IOP Publishing

A Practical Introduction to Beam Physics and Particle Accelerators (Third Edition)

Santiago Bernal

Chapter 1

Rays, matrices, and transfer maps

To understand particle accelerators and their components, we need to study and solve the equations of motion of charged particles in external electromagnetic fields. In their most general form, these equations contain terms involving all orders of the coordinates and velocities of the particles [1, 2]. However, if the particles are confined to small distances from a *reference orbit* and have small angles relative to the same orbit, a *linear equation of motion* resembling a simple pendulum (or mass-on-a-spring) equation results. This is the basis for the *paraxial approximation* presented in section 1.1. The 'spring constant' in the equation of motion can be generalized to include a piecewise *focusing function* whose form depends on the types of external fields present. In addition to the paraxial approximation, the focusing elements of particle accelerators can be treated as *thin* or *thick lenses*, depending on whether the focal length is large or comparable to the physical axial extent of the external fields. We discuss thin and thick lenses in sections 1.2 and 1.3. For both thin and thick lenses, a *matrix* description of particle motion is both convenient and powerful. Finally, in section 1.4 we mention the existence of powerful methods based on Lie operators that address higher-order particle tracking.

1.1 Paraxial approximation

The linearized equation of motion for a charged particle in the presence of external fields represented by the focusing function $\kappa(s)$ is

$$x''(s) + \kappa(s)x(s) = 0, \tag{1.1}$$

where primes indicate derivatives with respect to the axial distance 's' measured along the reference trajectory. For now, we consider just one of the transverse coordinates for particle motion, the horizontal component (or the radial component in a local curvilinear coordinate system).

doi:10.1088/978-0-7503-4039-7ch1

Note that equation (1.1), known as Hill's equation, is a homogeneous, second-order ordinary differential equation (ODE) with a coefficient, the focusing function, which is not generally constant. The solution depends on the initial conditions $x(0)$, $x'(0)$, and on the form of $\kappa(s)$, but not on any higher orders of either $x(0)$ or $x'(0)$. This latter condition is the essence of the *paraxial approximation*; it is the equivalent of using θ instead of $\sin\theta$ in Snell's law of standard optics.

Here and in subsequent chapters we will discuss the solutions of Hill's equation and some of its non-homogeneous versions in some detail; additional considerations can be found in many textbooks (e.g., [2]). The general solution of equation (1.1) can be written as a superposition of *cosine-like* $C(s)$ and *sine-like* $S(s)$ functions

$$\begin{aligned} x(s) &= x(0)C(s) + x'(0)S(s), \\ x'(s) &= x(0)C'(s) + x'(0)S'(s), \end{aligned} \tag{1.2}$$

such that $C(0) = 1$, $S(0) = 0$, $C'(0) = 0$, $S'(0) = 1$. In matrix form, equation (1.2) reads

$$\begin{bmatrix} x(s) \\ x'(s) \end{bmatrix} = \begin{bmatrix} C(s) & S(s) \\ C'(s) & S'(s) \end{bmatrix} \begin{bmatrix} x(0) \\ x'(0) \end{bmatrix}. \tag{1.3}$$

In a more compact form we can write

$$[x(s),\, x'(s)]^T = \boldsymbol{R}[x(0),\, x'(0)]^T. \tag{1.4}$$

The space spanned by the coordinates $x(s)$, $x'(s)$ for an ensemble of particles is called *trace space* or, more commonly, *phase space*. Strictly speaking, this phase space would be a projection on the $(x,\ x')$ plane of full phase space. In the simple situations in which no radiation, acceleration, or particle losses occur, the area in phase space is conserved under transformations represented by $\boldsymbol{R}$. The latter statement is known as *Liouville's theorem*, which we will revisit in chapter 4. As a consequence of this theorem, the determinant of $\boldsymbol{R}$ is equal to one, i.e., $|\,\boldsymbol{R}\,| = 1$.

The simplest case of a focusing function is the 'point' lens, a mathematical construct whereby $\kappa(s) = \delta(s)/f$. 'δ' represents the Dirac delta function and 'f' is a constant. By an integration of equation (1.1) with this focusing function, it is easy to see that 'f' represents the focal length of a zero-length lens located at $s = 0$. Figure 1.1 illustrates the zero-length or point lens.

By studying equation (1.2) for a point lens, we can identify the *principal trajectories* or rays illustrated in figure 1.1: for the first ray, $x'_1(0) = 0$, $x_1(0) = x_{10}$, leading to $x'_1(s) = -x_{10}(1/f)$, $x_1(s) = x_{10}(1 - s/f)$ for $s > 0$; for the second ray, $x_2(0) = 0$, $x'_2(0) = x'_{20}$, leading to $x_2(s) = x'_{20}s$. Thus, in this case, the matrix $\boldsymbol{R}$ is:

$$\boldsymbol{R}_{\mathbf{f}} = \begin{bmatrix} 1 & 0 \\ -1/f & 1 \end{bmatrix}. \tag{1.5}$$

Note that the first column in $\boldsymbol{R}$ represents ray 1 (coefficients multiplying initial conditions), while the second column represents ray 2.

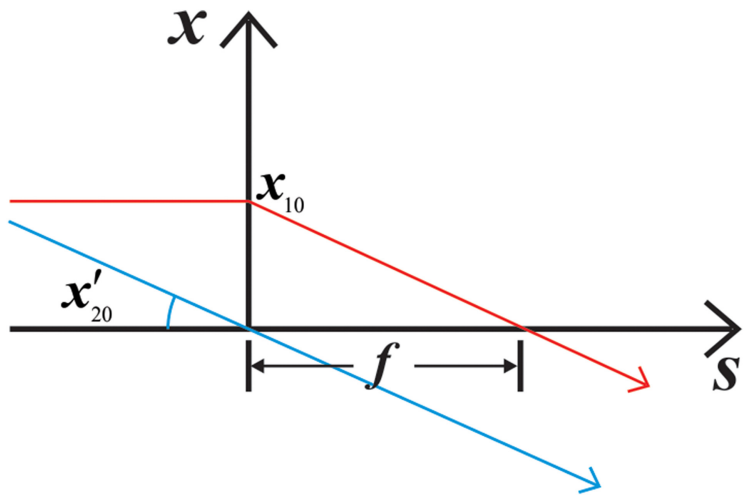

Figure 1.1. A 'point' lens at $s = 0$. The red and blue rays represent the principal trajectories or rays.

An equally important transfer matrix represents the ray transformation over a drift of length L:

$$\boldsymbol{R}_{\mathrm{L}} = \begin{bmatrix} 1 & L \\ 0 & 1 \end{bmatrix}. \tag{1.6}$$

The matrices in equations (1.5) and (1.6) are the building blocks for transforming rays in linear systems with multiple lenses.

1.2 Thin lenses

If the extent of a real lens, as quantified by the *effective length* l_{eff} of the profile of the focusing function $\kappa(s)$, is much smaller than the focal length, then the lens is considered to be *thin*. (The point lens introduced above is an extreme example of a thin lens, because it remains so for *any* $f > 0$.) A useful construction that applies especially to lenses whose profiles $\kappa(s)$ are not dominated by *fringe fields*, i.e., extended 'tails,' is the *hard-edge model*. Figure 1.2 shows an example of an actual axial profile (of a magnetic field) related to $\kappa(s)$ and the corresponding hard-edge model.

The effective length of the hard-edge model is defined by

$$l_{\mathrm{eff}} \equiv \frac{1}{\kappa_{\mathrm{peak}}} \int_{-s_1}^{+s_1} \kappa(s) ds, \qquad \frac{1}{f} = l_{\mathrm{eff}} \kappa_{\mathrm{peak}}. \tag{1.7}$$

Therefore, the condition for a thin lens is $\kappa_{\mathrm{peak}} \ll 1/l_{\mathrm{eff}}{}^2$, or $f \gg l_{\mathrm{eff}}$. According to this criterion, a very strong lens can have a focal length comparable to its effective length, in which case the lens is *not* thin; or else the lens can be weak and have a very long focal length, satisfying the condition of a thin lens.

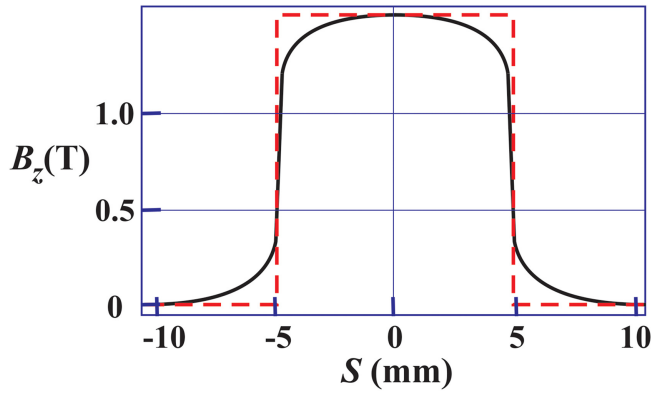

Figure 1.2. Actual axial profile of a magnetic field (related to the focusing function) and corresponding hard-edge model (red broken curve).

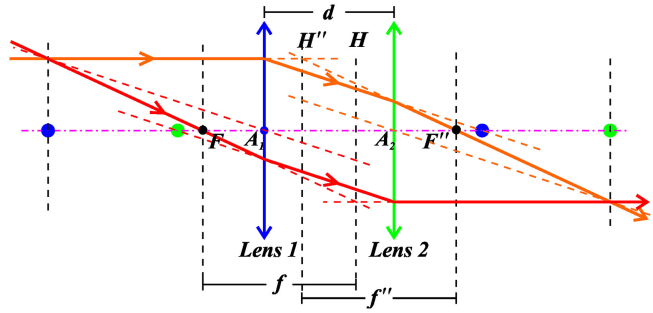

Figure 1.3. The system of two positive thin lenses is equivalent to a thick lens that has the principal planes H and H'' shown. Note that the focal lengths of the system are measured from the principal planes.

Ray tracing for a thin lens proceeds just as for the point lens (figure 1.1); so, for example, the exit x-coordinate of a given ray is unchanged after traversing the lens. Thus, the basic transfer matrix that connects ray coordinates on both sides of the lens is the same as that given in equation (1.5).

1.3 Thick lenses

Thick lenses in charged-particle optics can be treated just as thick lenses in standard optics. Concepts such as *principal planes*, *nodal points*, etc. also apply to magneto-static or electrostatic lenses [3]. The principal planes of a thick lens or a system of lenses are two hypothetical planes that connect locations of unit magnification, i.e., an object located on one plane is imaged on the second plane with the same size. Figure 1.3 illustrates the primary and secondary principal planes, H and H'', for a system of two focusing thin lenses separated by a distance d that have the same focal lengths. Ray tracing is done using the method of parallel rays as discussed in the standard optics book by Jenkins and White [4].

We can use equations (1.5) and (1.6) to find the matrix for the lens system shown in figure 1.3:

$$\boldsymbol{R} = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1/f_2 & 1 \end{bmatrix} \begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1/f_1 & 1 \end{bmatrix}, \tag{1.8}$$

or,

$$\boldsymbol{R} = \begin{bmatrix} 1 - d/f_1 & d \\ -\left(\dfrac{1}{f_1} + \dfrac{1}{f_2} - \dfrac{d}{f_1 f_2}\right) & 1 - d/f_2 \end{bmatrix}. \tag{1.9}$$

By multiplying $\boldsymbol{R}$ with a drift matrix (length l_1) on the left and another (length l_2) on the right, we obtain a new matrix $\overline{\boldsymbol{R}}$

$$\overline{\boldsymbol{R}} = \begin{bmatrix} \overline{R}_{11} & \overline{R}_{12} \\ \overline{R}_{21} & \overline{R}_{22} \end{bmatrix} = \begin{bmatrix} R_{11} + l_2 R_{21} & R_{12} + l_1 R_{11} + l_2 R_{22} + l_1 l_2 R_{21} \\ R_{21} & R_{22} + l_1 R_{21} \end{bmatrix}. \tag{1.10}$$

Now, by setting $\overline{R}_{11} = 0$ we obtain the distance from the second lens to the second focal point of the system, F'' [5]:

$$l_2 = -\frac{R_{11}}{R_{21}} = \left(1 - \frac{d}{f_1}\right)\left(\frac{1}{f_1} + \frac{1}{f_2} - \frac{d}{f_1 f_2}\right)^{-1}. \tag{1.11}$$

The second factor on the right of equation (1.11) gives the inverse effective *object and image focal lengths* f, f'' of the lens combination [4]:

$$\frac{1}{f} = \left(\frac{1}{f_1} + \frac{1}{f_2} - \frac{d}{f_1 f_2}\right) = \frac{1}{f''}. \tag{1.12}$$

The focal lengths f and f'' are *measured from the corresponding principal planes*, as shown in the figures.

Now assume that the second lens is divergent but that the two lenses have the same strength, i.e., the same focal lengths in magnitude: $f_1 = -f_2 > 0$. Then, from equation (1.12), the effective focal length is simply

$$f = \frac{f_1^2}{d} = \frac{f_2^2}{d}. \tag{1.13}$$

Furthermore, from equations (1.11) and (1.12), the second focal point of the system, F'', as measured from the second lens is

$$l_2 = f_1\left(\frac{f_1}{d} - 1\right). \tag{1.14}$$

This quantity is positive, indicating a convergent *equivalent thick lens* for the two-lens system if $f_1 > d$. Thus, we arrive at the important result that the combination of two lenses, one convergent (positive) and the other divergent (negative) that have the

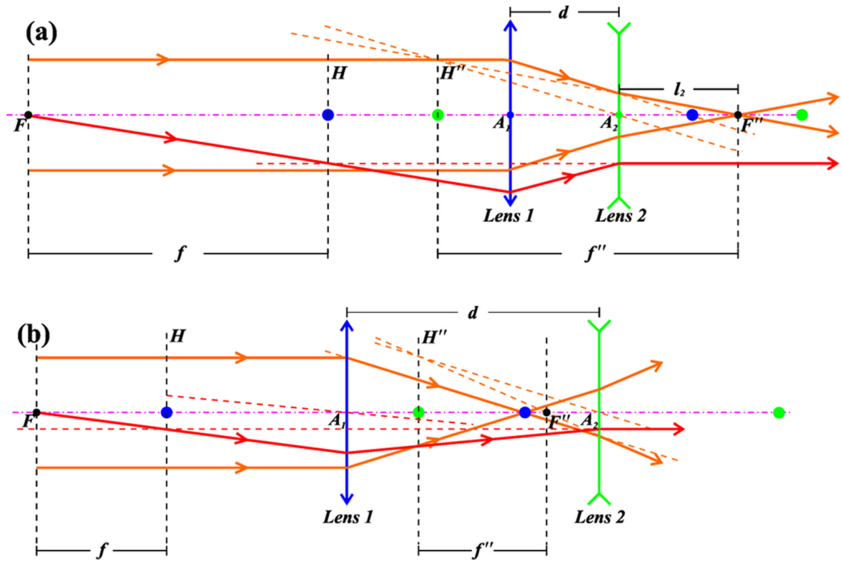

Figure 1.4. Combination of convergent (lens 1) and divergent (lens 2) lenses with the same strengths. In (a) the spacing d is *smaller* than the individual focal lengths ($f_1 = -f_2 > 0$); in (b) the spacing d is *larger* than the individual focal lengths. Note how the principal planes H and H'' are found using the parallel ray technique.

same strengths, leads to net focusing if their separation is smaller than the positive focal length f. This is relevant to the strong focusing scheme discussed in chapter 3 for particle accelerators.

Figure 1.4 shows the action of the focusing–defocusing lens combination for two values of their separation d. The geometrical construction that finds the principal planes uses the method of parallel rays mentioned above; examination of the figures may reveal the basic ideas. Note that the principal planes lie outside the space between the lenses in figure 1.4(a), unlike figure 1.3(a), for a pair of converging lenses with the same separation as in figure 1.4(a).

Additional relationships between the elements of the matrix in equation (1.10) and the important optical properties of a general lens system are discussed in an article by Halbach [6] and in appendix D of [5].

1.4 Transfer maps

The matrix formalism presented in the previous sections can be extended to higher orders. The complexity of the equations of motion to third order can be appreciated in Wiedemann's explicit treatment in his chapter 3 [2]; classic particle accelerator matrix software such as TRANSPORT [7], on the other hand, can perform only up to second-order calculations for most lattice elements. To third order, the ith component of the *six-dimensional* vector $\boldsymbol{\eta} = (x, x', y, y', z, \delta)^T$ (δ is the fractional momentum error—see section 6.5) is transformed between beamline planes '1' (exit) and '0' (input) according to [7]:

$$\eta_i(1) = \sum_j R_{ij}\eta_j(0) + \sum_{j,k} T_{ijk}\eta_j(0)\eta_k(0) + \sum_{j,k,l} U_{ijkl}\eta_j(0)\eta_k(0)\eta_l(0). \tag{1.15}$$

Most calculations in TRANSPORT employ up to the second-order matrix $\boldsymbol{T}$. To perform second- and higher-order calculations, a more elegant and powerful (but also more demanding—from a mathematical viewpoint) approach is provided by Lie algebraic methods. The latter are based on the Lie algebra structure of Hamiltonian mechanics, as briefly discussed in appendix A and in far more detail by Dragt in [8]. Essentially, a beamline is described by a transfer *map* consisting of the product of the linear-transformation matrix $\boldsymbol{R}$ and the exponentials of *Lie operators* containing polynomials f_m [9]. The polynomials f_3, f_4, for example, describe second- and third-order effects, respectively.

Furthermore, since standard light optics can be shown to have a Hamiltonian structure [10], the same Lie algebra methods can be applied to both particle trajectories and light rays. In classical mechanics, particle trajectories obey the *principle of least action* [11]. Light rays, on the other hand, travel through optical systems such that their paths take the least times. This is known as *Fermat's principle* in geometrical optics [4].

To conclude this chapter, we present an example of non-paraxial imaging. We have used the MARYLIE code to image the letters in 'UMER' through a system of four focusing–defocusing quadrupole doublets; additional details are discussed in appendix A and in the MARYLIE manual [12]. Figure 1.5(a) shows the object to be imaged by a unit-magnification system employing 20 TeV protons and superconducting magnets. Figure 1.5(b) displays the image obtained by tracing 2052 rays. The calculated transfer map differs slightly from a 6×6 unit matrix, which is enough to result in the blurring seen in figure 1.5(b). Most of the aberrations are caused by the quadrupole fringe fields; effects due to third-order distortion (see the letters 'U' and 'R'), astigmatism, and field curvature are present. Starting with a much smaller object and/or turning off the fringe fields considerably reduces the aberrations.

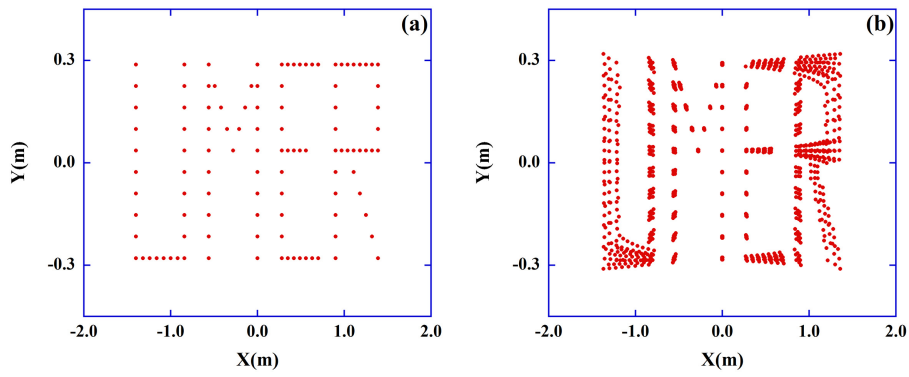

Figure 1.5. (a) The object to be imaged by four superconducting quadrupole doublets employing 20 TeV protons. (b) Image computed using the MARYLIE code; the effects of third-order aberrations are seen (see also appendix A and [12]).

1.5 Computer resources

Matrix multiplication, both symbolic and numeric, can be readily implemented in either *Mathcad* or *Matlab* or their freeware counterparts *SMath Studio* and *Octave* (see appendix A). Many standalone optics design programs are also available, but most are either too expensive or intended for professional use. Some software of this type is available as packages for codes such as *Mathematica* (e.g., Optics Lab).

We found the free program *Optgeo*, written by Jean-Marie Biansan, to be instructive and relatively easy to use (see appendix A for download information). The ray tracing in figures 1.3 and 1.4 can be reproduced without too much effort. We include two *Optgeo* files in this book's website. Alternatively, for *Mathcad* users we recommend a very well designed program for 2D optics called *2D Optical Ray Tracer*, which is written by Valter Kiisk of the University of Tartu (Estonia) and is available online for free. We include in our website a version with input that closely reproduces the results shown in figure 1.4(a). It is important to keep the lens thicknesses small relative to their focal lengths for comparison with figure 1.4, but we can also easily explore thick lens combinations. Figure 1.6 illustrates typical output.

For *Matlab*, there is at least one free app, *LensLab*, which can be used to reproduce the ray tracing of the lens doublets in this chapter. Another program that we found useful is *Ray Optics* for Android devices; there is a 'Pro' version that adds details of the calculations. *RayLab*, a free app for Apple's IOS devices, looks very professional and rather challenging as an educational tool for geometrical optics.

Thin magnetostatic lenses can be implemented in several of the most popular codes for particle accelerator design. (In fact, in some cases thin lenses *are* required for some operations, such as tracking in the code MAD.) Our website includes three example files for the program WINAGILE (see appendix A). The first file illustrates the implementation of one thin quadrupole lens; the other two files are examples of two thin quadrupole combinations. The files illustrate the concepts of this chapter and

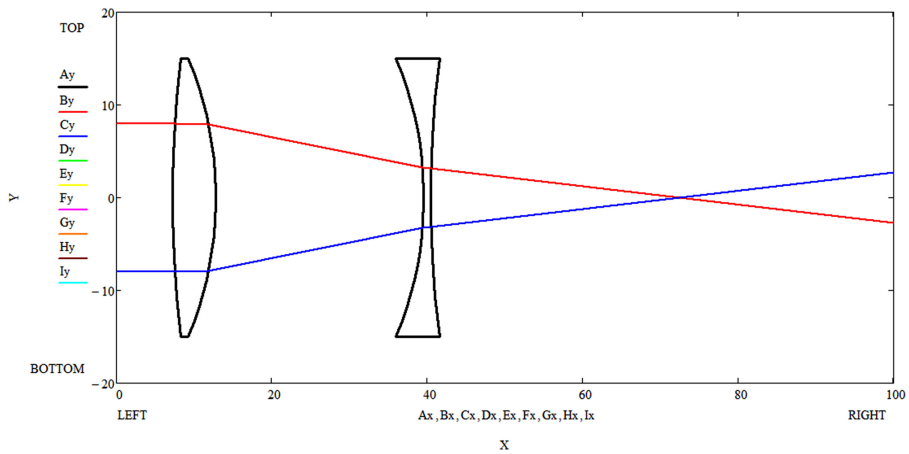

Figure 1.6. Graphics produced by the 2D optics *Mathcad* program *2D Optical Ray Tracer* by Valter Kiisk. Distances are in mm.

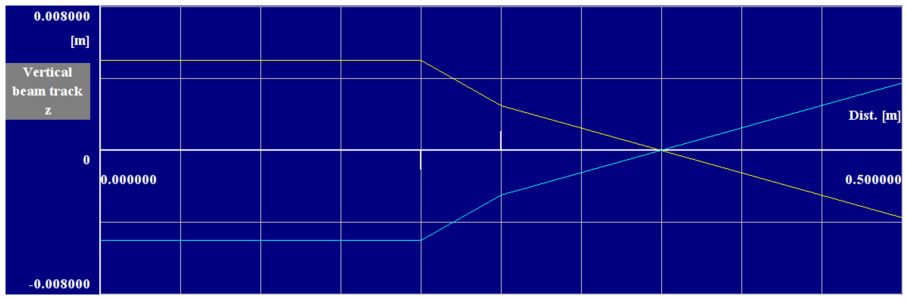

Figure 1.7. Trajectory tracking in WINAGILE for a combination of two thin quadrupoles of opposite strengths. The quadrupoles are located at $s = 20$ cm and 25 cm, while their individual focal lengths are $f_1 = -f_2 = 10$ cm. See also chapters 2 and 3 for additional important considerations.

serve as a preamble to quadrupole focusing (chapter 2) and to alternating gradient, or strong focusing (chapter 3). In appendix A, we indicate how to download, install, and run WINAGILE. Figure 1.7 is a screenshot from WINAGILE that shows trajectory tracking with a two thin quadrupole combination.

The example of non-paraxial imaging illustrated in figure 1.5 relies on the MARYLIE code [12] briefly described in section 1.4. Additional details pertaining to this code and the example can be found in appendix A.

References

[1] Steffen K G 1965 *High Energy Beam Optics,* (New York: Interscience Publishers/Wiley)
[2] Wiedemann H 2007 *Particle Accelerator Physics* 3rd edn (Berlin: Springer)
[3] Wollnik H 1987 *Optics of Charged Particles* (New York: Academic)
[4] Jenkins F A and White H E 1981 *Fundamentals of Optics* 4th ed (New York, NY: McGraw-Hill) 5
[5] Conte M and MacKay W W 1991 *An Introduction to the Physics of Particle Accelerators* (Singapore: World Scientific)
[6] Halbach K 1964 Matrix Representation of Gaussian Optics *Am. J. Phys.* **32** 90–108
[7] Carey D C, Brown K L and Rothacker F 1995 Third-order transport—a computer program for designing charged particle beam transport systems, SLAC-R-95-462, fermilab-Pub-95/068, UC-414
[8] Dragt, A J 1982 Lectures on nonlinear orbit dynamics *AIP Conf. Proc.* 87
[9] Dragt A J 2013 An overview of lie methods for accelerator physics *Proc. of PAC* (Pasadena, CA) 1134
[10] Dragt A J 1982 Lie algebraic theory of geometrical optics and optical aberrations *J. Opt. Soc. Am.* **72** 372–379
[11] Goldstein H 1980 *Classical Mechanics* 2nd edn (Reading, MA: Addison-Wesley)
[12] Dragt A *et al* 2003 *MARYLIE 3.0 Users' Manual: A Program for Charged Particle Beam Transport based on Lie Algebraic Methods* (College Park: University of Maryland) http://www.physics.umd.edu/dsat/docs/MaryLieMan.pdf