THE PENDULUM PERIOD EXPANSION VIA CANONICAL PERTURBATION THEORY

TERRY R. MCCONNELL
SYRACUSE UNIVERSITY

Abstract. This expository paper obtains the expansion of the period of a plane pendulum for small amplitudes to second order in the square of the amplitude, using both time independent and time dependent perturbation techniques. An appendix outlines the steps needed to supply full mathematical rigor.

1. Introduction

The pendulum, with its metronomic beat, has measured out countless lives. Used for centuries to regulate timepieces, it has come to symbolize the steady and inexorable passage of time. Even early clockmakers, though, recognized that the pendulum on its own is an imperfect timekeeper, its vibrations speeding up slightly as their amplitude decreases. Some experimenters, notably Christiaan Huygens, used carefully designed collars to compensate for this effect. (See [7], especially chapter 10, for a survey of clockmaking, including the role of the plane pendulum and its variations.)

Virtually every mechanics text includes the plane pendulum as an example of a familiar system that can be analyzed, at least approximately, using Newtonian mechanics. (See, e.g., [11], [6] or [10].) Consider an idealized pendulum that has a massless string or arm of length l and a bob that is a point mass of size m. Assume the arm pivots around a fixed frictionless axis that is perpendicular to the plane of motion and let θ denote the angle at the pivot in that plane, measured counterclockwise from the vertical to the arm. Assume the weight mg of the bob is the only[1] external force acting. Following [11], if one denotes by T the magnitude of the tension in the string, then resolving the tension into horizontal and vertical components and applying Newton's third law yields the equations of motion

$$m\ddot{x} = -T\sin\theta \text{ and } m\ddot{y} = T\cos\theta - mg.$$

(Here the y axis is vertical in the plane of motion and the origin is placed at the pivot, so that $x = l\sin\theta$.) If we assume the amplitude (maximum value of θ) is very small, so that θ never varies much from zero, then T has approximately the constant value mg. Thus the x equation becomes

$$\ddot{x} + \omega_0^2 x = 0,$$

[1] This assumption is violated by clock pendulums, which are subject to a periodic impulse from the *escapement,* a ratchet that serves to transmit energy from a spring or falling weight to the pendulum. It is adjusted so as to compensate for friction.

where $\omega_0 = \sqrt{\frac{g}{l}}$. The general solution is, of course, $x(t) = A\cos\omega_0 t + B\sin\omega_0 t$, where A and B can be determined from initial conditions. Thus, to this approximation, the motion is simple harmonic at frequency $\nu_0 = \frac{\omega_0}{2\pi}$.

In addition to seeking an exact solution when the amplitude is not small, a more accurate model would treat the pendulum as an extended rigid body and would account for the effects of friction; still more accurate models might consider elastic and thermal properties of the materials, etc. (See [8] for a detailed discussion of how to account for some of these effects in the laboratory.) Even the idealized problem cannot be solved in closed form, i.e., in terms of elementary functions, though it is possible to express the exact solution in terms of elliptic integrals. See, e.g., [12]. It is easy, however, to derive some general features of the motion, such as that it is strictly periodic in time with a frequency ν that depends on the total energy (equivalently amplitude,) and approaches ν_0 in the low energy limit. This is most easily seen in the *Hamiltonian* formulation of the problem. (See, e.g., [4] for a general introduction to Hamiltonian mechanics.)

Recall that the Hamiltonian, H, is the total energy expressed as a function of the angle coordinate θ and the angular momentum p. In this case, it reduces to

$$H(p,\theta) \;=\; \frac{p^2}{2ml^2} + mgl(1-\cos\theta).$$

The two terms here correspond to kinetic and potential energy, respectively. The time evolution of the system is governed by Hamilton's equations:

$$\dot\theta = \frac{\partial H}{\partial p}(p,\theta), \text{ and } \dot p = -\frac{\partial H}{\partial \theta}(p,\theta).$$

Here dot denotes differentiation with respect to time and the initial values $\theta(0) = \theta_0$ and $p(0) = p_0$ are given by the angle (θ_0) and angular momentum (p_0) at time zero.

Assume, for example, that $p_0 = 0$ and $0 < \theta_0 < \pi$, so that the pendulum is instantaneously at rest at time zero. The energy is then $E = mgl(1-\cos\theta_0)$. Since energy is conserved (an easy consequence of Hamilton's equations,) the pendulum subsequently can be at rest only when $\theta = \pm\theta_0$. Again, the equations of motion imply that the motion of the pendulum for small positive times can be described qualitatively as follows: Both p and θ decrease monotonically, with p reaching a negative minimum when $\theta = 0$. Thereafter, θ continues to decrease, while p increases towards zero.

The conclusion[2] is that p reaches zero again at some finite time, which we may designate as $\frac{1}{2\nu}$. By symmetry, the motion during $[\frac{1}{2\nu}, \frac{1}{\nu}]$ is the mirror image, with the angle and momentum returning to their original values at time $\frac{1}{\nu}$. Since the governing equations are first order in time, standard uniqueness results from the theory of ordinary differential equations now ensure that this initial cycle will repeat forever with period $\tau = \frac{1}{\nu}$.

Level curves of H in the two-dimensional (p,θ) space (known as *phase space*) provide pictures of the possible orbits of a pendulum system. Unbounded orbits occur when the pendulum has enough energy to swing over the top ($E > 2mgl$.) Other orbits are ovoid loops around the origin, with the point representing the values of p and θ at time t cycling around them. It turns out to be useful and elegant to reparametrize the level curves so that the system point coordinate moves at uniform

---

[2] To be rigorous, one must rule out the (physically unexpected) possibility that p approaches zero asymptotically. This, and other purely mathematical issues, are discussed in the appendix.

speed. This variable, named the *angle variable,* and its companion momentum, the *action variable,* form the basis of an elegant formulation of mechanics. We review this theory in section 2 along with some of the main ideas of the Hamilton-Jacobi formulation on which it is based.

The simple harmonic oscillator, together with its generalizations to systems with more degrees of freedom, plays a fundamental role in physics. For example, all reasonable potentials V whose bowl shapes allow for periodic bound motion can be viewed as perturbations of the harmonic oscillator potentials. (This is seen from the Taylor expansion of V.)

Inasmuch as the plane pendulum is a familiar, yet nontrivial, perturbation of the harmonic oscillator, it is often used in examples illustrating the methods of perturbation theory. (See, e.g., [4], [6], or [10].) All of these texts show that to first order in the amplitude θ_0 squared,

$$\nu = \nu_0 - \frac{1}{16}\nu_0\theta_0^2,$$

or give an equivalent result about the period. Thus, pendulums run a bit slower than expected as the amplitude increases.

Goldstein's classic text [4], in particular, assigns the working out of the next higher order correction as an exercise (chapter 11, exercise 5.) This paper merely represents the author's own solution of this exercise, fleshed out with enough background material to make it reasonably self contained.

We were not able to locate in the literature the details of the second order calculation via standard perturbation techniques. (However, see [3], where the calculation is done using a method due to Kryloff and Bogoliuboff.) Moreover, there is at least one subtle trap in the second order calculation that is not present in the first. For these reasons, it seemed desirable to record a detailed example of such a calculation.

In section 3 of the paper we present the second order calculation using a time independent method that Goldstein attributes to von Zeipel. In section 4 we redo the calculation using time-dependent techniques. Finally, the appendix collects for easy reference some of the mathematical results that would be needed to make these calculations rigorous.

2. The Pendulum Problem in Action-Angle Variables

Perturbation theory is applicable when the Hamiltonian H for a system can be written $H = H_0 + \epsilon\Delta H$, where ϵ is a small parameter and the system with Hamiltonian H_0 is integrable in closed form. For the plane pendulum it is most natural to describe the configuration using the angle θ between the pendulum arm and the vertical. Recall that the arm itself is massless of length l, and that the bob is a point mass of size m. The conjugate momentum is then the angular momentum of the bob, $p = ml^2\dot{\theta}$, and the Hamiltonian is

$$H = \frac{p^2}{2ml^2} + mgl(1 - \cos\theta) = \frac{p^2}{2I} + \omega_0^2 I(1 - \cos\theta).$$

The constants $I = ml^2$ and $\omega_0 = \sqrt{\frac{g}{l}}$ are introduced both to simplify the appearance of H and because of their physical significance: I is the moment of inertia of the pendulum bob about the axis of rotation and ω_0 is the limiting value of the angular frequency (radians/sec) as the amplitude approaches zero.

By the Maclaurin expansion of the cosine function we may write

$$H = H_0 \ + \ \theta_0^2 H_1 \ + \ \theta_0^4 H_2 \ + \ o(\theta_0^4), \tag{1}$$

where

$$H_0 = \frac{(p^2 + \omega_0^2 I^2 \theta^2)}{2I}, \tag{2}$$

$$H_1 = -\frac{\omega_0^2 I \theta^4}{24\theta_0^2},$$

and

$$H_2 = \frac{\omega_0^2 I \theta^6}{720\,\theta_0^4}.$$

Following [4], we shall take θ_0 to be the amplitude the harmonic oscillator with Hamiltonian H_0 would have at whatever value of energy is given. We also use the mathematician's "small o" notation, according to which the symbol $o(h)$ represents a quantity such that $\frac{o(h)}{h}$ tends to zero as h tends to zero. For the purpose of perturbation analysis it is convenient to replace θ_0^2 with a dimensionless parameter ϵ that can be varied continuously from 0. At the end we shall set $\epsilon = \theta_0^2$ to obtain results that are physically meaningful for pendulum motion. Accordingly, we shall study the parametrized family of Hamiltonians given by

$$H = H_0 + \epsilon H_1 + \epsilon^2 H_2 + \dots.$$

Let us briefly recall some of the ideas of Hamilton-Jacobi mechanics. We seek to change variables from (p, θ) to a new pair of canonical variables (J, w) such that H depends only on J. Canonical means that the form of Hamilton's equations of motion is unchanged under the change of variables. It then follows from the fact that H is independent of w that w is a linear function of time, $w = \nu t + \beta$, where

$$\nu = \frac{\partial H}{\partial J}. \tag{3}$$

One way to generate canonical changes of variable is via a function $Y = Y(J, \theta)$ of the old spatial coordinate and new momentum known as the generating function for the transformation. If for such a Y we set $\frac{\partial Y}{\partial \theta} = p$, and $\frac{\partial Y}{\partial J} = w$, then provided only the solution of these equations for one pair of variables or the other is a *bona fide* change of variables, it is automatically canonical. This can be seen in various ways. Here is a sketch of one approach using calculus of variations. Consider the function

$$\Lambda(p, \theta, \dot{p}, \dot{\theta}) = p\dot{\theta} - H(p, \theta).$$

This is to be viewed as a function defined on the tangent bundle of the phase space, i.e., $\dot{p}, \dot{\theta}$ denote formal variables here. They are not to be found from the ultimate time dependence of p and θ. The actual system path γ in phase space is a fixed endpoint extremum of $\int \Lambda(\gamma, \dot{\gamma})\, dt$ (with obvious abuse of notation,) since the Euler-Lagrange equations for this variational problem are precisely Hamilton's equations. Now if Λ is replaced by $\Lambda - \frac{\partial Y}{\partial \theta}\dot{\theta} - \frac{\partial Y}{\partial J}\dot{J}$, the same path is still an extremum since the two integrands differ by an exact differential. Since $\frac{\partial Y}{\partial J} = w$ and $\frac{\partial Y}{\partial \theta} = p$, we have

$$\Lambda - \frac{\partial Y}{\partial \theta}\dot{\theta} - \frac{\partial Y}{\partial J}\dot{J} = -\dot{J}w - K(J, w),$$

where $K(J,w) = H(p(J,w),\theta(J,w))$ is the Hamiltonian in terms of the new variables. Finally, one observes that the Euler-Lagrange equations for the functional $-\dot{J}w - K(J,w)$ also reduce to Hamilton's equations in the new variables J and w.

One can also consider generating functions, hence changes of variable, that depend on time. (See Appendix.)

The pair (J,w), called *action-angle* variables, are particularly convenient for perturbation theory. There are several approaches to finding expressions for θ and p in terms of w and J. One general method is to solve the nonlinear partial differential equation

$$H(\frac{\partial Y}{\partial \theta},\theta) = E \tag{4}$$

for the generating function $Y = Y(J,\theta)$. In equation (4), known as the *Hamilton-Jacobi* equation, E is the total energy. It can be expressed in terms of the parameter θ_0 by noting that $p = 0$ at maximum amplitude. Thus,

$$E \ = \ H_0(0,\theta_0) \ = \ \frac{\omega_0^2 I}{2}\theta_0^2 \ = \ 2\pi^2\nu_0^2 I\theta_0^2. \tag{5}$$

Here ν_0 represents the harmonic oscillator ($\epsilon = 0$) angular frequency in cycles per second.

To understand why a solution of the Hamilton-Jacobi equation generates a transformation to coordinates in which H depends only on the canonical momentum J, one must examine more carefully the nature of the solution. First note that the quantity E in (4) serves as a parameter that determines a particular level curve of the Hamiltonian. Since H does not explicitly depend on J, one may also treat J as a parameter. There results a parametric family of functions of J and θ, $Y(J,\theta;E)$, which satisfy (4). In perturbation theory, one encounters transformations which, for small ϵ, differ very little from the identity. The identity transformation is generated by θJ (plus an arbitrary constant), which allows us to assume that $\frac{\partial^2 Y}{\partial J \partial \theta}$ is nonzero, since it equals one for θJ. (See the appendix for further discussion.) Now define the coordinate w as $w = \frac{\partial Y}{\partial J}(J,\theta;E)$. This equation, together with $p = \frac{\partial Y}{\partial \theta}(J,\theta;E)$ and (4) itself, can be used (in principle) to eliminate E and express the pairs (p,θ) and (J,w) in terms of each other. As we noted above, assuming this can all be done successfully, the resulting pair (J,w) is necessarily canonical.

Let K denote the Hamiltonian expressed in terms of J and w[3]: $K(J,w) = H(p(J,w),\theta(J,w))$. Since $p(J,w) = \frac{\partial Y}{\partial \theta}(J,\theta(J,w))$, we have

$$K(J,w) \ = \ H\left(\frac{\partial Y}{\partial \theta}(J,\theta(J,w)),\theta(J,w)\right). \tag{6}$$

If we fix J and allow θ to vary, then by (4) the values of p assumed by $\frac{\partial Y}{\partial \theta}$ are such as to keep (p,θ) on some level curve of H. Noting that in (6) K depends on w only through θ, it follows that K is, in fact, independent of w.

Hamilton's equation $\dot{J} = \frac{\partial K}{\partial w} = 0$ yields that J is dynamically constant, the actual value being determined by E (and, ultimately, the initial conditions.) Another expression for $J = J(E)$ can be found from the general formula ([4], 460-461,)

$$J \ = \ \oint p\, d\theta, \tag{7}$$

[3] That the appropriate Hamiltonian in the new variables may be found this way is due to the fact that the canonical transformation is independent of time.

where the integration is over a full cycle of θ, from θ_0 to $-\theta_0$ and back again. In the case $\epsilon = 0$ it is more convenient to take advantage of the freshman physics solution of the harmonic oscillator equation of motion. To distinguish quantities associated with Hamiltonian H_0 from those associated with H, we shall henceforth attach to the former a subscript zero.

From (3) and the fact that frequency is independent of amplitude, we obtain

$$H_0(J_0) = \nu_0 J_0. \tag{8}$$

(The additive constant of integration is not physically meaningful. We take it to be zero.) Since w_0 is a linear function of time and θ is a simple periodic function of time, we must clearly have

$$\theta(J_0, w_0) \;=\; A(J_0)\cos(2\pi w_0 + \phi),$$

for some suitable phase angle ϕ and amplitude A. It follows that $p(J_0, w_0)$ is equal to $-2\pi\nu_0 I A \sin(2\pi w_0 + \phi)$. By (2),(8), and (5) we have $\nu_0 J_0 = 2\pi^2\nu_0^2 A^2 I$. Hence, $A = \frac{1}{\pi}\sqrt{\frac{J_0}{2\nu_0 I}}$, or

$$\theta = \frac{1}{\pi}\sqrt{\frac{J_0}{2\nu_0 I}}\cos(2\pi w_0 + \phi). \tag{9}$$

(We shall not need the transformation equation for p.) For later reference, here is the coefficient H_n, of ϵ^n in the expansion of H expressed in terms of J_0 and w_0:

$$H_n(J_0, w_0) \;=\; \frac{(-1)^n 2^{1-n} J_0^{n+1}\cos^{2n+2}(2\pi w_0 + \phi)}{\pi^{2n}(2n+2)!\nu_0^{n-1} I^n \theta_0^{2n}}. \tag{10}$$

3. Period Expansion via Time Independent Perturbation

Let J, w (without subscripts) denote the action-angle pair for the true Hamiltonian H. Then H is constant on the curves of constant J so there is a function α of J only such that $\alpha(J) = H(J, w)$, for all w. The true frequency of vibration, ν, is given by $\nu = \frac{d\alpha}{dJ}$, since $w = \nu t + w(0)$ satisfies Hamilton's equation $\dot{w} = \frac{d\alpha}{dJ}$. Of course, ν is a function of J, and J is determined by E, or equivalently θ_0.

Assuming α is analytic[4] in the perturbation parameter ϵ, ν can be expanded in a power series

$$\nu \;=\; \nu_0(1 + \alpha_1\epsilon + \alpha_2\epsilon^2 + \dots)$$

with coefficients $\alpha_i = \alpha_i(J)$. Evaluating at $\epsilon = \theta_0^2$ gives

$$\nu \;=\; \nu_0 \;+\; \nu_0\alpha_1\theta_0^2 \;+\; \nu_0\alpha_2\theta_0^4 \;+\; \dots,$$

an expansion that converges when the amplitude is sufficiently small. (We assume from now on without explicit comment that θ_0 is taken sufficiently small.)

The purpose of this section is to derive the explicit numerical values of α_1 and α_2. (In principle the same methods can be used to find higher order terms in the expansion of ν.) These turn out to be $\alpha_1 = -\frac{1}{16}$ and $\alpha_2 = -\frac{5}{1024}$. Following [4], let $Y = Y(\epsilon, J, w_0)$ be a generating function for the canonical transformation $(J_0, w_0) \to (J, w)$. Assuming Y is analytic,

$$Y \;=\; w_0 J + \epsilon Y_1(J, w_0) + \epsilon^2 Y_2(J, w_0) + \dots.$$

[4]See the appendix for sketches of proofs of all the analyticity assumptions.

(Recall that $w_0 J$ generates the identity transformation.) Here Y_1 and Y_2 are certain analytic functions. The Hamilton-Jacobi equation is

$$H\left(\frac{\partial Y}{\partial w_0}, w_0\right) = \alpha(J).$$

Since the right side is independent of w, and H is analytic,

$$\alpha(J) = \frac{1}{2\pi}\int_0^{2\pi} H(J + \epsilon\frac{\partial Y_1}{\partial w_0} + \ldots, w_0)\, dw_0 = \sum_{j=0}^{\infty} \epsilon^j \frac{1}{2\pi}\int_0^{2\pi} H_j(J + \epsilon\frac{\partial Y_1}{\partial w_0} + \ldots, w_0)\, dw_0. \tag{11}$$

We must expand the integrands in powers of ϵ and collect coefficients of ϵ and ϵ^2 to identify α_1 and α_2. Using angled brackets to denote averaging the expression contained inside on w_0, and displaying only terms relevant to the coefficients of interest,

$$\alpha(J) = \langle H(J, w_0)\rangle + \left\langle \frac{\partial H}{\partial J}(J, w_0)\left(\epsilon\frac{\partial Y_1}{\partial w_0} + \epsilon^2\frac{\partial Y_2}{\partial w_0} + \ldots\right)\right\rangle + \frac{1}{2}\left\langle \frac{\partial^2 H}{\partial J^2}\left(\epsilon\frac{\partial Y_1}{\partial w_0} + \ldots\right)^2\right\rangle + \ldots. \tag{12}$$

Recall that $H = \nu_0 J + \epsilon H_1 + \epsilon^2 H_2 + \ldots$. Since the leading term here is the only one that contributes to the last term in (12) at order ϵ^2, the latter term may be dropped. Differentiating both sides of (12) with respect to J we have

$$\nu(J) = \left\langle \frac{\partial H}{\partial J}\right\rangle + \left\langle \frac{\partial^2 H}{\partial J^2}\left(\epsilon\frac{\partial Y_1}{\partial w_0} + \epsilon^2\frac{\partial Y_2}{\partial w_0} + \ldots\right)\right\rangle + \left\langle \frac{\partial H}{\partial J}\left(\epsilon\frac{\partial^2 Y_1}{\partial J\partial w_0} + \epsilon^2\frac{\partial^2 Y_2}{\partial w_0\partial J} + \ldots\right)\right\rangle + \ldots. \tag{13}$$

Using the expansion of H we have

$$\nu = \nu_0 + \epsilon\left\langle \frac{\partial H_1}{\partial J}(J, w_0)\right\rangle + \epsilon^2\left\langle \frac{\partial H_2}{\partial J}(J_0, w_0)\right\rangle + \epsilon^2\left\langle \frac{\partial^2 H_1}{\partial J^2}(J_0, w_0)\frac{\partial Y_1}{\partial w_0}(J_0, w_0)\right\rangle + \epsilon^2\left\langle \frac{\partial H_1}{\partial J}(J_0, w_0)\frac{\partial^2 Y_1}{\partial w_0\partial J}(J_0, w_0)\right\rangle + o(\epsilon^2), \tag{14}$$

where terms on each line of the display come from the same term of (13).

Note that the second term is evaluated at J, while the subsequent terms are all evaluated at J_0, the value J has when $\epsilon = 0$. This difference turns out to be quite important. In the second term of (13) we used that H_0 is linear in J, and in the third that $\frac{\partial Y_1}{\partial w_0}$, and indeed $\frac{\partial Y_j}{\partial w_0}$ for $j \geq 1$, have mean zero. This is shown in [4], pp 517-518, along with the derivation of (14), but we also show this for Y_1 below.

For the (first order) calculation of α_1 we may evaluate $\frac{\partial H_1}{\partial J}$ in (14) at $J = J_0$. Using the first of the following three trigonometric averages

$$\langle\cos^4\rangle = \frac{3}{8},\ \langle\cos^6\rangle = \frac{5}{16},\ \text{and}\ \langle\cos^8\rangle = \frac{35}{128} \tag{15}$$

we find that

$$\left\langle \frac{\partial H_1}{\partial J}(J_0, w_0) \right\rangle \;=\; -\frac{3J_0}{96\pi^2 I\theta_0^2}.$$

But $J_0 = 2\pi^2 I\nu_0\theta_0^2$ (see (5) and (8),) so the above reduces to $-\frac{\nu_0}{16}$. Thus,

$$\alpha_1 = -\frac{1}{16}. \tag{16}$$

The fact that Y_1 has mean zero is a consequence of a general property of the angle variables w_0 and w: if J_0 is held fixed and w_0 increases by 1, then w also increases by 1; vice versa, if J is held fixed and w increases by 1, then w_0 increases by 1. This follows from (7) with p and θ replaced by J_0 and w_0. At constant J_0, let Δw denote the change in w when w_0 increases from some value a to $a+1$. Then

$$\Delta w = \int_a^{a+1} \frac{\partial w}{\partial w_0}\,dw_0 \;=\; \int_a^{a+1} \frac{\partial^2 Y}{\partial J \partial w_0}\,dw_0 \;=\; \int_a^{a+1} \frac{\partial J_0}{\partial J}\,dw_0 \;=\; \frac{\partial}{\partial J}\int_a^{a+1} J_0\,dw_0. \tag{17}$$

A similar calculation involving the original form of (7) shows that the range $[a, a+1]$ describes exactly one full period of motion, so $\int_a^{a+1} J_0\,dw_0 = J$. Hence $\Delta w = 1$. Recalling that brackets denote an average over one full cycle of w_0, we have shown that $\frac{\partial}{\partial J}\langle \frac{\partial Y}{\partial w_0}\rangle = 1$, so $\langle \frac{\partial Y}{\partial w_0}\rangle = J + c(\epsilon)$. Since $\frac{\partial Y}{\partial w_0} = J_0$ and J and J_0 vanish together, $c(\epsilon) = 0$. On the other hand,

$$Y = w_0 J + \epsilon Y_1(J, w_0) + o(\epsilon),$$

from which it follows immediately that $\langle \frac{\partial Y_1}{\partial w_0}\rangle = 0$, the fact needed in the derivation of (14) above.

Next, we use the fact that $\frac{\partial Y_1}{\partial w_0}$ has mean zero to find an explicit formula for it. First, it follows from (12) that

$$\alpha(J) \;=\; \nu_0 J + \epsilon\langle H_1(J_0, w_0)\rangle + o(\epsilon).$$

On the other hand,

$$\alpha(J) \;=\; H(J_0(J,w), w_0(J,w)) \;=\; \nu_0 J_0 + \epsilon H_1(J_0, w_0) + o(\epsilon).$$

Combining these, we have

$$\nu_0 J + \epsilon\langle H_1(J_0, w_0)\rangle \;=\; \nu_0 J_0 + \epsilon H_1(J_0, w_0) + o(\epsilon).$$

Together with the transformation equation $J_0 = J + \epsilon\frac{\partial Y_1}{\partial w_0} + o(\epsilon)$ this yields

$$\frac{\partial Y_1}{\partial w_0} = \frac{\langle H_1\rangle - H_1}{\nu_0}. \tag{18}$$

Let us turn now to the calculation of α_2. In addition to the 3 terms of order ϵ^2 in (14) there is a contribution of

$$\epsilon\left\langle \frac{\partial^2 H_1}{\partial J^2}(J_0, w_0)(J - J_0)\right\rangle \tag{19}$$

from the ϵ term there.

Here it is important to stress the difference between an arbitrary value of J, as given by the coordinate transformation equation

$$J = J_0 - \epsilon\frac{\partial Y_1}{\partial w_0} + o(\epsilon),$$

and the *physical* value J has at the given energy $E = J_0\nu_0$. It is the latter that is needed for $J - J_0$ here. Note that we only need this increment to order ϵ.

For the physical value of J, we have that $\alpha(J)$ equals

$$\begin{aligned}(20)\quad H(J_0(J,w),w_0(J,w)) &= \nu_0 J_0(J,w) + \epsilon H_1(J_0(J,w),w_0(J,w)) + o(\epsilon)\\ &= \nu_0 J_0 + \epsilon H_1(J,w_0) + o(\epsilon) = \nu_0\left(J + \epsilon\frac{\partial Y}{\partial w_0}\right) + \epsilon H_1(J,w_0) + o(\epsilon),\end{aligned}$$

which is equal to $\nu_0 J + \epsilon\langle H_1(J,w_0)\rangle + o(\epsilon)$ using the transformation equation. Thus, by (10),

$$(21)\qquad \alpha(J) = \nu_0 J - \frac{J^2\epsilon}{64\pi^2 I\theta_0^2} + o(\epsilon) = \nu_0 J - \frac{\epsilon\nu_0 J^2}{32J_0} + o(\epsilon).$$

On the other hand, $\alpha(J) = E$, i.e.,

$$\nu_0 J_0 = E = \nu_0 J - \frac{\epsilon}{32J_0}\nu_0 J^2 + o(\epsilon).$$

Solving for J yields $J - J_0 = \frac{J_0\epsilon}{32} + o(\epsilon)$. Using this and $J_0 = 2\pi^2 I\nu_0\theta_0^2$ we obtain finally the desired estimate of (19):

$$(22)\quad \epsilon\left\langle\frac{\partial^2 H_1}{\partial J^2}\right\rangle(J - J_0) = \left(\frac{-2}{24\pi^2 I\theta_0^2}\right)\left(\frac{3}{8}\right)\left(\frac{1}{32}\right)J_0\epsilon^2 + o(\epsilon^2) = -\frac{\nu_0\epsilon^2}{2^9} + o(\epsilon^2).$$

For all other terms in (14) we may evaluate at $J = J_0$. By (10) and (15) we have

$$(23)\qquad \epsilon^2\left\langle\frac{\partial H_2}{\partial J}\right\rangle = \left(\frac{3J_0^2\epsilon^2}{1440\pi^4 I^2\nu_0\theta_0^4}\right)\left(\frac{5}{16}\right) = \frac{1}{3\times 2^7}\nu_0\epsilon^2.$$

By (18), (10), and (15),

$$\begin{aligned}(24)\quad \epsilon^2\left\langle\frac{\partial^2 H_1}{\partial J^2}\frac{\partial Y_1}{\partial w_0}\right\rangle &= \frac{\epsilon^2}{\nu_0}\left\{\left\langle\frac{\partial^2 H_1}{\partial J^2}\right\rangle\langle H_1\rangle - \left\langle\frac{\partial^2 H_1}{\partial J^2}H_1\right\rangle\right\}\\ &= -\left(\frac{2J_0^2\epsilon^2}{24^2\pi^4\nu_0 I^2\theta_0^4}\right)\left(\frac{17}{128}\right) = -\frac{17}{9\times 2^{10}}\nu_0\epsilon^2.\end{aligned}$$

Similarly,

$$\epsilon^2\left\langle\frac{\partial H_1}{\partial J}\frac{\partial^2 Y_1}{\partial J\partial w_0}\right\rangle = -\frac{17}{9\times 2^9}\nu_0\epsilon^2.$$

Setting $\epsilon = \theta_0^2$ and combining all terms,

$$\nu = \nu_0 - \frac{1}{16}\nu_0\theta_0^2 - \frac{5}{2^{10}}\nu_0\theta_0^4 + o(\theta_0^4),$$

i.e., $\alpha_2 = -\frac{5}{2^{10}}$.

To compare this result with the literature, it is necessary to make two further adjustments. First, we obtain the corresponding expansion of the period, τ, by taking reciprocals:

$$\tau = \tau_0 + \frac{1}{16}\tau_0\theta_0^2 + \frac{9}{2^{10}}\tau_0\theta_0^4 + o(\theta_0^4),$$

where $\tau_0 = \frac{1}{\nu_0}$.

Secondly, the parameter θ_0 is the amplitude of an idealized harmonic oscillator. To obtain an expansion in powers of the square of the actual amplitude θ, set the two expressions for potential energy equal: $\frac{1}{2}\omega_0^2 I\theta_0^2 = mgl(1-\cos\theta)$, or, since $\omega_0^2 = \frac{g}{l}$ and $I = ml^2$,

$$\frac{1}{2}\theta_0^2 = 1 - \cos(\theta) = \frac{1}{2}\theta^2 - \frac{1}{24}\theta^4 + \dots.$$

Thus,

$$\theta_0^2 = \theta^2 - \frac{\theta^4}{12} + \dots.$$

Substituting this for θ_0^2 above, we obtain finally the desired period expansion to second order in the actual amplitude θ:

$$\tau = \tau_0\left(1 + \frac{1}{16}\theta^2 + \frac{11}{3072}\theta^4 + \dots\right).$$

4. Period Expansion via Time Dependent Perturbation

We shall assume that the true pendulum motion is time periodic with frequency ν. (The fact that the motion must be strictly periodic was discussed in the introduction.) The frequency depends on the value of the (constant) action, J. If we express H in terms of the action-angle pair, then $H = H(J)$, as explained in section 2. The equation of motion for the angle variable w, $w(t) = \frac{\partial H}{\partial J}$, yields the solution $w(t) = \nu t + \beta$, where $\nu = \frac{\partial H}{\partial J}$ and β is the constant of integration. As in the previous section, we attach subscript 0 to denote analogous quantities for the harmonic oscillator Hamiltonian, H_0.

In time dependent perturbation theory we view the small change in Hamiltonian from H_0 to H as the generator of a slow time variation in the quantities β_0 and J_0, which are constant in the unperturbed system. The resultant rate of change in the phase angle β_0 can then be incorporated as a small correction to the unperturbed frequency.

For the purposes of this perturbation theory it is convenient to "factor out" the unperturbed motion by writing

$$w(t) \;=\; \nu_0 t + \hat{\beta}(t),$$

thus defining a certain function $\hat{\beta}$. Clearly $\hat{\beta}$'s equation of motion can be written in terms of a *perturbation Hamiltonian,* $\Delta H = H - H_0 = \epsilon H_1 + \epsilon^2 H_2 + \dots$, as

$$\dot{\hat{\beta}} \;=\; \frac{\partial \Delta H}{\partial J}(J, \nu_0 t + \hat{\beta}), \quad \hat{\beta}(0) = \beta. \tag{25}$$

The notation $\hat{\beta}$ will be short-lived, as we shortly consider certain approximate solutions. While the exact $\dot{\hat{\beta}}$ is constant, the corresponding quantity for the approximations is not. It is therefore simpler to consider their time averages. In this section we denote the average with respect to time over one full period $(\frac{1}{\nu})$ using angled brackets. In this notation the exact result could be written

$$\nu = \nu_0 + \langle\dot{\hat{\beta}}\rangle = \nu_0 + \nu\int_0^{\frac{1}{\nu}} \dot{\hat{\beta}}(s)\,ds. \tag{26}$$

Similarly, $J - J_0$ is a time-dependent function $\hat{J}$ that satisfies

$$\dot{\hat{J}} \;=\; -\frac{\partial \Delta H}{\partial w}(J_0 + \hat{J}, \nu_0 t + \hat{\beta}), \;\; \hat{J}(0) = J - J_0. \tag{27}$$

Of course differential equations (25) and (27) are coupled since, e.g., the right side of (25) depends on both $\hat{\beta}$ and $\hat{J}$.

Time dependent perturbation theory uses a method of successive approximations to solve the system (25) and (27) approximately. More precisely, for $n \geq 1$ one defines functions β_n and J_n inductively by

$$\beta_n(t) = \beta_n(0) + \int_0^t \frac{\partial \Delta H}{\partial J} \left(J_{n-1}(s), \nu_0 s + \beta_{n-1}(s)\right)\right) ds,$$

and

$$J_n(t) = J_n(0) - \int_0^t \frac{\partial \Delta H}{\partial w} \left(J_{n-1}(s), \nu_0 s + \beta_{n-1}(s)\right)\right) ds.$$

The constants $\beta_n(0)$ and $J_n(0)$ need to be chosen so that we have convergence to the physical values at time zero. Standard analytic methods (see appendix) show that on any compact time interval $[0, T]$, we then have $\beta_n \to \hat{\beta}$, a solution of (25), uniformly provided ϵ is sufficiently small depending on T; indeed, if we choose the constants appropriately, $|\beta_n(t) - \hat{\beta}(t)| \leq C\epsilon^{n+1}$, where $C = C(T)$. Thus, for the calculation of ν at order ϵ^2 via (26) it suffices to compute β_1 and β_2. As an additional simplification, we may ignore terms in the perturbation Hamiltonian that are too small to affect the results at a given order of approximation. Thus, β_1 may be computed using only the first term in ΔH, β_2 using only the first 2 terms, and so on.

It is important to bear in mind that $\beta_1, \beta_2, \ldots,$ do not represent the physical motion of *any* system. For example, the explicit form of β_2 given below shows that it is, in general, a quasi-periodic function, containing periodic components with incommensurate frequencies for most values of ϵ. Since such functions have no identifiable finite period, one might consider, erroneously, the long-time or *ergodic* average

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \dot{\beta}_2(s)\, ds.$$

This has, however, a quite different value from the average $\langle \dot{\beta}_2 \rangle$ due to cancellation effects. Indeed, one should not expect good agreement since β_2 only approximates $\hat{\beta}$ well over short time intervals.

Now $\dot{\beta}_1 = \epsilon \frac{\partial H_1}{\partial J}(J_0, \beta_0 + \nu_0 t) + o(\epsilon)$, so to first order in ϵ we have (see (10))

$$\langle \dot{\beta}_1 \rangle = -\frac{J_0 \epsilon}{12\pi^2 I \theta_0^2} \langle \cos^4 2\pi(\nu_0 t + \beta_0) \rangle.$$

(It is helpful to recall that $J_0 = 2\pi^2 I \nu_0 \theta_0^2$.) (Here, and in the rest of this section, one should interpret the $o(\epsilon)$ and similar estimates as holding uniformly over some fixed compact time interval containing $[0, \frac{1}{\nu}]$.) Noting that

$$\begin{aligned} (28) \quad \langle \cos^4 2\pi(\nu_0 t + \beta_0) \rangle &= \nu \int_0^{\frac{1}{\nu}} \cos^4 2\pi(\nu_0 t + \beta_0)\, dt \\ &= \nu_0 \int_0^{\frac{1}{\nu_0}} \cos^4 2\pi(\nu_0 t + \beta_0)\, dt + o(1), \end{aligned}$$

we obtain $\langle \dot{\beta}_1 \rangle = -\frac{\nu_0 \epsilon}{16} + o(\epsilon)$, in agreement with the results of section 3.

For the calculation of $\langle\dot{\beta}_2\rangle$ we need the explicit formulas for $\beta_1(t)$ and $J_1(t)$. They are

$$\beta_1(t) = \beta_1(0) - \frac{1}{6}\nu_0\epsilon \int_0^t \cos^4 2\pi(\nu_0 s + \beta_0)\, ds, \tag{29}$$

and

$$J_1(t) = J_1(0) - \frac{1}{3}\frac{J_0^2\epsilon}{\pi I\theta_0^2}\int_0^t \cos^3 2\pi(\nu_0 s+\beta_0)\sin 2\pi(\nu_0 s+\beta_0)\, ds \tag{30}$$
$$= J_1(0) + \frac{J_0\epsilon}{12}\cos^4 2\pi(\nu_0 t+\beta_0).$$

We obtain β_2 by integrating

$$\dot{\beta}_2 = \epsilon\frac{\partial H_1}{\partial J}(J_1(t), \nu_0 t + \beta_1(t)) + \epsilon^2\frac{\partial H_2}{\partial J}(J_0, \nu_0 t+\beta_0).$$

The equation of motion does not determine the constant terms $J_1(0)$ and $\beta_1(0)$. This subtlety arose already in the previous section when we needed to determine the physical value of J to order ϵ. We found there that $J = J_0(1+\frac{\epsilon}{32}) + o(\epsilon)$. The solution consistent with $\langle\dot{J}\rangle = J_0(1+\frac{\epsilon}{32}) + o(\epsilon)$ is

$$J_1 = J_0\left(1+\frac{\epsilon}{32}+\frac{1}{12}\epsilon\left[\cos^4 2\pi(\nu_0 t+\beta_0) - \frac{3}{8}\right]\right).$$

The actual value of $\beta_1(0)$ is irrelevant to the time averages we wish to compute, and depends, ultimately, on the position of the pendulum at time zero. Let us merely assume the time origin has been chosen in such a way that we can continue to designate $\beta_1(0)$ as β_0.

By formula (10) for H_1 and (29), the H_1 contribution to $\dot{\beta}_2$ is

$$-\frac{\epsilon\nu_0}{6}\left(1+\frac{\epsilon}{32}+\frac{1}{12}\epsilon(\cos^4 2\pi(\nu_0 t+\beta_0) - \frac{3}{8})\right)\cos^4[2\pi(\nu_1 t+\beta_0 \tag{31}$$
$$-\frac{\epsilon}{48\pi}\sin 2\pi(2\nu_0 t+2\beta_0) - \frac{\epsilon}{384\pi}\sin 2\pi(4\nu_0 t+4\beta_0)] + o(\epsilon^2),$$

where $\nu_1 = \nu_0 - \frac{1}{16}\epsilon\nu_0$. The contribution of H_2 is the same as in the time independent approach - a spatial average being replaced by a time average. Expanding $\cos^4$ around $2\pi(\nu_1 t+\beta_0)$ and retaining only the first order terms produces (see (23) above,)

$$\langle\dot{\beta}_2\rangle = -\frac{1}{16}\nu_0\epsilon + \epsilon^2\nu_0\{\frac{1}{3\times 2^7} - \frac{1}{3\times 64}\langle\cos^4 2\pi(\nu_1 t+\beta_0)\rangle \tag{32}$$
$$-\frac{1}{36}\langle\cos^3 2\pi(\nu_1 t+\beta_0)\sin 2\pi(\nu_1 t+\beta_0)\sin 2\pi(2\nu_0 t+2\beta_0)\rangle$$
$$-\frac{1}{288}\langle\cos^3 2\pi(\nu_1 t+\beta_0)\sin 2\pi(\nu_1 t+\beta_0)\sin 2\pi(4\nu_0 t+4\beta_0)\rangle -$$
$$\frac{1}{72}[\langle\cos^4 2\pi(\nu_0 t+\beta_0)\cos^4 2\pi(\nu_0 t+\beta_0)\rangle - \frac{3}{8}\langle\cos^4 2\pi(\nu_1 t+\beta_0)\rangle]\} + o(\epsilon^2).$$

The frequency difference $\nu_0 - \nu_1$ is of order ϵ, and so it can be ignored at order ϵ^2. Setting $\nu_1 = \nu_0$ and using the double angle formula and (15), a calculation shows that the expression in brackets equals

$$\frac{1}{3\times 2^7} - \frac{1}{2^9} - \frac{1}{9\times 2^5} - \frac{1}{9\times 2^9} - \frac{17}{9\times 2^{10}} = -\frac{5}{2^{10}},$$

in agreement with the result of section 2.

5. Appendix

In this section we discuss some mathematical issues that arose in the main text of the article and attempt to point the reader toward results in the literature that can be used to make all of the arguments mathematically rigorous.

A rigorous proof, based on the equations of motion alone, that the pendulum's motion is strictly periodic in time must include showing that it will reach the maximum angle allowed by conservation of energy in finite time. More precisely, if the initial angular momentum and angle are $p_0 > 0$ and 0 respectively, with $\frac{p_0^2}{2ml^2} < 2mgl$, then $p(t) = 0$ at some $t > 0$. To see this, note that Hamilton's equations are $\dot{p} = -mgl\sin\theta$, and $\dot{\theta} = \frac{p}{ml^2}$. If the claimed behavior did not hold, the second equation would force θ to increase monotonically to some limiting value $\theta_\infty \leq \infty$. Energy conservation would then imply that $\theta_\infty < \pi$. But then the equation for $\dot{p}$ forces p to $-\infty$. (Note that if the initial momentum is such that the kinetic energy is exactly $2mgl$ the pendulum will approach the state of perfect balance in the vertical position, but this will take infinite time.)

We used several times implicitly the fact that the generating function, W, of the canonical transformation from the original pair of variables, p and θ, to the action-angle pair J, w is analytic in a neighborhood of 0. This generating function, called *Hamilton's characteristic function*, is closely tied to another generating function, S, known as *Hamilton's principal function,* which governs a transformation to a set of variables which are both constant in time. (Recall that J, the action variable, is constant, but w increases linearly with time. In the latter canonical transformation w is replaced by its value at time zero, i.e., the initial phase angle.)

Let us consider more generally a Hamiltonian $H(x, y)$ that does not depend explicitly on time and which is analytic in both variables near the origin. We begin by showing that S is also analytic near the origin. The idea is to construct S formally from a parametric family of solutions of the equations of motion.

Since H is analytic, there is an analytic *Hamiltonian flow* in a neighborhood of the origin, i.e., there is a function $F : \mathbb{C}^3 \to \mathbb{C}^2$ which is analytic in a suitable neighborhood of the origin in $\mathbb{C}^3$ (and real for real values of its variables,) such that $(x(t), y(t)) = F(\xi, \eta, t)$ satisfy Hamilton's equations with initial conditions $x(0) = \xi, y(0) = \eta$. Moreover, the transformation $(x, y, t) \to (\xi, \eta, t)$ is canonical and has generating function $S = S(x, \eta, t)$ that satisfies $y = \frac{\partial S}{\partial x}, \xi = \frac{\partial S}{\partial \eta}$, and the Hamilton-Jacobi equation

$$\frac{\partial S}{\partial t} = -H(x, \frac{\partial S}{\partial x}). \tag{33}$$

(See section 4 of [9] for a proof of the analyticity of F in the time variable. For analyticity in the parameters ξ and η see, e.g., [2], equation 21 on page 120. For the remaining facts, see sections 2 and 3 of [9].)

These results hold on a possibly smaller neighborhood where $\frac{\partial x}{\partial \xi} \neq 0$. Since $\frac{\partial x}{\partial \xi}$ is identically equal to 1 when $t = 0$, such a neighborhood exists, and one may solve for ξ as $\xi = \xi(x, \eta, t)$ by the implicit function theorem. (See e.g. [5], Theorem 2.1.2, for the requisite analytic version of the implicit function theorem.) Since $y = \frac{\partial S}{\partial x}$, we have

$$\frac{\partial S}{\partial x}(x, \eta, t) = y(\xi, \eta, t) = y(\xi(x, \eta), \eta, t),$$

showing that $\frac{\partial S}{\partial x}$ is analytic in a neighborhood of the origin. Finally, $S(x, \eta, 0) = c + x\eta$, a generator of the identity transformation, from which

$$S(x, \eta, t) = c + x\eta + \int_0^t \frac{\partial S}{\partial s}(x, \eta, s)\, ds.$$

This, together with (33) and the foregoing, yields the desired analyticity of S.

In the case of the pendulum Hamiltonian, it follows from (33) that S is also analytic in the expansion parameter ϵ. Let us continue to denote by (ξ, η) the canonical variables that arise in the dynamical solution $(p, \theta, t) \to (\eta, \xi, t)$ of the pendulum equations of motion.

Next we establish that the frequency, ν, (hence also the period τ,) is an analytic function of ϵ and the initial conditions ξ and η. Since $x(t)$ is a continuous periodic function it has a Fourier expansion

$$x(t) = \sum_{j=-\infty}^{\infty} a_j e^{2\pi i j \nu t}$$

with coefficients $a_j = a_j(\xi, \eta, \epsilon)$. The Hamiltonian flow argument can be used in a neighborhood of any t, so x is actually analytic in a neighborhood of the entire real t axis. It follows that the coefficients are analytic and tend to zero exponentially fast as $j \to \pm\infty$. (See, e.g., [1], pp. 80-81.) The desired analyticity of ν then follows, for example, from the expression

$$\nu = \frac{\dot{x}(0)}{2\pi i \sum_{j \neq 0} j a_j}.$$

Next, by (7) we have

$$J = \int_0^\tau p(t)\dot{\theta}(t)\, dt,$$

so that the action, J, is also analytic. The Hamiltonian expressed in terms of J and ϵ alone, $\alpha(J, \epsilon)$, is then easily seen to be analytic in J and ϵ; the frequency $\nu = \frac{\partial \alpha}{\partial J}$ is analytic in J also.

Consider the function $W(J, \theta) = S(J, \theta, t) + \alpha(J, \epsilon)t$, where S is Hamilton's principal function constructed above. Then W is an analytic function of J, θ, and ϵ, and since S satisfies the Hamilton-Jacobi equation (33) it follows that W satisfies the Hamilton-Jacobi equation (4) and generates the canonical coodinate transformation from (p, θ) to the action-angle pair (J, w). That is, W is Hamilton's characteristic function. Since the angle variable $w = \frac{\partial W}{\partial J}$, we have that the sets of coordinates $(p, \theta, t), (\eta, \xi, t)$ and (J, w, t) are analytic functions of each other and ϵ in a suitable neighborhood of the origin. To these we may add (J_0, w_0, t) (independent of ϵ,) since we have given their explicit transformation equations with (p, θ, t) above.

Finally, consider the generator, Y, of $(J_0, w_0, t) \to (J, w, t)$. Since $w = \frac{\partial Y}{\partial J}$ and $J_0 = \frac{\partial Y}{\partial w_0}$, the partials of Y are analytic because the coordinate changes are. Thus Y itself is analytic. (The coordinate transformations determine Y up to an additive function of time only. The fact that we have chosen to use a new Hamiltonian that does not depend on time further limits the indeterminacy of Y to an additive constant. This remaining indeterminacy is irreducible and can be traced to the fact that the Hamiltonian itself is indeterminate up to an additive constant.)

Aside from the these analyticity issues, the only other obvious mathematical *faux pas* in section 3 were differentiations under integral signs and interchanges amongst various integrations, summations, and differentiations. These can all easily be justified using analyticity of the various integrands.

The only real issue in section 4 was the convergence of the iterative scheme[5] for solving the equations of motion. Since ΔH is nicely Lipschitz in its argument, virtually any text on ordinary differential equations should contain adequate results. For example, see [2], Theorem 6, pp 112-115.

References

[1] N.K. Bari, *A Treatise on Trigonometric Series,* Volume I, MacMillan, NY, 1964.

[2] G. Birkhoff and G. Rota, *Ordinary Differential Equations* , Ginn, Boston, 1962.

[3] L.P. Fulcher and B.F. Davis, Theoretical and experimental study of the motion of the simple pendulum, *Am. J. Phys.,* **44**(1976),51-53.

[4] H. Goldstein, *Classical Mechanics,* 2nd Ed., Addison Wesley, Reading, 1980.

[5] L. Hörmander, *An Introduction to Complex Analysis in Several Variables,* Van Nostrand, Princeton, 1966.

[6] J.B. Marion, *Classical Dynamics of Particles and Systems*, Academic Press, New York, 1965.

[7] W. Milham, *Time and Timekeepers*, MacMillan, New York, 1929.

[8] R.A. Nelson and M.G. Olsson, The pendulum - rich physics from a simple system, *Am. J. Phys.,* **54**(1986),1-11.

[9] C.L. Siegel and J.K. Moser, *Lectures on Celestial Mechanics,* Springer Verlag, New York, 1971.

[10] K.R. Symon, *Mechanics*, 3rd Ed., Addison Wesley, Reading, 1971.

[11] J.L. Synge and B.A. Griffith, *Principles of Mechanics,* 3rd Ed., McGraw Hill, New York, 1959.

[12] E.T. Whittaker, *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies,* 4th Ed., Cambridge University Press, New York, 1937.

[5] This scheme is known as *Picard's method* in the literature.