The Conjugate Variable Method in Hamilton-Lie Perturbation Theory −Applications to Plasma Physics−

Shinji TOKUDA

Japan Atomic Energy Agency, Naka, Ibaraki, 311-0193 Japan

(Received 21 February 2008 / Accepted 29 July 2008)

The conjugate variable method, an essential ingredient in the path-integral formalism of classical statistical dynamics, is used to apply the Hamilton-Lie perturbation theory to a system of ordinary differential equations that does not have the Hamiltonian dynamic structure. The method endows the system with this structure by doubling the unknown variables; hence, the canonical Hamilton-Lie perturbation theory becomes applicable to the system. The method is applied to two classical problems of plasma physics to demonstrate its effectiveness and study its properties: a non-linear oscillator that can explode and the guiding center motion of a charged particle in a magnetic field.

Keywords: one-form, Hamilton-Lie perturbation method, conjugate variable, non-linear oscillator, guiding center motion, plasma physics

DOI: 10.1585/pfr.3.057

1. Introduction

We begin by considering the following two non-linear oscillators

$$\frac{dx}{dt} = y, \quad \frac{dy}{dt} = -x - \epsilon x^3, \tag{1}$$

and

$$\frac{dx}{dt} = y, \quad \frac{dy}{dt} = -x - \epsilon y^3, \tag{2}$$

where t is the time variable. When $\epsilon = 0$, both oscillators reduce to an identical harmonic oscillator. When $\epsilon \neq 0$, the equations have quite different properties. Equation (1) has the variational principle in the phase space (x, y) [1], and performs the Hamiltonian motion under the Hamiltonian

$$h(x, y) = \frac{x^2}{2} + \frac{y^2}{2} + \frac{\epsilon}{4}x^4. \tag{3}$$

One of the properties peculiar to the Hamiltonian motion is, of course, the volume invariance of the flow of motion in the phase space (x, y). Equation (2) does not have the Hamiltonian structure and loses the invariant property of the volume in the phase space (x, y).

Next, an important property of Eq. (1) from a practical point of view is that it can be analyzed by a powerful tool, the Hamilton-Lie perturbation method [2]. This method, however, is not applicable for Eq. (2); therefore, different approximation methods have been devised for Eq. (2), such as the multi-scale expansion method [3]. These methods abandon the viewpoint of Hamiltonian dynamics.

Since the Hamilton-Lie perturbation method is a powerful analytical tool that enables us to investigate the problem deeply, it would be a significant contribution to the development of approximation methods in physics if it were made applicable for equations such as Eq. (2), which do not have the Hamiltonian structure. This seems impossible since the Hamilton-Lie perturbation method assumes the Hamiltonian structure of the equations. However, that is not the case. We can endow any ordinary differential equations with the Hamiltonian structure by doubling the unknown variables. This technique is well known as the *conjugate variable method*, and plays an essential role in the path-integral formalism of classical statistical dynamics [4]. Once we make the Hamiltonian for Eq. (2) by the conjugate variable method, we can systematically apply the Hamilton-Lie perturbation method to the equation. In the present paper, we report and also advocate the conjugate variable method for the Hamilton-Lie perturbation method.

In Sec. 2, after briefly reviewing the conjugate variable method, we apply the Hamilton-Lie perturbation method to Eq. (2) up to the second order $O(\epsilon^2)$. We demonstrate that the equations for the original variables are obtained in a closed form that does not include conjugate variables, although the conjugate variable method doubles the number of variables. In Sec. 3, we apply the conjugate variable method to the guiding center motion of a charged particle in a magnetic field. This is a classical problem with a long history in plasma physics [5, 6]. The most elegant and profound solution method will be the *non-canonical* perturbation method [7]. We show that when the conjugate variable

author's e-mail: tokuda.shinji@jaea.go.jp

method is used, the guiding center problem can be solved by the *canonical* perturbation method. Our conclusions are given in Sec. 4, with discussion on further applications of the present method to plasma physics.

2. Perturbation Analysis of a Non-Linear Oscillator by Introducing Conjugate Variables

Following Ref. [4], we can construct a fundamental 1-form for Eq. (2) as follows

$$\begin{aligned} \gamma &= q(\mathrm{d}x - y\mathrm{d}t) + p[\mathrm{d}y + (x + \epsilon y^3)\mathrm{d}t] \\ &= q\mathrm{d}x + p\mathrm{d}y - h\mathrm{d}t. \end{aligned} \tag{4}$$

Here, q(res. p) is called the conjugate variable with respect to x(res. y); h is the Hamiltonian for the 1-form and is given by

$$h = h^{(0)} + \epsilon h^{(1)}, \tag{5}$$

$$h^{(0)} = qy - px, \tag{6}$$

and

$$h^{(1)} = -py^3. \tag{7}$$

Since Eq. (4) is a canonical 1-form, the equations of motion for the phase space coordinates (x, q, y, p) are obtained from the Hamilton equation

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\partial h}{\partial q} = y, \tag{8}$$

$$\frac{\mathrm{d}q}{\mathrm{d}t} = -\frac{\partial h}{\partial x} = p, \tag{9}$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\partial h}{\partial p} = -x - \epsilon y^3, \tag{10}$$

and

$$\frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial h}{\partial y} = -q + 3\epsilon p y^2. \tag{11}$$

The equation of motion for any function of the phase space coordinates is also expressed by the Hamilton equation. For example, the equation of motion for the energy of the oscillator,

$$\mathcal{E} = \frac{1}{2}(x^2 + y^2), \tag{12}$$

is given by

$$\frac{\mathrm{d}\mathcal{E}}{\mathrm{d}t} = \{\mathcal{E}, h\} = -\epsilon y^4 \leq 0, \tag{13}$$

where $\{\ ,\ \}$ is the Poisson bracket in the phase space (x, q, y, p).

The 1-form, Eq. (4), appears to be only nominal and therefore useless. It would be so if the only solution were to solve the differential equations, Eqs. (8)-(11), directly. However, the Hamiltonian structure provides other powerful devices to solve the equation of motion: coordinate transformations and gauge functions in the phase space. We will show them for the non-linear oscillator, Eq. (2), below.

2.1 Harmonic oscillator with conjugate variables

We first examine the case of $\epsilon = 0$, the harmonic oscillator. Let us define the transformation from (x, y) to (a, θ) by

$$x = a\cos\theta, \quad y = -a\sin\theta. \tag{14}$$

Then we have

$$\begin{aligned} q\mathrm{d}x + p\mathrm{d}y = [&q\cos\theta - p\sin\theta]\mathrm{d}a \\ &- a[q\sin\theta + p\cos\theta]\mathrm{d}\theta, \end{aligned} \tag{15}$$

and

$$h^{(0)} = -a(q\sin\theta + p\cos\theta). \tag{16}$$

These results imply the transformation from (a, θ, q, p) to (a, θ, Q, P) defined by

$$Q = Q(q, p, a, \theta) = q\cos\theta - p\sin\theta, \tag{17}$$

and

$$P = P(q, p, a, \theta) = -a(q\sin\theta + p\cos\theta). \tag{18}$$

Then, we obtain the 1-form of the harmonic oscillator in the phase space (a, Q, θ, P) given by

$$\gamma = Q\mathrm{d}a + P\mathrm{d}\theta - h^{(0)}\mathrm{d}t, \tag{19}$$

with the Hamiltonian

$$h^{(0)} = P. \tag{20}$$

Since $h^{(0)}$ is independent of a, Q and θ, the conjugate variables Q, a, P with respect to them are constants of motion. On the other hand, the equation of motion for θ is

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \frac{\partial h^{(0)}}{\partial P} = 1. \tag{21}$$

Consequently, we have

$$a(t) = a_0, \quad \theta(t) = t + \theta_0, \tag{22}$$

and

$$Q(t) = Q_0, \quad P(t) = P_0, \tag{23}$$

where a_0, θ_0, Q_0 and P_0 are arbitrary initial values, and from Eq. (14) we have

$$x(t) = a_0\cos(t + \theta_0), \tag{24}$$

and

$$y(t) = -a_0\sin(t + \theta_0). \tag{25}$$

The inverse transformations of Eqs. (17) and (18) are

$$q = Q\cos\theta - \frac{P}{a}\sin\theta, \tag{26}$$

and

$$p = -Q\sin\theta - \frac{P}{a}\cos\theta, \tag{27}$$

and $q(t)$ and $p(t)$ are easily obtained. Note that to perform the transformations of Eqs. (17) and (18), it is indispensable to increase the variable (x, y) to (x, q, y, p).

For applying the Hamilton-Lie perturbation method, it is convenient to introduce

$$A = \frac{1}{2}a^2, \tag{28}$$

and to replace Q/a with Q. The resultant 1-form remains canonical,

$$\gamma = QdA + P\mathrm{d}\theta - h^{(0)}\mathrm{d}t, \tag{29}$$

with (A, θ, Q, P) being canonical coordinates.

Now, we return to Eq. (2) with $\epsilon \neq 0$. The perturbed Hamiltonian $h^{(1)}$ written in the phase space (A, θ, Q, P) is

$$\begin{aligned} h^{(1)} &= -py^3 = -\frac{3}{2}QA^2 + QA^2 f_{\mathrm{c}}(\theta) + PAf_{\mathrm{s}}(\theta) \\ &= \langle h^{(1)}\rangle + \tilde{h}^{(1)}, \end{aligned} \tag{30}$$

and

$$\langle h^{(1)}\rangle = -\frac{3}{2}QA^2, \tag{31}$$

where $f_{\mathrm{c}}(\theta)$ and $f_{\mathrm{s}}(\theta)$ are periodic functions

$$f_{\mathrm{c}}(\theta) = 2\cos 2\theta - \frac{1}{2}\cos 4\theta, \tag{32}$$

and

$$f_{\mathrm{s}}(\theta) = -\frac{1}{2}\sin 2\theta + \frac{1}{4}\sin 4\theta, \tag{33}$$

and $\langle\ \rangle$ represents the average with respect to θ.

The Hamilton equations of motion developed from the total Hamiltonian $h = h^{(0)} + \epsilon h^{(1)}$ are

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \frac{\partial h}{\partial Q} = -\frac{3}{2}\epsilon A^2 + \epsilon A^2 f_{\mathrm{c}}(\theta), \tag{34}$$

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \frac{\partial h}{\partial P} = 1 + \epsilon A f_{\mathrm{s}}(\theta), \tag{35}$$

$$\begin{aligned} \frac{\mathrm{d}Q}{\mathrm{d}t} = -\frac{\partial h}{\partial A} &= 3\epsilon QA \\ &\quad - \epsilon[Pf_{\mathrm{s}}(\theta) + 2QAf_{\mathrm{c}}(\theta)], \end{aligned} \tag{36}$$

and

$$\frac{\mathrm{d}P}{\mathrm{d}t} = -\frac{\partial h}{\partial \theta} = -\epsilon A[Pf_{\mathrm{s}}'(\theta) + QAf_{\mathrm{c}}'(\theta)]. \tag{37}$$

These equations of motion are exact, and prove the effectiveness of the conjugate variable method and the transformations among phase space coordinates. The Hamiltonian h has a formal characteristic that h is linear in the conjugate variables Q and P. Consequently, the equations for A and θ, Eqs. (34) and (35), are closed; they do not contain the conjugate variables. In contrast, the equations for Q and P, Eqs. (36) and (37), are linear with coefficients that are functions of time t determined implicitly by A and θ. We also see that Eq. (34) predicts an explosion of amplitude A for $\epsilon < 0$.

Equations (34)-(37) can be systematically solved by applying the canonical perturbation method, where the gauge functions are thoroughly exploited; a summary is given in Appendix A. It is shown that the canonical perturbation method is applicable even if the amplitude A might explode.

2.2 First-order canonical perturbation analysis

Following Ref. [2], we determine the gauge function $S^{(1)}$ such that the Hamiltonian after the Lie transformation is $H^{(1)} = \langle h^{(1)}\rangle$. From Eq. (A9) in Appendix A, the condition for this yields the partial differential equation for $S^{(1)}$

$$\frac{\partial S^{(1)}}{\partial\theta} = \tilde{h}^{(1)}. \tag{38}$$

Here, we have used the conditions that $S^{(1)}$ is independent of the time t, and have used

$$\{S^{(1)}, h^{(0)}\} = \{S^{(1)}, P\} = \partial_\theta S^{(1)}. \tag{39}$$

Equation (38) is easily solved,

$$S^{(1)} = QA^2F_{\mathrm{s}}(\theta) + PAF_{\mathrm{c}}(\theta), \tag{40}$$

where

$$F_{\mathrm{s}}(\theta) = \sin(2\theta) - \frac{1}{8}\sin(4\theta), \tag{41}$$

and

$$F_{\mathrm{c}}(\theta) = \frac{1}{4}\cos(2\theta) - \frac{1}{16}\cos(4\theta). \tag{42}$$

The Hamiltonian in the phase space $(\bar{A}, \bar{\theta}, \bar{Q}, \bar{P})$ after the Lie transformation is

$$\mathcal{H} = \bar{P} - \frac{3\epsilon}{2}\bar{Q}\bar{A}^2, \tag{43}$$

and the equations of motion in the phase space $(\bar{A}, \bar{\theta}, \bar{Q}, \bar{P})$ are

$$\frac{\mathrm{d}\bar{A}}{\mathrm{d}t} = \frac{\partial\mathcal{H}}{\partial\bar{Q}} = -\frac{3}{2}\epsilon\bar{A}^2, \tag{44}$$

$$\frac{\mathrm{d}\bar{\theta}}{\mathrm{d}t} = \frac{\partial\mathcal{H}}{\partial\bar{P}} = 1, \tag{45}$$

$$\frac{\mathrm{d}\bar{Q}}{\mathrm{d}t} = -\frac{\partial\mathcal{H}}{\partial\bar{A}} = 3\epsilon\bar{Q}\bar{A}, \tag{46}$$

and

$$\frac{\mathrm{d}\bar{P}}{\mathrm{d}t} = -\frac{\partial\mathcal{H}}{\partial\bar{\theta}} = 0. \tag{47}$$

The solutions of these equations are easily obtained; for $\epsilon < 0$

$$\bar{A}(t) = \bar{A}_0 \frac{1}{1 - (3/2)\bar{A}_0|\epsilon|t}, \qquad (48)$$

where $\bar{A}_0$ is the initial value of $\bar{A}$. On the other hand, the solution

$$\bar{Q}(t) = \bar{Q}_0 \left(1 - \frac{3}{2}\bar{A}_0|\epsilon|t\right)^2, \qquad (49)$$

becomes zero at the time when $\bar{A}$ explodes ($\bar{Q}_0$ is the initial value of $\bar{Q}$); this reflects the conservation property of the phase volume in the (A, θ, Q, P) space.

Let us observe the motion on the original phase space (A, θ, Q, P). We need the generators $g^1_{(1)}$ and $g^2_{(1)}$ of the Lie transformation $(A, \theta, Q, P) \to (\bar{A}, \bar{\theta}, \bar{Q}, \bar{P})$. They are

$$g^1_{(1)} = -\frac{\partial S^{(1)}}{\partial Q} = -A^2 F_{\rm s}(\theta), \qquad (50)$$

and

$$g^2_{(1)} = -\frac{\partial S^{(1)}}{\partial P} = -A F_{\rm c}(\theta). \qquad (51)$$

The backward transformations are therefore given by

$$A = \bar{A} - \epsilon g^1_{(1)} = \bar{A} - \epsilon \bar{A}^2 F_{\rm s}(\bar{\theta}), \qquad (52)$$

and

$$\theta = \bar{\theta} - \epsilon g^2_{(1)} = \bar{\theta} - \epsilon \bar{A} F_{\rm s}(\bar{\theta}). \qquad (53)$$

By applying them to $x(t) = \sqrt{2A(t)}\cos\theta(t)$, we have

$$x(t) = a\left(1 - \frac{3}{4}a^2|\epsilon|t\right)^{-1/2} \cos(t) + O(\epsilon, |\epsilon|^2 t), \qquad (54)$$

where we set $a = \sqrt{2A_0}$. We found that the solution given by Eq. (54) accords with the solution given by the multi-scale expansion method [3].

2.3 Second-order canonical perturbation analysis

From Eq. (31) and Eq. (40), we have for Eq. (A22) in Appendix A

$$\{S^{(1)}, \langle h^{(1)}\rangle\} = -\frac{3}{2}F_{\rm c}(\theta)\{PA, QA^2\} = -\frac{3}{2}F_{\rm c}(\theta)PA^2.$$

Notice that the Poisson bracket results in a linear function of P. Similar but lengthy algebra with Eq. (A22) gives

$$\begin{aligned}\xi^{(2)}_0 = &-\frac{PA^2}{2}[3F_{\rm c}(\theta) + f_{\rm s}(\theta)F_{\rm s}(\theta) \\ &- f_{\rm c}(\theta)F_{\rm c}(\theta) + F_{\rm c}(\theta)f'_{\rm s}(\theta) - (f_{\rm s}(\theta))^2] \\ &- \frac{QA^3}{2}[F_{\rm c}(\theta)f'_{\rm c}(\theta) - f_{\rm s}(\theta)f_{\rm c}(\theta)]. \end{aligned} \qquad (55)$$

Due to the properties of the Poisson bracket, the second-order Hamiltonian is still *linear* in the conjugate variables P and Q, while $\xi^{(2)}_0$ has powers of A higher than those in $h^{(1)}$. Equation (55) can be re-arranged as

$$\xi^{(2)}_0 = -\frac{PA^2}{2}F^{(2)}_P(\theta) - \frac{QA^3}{2}F^{(2)}_Q(\theta) + \langle\xi^{(2)}_0\rangle, \qquad (56)$$

where $F^{(2)}_P(\theta)$ and $F^{(2)}_Q(\theta)$ are periodic functions whose averages are zero, and the average of $\xi^{(2)}_0$ is

$$\langle\xi^{(2)}_0\rangle = \frac{27}{64}PA^2. \qquad (57)$$

Therefore, when we choose the gauge function in Eq. (A21) as

$$S^{(2)} = \frac{PA^2}{2}\int F^{(2)}_P(\theta)\mathrm{d}\theta + \frac{QA^3}{2}\int F^{(2)}_Q(\theta)\mathrm{d}\theta, \qquad (58)$$

the averaged Hamiltonian in the second order is

$$\mathcal{H} = \bar{P} - \frac{3\epsilon}{2}\bar{Q}\bar{A}^2 - \frac{27}{64}\epsilon^2\bar{P}\bar{A}^2. \qquad (59)$$

Let us first notice that since the averaged Hamiltonian $\mathcal{H}$ is linear in $\bar{P}$ and $\bar{Q}$, they are not incorporated in the equations for $\bar{A}$ and $\bar{\theta}$. Next, the Hamiltonian $\mathcal{H}$ does not have $\bar{\theta}$ due to the averaging, the conjugate variable $\bar{P}$ is the constant of motion. We also see that the second-order perturbation does not influence the behavior of amplitude $\bar{A}$, which remains Eq. (48), but influences the behavior of $\bar{\theta}$, the equation of which is

$$\frac{\mathrm{d}\bar{\theta}}{\mathrm{d}t} = \frac{\partial\mathcal{H}}{\partial\bar{P}} = 1 - \frac{27}{64}\epsilon^2\bar{A}^2, \qquad (60)$$

and using Eq. (48), we have

$$\bar{\theta}(t) = \theta_0 + t - \frac{9}{32}(|\epsilon|\bar{A}_0)\left[\frac{1}{1 - 3/2|\epsilon|\bar{A}_0 t} - 1\right]. \qquad (61)$$

The second-order perturbation delays the phase because $\mathrm{d}\bar{\theta}/\mathrm{d}t - 1 < 0$, irrespective of the sign of ϵ; the phase delay becomes remarkable as t approaches the explosion time.

Finally, let us obtain the Lie transformation $(A, \theta, Q, P) \to (\bar{A}, \bar{\theta}, \bar{Q}, \bar{P})$. The generators produced by $S^{(2)}$ are

$$g^1_{(2)} = -\frac{\partial S^{(2)}}{\partial Q} = -\frac{A^3}{2}\int F^{(2)}_Q(\theta)\mathrm{d}\theta, \qquad (62)$$

and

$$g^2_{(2)} = -\frac{\partial S^{(1)}}{\partial P} = -\frac{A^2}{2}\int F^{(2)}_P(\theta)\mathrm{d}\theta. \qquad (63)$$

Next, let us investigate the second-order effects of the generators $g^j_{(1)}$. By substituting Eqs. (50) and (51) into the fourth term in the right-hand side of Eq. (A24), we get

$$\begin{aligned}(g^{\nu}_{(1)}\partial_{\nu})g^1_{(1)} &= g^{\nu}_{(1)}\partial_{\nu}(-A^2F_{\rm s}) \\ &= A^3[2(F_{\rm s})^2 + F_{\rm c}f_{\rm c}], \end{aligned} \qquad (64)$$

and

$$(g^{\nu}_{(1)}\partial_{\nu})g^{2}_{(1)} = g^{\nu}_{(1)}\partial_{\nu}(-AF_{\mathrm{c}}) \\ = A^{2}[F_{\mathrm{s}}F_{\mathrm{c}} + F_{\mathrm{c}}f_{\mathrm{s}}]. \tag{65}$$

Here, $g^{3}_{(1)}$, $g^{4}_{(1)}$ that contain the conjugate variables do not contribute to the above results. Consequently, the backward Lie transformation does not mix the original variables A, θ with the conjugate variables Q, P.

Equation (65) implies that the Lie transformation does not yield any secular terms in θ, but Eq. (64) produces a secular term

$$\langle 2(F_{\mathrm{s}})^{2} + F_{\mathrm{c}}f_{\mathrm{c}}\rangle = \frac{83}{64}, \tag{66}$$

and then the averaged solution of A is

$$A(t) = \bar{A}(t)\left(1 + \frac{83}{128}(\epsilon\bar{A})^{2}\right), \tag{67}$$

in the original phase space.

Although the conjugate variable method doubles the variables, we obtain the equations of A and θ and their Lie transformations unmixed with the conjugate variables P and Q until the second order. These are the consequences of the Poisson bracket's properties and should be valid for any orders, although a strong proof is not provided here. If mixing occurred at each order, the conjugate variable method would be complicated, and therefore, would not be an effective tool.

3. Analysis of Charged Particle Motion

In the Cartesian coordinate system (x_1, x_2, x_3), the equation of motion for a charged particle with mass m and electrical charge e in a magnetic field $\boldsymbol{B}$ (the Lorentz equation) is

$$\frac{\mathrm{d}x_j}{\mathrm{d}t} = u_j, \quad \frac{\mathrm{d}u_j}{\mathrm{d}t} = \frac{e}{m}(\boldsymbol{u}\times\boldsymbol{B})_j, \quad j = 1, 2, 3. \tag{68}$$

By introducing the conjugate variables q_j(res. y_j) for x_j(res. u_j), the fundamental 1-form for the Lorentz equation is constructed as

$$\gamma = q_j\mathrm{d}x_j + y_j\mathrm{d}u_j - \mathcal{H}\mathrm{d}t, \tag{69}$$

where the Hamiltonian $\mathcal{H}$ is

$$\mathcal{H} = u_j q_j + \frac{e}{m}\boldsymbol{y}\cdot(\boldsymbol{u}\times\boldsymbol{B}). \tag{70}$$

Here, the summation convention is implied for the repeated index. Since the Hamiltonian $\mathcal{H}$ is linear in $\boldsymbol{B}$, $\mathcal{H}$ is easily decomposed into unperturbed and perturbed parts when expressing $\boldsymbol{B} = \boldsymbol{B}_0 + \epsilon\tilde{\boldsymbol{B}}$. Therefore, as long as the unperturbed Hamiltonian made by $\boldsymbol{B}_0$ is solvable, the canonical perturbation method can be systematically applied to Eq. (69).

We first study the guiding center motion in a uniform magnetic field, and demonstrate that doubling the variables enables us to perform a canonical transformation in such a manner that the guiding center motion looks straight on the transformed phase space [Eqs. (91) and (93)]. This is a nontrivial result and is the starting point for the perturbation analysis. In order to make the analysis intelligible, we adopt a simple configuration of magnetic fields, Eq. (98).

3.1 Guiding center motion in a uniform magnetic field

Let a uniform magnetic field be given by

$$\boldsymbol{B} = B\boldsymbol{b} = B\boldsymbol{e}_3, \tag{71}$$

where $\boldsymbol{e}_3$ is the unit vector along the z-axis. Also, let $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ be unit vectors along the x- and y- axes, respectively. Then, the Hamiltonian in Eq. (70) reads

$$\mathcal{H} = u_j q_j + \Omega\boldsymbol{y}\cdot(\boldsymbol{u}\times\boldsymbol{b}), \tag{72}$$

where $\Omega = eB/m$. Following the theory of guiding center motion [7], we introduce the auxiliary vectors

$$\boldsymbol{a} = \boldsymbol{e}_1\cos\theta - \boldsymbol{e}_2\sin\theta, \tag{73}$$

and

$$\boldsymbol{c} = -\boldsymbol{e}_1\sin\theta - \boldsymbol{e}_2\cos\theta, \tag{74}$$

where the angle θ is defined by

$$\theta = \tan^{-1}\left(\frac{\boldsymbol{u}\cdot\boldsymbol{e}_1}{\boldsymbol{u}\cdot\boldsymbol{e}_2}\right). \tag{75}$$

The inverse transformations for Eqs. (73) and (74) are

$$\boldsymbol{e}_1 = \boldsymbol{a}\cos\theta - \boldsymbol{c}\sin\theta, \tag{76}$$

and

$$\boldsymbol{e}_2 = -\boldsymbol{a}\sin\theta - \boldsymbol{c}\cos\theta. \tag{77}$$

Then we have

$$\boldsymbol{u} = u_{\perp}\boldsymbol{c} + u_{\parallel}\boldsymbol{b}, \tag{78}$$

$$\boldsymbol{u}\times\boldsymbol{b} = -u_{\perp}\boldsymbol{a}, \tag{79}$$

and

$$\mathrm{d}\boldsymbol{a} = \boldsymbol{c}\mathrm{d}\theta, \quad \mathrm{d}\boldsymbol{c} = -\boldsymbol{a}\mathrm{d}\theta. \tag{80}$$

Here, let us introduce the guiding center coordinates $(\boldsymbol{R}, u_{\perp}, \theta, u_{\parallel})$ by the transformation

$$\boldsymbol{x} = \boldsymbol{R} + \frac{u_{\perp}}{\Omega}\boldsymbol{a}, \quad \boldsymbol{R} = (X, Y, Z). \tag{81}$$

Then we have

$$\mathrm{d}\boldsymbol{x} = \mathrm{d}\boldsymbol{R} + \frac{\mathrm{d}u_{\perp}}{\Omega}\boldsymbol{a} + \frac{u_{\perp}}{\Omega}\mathrm{d}\theta\boldsymbol{c}, \tag{82}$$

and

$$\mathrm{d}\boldsymbol{u} = \mathrm{d}u_\perp \boldsymbol{c} + \mathrm{d}u_\| \boldsymbol{b} - u_\perp \boldsymbol{a}\mathrm{d}\theta. \tag{83}$$

These are well-known procedures; however, by applying them in Eqs. (69) and (72), we have

$$\begin{aligned}\boldsymbol{q}\cdot\mathrm{d}\boldsymbol{x} + \boldsymbol{y}\cdot\mathrm{d}\boldsymbol{u} = \boldsymbol{q}\cdot\mathrm{d}\boldsymbol{R} &+ y_\|\mathrm{d}u_\| \\ &+ \left(\frac{q_a}{\Omega} + y_c\right)\mathrm{d}u_\perp \\ &+ \left(\frac{q_c}{\Omega} - y_a\right)u_\perp \mathrm{d}\theta,\end{aligned} \tag{84}$$

and

$$\mathcal{H} = u_\| q_\| + \Omega\left(\frac{q_c}{\Omega} - y_a\right)u_\perp, \tag{85}$$

where

$$q_\| = \boldsymbol{q}\cdot\boldsymbol{b}, \quad q_a = \boldsymbol{q}\cdot\boldsymbol{a}, \quad q_c = \boldsymbol{q}\cdot\boldsymbol{c}, \tag{86}$$

and $y_\|$, y_a and y_c are similarly defined. Equations (84) and (85) imply the transformations

$$\frac{q_a}{\Omega} + y_c = \frac{u_\perp}{\Omega}P, \tag{87}$$

and

$$\left(\frac{q_c}{\Omega} - y_a\right)u_\perp = M. \tag{88}$$

These are the transformations $(y_a, y_c) \mapsto (P, M)$

$$y_c = \frac{u_\perp}{\Omega}P - \frac{q_a}{\Omega}, \tag{89}$$

and

$$y_a = -\frac{M}{u_\perp} + \frac{q_c}{\Omega}, \tag{90}$$

and we get a canonical 1-form

$$\gamma = \boldsymbol{q}\cdot\mathrm{d}\boldsymbol{R} + y_\|\mathrm{d}u_\| + P\mathrm{d}\mu + M\mathrm{d}\theta - \mathcal{H}\mathrm{d}t, \tag{91}$$

where

$$\mu = \frac{u_\perp^2}{2\Omega} \tag{92}$$

is the magnetic moment, and the Hamiltonian is given by

$$\mathcal{H} = u_\| q_\| + \Omega M. \tag{93}$$

Consequently, the transformation from $(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{q}, \boldsymbol{y})$ to $(\boldsymbol{R}, u_\|, \mu, \theta, \boldsymbol{q}, y_\|, P, M)$ is a canonical transformation that has a trivially soluble Hamiltonian; the equations are

$$\frac{\mathrm{d}\boldsymbol{R}}{\mathrm{d}t} = \frac{\partial\mathcal{H}}{\partial\boldsymbol{q}} = u_\|\boldsymbol{b}, \tag{94}$$

$$\frac{\mathrm{d}u_\|}{\mathrm{d}t} = \frac{\partial\mathcal{H}}{\partial y_\|} = 0, \tag{95}$$

$$\frac{\mathrm{d}\mu}{\mathrm{d}t} = \frac{\partial\mathcal{H}}{\partial P} = 0, \tag{96}$$

and

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \frac{\partial\mathcal{H}}{\partial M} = \Omega. \tag{97}$$

These results again demonstrate the effectiveness of the conjugate variable method.

3.2 Canonical perturbation analysis of the guiding center motion

We adopt a simple model where the magnetic fields consists of

$$\boldsymbol{B} = B\boldsymbol{b} + \epsilon\tilde{\boldsymbol{B}}(\boldsymbol{x}), \tag{98}$$

and $\tilde{\boldsymbol{B}}$ is assumed to be independent of z and perpendicular to $\boldsymbol{b}$,

$$\tilde{\boldsymbol{B}}(\boldsymbol{x})\cdot\boldsymbol{b} = 0. \tag{99}$$

Following Ref. [8], we define the transformation from the Cartesian coordinate system to the guiding center coordinate system $(\boldsymbol{R}, u_\|, \mu, \theta, \boldsymbol{q}, y_\|, P, M)$ using the uniform field; the transformation is given by Eq. (81), and the vectors $\boldsymbol{a}$ and $\boldsymbol{c}$ are defined by Eqs. (73) and (74). Consequently the 1-form remains Eq. (91), whereas the Hamiltonian consists of

$$\mathcal{H} = \mathcal{H}^{(0)} + h, \tag{100}$$

where the unperturbed Hamiltonian is

$$\mathcal{H}^{(0)} = u_\| q_\| + \Omega M, \tag{101}$$

and the perturbed Hamiltonian is given by

$$h = \epsilon\Omega\boldsymbol{y}\cdot(\boldsymbol{u}\times\tilde{\boldsymbol{B}}(\boldsymbol{x})). \tag{102}$$

Here, the perturbed magnetic field is normalized to the uniform magnetic field B.

Let L_B be the characteristic length over which the perturbed magnetic field varies, and assume that the Larmor radius ρ is much shorter than L_B such that

$$\frac{\rho}{L_B} = O(\epsilon), \tag{103}$$

and we expand $\tilde{\boldsymbol{B}}(\boldsymbol{x})$ as

$$\tilde{\boldsymbol{B}}(\boldsymbol{x}) = \tilde{\boldsymbol{B}}(\boldsymbol{R}) + \epsilon\frac{u_\perp}{\Omega}(\boldsymbol{a}\cdot\nabla)\tilde{\boldsymbol{B}}(\boldsymbol{R}). \tag{104}$$

Consequently, we have

$$h = \epsilon h^{(1)} + \epsilon^2 h^{(2)}, \tag{105}$$

where the first-order Hamiltonian reads

$$h^{(1)} = \Omega\boldsymbol{y}\cdot(\boldsymbol{u}\times\tilde{\boldsymbol{B}}(\boldsymbol{R})), \tag{106}$$

and the second-order Hamiltonian reads

$$h^{(2)} = u_\perp\boldsymbol{y}\cdot\left[\boldsymbol{u}\times(\boldsymbol{a}\cdot\nabla)\tilde{\boldsymbol{B}}(\boldsymbol{R})\right]. \tag{107}$$

Based on the assumption, the perturbed field is written as

$$\tilde{\boldsymbol{B}}(\boldsymbol{R}) = \tilde{B}_x(X,Y)\boldsymbol{e}_1 + \tilde{B}_y(X,Y)\boldsymbol{e}_2. \tag{108}$$

3.3 First-order perturbation analysis

The perturbed Hamiltonian $h^{(1)}$ can be easily expressed by the canonical coordinates $(\boldsymbol{R}, u_\parallel, \mu, \theta, \boldsymbol{q}, y_\parallel, P, M)$,

$$h^{(1)} = \Omega u_\perp \left(y_\parallel - \frac{P}{\Omega}u_\parallel\right)\tilde{B}_a(\boldsymbol{R},\theta) - M\Omega\frac{u_\parallel}{u_\perp}\tilde{B}_c(\boldsymbol{R},\theta) + u_\parallel(q_1\tilde{B}_x(\boldsymbol{R}) + q_2\tilde{B}_y(\boldsymbol{R})) \tag{109}$$

$$= \tilde{h}^{(1)} + \langle h^{(1)}\rangle, \tag{110}$$

and

$$\langle h^{(1)}\rangle = u_\parallel(q_1\tilde{B}_x(\boldsymbol{R}) + q_2\tilde{B}_y(\boldsymbol{R})). \tag{111}$$

Here,

$$\tilde{B}_a(\boldsymbol{R},\theta) = \tilde{B}_x(\boldsymbol{R})\cos\theta - \tilde{B}_y(\boldsymbol{R})\sin\theta, \tag{112}$$

and

$$\tilde{B}_c(\boldsymbol{R},\theta) = -\tilde{B}_x(\boldsymbol{R})\sin\theta - \tilde{B}_y(\boldsymbol{R})\cos\theta. \tag{113}$$

These mean

$$\tilde{\boldsymbol{B}}(\boldsymbol{R}) = \tilde{B}_a(\boldsymbol{R},\theta)\boldsymbol{a} + \tilde{B}_c(\boldsymbol{R},\theta)\boldsymbol{c}. \tag{114}$$

Let us make the equations of motion from $\mathcal{H} = \mathcal{H}^{(0)} + \epsilon h^{(1)}$. They are

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \frac{\partial\mathcal{H}}{\partial q_1} = \epsilon u_\parallel\tilde{B}_x, \tag{115}$$

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = \frac{\partial\mathcal{H}}{\partial q_2} = \epsilon u_\parallel\tilde{B}_y, \tag{116}$$

$$\frac{\mathrm{d}Z}{\mathrm{d}t} = \frac{\partial\mathcal{H}}{\partial q_\parallel} = u_\parallel, \tag{117}$$

$$\frac{\mathrm{d}u_\parallel}{\mathrm{d}t} = \frac{\partial\mathcal{H}}{\partial y_\parallel} = \epsilon\Omega u_\perp\tilde{B}_a = \epsilon\Omega(\boldsymbol{u}\times\tilde{\boldsymbol{B}})\cdot\boldsymbol{b}, \tag{118}$$

$$\frac{\mathrm{d}\mu}{\mathrm{d}t} = \frac{\partial\mathcal{H}}{\partial P} = -\epsilon u_\perp u_\parallel\tilde{B}_a, \tag{119}$$

and

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \frac{\partial\mathcal{H}}{\partial M} = \Omega\left\{1 - \epsilon\frac{u_\parallel}{u_\perp}\tilde{B}_c\right\}. \tag{120}$$

The first three equations express the guiding center's motion along the magnetic field lines, and the other three equations are identical to the equation

$$\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} = \Omega\boldsymbol{u}\times(\boldsymbol{b} + \epsilon\tilde{\boldsymbol{B}}) \tag{121}$$

expressed in terms of $u_\parallel$, μ and θ.

The Poisson bracket in Eq. (A9) is, for the present problem,

$$\{S^{(1)}, \mathcal{H}^{(0)}\} = u_\parallel\frac{\partial S^{(1)}}{\partial Z} + \Omega\frac{\partial S^{(1)}}{\partial\theta} - q_\parallel\frac{\partial S^{(1)}}{\partial y_\parallel}. \tag{122}$$

Therefore, we obtain an equation for $S^{(1)}$ that yields the Hamiltonian averaged over θ

$$\Omega\frac{\partial S^{(1)}}{\partial\theta} - q_\parallel\frac{\partial S^{(1)}}{\partial y_\parallel} = \tilde{h}^{(1)}. \tag{123}$$

Here, we used the assumption that $S^{(1)}$ is independent of both time and Z. Equation (123) gives the solution

$$S^{(1)} = -\left(q_\parallel\frac{u_\perp}{\Omega} + M\frac{u_\parallel}{u_\perp}\right)\tilde{B}_a(\boldsymbol{R},\theta) - u_\perp\left(y_\parallel - \frac{u_\parallel}{\Omega}P\right)\tilde{B}_c(\boldsymbol{R},\theta). \tag{124}$$

The averaged Hamiltonian is therefore

$$H = \mathcal{H}^{(0)} + \epsilon\langle h^{(1)}\rangle, \tag{125}$$

and the equations of motion are

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \epsilon u_\parallel\tilde{B}_x, \tag{126}$$

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = \epsilon u_\parallel\tilde{B}_y, \tag{127}$$

$$\frac{\mathrm{d}Z}{\mathrm{d}t} = u_\parallel, \tag{128}$$

$$\frac{\mathrm{d}u_\parallel}{\mathrm{d}t} = \frac{\partial H}{\partial y_\parallel} = 0, \tag{129}$$

$$\frac{\mathrm{d}\mu}{\mathrm{d}t} = \frac{\partial H}{\partial P} = 0, \tag{130}$$

and

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \frac{\partial H}{\partial M} = \Omega. \tag{131}$$

Since the Hamiltonian H does not include the conjugate variables $y_\parallel$ and P, the corresponding variables $u_\parallel$ and μ (and consequently the kinetic energy $\mathcal{E} = \Omega\mu + u_\parallel^2/2$) are conserved.

3.4 Second-order perturbation analysis

When we consider the averaged Hamiltonian, Eq. (A22) is reduced to

$$\langle\xi_0^{(2)}\rangle = -\langle h^{(2)}\rangle + \frac{1}{2}\langle\{S^{(1)}, \tilde{h}^{(1)}\}\rangle, \tag{132}$$

since $\{S^{(1)}, \langle h^{(1)}\rangle\}$ is periodic with respect to θ, and consequently its average vanishes. Our work is then reduced to calculating $\langle\{S^{(1)}, \tilde{h}^{(1)}\}\rangle$ and $\langle h^{(2)}\rangle$. For calculating $\langle\{S^{(1)}, \tilde{h}^{(1)}\}\rangle$, it is convenient to express $\tilde{h}^{(1)}$ and $S^{(1)}$ as

$$\tilde{h}^{(1)} = h_a\tilde{B}_a + h_c\tilde{B}_c, \tag{133}$$

and

$$S^{(1)} = S_a\tilde{B}_a + S_c\tilde{B}_c. \tag{134}$$

Here

$$h_a = \Omega u_\perp\left(y_\parallel - \frac{P}{\Omega}u_\parallel\right), \tag{135}$$

$$h_c = -M\Omega\frac{u_\parallel}{u_\perp}, \tag{136}$$

$$S_a = -\left(q_\parallel\frac{u_\perp}{\Omega} + M\frac{u_\parallel}{u_\perp}\right) = -q_\parallel\frac{u_\perp}{\Omega} + \frac{h_c}{\Omega}, \tag{137}$$

and

$$S_c = -u_\perp\left(y_\parallel - \frac{u_\parallel}{\Omega}P\right) = \frac{h_a}{\Omega}. \tag{138}$$

Using Eqs. (133) and (134), we have

$$\begin{aligned}\{S^{(1)}, \tilde{h}^{(1)}\} = \{S_a\tilde{B}_a, h_a\tilde{B}_a\} + \{S_a\tilde{B}_a, h_c\tilde{B}_c\} \\ + \{S_c\tilde{B}_c, h_a\tilde{B}_a\} + \{S_c\tilde{B}_c, h_c\tilde{B}_c\}.\end{aligned} \tag{139}$$

Averaging over θ operates on $\tilde{B}_a$ and $\tilde{B}_c$, resulting in

$$\langle(\tilde{B}_a)^2\rangle = \langle(\tilde{B}_c)^2\rangle = \frac{1}{2}|\tilde{\boldsymbol{B}}|^2, \tag{140}$$

and

$$\langle\tilde{B}_a\tilde{B}_c\rangle = 0. \tag{141}$$

After lengthy algebraic computation for the right-hand side of Eq. (139), the details of which are shown in Appendix B, we have

$$\langle\{S^{(1)}, \tilde{h}^{(1)}\}\rangle = \frac{1}{2}|\tilde{\boldsymbol{B}}|^2\left[2q_\parallel u_\parallel - \frac{M}{\mu}(u_\parallel)^2\right]. \tag{142}$$

Now let us rewrite $h^{(2)}$ as

$$h^{(2)} = u_\perp\boldsymbol{y}\cdot[\boldsymbol{u}\times\boldsymbol{W}]. \tag{143}$$

Here,

$$\begin{aligned}\boldsymbol{W} &= (\boldsymbol{a}\cdot\nabla)\tilde{\boldsymbol{B}}(\boldsymbol{R}) \\ &= W_1\boldsymbol{e}_1 + W_2\boldsymbol{e}_2 &(144)\\ &= W_a\boldsymbol{a} + W_c\boldsymbol{c}, &(145)\\ W_1 &= (\cos\theta\partial_x - \sin\theta\partial_y)\tilde{B}_x(\boldsymbol{R}), &(146)\\ W_2 &= (\cos\theta\partial_x - \sin\theta\partial_y)\tilde{B}_y(\boldsymbol{R}), &(147)\\ W_a &= W_1\cos\theta - W_2\sin\theta, &(148)\end{aligned}$$

and

$$W_c = -W_1\sin\theta - W_2\cos\theta. \tag{149}$$

Averaging them over θ gives

$$\langle W_1\rangle = \langle W_2\rangle = 0, \tag{150}$$

$$\langle W_a\rangle = \frac{1}{2}(\partial_x\tilde{B}_x + \partial_y\tilde{B}_y) = 0, \tag{151}$$

and

$$\langle W_c\rangle = \frac{1}{2}(\partial_y\tilde{B}_x - \partial_x\tilde{B}_y) = -\frac{1}{2}J_z, \tag{152}$$

where J_z is the current density

$$J_z = \partial_x\tilde{B}_y - \partial_y\tilde{B}_x. \tag{153}$$

Applying Eqs. (148) and (149) to Eq. (143), we have

$$\boldsymbol{u}\times\boldsymbol{W} = u_\perp W_a\boldsymbol{b} - u_\parallel W_a\boldsymbol{c} + u_\parallel W_c\boldsymbol{a},$$

and then

$$h^{(2)} = y_\parallel(u_\perp)^2W_a + u_\perp u_\parallel(y_aW_c - y_cW_a). \tag{154}$$

Applying the transformation formula, Eqs. (89) and (90), for Eq. (154) yields $h^{(2)}$ expressed in the phase space coordinates $(\boldsymbol{R}, u_\parallel, \mu, \theta, \boldsymbol{q}, y_\parallel, P, M)$

$$\begin{aligned}h^{(2)} = \frac{(u_\perp)^2}{\Omega}W_a(\Omega y_\parallel - Pu_\parallel) \\ - Mu_\parallel W_c + \frac{u_\perp}{\Omega}u_\parallel(q_1W_1 + q_2W_2),\end{aligned} \tag{155}$$

and averaging it gives

$$\langle h^{(2)}\rangle = -Mu_\parallel\langle W_c\rangle = \frac{Mu_\parallel}{2}J_z. \tag{156}$$

From Eqs. (142) and (156), we obtain the averaged second-order Hamiltonian

$$\begin{aligned}\langle H^{(2)}\rangle &= -\langle\xi_0^{(2)}\rangle = \langle h^{(2)}\rangle - \frac{1}{2}\langle\{S^{(1)}, \tilde{h}^{(1)}\}\rangle \\ &= \frac{Mu_\parallel}{2}J_z - \frac{1}{2}|\tilde{\boldsymbol{B}}|^2\left[q_\parallel u_\parallel - \frac{M}{2\mu}(u_\parallel)^2\right],\end{aligned} \tag{157}$$

and then the total Hamiltonian

$$H = H^{(0)} + \epsilon\langle H^{(1)}\rangle + \epsilon^2\langle H^{(2)}\rangle, \tag{158}$$

where

$$H^{(0)} = q_\parallel u_\parallel + \Omega M, \tag{159}$$

and

$$\langle H^{(1)}\rangle = u_\parallel(q_1\tilde{B}_x(\boldsymbol{R}) + q_2\tilde{B}_y(\boldsymbol{R})). \tag{160}$$

Again, we can observe that the Hamiltonian H retains linearity in the conjugate variables $(\boldsymbol{q}, y_\parallel, P, M)$. Consequently, the equations of motion for $(\boldsymbol{R}, u_\parallel, \mu, \theta)$ do not contain any conjugate variables; they are

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \frac{\partial H}{\partial q_1} = \epsilon u_\parallel\tilde{B}_x(\boldsymbol{R}), \tag{161}$$

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = \frac{\partial H}{\partial q_2} = \epsilon u_\parallel\tilde{B}_y(\boldsymbol{R}), \tag{162}$$

$$\frac{\mathrm{d}Z}{\mathrm{d}t} = \frac{\partial H}{\partial q_\parallel} = u_\parallel\left[1 - \frac{\epsilon^2}{2}|\tilde{\boldsymbol{B}}|^2\right], \tag{163}$$

$$\frac{\mathrm{d}u_\parallel}{\mathrm{d}t} = \frac{\partial H}{\partial y_\parallel} = 0, \tag{164}$$

$$\frac{\mathrm{d}\mu}{\mathrm{d}t} = \frac{\partial H}{\partial P} = 0, \tag{165}$$

and

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \frac{\partial H}{\partial M} = \Omega + \epsilon^2\left[\frac{J_z}{2}u_\parallel + \frac{(u_\parallel)^2}{2\mu}|\tilde{\boldsymbol{B}}|^2\right]. \tag{166}$$

It is important to study whether the conjugate variable method gives the same results as those of the usual canonical (or non-canonical) Hamilton-Lie perturbation method or the multi-scale expansion method. However, this would require rigorous investigation with deep mathematical analysis, and therefore remains open to future study.

4. Summary

It has been shown that the conjugate variable method enables us to apply the Hamilton-Lie perturbation method systematically to any system of ordinary differential equations that does not have the Hamiltonian dynamic structure. Although the method doubles the unknown variables, the equations for the original variables are derived without mixing with the conjugate variables at each order in the canonical perturbation analysis.

Several problems are open to future study. One important problem is to determine the equivalence of the method proposed in the present work with other methods, that is, whether the results by the conjugate variable method are the same as those by other methods, such as the usual canonical (or non-canonical) perturbation method, and the classical multi-scale expansion method. Another interesting problem in plasma physics is to apply the present method to eigenvalue equations related to MHD stability analysis, such as the MHD ballooning mode equation. Studies on these problems will be reported in the near future.

Acknowledgment

The author thanks Dr. M. Kikuchi, the Unit Manager of the Division of Advanced Plasma Research, Fusion Research and Development Directorate at Japan Atomic Energy Agency, for his encouragement.

Appendix A: Summary of Canonical Hamilton-Lie Perturbation Analysis

Following Ref. [2], let a 1-form be given by

$$\gamma = \gamma^{(0)} + \epsilon\gamma^{(1)} + \epsilon^2\gamma^{(2)} + \cdots, \tag{A1}$$

in a phase space $(z^1, \cdots, z^{2N})$ with $z^0 := t$, and let $\gamma^{(0)}$ be integrable. Recall that, with respect to the generators g^ν for the Lie transformation, the Lie derivative $\mathcal{L}_g$ for a 1-form γ is defined by

$$(\mathcal{L}_g\gamma)_\mu = g^\nu\omega_{\nu\mu}, \quad \mu, \nu = 0, \cdots, 2N. \tag{A2}$$

Here

$$\omega_{\nu\mu} = \partial_\nu\gamma_\mu - \partial_\mu\gamma_\nu \tag{A3}$$

is the Lagrangian tensor; the inverse of it, which is expressed by J^{kj}, is the Poisson tensor. The 1-form Γ after the Lie transformation

$$\Gamma = \gamma^{(0)} + \epsilon\Gamma^{(1)} + \epsilon^2\Gamma^{(2)} + \cdots, \tag{A4}$$

reads

$$\Gamma^{(1)} = -\mathcal{L}_{(1)}\gamma^{(0)} + \mathrm{d}S^{(1)} + \gamma^{(1)}, \tag{A5}$$

and

$$\Gamma^{(2)} = -\mathcal{L}_{(2)}\gamma^{(0)} + \mathrm{d}S^{(2)} + \xi^{(2)}, \tag{A6}$$

where $\xi^{(2)}$ is given by

$$\xi^{(2)} = \gamma^{(2)} - \mathcal{L}_{(1)}\gamma^{(1)} + \frac{1}{2}(\mathcal{L}_{(1)})^2\gamma^{(0)}. \tag{A7}$$

In this appendix, we give the expression for $\xi^{(2)}$ assuming that the 1-form is canonical. In the following, let $h^{(j)}$ be the Hamiltonian in the corresponding form $\gamma^{(j)}$.

In the first-order canonical perturbation analysis, the generators $g^j_{(1)}$ for space variables are given by

$$g^j_{(1)} = \partial_k S^{(1)} J^{kj}, \tag{A8}$$

where $S^{(1)}$ is the gauge function. The transformed Hamiltonian $H^{(1)}$ reads

$$-H^{(1)} = \partial_t S^{(1)} + \{S^{(1)}, h^{(0)}\} - h^{(1)}, \tag{A9}$$

where the Poisson bracket $\{S^{(1)}, h^{(0)}\}$ is defined by

$$\{S^{(1)}, h^{(0)}\} = \partial_k S^{(1)} J^{kj}\partial_j h^{(0)}. \tag{A10}$$

In Hamilton-Lie perturbation analysis, the gauge function $S^{(1)}$ is determined such that the Hamiltonian $H^{(1)}$ becomes simple. It was found that $\xi^{(2)}$ in Eq. (A7) can be expressed in a simple form, since $\gamma^{(1)}$ and $\gamma^{(2)}$ are canonical.

We start with the calculation of $\mathcal{L}_{(1)}\gamma^{(1)}$ in Eq. (A7). Let us tentatively write the gauge function defined according to the first-order perturbation analysis, Eq. (A9), as ϕ. Then the generators read

$$g^j = \partial_k\phi\, J^{kj}\ (j \neq 0), \tag{A11}$$

and the generator for the time is $g^0 = 0$. The Lie derivative of $\gamma^{(1)}$ with respect to g^j results in

$$[\mathcal{L}_\phi\gamma^{(1)}]_j = 0, \tag{A12}$$

and

$$[\mathcal{L}_\phi\gamma^{(1)}]_0 = -\partial_k\phi J^{kj}\partial_j h^{(1)}. \tag{A13}$$

Consequently, we have

$$\mathcal{L}_\phi\gamma^{(1)} = -\{\phi, h^{(1)}\}\mathrm{d}t. \tag{A14}$$

By calculation of $(\mathcal{L}_{(1)})^2\gamma^{(0)}$ in Eq. (A7) using a procedure similar to that used for Eq. (A14), we have

$$\Omega = \mathcal{L}_\phi\gamma^{(0)} = \partial_j\phi\, dz^j - \{\phi, h^{(0)}\}\mathrm{d}t. \tag{A15}$$

Next, let $\bar{g}^j = \partial_k\psi J^{kj}$ be generators defined by another function ψ, and operate $\mathcal{L}_\psi$ on Ω. Then we have

$$\mathcal{L}_\psi\mathcal{L}_\phi\gamma^{(0)} = -\{\psi, \{\phi, h^{(0)}\}\}\mathrm{d}t. \tag{A16}$$

By applying Eq. (A16) in Eq. (A7), we have

$$(\mathcal{L}_\phi)^2\gamma_0 = -\{\phi, \{\phi, h^{(0)}\}\}\mathrm{d}t. \tag{A17}$$

From Eqs. (A14) and (A17), and from the assumption that $\gamma^{(2)} = -h^{(2)}\mathrm{d}t$, we observe that $\xi^{(2)}$ is canonical, and obtain the second-order Hamiltonian

$$\xi^{(2)}_0 = -h^{(2)} + \{S^{(1)}, h^{(1)}\} - \frac{1}{2}\{S^{(1)}, \{S^{(1)}, h^{(0)}\}\}. \tag{A18}$$

Here, let us return to Eq. (30). The gauge function $S^{(1)}$ is determined such that

$$\{S^{(1)}, h^{(0)}\} = \tilde{h}^{(1)}. \tag{A19}$$

We have therefore in Eq. (A18)

$$\{S^{(1)}, h^{(1)}\} - \frac{1}{2}\{S^{(1)}, \{S^{(1)}, h^{(0)}\}\} = \{S^{(1)}, \langle h^{(1)}\rangle\} + \frac{1}{2}\{S^{(1)}, \tilde{h}^{(1)}\}. \tag{A20}$$

The second-order Hamiltonian is

$$-H^{(2)} = \{S^{(2)}, h^{(0)}\} + \xi_0^{(2)}, \tag{A21}$$

and

$$\xi_0^{(2)} = -h^{(2)} + \{S^{(1)}, \langle h^{(1)}\rangle\} + \frac{1}{2}\{S^{(1)}, \tilde{h}^{(1)}\}. \tag{A22}$$

We can apply the same procedure as the first-order analysis for Eq. (A21) to determine the gauge function $S^{(2)}$ that makes the transformed Hamiltonian $H^{(2)}$ as simple as possible. From $S^{(1)}$ and $S^{(2)}$ we get the generators $g^j_{(1)}$ and $g^j_{(2)}$, and we can construct the successive Lie transformation from $(\bar{z}^\mu)$ to (z^μ)

$$z^\mu = \exp(-\epsilon^2\mathcal{L}_{(2)})\exp(-\epsilon\mathcal{L}_{(1)})I^\mu(\bar{z}), \tag{A23}$$

where $I^\mu(\bar{z}) = \bar{z}^\mu$ is the coordinate function. By expanding the right-hand side to $O(\epsilon^2)$, we have

$$z^j = \bar{z}^j - \epsilon g^j_{(1)}(\bar{z}^\mu) - \epsilon^2 g^j_{(2)}(\bar{z}^\mu) + \frac{\epsilon^2}{2}[(g^\nu_{(1)}\partial_\nu)g^j_{(1)}](\bar{z}^\mu). \tag{A24}$$

When we substitute $\bar{z}^j(t)$, the solutions of the second-order analysis, into the right-hand side of Eq. (A24), we obtain the solution $z^j(t)$ in the original phase space. Note that when $S^{(1)}$ is used for the fourth term of Eq. (A24), we have

$$(g^\nu_{(1)}\partial_\nu)g^j_{(1)} = \partial_\mu S^{(1)}\partial_\nu\partial_\sigma S^{(1)}J^{\mu\nu}J^{\sigma j}. \tag{A25}$$

Appendix B: Calculation of $\{S^{(1)}, \tilde{h}^{(1)}\}$

By calculating the Poisson bracket in the right-hand side of Eq. (139) and using Eqs. (140) and (141), we obtain

$$\langle\{S_a\tilde{B}_a, h_a\tilde{B}_a\}\rangle = \frac{1}{2}|\tilde{\boldsymbol{B}}|^2\{S_a, h_a\}, \tag{B1}$$

$$\langle\{S_a\tilde{B}_a, h_c\tilde{B}_c\}\rangle = \frac{1}{2}|\tilde{\boldsymbol{B}}|^2\partial_M(S_a h_c), \tag{B2}$$

$$\langle\{S_c\tilde{B}_c, h_a\tilde{B}_a\}\rangle = \langle\tilde{B}_c\tilde{B}_a\rangle\{S_c, h_a\} = 0, \tag{B3}$$

and

$$\langle\{S_c\tilde{B}_c, h_c\tilde{B}_c\}\rangle = \frac{1}{2}|\tilde{\boldsymbol{B}}|^2\{S_c, h_c\}. \tag{B4}$$

By adding Eqs. (B1), (B2), and (B4), we get

$$\langle\{S^{(1)}, \tilde{h}^{(1)}\}\rangle = \frac{1}{2}|\tilde{\boldsymbol{B}}|^2\left[\{S_a, h_a\} + \{S_c, h_c\} + \partial_M(S_a h_c)\right]. \tag{B5}$$

Next, from Eqs. (136) and (138), we get

$$\{S_a, h_a\} = -\frac{1}{\Omega}\{q_\parallel u_\perp, h_a\} + \frac{1}{\Omega}\{h_c, h_a\},$$

and

$$\{S_c, h_c\} = \frac{1}{\Omega}\{h_a, h_c\}.$$

Then Eq. (B5) is further reduced to

$$\langle\{S^{(1)}, \tilde{h}^{(1)}\}\rangle = \frac{|\tilde{\boldsymbol{B}}|^2}{2}\left[-\frac{\{q_\parallel u_\perp, h_a\}}{\Omega} + \partial_M(S_a h_c)\right]. \tag{B6}$$

Finally, by substituting

$$\{q_\parallel u_\perp, h_a\} = -\Omega q_\parallel u_\parallel, \tag{B7}$$

and

$$\partial_M(S_a h_c) = q_\parallel u_\parallel - \frac{(u_\parallel)^2}{\mu}M, \tag{B8}$$

in Eq. (B6), we obtain Eq. (142).

[1] E.C.G. Sudarshan and N. Mukunda, *Classical Dynamics: A Modern Perspective* (Rober E. Krieger Publishing Company, Florida, U.S.A, 1983).

[2] J.R. Cary and R.G. Littlejohn, Ann. Phys. **151**, 1 (1983).

[3] C.M. Bender and S.A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw-Hill, New York, U.S.A, 1978) Chap.11.

[4] B. Jouvet and R. Phytian, Phys. Rev. A **19**, 1350 (1979).

[5] T.G. Northrop, *The Adiabatic Motion of Charged Particles* (Interscience, New York, U.S.A, 1963).

[6] A.I. Morozov and L.S. Solov'ev, Motion of Charged Particles in Electromagnetic Fields, in *Reviews of Plasma Physics* Vol.2, 201-297, Edited by M.A. Leontovich, Consultants Bureau, New York (1966).

[7] R.G. Littlejohn, J. Plasma Phys. **29**, 111 (1983).

[8] T.S. Hahm, W.W. Lee and A. Brizard, Phys. Fluids **31**, 1940 (1988).