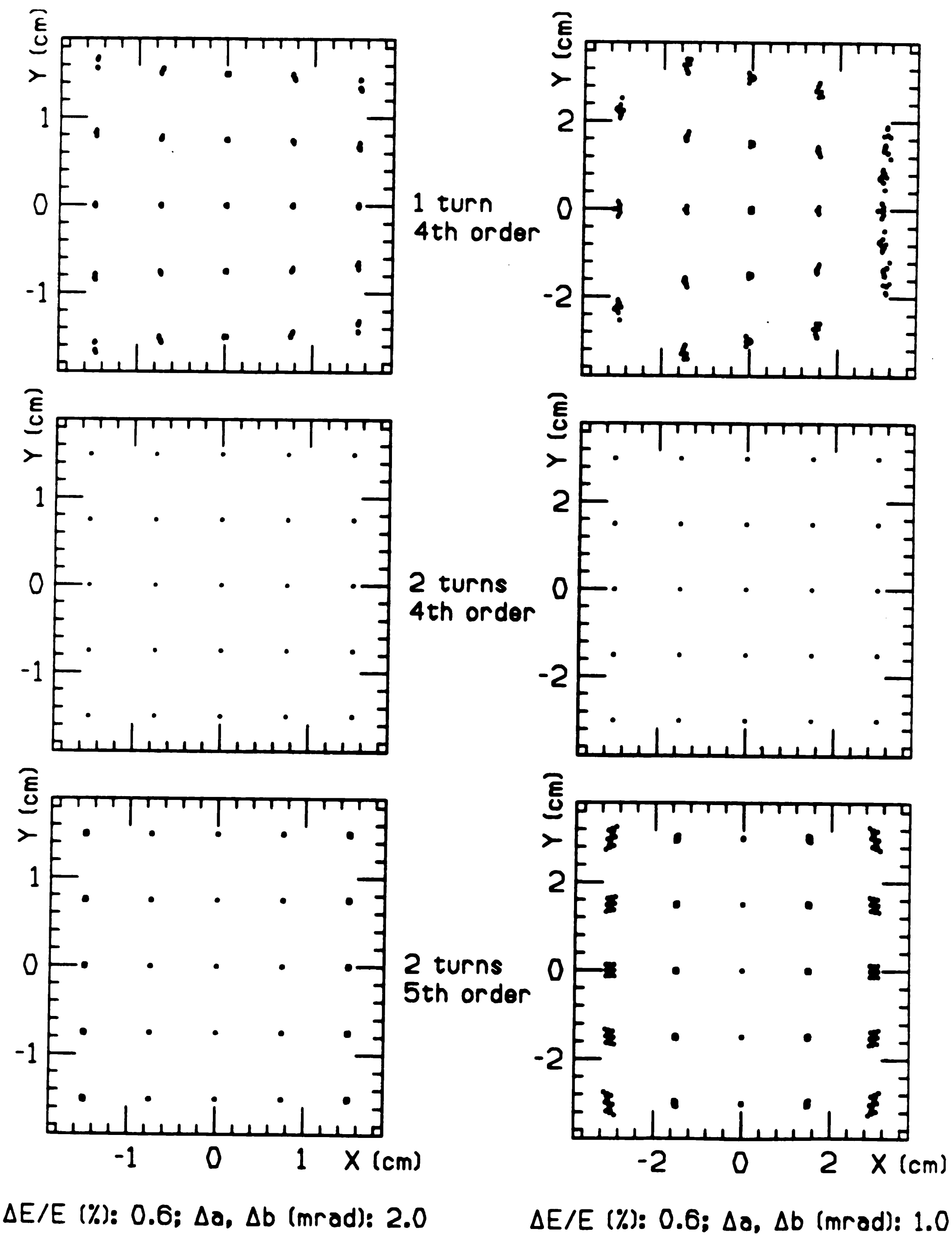

Figure 5.10: The FRFR fourth-order achromat: Beam spots of different emittances. The top two rows show that an achromat is reached after two turns. The bottom row shows the remaining higher-order aberrations.

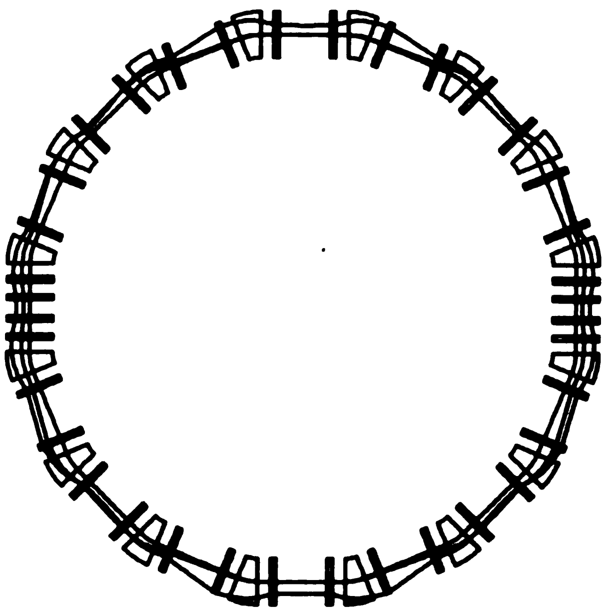

Figure 5.11: The FRFR fifth-order achromat: The layout, beam envelope and dispersive ray. The phase advances per cell are $\mu_x = \mu_y = \pi/2$. The circumference is 266.64 m; the emittance is 30π mm mrad; and the dispersion is 0.3%.

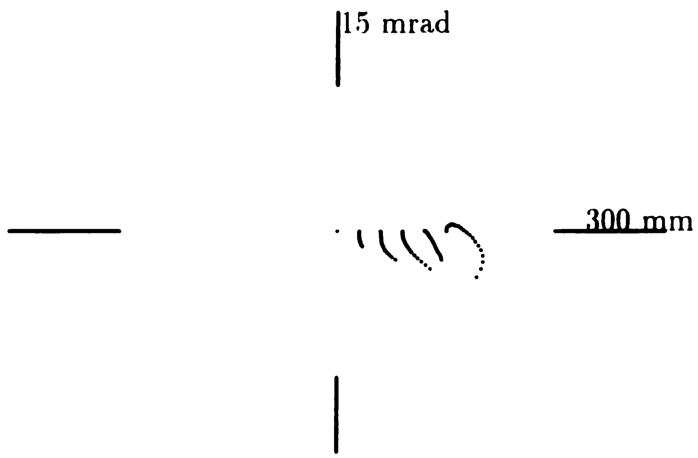

Figure 5.12: The FRFR fifth-order achromat: 1000-turn tracking of the x-a motion of on-energy particles.

Strengths of the Multipoles (Aperture 10 cm)			
Octupoles		Decapoles	
Gradient (kG/cm^3)	Field (kG)	Gradient (kG/cm^4)	Field (kG)
-0.996975E-06	-0.124622E-03	-0.391808E-06	-0.244880E-03
-0.246999E-05	-0.308749E-03	0.239260E-06	0.149538E-03
0.204723E-05	0.255903E-03	-0.346336E-07	-0.216460E-04
-0.135901E-05	-0.169876E-03	-0.413315E-07	-0.258322E-04
0.951498E-06	0.118937E-03	0.100518E-06	0.628240E-04
-0.228548E-04	-0.285685E-02	-0.501265E-07	-0.313291E-04
0.177119E-04	0.221399E-02	-0.953086E-07	-0.595678E-04
-0.158309E-04	-0.197886E-02	0.511256E-06	0.319535E-03
0.420261E-05	0.525326E-03	-0.305803E-07	-0.191127E-04
0.871498E-07	0.108937E-04	-0.775351E-07	-0.484594E-04
0.377365E-06	0.471706E-04	0.506782E-08	0.316738E-05
0.533332E-05	0.666665E-03	0.153783E-07	0.961144E-05
0.321821E-05	0.402276E-03	-0.152854E-07	-0.955335E-05
0.191867E-05	0.239833E-03	0.159598E-06	0.997489E-04
-0.130343E-05	-0.162929E-03	-0.317045E-06	-0.198153E-03

Table 5.12: 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.

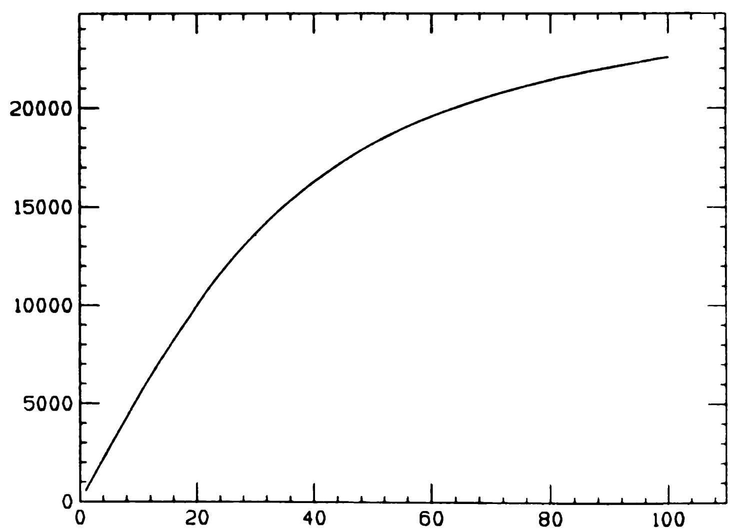

Figure 5.13: The FRFR fifth-order achromat: Resolution vs numbers of turns at the acceptance. The saturation comes from the accumulation of higher-order aberrations over turns.

Strengths of the Duodecapoles (Aperture 10 cm)			
Gradient (kG/cm^5)	Field (kG)	Gradient (kG/cm^5)	Field (kG)
0.260526E-06	0.162829E-03	0.143366E-06	0.896036E-04
-0.141949E-06	-0.887180E-04	0.111585E-06	0.697405E-04
-0.602391E-07	-0.376494E-04	-0.392296E-06	-0.245185E-03
0.115200E-06	0.720003E-04	0.426602E-06	0.266626E-03
-0.129574E-06	-0.809839E-04	-0.251765E-06	-0.157353E-03
0.167172E-06	0.104483E-03	0.101758E-06	0.635989E-04
-0.146698E-06	-0.916861E-04	-0.812971E-07	-0.508107E-04
0.109038E-07	0.681489E-05	0.113277E-06	0.707979E-04
-0.897166E-07	-0.560728E-04	-0.423092E-07	-0.264433E-04
0.905100E-07	0.565687E-04	-0.733480E-07	-0.458425E-04
0.422171E-07	0.263857E-04	0.173217E-07	0.108261E-04
-0.119032E-06	-0.743948E-04	0.970192E-07	0.606370E-04
0.812032E-07	0.507520E-04	0.745327E-07	0.465829E-04
-0.859254E-07	-0.537034E-04	-0.158631E-06	-0.991446E-04
0.143652E-06	0.897825E-04	0.230450E-06	0.144031E-03
-0.192421E-06	-0.120263E-03	-0.172798E-06	-0.107999E-03
0.231122E-06	0.144451E-03	0.923330E-07	0.577081E-04
-0.729862E-07	-0.456164E-04	0.126337E-06	0.789607E-04
-0.102382E-06	-0.639889E-04	-0.256941E-06	-0.160588E-03
-0.913997E-07	-0.571248E-04		

Table 5.13: 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.

Summary

The theoretical and experimental development of achromats is reviewed in detailed studies of Brown's second-order achromat theory and Dragt's third-order achromat theory. It is shown that Dragt's theory provides a complete proof of Brown's theory in a different way, thus demonstrating that Dragt's theory is more general.

The second- and third-order theories are extended to an arbitrary-order theory where detailed proofs of all conclusions are given. As opposed to repetition, this theory explores the role of mirror symmetry in building an achromatic system. It is shown that two- and three-cell systems are not the best choices for making achromats, because they require more conditions than all efficient 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 are the best choices for building arbitrary-order achromats. For four-cell systems, the best solution is found. 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; it is proved analytically up to the fourth order and computational results suggest that it is valid up to the sixth order. This is close to the best solution that can be obtained from this theory.

Four examples of achromats of the third, fourth, and fifth orders are presented. An "S"-shaped third-order achromat shows the possibility for use as a single-pass time-of-flight spectrometer. A circular third-order isochronous achromat can be used

as a multi-pass time-of-flight mass spectrometer for studying very short-lived nuclei. A fourth-order and a fifth-order achromats verify the analytical theory to the sixth order.

In conclusion, an understanding of arbitrary-order, mirror symmetrical achromats, particularly four-cell systems, has been developed.

Appendix A

Symplectic Properties of Matrices R and S

Theorem A.1 *R is antisymplectic and S is symplectic.*

Proof:

To be consistant with the definition of $\hat{J}$ in eq. (2.2), we have

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

on the basis of $I = (x, y, a, b)$. Therefore, we have

$$\begin{aligned} R\hat{J}R^t &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \\ &= \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \\ &= \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \begin{pmatrix} 0 & -I \\ -I & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} = -\hat{J}, \end{aligned}$$

which shows that R is antisymplectic.

Similarly, we have

$$S = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Now let us define $\sigma = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$. Since

$$\begin{aligned} S\hat{J}S^t &= \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \sigma & 0 \\ 0 & \sigma \end{pmatrix} \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} \sigma & 0 \\ 0 & \sigma \end{pmatrix} \\ &= \begin{pmatrix} \sigma & 0 \\ 0 & \sigma \end{pmatrix} \begin{pmatrix} 0 & \sigma \\ -\sigma & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & \sigma^2 \\ -\sigma^2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} = \hat{J}, \end{aligned}$$

S is symplectic.

Appendix B

The Proof of Equation (4.60)

Recall eq. (4.59), we have

$$\begin{cases} m_{11}^{(1)} m_{22}^{(1)} = -1 \\ m_{11}^{(3)} = -1 \end{cases}.$$

Next we show that eq. (4.60) holds, i.e.,

$$m_{11}^{(1)} = -m_{22}^{(1)}. \tag{B.1}$$

Since $m_{11}^{(1)}$ = -$m_{11}^{(2)}$ and $m_{22}^{(1)}$ = -$m_{22}^{(2)}$,

$$m_{11}^{(1)} m_{22}^{(1)} = -1 \quad \text{means that} \quad m_{11}^{(2)} m_{22}^{(2)} = -1. \tag{B.2}$$

To keep the generality, we have to consider the other cases because when the specific knowledge of the odering the cells are needed, as shown below, the permutational symmetry is broken. Therefore, the conclusion drown from eq. (B.2) is that only one out of the last three cells satisfies $m_{11}^{(i)} = m_{22}^{(i)} = -1$ ($i = 1, 2, 3$).

From eq. (4.70) and (4.73) in Section 4.2.4, the second cell has to be either R or C when $m_{11}^{(1)} m_{22}^{(1)} = -1$.

First let the second cell be R. From eq. (4.70), we have

$$M_1^{(1)} = L_1^{-1} R_1^{-1} \cdot L_1 = \begin{pmatrix} ad + bc & 2bd & 2b\eta' \\ -2ac & -(ad + bc) & -2a\eta' \\ 0 & 0 & 1 \end{pmatrix}.$$

Under the condition that $m_{12}^{(1)} = m_{21}^{(1)} = m_{15}^{(1)} = m_{25}^{(1)} = 0$, we have

$$M_1^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which means that $m_{11}^{(1)} = -m_{22}^{(1)}$. If the second cell is C, the result is the same as for the case of R.

Similarly, the second cell has to be either F or S when $m_{11}^{(1)} = m_{22}^{(1)} = -1$.

When it is F, we have

$$L_1 = M_1^{(1)} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

This implies that the third cell must satisfy $m_{11}^{(2)} m_{22}^{(2)} = -1$. Thus for R and C, $M_1^{(2)}$ is

a) $$M_1^{(2)} = L_1^{-1} R_1^{-1} \cdot L_1 L_1 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{and}$$

b) $$M_1^{(2)} = L_1^{-1} R_1^{-1} S_1^{-1} \cdot L_1 L_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

respectively. This shows that $m_{11}^{(2)} = -m_{22}^{(2)}$. If the second cell is S, the result is the same as the case above, except that $L_1 = \hat{I}$.

Bibliography

[Arnold89] V. I. Arnold, *Mathematical Methods of Classical Mechanics* (GTM 60), (Springer-Verlag, New York, 1991), Chap 8.

[Berz85] H. Wollnik and M. Berz, *Nucl. Instr. Methods* **A238,** 127 (1985).

[Berz87a] M. Berz, H. C. Hoffmann, and H. Wollnik, *Nucl. Instr. Methods* **A258,** 402 (1987).

[Berz89] M. Berz, *Particle Accelerators,* **24,** 109 (1989).

[Berz90] M. Berz, *Nucl. Instr. Methods* **A298,** 426 (1990).

[Berz92a] M. Berz, *Proc. Int. Workshop Nonlinear Problems in Accelerator Phys.*, Berlin (1992). M. Berz, S. Martin and K. Ziegler (Eds.), IOP Conf. Ser. **131,** 77 1993.

[Berz92b] M. Berz, *Physics of Particle Accelerators,* AIP Conf. Proc. **249,** 143 (1992).

[Berz92c] M. Berz, *Proc. XIII Inter. Par. Accel. Conf.,* Dubna (1992). Technical Report, Michigan State University, MSUCL-807, (1991), East Lansing.

[Berz93] M. Berz, Technical Report, Michigan State University, MSUCL-869 (1993), East Lansing.

[Brown73] K. L. Brown, D. C. Darey, Ch. Islin, and F. Rothacker, Technical Reports, SLAC-91 (1973 rev.), NAL-91, CERN 80-04 (1980).

[Brown79] K. L. Brown, *IEEE Transactions on Nuclear Science,* NS-**26**, No. 3. p3490 (1979).

[Brown82a] K. L. Brown and R. V. Servranckx, *Nucl. Instr. Methods* **203,** 73 (1982).

[Brown82b] K. L. Brown, Technical Report, SLAC-75 (1982).

[Brown85] K. L. Brown and R. V. Servranckx, *IEEE Transactions on Nuclear Science,* NS-**32**, No. 5, p2288 (1985).

[Brown87a] G. E. Fischer, K. L. Brown, F. Bulos etc., *Proc. 1987 Particle Accelerator Conference,* p139.

[Brown87b] J. J. Murray, K. L. Brown and T. Fieguth, *Proc. 1987 Particle Accelerator Conference*, p1331.

[Carey81] David C. Carey, *Nucl. Instr. Methods* **189,** 365 (1981).

[Cour58] E. D. Courant and H. S. Snyder, *Ann. Phys.,* **3** 1 (1958).

[Dragt76] A. J. Dragt and J. M. Finn, *J. Math. Phys.,* **17**, 2215 (1976).

[Dragt81] A. J. Dragt, *Physics of High Energy Particle Accelerators* (Fermi Summer School 1981), AIP Conf. Proc. **87,** 147 (1992).

[Dragt85] A. J. Dragt, L. M. Healy, F. Neri, and R. D. Ryne, *IEEE Transactions on Nuclear Science*, NS-**32**, No. 5, p2311 (1985).

[Dragt87] A. J. Dragt, *Nucl. Instr. Methods* **A258,** 339 (1987).

[Dragt79] A. J. Dragt and J. M. Finn, *J. Math. Phys.,* **21**, 2649 (1976).

[Dugan91] G. Dugan, *Proc. 1989 Particle Accelerator Conference,* p426.

[Dynkin62] E. B. Dynkin, *Amer. Math. Soc. Transl. (1),* **9**, 470 (1962).

[Finley91] D. A. Finley, *Proc. 1993 Particle Accelerator Conference,* p3721.

[Flanz89a] J. B. Flanz, *Proc. 1989 Particle Accelerator Conference,* p1349.

[Flanz89b] 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, *Proc. 1989 Particle Accelerator Conference*, p1349.

[Forest89] E. Forest, M. Berz, J. Irwin, *Particle Accelerators,* **24,** 91 (1989).

[Franz87] B. Franzke, *Nucl. Instr. Methods* **B24/25,** 18 (1987).

[Garren75] A. A. Garren and J. W. Eusebio, LBL Report, UCID-10153 (1975).

[Gold80] H. Goldstein, *Classical Mechanics,* (Addison-Wesley, Reading, MA, 1980), Chap 9.

[Gordon59] M. M. Gordon and T. A. Welton, Oak Ridge National Lab, ORNL Report, 2765, (1959).

[Iselin85] Ch. Iselin, Technical Report, CERN LEP-TH/85-15 (1985).

[Iselin88] Ch. Iselin and J. Niederer, Technical Report, CERN LEP-TH/88-38 (1988).

[Jack87] A. Jackson, *Particle Accelerators,* **22,** 111 (1987).

[Kowal85] S. Kowalski and H. Enge, Technique Report, MIT, 156 (1985), Cambridge.

[Law32] E. O. Lawrence, M. S. Livingston, *Phys. Rev.,* **40**, 19 (1932).

[Litv93] V. N. Litvinenko and J. M. J. Madey, *Proc. 1993 Particle Accelerator Conference*, p1442.

[Matsuo76] T. Matsuo, H. Matsuda, Y. Fujita, and H. Wollnik, *Mass Spectrosc. Jpn.*, **24**, 19 (1976).

[Neri91] F. Neri, in *Proc. Workshop on High Order Effects.* M. Berz and J. McIntyre (Eds.), Technical Report, Michigan State University, MSUCL-767, 131 (1991), East Lansing.

[Vara84] V. S. Varadarajan, *Lie Groups, Lie Algebras, and Their Representations* (GTM 102), (Springer-Verlag, New York, 1984), Chap 2.

[Meyer91] K. R. Meyer and G. R. Hall, *Introduction to Hamiltonian Dynamical Systems and the N-body Problem* (AMS 90), (Springer-Verlag, New York, 1991).

[Murphy92] J. Murphy, *Synchrotron Light Source Data Book*, Technical Report, Bnl 42333, (1992).

[Schwar94] Bertram Schwarzschild, *Physics Today,* **47**, No. 7, 22 (1994).

[Serv83] J. C. Bergstrom, H. S. Caplan, and R. V. Servranckx, *IEEE Transactions on Nuclear Science,* NS-**30**, No. 4, p3226 (1983).

[Serv85] R. Servranckx, K. L. Brown, Lindsay Schachinger, and David Douglas, Technical Report, SLAC-Report-285 (1985).

[Serv89] R. V. Servranckx, U. Wienands, and M. K. Craddok, *Proc. 1989 Particle Accelerator Conference,* p1355.

[Tal87] L. Schachinger and R. Talman, *Particle Accelerators,* **22,** 35 (1987).

[Thiess72] H. A. Thiessen and M. Klein, in *Proc. of the IV Int. Conf. on Magnet Technology,* (1972) (unpublished).

[Wan92] W. Wan, E. Goldmann and M. Berz, *Proc. Int. Workshop Nonlinear Problems in Accelerator Phys.*, Berlin (1992). M. Berz, S. Martin and K. Ziegler (Eds.), IOP Conf. Ser. **131,** 201 1993.

[Wan93a] W. Wan, E. Goldmann and M. Berz, *Proc. Comput. Accel. Phys. Conf.* (1993). AIP Conf. Proc. **297,** 143 (1994).

[Wan93b] W. Wan and M. Berz, *Proc. 1993 Particle Accelerator Conference,* p155.

[Wied93] Helmut Wiedemann, *Particle Accelerator Physics,* (Springer-Verlag, New York, 1993), Chap 1-3.

[Woll87a] H. Wollnik, J. Brezina, and M. Berz, *Nucl. Instr. Methods* **A258,** 408 (1987).

[Woll87b] H. Wollnik, *Nucl. Instr. Methods* **B26,** 267 (1987).

[Woll87c] H. Wollnik, *Nucl. Instr. Methods* **A258,** 289 (1987).

[Wood83] M. D. Woodley, M. J. Lee, J. Jäger, and A. S. King, *IEEE Transactions on Nuclear Science*, NS-**30**, No. 4, p2367 (1983).

[Wouter85] J. M. Wouters, D. J. Vieira, H. Wollnik, H. A. Enge, S. Kowalski and K. L. Brown, *Nucl. Instr. Methods* **A240,** 77 (1985).

[Wouter87] J. M. Wouters, H. Wollnik and D. J. Vieira, *Nucl. Instr. Methods* **A258,** 331 (1987).

[Wu93a] Y. Wu, V. N. Litvinenko, and J. M. J. Madey, *Proc. 1993 Particle Accelerator Conference*, p218.

[Wu93b] Y. Wu, V. N. Litvinenko, E. Forest, and J. M. J. Madey, *Nucl. Instr. Methods* **A331,** 287 (1993).

[Yan90] Y. Yan and C. Yan, Technical Report, SSCL-300 (1990).

[Zeijts92] J. B. J. van Zeijts and F. Neri, *Proc. Int. Workshop Nonlinear Problems in Accelerator Phys.*, Berlin (1992). M. Berz, S. Martin and K. Ziegler (Eds.), IOP Conf. Ser. **131,** 135 1993.