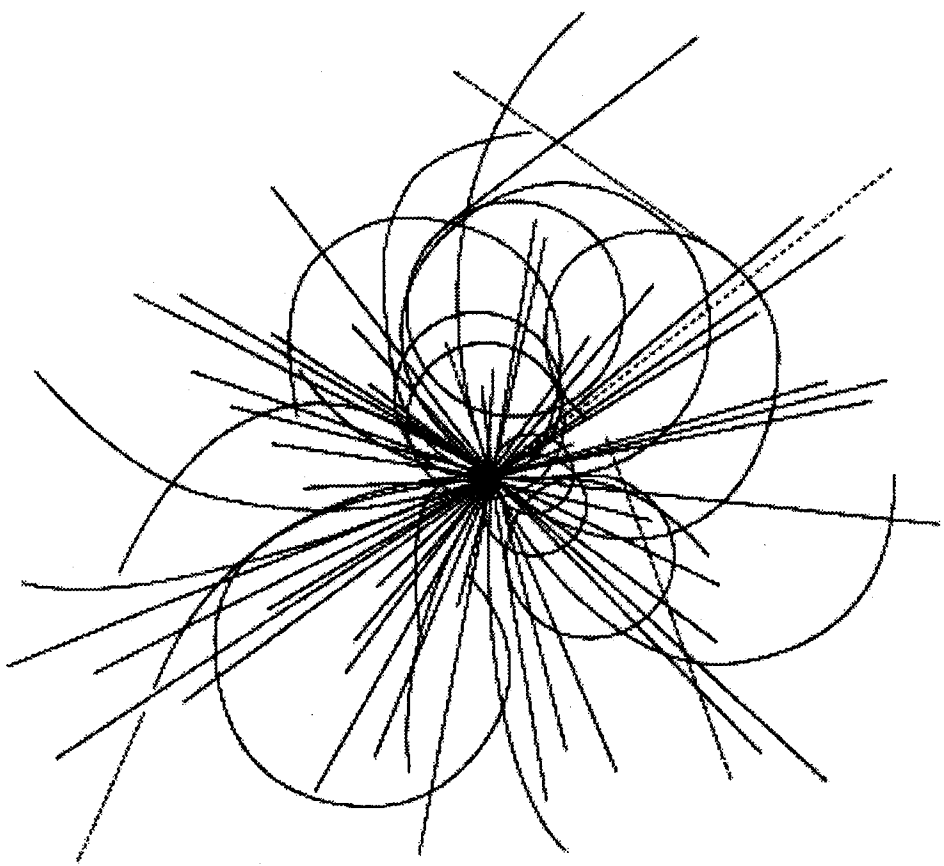
SSCL-Preprint-570

SSCL-Preprint-570
August 1994
Distribution Category:

J. Ellison
H.-J. Shih

The Method of Averaging in Beam Dynamics



Superconducting Super Collider Laboratory

SSCL-Preprint-570

The Method of Averaging in Beam Dynamics*

J. Ellison† and H.-J. Shih

Superconducting Super Collider Laboratory‡
2550 Beckleymeade Avenue
Dallas, Texas 75237

August 1994

* To be published in *Accelerator Physics at the Superconducting Super Collider.*

† Also with the Department of Mathematics, University of New Mexico, Albuquerque, New Mexico 87131.

‡ Operated by the Universities Research Association, Inc., for the U.S. Department of Energy under Contract No. DE-AC35-89ER40486.

THE METHOD OF AVERAGING IN BEAM DYNAMICS*

JAMES A. ELLISON† and H.-JENG SHIH
Superconducting Super Collider Laboratory,‡ 2550 Beckleymeade Avenue
Dallas, Texas 75237

ABSTRACT

This paper introduces the method of averaging and applies it to several beam dynamics problems. Averaging is an important tool in the rigorous study of ODEs with a small parameter, and it leads to a systematic perturbation expansion complete with error bounds. First- and second-order averaging theorems are presented and applied; in addition, proofs are included for the interested reader. Resonance at first and second order is treated.

1. Introduction

Perturbation theory is an important tool in beam dynamics. Regular perturbation theory is not satisfactory because of the existence of secular terms. Secular terms first arose in celestial mechanics, and this led to long time perturbation techniques that eliminate secular terms and are valid on longer time intervals. Examples are the method of averaging, multiple time scales,[1] and canonical perturbation methods[2,3] (mixed generating function and Lie transformation). Here we introduce the method of averaging and show how it applies in several not-too-complicated examples of beam dynamics, both with and without resonance.

The method of averaging is an important tool in the rigorous study of differential equations with a small parameter. Since the publication of the now classic book by Bogoliubov and Mitropolskii,[4] the literature on averaging has grown immensely. A good introduction can be found in Murdock.[1] For additional background and results, the interested reader is referred to the books by Sanders and Verhulst[5] and by Lochak and Meunier.[6]

At its heart, averaging is a transformation procedure leading to a systematic perturbation expansion complete with error bounds on the difference between exact and approximate solutions. In addition, it is a tool for proving properties of the exact problem based on properties of the approximate problem; for example, existence of periodic solutions can be proved using averaging together with the implicit function theorem, and the existence of invariant tori can be proved using averaging together with the Moser twist theorem. In addition, averaging is very robust as, for example, it applies to both Hamiltonian and dissipative systems.

In Section 2, we discuss both regular perturbation theory and the method of averaging. We begin by discussing the Duffing equation and show how secular terms arise in regular perturbation theory and show that the method of averaging circumvents this problem. We then prove a first-order regular perturbation theorem and a first-order averaging theorem, which shows that averaging gives results on longer time intervals than regular perturbation results. These theorems make precise the statement that averaging is a long time perturbation theory. Basically,

---

*To be published in *Accelerator Physics at the Superconducting Super Collider.*

†Also with the Department of Mathematics, University of New Mexico, Albuquerque, New Mexico 87131.

‡Operated by the Universities Research Association, Inc., for the U.S. Department of Energy under Contract No. DE-AC35-89ER40486.

long time perturbation theories give results on time intervals of order $1/\epsilon$, where ϵ is the size of the perturbation. In Section 3, we discuss the basic equations of transverse beam dynamics and show how the perturbed transverse beam dynamics equations can be transformed into a standard form for the method of averaging. We also discuss how a simple resonance at the first order can be treated using averaging.

In Section 4, we discuss several beam dynamics problems in the context of first-order averaging: (A) a special class of perturbations that includes chromaticity, sextupole, octupole, and beam-beam perturbations; (B) 1-D sextupole near the 1:3 resonance; (C) the beam-beam with x-y coupling; and (D) rf phase modulation with damping. In Section 5, we present a second-order averaging theorem that improves the approximation on $O(1/\epsilon)$ time intervals. We also present a theorem that extends the time interval of validity to $O(1/\epsilon^2)$ in a special case, and we discuss resonances that appear at second order. In Section 6, we discuss several examples in the context of second-order averaging: (E) 1-D sextupole (nonresonant); (F) 1-D sextupole near the 1:4 resonance; (G) 2-D sextupole with x-y coupling (nonresonant); and (H) 1-D sextupole with dipole ripple.

In Section 7, we make some concluding remarks and discuss extensions of the averaging method. The extensions include higher-order averaging results, averaging theorems for stochastic perturbations, and adiabatic invariant results in both the deterministic and stochastic cases. We believe these extensions will prove useful in the study of beam dynamics, but we do not present examples here.

2. First-Order Perturbation Theory

A standard test problem in perturbation theory is the Duffing oscillator:

$$\ddot{y} + y + \epsilon y^3 = 0\,, \qquad y(0) = y_0\,, \qquad \dot{y}(0) = \dot{y}_0\,, \tag{2.1}$$

where $0 \le \epsilon \ll 1$. This has a conservation law:

$$E = \frac{1}{2}\dot{y}^2 + \frac{1}{2}y^2 + \frac{1}{4}\epsilon y^4\,, \tag{2.2}$$

and since $\epsilon \ge 0$, (2.2) defines a closed curve in the phase space for each $E > 0$, and this implies that all solutions of (2.1) are periodic. Regular perturbation theory assumes an expansion of the solution of (2.1) of the form

$$y(t,\epsilon) = y^{(0)}(t) + \epsilon y^{(1)}(t) + \ldots + \epsilon^n y^{(n)}(t) + \ldots\,. \tag{2.3}$$

Inserting (2.3) into (2.1) and equating powers of ϵ yields

$$O(1) : \ddot{y}^{(0)} + y^{(0)} = 0\,, \qquad y^{(0)}(0) = y_0,\ \dot{y}^{(0)}(0) = \dot{y}_0\,, \tag{2.4a}$$

$$O(\epsilon) : \ddot{y}^{(1)} + y^{(1)} = -y^{(0)^3}\,, \qquad y^{(1)}(0) = \dot{y}^{(1)}(0) = 0\,, \tag{2.4b}$$

and so on. In fact, at every order $n \ge 1$ we obtain

$$\ddot{y}^{(n)} + y^{(n)} = f_n\big(y^{(0)}(t), \ldots, y^{(n-1)}(t)\big), \qquad y^{(n)}(0) = \dot{y}^{(n)}(0) = 0\,, \tag{2.4c}$$

which is a *linear* nonhomogeneous equation that can be reduced to quadratures. Thus, the regular perturbation method is a linearization as it reduces the nonlinear problem (2.1) to a sequence of linear problems. Solving (2.4a) and then (2.4b) with $\dot{y}_0 = 0$ yields

$$y(t,\epsilon) = y_0 \cos t + \epsilon y_0^3 \left(-\frac{1}{32} \cos t + \frac{1}{32} \cos 3t - \frac{3}{8} t \sin t \,, \right) + \ldots \,, \tag{2.5}$$

where the $t \sin t$ term arises because of the resonant forcing term in (2.4b) in an apparent contradiction of the fact that all solutions of (2.1) are periodic. These are the so-called secular terms that inspired a significant portion of modern perturbation theory. Even though regular perturbation theory is qualitatively incorrect, it is quantitatively correct in the sense that for every $T > 0$ there exists a C depending only on T such that $\left| y(t,\epsilon) - y^{(0)}(t) \right| \le C(T)\epsilon$ and $\left| y(t,\epsilon) - y^{(0)}(t) - \epsilon\, y^{(1)}(t) \right| \le C(T)\epsilon^2$ for $0 \le t \le T$ and for ϵ sufficiently small, and thus it does give a good approximation on *finite* time intervals. We will prove the first inequality after a brief discussion of the method of averaging.

The method of averaging circumvents the problem of secular terms and yields a result on a longer time interval. One of the standard forms for the method is

$$\dot{z} = \epsilon\, f(z,t), \qquad z(0,\epsilon) = z_0, \tag{2.6}$$

and the so-called averaged problem is

$$\dot{v} = \epsilon\, \bar{f}(v), \qquad v(0,\epsilon) = z_0 \,, \tag{2.7}$$

where

$$\bar{f}(v) = \lim_{T\to\infty} \frac{1}{T} \int_0^T f(v,t)\, dt.$$

Under certain conditions on f, we can show that

$$\left\| z(t,\epsilon) - v(t,\epsilon) \right\| \le C(T)\, \epsilon$$

for $0 \le t \le T/\epsilon$ and for ϵ sufficiently small. Thus, one says that at first order regular perturbation theory gives an $O(\epsilon)$ approximation on $O(1)$ time intervals, whereas averaging gives an $O(\epsilon)$ approximation on $O(1/\epsilon)$ time intervals. Furthermore, averaging usually gives the "correct" qualitative behavior, although in general it cannot be extended to time intervals longer than $O(1/\epsilon)$, as will be seen in the following example using the Duffing oscillator.

In order to apply averaging to (2.1) it must be put in the standard form. Using the variation of parameters transformation,

$$\begin{aligned} y &= e^{it} z_1 + e^{-it} z_2 \,, \\ \dot{y} &= i e^{it} z_1 - i e^{-it} z_2 \,, \end{aligned} \tag{2.8}$$

based on the solutions e^{it} and e^{-it} of the unperturbed problem (2.1), we obtain (2.6) with

$$f(z,t) = \frac{i}{2} \begin{pmatrix} z_1^3 e^{2it} + 3z_1^2 z_2 + 3z_1 z_2^2 e^{-2it} + z_2^3 e^{-4it} \\ -z_1^3 e^{4it} - 3z_1^2 z_2 e^{2it} - 3z_1 z_2^2 - z_2^3 e^{-2it} \end{pmatrix}, \quad z_0 = \frac{1}{2} \begin{pmatrix} y_0 - i\dot{y}_0 \\ y_0 + i\dot{y}_0 \end{pmatrix}. \tag{2.9}$$

The averaged problem becomes

$$\dot{v}_1 = \epsilon \frac{3i}{2} v_1^2 v_1^* , \quad v_1(0) = \frac{1}{2}(y_0 - i\dot{y}_0) =: Ae^{i\phi} \tag{2.10a}$$

$$v_2(t) = v_1^*(t) . \tag{2.10b}$$

Notice that one consequence of using the form (2.8) is that the solution of the averaged problem comes in complex conjugate pairs. This can be of great practical advantage, as it is often easier to use complex exponentials rather than cosines and sines. If the averaging approach is to be fruitful, the averaged equations (2.10) must entail a simplification over the original problem, and this is always the case in our experience. Here this manifests in the form of a conservation law for (2.10a), namely

$$v_1(t)v_1^*(t) = v_1(0)v_1^*(0) = A^2 ,$$

which allows the solution of (2.10a) by elementary means:

$$v_1(t) = A \exp \left[i \left(\epsilon \frac{3}{2} A^2 t + \phi \right) \right] . \tag{2.11}$$

Using (2.8) this gives

$$y(t) = y_0 \cos \left(1 + \frac{3}{8} \epsilon y_0^2 \right) t + O(\epsilon) , \tag{2.12}$$

for $\dot{y}_0 = 0$ and for $0 \le t \le T/\epsilon$. If we restrict our attention to $0 \le t \le T$, then we can expand in ϵ, which yields the regular perturbation result (2.5) to $O(\epsilon)$ as it must. Using the second-order averaging procedure and Theorem 3 as discussed in Section 5, we obtain

$$y(t, \epsilon) = y_0 \cos\big(\omega(\epsilon) t\big) + \epsilon y_0^3 \left(-\frac{1}{32} \cos\big(\omega(\epsilon)t\big) + \frac{1}{32} \cos\big(3\omega(\epsilon)t\big) \right) + O(\epsilon^2) , \tag{2.13a}$$

$$\omega(\epsilon) = 1 + \frac{3}{8} \epsilon y_0^2 - \frac{21}{256} \epsilon^2 y_0^4 \tag{2.13b}$$

for $\dot{y}_0 = 0$ and for $0 \le t \le T/\epsilon$. Expanding in ϵ for $0 \le t \le T$ yields the regular perturbation result to $O(\epsilon^2)$, again as it must.

First-Order Regular Perturbation and Averaging Theorems

To make the above rigorous we state and prove a regular perturbation theorem and an averaging theorem at first order. An understanding of the proofs of the theorems is not necessary for understanding the paper; they are included for the interested reader.

For the regular perturbation result, consider the IVP

$$\dot{z} = f_0(z, t) + \epsilon f_1(z, t) , \qquad z(0, \epsilon) = z_0 , \tag{2.14}$$

and let

$$\dot{v} = f_0(v, t) , \qquad v(0) = z_0 . \tag{2.15}$$

Theorem 1 (Regular Perturbation Theorem): *Let $f_0(z,t)$ and $f_1(z,t)$ be locally z-Lipschitz on an open set $\mathcal{U} \subset \mathbb{R}^d$ containing z_0, continuous in z and t on $\mathcal{U} \times \mathbb{R}$, and assume that the solution of* (2.15) *exists in $\mathcal{U}$ and on $[0,T]$. Then there exist $C(T)$ and $\epsilon_0(T)$ such that for $0 < \epsilon \leq \epsilon_0$ the solution of* (2.14) *exists in $\mathcal{U}$ on $[0,T]$ and on this same time interval $\left\| z(t,\epsilon) - v(t) \right\| \leq C(T)\epsilon$.*

Remarks:

(1) Here and in the following, $\| \quad \|$ will be a fixed vector norm.

(2) $f(z,t)$ is locally z-Lipschitz on $\mathcal{U}$ if for every $z \in \mathcal{U}$ there exists a neighborhood of z and a constant L such that $\|f(x,t) - f(y,t)\| \leq L\|x-y\|$ for all x and y in the neighborhood and all $t \in \mathbb{R}$. It can be shown that this is equivalent to f being z-Lipschitz on every compact subset of $\mathcal{U}$.

(3) A sufficient condition for $f(z,t)$ being locally z-Lipschitz is differentiability of $f(z,t)$ with respect to z on $\mathcal{U}$.

Proof: Let $\mathcal{S} = \{v(t) \,|\, 0 \leq t \leq T\}$ and let $\mathcal{U}_1$ be an open bounded set such that $\mathcal{S} \subset \mathcal{U}_1 \subset \bar{\mathcal{U}}_1 \subset \mathcal{U}$. Here $\bar{\mathcal{U}}_1$ denotes the closure of $\mathcal{U}_1$. Since $f_0(z,t)$ and $f_1(z,t)$ are locally z-Lipschitz on $\mathcal{U}$, f_0 is z-Lipschitz on $\bar{\mathcal{U}}_1$ with Lipschitz constant L, and f_1 is bounded on $\bar{\mathcal{U}}_1$ with bound M. Let $[0,\beta(\epsilon))$ be the maximum forward interval of existence of (2.14) in $\mathcal{U}_1$ and $J = [0,\beta(\epsilon)) \cap [0,T]$. Then subtracting (2.15) from (2.14), integrating, and using the triangle inequality yields

$$\begin{aligned}
\left\| z(t,\epsilon) - v(t) \right\| &\leq \int_0^t \left\| f_0\big(z(s,\epsilon),s\big) - f_0\big(v(s),s\big) \right\| ds + \epsilon \int_0^t \left\| f_1\big(z(s,\epsilon),s\big) \right\| ds \\
&\leq L \int_0^t \left\| z(s,\epsilon) - v(s) \right\| ds + \epsilon M t
\end{aligned}$$

for $t \in J$. The Gronwall inequality (see Appendix) then gives

$$\left\| z(t,\epsilon) - v(t) \right\| \leq \epsilon M t e^{Lt} \leq \epsilon M T e^{LT} =: C(T)\epsilon \,,$$

for $t \in J$. Choose $C(T)\epsilon_0(T) < \operatorname{dist}(\mathcal{S}, \partial\bar{\mathcal{U}}_1)$ and $0 < \epsilon \leq \epsilon_0$, then z does not approach the boundary of $\bar{\mathcal{U}}_1$ on J and the continuation theorem gives $\beta(\epsilon) > T$. Thus $J = [0,T]$ and the theorem is proven. □

Remarks:

(1) The continuation theorem states that either $z(t,\epsilon)$ exists for all forward time in $\mathcal{U}_1$, or it permanently leaves every compact subset of $\mathcal{U}_1$ as $t \nearrow \beta(\epsilon)$.

(2) Writing (2.1) in the system form of (2.14) by defining $(z_1, z_2) = (y, \dot{y})$ and then applying the theorem yields the first result mentioned after (2.5).

Next we state and prove a first-order averaging theorem for

$$\dot{z} = \epsilon f(z,t), \quad z(0,\epsilon) = z_0 \tag{2.16}$$

and the associated averaged problem

$$\dot{v} = \epsilon \bar{f}(v), \quad v(0,\epsilon) = z_0 \,, \tag{2.17}$$

in the case where f is quasiperiodic; that is,

$$f(z,t) = g\Big(z, \theta(t)\Big) \,, \tag{2.18}$$

where $\theta(t) = (\omega_1 t, \ldots, \omega_k t)$ and $g(z,\theta)$ is periodic of period 2π in each component of $\theta = (\theta_1, \theta_2, \ldots, \theta_k)$. We basically follow the approach in Remark 3, p. 393, of Reference 7 and Section VIIB of Reference 8. The ω_i are the base frequencies of the quasiperiodic function $f(z,\cdot)$, and the Fourier representation of g will be written

$$g(z,\theta) = \sum_{m \in \mathbf{Z}^d} g_m(z)\, e^{i\langle m,\theta\rangle} \,, \tag{2.19}$$

where $m = (m_1, \ldots, m_k)$ is an integer vector and $\langle m,\theta\rangle = m_1\theta_1 + \ldots + m_k\theta_k$. Thus we can write

$$f(z,t) = \sum_{m \in \mathbf{Z}^d} g_m(z)\, e^{i\langle m,\omega\rangle t} \tag{2.20}$$

and

$$\bar{f}(z) = \sum_{m \in \mathcal{M}} g_m(z) = \lim_{T\to\infty} \frac{1}{T} \int_0^T f(z,t)dt \,, \tag{2.21}$$

where

$$\mathcal{M} = \Big\{ m \in \mathbf{Z}^d \Big| \langle m,\omega\rangle = 0 \Big\} \,. \tag{2.22}$$

We define the "guiding solution" by the ϵ-independent IVP

$$\frac{du}{d\tau} = \bar{f}(u), \ u(0) = z_0 \,, \tag{2.23}$$

and note that $v(t,\epsilon) = u(\epsilon t)$ is the solution of (2.17).

Theorem 2 (First-Order Averaging Theorem): *Let $f(z,t)$ be locally z-Lipschitz on an open set $\mathcal{U} \subset \mathbf{C}^d$ containing z_0, continuous in z and t on $\mathcal{U} \times \mathbf{R}$ and assume that the solution of* (2.23) *exists in $\mathcal{U}$ on* $[0,T]$. *Let $\mathcal{S} = \Big\{ u(\tau) \Big| 0 \le \tau \le T \Big\}$ and assume for $z \in \mathcal{S}$ that*

$$\sum_{m \notin \mathcal{M}} 2 \Big\| g_m(z) \Big\| \Big/ \Big| \langle m,\omega\rangle \Big| \le M_1 < \infty \,, \tag{2.24}$$

and

$$\sum_{m \notin \mathcal{M}} 2 \Big\| D g_m(z) \Big\| \Big/ \Big| \langle m,\omega\rangle \Big| \le M_2 < \infty \,. \tag{2.25}$$

Then there exist $C(T)$ and $\epsilon_0(T)$ such that for $0 \le \epsilon \le \epsilon_0$ the solution of (2.16) exists in $\mathcal{U}$ on $[0, T/\epsilon]$ and on this same interval $\left\| z(t,\epsilon) - v(t,\epsilon) \right\| \le C(T)\epsilon$.

Proof: Let $\mathcal{U}_1$ be an open bounded set such that $\mathcal{S} \subset \mathcal{U}_1 \subset \bar{\mathcal{U}}_1 \subset \mathcal{U}$, let $\left[0, \beta(\epsilon)\right)$ be the maximum forward interval of existence of (2.16) in $\mathcal{U}_1$, and let $J = [0, T/\epsilon] \cap \left[0, \beta(\epsilon)\right)$. Let L be the Lipschitz constant of $f(z,t)$ for $z \in \bar{\mathcal{U}}_1$; then for $t \in J$,

$$\left\| z(t,\epsilon) - v(t,\epsilon) \right\| \le \epsilon L \int_0^t \left\| z(s,\epsilon) - v(s,\epsilon) \right\| ds + \epsilon \left\| \int_0^t \left\{ f\left(v(s,\epsilon), s\right) - \bar{f}\left(v(s,\epsilon)\right) \right\} ds \right\|. \tag{2.26}$$

Here we have subtracted (2.17) from (2.16), integrated, and added and subtracted $f(v,s)$ before using the triangle inequality. Let M_3 be the bound for $\left\| \bar{f}(v) \right\|$ for $v \in \mathcal{S}$ and $M := M_1 + M_2 M_3 T$; then we prove shortly that

$$\left\| \int_0^t \left\{ f\left(v(s,\epsilon), s\right) - \bar{f}\left(v(s,\epsilon)\right) \right\} ds \right\| \le M \tag{2.27}$$

for $t \in [0, T/\epsilon]$. An application of the Gronwall inequality then gives

$$\left\| z(t,\epsilon) - v(t,\epsilon) \right\| \le \epsilon M \exp \epsilon L t \le \epsilon M \exp LT =: C(T)\epsilon$$

for $t \in J$. Choose ϵ_0 such that $C(T)\epsilon_0(T) < \operatorname{dist}(\mathcal{S}, \partial\bar{\mathcal{U}}_1)$, then z does not approach the boundary of $\bar{\mathcal{U}}_1$ on J, and the continuation theorem gives $\beta(\epsilon) > T/\epsilon$ and the theorem is proven. It remains to prove (2.27). The left hand side of (2.27) is given by

$$\begin{aligned}
&\left\| \int_0^t \sum_{m \notin \mathcal{M}} g_m\left(v(s,\epsilon)\right) e^{i\langle m,\omega\rangle s} ds \right\| \\
&= \left\| \sum_{m \notin \mathcal{M}} \left\{ g_m\left(v(t,\epsilon)\right) \frac{e^{i\langle m,\omega\rangle t} - 1}{i\langle m,\omega\rangle} - \int_0^t Dg_m\left(v(s,\epsilon)\right) \frac{dv(s,\epsilon)}{ds} \frac{e^{i\langle m,\omega\rangle s} - 1}{i\langle m,\omega\rangle} ds \right\} \right\| \\
&\le M_1 + \epsilon M_3 M_2 t \le M_1 + M_2 M_3 T = M\,,
\end{aligned}$$

where the equality follows from integration by parts and the first inequality from Eqs. (2.24), (2.25), and (2.7). Equation (2.27) follows. □

Remarks:

(1) Notice there is one restriction on ϵ given by $\epsilon \le \epsilon_0 < \operatorname{dist}(\mathcal{S}, \partial\mathcal{U}_1)/C(T)$.

(2) The idea of estimating (2.27) is due to Besjes; see References 7 and 8.

(3) Ideas of resonance and nonresonance in relation to this theorem will be discussed at the end of Section 3. However, notice that (2.24) and (2.25) cannot be satisfied if $g_m \neq 0$ for $\langle m,\omega\rangle = 0$.

(4) More generally, an $O(\epsilon^2)$ term could be added to the rhs of (2.16). This would not affect the averaged problem or the approximation.

3. Transverse Motion in Averaging Form

A basic problem of transverse beam dynamics in a storage ring is to understand the motion defined by

$$x'' + K_x(s)x = \epsilon h_x(x,y,s)\,, \tag{3.1}$$
$$y'' + K_y(s)y = \epsilon h_y(x,y,s)\,, \tag{3.2}$$

for various perturbations h_x and h_y and for ϵ sufficiently small. In order to proceed we need to transform these equations to a standard form for the method of averaging. We will again use variation of parameters, and for that we need to discuss the solution of the unperturbed problem

$$x'' + K(s)x = 0\,, \quad x(0) = x_0\,, \quad x'(0) = x_0'\,, \tag{3.3}$$

which can be written in system form as

$$\underline{x}' = A(s)\underline{x}, \quad \underline{x}(0) = \begin{pmatrix} x_0 \\ x_0' \end{pmatrix}, \tag{3.4}$$

where $\underline{x} = (x_1, x_2)^T = (x, x')^T$ and $A(s) = \begin{pmatrix} 0 & 1 \\ -K(s) & 0 \end{pmatrix}$. Here $K(s)$ and thus $A(s)$ are periodic with period C, the circumference of the storage ring. Two linearly independent solutions of (3.3) are

$$\sqrt{\beta(s)}\, e^{\pm i\psi(s)}\,, \tag{3.5}$$

with Wronskian $-2i$, where

$$\psi(s) := \int_0^s \frac{1}{\beta(t)} dt \tag{3.6}$$

and $\beta(s)$ is a solution of

$$2\beta\beta'' - (\beta')^2 + 4\beta^2 K = 4\,. \tag{3.7}$$

We assume K is such that (3.7) has C-periodic solutions and take β to be C-periodic. Presumably it can be proven that (3.7) has a unique C-periodic solution, but uniqueness is not necessary for our analysis. The tune ν is defined by

$$2\pi\nu := \int_0^C \frac{1}{\beta(t)} dt = \psi(C)\,, \tag{3.8}$$

and if we define $\psi_p(s)$ by

$$\psi(s) = \psi_p(s) + 2\pi\nu s/C\,, \tag{3.9}$$

then $\psi_p(s)$ has period C and thus (3.5) is seen to be quasiperiodic with base frequencies $\omega_1 := 2\pi/C$ and $\omega_2 := \nu\omega_1$ (periods C and C/ν).

Choosing the right coordinates is a significant issue for perturbation calculations, and here it is convenient to define the following fundamental solution matrix of (3.4):

$$\Psi(s) = \sqrt{\beta(s)} \begin{pmatrix} e^{i\psi(s)} & e^{-i\psi(s)} \\ \frac{1}{\beta(s)}\left(\frac{1}{2}\beta'(s) + i\right)e^{i\psi(s)} & \frac{1}{\beta(s)}\left(\frac{1}{2}\beta'(s) - i\right)e^{-i\psi(s)} \end{pmatrix} . \tag{3.10}$$

Note that $\det \Psi(s) = -2i$, that

$$\Psi^{-1}(s) = \frac{i}{2}\sqrt{\beta(s)} \begin{pmatrix} \frac{1}{\beta(s)}\left(\frac{1}{2}\beta'(s) - i\right)e^{-i\psi} & -e^{-i\psi} \\ -\frac{1}{\beta(s)}\left(\frac{1}{2}\beta'(s) + i\right)e^{i\psi} & e^{i\psi} \end{pmatrix} , \tag{3.11}$$

and that the Floquet decomposition of (3.10) can be written by observation as

$$\Psi(s) = \Psi_p(s)\, e^{iB\omega_2 s} , \tag{3.12}$$

where Ψ_p is Ψ with ψ replaced by ψ_p and B is the diagonal matrix with diagonal elements ± 1.

We now consider the perturbed problem for (3.3), namely

$$x'' + K(s)x = \epsilon h(x, \omega_1 s, \omega_3 s) , \quad x(0) = x_0 , \quad x'(0) = x_0' \tag{3.13}$$

where $h(x, \theta_1, \theta_2)$ is 2π-periodic in θ_1 and in θ_2. Here θ_1 will represent perturbations with the period of the lattice, and θ_2 will represent external periodic perturbations due to, for example, power supply ripple. It is a simple matter to add more frequency components to h. In the system form of (3.4) this becomes

$$\underline{x}' = A(s)\underline{x} + \epsilon H(\underline{x}, \omega_1 s, \omega_3 s), \quad \underline{x}(0) = \begin{pmatrix} x_0 \\ x_0' \end{pmatrix} , \tag{3.14}$$

where $H = (0, h)^T$. To put this in the standard form for the method of averaging we define a transformation from $\underline{x}$ to z via

$$\underline{x} = \Psi(s) z . \tag{3.15}$$

The IVP for z becomes

$$z' = \epsilon f(z, s), \quad z(0) = z_0 = \begin{pmatrix} Ae^{i\phi} \\ cc \end{pmatrix} , \tag{3.16}$$

where cc denotes the complex conjugate of the first component,

$$f(z, s) = \Psi^{-1}(s) H\left(\Psi(s)z, \omega_1 s, \omega_3 s\right) , \tag{3.17}$$

and

$$Ae^{i\phi} = \frac{1}{2}\beta_0^{-1/2}\left[x_0 + i\left(\frac{1}{2}\beta_0' x_0 - \beta_0 x_0'\right)\right] . \tag{3.18}$$

Equation (3.16) is in a standard form for the method of averaging.

Since $\Psi^{-1}(s)\Psi^*(s) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $\underline{x}$ is real, we have $z_1 = z_2^*$ and $f_1(z,s) = f_2(z,s)^*$. Thus it suffices to work with

$$f_1(z,s) = -\frac{i}{2}\sqrt{\beta(s)}e^{-i\psi(s)}h(x,\omega_1 s,\omega_3 s)\,, \tag{3.19a}$$

where x on the rhs must be replaced by

$$x = \sqrt{\beta(s)}\left(e^{i\psi(s)}z_1 + e^{-i\psi(s)}z_2\right)\,. \tag{3.19b}$$

It is thus clear that we are dealing with $f(z,\cdot)$, which is quasiperiodic with base frequencies ω_1, ω_2 and ω_3. The averaged problem will be written

$$v' = \epsilon\bar{f}(v)\,, \qquad v(0) = z_0\,, \tag{3.20}$$

and the first-order averaging result of Section 2 now applies. Before presenting some examples we briefly discuss resonance in first-order perturbation theory.

Nonresonance and Resonance at First Order

Let $f(z,s) = g(z,\omega_1 s,\omega_2 s,\omega_3 s)$ as in (2.18), then $\omega = (\omega_1,\omega_2,\omega_3)$. The case where $\bar{f} = g_0$, in (2.21), is called the nonresonant case, at first order. It is of practical significance to notice that in this case, $\bar{f}(z)$ can be constructed by averaging over each "frequency" separately; that is, $\bar{f}(z)$ is obtained by averaging $g(z,\theta_1,\theta_2,\theta_3)$ over each θ_i independently. Resonance is defined as not nonresonant, or equivalently there exists nonzero m such that $\langle m,\omega\rangle = 0$ and $g_m \neq 0$. In the resonant case, it is often of interest to analyze the motion in a neighborhood of the resonance rather than right on resonance. Let us fix ω_1 and ω_3 and suppose that $(\omega_1,\omega_2,\omega_3)$ is resonant at $\omega_2 = \omega_{20} := \nu_0\omega_1$; that is, there exists a nonzero m such that $m_1\omega_1 + m_2\omega_{20} + m_3\omega_3 = 0$. Let

$$\omega_2 = \omega_{20} + \epsilon a \tag{3.21}$$

and define

$$f_r(z,\tau,s) := g(z,\omega_1 s,\omega_{20}s + a\tau,\omega_3 s)\,. \tag{3.22}$$

The quantity a measures the distance from resonance and will be called the resonance parameter. We now show that this case can be handled in the context of Theorem 2.

The IVP $z' = \epsilon f_r(z,\epsilon s,s)$, $z(0) = z_0$ can be written

$$\begin{aligned} z' &= \epsilon f_r(z,\tau,s), \quad z(0) = z_0 \\ \tau' &= \epsilon, \quad \tau(0) = 0\,, \end{aligned} \tag{3.23}$$

which is in the form of (2.16) with z replaced by $(z,\tau)^T$ and f replaced by $(f_r,1)^T$. Averaging does not change the τ-equation, and the averaged IVP can be written

$$v' = \epsilon\bar{f}_r(v,\epsilon s), \quad v(0) = z_0\,, \tag{3.24}$$

where

$$\bar{f}_r(v,\tau) := \lim_{T\to\infty} \frac{1}{T}\int_0^T f_r(v,\tau,s)ds,$$

and where it is important to note that both v and τ are fixed in computing the average. It is easy to see that the averaging theorem gives $\left\|z(s,\epsilon)-v(s,\epsilon)\right\| \le C(T)\epsilon$ on $O(1/\epsilon)$ s-intervals as before.

4. Beam Dynamics Examples at First Order

Here we discuss four beam dynamics problems in the context of first-order averaging.

Example A: $h(x,\omega_1 s)$ is a product.

Several problems in beam dynamics have the form

$$h(x,\omega_1 s) = d(x)e(s)\,, \tag{4.1}$$

where $e(s)$ is C-periodic. For example,

$$h(x,\omega_1 s) = \begin{cases} \frac{1}{2}S(s)x^2, & \text{sextupole} \\ \frac{D}{x}(1-e^{-x^2/2\sigma^2})e(s), & \text{beam-beam} \\ L(s)x, & \text{chromaticity} \\ \frac{1}{6}O(s)x^3, & \text{octupole}. \end{cases}$$

Thus from (3.19),

$$f_1(z,s) = -\frac{1}{2}i\sqrt{\beta(s)}\,e(s)e^{-i\psi(s)}d(x)\,. \tag{4.2}$$

Expanding d in a Taylor series and using (3.19b) gives

$$d(x) = \sum_{n=0}^{\infty} d_n x^n = \sum_{n=0}^{\infty}\beta(s)^{n/2}d_n\sum_{k=0}^{n}\binom{n}{k}e^{i(2k-n)\psi(s)}z_1^k z_2^{n-k}\,, \tag{4.3}$$

and thus

$$f_1(z,s) = -\frac{1}{2}i\sum_{n=0}^{\infty} d_n\beta(s)^{\frac{n+1}{2}}e(s)\sum_{k=0}^{n}\binom{n}{k}e^{i(2k-n-1)\psi(s)}z_1^k z_2^{n-k}\,. \tag{4.4}$$

Here we will consider the nonresonant case and treat the chromaticity, sextupole, octupole, and beam-beam as special cases.

In the nonresonant case, we can average the terms in $\omega_1 s$ and $\omega_2 s$ in (4.4) separately. Note that $\omega_2 s$ occurs only in the term $e^{i(2k-n-1)\omega_2 s}$, which has zero average unless $k=\frac{n+1}{2}$. Thus, the inner sum only contributes for odd $n=2l+1$ and thus $k=l+1$. The second average with respect to $\omega_1 s$ then gives

$$\bar{f}_1(v) = -\frac{1}{2}i\sum_{l=0}^{\infty}\binom{2l+1}{l+1}d_{2l+1}\,\overline{\beta(s)^{l+1}e(s)}\,|v_1|^{2l}v_1 =: i\gamma(|v_1|)v_1\,, \tag{4.5}$$

where

$$\gamma(A) = -\frac{1}{2}\sum_{l=0}^{\infty}(2l+1)!\frac{d_{2l+1}}{l!(l+1)!}\,\overline{\beta(s)^{\ell+1}e(s)}\,A^{2l},$$

and we have used $v_2 = v_1^*$. The averaged IVP becomes

$$v_1' = i\epsilon\gamma(|v_1|)v_1\,, \quad v_1(0) = Ae^{i\phi}\,, \tag{4.6}$$

and the complex conjugate for v_2. Note that γ is real. It is easy to see that $|v_1|$ is conserved, and the solution of (4.6) is therefore $v_1 = Ae^{i(\epsilon\gamma(A)s+\phi)}$. Under the hypotheses of the averaging theorem, we have

$$x(s) = 2\sqrt{\beta(s)}A\cos\Big(\psi(s) + \epsilon\gamma(A)s + \phi\Big) + O(\epsilon) \tag{4.7}$$

for $0 \leq s \leq T/\epsilon$. Thus to $O(\epsilon)$ the motion follows the betatron motion with a tune shift of $\epsilon C\gamma(A)/2\pi$.

Since $\omega_2 = \nu\omega_1$, the nonresonance condition is satisfied if $(2k-n-1)\nu$ is noninteger for $0 \leq k \leq n$ and all n such that $d_n \neq 0$. It is straightforward but cumbersome to derive conditions on the Fourier coefficients of $\beta(s)^{\frac{n+1}{2}}\,e(s)\,e^{i(2k-n-1)\psi_p(s)}$ so that (2.24) and (2.25) are satisfied.

In the sextupole case, f_1 in (3.19) is given by

$$f_1(z,s) = i\Big[e_1(s)e^{i\omega_2 s}z_1^2 + e_{-1}(s)e^{-i\omega_2 s}2z_1z_2 + e_{-3}(s)e^{-i3\omega_2 s}z_2^2\Big]\,, \tag{4.8a}$$

where

$$e_\ell(s) = -\frac{1}{4}\beta^{3/2}(s)S(s)e^{i\ell\psi_p(s)} =: \sum_n e_{\ell n}e^{in\omega_1 s} \tag{4.8b}$$

is C-periodic. It is easy to check that $\gamma(A) = 0$ and thus $\bar{f}(z) = 0$ if ν and 3ν are not integers. Thus, if the conditions of Theorem 2 are satisfied, then $z(s) = z_0 + O(\epsilon)$ and $\underline{x}(s) = \Psi(s)z(s)$ follows the betatron motion defined by (3.3) to $O(\epsilon)$ on $O(1/\epsilon)$ time intervals. To satisfy the conditions of Theorem 2, $\mathcal{U}$ can be any open set containing z_0, and T any positive number. Notice that $\mathcal{S}$ becomes a point and that (2.24) and (2.25) are satisfied if $\beta(s)$ and $S(s)$ are such that

$$\sum_n \frac{|e_{\ell n}|}{|n+\ell\nu|} < \infty$$

for $\ell = 1, -1$, and -3. But since ν and 3ν are noninteger, this is satisfied if $\sum_n |e_{\ell n}| < \infty$.

In the beam-beam case, it is reasonable to take $e(s) = \delta_p(s)$, where δ_p is the periodic delta function with period C. Then we obtain

$$\gamma(A) = -\frac{D}{8\pi\,A^2\,C}\int_0^{2\pi}\left[1 - \exp\left(\frac{-4\beta_0A^2\sin^2\theta}{2\sigma^2}\right)\right]d\theta\,,$$

which checks with References 9 and 10. It should be pointed out, however, that the averaging theorem does not apply in this case because the smoothness conditions on the vector field are not satisfied. It would be interesting to extend the averaging error analysis to vector fields with delta functions.

In the other cases we obtain the well-known results:

$$\gamma(A) = \begin{cases} -\frac{1}{2}\,\overline{\beta(s)L(s)}, & \text{chromaticity} \\ -\frac{1}{24}\,\overline{\beta^2(s)O(s)}\;A^2, & \text{octupole}\;. \end{cases}$$

Example B: Sextupole near the 1:3 resonance.

We now consider the sextupole third-integer resonance. We know that $\bar{f}$ is easily computed if 3ν (and/or ν) is an integer, but to make it more interesting we investigate the case of near resonance, where

$$\omega_2 = \omega_{20} + \epsilon a; \qquad 3\omega_{20} = 3\nu_0\omega_1 = M\omega_1\,, \tag{4.9}$$

where M is an integer but $\nu_0 = M/3$ is not. We follow the resonance discussion at the end of Section 3. From (3.22) and (4.8a),

$$f_{r1}(z,\tau,s) = ie_{-3}(s)\,e^{-iM\omega_1 s}\,e^{-i3a\tau}z_2^2 + \text{zero } s\text{-mean terms}\,.$$

Thus

$$\bar{f}_{r1}(z,\tau) = i\alpha_M\,e^{-i3a\tau}z_2^2\,, \quad \alpha_M = e_{-3,M}\,,$$

and the averaged IVP becomes

$$\begin{aligned} v_1' &= i\epsilon\,\alpha_M\,e^{-i\epsilon 3as}v_2^2\,, \\ & & v(0) = z_0 \\ v_2' &= -i\epsilon\,\alpha_M^*\,e^{i\epsilon 3as}v_1^2\,, \end{aligned}$$

or

$$v_1' = i\epsilon\alpha_M e^{-i\epsilon 3as}v_1^{*2}, \qquad v_1(0) = Ae^{i\phi}\,.$$

These equations are nonautonomous but only in a trivial way. If we let $v_1 = e^{i(-\epsilon as)}\zeta$, then

$$\zeta' = i\epsilon\left[a\zeta + \alpha_M\zeta^{*2}\right]\,, \qquad \zeta(0) = Ae^{i\phi}\,. \tag{4.10}$$

It is easy to see that this equation has four equilibrium solutions, $\zeta = 0$ and three others, as is to be expected for the sextupole 1:3 resonance. Also (4.10) has a Hamiltonian type structure; see (6.24). If we let $\zeta = X + iY$ and $\alpha_M = \gamma + i\delta$, with X, Y, γ, and δ real, then

$$\frac{dX}{d\tau} = \left[-aY - \delta(X^2 - Y^2) + 2\gamma XY\right]\,, \quad X(0) = A\cos\phi \tag{4.11a}$$

$$\frac{dY}{d\tau} = \left[aX + \gamma(X^2 - Y^2) + 2\delta XY\right]\,, \quad Y(0) = A\sin\phi\,, \tag{4.11b}$$

where $\tau = \epsilon s$. These equations are autonomous and Hamiltonian and thus are easily analyzed in the phase plane. Since

$$\underline{x} = \Psi(s)z \simeq \Psi(s)\begin{pmatrix} e^{-i\epsilon as}(X + iY) \\ e^{i\epsilon as}(X - iY) \end{pmatrix}\,,$$

we obtain

$$x(s,\epsilon) = x_1(s,\epsilon) = 2\sqrt{\beta(s)}\Big[X(\epsilon s)\cos\big(\psi(s)-\epsilon a s\big) - Y(\epsilon s)\sin\big(\psi(s)-\epsilon a s\big)\Big] + O(\epsilon) \qquad (4.12)$$

as our complete approximate third-integer resonance solution on $0 \le s \le T/\epsilon$. Note that this is the betatron motion with a frequency shift of $\Delta\omega_2 = -\epsilon a$ and slowly varying coefficients defined by the IVP (4.11).

To check the conditions of the theorem, we pick A, ϕ, and T so that the solution of (4.11) exists for $0 \le \tau \le T$. Then $\mathcal{U}$ and $\mathcal{U}_1$ are chosen appropriately. Equations (2.24) and (2.25) will then be satisfied if

$$\sum_n \frac{|e_{\ell n}|}{|n+\ell\nu_0|} < \infty \quad \text{for} \quad \ell = \pm 1\,,$$

and

$$\sum_{n\neq M} \frac{|e_{\ell n}|}{|n-M|} < \infty \quad \text{for} \quad \ell = -3\,.$$

Again this will be true if $\sum_n |e_{\ell n}| < \infty$.

Example C: Beam-beam with x-y coupling.

The equations of motion (3.1) and (3.2) can be written in system form as

$$\underline{x}' = A(s)\underline{x} + \epsilon H(\underline{x},\omega_1 s)\,, \qquad (4.13)$$

where $\underline{x} = (x_1,x_2,x_3,x_4)^T = (x,x',y,y')^T$,

$$A(s) = \begin{pmatrix} A_x(s) & 0 \\ 0 & A_y(s) \end{pmatrix}, \qquad A_{x,y} = \begin{pmatrix} 0 & 1 \\ -K_{x,y}(s) & 0 \end{pmatrix},$$

and

$$H(\underline{x},\omega_1 s) = \Big(0, h_x(x_1,x_3,\omega_1 s), 0, h_y(x_1,x_3,\omega_1 s)\Big)^T. \qquad (4.14)$$

The variation of parameters transformation

$$\underline{x} = \Psi(s) z, \qquad \Psi(s) = \begin{pmatrix} \Psi_x(s) & 0 \\ 0 & \Psi_y(s) \end{pmatrix} \qquad (4.15)$$

gives

$$z' = \epsilon f(z,s) := \epsilon \Psi^{-1}(s) H\Big(\Psi(s)z,\omega_1 s\Big) =: \epsilon g(z,\omega_1 s,\omega_x s,\omega_y s)\,, \qquad (4.16)$$

where $\omega_x = 2\pi\nu_x/C$, $\omega_y = 2\pi\nu_y/C$, and ν_x and ν_y are the x and y tunes. Now

$$\begin{aligned} f_1(z,s) &= -\frac{i}{2}\sqrt{\beta_x(s)}e^{-i\psi_x(s)} h_x(x_1,x_3,\omega_1 s)\,, \\ f_3(z,s) &= -\frac{i}{2}\sqrt{\beta_y(s)}e^{-i\psi_y(s)} h_y(x_1,x_3,\omega_1 s)\,, \end{aligned} \qquad (4.17)$$

where x_1 and x_3 must be replaced by

$$\begin{aligned} x_1 &= \sqrt{\beta_x(s)}\left(e^{i\psi_x(s)}z_1 + e^{-i\psi_x(s)}z_2\right) \\ x_3 &= \sqrt{\beta_y(s)}\left(e^{i\psi_y(s)}z_3 + e^{-i\psi_y(s)}z_4\right) . \end{aligned} \tag{4.18}$$

In the beam-beam

$$\begin{aligned} h_x(x_1, x_3, \omega_1 s) &= x_1 d(x_1, x_3) e(s) \\ h_y(x_1, x_3, \omega_1 s) &= x_3 d(x_1, x_3) e(s) , \end{aligned} \tag{4.19}$$

where

$$\begin{aligned} d(x_1, x_3) &= D(x_1^2 + x_3^2)^{-1}\left[1 - \exp\left(-(x_1^2 + x_3^2)/2\sigma^2\right)\right] \\ &= \sum_{k,\ell=0}^{\infty} d_{k\ell} x_1^k x_3^\ell , \end{aligned} \tag{4.20}$$

and the second equality defines the $d_{k\ell}$. Combining (4.17)–(4.20) and using the binomial expansion gives

$$\begin{aligned} f_1(z,s) &= -\frac{i}{2}\sum_{k,\ell=0}^{\infty} d_{k\ell}\beta_x(s)^{\frac{k+2}{2}}\beta_y(s)^{\frac{\ell}{2}} e(s) \sum_{m=0}^{k+1}\binom{k+1}{m} z_1^m z_2^{k+1-m} e^{i(2m-k-2)\psi_x(s)} \\ &\quad \times \sum_{n=0}^{\ell}\binom{\ell}{n} z_3^n z_4^{\ell-n} e^{i(2n-\ell)\psi_y(s)} , \end{aligned} \tag{4.21}$$

and a similar expression for $f_3(z,s)$. In the nonresonance case, where we can average the terms with $\omega_1 s$, $\omega_x s$, and $\omega_y s$ separately, the terms in the sum of (4.21) are zero unless k and ℓ are even and $m = \frac{k+2}{2}$ and $n = \frac{\ell}{2}$. Letting $k = 2p$ and $\ell = 2q$ then gives

$$\bar{f}_1(z) = i\gamma_x(|z_1|, |z_3|)z_1 , \tag{4.22}$$

where

$$\gamma_x(A_x, A_y) = -\frac{1}{2}\sum_{p,q=0}^{\infty} d_{2p,2q}\binom{2p+1}{p+1}\binom{2q}{q}\overline{\beta_x^{p+1}(s)\,\beta_y^q(s)\,e(s)}\;A_x^{2p}A_y^{2q} , \tag{4.23}$$

and

$$\bar{f}_3(z) = i\gamma_y(|z_1|, |z_3|)z_3 , \tag{4.24}$$

where

$$\gamma_y(A_x, A_y) = -\frac{1}{2}\sum_{p,q=0}^{\infty} d_{2p,2q}\binom{2p}{p}\binom{2q+1}{q+1}\overline{\beta_x^p(s)\beta_y(s)^{q+1}\,e(s)}\;A_x^{2p}A_y^{2q} . \tag{4.25}$$

The averaged equations become

$$v_1' = i\epsilon\gamma_x\Big(|v_1|, |v_3|\Big)v_1\,, \qquad v_1(0) = A_x e^{i\phi_x}\,,$$
$$v_3' = i\epsilon\gamma_y\Big(|v_1|, |v_3|\Big)v_3\,, \qquad v_3(0) = A_y e^{i\phi_y}\,, \tag{4.26}$$

and recall that $v_2(s) = v_1^*(s)$ and $v_4(s) = v_2^*(s)$.

It is easy to check that $|v_1|$ and $|v_3|$ are conserved and thus the solution of (4.26) is

$$\begin{aligned} v_1(s) &= A_x e^{i\left(\epsilon\gamma_x(A_x,A_y)s+\phi_x\right)}\,, \\ v_3(s) &= A_y e^{i\left(\epsilon\gamma_y(A_x,A_y)s+\phi_y\right)}\,, \end{aligned} \tag{4.27}$$

and the averaging theorem gives

$$\begin{aligned} x_1(s) &= 2\sqrt{\beta_x(s)}A_x\cos\Big(\psi_x(s) + \epsilon\gamma_x(A_x, A_y)s + \phi_x\Big) + O(\epsilon)\,, \\ x_3(s) &= 2\sqrt{\beta_y(s)}A_y\cos\Big(\psi_y(s) + \epsilon\gamma_y(A_x, A_y)s + \phi_y\Big) + O(\epsilon)\,, \end{aligned} \tag{4.28}$$

for $0 \le s \le T/\epsilon$. Thus to $O(\epsilon)$ the motion follows the betatron motion on $O(1/\epsilon)$ s-intervals with tune shifts that can be calculated from (4.23) and (4.25).

For the special case of $e(s) = \delta_p(s-s_b)$, where s_b is the location of beam-beam interaction and $\beta_x(s_b) = \beta_y(s_b) =: \beta_b$, one can show that

$$\begin{aligned} \gamma_x(A_x, A_y) &= -\frac{D}{16\pi^2 A_x^2 C^2}\int_0^{2\pi} d\theta_x \int_0^{2\pi} d\theta_y \frac{x^2}{x^2+y^2}\left[1 - \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)\right]\,, \\ \gamma_y(A_x, A_y) &= \gamma_x(A_y, A_x)\,, \end{aligned}$$

where $x = 2A_x\sqrt{\beta_b}\sin\theta_x$ and $y = 2A_y\sqrt{\beta_b}\sin\theta_y$ in the integrand. This checks with References 9 and 10. Again, the averaging theorem does not apply in this special case because of the non-smooth vector field.

Example D: Rf phase modulation with damping.

A recent experiment at the Indiana University Cyclotron Facility (IUCF) with electron cooling showed that rf phase modulation near the 1:1 resonance leads to longitudinal beam splitting. In Reference 11, we elucidate this by applying the method of averaging to a pendulum equation with small damping and periodic forcing. This gives a detailed picture of the dynamics in terms of a Poincaré map with two attractors. Here we summarize the perturbation part of that paper.

The equation of motion for the longitudinal dynamics in the IUCF experiment is

$$\ddot{\phi} + 2\alpha\dot{\phi} + \omega_s^2\sin\phi = 2\alpha\dot{\psi} + \ddot{\psi}\,, \tag{4.29}$$

where

$$\psi(t) = F\frac{\omega_s}{\omega_m}\sin(\omega_m t + \Theta)$$

is the phase modulation, ϕ the rf phase, α the damping due to the electron cooling, ω_s the synchrotron frequency, and ω_m the modulation frequency. Because we are interested in the Poincaré map (PM) we change independent variable by

$$\tau = \omega_m t + \Theta\,,$$

which gives the IVP:

$$\phi'' + 2\frac{\alpha}{\omega_m}\phi' + \frac{\omega_s^2}{\omega_m^2}\sin\phi = F\frac{\omega_s}{\omega_m}\left(2\frac{\alpha}{\omega_m}\cos\tau - \sin\tau\right)\,, \tag{4.30}$$

$$\phi(\tau_0) = \phi_0, \qquad \phi'(\tau_0) = \phi_0'\,.$$

The τ_0-PM is then the map $(\phi_0, \phi_0') \longrightarrow \big(\phi(\tau_0 + 2\pi), \phi'(\tau_0 + 2\pi)\big)$. Typical parameters for this problem are $\omega_m = 2\pi(240)\ \mathrm{s}^{-1}$, $\omega_s = 2\pi(262)\ \mathrm{s}^{-1}$, $\alpha = 2.5\ \mathrm{s}^{-1}$, and $F = 0.0195$. Numerical experiments (as well as the experiment) indicate the existence of two attracting periodic orbits in the dynamics of (4.30). In the $\tau_0 = -\pi/2$ PM these appear as fixed points (FPs) at $(\phi_e, \phi_e') = (-1.1957, 0.2146)$ and $(0.119, 0.0023)$.

We have found these stable FPs of the PM by numerically integrating (4.30) and following two orbits until they settle down. It appears that each point in a region surrounding the FPs limits on one of the two fixed points; however, the basins of attraction are not clear, as nearby points can go to different attractors. Our goal is to explain this using the method of averaging, which will interpret the complicated dynamics of (4.30) in terms of an autonomous system in the plane. The autonomous system in the plane is easily understood in terms of its phase-plane portrait, and the portrait characterizes the basins of attraction and elucidates the beam splitting.

To put (4.30) in a form for the method of averaging we need to introduce scaling and transformations. From the numerical experiments it appears that the effects of nonlinearity, forcing, and damping are all important. If there were no forcing, (4.30) would be a damped pendulum, and there would be only one FP at the origin, and it would be a stable spiral point. (The FPs at $(\pm\pi, 0)$ are not in the region of interest.) If nonlinearity were not important we could replace $\sin\theta$ by θ, but then again there would be only one FP, a stable spiral. The fact that there appear to be asymptotically stable FPs shows that the damping is significant. We will treat all three effects as perturbations and introduce a small parameter ϵ in a way that brings in each effect at the same order: $\phi = \epsilon^{1/2}\varphi_1$, $F = \epsilon^{3/2}\hat{F}$, and $\alpha/\omega_s = \epsilon\hat{\alpha}$. Here we assume φ_1, $\hat{F}$, and $\hat{\alpha}$ to be $O(1)$. Also because ω_m and ω_s are close we expect a 1:1 resonance, and so we take $\omega_m/\omega_s =: 1 + \beta =: 1 + \epsilon\hat{\beta}$, which yields

$$\begin{aligned} \varphi_1' &=: \varphi_2 \\ \varphi_2' &= -\varphi_1 + \epsilon g_2(\varphi,\tau) + \epsilon^2 h_2(\varphi,\tau,\epsilon), \end{aligned} \tag{4.31}$$

where $g_2(\varphi,\tau) = +2\hat{\beta}\varphi_1 + \frac{1}{6}\varphi_1^3 - 2\hat{\alpha}\varphi_2 - \hat{F}\sin\tau$ and $h_2(\varphi,\tau,\epsilon) = O(1)$, *i.e.*, h_2 is bounded as $\epsilon \to 0$. With $h_2 = 0$, this is just the Duffing equation with small damping, forcing, and nonlinearity near resonance, an equation which has been extensively studied in the dynamical systems literature (see References 1 and 12). If we use the previous parameters with $\hat{F} = 1$, then $\epsilon = 0.07239$, $\hat{\alpha} = 0.02098$, $\beta = -0.0840$, and $\hat{\beta} = -1.1600$.

Regular perturbation theory gives the $O(1)$ periodic solutions of (4.31) as $\varphi_1 = r\cos(\tau - \chi) + O(\epsilon)$, where

$$\begin{aligned} \hat{\alpha} r - \frac{1}{2}\hat{F}\cos\chi &= 0, \\ \hat{\beta} r + \frac{1}{16} r^3 - \frac{1}{2}\hat{F}\sin\chi &= 0, \end{aligned} \tag{4.32}$$

which follows from the expansion $\varphi_1 = y_0 + \epsilon\, y_1 + O(\epsilon^2)$ and the removal of secular terms from the y_1 equation since y_1 must be periodic. In the PM, periodic solutions are fixed points. Solving (4.32) for the above parameters gives, in the coordinates of the PM,

$$\begin{pmatrix} \phi_e \\ \phi_e' \end{pmatrix} = \begin{pmatrix} -1.1905 \\ 0.2292 \end{pmatrix}, \begin{pmatrix} 0.1171 \\ 0.0021 \end{pmatrix}, \begin{pmatrix} 1.0809 \\ 0.1877 \end{pmatrix}.$$

Two correspond very well with the two apparent asymptotically stable FPs found by iteration, and the other, we shall see later, corresponds to an unstable FP of the PM.

We now proceed, using the method of averaging, to find an approximation to the PM for (4.30). The initial value problem for (4.31) can be written in vector form $\varphi' = J\varphi + \epsilon g(\varphi,\tau) + \epsilon^2 h(\varphi,\tau,\epsilon)$, $\varphi(\tau_0,\tau_0,\xi,\epsilon) = \xi$, where $\varphi = \varphi(\tau,\tau_0,\xi,\epsilon)$, $J = (0,1;-1,0)$, $\varphi = (\varphi_1,\varphi_2)^T$, $g = (0,g_2)^T$, and $h = (0,h_2)^T$. Defining x by the variation-of-parameters transformation,

$$\varphi = e^{J(\tau-\tau_0)}x, \tag{4.33}$$

yields the IVP for $x = x(\tau,\tau_0,\xi,\epsilon)$:

$$x' = \epsilon f(x,\tau,\tau_0) + \epsilon^2 R, \quad x(\tau_0,\tau_0,\xi,\epsilon) = \xi, \tag{4.34}$$

$$f(x,\tau,\tau_0) = e^{-J(\tau-\tau_0)} g(e^{J(\tau-\tau_0)}x,\tau), \tag{4.35}$$

and $R = R(x,\tau,\tau_0,\epsilon)$ defined analogously. We have introduced τ_0 in (4.33) so that in the τ_0 section $\varphi = x$. This problem is now in a standard form for the method of averaging. The averaged problem is

$$v' = \epsilon \bar{f}(v,\tau_0), \quad v(\tau_0,\tau_0,\xi,\epsilon) = \xi, \tag{4.36}$$

$$\begin{aligned} \bar{f}(v,\tau_0) &= \frac{1}{2\pi}\int_0^{2\pi} f(v,\tau,\tau_0)\,d\tau \\ &= \begin{pmatrix} -\hat{\alpha} v_1 - \hat{\beta} v_2 - \frac{1}{16}(v_1^2+v_2^2)v_2 + \frac{1}{2}\hat{F}\cos\tau_0 \\ \\ -\hat{\alpha} v_2 + \hat{\beta} v_1 + \frac{1}{16}(v_1^2+v_2^2)v_1 - \frac{1}{2}\hat{F}\sin\tau_0 \end{pmatrix}. \end{aligned}$$

A modification of Theorem 2 gives

$$x = v + \epsilon P\big(v,\tau,\tau_0\big) + O\big(\epsilon^2 + \epsilon^2(\tau-\tau_0)\big) \tag{4.37}$$

for $0 \leq \tau - \tau_0 \leq O(1/\epsilon)$. Here x and v have the same arguments $(\tau, \tau_0, \xi, \epsilon)$ and $P(v, \tau, \tau_0) := \int_{\tau_0}^{\tau} \big(f(v, s, \tau_0) - \bar{f}(v, \tau_0)\big) ds$. Note that this gives an improved approximation to Theorem 2 on $O(1)$ τ-intervals but gives the same estimate on $O(1/\epsilon)$ intervals. This result is useful for proving the existence of periodic orbits and invariant tori, as will be pointed out later, because

$$\varphi(\tau_0 + 2\pi n, \tau_0, \xi, \epsilon) = v(\tau_0 + 2\pi n, \tau_0, \xi, \epsilon) + O(\epsilon^2 n). \tag{4.38}$$

That is, the PM is actually defined by the averaged problem to $O(\epsilon^2)$ instead of $O(\epsilon)$, as given by Theorem 2. To see that (4.37) is to be expected, let $w = v + \epsilon P(v, \tau)$ where we suppress the τ_0 dependence. Differentiating along the solutions of (4.36) gives

$$\begin{aligned} w' &= \big(I + \epsilon D_1 P(v, \tau)\big)\, \epsilon \bar{f}(v) + \epsilon\big(f(v, \tau) - \bar{f}(v)\big) \\ &= \epsilon f(w, \tau) - \epsilon\big(f(w, \tau) - f(v, \tau)\big) + \epsilon^2 D_1 P(v, \tau) \bar{f}(v). \end{aligned}$$

This yields $w' = \epsilon f(w, \tau) + O(\epsilon^2)$, which is $O(\epsilon^2)$ close to (4.34). Subtracting this from (4.34), integrating, and applying the triangle and Gronwall inequalities gives (4.37). To make this rigorous one would proceed as in the proof of Theorem 2. We will pursue this type of argument in Section 5 when we discuss second-order averaging.

The equilibrium solutions of the averaged problem, $v = v_e$, are precisely those given by (4.32). To see this, let $v_{e1} = r \cos\theta$ and $v_{e2} = r \sin\theta$; then $\bar{f}(v_e, \tau_0) = 0$ gives (4.32), with $\chi = \tau_0 + \theta$ (and (4.33), with x replaced by v_e, gives the periodic solutions obtained from regular perturbation theory). Linearizing about the equilibrium solutions and solving the associated eigenvalue problems gives the eigenvalues

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} a + ib \\ a - ib \end{pmatrix}, \begin{pmatrix} a + ic \\ a - ic \end{pmatrix}, \begin{pmatrix} -0.0369 \\ 0.0335 \end{pmatrix},$$

where $a = -0.0017$, $b = 0.0389$, and $c = 0.0822$. Thus the linearization has two asymptotically stable spirals and a saddle point. Since the equilibria are hyperbolic, (4.36) has the same structure as the linearization near them. This is illustrated in Figure 1, where we show the phase-plane portrait for $\tau_0 = -\pi/2$. The four non-constant trajectories shown are the stable and unstable manifolds of the saddle point. The manifolds were computed numerically, using initial conditions near the saddle point on the eigenvectors: $(-0.4411, 1.0)$ and $(0.0717, 1.0)$. The basins of attraction associated with the two attractors (one is shown by the shading) are clearly defined by the stable manifolds. The evolution defined by (4.36) for an arbitrary initial condition is now easily inferred.

The averaging theorem says the phase-plane portrait of (4.36) and the Poincaré section of (4.31) will be close for ϵ small (and n not too large). However, we can say more about the original problem in terms of the PM: $\mathcal{P}_{\tau_0}(\xi, \epsilon) := \varphi(\tau_0 + 2\pi, \tau_0, \xi, \epsilon) = v(\tau_0 + 2\pi, \tau_0, \xi, \epsilon) + O(\epsilon^2)$. For ϵ sufficiently small, an application of the implicit function theorem shows that the PM has fixed points $O(\epsilon^2)$ close to the equilibrium solutions of the averaged problem (and, of course, FPs of the PM are periodic solutions of (4.31)). Since the equilibrium solutions of (4.36) are hyperbolic, Theorem 4.1.1 in Reference 12 asserts that for ϵ small the phase-plane structure persists in the PM (see also References 1 and 13). Thus we have explained the dynamics of (4.30) in terms of the phase-plane portrait in Figure 1, and this in turn explains the beam splitting in the IUCF experiment in terms of the two attractors of the averaged system and their basins of attraction. We have taken $\tau_0 = -\pi/2$ in Figure 1, but clearly the beam splitting should not depend on a particular Poincaré section. In fact, the phase-plane portraits of (4.36) for different values of τ_0 are obtained simply by rotation, since the coordinate transformation $w = e^{J\gamma}v$, which is a rotation, gives $\dot{w} = \bar{f}(w, \tau_0 + \gamma)$, as can be seen from (4.35).

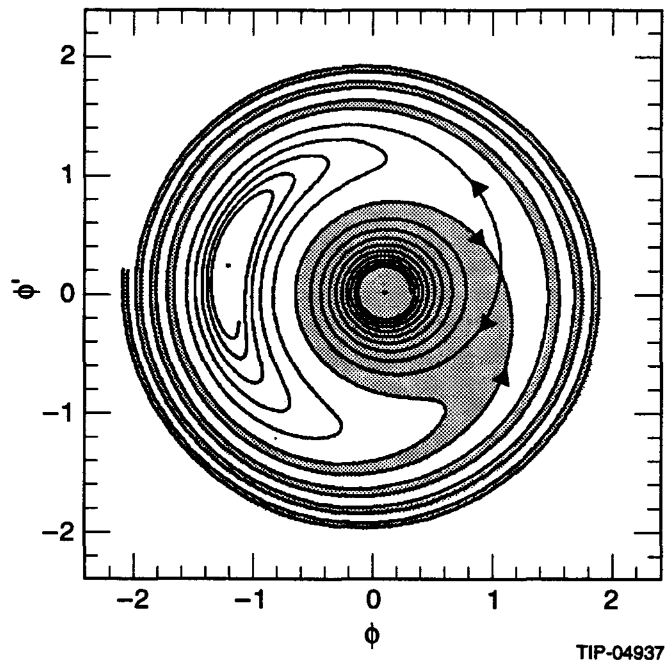


Figure 1. Phase-plane portrait of Eq. (4.36) for $\tau_0 = -\pi/2$. See text for the parameters used.

Finally, we discuss the case $\alpha = 0$ for F of Figure 1, which is relevant to the IUCF experiment as well as to the Fermilab nonlinear dynamics experiment E778. In Figure 2(a) we show the phase-plane portrait of the averaged system for $\beta = -0.0840$. The implicit function theorem can again be applied to show that the equilibrium solutions correspond to periodic solutions of (4.31). The persistence of the invariant circles for ϵ small is a deeper result that follows from the Moser Twist theorem if the frequency of the periodic solutions of the averaged system as a function of action has a non-zero derivative at zero action. We have verified that this is the case. Also there are two homoclinic orbits (dashed curves in Figure 2(a)), and these presumably do not persist in the PM. The stable and unstable manifolds most likely intersect transversely, with transcendentally small angle, giving rise to a thin stochastic layer. However, this is very difficult to prove, as the literature on the rapidly forced pendulum, a prototype problem, shows. (See, for example, Reference 14.)

As β increases from its value in Figure 2(a), a bifurcation to one equilibrium solution occurs at $\beta = \beta_c := -3(\sqrt{2}F)^{2/3}/8 = -0.0342$. Figure 2(b) shows the phase-plane portrait for $\beta = \beta_c$ and indicates a cusp structure at the bifurcation point. Figure 2(c) shows the on-resonance case ($\beta = 0$), and all solutions are periodic. Again the Moser Twist theorem can be applied to determine stability of the associated FP of the PM.

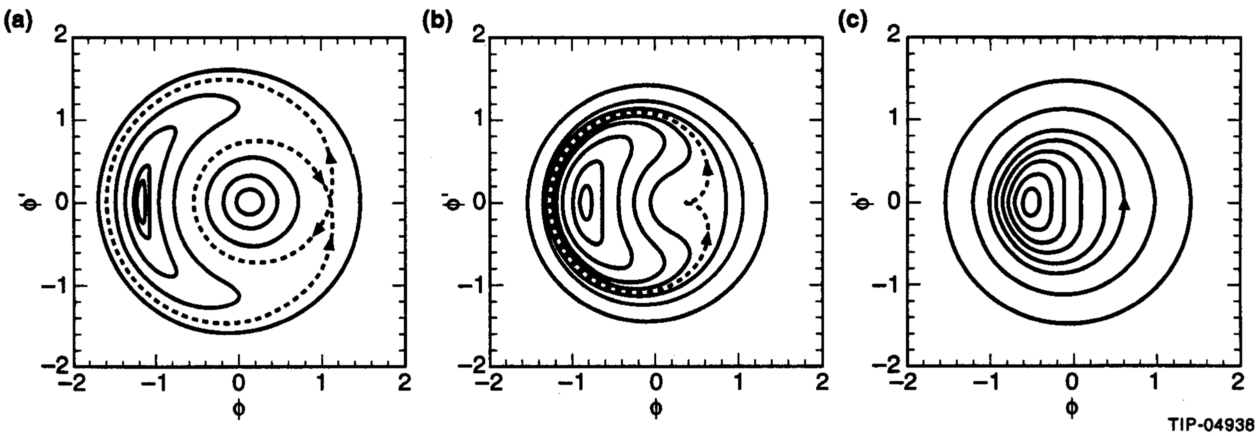


Figure 2. As in Figure 1 for $\alpha = 0$. (a) $\beta = -0.0840$; (b) $\beta = \beta_c = -0.0342$; (c) $\beta = 0$.

5. Two Second-Order Averaging Theorems

We discuss second-order averaging for the IVP:

$$z' = \epsilon f(z,s) + \epsilon^2 F(z,s), \qquad z(0,\epsilon) = z_0 . \tag{5.1}$$

Here f and F are defined on $\mathcal{U} \times \mathbb{R}$, where $\mathcal{U}$ is an open subset of $\mathbb{C}^d$ containing z_0, and are quasiperiodic in s; that is,

$$f(z,s) = \sum_{n \in \mathbf{Z}^d} \hat{f}_n(z) e^{i\langle n,\omega\rangle s} , \tag{5.2}$$

$$F(z,s) = \sum_{n \in \mathbf{Z}^d} \hat{F}_n(z) e^{i\langle n,\omega\rangle s} . \tag{5.3}$$

In what follows $\bar{f}$ and $\tilde{f}$ will denote the average and zero mean parts of a quasiperiodic function and $\mathcal{I}\tilde{f}$ will denote the zero mean integral of a zero mean quasiperiodic function. Smoothness conditions will be discussed later. We include F for two reasons: (1) it is necessary in the resonance discussions and (2) it is easy to handle and we obtain a more general result.

Our first goal is to understand a special class of solutions $z(s,\epsilon)$ of (5.1) on $O(1/\epsilon)$ time intervals in terms of a simpler autonomous system constructed from f and F. In addition, we discuss an extension to $O(1/\epsilon^2)$ time intervals and resonance. Averaging is essentially a transformation procedure that leads to a systematic perturbation expansion of certain solutions of (5.1), and leads to better and better approximations on time intervals of $O(1/\epsilon)$ as discussed in References 7, 8 and 15.

We look for a transformation

$$z = y + \epsilon P(y,s) , \qquad P(y,0) = 0 , \tag{5.4}$$

quasiperiodic in s, which transforms (5.1) into the IVP

$$y' = \epsilon V^{(1)}(y) + \epsilon^2 V(y,s) + \epsilon^3 H(y,s,\epsilon), \qquad y(0,\epsilon) = z_0\,, \tag{5.5}$$

whose solution we denote by $y(s,\epsilon)$ and where we require $V(y,s)$ to be quasiperiodic in s. The averaging approximation will be given by

$$z(s,\epsilon) \cong w(s,\epsilon) := v(s,\epsilon) + \epsilon P\Big(v(s,\epsilon),s\Big)\,, \tag{5.6}$$

where v is defined by the IVP

$$v' = \epsilon V^{(1)}(v) + \epsilon^2 \bar{V}(v), \qquad v(0,\epsilon) = z_0\,. \tag{5.7}$$

Equation (5.7) is obtained from (5.5) by ignoring the $O(\epsilon^3)$ term and replacing V by its average. We will show, under suitable restrictions on f and F, that $y(s,\epsilon) = v(s,\epsilon) + O(\epsilon^2)$ uniformly in s on $O(1/\epsilon)$ s-intervals. It will follow from (5.4) and (5.6) that $z(s,\epsilon) = w(s,\epsilon) + O(\epsilon^2)$ uniformly in s on $O(1/\epsilon)$ s-intervals; thus we have a second-order perturbation procedure based on a solution of the autonomous problem (5.7).

To determine $V^{(1)}$, V, P and H we differentiate (5.4) along the solutions of (5.5) and insert into (5.1) to obtain

$$\begin{aligned}\Big[1 + \epsilon D_1 P(y,s)\Big]\Big[\epsilon V^{(1)}(y) + \epsilon^2 V(y,s) + \epsilon^3 H(y,s,\epsilon)\Big] + \epsilon D_2 P(y,s) \\ = \epsilon f\Big(y + \epsilon P(y,s), s\Big) + \epsilon^2 F\Big(y + \epsilon P(y,s), s\Big)\,. \end{aligned} \tag{5.8}$$

Here and in the following D_i^j denotes the jth derivative of the ith argument, be it vector or scalar. We now show that (5.8) determines the four unknown functions. Expanding f and F on the right hand side of (5.8) by Taylor's theorem, we obtain

$$f\Big(y + \epsilon P(y,s), s\Big) = f(y,s) + \epsilon D_1 f(y,s) P(y,s) + \epsilon^2 \rho_1(y,s,\epsilon) \tag{5.9a}$$

$$F\Big(y + \epsilon P(y,s), s\Big) = F(y,s) + \epsilon \rho_2(y,s,\epsilon)\,. \tag{5.9b}$$

These equations can be viewed as defining ρ_1 and ρ_2; then under suitable conditions Taylor's theorem gives ρ_1 and ρ_2 as order one (in ϵ) functions. Using (5.9) in (5.8) and equating powers of ϵ gives

$$V^{(1)}(y) + D_2 P(y,s) = f(y,s)\,, \tag{5.10a}$$

$$V(y,s) = D_1 f(y,s) P(y,s) + F(y,s) - D_1 P(y,s) V^{(1)}(y)\,, \tag{5.10b}$$

$$\Big(1 + \epsilon D_1 P(y,s)\Big) H(y,s,\epsilon) = -D_1 P(y,s) V^{(2)}(y) + \rho_1(y,s,\epsilon) + \rho_2(y,s,\epsilon)\,. \tag{5.10c}$$

Equations (5.10) are solved sequentially, treating y and s as independent variables. Since P is to be quasiperiodic in s, $D_2 P$ has zero s-mean; therefore, $V^{(1)}(y)$ must be the average of f (whence the term "method of averaging"). Equation (5.10a) along with $P(y,0)=0$ then determines P, and V is then defined by (5.10b). Finally, H is determined by (5.10c) in combination with (5.9) as long as $1+\epsilon D_1 P$ is invertible. Thus, recalling (2.22), we have

$$V^{(1)}(v)=\bar{f}(v)=\sum_{n\in\mathcal{M}}\hat{f}_n(v)\,, \tag{5.11a}$$

$$\begin{aligned} P(v,s) &= \int_0^s \tilde{f}(v,t)dt=\sum_{n\notin\mathcal{M}}\frac{\hat{f}_n(v)}{i\langle n,\omega\rangle}(e^{i\langle n,\omega\rangle s}-1) \\ &=: \sum_{n\notin\mathcal{M}}\frac{\hat{f}_n(v)}{i\langle n,\omega\rangle}\,e^{i\langle n,\omega\rangle s}+p(v)\,, \end{aligned} \tag{5.11b}$$

$$\begin{aligned} \bar{V}(v) &= \overline{D_1 f(v,s)\,P(v,s)}-\overline{D_1 P(v,s)}\,V^{(1)}(v)+\overline{F(v,s)} \\ &= \sum_{n\in\mathcal{M}}\sum_{m\notin\mathcal{M}}\frac{D\hat{f}_{n-m}(v)\hat{f}_m(v)}{i\langle m,\omega\rangle}+DV^{(1)}(v)p(v)-Dp(v)V^{(1)}(v)+\sum_{n\in\mathcal{M}}\hat{F}_n(v)\,. \end{aligned} \tag{5.11c}$$

Here we have made use of the fact that

$$D_1 f(v,s)P(v,s)=D_1 f(v,s)p(v)+\sum_{n\in\mathbf{Z}^d}\left(\sum_{m\notin\mathcal{M}}\frac{D\hat{f}_{n-m}(v)\hat{f}_m(v)}{i\langle m,\omega\rangle}\right)e^{i\langle n,\omega\rangle s}\,.$$

We now have completely determined a candidate w for an approximate solution of the IVP (5.1), namely,

$$z(s,\epsilon)\simeq w(s,\epsilon)=v(s,\epsilon)+\epsilon P\Big(v(s,\epsilon),\epsilon\Big)\,, \tag{5.12}$$

where P is defined by (5.11b), and $v(s,\epsilon)$ is the solution of the autonomous IVP (5.7) with $V^{(1)}$ and $\bar{V}$ defined by (5.11a) and (5.11c).

In order to state the theorem that specifies the relation between $z(s,\epsilon)$ and $w(s,\epsilon)$, we need some general conditions on f and F as defined by (5.2) and (5.3). We make the following rather mild assumptions:

(A) $\hat{f}_n$, $D\hat{f}_n$, $D^2\hat{f}_n$, $\hat{F}_n$, and $D\hat{F}_n$ exist on $\mathcal{U}$, and for each compact subset $K\subset\mathcal{U}$ each of the following series converge: $\sum_{n\in\mathbf{Z}^d}\sup_{z\in K}|D^j\hat{f}_n(z)|$ and $\sum_{n\notin\mathcal{M}}\sup_{z\in K}D^j\hat{f}_n(z)/|\langle n,\omega\rangle|$ for $j=0,1,2$ and $\sum_{n\in\mathbf{Z}^d}\sup_{z\in K}\left|D^j\hat{F}_n(z)\right|$ for $j=0,1$.

(B) $g(v,s):=\tilde{V}(y,s)$ satisfies (2.24) and (2.25), where V and thus $\tilde{V}$ are defined by (5.10b).

(C) H is well-defined by (5.10c) and bounded on compact $\bar{\mathcal{U}}_2\subset\mathcal{U}$, where $\mathcal{U}_2$ is defined in the proof of the theorem.

Note that (A) implies that the rhs of the ODE in (5.1) is locally z-Lipschitz on $\mathcal{U}$ and continuous on $\mathcal{U} \times \mathbb{R}$, hence the IVP (5.1) has a unique maximal solution in $\mathcal{U}$.

An important concept in our theorem and its proof is the guiding solution,[7,15] which is defined by the ϵ-independent IVP,

$$\frac{du}{d\tau} = V^{(1)}(u), \qquad u(0) = z_0 . \tag{5.13}$$

By (A), $V^{(1)}$ is locally Lipschitz on $\mathcal{U}$. Therefore, there exists a function $T_1 : \mathcal{U} \longrightarrow (0, \infty]$, such that $[0, T_1(z_0))$ is the maximum forward interval of existence of the solution of the IVP (5.13). Standard continuation arguments show that for $T_1(z_0) < \infty$, $u(\tau)$ approaches the boundary of $\mathcal{U}$ as $\tau \nearrow T_1(z_0)$. We can now state the basic existence, uniqueness, and approximation theorem of second-order averaging.

Theorem 3 (Second-Order Averaging): *Fix $z_0 \in \mathcal{U}$ and let $T < T_1(z_0)$. Then there exist positive numbers $\epsilon_0 = \epsilon_0(z_0, T)$ and $C(z_0, T)$ such that for $0 < \epsilon < \epsilon_0$ and for $0 \leq t \leq T/\epsilon$, the IVPs* (5.1) *and* (5.7) *have unique solutions in $\mathcal{U}$ and*

$$\|z(s, \epsilon) - w(s, \epsilon)\| \leq C\epsilon^2 .$$

Proof: Let $\mathcal{S} = \{z \in \mathbb{C}^n | z = u(\tau), 0 \leq \tau \leq T\} \subset \mathcal{U}$ be the compact set defined by the guiding solution. Let $\mathcal{U}_1$ and $\mathcal{U}_2$ be open-bounded subsets of $\mathcal{U}$ satisfying

$$\mathcal{S} \subset \mathcal{U}_1 \subset \bar{\mathcal{U}}_1 \subset \mathcal{U}_2 \subset \bar{\mathcal{U}}_2 \subset \mathcal{U} .$$

The important point here is that there is a positive distance between the boundaries of these sets. The proof has three steps, involving several restrictions on ϵ: (i) $v(s, \epsilon)$ exists uniquely in $\mathcal{U}_1$ for $0 \leq s \leq T/\epsilon$ and for ϵ sufficiently small; (ii) $y(s, \epsilon)$ as defined by (5.5) exists uniquely in $\mathcal{U}_2$ for $0 \leq s \leq T/\epsilon$ and for ϵ sufficiently small and satisfies $\|y(s, \epsilon) - v(s, \epsilon)\| = O(\epsilon^2)$; (iii) $z(s, \epsilon)$ as defined by (5.4) and (5.5) is the unique solution of (5.1) in $\mathcal{U}$ for $0 \leq s \leq T/\epsilon$ and satisfies $\|z(s, \epsilon) - w(s, \epsilon)\| \leq C\epsilon^2$. The restrictions on ϵ will be indicated in the proof, and the ϵ_0 of the theorem will be the minimum ϵ in the restrictions.

To prove (i) we first note that (A) implies that $\bar{V}$ is locally Lipschitz on $\mathcal{U}$ since $D_1 f$, P, F, $D_1 P$, and $V^{(1)}$ are. Therefore $v(s, \epsilon)$ exists uniquely in $\mathcal{U}_1$ on its maximum forward interval of existence $\big[0, \beta_1(\epsilon)\big)$. A comparison of the guiding solution u with v obtained by subtracting (5.13) from (5.7), integrating, using the triangle inequality, using the facts that $V^{(1)}$ is Lipschitz and $\bar{V}$ is bounded on $\mathcal{U}_1$, and applying the Gronwall inequality, shows there exists an ϵ_1 such that for $\epsilon < \epsilon_1$, $\beta_1(\epsilon) > T/\epsilon$, and (i) is proved.

To prove (ii) we first note that by (A) $P(y, s)$ is bounded on $\bar{\mathcal{U}}_2$; thus there exists an ϵ_2 such that for $\epsilon < \epsilon_2$, $y + \epsilon P(y, s) \in \mathcal{U}$ for $y \in \bar{\mathcal{U}}_2$; thus ρ_1 and ρ_2 in (5.9) are well-defined on $\bar{\mathcal{U}}_2$. Since P is bounded, there exists an ϵ_3 such that for $\epsilon < \epsilon_3$, $1 + \epsilon D_1 P(y, s)$ is invertible on $\bar{\mathcal{U}}_2$; thus H is well-defined on $\bar{\mathcal{U}}_2$ by assumption (C). Thus the rhs of (5.5) is well-defined and locally y-Lipschitz on $\mathcal{U}_2$, and we let $\big[0, \beta(\epsilon)\big)$ denote the maximal forward interval of existence of the unique solution of (5.5) in $\mathcal{U}_2$ and $J = \big[0, \beta(\epsilon)\big) \cap [0, T/\epsilon]$. Subtracting (5.7) from (5.5), integrating and using the triangle inequality gives

$$
\begin{aligned}
\Big\|y(s,\epsilon) - v(s,\epsilon)\Big\| \leq \epsilon \int_0^s \Big\|V^{(1)}\big(y(t,\epsilon)\big) - V^{(1)}\big(v(t,\epsilon)\big)\Big\|\, dt \\
+\epsilon^2 \int_0^s \Big\|\bar{V}\big(y(t,\epsilon)\big) - \bar{V}\big(v(t,\epsilon)\big)\Big\|\, dt + \epsilon^2 \Big\|\int_0^s g\big(y(t,\epsilon),t\big)\, dt\Big\| \\
+\epsilon^3 \int_0^s \Big\|H\big(y(t,\epsilon),t,\epsilon\big)\Big\|\, dt\,. \qquad (5.14)
\end{aligned}
$$

Using (B) and the result at the end of Section 2 gives $\Big\|\int_0^s g\big(y(t,\epsilon),t\big)dt\Big\|$ bounded on J, and (C) gives H bounded on $\mathcal{U}_2$; thus the last two terms in (5.14) are bounded by $M\epsilon^2$ on J for some constant M. Letting L be the Lipschitz constant for $V^{(1)}$ and $\bar{V}$ on $\mathcal{U}_2$ gives

$$\Big\|y(s,\epsilon) - v(s,\epsilon)\Big\| \leq M\epsilon^2 + L(\epsilon+\epsilon^2)\int_0^s \Big\|y(t,\epsilon) - v(t,\epsilon)\Big\|\, dt$$

for $s \in J$. An application of the Gronwall inequality yields $\Big\|y(s,\epsilon) - v(s,\epsilon)\Big\| \leq M\epsilon^2 \exp L(1+\epsilon)\epsilon s$ for $s \in J$, and ϵ restricted as above. Thus there exists an ϵ_4 such that for $\epsilon < \epsilon_4$, $\Big\|y(s,\epsilon)-v(s,\epsilon)\Big\|$ is less than the distance between $\partial\mathcal{U}_1$ and $\partial\mathcal{U}_2$. Therefore, y stays inside $\mathcal{U}_2$, which by the continuation theorem implies $\beta(\epsilon) > T/\epsilon$. Thus (ii) is proved.

To prove (iii) we note from (5.4) and (5.8) that

$$
\begin{aligned}
z' &= \Big[1 + \epsilon P(y,s)\Big]\Big[\epsilon V^{(1)}(y) + \epsilon^2 V(y,s) + \epsilon^3 H(y,s,\epsilon)\Big] + \epsilon D_2 P(y,s) \\
&= \epsilon f\Big(y + \epsilon P(y,s), s\Big) + \epsilon^2 F\Big(y + \epsilon P(y,s), s\Big)\,,
\end{aligned}
$$

and for $y \in \mathcal{U}_2$, $y + \epsilon P(y;s) \in \mathcal{U}$, which proves the first part of (iii). Finally, $\Big\|z(s,\epsilon) - w(s,\epsilon)\Big\| \leq \Big\|y(s,\epsilon) - v(s,\epsilon)\Big\| + \epsilon\Big\|P\big(y(s,\epsilon),s\big) - P\big(v(s,\epsilon),s\big)\Big\| \leq (1+\epsilon L)\Big\|y(s,\epsilon) - v(s,\epsilon)\Big\|$, where L is the Lipschitz constant for P on $\mathcal{U}_2$. Thus, using (ii), (iii) is proved and this completes the proof of the theorem with $\epsilon_0 = \mathrm{Min}(\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4)$. □

Remarks:

(1) In summary, the solution of the IVP (5.1) is given by

$$z(s,\epsilon) = v(s,\epsilon) + \epsilon P\Big(v(s,\epsilon), s\Big) + O(\epsilon^2) \qquad (5.15)$$

for $\epsilon < \epsilon_0$ and $0 \leq s \leq T/\epsilon$, where T is defined by the solution of (5.13) in $\mathcal{U}$; P, $V^{(1)}$ and $\bar{V}$ are defined by (5.11) and $v(s,\epsilon)$ is defined by the IVP (5.7).

(2) In the case where $\bar{f} = 0$, $V^{(1)} = 0$, and $\bar{V}(v) = \overline{D_1 f(v,s)\mathcal{I} f(v,s)} + \overline{F(v,s)} = \sum_{n\in\mathcal{M}, m\notin\mathcal{M}} \frac{D\hat{f}_{n-m}(v)\hat{f}_m(v)}{i\langle m,\omega\rangle} + \sum_{n\in\mathcal{M}} \hat{F}_n(v)$. Thus, v is defined by

$$v' = \epsilon^2 \bar{V}(v)\,, \qquad v(0) = z_0\,. \tag{5.16}$$

The approximation for z is again given by (5.15).

(3) It is easy to see that the first-order averaging theorem in Section 2 is a special case of the second-order theorem under sufficient smoothness. For example, when $\bar{f} = 0$ we obtain from Theorem 2 that $z(s,\epsilon) = z_0 + O(\epsilon)$ for $0 \le s \le T/\epsilon$, which is contained in Remark (2) above, since $v(s,\epsilon)$ changes by no more than $O(\epsilon)$ on $O(1/\epsilon)$ s-intervals.

Averaging on $O(1/\epsilon^2)$ s-intervals

A natural question in averaging is when can results be obtained on time intervals longer than $O(1/\epsilon)$. We now show this to be the case when $\bar{f} = 0$. Our goal is to obtain an $O(\epsilon)$ approximation on $O(1/\epsilon^2)$ s-intervals.

Here we define our basic approximation $v(s,\epsilon) = u(\epsilon^2 s)$, where u is defined by the ϵ-independent IVP:

$$\frac{du}{d\tau} = \bar{V}(u)\,, \qquad u(0) = z_0\,, \tag{5.17}$$

which is (5.7) in scaled time. Using $\bar{f} = 0$, (5.11c) becomes

$$\bar{V}(u) = \overline{V(u,s)} = \overline{D_1 f(u,s)\mathcal{I} f(u,s)} + \overline{F(u,s)}\,.$$

We make the same assumptions as in the previous theorem; thus, $\bar{V}$ is locally Lipschitz on $\mathcal{U}$ and there exists a function $T_2 : \mathcal{U} \longrightarrow (0,\infty]$, where $[0, T_2(z_0))$ is the maximum forward interval of existence of the solution of the IVP (5.17). Standard continuation arguments show that for $T_2(z_0) < \infty$, $u(\tau)$ approaches the boundary of $\mathcal{U}$ as $\tau \nearrow T_2(z_0)$. We can now state the basic existence, uniqueness, and approximation theorem on $O(1/\epsilon^2)$ s-intervals in this special case.

Theorem 4 (Second-Order Averaging for $\bar{f} = 0$): *Fix $z_0 \in \mathcal{U}$ and let $T < T_2(z_0)$. Then there exist positive numbers $\epsilon_0 = \epsilon_0(z_0,T)$ and $C(z_0,T)$ such that $0 < \epsilon < \epsilon_0$ implies that for $0 \le s \le T/\epsilon^2$, the IVP (5.1) has a unique solution in $\mathcal{U}$ and*

$$\|z(s,\epsilon) - u(\epsilon^2 s)\| \le C\epsilon\,.$$

Proof: Let $\mathcal{S} = \{z \in \mathbf{C}^n | z = u(\tau), 0 \le \tau \le T\}$ be the compact set defined by the approximating solution (5.17), and let $\mathcal{U}_1$ be an open-bounded subset of $\mathcal{U}$ satisfying $\mathcal{S} \subset \mathcal{U}_1 \subset \bar{\mathcal{U}}_1 \subset \mathcal{U}$. Also the assumptions A, B, and C are in force.

The proof is quite similar to the proof of Theorem 3 and has two steps that involve restrictions on ϵ: (i) $y(s,\epsilon)$ as defined by (5.5) exists uniquely in $\mathcal{U}_1$ for $0 \le s \le T/\epsilon^2$ and for ϵ sufficiently small and satisfies $\left\|y(s,\epsilon) - u(\epsilon^2 s)\right\| = O(\epsilon)$ on the same interval; (ii) z as defined by (5.4) and (5.5) is the unique solution of (5.1) in $\mathcal{U}$ for $0 \le s \le T/\epsilon^2$ and satisfies $\left\| z(s,\epsilon) - u(\epsilon^2 s) \right\| \le C\epsilon$ at each such s for ϵ sufficiently small.

The proof of (i) is basically the same as the proof of (ii) in Theorem 3, except that (1) $V^{(1)}$ is missing in (5.14), and (2) the argument for the bound on the g term is different. The latter uses the fact that $y' = O(\epsilon^2)$ in the argument at the end of Section 2 to give the boundedness of the g term on $O(1/\epsilon^2)$ s-intervals. Equation (5.14) then yields $\left\| y(s,\epsilon) - u(\epsilon^2 s) \right\| \le M\epsilon + L\epsilon^2 \int_0^s \left\| y(t,\epsilon) - u(\epsilon^2 t) \right\| dt$, and the result follows. The proof of (ii) is identical to the proof of (iii) in Theorem 3. $\square$

Remark: In summary, the solution of the IVP (5.1) is given by

$$z(s,\epsilon) = u(\epsilon^2 s) + O(\epsilon)$$

for $\epsilon < \epsilon_0$ and $0 \le s \le T/\epsilon^2$, where T and u are defined by the IVP (5.17).

Resonance at Second Order

We now consider the case of resonance at second order for (5.1) with $F = 0$ and (5.2) with $\hat{f}_0 = 0$. Recall that $\bar{V}(y) = \overline{D_1 f(u,s) \mathcal{I} f(u,s)}$, and we proceed as in the case of first-order resonance near the end of Section 3. We suppose that $D_1 f(y,s) \mathcal{I} f(y,s)$ has resonant terms when $\omega_2 = \omega_{20}$ but f does not. As in the case of first-order resonance, we analyze the motion in a neighborhood of ω_{20}. Because we are at second order we let

$$\omega_2 = \omega_{20} + \epsilon^2 a ,$$

and in analog with (3.22) define

$$f_r(z,\tau,s) := g(z, \omega_1 s, \omega_{20} s + a\tau, \omega_3 s) . \tag{5.18}$$

The IVP (5.1) can now be written

$$z' = \epsilon f_r(z,\tau,s) , \qquad z(0) = z_0 , \tag{5.19a}$$

$$\tau' = \epsilon^2 , \qquad \tau(0) = 0 , \tag{5.19b}$$

which is analogous to (3.23) and is in the form of (5.1) with z replaced by $(z,\tau)^T$, f replaced by $(f_r, 0)^T$, and F replaced by $(0,1)^T$. We now apply Theorem 4. Averaging does not change the τ equation (5.19b), and after solving the averaged equation corresponding to τ the averaged IVP can be replaced by

$$\frac{du}{d\tau} = V(u,\tau) , \qquad u(0) = z_0 , \tag{5.20}$$

where

$$V(u,\tau) = \overline{D_1 f_r(u,\tau,s) \mathcal{I} f_r(u,\tau,s)} . \tag{5.21}$$

Here the indicated averaging is taken over s with u and τ fixed. Theorem 4 then gives $z(s,\epsilon) = u(\epsilon^2 s) + O(\epsilon)$ on $O(1/\epsilon^2)$ s-intervals. Theorem 3 can also be applied to give an $O(\epsilon^2)$ approximation on $O(1/\epsilon)$, but we won't pursue that here.

6. Beam Dynamics Examples at Second Order

Here we discuss four examples in the context of second-order averaging.

Example E: Sextupole with ν, 2ν, 3ν, and 4ν non-integers.

We now continue our discussion of the sextupole:

$$\begin{aligned} x'' + K(s)x &= \epsilon\frac{1}{2}S(s)x^2 , \\ x(0) = x_0, \qquad x'(0) &= x_0' , \end{aligned} \tag{6.1}$$

which began in Section 4. Letting $\underline{x} = (x_1, x_2)^T = (x, x')^T$ and $\underline{x} = \Psi(s)z$, we obtain $\dot{z} = \epsilon f(z,s)$, $z(0) = z_0$, where

$$f_1(z,s) = i\left[e_1(s)e^{i\omega_2 s}z_1^2 + e_{-1}(s)e^{-i\omega_2 s}2z_1z_2 + e_{-3}(s)e^{-3i\omega_2 s}z_2^2\right] , \tag{6.2}$$

and

$$e_\ell(s) = -\frac{1}{4}\beta^{3/2}(s)S(s)e^{i\ell\psi_p(s)} =: \sum_n e_{\ell n}e^{in\omega_1 s} \tag{6.3}$$

as before. Clearly $\bar{f} = 0$ for ν and 3ν non-integer. The second-order averaging Theorems 3 and 4 give

$$\underline{x}(s,\epsilon) = \Psi(s)z(s,\epsilon) = \begin{cases} \Psi(s)u(\epsilon^2 s) + \epsilon\Psi(s)P\Big(u(\epsilon^2 s), s\Big) + O(\epsilon^2), & 0 \le s \le T/\epsilon \\ \Psi(s)u(\epsilon^2 s) + O(\epsilon), & 0 \le s \le T/\epsilon^2 . \end{cases} \tag{6.4}$$

Since $\bar{f}$ and F are zero, u and P are defined by

$$\frac{du}{d\tau} = V(u) , \qquad u(0) = z_0 = \begin{pmatrix} Ae^{i\phi} \\ cc \end{pmatrix} , \tag{6.5}$$

$$P(u,s) = \int_0^s f(u,t)dt = \mathcal{I}f(u,s) + p(u) , \tag{6.6}$$

and

$$V(u) := V^{(2)}(u) = \overline{D_1 f(u,s)\mathcal{I}f(u,s)} , \tag{6.7}$$

where (6.5) is (5.7) with $V^{(1)} = 0$ and with scaled time $\tau = \epsilon^2 s$, and (6.6) and (6.7) are obtained from (5.11b) and (5.11c).

In the nonresonant case, where ν, 2ν, 3ν, and 4ν are non-integers the calculation in (6.7) yields

$$V_1(v) = i\gamma v_1^2 v_2 , \tag{6.8}$$

where γ is a real constant. Before outlining the calculation of (6.8) we complete the discussion of the approximate solution. The averaged IVP (6.5) becomes $v_1' = i\gamma v_1^2 v_1^*$, $v_1(0) = Ae^{i\phi}$ and its complex conjugate. It is easy to check that $v_1 v_1^*$ is conserved; therefore,

$$v_1(\epsilon^2 s) = A \exp i\Big(\epsilon^2 \gamma A^2 s + \phi\Big) . \tag{6.9}$$

Thus, we obtain from (6.4) that

$$\begin{aligned}
x(s) = x_1(s) &= 2\sqrt{\beta(s)} A \cos\Big(\psi(s) + \epsilon^2\gamma A^2 s + \phi\Big) \\
&+ \frac{1}{2}\epsilon\sqrt{\beta(s)} A^2 \int_0^s \beta^{3/2}(\tau) S(\tau) \Big[\sin\Big(\psi(\tau) + \psi(s) + 2\epsilon^2\gamma A^2 s + 2\phi\Big) \\
&+ 2\sin\Big(-\psi(\tau) + \psi(s)\Big) + \sin\Big(-3\psi(\tau) + \psi(s) - 2\epsilon^2\gamma A^2 s - 2\phi\Big)\Big] d\tau + O(\epsilon^2) ,
\end{aligned} \tag{6.10}$$

for $0 \le s \le T/\epsilon$ and

$$x(s) = 2\sqrt{\beta(s)} A \cos\Big(\psi(s) + \epsilon^2\gamma A^2 s + \phi\Big) + O(\epsilon) , \tag{6.11}$$

for $0 \le s \le T/\epsilon^2$. In (6.10) notice the tune shift correction in the first term, but also notice that a consistent expansion to $O(\epsilon^2)$ on $0 \le s \le T/\epsilon$ must also include the second term. Even though the tune shift correction is only $O(\epsilon)$ on $O(1/\epsilon)$ s-intervals, it is needed to obtain the $O(\epsilon^2)$ approximation. In (6.11) the tune shift can be ignored on $O(1/\epsilon)$ s-intervals because the approximation is good only to $O(\epsilon)$. These results should be contrasted with the octupole, where the tune shift becomes important when s is of the order of the reciprocal of the octupole strength and there apparently is no result analogous to (6.11).

To compute $V_1(z)$ we note from (6.7) that

$$V_1(z) = \overline{\frac{\partial f_1(z,s)}{\partial z_1} \mathcal{I} f_1(z,s)} + \overline{\frac{\partial f_1}{\partial z_2} \Big(\mathcal{I} f_1(z,s)\Big)^*} , \tag{6.12}$$

where we have used the fact that $\mathcal{I} f_2(z,s) = \Big(\mathcal{I} f_1(z,s)\Big)^*$. These averages require averages of

$$\begin{aligned}
a_{k\ell}(s) &:= e_k(s) e^{ik\omega_2 s} \mathcal{I}\Big(e_\ell(s) e^{i\ell\omega_2 s}\Big) \\
&= \sum_{n,n'} \frac{e_{kn} e_{\ell n'}}{i(n'\omega_1 + \ell\omega_2)} e^{i[(n+n')\omega_1 + (\ell+k)\omega_2]s} .
\end{aligned} \tag{6.13}$$

Notice that $e_{-\ell}(s) = e_\ell(s)^*$, which implies that $e_{kn}^* = e_{-k,-n}$ and $a_{k,n}(s)^* = a_{-k,-n}(s)$. Also note that k and ℓ will be ± 1 or ± 3. Now we have assumed that $\bar{f} = 0$, and so ν and 3ν must not be integers; thus, the denominators in (6.13) are nonzero and the $a_{k\ell}$ are well-defined, assuming $\sum_{nn'} |e_{kn}||e_{\ell n'}|$ is finite. Furthermore, $\ell + k$ can take the values ± 2 and ± 4, and thus if 2ν and 4ν are not integers we can do the $\omega_1 s$ and $\omega_2 s$ averages in (6.13) separately, which yields

$$\overline{a_{k\ell}} = \begin{matrix} 0, & \ell \ne -k \\ i\gamma_k, & \ell = -k , \end{matrix} \tag{6.14a}$$

where

$$\gamma_k = \sum_n \frac{|e_{kn}|^2}{n\omega_1 + k\omega_2} = \frac{1}{\omega_1}\sum_n \frac{|e_{kn}|^2}{n+\nu k} = -\gamma_{-k} \tag{6.14b}$$

is real and k and ℓ take the values ± 1 and ± 3. Multiplying out the terms in (6.12) and making use of (6.14) yields (6.8) with $\gamma = -6\gamma_1 - 2\gamma_3$.

In the case where $S(s) = S_0\delta_p(s)$, $e_\ell(s) = -\frac{1}{4}\beta_0^{3/2} S_0\,\delta_p(s)$ and $e_{kn} = -\frac{1}{4C}\beta_0^{3/2} S_0$; that is, the latter is independent of both k and n. From (6.14b),

$$\gamma_k = \frac{1}{16C^2\omega_1}\beta_0^3 S_0^2 \sum_n \frac{1}{n+k\nu} = \frac{\beta_0^3 S_0^2}{32C}\cot \pi k\nu\,, \tag{6.15}$$

and thus

$$\gamma = -\frac{\beta_0^3 S_0^2}{16C}[3\cot\pi\nu + \cot 3\pi\nu]\,.$$

In addition, the integral term in (6.10) can be evaluated. Again it should be pointed out that the averaging theorems do not apply for delta function perturbations.

Example F: Sextupole near the fourth integer resonance.

Here we consider the sextupole in the neighborhood of the fourth integer resonance. Let

$$\omega_2 = \omega_{20} + \epsilon^2 a; \qquad 4\omega_{20} = 4\nu_0\omega_1 = M\omega_1\,, \tag{6.16}$$

where M is an integer, but $\nu_0 = M/4$, $2\nu_0 = M/2$, and $3\nu_0 = M/3$ are not. We follow the resonance discussion at the end of Section 5. From (5.18) and (6.2),

$$f_{r1}(z,\tau,s) = i\Big[e_1(s)\,e^{i\omega_{20}s}\,e^{ia\tau}z_1^2 + e_{-1}(s)\,e^{-i\omega_{20}s}\,e^{-ia\tau}2z_1z_2 + e_{-3}(s)\,e^{-3i\omega_{20}s}\,e^{-i3a\tau}z_2^2\Big]\,, \tag{6.17}$$

which is similar to the equation after (4.9). Equation (5.21) (see also (6.12)) now becomes

$$\begin{aligned} V_1(z,\tau) &= \overline{\frac{\partial f_1(z,\tau,s)}{\partial z_1}\mathcal{I}f_1(z,\tau,s)} + \overline{\frac{\partial f_1(z,\tau,s)}{\partial z_2}\big(\mathcal{I}f_1(z,\tau,s)\big)^*} \\ &= I + II\,, \end{aligned} \tag{6.18}$$

where $\mathcal{I}$ is as before the zero mean integral with respect to s (holding τ and z fixed) and the average is with respect to s (holding τ and z fixed). If we replace ω_2 by ω_{20} in (6.13) and define γ_{kl} by

$$\overline{a_{kl}} =: i\gamma_{kl}\,, \tag{6.19}$$

then $\gamma_{kl}^* = -\gamma_{-k-l}$, the nonzero γs are $\gamma_{k,-k}$, $\gamma_{1,3}$, $\gamma_{-1,-3}$, $\gamma_{3,1}$, $\gamma_{-3,-1}$ and

$$\begin{cases} \gamma_{k,-k} =: \gamma_k\,, \\ \gamma_{13} = -\sum_n \dfrac{e_{1,-M-n}e_{3n}}{(n+3\nu_0)\omega_1} = -\gamma^*_{-1,-3}\,, \\ \gamma_{31} = -\sum \dfrac{e_{3,-M-n}e_{1n}}{(n+\nu_0)\omega_1} = -\gamma^*_{-3,-1}\,, \end{cases} \tag{6.20}$$

where γ_k is defined by (6.14b). It is straightforward to calculate I and II as

$$
\begin{aligned}
I &= \left(-4\overline{a_{1,-1}(s)} - 2\overline{a_{-1,1}(s)}\right) z_1^2 z_2 - 2\overline{a_{-1,-3}(s)}\, e^{-4ia\tau} z_2^3 \\
II &= \left(4\overline{a_{-1,1}(s)} + 2\overline{a_{-3,3}(s)}\right) z_1^2 z_2 + 2\overline{a_{-3,-1}(s)}\, e^{-4ia\tau} z_2^3 \,,
\end{aligned}
$$

and thus

$$
V_1(z,\tau) = i\left(\gamma z_1^2 z_2 + \alpha e^{-4ia\tau} z_2^3\right), \tag{6.21}
$$

where $\gamma = -6\gamma_1 - 2\gamma_3$ as in Example E, and $\alpha = -2\gamma_{-1,-3} + 2\gamma_{-3,-1}$. The averaged IVP now becomes

$$
v_1' = \epsilon^2 i\left[\gamma v_1^* v_1^2 + \alpha e^{-i4\epsilon^2 as} v_1^{*3}\right], \qquad v_1(0) = Ae^{i\phi}, \tag{6.22}
$$

where γ is real and α is in general complex. As in the third integer resonance case, this can be made autonomous by the transformation $v_1 = e^{-i\epsilon^2 as}\zeta$, which gives

$$
\zeta' = i\epsilon^2[a\zeta + \gamma\zeta^*\zeta^2 + \alpha\zeta^{*3}], \qquad \zeta(0) = Ae^{i\phi}. \tag{6.23}
$$

This is analogous to (4.10) and it is easy to see that this equation has five equilibrium solutions, $\zeta = 0$ and four others, as is to be expected for the sextupole 1:4 resonance. The IVP (6.23) is equivalent to

$$
\begin{aligned}
\zeta' &= i\epsilon^2 \frac{\partial H(\zeta,\zeta^*)}{\partial \zeta^*}, \\
\zeta^{*\prime} &= -i\epsilon^2 \frac{\partial H(\zeta,\zeta^*)}{\partial \zeta},
\end{aligned} \tag{6.24}
$$

where

$$
H(\zeta,\zeta^*) = a\zeta\zeta^* + \frac{1}{2}\gamma(\zeta\zeta^*)^2 + \frac{1}{4}\alpha^*\zeta^4 + \frac{1}{4}\alpha\zeta^{*4}. \tag{6.25}
$$

It is easy to check that (6.25) is a conservation law for (6.23). If we let $\zeta(s) = \frac{1}{2}\left(X(\epsilon^2 s) + iY(\epsilon^2 s)\right)$ and $\mathcal{H}(X,Y) = H(\zeta,\zeta^*)$ then (6.24) becomes

$$
\begin{aligned}
\frac{dX}{d\tau} &= -\frac{\partial \mathcal{H}}{\partial Y}, \\
\frac{dY}{d\tau} &= \frac{\partial H}{\partial X}.
\end{aligned}
$$

These are real, Hamiltonian, and easily analyzed in the phase plane in analogy to the situation in the 1:3 resonance of Example B. Finally we apply Theorem 4 to obtain

$$
\underline{x}(s) = \Psi(s)z(s) = \Psi(s)\begin{pmatrix} e^{-i\epsilon^2 as}\left(X(\epsilon^2 s) + iY(\epsilon^2 s)\right) \\ cc \end{pmatrix} + O(\epsilon) \tag{6.26}
$$

for $0 \le s \le T/\epsilon^2$, which gives

$$x_1(s) = 2\sqrt{\beta(s)}\Big[X(\epsilon^2 s)\cos\big(\psi(s) - \epsilon^2 as\big) - Y(\epsilon^2 s)\sin\big(\psi(s) - \epsilon^2 as\big)\Big] + O(\epsilon) \tag{6.27}$$

on the same s-interval. An improved approximation on the shorter interval $[0, T/\epsilon]$ can be obtained from Theorem 3, as in Example E, by including the P function.

Example G: Sextupole with x-y coupling.

The equations of motion for the coupled sextupole can be written

$$\underline{x}' = A(s)\underline{x} + \epsilon H(\underline{x}, \omega_1 s)\,, \tag{6.28}$$

where $\underline{x} = (x_1, x_2, x_3, x_4)^T = (x, x', y, y')^T$ and

$$H(\underline{x}, \omega_1 s) = \left(0, \frac{1}{2}S(s)(x_1^2 - x_3^2), 0, -S(s)x_1x_3\right)^T . \tag{6.29}$$

The variation of parameters transformation

$$\underline{x} = \Psi(s)z, \qquad \Psi(s) = \begin{pmatrix} \Psi_x(s) & 0 \\ 0 & \Psi_y(s) \end{pmatrix} \tag{6.30}$$

gives

$$z' = \epsilon f(z,s) = \epsilon \Psi^{-1}(s) H\Big(\Psi(s)z, \omega_1 s\Big) =: \epsilon g(z, \omega_1 s, \omega_x s, \omega_y s)\,, \tag{6.31}$$

where $\omega_\ell = 2\pi\nu_\ell/C$ and ν_x and ν_y are the x and y tunes. Now

$$\begin{aligned} f_1(z,s) &= -\frac{i}{4}\sqrt{\beta_x(s)}S(s)e^{-i\psi_x(s)}(x_1^2 - x_3^2) \\ f_3(z,s) &= \frac{i}{2}\sqrt{\beta_y(s)}S(s)e^{-i\psi_y(s)}x_1x_3\,, \end{aligned} \tag{6.32}$$

where x_1 and x_3 must be replaced by

$$\begin{aligned} x_1 &= \sqrt{\beta_x(s)}\left(e^{i\psi_x(s)}z_1 + e^{-i\psi_x(s)}z_2\right) \\ x_3 &= \sqrt{\beta_y(s)}\left(e^{i\psi_y(s)}z_3 + e^{-i\psi_y(s)}z_4\right) . \end{aligned} \tag{6.33}$$

This gives

$$\begin{aligned} f_1(z,s) &= i\Big[e_1(s)e^{i\omega_x s}z_1^2 + e_{-1}(s)e^{-i\omega_x s}2z_1z_2 + e_{-3}(s)e^{-i3\omega_x s}z_2^2 \\ &+ d_{-1,2}(s)e^{-i(\omega_x - 2\omega_y)s}z_3^2 + d_{-1,0}(s)e^{-i\omega_x s}2z_3z_4 + d_{-1,-2}(s)e^{-i(\omega_x + 2\omega_y)s}z_4^2\Big] \end{aligned} \tag{6.34a}$$

and

$$\begin{aligned} f_3(z,s) &= 2i\Big[d_{1,0}(s)e^{i\omega_x s}z_1z_3 + d_{1,-2}(s)e^{i(\omega_x-2\omega_y)s}z_1z_4 \\ &\quad + d_{-1,0}(s)e^{-i\omega_x s}z_2z_3 + d_{-1,-2}(s)e^{-i(\omega_x+2\omega_y)s}z_2z_4\Big]. \end{aligned} \tag{6.34b}$$

Here the e_ℓ are defined by (6.3), and

$$\begin{aligned} d_{k\ell}(s) &= \frac{1}{4}S(s)\beta_y(s)\beta_x^{1/2}(s)\,e^{i(k\psi_{xp}(s)+\ell\psi_{yp}(s))} \\ &= \sum_m d_{k\ell m}\,e^{im\omega_1 s}. \end{aligned}$$

Since $z_2^* = z_1$ and $z_4^* = z_3$, $f_2(z,s) = f_1(z,s)^*$ and $f_4(z,s) = f_3(z,s)^*$, and it is easy to check that $\bar{f} = 0$ if ν_x, $3\nu_x$, and $\nu_x \pm 2\nu_y$ are non-integer.

Before proceeding, however, we analyze the $\nu_x + 2\nu_y$ near-integer resonance. Recall the resonance discussion at the end of Section 3 and the 1:3 resonance in Example B. Letting $\omega_x = \omega_{x0} + a_x\epsilon$, $\omega_y = \omega_{y0} + a_y\epsilon$, $\omega_{x0} = \nu_{x0}\omega_1$, $\omega_{y0} = \nu_{y0}\omega_1$, where $\nu_{x0} + 2\nu_{y0} = M$ (an integer), ν_{x0}, $3\nu_{x0}$, and $\nu_{x0} - 2\nu_{y0}$ (nonintegers) gives $\omega_x + 2\omega_y = M\omega_1 + a\epsilon$. The averaged vector field becomes

$$\begin{aligned} \bar{f}_{r1}(z,\tau) &= i\,\overline{d_{-1,-2}(s)e^{-iM\omega_1 s}}\;e^{-ia\tau}z_4^2 \\ \bar{f}_{r3}(z,\tau) &= i\,\overline{2d_{-1,-2}(s)e^{-iM\omega_1 s}}\;e^{-ia\tau}z_2z_4\,, \end{aligned} \tag{6.35}$$

and the averaged equations become

$$\begin{aligned} v_1' &= i\epsilon\gamma_M e^{-ia\epsilon s}v_3^{*2}\,, \\ v_3' &= i\epsilon 2\gamma_M e^{-ia\epsilon s}v_1^*v_3^*\,, \end{aligned} \tag{6.36}$$

where $\gamma_M = d_{-1,-2,M}$ and may be complex. Letting $v_1 = e^{ia\epsilon s}\zeta_1$ and $v_3 = e^{-ia\epsilon s}\zeta_3$ autonomizes as in Examples B and F, giving

$$\begin{aligned} \zeta_1' &= i\epsilon[-a\zeta_1 + \gamma_M\zeta_3^{*2}] \\ \zeta_3' &= i\epsilon[a\zeta_3 + 2\gamma_M\zeta_1^*\zeta_3^*]\,. \end{aligned} \tag{6.37}$$

Letting $\zeta_1 = \frac{1}{2}(X_1+iY_1)$ and $\zeta_3 = \frac{1}{2}(X_2+iY_2)$, (6.37) is transformed into a two-degree of freedom autonomous Hamiltonian system which can be analyzed using the standard techniques for such systems. Theorem 2 is then applied to obtain the approximation.

In the nonresonant case P is easily constructed as $\int_0^s f(z,\tau)d\tau$ and the averaged problem becomes

$$\begin{aligned} v' &= \epsilon^2 V(v)\,, \\ v(0) &= z_0 = \left(A_x e^{i\phi_x}, cc, A_y e^{i\phi_y}, cc\right)^T \end{aligned} \tag{6.38}$$

where

$$V(z) = \overline{D_1 f(z,s) \mathcal{I} f(z,s)}\,. \tag{6.39}$$

To calculate V_1 and V_3 we proceed as in Example E. For example, V_1 is found from (6.39) by averaging $\sum_{j=1}^{4} \frac{\partial f_1(z,s)}{\partial z_j} \mathcal{I} f_j(z,s)$ using (6.34), the nonresonance condition, the definition of $d_{k\ell}$ and $d_{k\ell m}$ after (6.34) and the fact that $f_2(z,s) = f_1(z,s)^*$ and $f_4(z,s) = f_2(z,s)^*$. Because of the nonresonance condition the only terms that appear in $V_1(z)$ are $z_1^2 z_2$ and $z_1 z_3 z_4$ and in $V_3(z)$ are $z_1 z_2 z_3$ and $z_3^2 z_4$. After considerable computation, we find

$$\begin{aligned} v_1' &= i\epsilon^2\Big(\gamma_x |v_1|^2 + \gamma_c |v_3|^2\Big) v_1\,, \qquad v_1(0) = A_x e^{i\phi_x} \\ v_3' &= i\epsilon^2\Big(\gamma_c |v_1|^2 + \gamma_y |v_3|^2\Big) v_3\,, \qquad v_3(0) = A_y e^{i\phi_y} \end{aligned} \tag{6.40}$$

where $\gamma_x = \gamma$ as given after (6.14b), the coupling constant, γ_c, is given by

$$\gamma_c = \frac{4}{\omega_1} \sum_n \left\{ 2\frac{\mathrm{Re}\,[e_{1,n} d^*_{1,0,n}]}{n+\nu_x} + \frac{|d_{-1,2,-n}|^2}{n+\nu_x-2\nu_y} + \frac{|d_{-1,-2,n}|^2}{n-\nu_x-2\nu_y} \right\}$$

and

$$\gamma_y = -\frac{2}{\omega_1} \sum_n \left\{ 4\frac{|d_{1,0,n}|^2}{n+\nu_x} + \frac{|d_{1,-2,n}|^2}{n+\nu_x-2\nu_y} - \frac{|d_{-1,-2,n}|^2}{n-\nu_x-2\nu_y} \right\}.$$

Since γ_x, γ_c, and γ_y are each real, it is easy to see that $|v_1|$ and $|v_2|$ are conserved in (6.40) and thus

$$\begin{aligned} v_1(s) &= A_x \exp\Big[i\Big(\epsilon^2(\gamma_x A_x^2 + \gamma_c A_y^2)s + \phi_x\Big)\Big]\,, \\ v_3(s) &= A_y \exp\Big[i\Big(\epsilon^2(\gamma_c A_x^2 + \gamma_y A_y^2)s + \phi_y\Big)\Big]\,, \end{aligned} \tag{6.41}$$

and the approximation can be constructed as in (6.4), noting that (6.5) is in the scaled time whereas (6.38) is not.

In the case of "thin" sextupoles distributed around the ring, $S(s)$ can be represented by a sum of delta functions. This case has been treated by Collins, Ng, and Ohnuma (see Reference 16), and we are now in a position to compare results.

Example H: Dipole ripple and sextupole (nonresonance).

In this section, we extend the result in Example E to include dipole ripple. This problem was also discussed in Reference 17, where another standard form of the method of averaging was used. The perturbation term h in (3.13) in this case can be written

$$h(x, \omega_1 s, \omega_3 s) = h_s(x, \omega_1 s) + h_r(\omega_1 s, \omega_3 s)\,, \tag{6.42a}$$

where

$$h_s(x, \omega_1 s) = \frac{1}{2} S(s) x^2 \,, \tag{6.42b}$$

$$h_r(\omega_1 s, \omega_3 s) = A_r(s) \cos\Big(\omega_3 s + \alpha_r(s)\Big) \,, \tag{6.42c}$$

and $\omega_3 = \nu_r \omega_1$, with $\nu_r = f_r / f_0$ and f_0 is the beam frequency and f_r the ripple frequency. Thus, the first component of $f(z,s)$ in (3.16) can be written as

$$f_1 = f_{s1} + f_{r1} \,, \tag{6.43}$$

where $f_{s1}(z,s)$ is given by (6.2) and

$$f_{r1}(z,s) = ir(s)\Big[e^{i\alpha_r(s)} e^{-i(\omega_2 - \omega_3)s} + e^{-i\alpha_r(s)} e^{-i(\omega_2 + \omega_3)s}\Big] \,, \tag{6.44}$$

where

$$r(s) = -\frac{1}{4} \beta^{1/2}(s) e^{-i\psi_p(s)} A_r(s) \,.$$

We assume ν, 3ν, and $\nu_r \pm \nu$ are nonintegers; thus $\bar{f} = 0$ and the second-order averaging Theorems 3 and 4 give (6.4)–(6.7) as before. Making use of (6.43), (6.7) becomes

$$V(z) = \overline{D_1 f_s(z,s) \mathcal{I} f_s(z,s)} + \overline{D_1 f_s(z,s) \mathcal{I} f_r(s)} \,, \tag{6.45}$$

since $D_1 f_r(z,s) = 0$. The first term was computed in Example E, and the second term is zero if $\nu_r \pm \ell\nu$ is noninteger for $\ell = 0, 2, 4$. To see this, notice that f_s contains the terms $e^{\pm i\ell_1 \omega_2 s}$ for $\ell_1 = 1, 3$ and f_r contains the terms $e^{-i(\omega_2 \pm \omega_3)s}$, and products of these must not give $e^{iM\omega_1 s}$ for integer M. Therefore, $V(z)$ is as in Example E, where it is defined by (6.8), (6.14b), and $\gamma = -6\gamma_1 - 2\gamma_3$.

Thus by Theorem 4, the ripple effect is at most $O(\epsilon)$ on $O(1/\epsilon^2)$ s-intervals and does not affect the approximation given by Theorem 4 in (6.4b) and (6.11). However, the effect of the ripple does enter through the P function,

$$P(z,s) = P_s(z,s) + P_r(z,s) = \int_0^s f_s(z,t)dt + \int_0^s f_r(z,t)dt \,. \tag{6.46}$$

Equation (6.11) doesn't change, and

$$\sqrt{\beta(s)}\Big[e^{i\psi(s)} P_{r1}(s) + e^{-i\psi(s)} P_{r2}(s)\Big]$$

must be added to (6.10). This term can be written

$$\beta(s)^{1/2} \int_0^s \beta^{1/2}(t) A_r(t) \sin\Big(\psi(s) - \psi(t)\Big) \cos\Big(\omega_3 t + \alpha_r(t)\Big) dt \,, \tag{6.47}$$

and this completes the approximation.

In the delta function case, where $S(s) = S_0\delta_p(s-s_0)$ and $A_r(s) = A_r\delta_p(s-s_r)$, we obtain a simple expression for γ as discussed in Example E. In addition we obtain

$$\begin{aligned}
(\tilde{P}_s)_1(v,s) &= -\frac{\beta_0^{3/2}S_0}{8\pi}\left\{v_1^2 e^{i\left(\omega_2 s+\psi_p(s_0)\right)}\sum_n \frac{e^{in\omega_1(s-s_0)}}{n+\nu} + 2v_1v_2 e^{-i\left(\omega_2 s+\psi_p(s_0)\right)}\sum_n \frac{e^{in\omega_1(s-s_0)}}{n-\nu}\right. \\
&\quad \left. + v_2^2 e^{-3i\left(\omega_2 s+\psi_p(s_0)\right)}\sum_n \frac{e^{in\omega_1(s-s_0)}}{n-3\nu}\right\} \\
&= -i\frac{\beta_0^{3/2}S_0}{4}\left\{v_1^2 e^{i\left(\omega_2 s+\psi_p(s_0)\right)}\frac{e^{-i\nu\langle\theta-\theta_0\rangle}}{1-e^{-i2\pi\nu}} + 2v_1v_2 e^{-i\left(\omega_2 s+\psi_p(s_0)\right)}\frac{e^{i\nu\langle\theta-\theta_0\rangle}}{1-e^{i2\pi\nu}}\right. \\
&\quad \left. + v_2^2 e^{-3i\left(\omega_2 s+\psi_p(s_0)\right)}\frac{e^{i3\nu\langle\theta-\theta_0\rangle}}{1-e^{i6\pi\nu}}\right\}, \qquad (6.48)
\end{aligned}$$

where $\theta \equiv \omega_1 s = 2\pi s/C$, $\theta_0 \equiv \omega_1 s_0$, and $0 < \langle\theta\rangle < 2\pi$, and

$$\begin{aligned}
(\tilde{P}_r)_1(z,s) &= -\frac{1}{8\pi}\beta_r^{1/2}A_r e^{-i\left(\psi_p(s_r)+\omega_2 s\right)} \times \left\{e^{i\left(\alpha_r+\omega_3 s\right)}\sum_n \frac{e^{in\omega_1(s-s_r)}}{n+\nu_r-\nu}\right. \\
&\quad \left. + e^{-i(\alpha_r+\omega_3 s)}\sum_n \frac{e^{in\omega_1(s-s_r)}}{n-\nu_r-\nu}\right\} \\
&= -\frac{i}{4}\beta_r^{1/2}A_r e^{-i\left(\psi_p(s_r)+\omega_2 s\right)} \\
&\quad \times \left\{\frac{e^{i[(\nu-\nu_r)\langle\theta-\theta_r\rangle+\alpha_r+\nu_r\theta]}}{1-e^{i2\pi(\nu-\nu_r)}} + \frac{e^{i[(\nu+\nu_r)\langle\theta-\theta_r\rangle-\alpha_r-\nu_r\theta]}}{1-e^{i2\pi(\nu+\nu_r)}}\right\}. \qquad (6.49)
\end{aligned}$$

7. Conclusions and Discussions

We have presented first- and second-order averaging theorems in the quasiperiodic case and applied these theorems to several beam dynamics problems. We have indicated how the conditions of the theorems can be satisfied in some of the examples, but a more detailed study of this is necessary. Furthermore, the theorems do not apply in the case of delta function perturbations, and this seems like a good problem for future work. We have selected a variety of problems. In the first-order case, Example A discusses a fairly general situation that includes chromaticity, sextupole, octupole, and beam-beam perturbations. Example B shows how near resonance fits into the averaging framework and the resonant normal form appears quite naturally. Example C is a higher-dimensional problem that illustrates the treatment of transverse coupling in the important beam-beam case. A final first-order example involves longitudinal beam dynamics in the case of rf phase modulation and electron cooling. This illustrates the robustness of the averaging method, as the system is dissipative and thus non-Hamiltonian. Furthermore, we indicate how the averaging theorems can be used to prove the existence of periodic solutions and invariant tori.

In the second-order case, Example E illustrates the method on the sextupole, a standard beam dynamics example; Example F illustrates near resonance at second order, and again the resonant normal form appears quite naturally. Example G is a higher-dimensional example, the sextupole with transverse coupling. We had hoped to compare these results with the work of Ng, Collins, and Ohnuma,[16] but there wasn't time. Finally, in Example H, we discuss the important case of the combined sextupole-dipole ripple calculation that we had investigated previously.[17] Regular perturbation theory was discussed, primarily to make clear the meaning of averaging as a long time perturbation theory.

Another example in the spirit of this paper is the work of Reference 18. The authors study the evolution of an ensemble of forced duffing oscillators as a model for a beam with nonlinearity and dipole ripple. They consider the case when the external frequency is near the linearized natural frequency, and nonlinearity and forcing are small. They argue, based on the method of averaging, that the beam equilibrates in a coarse-grained sense, and they calculate the equilibrated beam characteristics from the averaging approximation. A loose end in this paper, in the context of averaging, is that the averaging results are on $O(1/\epsilon)$ or $O(1/\epsilon^2)$ time intervals, whereas "equilibrium" considerations may require longer times. (Mathematically, of course, equilibrium is an infinite time concept.)

In any perturbation problem there are two important choices: the starting coordinates and the method. We think that in our transverse motion examples, we have chosen the optimal coordinates for calculations, and we are grateful to Kummer[19] for pointing these out to us. Not only do these coordinates considerably simplify the calculation of averages, they also make it possible to determine the approximations by consideration of one-half of the vector field.

As mentioned in the introduction there are several long time perturbation procedures from which to choose. All legitimate ones should lead to the same asymptotic expansions, although Murdock's thoughtful remarks on pages 14–15 of Reference 1 are significant; furthermore, it should be remembered that asymptotic approximations are "valid for small ϵ," but how small is usually a difficult question. Our emphasis in the examples has been to obtain complete approximate solutions along with error estimates. We are in the process of deepening our understanding of the other methods so that we can make detailed comparisons; however, at this point we make the following tentative remarks. The advantages of averaging are that: (1) it is robust in the sense that it can handle any problem the other methods handle as well as other problems. For example, canonical perturbation theory cannot handle dissipative systems, whereas averaging can; (2) it is set up for an easy return to the original variables. Tune shifts are a consequence of the calculation but not the primary focus. If tune shifts are the primary focus, then other methods are most likely better, particularly as the order increases; (3) it is set up for ease in error estimation.

The disadvantages of averaging are: (1) the problem needs to be put into a standard form for the method. Example D used scaling in combination with variation of parameters, whereas the other examples used just variation of parameters; (2) bookkeeping may be cumbersome. At first order all methods are easy. At second order, the easiest method is probably the one of greatest familiarity. At higher order, other methods may be easier. We hope to do a detailed comparison of averaging with other methods in the future. For example, in Hamiltonian problems canonical perturbation methods have the advantage of working on the Hamiltonian (a scalar) rather than the vector field, and this is a considerable simplification. What is not clear to us, however, is the level of difficulty in going back through the transformations to find the approximation in the original variables and the level of difficulty in setting up for the error analysis. Perhaps a good test case would be the 1:5 sextupole resonance, which appears in third-order perturbation theory.

We now indicate some extensions of the type of averaging results presented in this paper in the context of

$$\dot{x} = \epsilon f(x,t,\omega) = \epsilon\Big[\bar{f}(x) + p(x,t) + q(x,t,\omega)\Big]\,. \tag{7.1}$$

Here p will be deterministic with zero t-mean and q will be stochastic with zero stochastic mean and satisfy a so-called mixing condition that specifies the rate at which $q(x,t,\omega)$ and $q(x,s,\omega)$ become independent as $|t-s|$ grows. We use the symbol ω to denote a random function, as is standard in probability and stochastic processes. The unperturbed problem will be $\dot{v} = \epsilon\bar{f}(v)$, and thus p and q will be viewed as perturbations. In a future publication we hope to explore applications of these results in the beam dynamics context, although some discussion of this can be found in the context of longitudinal beam dynamics with rf noise.[20]

We first consider the deterministic case when $q = 0$. All our examples were of first- and second-order averaging, and the first extension is to nth-order averaging. This is discussed in References 7 and 8 and fully reveals averaging as a systematic perturbation expansion yielding an $O(\epsilon^n)$ approximation on $O(1/\epsilon)$ time intervals. In general, it is difficult to obtain approximations on longer intervals although in the case where $\bar{f} = 0$ it is fairly straightforward, as discussed in Section 5. Usual proofs of averaging theorems use the Gronwall inequality, and the error bounds contain a factor of $\exp C(\epsilon)t$. $C(\epsilon)$ is $O(\epsilon)$ unless $\bar{f} = 0$, thus giving error bounds that are transcendentally large (and thus useless) at times larger than $O(1/\epsilon)$. For $\bar{f} = 0$, $C(\epsilon) = O(\epsilon^2)$, and this allowed the extension to $O(1/\epsilon^2)$ time intervals.

If the unperturbed problem has an invariant, then it may be possible to obtain a result on an interval longer than $O(1/\epsilon)$ even though $\bar{f} \neq 0$. More specifically, if $I(x)$ is an invariant or family of invariants for the unperturbed problem, that is, if $I'(x)\bar{f}(x) = 0$, then under suitable conditions $y(t,\epsilon) = I\big(x(t,\epsilon)\big)$ evolves approximately according to an ODE,

$$\dot{J} = \epsilon^2\mu(J) \tag{7.2}$$

on $O(1/\epsilon^2)$. This is a special case of a more general stochastic theorem proved in Reference 21. The following example taken from that reference illustrates this:

$$\begin{aligned}
\dot{x}_1 &= \epsilon x_2 \\
\dot{x}_2 &= -\epsilon U'(x_1) + \epsilon(x_2^2\cos\lambda t + \lambda\alpha\sin\lambda t)\,,
\end{aligned} \tag{7.3}$$

where $U(x_1)$ is a bowl potential and α and λ are constants. Let $I(x)$ be the action of the unperturbed problem; then $\mu(J) = -\alpha J$ and the theorem of Reference 21 gives

$$I\big(x(t,\epsilon)\big) = I_0 \exp(-\alpha\epsilon^2 t) + o(1) \tag{7.4}$$

for $0 \le t \le T/\epsilon^2$, where $o(1)$ denotes a function that goes to zero as ϵ goes to zero.

In the stochastic case of (7.1), Khas'minskii[22] proved that under suitable conditions

$$x(t,\epsilon) \cong u(\epsilon t) + \sqrt{\epsilon}Y_0(\epsilon t,\omega)\,, \tag{7.5}$$

where $u(\tau)$ is defined by $u' = \bar{f}(u)$, $u(0) = x_0$, $Y_0(\tau,\omega)$ is a Gauss-Markov process defined by the Itô stochastic differential equation

$$dY_0 = D\bar{f}\big(u(\tau)\big)Y_0 d\tau + \sigma\big(u(\tau)\big)dW\,, \qquad Y_0(0) = 0\,, \tag{7.6}$$

where $W = W(\tau)$ is standard Brownian motion and σ is determined from the stochastic perturbation q. The approximation in (7.5) is in the sense of weak convergence;[23] that is, if $Y(\tau,\omega) := \big(x(\tau/\epsilon,\epsilon) - u(\tau)\big)/\sqrt{\epsilon}$ then Y converges weakly to Y_0 as $\epsilon \to 0$ for $0 \le \tau \le T$. Note that the scaling makes this an approximation on $O(1/\epsilon)$ t-intervals. If $\bar{f} = 0$ then Khas'minskii[24] also proved that

$$x(t,\epsilon) \cong X_0(\epsilon^2 t,\omega)\,, \tag{7.7}$$

where $X_0(\tau,\omega)$ is a Markov process defined by the Itô stochastic differential equation

$$dX_0 = b(X_0)d\tau + \sigma(X_0)dW\,, \tag{7.8}$$

and the approximation in (7.7) means that $x(\tau/\epsilon^2,\epsilon)$ converges weakly as $\epsilon \to 0$ to $X_0(\tau,w)$ for $0 \le \tau \le T$. The functions b and σ are defined in Reference 24. Note that this scaling makes this an approximation on $O(1/\epsilon^2)$ t-intervals. These theorems are also discussed in the book by Freidlin and Wentzel.[25]

Cogburn and Ellison[21] extended the latter result to the case where $\bar{f} \neq 0$ by showing that if $I(x)$ is a suitable vector of invariants of the unperturbed problem, then under suitable regularity and ergodicity conditions

$$Z(\tau,\omega,\epsilon) := I\big(x(\tau/\epsilon^2,\epsilon)\big) \tag{7.9}$$

converges weakly to a Markov diffusion process for $0 \le \tau \le T$. As before, this scaling gives an approximation on $O(1/\epsilon^2)$ t-intervals. In Reference 26, a related phase randomization result on $O(\epsilon^{-4/3})$ t-intervals is discussed.

To make the stochastic results more concrete, consider

$$\begin{aligned} \dot{x}_1 &= \epsilon x_2 \\ \dot{x}_2 &= \epsilon\big[-U'(x_1) + P(x_1,t) + Q(x_1,t,\omega)\big]\,, \end{aligned} \tag{7.10}$$

where U is a symmetric bowl type potential so that all solutions of the unperturbed problem are periodic. Then Khas'minskii's result[22] is

$$x(t,\epsilon) \cong u(\epsilon t) + \sqrt{\epsilon}Y_0(\epsilon t,\omega), \quad 0 \le \epsilon t \le T\,, \tag{7.11}$$

where $u(\tau)$ is defined by $u_1' = u_2$, $u_2' = -U'(u_1)$, and $Y_0(\tau, \omega)$ by

$$dY_0 = \begin{pmatrix} 0 & 1 \\ -U''(u_1(\tau)) & 0 \end{pmatrix} Y_0 d\tau + \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{C(u_1(\tau))} \end{pmatrix} dW \,, \quad Y_0(0) = 0 \,, \tag{7.12}$$

where $C(x_1) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \int_0^T E\Big(Q(x_1,t)Q(x_1,s)\Big) dt\, ds$. Thus on $O(\epsilon^{-1})$ t-intervals x follows the deterministic and periodic unperturbed motion with an $O(\sqrt{\epsilon})$ Gauss-Markov correction. While (7.11) is probably not valid on longer intervals, Reference 26 obtains a result on $O(\epsilon^{-4/3})$ t-intervals. Basically it says that under a nonlinearity assumption the process becomes uniform on thin energy shells at $O(\epsilon^{-4/3})$ times; that is, in action-angle variables the angle has become randomized. Let $J(h)$ be the unperturbed action as a function of energy, $h(x) = \frac{1}{2}x_2^2 + U(x_1)$, and θ the angle canonically conjugate to J. Then it is shown for $0 \le \epsilon^{4/3} t \le T$ that

1. the action $J\Big(h\big(x(t,\epsilon)\big)\Big)$ behaves like Brownian motion and changes in action are $o(1)$ as $\epsilon \to 0$, and
2. the angle behaves like

$$\theta(t,\epsilon) \cong \theta_0 + \epsilon t \Omega(J_0) + \Theta(\epsilon^{4/3} t, \omega) \,,$$

where $\Theta(\tau, \omega)$ is Gauss-Markov with zero mean and covariance $E\Big(\Theta(\tau_1), \Theta(\tau_2)\Big) = \frac{1}{2}\tau_1^2(3\tau_2 - \tau_1)\Omega'(J_0)^2\, \sigma^2(J_0)$, J_0 is the initial action, $\Omega(J)$ is the frequency of the unperturbed oscillator, and

$$\sigma^2(J) = 2\pi\Omega(J)^{-1} 4\sqrt{2} \int_0^{a(J)} \sqrt{U\big(a(J)\big) - U(x_1)}\, C(x_1)\, dx_1 \,.$$

If $\Omega'(J_0) \neq 0$, then the second result gives uniformity on thin energy shells. This phase randomization can also occur in a coarse-grained sense without the stochastic perturbation, as discussed in References 18 and 27. Because θ is now approximately uniform, it is possible for the action to be approximately Markovian. In fact, the theorem in Reference 21 gives that the changes in action remain small until $O(\epsilon^{-2})$ times, and at these times changes can be $O(1)$, and the action behaves like a Markov diffusion process. More specifically, $J\Big(h\big(x(\tau/\epsilon^2, \epsilon)\big)\Big)$ converges weakly on $0 \le \tau \le T$ to a Markov diffusion process $Z_0(\tau, \omega)$ defined by

$$dZ_0 = \mu(Z_0) d\tau + \sigma(Z_0) dW \,,$$

where σ^2 is defined above and $\mu(y) = \frac{1}{2}\frac{d}{dy}\sigma^2(y)$.

8. Acknowledgements

Thanks to Yiton Yan for inviting us to speak about the method of averaging in the SSCL Accelerator Physics Seminar Series; it was the genesis of this paper. Thanks also to Mike Syphers for his support of this work and Sylvia Lee for her expert and tireless preparation of this manuscript. One of us (JAE) acknowledges many enjoyable discussions on resonance and perturbation methods with Martin Kummer, as well as the general enthusiasm of Leo Michelotti for nonlinear dynamics, beam dynamics, and modern perturbation theories. Thanks also to Brenda Ramsey, Liana Baritchi, and Cameron Owen of SSCL and Bill Gordon, Alex Stone, and Ellen Goldberg of UNM for helping this work to go forward at a critical time. Finally, this work is dedicated to the dreams of what might have been—and to those who came to Dallas, full of enthusiasm, and who devoted a significant part of their lives to the manifestation of those dreams.

9. References

1. J.A. Murdock, *Perturbations: Theory and Methods*, Wiley, New York, 1991.
2. L. Michelotti, *Intermediate Classical Mechanics with Applications to Particle Accelerator Physics*, Wiley, New York, 1994.
3. R.D. Ruth, "Single Particle Dynamics in Circular Accelerators," in *Physics of Particle Accelerators*, edited by M. Month and M. Dienes, AIP Conference Proceedings 153, 1987.
4. N.N. Bogoliubov and Y.A. Mitropolski, *Asymptotic Methods in the Theory of Nonlinear Oscillations*, Hindustan Pub. Corp., Delhi, 1961.
5. J.A. Sanders and F. Verhulst, *Averaging Methods in Nonlinear Dynamical Systems*, Springer-Verlag, New York, 1985.
6. P. Lochak and C. Meunier, *Multiphase Averaging for Classical Systems*, Springer-Verlag, New York, 1988.
7. J.A. Ellison, A.W. Sáenz, and H.S. Dumas, "Improved nth order averaging theory for periodic systems," *J. Diff. Eq.*, **84** (1990), p. 383.
8. A.W. Saénz, "Higher-order averaging for nonperiodic systems," *J. Math. Phys.* **32**, 2679 (1991).
9. D. Neuffer and S. Peggs, "Beam-Beam Tune Shifts and Spreads in the SSC," SSC-63, April 1986.
10. G. López, "Head-On and Long Range Beam-Beam Tune Shifts Spread in the SSC," SSCL-442, May 1991.
11. J.A. Ellison, H.-J. Shih, and M. Kummer, "Theoretical study of longitudinal beam splitting and related phenomena," *Phys. Rev. E* **49**, 2484 (1994).
12. J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields*, Springer-Verlag, New York, 1986.
13. J.K. Hale, *Ordinary Differential Equations*, Wiley-Interscience, New York, 1969.
14. J.A. Ellison, M. Kummer, and A.W. Sáenz, "Transcendentally small transversality in the rapidly forced pendulum," *J. Dyn. Differential Eq.* **5**, 241 (1993).
15. H.S. Dumas, J.A. Ellison, and A.W. Sáenz, "Axial channeling in perfect crystals, the continuum model and the method of averaging," *Annals of Physics*, **209** (1991), pp. 97–123.
16. K.Y. Ng, "Simple Derivation of Distortion Functions," Fermilab-FN-586, March 1992; T. Collins, *Proceedings of the 1984 Summer Study on the Design and Utilization of the*

Superconducting Super Collider, Snowmass, Colorado, 1984; S. Ohnuma, *Proceedings of the Conference on the Interactions Between Particle and Nuclear Physics*, Steamboat Springs, Colorado, 1984.

17. H.-J. Shih, J.A. Ellison, M. Syphers and B.S. Newberger, "Emittance growth due to dipole ripple and sextupole," *Proceedings of the 1993 IEEE Particle Accelerator Physics Conference*, DC, 3588–3590.
18. T. Sen, J.A. Ellison, S.K. Kauffmann, "Collective behavior of an ensemble of forced duffing oscillators near the 1:1 resonance," SSCL-Preprint-561, 1994, submitted for publication.
19. M. Kummer, private communication.
20. H.-J. Shih, J.A. Ellison, B.S. Newberger, and R. Cogburn, "Longitudinal Beam Dynamics with RF Noise," *Part. Accel.* **43**, 159 (1994).
21. R. Cogburn and J.A. Ellison, "A stochastic theory of adiabatic invariance," *Commun. Math. Phys.* **149**, 97 (1992).
22. R.Z. Khas'minskii, "On processes defined by differential equations with a small parameter," *Theory Prob. Appl.*, **11** (1966), pp. 211–228.
23. S.N. Ethier and T.G. Kurtz, *Markov Processes Characterization and Convergence*, Wiley-Interscience, New York, 1986.
24. R.Z. Khas'minskii, "A limit theorem for solutions of differential equations with random right-hand side," *Theory Prob. Appl.*, **11** (1966), pp. 390–406.
25. M.I. Freidlin and A.D. Wentzell, *Random Perturbations of Dynamical Systems*, Springer-Verlag, New York, 1984.
26. R. Cogburn and J.A. Ellison, "A Four-Thirds Law for Phase Randomization of Stochastically Perturbed Oscillators and Related Phenomena," to be published in *Commun. Math. Phys.*
27. J.A. Ellison and T. Guinn, "Statistical Equilibrium, Planar Channeling and the Continuum Model," *Phys. Rev. B* **13**, 1880 (1976).

10. Appendix—Gronwall Inequality

If $g(t)$ is non-negative, and continuous and satisfies

$$g(t) \leq a + b(t - t_0) + L \int_{t_0}^{t} g(s)ds \tag{A1}$$

for a, b non-negative, and L positive, then

$$g(t) \leq \left(a + \frac{b}{L}\right) e^{L(t-t_0)} - \frac{b}{L} \leq \left[a + b(t - t_0)\right] e^{L(t-t_0)}, \tag{A2}$$

for $t \geq t_0$. The second inequality in (A2) follows from the rather crude estimate $1 - e^{-L(t-t_0)} \leq L(t - t_0)$. To obtain the first inequality let $R(t)$ denote the rhs of (A1). Then

$$R'(t) = b + Lg(t) \leq b + LR,$$

which is equivalent to

$$\frac{d}{dt}\left(e^{-L(t-t_0)}R(t)\right) \leq be^{-L(t-t_0)} .$$

Integrating over $[t_0, t]$ gives

$$e^{-L(t-t_0)}R(t) - a \leq \frac{b}{L}\left(1 - e^{-L(t-t_0)}\right) ,$$

where we have used $R(t_0) = a$. Since $g(t) \leq R(t)$, (A2) follows.