Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

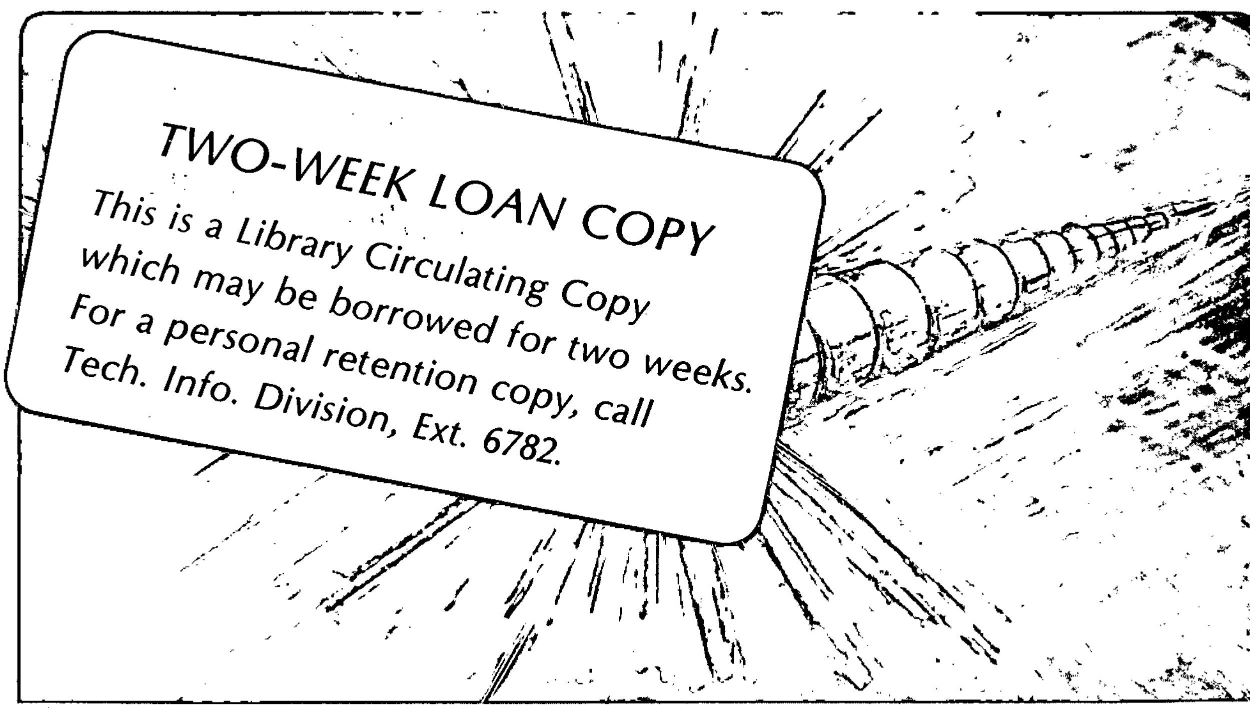
Accelerator & Fusion Research Division

Submitted to Physica D

CLASSICAL HAMILTONIAN PERTURBATION THEORY WITHOUT SECULAR TERMS OR SMALL DENOMINATORS

Henry D. I. Abarbanel

October 1979



Prepared for the U.S. Department of Energy under Contract W-7405-ENG-48

DISCLAIMER

Classical Hamiltonian Perturbation Theory Without Secular Terms or Small Denominators*

Henry D. I. Abarbanel

Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720

ABSTRACT

We consider perturbation theory in ε for the classical Hamiltonian $H = H_0 + \varepsilon H_1$ where H_0 gives rise to a known motion and ε is small. First we demonstrate how the usual secular terms and small denominators arise from a straightforward expansion in ε and argue that they are artifacts of the method. Then we present an alternative perturbation theory based on an analysis of the operator $(s-L)^{-1}$ where s is a complex number and L is the Liouville operator corresponding to H. This perturbation series contains neither secular terms nor small denominators. In the case of almost multiply periodic systems we show, to lowest non-trivial order in ε , how our series reproduces the standard results both in the resonant and non-resonant regions -- all in one analytic formula. As a final exercise we demonstrate that energy is conserved at order ε^{n+1} when the accuracy of the theory is order ε^n.

*Work supported by U. S. Department of Energy, under Contract W-7405-ENG-48.

I. Introduction

Perturbation theory in classical Hamiltonian systems has concentrated on various methods of effecting canonical transformations which transform away the perturbation order by order in the perturbation itself.[1] These conventional and Lie transform methods always suffer from problems with (a) secularity--which are overcome by some form of averaging technique[1] and (b) small denominators--which are overcome by a clever choice of Lie transform generating function[2] or a form of Kolmogorov's superconvergent perturbation theory[3].

In this paper we develop a perturbation theory for classical Hamiltonian mechanics which simultaneously avoids these two problems. If the Hamiltonian is written as

$$H = H_0 + \varepsilon H_1$$

with H_0, the unperturbed system, then our technique in some sense consists of writing the evolution of any function on phase space, $A(p_i,q_i)$ $i = 1,N$, as a ratio of power series in ε. The technique employed is known as Fredholm perturbation theory[4] as it resembles the familiar analysis of Fredholm integral equations. The denominator of the ratio of series in ε is the Fredholm determinant so commonly encountered in problems of quantum mechanical scattering theory.[5] The Fredholm determinant contains the exact eigenvalues of the Liouville operator

$$L = -\frac{\partial H}{\partial q_j}\frac{\partial}{\partial p_j} + \frac{\partial H}{\partial p_j}\frac{\partial}{\partial q_j}$$

which determine the time evolution of the system with Hamiltonian H. The eigenvalues of the unperturbed Liouville operator

$$L_0 = -\frac{\partial H_0}{\partial q_j}\frac{\partial}{\partial p_j} + \frac{\partial H_0}{\partial p_j}\frac{\partial}{\partial q_j}$$

are responsible for the small denominators, and the repeated small denominators in higher orders of conventional perturbation theory are responsible for the secular terms. In the second section of this paper we show how our perturbation theory avoids these problems.

In a third section we address the questions of motion of the system near a resonance of the unperturbed system--namely near the usual small denominators. We show how the frequencies of the motion in the neighborhood of the resonance are $O(\sqrt{\varepsilon})$ as are the deviations of momenta or action variables from their resonant values. Off resonance, as usual, the frequencies are order unity and the changes in momenta or action are $O(\varepsilon)$. The attractive feature of the present perturbation theory is that the same formulae can be used in both resonant and non-resonant regimes, providing, as it were, a smooth analytical interpolating method. Equally, the method is smooth near separatrices so may be used to construct adiabatic invariants and use them to examine motion through singular points. This may prove quite useful in the study of particle trapping by time varying electromagnetic waves. Finally, since we are dealing with approximations to the exact eigenfrequencies of the Liouville operator so that no secularities will be present, the perturbation theory developed here will be useful for studying the long time behavior of mechanical systems.

The language of this paper is couched in the mode of classical particle mechanics. Since eikonal ray trajectories obey particle like equations of motion with the dispersion relation playing the role of the Hamiltonian,[6] the perturbation theory developed here is useful for that situation as well.

II. Fredholm Perturbation Theory

We want to determine the evolution in time of a function $A(p_j,q_j)$ on the phase space p_j, q_j $j = 1,\ldots,N$ of particles whose motion is governed by the Hamiltonian

$$H = H_0(p_j) + \varepsilon H_1(p_j,q_j). \tag{1}$$

We take the system determined by H_0 to be soluble and have made a canonical transformation to make H_0 a function of the p_j only. The question then is to make a calculation of the time development of A treating ε as a small dimensionless parameter setting the scale of perturbation of the original system.

Central to the time development of A (p_j,q_j) is the Liouville operator[7]

$$L = \sum_{j=1}^{n}\left(\frac{\partial H}{\partial p_j}\frac{\partial}{\partial q_j} - \frac{\partial H}{\partial q_j}\frac{\partial}{\partial p_j}\right) \tag{2}$$

$$= \frac{\partial H_0(p_\ell)}{\partial p_j}\frac{\partial}{\partial q_j} + \varepsilon\left[\frac{\partial H_1}{\partial p_j}\frac{\partial}{\partial q_j} - \frac{\partial H_1}{\partial q_j}\frac{\partial}{\partial p_j}\right] \tag{3}$$

$$= L_0 + \varepsilon L_1. \tag{4}$$

As is carefully explained in Reference 1, the time development of $A(p_j,q_j)$ is given by operating on A by the time evolution operator $T(t)$ satisfying

$$\frac{dT(t)}{dt} = LT(t). \tag{5}$$

The function T(t)A evaluated at p_j, q_j yields our answer, $A(P_j(p_j,q_j,t), Q_j(p_j,q_j,t))$, where P_j and Q_j are the momenta and co-ordinates which develop at time t from p_j and q_j at time t = 0:

$$P_j(p_j,q_j,t = 0) = p_j, \tag{6}$$

and

$$Q_j(p_j,q_j,\ t = 0) = q_j. \tag{7}$$

We take H to be time independent and from (2) see that L is also time independent. We are led then to Laplace transform (5); with the operator

$$U(s) = \int_0^\infty dt\ e^{-st}T(t) \tag{8}$$

we have the operator equation

$$(s-L_0-\varepsilon L_1)U(s) = \mathbf{1} \tag{9}$$

with the identity operator on the right. The desired time dependent phase function is

$$A(P_j(p_j,q_j,t),Q_j(p_j,q_j,t)) = \int_{c-i\infty}^{c+i\infty} \frac{ds}{2\pi i}\, e^{st}\left[\frac{1}{s-L_0-\varepsilon L_1}\right]A(p_j,q_j) \tag{10}$$

where c is to the right of all singularities of U(s) in the s-plane.

Repeated use of the operator identity

$$\frac{1}{A-B} = \frac{1}{A} + \frac{1}{A}\, B\, \frac{1}{A-B} \tag{11}$$

generates the usual perturbation theory in ε as

$$\frac{1}{s-L_0-\varepsilon L_1} = \frac{1}{s-L_0} + \frac{1}{s-L_0}\,\varepsilon L_1\,\frac{1}{s-L_0} + \frac{1}{s-L_0}\,\varepsilon L_1\,\frac{1}{s-L_0}\,\varepsilon L_1\,\frac{1}{s-L_0} + \ldots \tag{12}$$

The secular terms of usual perturbation theory are located in the repeated poles of this formula at the eigenvalues λ_0 of the unperturbed Liouville operator L_0.

Consider the second term of this series. Introduce the eigenfunctions $\psi_0{}^{(a)}(p_j,q_j)$ and eigenvalues $\lambda_0^{(a)}$ of L_0

$$L_0\psi_0{}^{(a)}(p_j,q_j) = \lambda_0{}^{(a)}\psi_0{}^{(a)}(p_j,q_j)$$

which we'll assume form a complete set on phase space.

The second term in (12) is written (using Dirac bra-ket notation)

$$\sum_{a,a'} |a\rangle \frac{1}{s-\lambda_0^{(a)}} \langle a|\varepsilon L_1|a'\rangle \frac{1}{s-\lambda_0^{(a')}} \langle a'|, \tag{13}$$

where

$$\langle a|L_1|a'\rangle = \int d^Nq\,d^Np\,\bar{\psi}_0^{(a)}(p_j,q_j)L_1\psi_0^{(a')}(p_j,q_j) \tag{14}$$

In doing the Laplace inversion one encounters a double pole at the term in the a, a' sums where a = a'. Unless $\langle a|L_1|a\rangle = 0$, this will generate a term in $A(P_j,Q_j)$ linear in t. If the diagonal element $\langle a|L_1|a\rangle$ vanishes, as it well may, then the third term of the series in (12) yields

$$\varepsilon^2 \sum_{a,b,c} |a\rangle\frac{1}{s-\lambda_0^{(a)}}\langle a|L_1|b\rangle\frac{1}{s-\lambda_0^{(b)}}\langle b|L_1|c\rangle\frac{1}{s-\lambda_0^{(c)}}\langle c| \tag{15}$$

which has a double pole at, say, a = c with residue

$$\sum_{b\neq a} \langle a|L_1|b\rangle \frac{1}{\lambda_0^{(a)}-\lambda_0^{(b)}} \langle b|L_1|a\rangle \tag{16}$$

which is unlikely to vanish. Were this also to vanish, some term along the line would have a non-vanishing residue generating powers of t--i.e. secular terms, in the expansion of $A(P_j,Q_j)$.

The small denominators are also exhibited in (12). To illustrate this, take p_j and q_j to be action and angle variables I_j and θ_j. The eigenfunctions of L_0 are now just $\exp i\, \vec{m}\cdot\vec{\theta}$ where $\vec{m}$ is an N-vector with integer components $\vec{m} = (m_1,\ldots,m_N)$, $-\infty < m_i < \infty$. The eigenvalues are $\lambda_0^{\vec{m}} = i\vec{m}\cdot\vec{\omega}_0(I_j)$, $\omega_{0\ell}(I_j) = \partial H_0/\partial I_\ell$. The term of order ε in the perturbation series is

$$\varepsilon \sum_{m,n} |\vec{m}\rangle \frac{1}{s-i\vec{\omega}_0\cdot\vec{m}} \langle\vec{m}|L_1|\vec{n}\rangle \frac{1}{s-i\vec{\omega}_0\cdot\vec{n}} \langle\vec{n}| \; . \tag{17}$$

In the integration over s we pick up the pole at $s = i\vec{\omega}_0\cdot\vec{m}$, say, with residue

$$\varepsilon \sum_n |\vec{m}\rangle\langle\vec{m}|L_1|\vec{n}\rangle \frac{1}{i\vec{\omega}_0\cdot(\vec{m}-\vec{n})} \langle\vec{n}| \; , \tag{18}$$

and for resonant values of $\vec{I}$, on which $\vec{\omega}_0$ depends, the denominator will vanish.

Both of these problems with conventional perturbation theory arise from the repeated appearance of the operator $(s-L_0)^{-1}$. However, it is clear from (9) that the function U(s) is singular not at s = eigenvalue of L_0 but at s = eigenvalue of $L_0 + \varepsilon L_1$, the full Liouville operator. These eigenvalues are in general not at $\lambda_0^{(a)}$ but are "renormalized" or shifted by the action of L_1 which mixes the eigenmodes of L_0. Secular terms and small denominators are artifacts of the method of perturbation theory, they are not present in the full answer.

In the inversion of an operator, such as $s-L_0-\varepsilon L_1$, one expects to encounter something like

$$(s-L_0-\varepsilon L_1)^{-1} = \text{Operator}/\det(s-L_0-\varepsilon L_1)$$

by analogy with the usual rule for matrix inversion. The determinant of an operator is to be interpreted by the rule

$$\det(s-L_0-\varepsilon L_1) = \exp \operatorname{tr} \log(s-L_0-\varepsilon L_1) \tag{19}$$

and, expanding the logarithm in ε, we encounter the trace of products of the operator $\mathcal{O} = (s-L_0)^{-1}\varepsilon L_1$. The trace of an operator is independent of basis and can be evaluated in terms of any complete set, call it $F_n(p_j,q_j)$ as

$$\operatorname{tr} \mathcal{O} = \sum_n \langle n|\mathcal{O}|n\rangle, \tag{20}$$

where

$$\langle n|\mathcal{O}|n'\rangle = \int d^Npd^Nq\; \bar{F}_n(p_j,q_j)\; \{\mathcal{O}F_{n'}(p_j,q_j)\}. \tag{21}$$

Fredholm perturbation theory represents $(s-L_0-\varepsilon L_1)^{-1}$ as a ratio of power series in ε. Each series is entire in ε with the denominator vanishing at the poles of $(s-L_0-\varepsilon L_1)^{-1}$, which is to say at the true eigenvalues of $L = L_0 + \varepsilon L_1$.

To proceed we write

$$\beta(s) = (s-L_0)U(s) \tag{22}$$

so

$$\left(1-\varepsilon L_1 \frac{1}{s-L_0}\right)\beta(s) = 1, \tag{23}$$

and call

$$M = L_1 \frac{1}{s-L_0} \tag{24}$$

We'll explore this operator in a moment. We seek the operator

$$\beta(s) = \frac{1}{1-\varepsilon M} \tag{25}$$

Now write

$$\frac{1}{1-\varepsilon M} = \frac{[\mathfrak{h}(\varepsilon)/1-\varepsilon M]}{\mathfrak{h}(\varepsilon)} \quad , \tag{26}$$

where $\mathfrak{h}(\varepsilon)$ is to be entire in ε and have zeroes at the poles of $(1-\varepsilon M)^{-1}$. It is the Fredholm determinant. The standard choice for $\mathfrak{h}(\varepsilon)$ is

$$\mathfrak{h}(\varepsilon) = \exp -\int_0^\varepsilon dx \ \mathrm{tr}\left[M \frac{1}{1-xM}\right] \tag{27}$$

though as discussed by Sugar and Blankenbecler[8] many choices are possible.

Next write $\mathfrak{h}(\varepsilon)$ as a power series in ε

$$\mathfrak{h}(\varepsilon) = \sum_{n=0}^{\infty} z_n \varepsilon^n. \tag{28}$$

By expanding

$$\mathrm{tr}\ M \frac{1}{1-xM} = \sum_{n=0}^{\infty} x^n \ \mathrm{tr}(M^{n+1}), \tag{29}$$

we derive the following recursion relation for the z_n

$$(n+1)z_{n+1} + \sum_{k=0}^{n-1} z_k \ \mathrm{tr}(M^{n-k+1}) = 0 \tag{30}$$

and $z_0 = 1$, $z_1 = -\mathrm{tr}\, M$. The first few terms are

$$z_2 = \frac{1}{2}\{(\mathrm{tr}\ M)^2 - \mathrm{tr}\ M^2\}, \tag{31}$$

$$z_3 = -\frac{1}{6}((\mathrm{tr}\ M)^3 - 3\ \mathrm{tr}\ M\ \mathrm{tr}\ M^2 + 2\ \mathrm{tr}\ M^3). \tag{32}$$

Of course, the same result follows from

$$\det(1-\varepsilon M) = \exp\ \mathrm{tr}\ \log\ (1-\varepsilon M) \tag{33}$$

upon expansion of the right hand side in ε.

In a similar fashion we expand the numerator of (26) in ε

$$N(\varepsilon) = \mathfrak{h}(\varepsilon)/_{1-\varepsilon M} = \sum_{n=0}^{\infty} \varepsilon^n w_n \quad . \tag{34}$$

The w_n are operators whose relation to the z_n are

$$w_n = z_n + \sum_{\ell=1}^{n} z_{n-\ell} M^\ell \tag{35}$$

This means we may write

$$\beta(s) = \left(1 + \frac{\sum_{n=1}^{\infty} b_n \varepsilon^n}{1 + \sum_{n=1}^{\infty} z_n \varepsilon^n}\right)1 \tag{36}$$

with $b_n = \sum_{\ell=1}^{n} z_{n-\ell} M^\ell$. The low order b_n operators are

$$b_1 = M \tag{37}$$

$$b_2 = M^2-(\mathrm{tr}\ M)M \tag{38}$$

For the operator U(s) we finally have

$$U(s) = \frac{1}{s-L_0}\left\{1 + \frac{\sum_{n=1}^{\infty} b_n \varepsilon^n}{1 + \sum_{n=1}^{\infty} z_n \varepsilon^n}\right\} \tag{39}$$

which is our central result. The inversion of the operator $s-L_0$ is straightforward, by assumption, since we took H_0 to be soluble. It is also straightforward in actual practice.

In using this result we must be careful when doing the contour integration in the s plane to recover $A(P_j,Q_j)$ not to pick up the residue at the apparent pole of $(s-L_0)^{-1}$ since formally it is absent. That is clear from (7) and the appearance of $(s-L_0)^{-1}$ is due to our writing

$$(s-L_0-\varepsilon L_1)^{-1} = (s-L_0)^{-1}(1-\varepsilon L_1 \frac{1}{s-L_0})^{-1} \quad (40)$$

which, since it is an identity has no pole at the eigenvalues of L_0. Formally, the series, both numerator and denominator in (39) diverge at s = eigenvalues of L_0--and diverge in such a way as to cancel the apparent pole in $(s-L_0)^{-1}$. The only singularities present in the s plane are those of $\zeta(\varepsilon) = 1 + \sum_{n=1}^{\infty} z_n(s)\varepsilon^n$ where we note that z_n depends on s since M does.

III. Illustrating the Perturbation Series

In this section we give some illustration of our perturbation series by considering first its application to small perturbations around a multiply periodic system. The system is described by N actions, I_j, and N conjugate angles, θ_j which lie in $0 \leqslant \theta_j \leqslant 2\pi$. The problem is periodic in θ_j, and $H = H_0(\vec{I}) + \varepsilon H_1(\vec{I},\vec{\theta})$.

The L_0 operator is

$$L_0 = \omega_j(\vec{I})\frac{\partial}{\partial\theta_j} \quad , \quad \omega_j(\vec{I}) = \frac{\partial}{\partial I_j} H_0(\vec{I}), \quad (41)$$

and its eigenvalues are

$$\lambda_0^{(\vec{m})}(\vec{I}) = i\vec{m}\cdot\vec{\omega}(\vec{I}) \quad (42)$$

The quantity of direct interest to us is (see Equation (10))

$$(s-L_0-\varepsilon L_1)^{-1}A(\vec{I},\vec{\theta}). \quad (43)$$

We use the periodicity in θ_j to expand A in a fourier series

$$A(\vec{I},\vec{\theta}) = \sum_{\vec{n}} A(\vec{n},\vec{I})e^{i\vec{n}\cdot\vec{\theta}} \quad (44)$$

and write (43) as

$$\sum_{\vec{n},\vec{m}} \frac{e^{i\vec{n}\cdot\vec{\theta}}}{s-i\vec{\omega}(\vec{I})\cdot\vec{n}} [(1-\varepsilon M)^{-1}]_{\vec{n},\vec{m}} A(\vec{m},\vec{I}), \quad (45)$$

where $M = L_1/s-L_0$ has the $\vec{n},\vec{m}$ matrix element

$$M_{\vec{n},\vec{m}} = \int \frac{d^N\theta}{(2\pi)^N} e^{-i\vec{n}\cdot\vec{\theta}} M e^{i\vec{m}\cdot\vec{\theta}} \quad (46)$$

$$= \frac{1}{s-i\vec{m}\cdot\vec{\omega}(\vec{I})}\Big\{ im_\ell\Big[\frac{\partial}{\partial I_\ell} h(\vec{n}-\vec{m},\vec{I})\Big] + ih(\vec{n}-\vec{m},\vec{I})(m-n)_\ell \frac{\partial}{\partial I_\ell} + h(\vec{n}-\vec{m},\vec{I})(n-m)_\ell m_k \mathcal{H}_{\ell k}(\vec{I})/s-i\vec{m}\cdot\vec{\omega}(\vec{I})\Big\}. \quad (47)$$

In this we have written the fourier component of $H_1(\vec{I},\vec{\theta})$ as

$$h(\vec{n},\vec{I}) = \int\frac{d^N\theta}{(2\pi)^N} e^{-i\vec{n}\cdot\vec{\theta}} H_1(\vec{I},\vec{\theta}). \quad (48)$$

$h(\vec{0},\vec{I})=0$, since that term is in $H_0(\vec{I})$. Further we have introduced the Hessian tensor

$$\mathcal{H}_{ab}(\vec{I}) = \frac{\partial\omega_a(\vec{I})}{\partial I_b} = \frac{\partial^2 H_0(\vec{I})}{\partial I_a \partial I_b}. \quad (49)$$

Let us denote the fourier component of $(s-L_0-\varepsilon L_1)^{-1}A(\vec{I},\vec{\theta})$ by $B(\vec{n},\vec{I},s)$

$$(s-L_0-\varepsilon L_1)^{-1}A(\vec{I},\vec{\theta}) = \sum_{\vec{n}} e^{i\vec{n}\cdot\vec{\theta}}B(\vec{n},\vec{I},s). \tag{50}$$

The application of Fredholm perturbation theory to lowest non-trivial order in ε gives for B (see Equation (39)).

$$B(\vec{n},\vec{I},s) = \frac{1}{s-i\vec{n}\cdot\vec{\omega}(\vec{I})}\sum_{\vec{n}}\left[\delta_{\vec{m},\vec{n}} + \frac{\varepsilon M_{\vec{n},\vec{m}}}{1-\frac{\varepsilon^2}{2}\mathrm{tr}\,M^2}\right]A(\vec{m},\vec{I}), \tag{51}$$

since

$$\mathrm{tr}\,M = 0 \tag{52}$$

in our case.

Suppose we choose the phase function to be I_a, one of the actions. Then $A(\vec{n},\vec{I})$ is $\delta_{\vec{n},\vec{0}}I_a$ and

$$B(\vec{n},\vec{I},s) = \frac{I_a}{s}\delta_{\vec{n},\vec{0}} - \frac{i\varepsilon h(\vec{n},\vec{I})n_a}{(1-\frac{\varepsilon^2}{2}\mathrm{tr}\,M^2)s(s-i\vec{n}\cdot\vec{\omega}(\vec{I}))} \tag{53}$$

We first consider the case of non-resonant values of $\vec{I}$; namely there is no vector of integers $\vec{L} = (\ell_1, \ell_2, \ldots, \ell_N)$ for which $\vec{L}\cdot\vec{\omega}(\vec{I}) = 0$. We want to find the zeroes of the Fredholm determinant det $(1-\varepsilon M)$. To the order in ε shown in Equation (53) we have written det $(1-\varepsilon M) = 1-\frac{1}{2}\varepsilon^2\,\mathrm{tr}\,M^2$. So we need to find values of s for which tr $M^2 \approx \varepsilon^{-2}$ to balance the term of order unity. We can see from the form of $M_{\vec{n},\vec{m}}$ given in Equation (47) that tr $M^2 = \sum_{\vec{m},\vec{n}} M_{\vec{m},\vec{n}}M_{\vec{n},\vec{m}}$ has denominators of the form $[(s-i\vec{m}\cdot\vec{\omega})(s-i\vec{n}\cdot\vec{\omega})]^{-2}$ as well as denominators of lower order.

We expect that for off resonant values the vanishing of det $(1-\varepsilon M)$ will occur at s near the unperturbed eigenvalues $\lambda_0^{\vec{k}} = i\vec{k}\cdot\vec{\omega}$ of L_0.

Near $i\vec{k}\cdot\vec{\omega}$ we write $s = i\vec{k}\cdot\vec{\omega} + \varepsilon\Delta s_k$ and note that Δs_k is given by

$$1 - \frac{1}{2}\sum_{\vec{n}\neq\vec{k}}\left\{\frac{|h(\vec{n}-\vec{k},\vec{I})|^2[(n-k)_a k_b \mathcal{H}_{ab}][(k-n)_j n_\ell \mathcal{H}_{j\ell}]}{(\Delta s_k)^2[\varepsilon\Delta s_k + i(\vec{k}-\vec{n})\cdot\vec{\omega}(\vec{I})]^2}\right\} = 0$$

or to lowest order in ε

$$(\Delta s_k)^2 = \frac{1}{2}\sum_{\vec{n}\neq\vec{k}} |h(\vec{n}-\vec{k},\vec{I})|^2[(n-k)_a k_b \mathcal{H}_{ab}][(n-k)_j n_\ell \mathcal{H}_{j\ell}] \Big/ [(\vec{k}-\vec{n})\cdot\vec{\omega}]^2 .$$

In reaching this result we have used $h(0,\vec{I}) = 0$ to eliminate contributions when $\vec{n} = \vec{k}$, since that would lead to Δs_k which is not of order unity. There is a potential small denominator in this last formula when $(\vec{k}-\vec{n})\cdot\vec{\omega}(\vec{I})$ is close to zero. If this occurs, then neglecting $\varepsilon\Delta s_k$ with respect to $(\vec{k}-\vec{n})\cdot\vec{\omega}$ was inaccurate and we must use the previous formula for Δs_k.

When we have no resonance then, the eigenvalues of the full Liouville operator $L_0 + \varepsilon L_1$ are shifted from their unperturbed values by $O(\varepsilon)$.

In the case of a resonant value of $\vec{I}$, so $\vec{L}\cdot\vec{\omega}(\vec{I}) = 0$ for some $\vec{L}$, the situation is changed. In evaluating tr M^2, we encounter terms with $\vec{m}$ = integer × $\vec{L}$ and $\vec{n}$ = different integer × $\vec{L}$ {for example $\vec{m} = \vec{L}$, $\vec{n} = 2\vec{L}$}. We find then a term behaving as s^{-4} in tr M^2, and this means s is $O(\sqrt{\varepsilon})$ if $1-\frac{\varepsilon^2}{2}\mathrm{tr}\,M^2$ is to vanish. More precisely we have $s = [\varepsilon|h(\vec{L},\vec{I})|^2 L_a\mathcal{H}_{ab}(\vec{I})L_b]^{1/2}$. From (53) we see that along the $\vec{L}$ direction, the shift in $\vec{I}$ from the resonant value is of order $(\varepsilon|h(\vec{L},\vec{I})|/L_a\mathcal{H}_{ab}(\vec{I})L_b)^{1/2}$, which, for small ε, is much larger than the $O(\varepsilon)$ shift in the non-resonant case.

One can see this immediately by expanding the Hamiltonian about its resonant value keeping only the $\vec{L}\cdot\vec{\theta}$ term in H_1. One finds[6]

$$H(\vec{I},\vec{\theta}) = H_0(\vec{I}_R) + \frac{1}{2}(I-I_R)_a(I-I_R)_b\mathcal{H}_{ab}(I_R) + 2\varepsilon\, h(\vec{L},\vec{I}_R)\cos\psi + O(\varepsilon^2) \quad (55)$$

with

$$\psi = \vec{L}\cdot\vec{\theta} \quad (56)$$

when $h(\vec{L},\vec{I}_R)$ is real. Then the fundamental frequencies of motion about the resonance are precisely $O(\sqrt{\varepsilon h(\vec{L},\vec{I}_R)L_a\mathcal{H}_{ab}L_b})$ and the deviation of $\vec{I}$ from $\vec{I}_R$ along $\vec{L}$ is $O(\sqrt{\varepsilon h(\vec{L},I_R)/L_a\mathcal{H}_{ab}L_b})$. The important feature of our Fredholm perturbation formalism is that this resonant motion as well as the smoother non-resonant motion is totally contained in the same analytic formula (53) or its expression at higher orders in ε.

As a second illustration we show that energy is formally conserved by this perturbation theory to the next higher order in ε than the one being computed. First keep $O(\varepsilon)$ only. We can always arrange for $\mathrm{tr}\, M = 0$ by incorporating the diagonal elements of H_1 into H_0, so to $O(\varepsilon)$ we have (see (9), (10), and (39))

$$H(p_j,q_j,s) = \frac{1}{s-L_0}[1 + \varepsilon M]H(p_j,q_j,t = 0). \quad (57)$$

Since $LH = 0$, we have

$$\varepsilon MH = -\frac{L_0H}{s} + \varepsilon^2 M\,\frac{L_0H_1}{s}, \quad (58)$$

so

$$H(p_j,q_j,s) = \frac{1}{s-L_0}\left[H - \frac{L_0H}{s} + O(\varepsilon^2)\right] = H(p_j,q_j,t=0)/s + O(\varepsilon^2). \quad (59)$$

To second order in ε we write

$$H(p_j,q_j,s) = \frac{1}{s-L_0}\left[1 + \frac{\varepsilon M + \varepsilon^2 M^2}{1-\frac{\varepsilon^2}{2}\mathrm{tr}\, M^2}\right]H(p_j,q_j,t=0) \quad (60)$$

Now $\frac{1}{s-L_0} = \frac{1}{s-L_0-\varepsilon L_1}(1 - \varepsilon M)$, so

$$H(p_j,q_j,s) = \frac{1}{s-L_0-\varepsilon L_1}[1 + O(\varepsilon^3)]H(p_j,q_j,t = 0) \quad (61)$$

$$= \frac{1}{s}[1 + O(\varepsilon^3)]H(p_j,q_j,\ t = 0). \quad (62)$$

So it is clear how at order ε^n, energy is conserved by our perturbation series to order ε^{n+1}. If one requires exact energy conservation at any given order of ε, then solving for, say, $p_1,\ldots,p_N$ and $q_1,\ldots,q_{N-1}$ by the Fredholm series and then for q_N from $H(p_j,q_j) = E$ will guarantee the desired result.

IV. Discussion and Comments

We have presented a perturbation formalism for classical Hamiltonian mechanics based on recognizing that the essential problem is the inversion of the operator $s-L_0-\varepsilon L_1$ where s is a complex variable, L_0 is the Liouville operator for the unperturbed Hamiltonian H_0, while L_1 is the same operator for the perturbation H_1; $H = H_0 + \varepsilon H_1$. By writing $(s-L_0-\varepsilon L_1)^{-1}$ = [entire power series in ε]/det $(s-L_0-\varepsilon L_1)$ we have shown explicitly how one avoids the secular terms arising from a straightforward expansion of $(s-L_0-\varepsilon L_1)^{-1}$ in ε. Further, if the system is almost multiply periodic this device avoids the appearance of the notorious small denominators that plague the usual perturbation theories. Actually we have demonstrated a further useful result in the case of almost multiply periodic systems by showing how the perturbation series we employ smoothly contains the behavior of the perturbed system both near, at, and away from resonances where the unperturbed frequencies are commensurate- or equivalently where L_0 has zero eigenvalues.

A "poor person's" version of the KAM theorem [9] is lurking in our point of view. By focusing on the zeroes of $\det(s-L_0-\varepsilon L_1)$ as containing the true frequencies of the perturbed system, we see that for unperturbed orbits which correspond to a non-zero (i.e. non-resonant) eigenvalue of L_0 the motion is modified only slightly, namely the frequencies are shifted by $O(\varepsilon)$. Thus most orbits are only slightly perturbed. However, orbits with zero eigenvalue of L_0 are shifted to orbits with non-zero eigenvalue of $L_0 + \varepsilon L_1$ and therefore their time dependence is fundamentally changed in character. Formally the inverse Laplace transform needed to define $A(P_j(p_j,q_j,t),Q_j(p_j,q_j,t))$ no longer has poles at zero frequency and constant or periodic motion in time is destroyed. Clearly in the neighborhood of a resonance some destruction takes place as well.

There are, as yet, numerous avenues unexplored by the work reported in this note. For example, since the Fredholm determinant contains the true frequencies of the perturbed situation, it also allows an exploration of the regime where the unperturbed motion is singular; this is the case near a separatrix and the present theory allows one, in an essentially analytic way, to explore the variation of adiabatic invariants in these singular regimes.

References

1. A general review of conventional and Lie transform techniques can be found in J. R. Cary "Lie Transforms and Their Use in Hamiltonian Perturbation Theory," LBL-Preprint 6350, June 1978; and in R. G. Littlejohn, "A Pedestrian's Guide to Lie Transforms: A New Approach to Perturbation Theory in Classical Mechanics," LBL-Preprint 8091, November 1978; and references in these papers.
2. A. N. Kaufman, private communication.
3. A. N. Kolmogorov, Dokl. Akad. Nauk. SSSR 98, 527 (1954); and Ref. 1.
4. P. M. Morse and H. Feshbach, Methods of Theoretical Physics, Vol II, p. 1018-1026, (McGraw-Hill Book Co., New York, 1953).
5. M. L. Goldberger and K. M. Watson, Collision Theory, Chapter 6, (John Wiley, 1964).
6. A. N. Kaufman, "Regular and Stochastic Particle Motion in Plasma Dynamics," LBL-report 9407, August 1979. Also B. V. Chirikov, "A Universal Instability of Many-Dimensional Oscillator Systems, Physics Reports 52 (1979).
7. The sign of L is opposite to that of Reference 1 and is convenient here. It no doubt will add to the confusion previously cleared up by Cary, Ref. 1, Sec. 5.
8. R. L. Sugar and R. Blankenbecler, Phys. Rev. 136, B472 (1964).
9. See the discussion of M. V. Berry in "Regular and Irregular Motion," Nonlinear Dynamics, ed. S. Jorna, Am. Inst. of Physics Conference Proceedings, Vol. 46, (A.I.P., New York, 1978).

TECHNICAL INFORMATION DEPARTMENT
LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720