Successive approximations for charged particle motion

G.H. Hoffstaetter*

Deutsches Elektronen–Synchrotron DESY, Notkestrasse 85, 22607 Hamburg, Germany

Received 29 June 1999; received in revised form 9 November 1999

Dedicated to Professor Harald Rose on the occasion of his 65th birthday

Abstract

Single-particle dynamics in electron microscopes, ion or electron lithographic instruments, particle accelerators, and particle spectrographs is described by weakly nonlinear ordinary differential equations. Therefore, the linear part of the equation of motion is usually solved and the nonlinear effects are then found in successive order by iteration methods. When synchrotron radiation is not important, the equation can be derived from a Hamiltonian or a Lagrangian. The Hamiltonian nature can lead to simplified computations of particle transport through an optical device when a suitable computational method is used. H. Rose and his school have contributed to these techniques by developing and intensively using the eikonal method [1–3]. Many ingenious microscopic and lithographic devices were found by Rose and his group due to the simple structure of this method [4–6]. The particle optical eikonal method is either derived by propagating the electron wave or by the principle of Maupertuis for time-independent fields. Maybe because of the time-dependent fields which are often required, in the area of accelerator physics the eikonal method has never become popular, although Lagrange methods had been used sometimes already in early days [7]. In this area classical Hamilitonian dynamics is usually used to compute nonlinear particle motion. Here the author will therefore derive the eikonal method from a Hamiltonian quite familiar to the accelerator physics community and reformulate it in a simplifying way. With the event of high-energy polarized electron beams [8] and plans for high-energy proton beams [9], nonlinear effects in spin motion have become important in high-energy accelerators. The author introduces a successive approximation for the nonlinear effects in the coupled spin and orbit motion of charged particles which resembles some of the simplifications resulting from the eikonal method for the pure orbit motion. © 2000 Elsevier Science B.V. All rights reserved.

1. Introduction

The well-known Lagrange variational principle requires

$$\delta \int \mathscr{L} \mathrm{d}t = \delta \int [\tilde{\boldsymbol{p}} \cdot \dot{\boldsymbol{q}} - \mathscr{H}] \mathrm{d}t = 0, \tag{1.1}$$

with the Lagrangian $\mathscr{L}$, Hamiltonian $\mathscr{H}$, and generalized momenta $\tilde{\boldsymbol{p}}$ and coordinates $\boldsymbol{q}$.

* Corresponding author. Tel.: + 49-408998-3430; fax: 49-408998-8904.
E-mail address: georg.hoffstaetter@desy.de (G.H. Hoffstaetter)

In this principle all variations of $q(t)$ are allowed and therefore the Euler–Lagrange equations of motion hold,

$$\frac{\mathrm{d}}{\mathrm{d}t}\partial_{\dot{q}}\mathcal{L}(q,\dot{q},t) = \partial_q\mathcal{L}(q,\dot{q},t). \tag{1.2}$$

For relativistic single-particle motion the Lagrangian is

$$\mathcal{L} = -mc\sqrt{c^2 - \dot{r}^2} + e\dot{r}\cdot A - e\Phi, \tag{1.3}$$

where the position $r(q)$ is a function of the generalized coordinates q. The Jacobian matrix $\underline{r}$ of this function can be written in the form $\underline{r} = (\partial_q r^{\mathrm{T}})^{\mathrm{T}}$ and has the elements $r_{ij} = \partial_{qj}r_i$. In this efficient notation r^{T} is the transpose of the 3×1 matrix r. The Jacobian matrix of the function $\dot{r}(\dot{q})$ is also $\underline{r}$ since $\dot{r} = \sum_{i=1}^3 \dot{q}_i\partial_{qi}r = \underline{r}\,\dot{q}$.

The generalized momentum is $\tilde{p} = \partial_q\mathcal{L} = \underline{r}^{\mathrm{T}}(m\gamma\dot{r} + eA)$ and the variational principle can thus be written as

$$\delta\int[\tilde{p}^{\mathrm{T}}\underline{r}^{-1}\dot{r} - \mathcal{H}]\,\mathrm{d}t = \delta\int[(m\gamma\dot{r} + eA)^{\mathrm{T}}\underline{r}\underline{r}^{-1}\dot{r} - \mathcal{H}]\,\mathrm{d}t = \delta\int[m\gamma v^2 + eA^{\mathrm{T}}\dot{r} - \mathcal{H}]\,\mathrm{d}t. \tag{1.4}$$

If only variations $\delta_{\mathcal{H}=E}$ are considered which keep the total energy $\mathcal{H} = E$ constant, the variational principle becomes

$$\delta_{\mathcal{H}=E}\int[\tilde{p}\cdot\dot{q} - \mathcal{H}]\,\mathrm{d}t = \delta_{\mathcal{H}=E}\int\tilde{p}\cdot\mathrm{d}q = \delta_{\mathcal{H}=E}\int[m\gamma v^2 + eA^{\mathrm{T}}\dot{r}]\,\mathrm{d}t = 0. \tag{1.5}$$

The variational principle for constant total energy is called the principle of Maupertuis. However, in Eq. (1.5) it does not lead to Euler–Lagrange equations of motion, since not all variations are allowed.

A particle optical device usually has an optical axis or some design curve along which a particle beam should travel. This design curve $R(l)$ is parameterized by a variable l and the position of a particle in the vicinity of the design curve has coordinates x and y along the unit vectors e_x and e_y in a plane perpendicular to this curve. This coordinate system is shown in Fig. 1. The third coordinate vector $e_1 = \mathrm{d}R/\mathrm{d}l$ is tangential to the design curve and the curvature vector is $\kappa = -\mathrm{d}e_1/\mathrm{d}l$.

The unit vectors e_x and e_y in the usual Frenet–Serret comoving coordinate system rotate with the torsion of the design curve. If this rotation is wound back, the equations of motion do not contain the torsion of the design curve. The position and the velocity are

$$r = xe_x + ye_y + R(l), \quad \dot{r} = \dot{x}e_x + \dot{y}e_y + h\dot{l}e_l, \tag{1.6}$$

with $h = 1 + x\kappa_x + y\kappa_y$. This method is described in Refs. [3,10] and is mentioned here since design curves with torsion are becoming important when considering particle motion in helical wigglers, undulators, and wavelength shifters [11], and for polarized particle motion in helical dipole Siberian Snakes [12].

The variational principle Eq. (1.5) for the three generalized coordinates $x(t)$, $y(t)$, and $l(t)$ can now be written for the two generalized coordinates $x(l)$ and $y(l)$. This has the following two advantages: (a) The particle trajectory along the design curve is usually more important than the particle position at a time t, and (b) Whereas $\delta_{\mathcal{H}=E}$ does not allow for all variations of the three coordinates, the total energy can be conserved for all variations of the two coordinates x and y by choosing for each position r the appropriate momentum with $m\gamma v = \sqrt{(E - e\Phi(r))^2/c^2 - (mc)^2}$. We obtain from Eq. (1.5):

$$\delta_{\mathcal{H}=E}\int\tilde{p}\cdot\mathrm{d}q = \delta\int\left[m\gamma v^2\frac{\mathrm{d}t}{\mathrm{d}l} + evA\cdot\frac{\mathrm{d}r}{\mathrm{d}l}\right]\mathrm{d}l = 0 \tag{1.7}$$
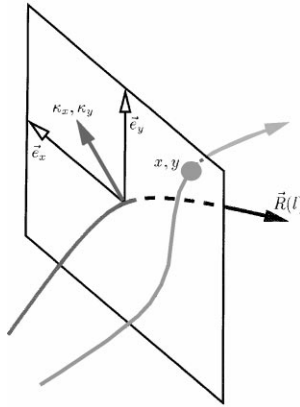
Fig. 1. Curvatures κ_x, κ_y of the design curve and generalized coordinates x, y, and l.

with $\mathrm{d}\mathbf{r}/\mathrm{d}l = x'\mathbf{e}_x + y'\mathbf{e}_y + h\mathbf{e}_l$ and $\mathrm{d}t/\mathrm{d}l = |\mathrm{d}\mathbf{r}/\mathrm{d}l|/v$. Since all variations are allowed, the integrand is a very simple new Lagrangian

$$\tilde{L} = m\gamma v\sqrt{x'^2 + y'^2 + h^2} + e(x'A_x + y'A_y + hA_l) \tag{1.8}$$

which leads to Euler–Lagrange equations of motion

$$\tilde{p}_x = \partial_{x'}\tilde{L}, \quad \tilde{p}'_x = \partial_x\tilde{L}, \tag{1.9}$$

$$\tilde{p}_y = \partial_{y'}\tilde{L}, \quad \tilde{p}'_y = \partial_y\tilde{L}. \tag{1.10}$$

The integral $\int_0^l \tilde{L}(\tilde{l})\,\mathrm{d}\tilde{l}$ is called the eikonal.

Since the Hamiltonian formulation is very common in the area of accelerator physics, we will show how the eikonal can be derived from a Hamiltonian formulation.

The equations of motion for the three generalized coordinates $x(t)$, $y(t)$, and $l(t)$ can be obtained from the Hamiltonian

$$\mathscr{H} = e\Phi + \sqrt{m^2c^2 + (\tilde{p}_x - \mathrm{e}A_x)^2 + (\tilde{p}_y - \mathrm{e}A_y)^2 + (\tilde{p}_l/h - \mathrm{e}A_l)^2}. \tag{1.11}$$

In the case of time-independent fields, $\mathscr{H}$ is the conserved total energy E and there are only five independent variables, rather than six. Note that the velocity dependent or non-holonomic [13] boundary condition $\mathscr{H}(\mathbf{p}(\mathbf{q},\dot{\mathbf{q}}),\mathbf{q},t) = E$ cannot be included in the Lagrange formalism directly. But in the Hamilton formalism this can be done. Furthermore, a switch of independent variable from t to l can easily be done in the Hamiltonian formulation. The Lagrange formulation is therefore abandoned (too easily, as will be shown later). In the variational condition

$$\delta \int [\dot{x}\tilde{p}_x + \dot{y}\tilde{p}_y + \dot{l}\tilde{p}_l - \mathscr{H}]\,\mathrm{d}t = 0, \tag{1.12}$$

one can change to the independent variable l as follows:

$$\delta \int [x'\tilde{p}_x + y'\tilde{p}_y + (-t')\mathscr{H} - (-\tilde{p}_l)]\,\mathrm{d}l = 0. \tag{1.13}$$

The six canonical coordinates are now x, $\tilde{p}_x$, y, $\tilde{p}_y$, $-t$, and $\mathcal{H}$, and the new Hamiltonian is given by $H = -\tilde{p}_l$ which has to be expressed as a function of the six coordinates [14,15],

$$\tilde{H} = -h[eA_l + \sqrt{(\mathcal{H} - e\Phi)^2 - (mc^2)^2 - (\tilde{p}_x - eA_x)^2 - (\tilde{p}_y - eA_y)^2}]. \tag{1.14}$$

In the Hamiltonion formalism it is simple to take advantage of the fact that the total energy is conserved for time-independent fields; $\mathcal{H}' = \partial_t \tilde{H} = 0$ leads to $\mathcal{H} = E$. Then from the six coordinates only the first four have to be considered, leading to the Lagrangian

$$\tilde{L} = x'\tilde{p}_x + y'\tilde{p}_y - \tilde{H}. \tag{1.15}$$

From $x' = \partial_{\tilde{p}x}\tilde{H} = (h/\sqrt{})(\tilde{p}_x - eA_x)$, $y' = \partial_{\tilde{p}_y}\tilde{H} = (h/\sqrt{})(\tilde{p}_y - eA_y)$ where $\sqrt{}$ is the square root in $\tilde{H}$, one obtains

$$\sqrt{} = h\sqrt{\frac{(E - e\Phi)^2 - (mc^2)^2}{x'^2 + y'^2 + h^2}} = m\gamma v \frac{h}{\sqrt{x'^2 + y'^2 + h^2}}, \tag{1.16}$$

$$\tilde{p}_x = \frac{\sqrt{}}{h}x' + eA_x, \quad \tilde{p}_y = \frac{\sqrt{}}{h}y' + eA_y, \tag{1.17}$$

$$\tilde{L} = m\gamma v\sqrt{x'^2 + y'^2 + h^2} + e(x'A_x + y'A_y + hA_l) \tag{1.18}$$

for $m\gamma v = \sqrt{(E - e\Phi)^2/c^2 - (mc)^2}$. The very simple Lagrangian $\tilde{L}$ agrees with the integrand (1.8) of the eikonal.

In the following, it will be shown how the Hamiltonian and the Lagrangian equations of motion for the particle trajectory $q(l)$ can be solved in an iterative way. We write a general equation of motion for a coordinate vector z in the form

$$z' = f^1(z, l) + f^{\geq 2}(z, l), \tag{1.19}$$

where we assume that $z = 0$ is a solution of the differential equation. Furthermore, we assume z to be small and let f^1 be linear in the coordinates. We assume that the nonlinear part of the equation of motion can be expanded in a Taylor series $f^{\geq 2}$. The linearized equation of motion is solved by a trajectory $z_1(l) = \underline{M}(l)z_i$ which depends linearly on the initial coordinates. For the transport matrix $\underline{M}(l)$ we therefore have

$$\underline{M}'z_i = \underline{f}^1\underline{M}z_i \tag{1.20}$$

for all coordinate vectors z_i; $\underline{f}^1$ being the Jacobian matrix of f^1.

One can write every solution of Eq. (1.19) as $z(l) = \underline{M}(l)\zeta(l)$, leading to the equation of motion

$$\underline{M}'\zeta + \underline{M}\zeta' = \underline{f}^1\underline{M}\zeta + f^{\geq 2}(z). \tag{1.21}$$

The Taylor coefficients of $\zeta(z_i, l)$ with respect to the initial coordinates $z_i = \zeta(0)$ are called aberration coefficients. With equation Eq. (1.20), one obtains

$$z(l) = \underline{M}(l)\left\{z_i + \int_0^l \underline{M}^{-1}(\tilde{l})f^{\geq 2}(z(\tilde{l}))\right\}d\tilde{l}. \tag{1.22}$$

Now, we assume that the general solution $z(z_i, l)$ can be expanded in a power series with respect to the initial coordinates. Then symbolizing the jth-order Taylor polynomial with $[\ldots]_j$, we write the orders up to

j as $z_j = [z(z_i, l)]_j$, i.e. we use lower indices to describe the order of z_i. The upper index in f describes the order in z, which is in turn a nonlinear function of z_i. When z_{n-1} is known, one can iterate the expansion up to order n with Eq. (1.22), since

$$z_n = \underline{M}(l)\left\{z_i + \int_0^l \underline{M}^{-1}(\tilde{l})[f^{\geqslant 2}(z_{n-1})]_n\right\}d\tilde{l}. \tag{1.23}$$

The zeroth order of the expansion with respect to the coordinates must vanish, which means that the trajectory $q = 0$ must satisfy the equation of motion for some momentum $p(l)$. Additionally, we require that the vector potential on the design curve is gauged to zero. This can always be achieved. The canonical momentum p then also vanishes for the trajectory $q = 0$. It then follows that the Hamiltonian and the Lagrangian have no components linear in the coordinates and momenta. When computing trajectories through a particle optical device, it is customary to normalize the momenta to the initial design momentum $p_0 = p(0)$. The following two-dimensional generalized coordinates are therefore used:

$$q = \begin{pmatrix} x \\ y \end{pmatrix}, \quad p = \begin{pmatrix} \tilde{p}_x/p_0 \\ \tilde{p}_y/p_0 \end{pmatrix}, \quad L(q, q', l) = \tilde{L}/p_0,$$

$$p = \partial_{q'}L, \quad p' = \partial_q L, \quad H(q, p, l) = \tilde{H}/p_0, \quad q' = \partial_p H, \quad p' = -\partial_q H. \tag{1.24}$$

The Euler–Lagrange equations lead to the second-order differential equations $(d/dl)\partial_{q'}L = \partial_q L$ for the two-dimensional vector q.

2. Successive approximation in terms of Hamiltonians

In the Hamilton formalism, one obtains first-order equations of motion for the four-dimensional vector $z^T = (q_1, q_2, p_1, p_2)$. With the antisymmetric matrix $\underline{J}$ one can write the equation of motion as

$$\underline{J} = \begin{pmatrix} \mathbf{0}_2 & \mathbf{1}_2 \\ -\mathbf{1}_2 & \mathbf{0}_2 \end{pmatrix}, \quad z' = \underline{J}\partial_z H, \tag{2.1}$$

with the 2×2 identity and zero matrixes $\mathbf{1}_2$ and $\mathbf{0}_2$. This structure implies special symmetries for the transport maps m of particle optics. These maps describe how the final phase-space coordinates $z_f = m(z_i)$ of a particle, after flying through an optical device, are related to the initial coordinates z_i. These maps are often weakly nonlinear and can be expanded in a Taylor expansion. The Hamiltonian nature implies that the Jacobian $\underline{m} = (\partial_z m^T)^T$ of any transport map $m(z)$ is symplectic [13], meaning that

$$\underline{m}\underline{J}\underline{m}^T = \underline{J}. \tag{2.2}$$

For the successive approximations we separate the equation of motion into its linear and nonlinear part,

$$z' = \underline{J}\partial_z(H^2 + H^{\geqslant 3}). \tag{2.3}$$

After we have solved for the linear transport matrix $z_1 = \underline{M}z_i$, we can iterate by Eq. (1.23) which takes the form

$$z_n = \underline{M}\left\{z_i + \int_0^l \underline{M}^{-1}[\underline{J}\partial_z H^{\geqslant 3}(z_{n-1})]_n \, d\tilde{l}\right\}. \tag{2.4}$$

With the relation $\underline{m}^{-1}\underline{J} = \underline{J}\underline{m}^T$ from Eq. (2.2) this can be written as

$$z_n = \underline{M}\left\{z_i + \underline{J}\int_0^l \underline{M}^T[\partial_z H^{\geqslant 3}(z_{n-1})]_n \, d\tilde{l}\right\}. \tag{2.5}$$

The corresponding equation for the aberrations $\zeta_n = \underline{M}^{-1}z_n$ becomes

$$\zeta_n = z_i + \underline{J}\int_0^l [\partial_\zeta H^{\geqslant 3}(\underline{M}\zeta_{n-1})]_n \, d\tilde{l}. \tag{2.6}$$

This form of the iteration equation is quite simple. However, since Hamiltonian (1.14) is a complicated function, the evaluation of the four integrals can become very cumbersome.

3. Successive approximation in terms of Lagrangians

In Ref. [3] Rose used a variational principle to derive a successive approximation to nonlinear motion based on the eikonal. This method iterates position q and momentum p in their nonlinear dependence on the initial position q_i and momentum p_i. Knowing the order $n-1$ dependence q_{n-1} and p_{n-1}, one has to compute q'_{n-1} by differentiation of q_{n-1} or by inversion of $p = \partial_{q'} L(q, q', l)$. Then the eikonal can be evaluated to compute the order n dependence q_n and p_n. In general, it can be cumbersome to compute q'_{n-1} and therefore here we derive a new version of the eikonal method, which iterates directly q'_n rather than the momentum.

In deriving the simple form of Eq. (2.6), advantage has only been taken of the symplectic first-order transfer matrix. We therefore wish to exploit this advantage again by working with new coordinates which are identical with the canonical q and p up to first order so that the new coordinates lead to the same first-order transport matrix $\underline{M}$. To first order, one obtains

$$p = \partial_{q'} L = \frac{p(s)}{p_0}q' + \frac{e}{p_0}\begin{pmatrix} A_x^1 \\ A_y^1 \end{pmatrix} + \mathcal{O}^2(q, q'), \tag{3.1}$$

where $p(s)$ is the momentum of a particle traveling on the design curve $q = 0$, and the upper index 1 specifies the part of the vector potential linear in x and y. We therefore work with the coordinates

$$Q = \begin{pmatrix} q \\ u \end{pmatrix} = \begin{pmatrix} q \\ \dfrac{p(s)}{p_0}q' + \dfrac{e}{p_0}\begin{pmatrix} A_y^1 \\ A_x^1 \end{pmatrix} \end{pmatrix}. \tag{3.2}$$

Moreover, it can be shown [3,16–18] that the contribution from the vector potential can be gauged to vanish whenever there is no longitudinal magnetic field $B_0 e_l$ on the design curve. Then if one investigates trajectories which start with momentum p_0 in a region free of such a field, we have the simple relation $u_i = q'_i$.

By splitting the Lagrangian into its second-order and its higher-order part, the equation of motion becomes

$$Q' = \begin{pmatrix} q' \\ u' \end{pmatrix} = \begin{pmatrix} \dfrac{p_0}{p(s)}u - \dfrac{e}{p(s)}\binom{A_y^1}{A_x^1} \\ \partial_q L^2 \end{pmatrix} - \begin{pmatrix} 0 \\ \dfrac{d}{d\tilde{l}}\partial_{q'}L^{\geqslant 3} - \partial_q L^{\geqslant 3} \end{pmatrix}. \tag{3.3}$$

After having solved the linearized equation of motion, we obtain with Eq. (1.23)

$$\mathbf{Q} = \underline{\mathbf{M}}\left\{\mathbf{Q}_i - \int_0^l \underline{\mathbf{M}}^{-1}\begin{pmatrix} 0 \\ \dfrac{\mathrm{d}}{\mathrm{d}\tilde{l}}\partial_{q'}L^{\geqslant 3} - \partial_q' L^{\geqslant 3} \end{pmatrix}\mathrm{d}\tilde{l}\right\}$$

$$= \underline{\mathbf{M}}\left\{\mathbf{Q}_i + \underline{\mathbf{J}}\int_0^l \underline{\mathbf{M}}^{\mathrm{T}}\begin{pmatrix} \dfrac{\mathrm{d}}{\mathrm{d}\tilde{l}}\partial_{q'}L^{\geqslant 3} - \partial_q L^{\geqslant 3} \\ 0 \end{pmatrix}\mathrm{d}\tilde{l}\right\}. \tag{3.4}$$

An integration by parts leads to

$$\mathbf{Q} = \underline{\mathbf{M}}\left\{\mathbf{Q}_i - \underline{\mathbf{J}}\int_0^l \left[\underline{\mathbf{M}}'^{\mathrm{T}}\begin{pmatrix} \partial_{q'}L^{\geqslant 3} \\ 0 \end{pmatrix} + \underline{\mathbf{M}}^{\mathrm{T}}\begin{pmatrix} \partial_q L^{\geqslant 3} \\ 0 \end{pmatrix}\right]\mathrm{d}\tilde{l} + \underline{\mathbf{J}}\left[\underline{\mathbf{M}}^{\mathrm{T}}\begin{pmatrix} \partial_{q'}L^{\geqslant 3} \\ 0 \end{pmatrix}\right]_0^l\right\}. \tag{3.5}$$

Writing the Jacobian as $\underline{\mathbf{M}}^{\mathrm{T}} = \partial_{\mathbf{Q}_i}\mathbf{Q}_1^{\mathrm{T}} = \partial_{\mathbf{Q}_i}(\mathbf{Q} - \mathbf{Q}_{\geqslant 2})^{\mathrm{T}}$ where $\mathbf{Q} = \mathbf{Q}_1 + \mathbf{Q}_{\geqslant 2}$ was split into parts which depend on $\mathbf{Q}_i$ linearly and nonlinearly, we obtain

$$\underline{\mathbf{M}}^{-1}\mathbf{Q} = \mathbf{Q}_i - \underline{\mathbf{J}}\int_0^l \left[(\partial_{\mathbf{Q}_i}(\mathbf{q}'^{\mathrm{T}} - \mathbf{q}'^{\mathrm{T}}_{\geqslant 2}))\partial_{q'}L^{\geqslant 3} + (\partial_{\mathbf{Q}_i}(\mathbf{q}^{\mathrm{T}} - \mathbf{q}^{\mathrm{T}}_{\geqslant 2}))\partial_q L^{\geqslant 3}\right]\mathrm{d}\tilde{l} + \underline{\mathbf{J}}[(\partial_{\mathbf{Q}_i}\mathbf{q}_1^{\mathrm{T}})\partial_{q'}L^{\geqslant 3}]_0^l$$

$$= \mathbf{Q}_i - \underline{\mathbf{J}}\int_0^l \left[\partial_{\mathbf{Q}_i}L^{\geqslant 3} - (\partial_{\mathbf{Q}_i}\mathbf{q}'^{\mathrm{T}}_{\geqslant 2})\partial_{q'}(L - L^2) - (\partial_{\mathbf{Q}_i}\mathbf{q}^{\mathrm{T}}_{\geqslant 2})\partial_q(L - L^2)\right]\mathrm{d}\tilde{l} + \underline{\mathbf{J}}[(\partial_{\mathbf{Q}_i}\mathbf{q}_1^{\mathrm{T}})\partial_q' L^{\geqslant 3}]_0^l$$

$$= \mathbf{Q}_i - \underline{\mathbf{J}}\int_0^l \left[\partial_{\mathbf{Q}_i}L^{\geqslant 3} + (\partial_{\mathbf{Q}_i}\mathbf{q}'^{\mathrm{T}}_{\geqslant 2})\partial_{q'}L^2 + (\partial_{\mathbf{Q}_i}\mathbf{q}'^{\mathrm{T}}_{\geqslant 2})\partial_{q'}L^2\right.$$

$$\left. - (\partial_{\mathbf{Q}_i}\mathbf{q}'^{\mathrm{T}}_{\geqslant 2})\partial_{q'}L - (\partial_{\mathbf{Q}_i}\mathbf{q}^{\mathrm{T}}_{\geqslant 2})\frac{\mathrm{d}}{\mathrm{d}\tilde{l}}\partial_{q'}L\right]\mathrm{d}\tilde{l} + \underline{\mathbf{J}}[(\partial_{\mathbf{Q}_i}\mathbf{q}_1^{\mathrm{T}})\partial_{q'}L^{\geqslant 3}]_0^l. \tag{3.6}$$

Note that the A'_x, A'_y, $p(s)$ and p_0 of Eq. (3.2) drop out of the right-hand side of Eq. (3.6) owing to the multiplications by the zeros in Eq. (3.5). The second order L^2 of the Lagrangian is a quadratic form in which every quadratic combination of the $\mathbf{q}$ and $\mathbf{q}'$ can occur. It can be written using a matrix $\underline{L}^2$ as $L^2 = \mathbf{Q}^{\mathrm{T}}\underline{L}^2\mathbf{Q}$. Part of the above integrand can be rewritten as

$$\mathbf{q}'^{\mathrm{T}}_{\geqslant 2}\partial_{q'}L^2 + \mathbf{q}^{\mathrm{T}}_{\geqslant 2}\partial_q L^2 = \mathbf{Q}'^{\mathrm{T}}_{\geqslant 2}\underline{L}^2\mathbf{Q} + \mathbf{Q}^{\mathrm{T}}\underline{L}^2\mathbf{Q}_{\geqslant 2}. \tag{3.7}$$

For convenience, we write $L^2(\mathbf{a}) = \mathbf{a}^{\mathrm{T}}\underline{L}^2\mathbf{a}$. Another integration by parts in Eq. (3.6) leads to

$$\underline{\mathbf{M}}^{-1}\mathbf{Q} = \mathbf{Q}_i - \underline{\mathbf{J}}\int_0^l \left[\partial_{\mathbf{Q}_i}(L^{\geqslant 3} + L^2(\mathbf{Q}_{\geqslant 2}))\right.$$

$$\left. + (\partial_{\mathbf{Q}_i}\mathbf{q}'^{\mathrm{T}}_{\geqslant 2})\partial_{q'_1}L^2(\mathbf{Q}_1) + (\partial_{\mathbf{Q}_i}\mathbf{q}^{\mathrm{T}}_{\geqslant 2})\partial_{q_1}L^2(\mathbf{Q}_1)\right]\mathrm{d}\tilde{l}$$

$$+ \underline{\mathbf{J}}[(\partial_{\mathbf{Q}_i}\mathbf{q}^{\mathrm{T}}_{\geqslant 2})\partial_{q'}L + (\partial_{\mathbf{Q}_i}\mathbf{q}_1^{\mathrm{T}})\partial_{q'}L^{\geqslant 3}]_0^l. \tag{3.8}$$

The first part of the integral contains $L_E = L^{\geqslant 3} + L^2(\mathbf{Q}_{\geqslant 2})$. The integral $\int_0^l L_E\,\mathrm{d}\tilde{l}$ is called the perturbation eikonal. This scheme embodies the essential requirement that the $(n + 1)$-order dependence of L_E on the

initial variables $\boldsymbol{Q}_i$ can be computed already when $\boldsymbol{Q}_{n-1}$ is known; $\boldsymbol{Q}_n$ does not need to be known. For an iteration of $\boldsymbol{Q}_n$, knowledge of $\boldsymbol{Q}_{n-1}$ is sufficient. Since $\boldsymbol{q}_1$ satisfies the first-order equation of motion, we can use the relation $\partial_{q_1} L^2(\boldsymbol{Q}_1) = (\mathrm{d}/\mathrm{d}l)\partial_{q'_1} L^2(\boldsymbol{Q}_1)$ to perform another integration by parts,

$$\underline{M}^{-1}\boldsymbol{Q} = \boldsymbol{Q}_i - \underline{J}\partial_{\boldsymbol{Q}_i}\int_0^l L_E \, \mathrm{d}\tilde{l} + \underline{J}[(\partial_{\boldsymbol{Q}_i}\boldsymbol{q}_{\geqslant 2}^{\mathrm{T}})\partial_{q'}L + (\partial_{\boldsymbol{Q}_i}\boldsymbol{q}_1^{\mathrm{T}})\partial_{q'}L^{\geqslant 3} - (\partial_{\boldsymbol{Q}_i}\boldsymbol{q}_{\geqslant 2}^{\mathrm{T}})\partial_{q'_1}L^2(\boldsymbol{Q}_1)]_0^l$$

$$= \boldsymbol{Q}_i - \underline{J}\partial_{\boldsymbol{Q}_i}\int_0^l L_E \, \mathrm{d}\tilde{l} + \underline{J}[(\partial_{\boldsymbol{Q}_i}\boldsymbol{q}^{\mathrm{T}})\partial_{q'}L - (\partial_{\boldsymbol{Q}_i}\boldsymbol{q}_1^{\mathrm{T}})\partial_{q'}L^2 - (\partial_{\boldsymbol{Q}_i}\boldsymbol{q}_{\geqslant 2}^{\mathrm{T}})\partial_{q'_1}L^2(\boldsymbol{Q}_1)]_0^l. \tag{3.9}$$

$\partial_{\boldsymbol{Q}_i} L_E$ is the part of $\partial_{\boldsymbol{Q}_i} L$ which up to order n in $\boldsymbol{Q}_i$ does not depend on $\boldsymbol{Q}_n$. Similarly the term outside the integral is simply the part of $(\partial_{\boldsymbol{Q}_i}\boldsymbol{q}^{\mathrm{T}})\partial_{q'}L$ which up to order n does not depend on $\boldsymbol{Q}_n$. We therefore write $(\partial_{\boldsymbol{Q}_i}\boldsymbol{q}^{\mathrm{T}})\partial_{q'}L - (\partial_{\boldsymbol{Q}_i}\boldsymbol{q}_1^{\mathrm{T}})\partial_{q'}L^2 - (\partial_{\boldsymbol{Q}_i}\boldsymbol{q}_{\geqslant 2}^{\mathrm{T}})\partial_{q'_1}L^2(\boldsymbol{Q}_1) = \{(\partial_{\boldsymbol{Q}_i}\boldsymbol{q}^{\mathrm{T}})\partial_{q'}L\}_E$. If we now express the Lagrangian in terms of the aberrations $\boldsymbol{\xi}$ with $\boldsymbol{Q} = \underline{M}\boldsymbol{\xi}$, we obtain the iteration equation

$$\boldsymbol{\xi}_n = {}_n\boldsymbol{Q}_i - \underline{J}\partial_{\boldsymbol{Q}_i}\int_0^l L_E(\underline{M}\boldsymbol{\xi}_{n-1}) \, \mathrm{d}\tilde{l} + \underline{J}[\{(\partial_{\boldsymbol{Q}_i}\boldsymbol{q}^{\mathrm{T}})\partial_{q'}L\}_E]_0^l. \tag{3.10}$$

When computing $\boldsymbol{\xi}_n$ from $\boldsymbol{\xi}_{n-1}$ with this iteration equation, all parts of the right-hand side which contribute to higher orders are neglected, as indicated by $= {}_n$. This iteration equation can have several advantages over the Hamiltonian iteration Eq. (1.23):

(a) The Lagrangian Eq. (1.8) is a much simpler function than the Hamiltonian Eq. (1.14).
(b) The derivative in Eq. (3.10) is performed after the integral has been evaluated. Therefore only one integral has to be computed and it describes all four coordinates of $\boldsymbol{\xi}_n$.
(c) The fact that the various coordinates are the derivatives with respect to initial conditions yields very simple relations [19] between the various expansion coefficients of $\boldsymbol{\xi}_n$, which are the so-called aberration coefficients of particle optical devices. These relations can be much simpler than relations entailed by the symplectic symmetry implicit in the Hamiltonian formulation.
(d) The second pair of coordinates in Eq. (3.2) can be calculated very easily. With Eq. (1.21) and (3.3), the equation of motion for $\boldsymbol{\xi}$ is

$$\underline{M}\boldsymbol{\xi}' = \begin{pmatrix} 0 \\ \dfrac{\mathrm{d}}{\mathrm{d}\tilde{l}}\partial_{q'}L^{\geqslant 3} - \partial_q L^{\geqslant 3} \end{pmatrix}. \tag{3.11}$$

After having computed $\boldsymbol{q}_n = \underline{M}_{2\times 4}\boldsymbol{\xi}_n$ by iteration, the derivative $\boldsymbol{q}'$ can then easily be computed as $\boldsymbol{q}'_n = \underline{M}'_{2\times 4}\boldsymbol{\xi}_n$ using Eq. (3.11). One thus only needs to iterate the two-dimensional vector $\boldsymbol{q}_n$ and not a four-dimensional vector $\boldsymbol{z}_n$ as in the Hamiltonian iteration procedure.

4. Successive approximation for spin–orbit motion

The time variation of a spin $\boldsymbol{s}$ in the rest frame of a particle is described by the so-called Thomas–BMT equation $\dot{\boldsymbol{s}} = \boldsymbol{\Omega}_{\mathrm{BMT}} \times \boldsymbol{s}$ [20,21] where

$$\boldsymbol{\Omega}_{\mathrm{BMT}} = -\frac{q}{m\gamma}\left\{(a\gamma + 1)\boldsymbol{B}_\perp + (1 + a)\boldsymbol{B}_\parallel - \frac{\gamma}{c}\boldsymbol{\beta} \times \boldsymbol{E}\left(a + \frac{1}{1 + \gamma}\right)\right\} \tag{4.1}$$

with the electric field $\boldsymbol{E}$, the parts of the magnetic field $\boldsymbol{B}$ which are perpendicular ($\perp$) and parallel ($\parallel$) to the particle's velocity, and the anomalous gyro-magnetic factor $a = (g - 2)/2$.

Changing to the comoving coordinate system of Fig. 1, we obtain $s = S_x e_x + S_y e_y + S_l e_l$ and
$s' = (S'_x - S_l \kappa_x)e_x + (S'_y - S_l \kappa_y)e_y + (S'_l + S_x \kappa_x + S_y \kappa_y)e_l$. The equation of motion for the vector S of these
spin components is then

$$S' = \Omega \times S, \quad \Omega = \Omega_{\text{BMT}} \frac{h}{v} \sqrt{x'^2 + y'^2 + h^2} - \kappa \times e_l. \tag{4.2}$$

The equations of motion for the phase-space vector z and the spin S have the form

$$z' = f(z, l), \quad S' = \Omega(z, l) \times S. \tag{4.3}$$

The general solutions transporting the coordinates along the optical system, starting at the initial values
z_i, S_i, is given by the transport map $\mathcal{m}$ and the rotation matrix $\underline{R} \in SO(3)$,

$$z(l) = \mathcal{m}(z_i, l), S(l) = \underline{R}(z_i, l)S_i. \tag{4.4}$$

In order to find the general solution, one could compute the nine coefficients of the rotation matrix by
solving the differential equation

$$R_{ij}(z_i, l)' = \varepsilon_{ilk} \Omega_l R_{kj}(z_i, l), \tag{4.5}$$

where the vector product was expressed by the totally antisymmetric tensor ε_{ilk}. However, computing the
nine components of the rotation matrix seems inefficient, since a rotation can be represented by three angles.
It has turned out [22] to be most efficient to represent the rotation of spins by the quaternion A which gives
the rotation transformation in the SU(2) representation as

$$A = a_0 \underline{1} - i a \cdot \underline{\sigma}. \tag{4.6}$$

Here $\underline{1}$ is the 2×2 identity matrix and the elements of the vector $\underline{\sigma}$ are the three two-dimensional Pauli
matrixes. When a rotation by an angle ϕ is performed around the unit vector e, the quaternion representation
of the rotation has $a_0 = \cos(\phi/2)$ and $a = \sin(\phi/2)e$. Therefore, $a_0^2 + a^2 = 1$ and the identity transformation is
represented by $a_0 = 1$.

If a particle traverses an optical element which rotates the spin according to the quaternion A and then
passes through an element which rotates the spin according to the quaternion B, the total rotation of the spin
is given by

$$C = c_0 \underline{1} - i c \cdot \underline{\sigma} = (b_0 \underline{1} - i b \cdot \underline{\sigma}) \cdot (a_0 \underline{1} - i a \cdot \underline{\sigma})$$

$$= (b_0 a_0 - b \cdot a)\underline{1} - i(b_0 a + b a_0 + b \times a) \cdot \underline{\sigma}. \tag{4.7}$$

The concatenation of quaternions can be written in matrix form as

$$C = \begin{pmatrix} c_0 \\ c \end{pmatrix} = \underline{B} \begin{pmatrix} a_0 \\ a \end{pmatrix}, \quad \underline{B} = \begin{pmatrix} b_0 & -b_1 & -b_2 & -b_3 \\ b_1 & b_0 & -b_3 & b_2 \\ b_2 & b_3 & b_0 & -b_1 \\ b_3 & -b_2 & b_1 & b_0 \end{pmatrix}. \tag{4.8}$$

This concatenation of two quaternions can be used to find a differential equation for the spin rotation.

While propagating along the design curve by a distance dl, spins are rotated by an angle $\Omega \, dl = |\Omega| \, dl$
around the vector Ω. After having been propagated to l by the quaternion A, a spin gets propagated from l to

$l + \mathrm{d}l$ by the quaternion with $b_0 = 1$ and $\boldsymbol{b} = \frac{1}{2}\boldsymbol{\Omega}\,\mathrm{d}l$. The resulting total rotation is given by $A + A'\,\mathrm{d}l$ and we obtain the differential equation

$$
\begin{pmatrix} a'_0 \\ \boldsymbol{a}' \end{pmatrix} = \frac{1}{2}\begin{vmatrix} 0 & -\Omega_1 & -\Omega_2 & -\Omega_3 \\ \Omega_1 & 0 & -\Omega_3 & \Omega_2 \\ \Omega_2 & \Omega_3 & 0 & -\Omega_1 \\ \Omega_3 & -\Omega_2 & \Omega_1 & 0 \end{vmatrix}\begin{pmatrix} a_0 \\ \boldsymbol{a} \end{pmatrix}.
\tag{4.9}
$$

Writing the matrix as $\underline{\boldsymbol{\Omega}}$ and the vector as A, the spin orbit equation of motion has the form

$$
z' = \boldsymbol{f}(z, l), \quad A' = \underline{\boldsymbol{\Omega}}(z, l)\,A.
\tag{4.10}
$$

The starting conditions are $z(0) = z_0, a_0 = 1$, and $\boldsymbol{a} = 0$. The quaternion A depends on the initial phase-space coordinates z_i and can be expanded in a Taylor series with respect to these coordinates. In the following, we want to devise an iteration method for A_n, which is the Taylor expansion to order n of A.

The rotation vector $\boldsymbol{\Omega}$ is split into its value on the design curve and its phase-space-dependent part as $\boldsymbol{\Omega}(z, l) = \boldsymbol{\Omega}^0(l) + \boldsymbol{\Omega}^{\geq 1}(z, l)$. The spin motion on the design curve is given by $A'_0(l) = \underline{\boldsymbol{\Omega}}^0 A_0(l)$. Similarly, to Eq. (1.21), spin aberrations are defined with respect to the leading order motion. Small phase-space coordinates will create a rotation which differs little from $A_0(l)$ and we write the phase-space-dependent rotation as a concatenation of A_0 and the z-dependent rotation $(1 + \delta, \boldsymbol{\delta})$ which reduces to the identity for $z = 0$ by requiring that the aberrations δ and $\boldsymbol{\delta}$ vanish on the design curve. With Eq. (4.8) we obtain

$$
A = \underline{A_0}\begin{pmatrix} 1 + \delta \\ \boldsymbol{\delta} \end{pmatrix}.
\tag{4.11}
$$

The quaternion A is now inserted in the differential Eq. (4.10) to obtain

$$
\underline{A}'_0\begin{pmatrix} 1 + \delta \\ \boldsymbol{\delta} \end{pmatrix} + \underline{A_0}\begin{pmatrix} \delta' \\ \boldsymbol{\delta}' \end{pmatrix} = (\underline{\boldsymbol{\Omega}}^0 + \underline{\boldsymbol{\Omega}}^{\geq 1})\underline{A_0}\begin{pmatrix} 1 + \delta \\ \boldsymbol{\delta} \end{pmatrix}.
\tag{4.12}
$$

Taking into account the equation on the design curve and the fact that $\underline{A}_0^{\mathrm{T}}$ describes the inverse rotation of $\underline{A}_0$, we obtain

$$
\begin{pmatrix} \delta' \\ \boldsymbol{\delta}' \end{pmatrix} = (\underline{A}_0^{\mathrm{T}}\underline{\boldsymbol{\Omega}}^{\geq 1}\underline{A_0})\begin{pmatrix} 1 + \delta \\ \boldsymbol{\delta} \end{pmatrix} = \underline{\tilde{\boldsymbol{\Omega}}}(z, l)\begin{pmatrix} 1 + \delta \\ \boldsymbol{\delta} \end{pmatrix}.
\tag{4.13}
$$

Writing the Taylor expansion to order n in z_i, one finally obtains the iteration equation

$$
\begin{pmatrix} \delta_n \\ \boldsymbol{\delta}_n \end{pmatrix} = {}_n\!\int_0^l \underline{\tilde{\boldsymbol{\Omega}}}(z_n)\begin{pmatrix} 1 + \delta_{n-1} \\ \boldsymbol{\delta}_{n-1} \end{pmatrix}\mathrm{d}\tilde{l}, \quad \begin{pmatrix} \delta_0 \\ \boldsymbol{\delta}_0 \end{pmatrix} = 0.
\tag{4.14}
$$

This iteration method was used for the spin transport in the program SPRINT [22] and was evaluated using MATHEMATICA in Ref. [23].

In the case of successive approximation in terms of the Hamiltonian, the various aberration coefficients were related by the symplectic symmetry. With the Lagrange formalism the various aberration coefficients were related by their being derivatives of a common integral with respect to different initial coordinates. In the case of the successive approximation for spin motion, the various aberration coefficients in $\boldsymbol{\delta}$ and δ are related by the relation $(1 + \delta)^2 + \boldsymbol{\delta}^2 = 0$.

Acknowledgements

Thanks are due to D. Barber, H. Mais, and M. Vogt for thoroughly reading the manuscript and for the resulting improvements.

References

[1] H. Rose, U. Petri, Optik 33 (1971) 151.
[2] E. Plies, D. Typke, Zeitschrift für Naturforschung 33a (1978) 1361–1377.
[3] H. Rose, Nucl. Instr. and Meth. A 258 (1987) 374–401.
[4] H. Rose, Nucl. Instr. Meth. 187 (1981) 187–199.
[5] H. Rose, Optic 85 (1) (1990) 19–24.
[6] H. Rose, M. Haider, K. Urban, Physikalische Blätter 54 (5) (1998) 411–416.
[7] P.A. Sturrock, Static and dynamic electron optics, Part II, Cambridge University Press, Cambridge, 1995, p. 149.
[8] D.P. Barber et al., Phys. Lett. B 343 (1995) 436–443.
[9] SPIN Collaboration and the DESY Polarization Team. Acceleration of polarized protons to 820 GeV at HERA. UM–HR 96–20, University of Michigan Report, 1996.
[10] G. H. Hoffstaetter, Nonlinear dependence of synchrotron radiation on beam parameters, in: Proceedings of PAC 94, Dallas/TX, 1995.
[11] G. Wüstefeld, Orbit maps for helical snake and helical undulator magnets. In: Proceedings of the workshop on Polarized Protons at High Energies, DESY, Hamburg, 1999.
[12] A. Luccio, T. Roser, Third workshop on Siberian Snakes and Spin Rotators. formal report BNL-52453, Brookhaven National Laboratory, Upton, New York, 1994.
[13] H. Goldstein, Classical Mechanics, 2nd edition, Addison-Wesley, Reading, MA, USA, 1980.
[14] E.D. Courant, H.S. Snyder, Ann. Phys. 3 (1958) 1–48.
[15] H. Mais, Some topics in beam dynamics of storage rings, formal report DESY-96-119, 1996.
[16] E. Plies, H. Rose, Optik 34 (2) (1971) 171–190.
[17] G.H. Hoffstaetter, Geometrische Elektronenoptik angewandt auf ein durch Hexapole korrigiertes Mikroskop mit sub Ångström Auflösung Diploma Thesis, Darmstadt University of Technology, 1991.
[18] G.H. Hoffstaetter, H. Rose, Nucl. Instr. A 328 (1993) 398–401.
[19] G.H. Hoffstaetter, Nucl. Instr. and Meth. A 427 (1999) 275–281.
[20] L.H. Thomas, Phil. Mag. 3 (1927) 1.
[21] V. Bargmann, L. Michel, V.L. Telegdi, Phys. Rev. Lett. 2 (10) (1959) 435–436.
[22] K. Heinemann, G.H. Hoffstaetter, Phys. Rev. E 54 (1996) 4240–4255 and formal report DESY-96-078.
[23] Ch. Weißbäcker, Nichtlineare Effekte der Spindynamik in Protonenbeschleunigern. Diploma Thesis, Darmstadt University of Technology, 1998.