X-641-70-403
PREPRINT

NASA TM X-65379

CLASSICAL ADIABATIC PERTURBATION THEORY

DAVID P. STERN

OCTOBER 1970

GSFC

GODDARD SPACE FLIGHT CENTER
GREENBELT, MARYLAND

CLASSICAL ADIABATIC PERTURBATION THEORY

David P. Stern
Laboratory for Space Physics
Goddard Space Flight Center
Greenbelt, Maryland 20771

Abstract

Methods of classical perturbation theory developed for small perturbations are extended to slowly (or adiabatically) perturbed systems, with slow dependence either on time or on dynamical variables. Specifically, the extension is performed for the canonical perturbation theory of Poincare and Von Zeipel, for the Krylov-Bogoliubov-Kruskal method of eliminating angle variables, for the general form of direct near-identity canonical transformations and for two of its realizations, based on the "conventional" generating function and on the Lie transform. In addition, the concepts of slow (or adiabatic) perturbations and of an implicit "small parameter" ε are clarified, as is the distinction between two alternative definitions of adiabatic invariance, and as an example the solution of the slowly perturbed harmonic oscillator up to and including $O(\varepsilon^3)$ is derived.

CONTENTS

INTRODUCTION

Perhaps the most widely studied perturbation problem in classical mechanics is that of perturbed periodic motion. If a motion is given that is soluble and periodic, the problem may be concisely defined as the derivation of an approximate solution for a motion that is slightly different.

This "slight change" applied to the motion is termed the perturbation and it usually belongs to one of two types: "small" perturbations and "slow" (or "adiabatic") ones. The difference between the two is best explained by assuming that the motion can be described by a Hamiltonian, although this condition is not essential. In a slightly perturbed motion the Hamiltonian may then be written

$$H = H^{(0)} + \varepsilon H^{(1)} + \varepsilon^2 H^{(2)} + \dots \qquad (1)$$

where $\varepsilon \ll 1$ is a small numerical parameter characterizing the magnitude of the perturbation and where the limit $\varepsilon \to 0$ corresponds to the unperturbed motion. A typical example would be the motion of a planet around the sun as perturbed by the planet Jupiter. In that case $H^{(0)}$ describes the planet's Keplerian motion in the sun's gravity field while $H^{(1)}$ describes the lowest order of the perturbation induced by Jupiter. The zero-order Hamiltonian is then proportional to the solar mass m_s while $\varepsilon H^{(1)}$ is proportional to the mass m_j of Jupiter: the ratio of the two terms will be of the order (m_j/m_s) (about 10^{-3}) and this dimensionless quantity provides a natural choice for ε .

To illustrate a slow perturbation, consider a Hamiltonian that is slowly dependent on the time t (slow dependence may also involve canonical variables):

$$H = H(\underline{p}, \underline{q}, t) \qquad (2)$$

Then the dependence is said to be <u>slow</u> if the terms produced by the operation $\partial/\partial t$ are by an order in ε smaller than the terms from which they are derived, e.g.

$$\partial H/\partial t = O(\varepsilon H) \qquad (3)$$

The preceding equation is not quite precise, since it implies that ε has the dimension of t^{-1}. In fact, one always requires some natural time period T against which the rapidity of the time variation may be gauged, this usually being the period of the unperturbed system. With this taken into account, (3) becomes

$$\partial H/\partial t = O(\varepsilon H/T) \qquad (4)$$

and ε is clearly dimensionless.

In either type of problem there generally exists a steadily increasing "angle variable" appearing in the argument of sines and cosines, describing the nearly-periodic part of the motion. One way of "solving" the problem then involves finding a transformation to new variables, such that the angle variable is eliminated from the equations of motion. If the system also possesses a Hamiltonian H, the absence of the angle variable from H implies that its conjugate "action variable" is a constant of the motion, and this eliminates an additional variable from consideration. In <u>slowly</u> perturbed systems, such constants are called <u>adiabatic invariants</u>[1]. In <u>slightly perturbed</u> systems, no generally accepted name exists (G. Contopoulos, who investigated the relation between the two types of constants[2] has termed them "third integrals") but they are well-known in celestial mechanics and may be derived in a variety of ways.

The purpose of this work is to show how the standard methods of celestial mechanics, designed to handle small perturbations, may be modified to deal with slow perturbation and lead to the derivation of adiabatic invariants. Two methods will be considered here: the Poincaré-Von Zeipel method[(3)-(7)] for solving the Hamilton-Jacobi equation and the Krylov-Bogoliubov procedure[(7)-(12)] (or the related method of Kruskal). In addition, it will be shown that the direct form of near-identity canonical transformations can also be adapted to cases in which some variables are slow.

EXPLICIT AND IMPLICIT ε

In the example of perturbed planetary motion the small parameter ε can be given an explicit numerical value. In problems of slowly perturbed motion this is often difficult to do and one may then speak of an implicit ε.

As the archtype of a slowly perturbed system, consider the "pulled-up pendulum."[(13)(14)]: a simple pendulum is suspended from a hole in the ceiling and its suspension string is pulled up (or released) at a slow, though not necessarily constant rate. Obviously, the angular frequency ω of the pendulum will vary and, since work is being done against the centrifugal force of the oscillation, so will its energy E. However, as long as the rate at which the string is withdrawn is sufficiently slow (and does not resonate with the oscillation of the pendulum) an adiabatic invariant may be found, equaling E/ω in the lowest order.

Two points should be noted here. First, the perturbation need not be small: by the time the withdrawal is complete, the length of the pendulum may well have changed by a large factor. Secondly, while one can devise an explicit ε for the problem — e.g. $\varepsilon = \omega\tau$, where τ is the time in which the length of the pendulum is reduced to 1/e of its value, at the given (instantaneous) rate -- its value nowhere enters the calculation.

A more complicated example is provided by the motion of a charged particle in a slightly inhomogeneous magnetic field $\underline{B}$. Here "slightly" means that the derivatives $\partial B_i/\partial x_j$ are all of order ε smaller than the components of the field intensity and its magnitude B . Thus the slowness is in the dependence on spatial coordinates and a scale-length for gauging it is provided by the gyration radius ρ , giving, in analogy to eq. (4)

$$\partial B_i/\partial x_j = O(\varepsilon B/\rho) \tag{5}$$

Again, the value of ε does not explicitely enter, except through the requirement that for the perturbation approach (known as the guiding center theory) to be valid the problem must satisfy "Alfven's criterion"

$$(\rho/B)(\partial B_i/\partial x_j) \ll 1$$

An implicit ε may be "made visible" by the following device. Consider a Hamiltonian with slow time dependence: one may artificially introduce ε into its time derivative by writing

$$\partial H/\partial t = \varepsilon\, \partial H/\partial(\varepsilon t) \tag{6}$$

Since

$$\partial H/\partial \varepsilon t = O(1)$$

this notation clearly displays the fact that the term is of order ε , and for this reason the Hamiltonian (2) is often written

$$H = H(\underline{p}, \underline{q}, \varepsilon t)$$

A similar device may be used when there exists a slow dependence on dynamical variables; this can be quite useful in arranging the terms according to their orders in ε, but two things must be remembered. First, because of the way in which ε is introduced, expressions of the k-th order which have a factor ε^k standing in front of them, will also have "hidden inside" a factor ε^{-k}. Secondly, because a definite value of ε is never stated, such factors must be cancelled out before the final result is obtained.

An example may be useful here. Suppose a one-dimensional motion is given with a Hamiltonian that has a slow dependence on t, and it is also given that if this dependence is "frozen" (limit $\varepsilon = 0$) the motion is periodic. The solution of such a motion usually begins with a canonical transformation to new variables $(\underline{P}, \underline{Q})$ which are the action-angle variables of the unperturbed motion. If S is the generating function of this transformation, which in general is also slowly dependent on t, then the new Hamiltonian H' is

$$\begin{aligned} H'(\underline{P}, \underline{Q}) &= H + \partial S/\partial t \\ &= H + \varepsilon\, \partial S/\partial(\varepsilon t) \qquad (7) \\ &= H'^{(0)} + \varepsilon H'^{(1)} \end{aligned}$$

In the transformed Hamiltonian, the first order correction $H'^{(1)}$ has a factor ε preceding it, but this factor is artificial and is balanced by a factor ε^{-1} that is "hidden inside" the term, as is evident from the derivation. In practice, these factors must be cancelled before, say, the canonical equations of motion are used.

THE POINCARÉ - VON ZEIPEL METHOD FOR SLOW TIME DEPENDENCE

Consider a canonical system with $2N$ variables which has a slow dependence on time. We assume that the Hamiltonian H may be expanded in powers of ε

$$H = \sum \varepsilon^k H^{(k)}(\underline{p}, \underline{q}, t) \qquad (8)$$

We further assume that the Hamilton-Jacobi equation for $\varepsilon \to 0$ has been solved and that the transformation derived by it has already been applied, deriving as action-angle variables for the unperturbed motion

$$(J, \Omega) = (p_1, q_1)$$

and giving

$$H^{(0)} = J\omega/2\pi \qquad (9)$$

with $\omega = \omega(t)$ the slowly varying angular velocity. In the limit $\varepsilon \to 0$, evidently, ω is a constant and so are all the canonical variables, except for Ω which is then linear in time.

To "solve" the motion we now seek a near-identity canonical transformation to new variables $(\underline{P}, \underline{Q})$, with

$$(J^*, \Omega^*) = (P_1, Q_1)$$

generated by

$$\sigma(\underline{P}, \underline{q}, t) = \sum P_i q_i + \sum \varepsilon^k \sigma^{(k)}(\underline{P}, \underline{q}, t) \qquad (10)$$

such that the new Hamiltonian H^* does not depend on Ω^*. This is somewhat similar to, but simpler than, an approach advocated by Gardner[15] and investigated by Contopoulos[2], in which the same result is obtained by a succession of canonical transformations, each of which pushes the elimination of Ω^* from H^* one order higher.

If H^* is expanded in a manner similar to (8) and the time derivative is expressed as in (7), one obtains

$$\sum \varepsilon^k H^{*(k)}(\underline{P}, \underline{Q}, t) = \sum \varepsilon^k H^{(k)}(\underline{p}, \underline{q}, t) + \\ + \sum \varepsilon^k (\partial \sigma^{(k-1)} / \partial \varepsilon t) \qquad (11)$$

This equation contains 4N canonical variables, but half of them can be eliminated by means of the transformation equations

$$p_i = P_i + \sum \varepsilon^m (\partial \sigma^{(m)} / \partial q_i) \qquad (12)$$

$$Q_i = q_i + \sum \varepsilon^m (\partial \sigma^{(m)} / \partial P_i) \qquad (13)$$

To facilitate the elimination it is best to follow a method introduced by Musen[7] and use expansion operators[12] ($*$ denotes operation, $\partial/\partial \underline{P}$ etc. are gradient-type operators):

$$\begin{aligned} H^{*(k)}(\underline{P}, \underline{Q}, t) &= H^{*(k)}(\underline{P}, \underline{q} + \sum \varepsilon^m \partial \sigma^{(m)} / \partial \underline{P}, t) \\ &= \exp\left\{ \sum_{m=1} \varepsilon^m (\partial \sigma^{(m)} / \partial \underline{P}) \cdot (\partial / \partial \underline{q}) \right\} * H^{*(k)}(\underline{P}, \underline{q}, t) \\ &= \sum_{m=0} \varepsilon^m T^{(m)} * H^{*(k)}(\underline{P}, \underline{q}, t) \end{aligned} \qquad (14)$$

where

$$
\begin{aligned}
T^{(0)} &= 1 \\
T^{(1)} &= \sum_i (\partial\sigma^{(1)}/\partial P_i)(\partial/\partial q_i) \\
T^{(2)} &= \sum_i (\partial\sigma^{(2)}/\partial P_i)(\partial/\partial q_i) + \\
&\quad + \tfrac{1}{2} \sum_{i,\,j} (\partial\sigma^{(1)}/\partial P_i)(\partial\sigma^{(1)}/\partial P_j)(\partial^2/\partial q_i \partial q_j)
\end{aligned}
\tag{15}
$$

etc. Similarly

$$
H^{(k)}(\underline{p},\ \underline{q},\ t) = \sum_{m=0} \varepsilon^m S^{(m)} * H^{(k)}(\underline{P},\ \underline{q},\ t) \tag{16}
$$

where

$$
\begin{aligned}
S^{(0)} &= 1 \\
S^{(1)} &= \sum_i (\partial\sigma^{(1)}/\partial \underline{q_i})(\partial/\partial P_i) \\
S^{(2)} &= \sum_i (\partial\sigma^{(2)}/\partial q_i)(\partial/\partial P_i) + \\
&\quad + \tfrac{1}{2} \sum_{i,j} (\partial\sigma^{(1)}/\partial q_i)(\partial\sigma^{(1)}/\partial q_j)(\partial^2/\partial P_i \partial P_j)
\end{aligned}
\tag{17}
$$

and so forth. Substituting all this in (11) and collecting terms associated with ε^k gives

$$
\sum_{m=0}^{k} T^{(m)} * H^{*(k-m)} = \sum_{m=0}^{k} S^{(m)} * H^{(k-m)} + \partial\sigma^{(k-1)}/\partial(\varepsilon t) \tag{18}
$$

The terms with $m=0$ simply equal $H^{*(k)}$ and $H^{(k)}$ and will be taken outside the summation. The terms with $m=k$ also have simple form, for in general

$$S^{(k)} = \sum_i (\partial\sigma^{(k)}/\partial q_i)(\partial/\partial P_i) + N^{(k)} \qquad (19)$$

where $N^{(k)}$ contains only terms with at least two differentiations. Substituting (9) then gives

$$S^{(k)} * H^{(0)} = (\omega/2\pi)\,\partial\sigma^{(k)}/\partial\Omega \qquad (20)$$

Because the transformation reduces to the identity transformation in the limit of vanishing ε, $H^{*(0)}$ equals $H^{(0)}$ and due to (9) it satisfies

$$T^{(k)} * H^{*(0)} = 0$$

since $T^{(k)}$ operators involve only differentiation by the q_i, which $H^{(0)}$ does not contain. One then obtains the basic recursion relation

$$(\omega/2\pi)\,\partial\sigma^{(k)}/\partial\Omega - H^{*(k)}(\underline{P}, \underline{q}, t) = \Lambda^{(k)}(\underline{P}, \underline{q}, t) \qquad (21)$$

with

$$\Lambda^{(k)} = \sum_{m=1}^{k-1}\left\{T^{(m)} * H^{*(k-m)} - S^{(m)} * H^{(k-m)}\right\} - H^{(k)} - \partial\sigma^{(k-1)}/\partial\varepsilon t \qquad (22)$$

depending only on orders lower than the k-th. If Ω enters only as an angle variable with period unity, any function $F(\underline{P}, \underline{q}, t)$ may be resolved into an "averaged" part

$$\langle F \rangle = \int_0^1 F \, d\Omega \qquad (23)$$

and a "purely periodic" part with zero average

$$(F)_{per} = F - \langle F \rangle$$

The derivative of a purely periodic function is also purely periodic and therefore, so is

$$\partial\sigma^{(k)}/\partial\Omega = \partial/\partial\Omega \left\{ (\sigma^{(k)})_{per} + \langle \sigma^{(k)} \rangle \right\} \qquad (24)$$

since the contribution of $\langle \sigma^{(k)} \rangle$ vanishes. On the other hand, $H^{*(k)}$ does not depend on Ω, so one gets the recursive relations

$$H^{*(k)} = - \langle \Lambda^{(k)} \rangle \qquad (25)$$

$$\partial\sigma^{(k)}/\partial\Omega = (2\pi/\omega) \; (\Lambda^{(k)})_{per} \qquad (26)$$

Once these are solved, the calculation may be advanced to the next order.

EXAMPLE : THE HARMONIC OSCILLATOR (16)

The Hamiltonian H" of a harmonic oscillator with a slowly time-dependent angular velocity $\omega(t)$ is

$$H'' = (1/2m) \left[P^2 + \omega^2 m^2 Q^2 \right] \qquad (27)$$

If one "freezes" the time dependence, one can solve the Hamilton-Jacobi equation and derive a canonical transformation to action-angle variables (J, Ω), generated by

$$W = \int \left\{ (J\omega m/\pi) - m^2 \omega^2 Q^2 \right\}^{\frac{1}{2}} dQ \qquad (28)$$

Following this transformation, the new Hamiltonian H becomes

$$H = J\omega/2\pi + \varepsilon J\,(\omega'/4\pi\omega)\sin(4\pi\Omega) \qquad (29)$$

where the dash henceforth signifies the operation $\partial/\partial(\varepsilon t)$. Let σ of (10) generate a transformation to (J^*, Ω^*) such that all orders $H^{*(k)}$ of the new Hamiltonian are independent of Ω^* . This, combined with the fact that in the present case the only differentiation performed by $T^{(m)}$ of (15) is $\partial/\partial\Omega$, allows all such operators to be ignored __except__ for $T^{(0)}$.

A further simplification is obtained by noting that H contains only two orders, both linear in J : using the argument of (19) this gives, for the terms of (18) depending on H

$$\sum S^{(m)}{}_* H^{(k-m)} = S^{(k)}{}_* H^{(0)} + S^{(k-1)}{}_* H^{(1)}$$

$$= (\omega/2\pi)\,\partial\sigma^{(k)}/\partial\Omega + (\partial\sigma^{(k-1)}/\partial\Omega)(\omega'/4\pi\omega)\sin(4\pi\Omega) \qquad (30)$$

In what follows, we will for conciseness __write__ J instead of J^* , restoring the superscript — if necessary — only at the end. In analogy with (21) we then obtain as the basic recursion relation, for $k > 1$

$$(\omega/2\pi)(\partial\sigma^{(k)}/\partial\Omega) - H^{*(k)} =$$

$$= -(\partial\sigma^{(k-1)}/\partial\Omega)(\omega'/4\pi\omega)\sin(4\pi\Omega) - (\sigma^{(k-1)})' \qquad (31)$$

Using (18) directly for k=0, one simply gets the equality of $H^{(0)}$ and $H^{*(0)}$, while for k=1 this yields

$$(\omega/2\pi)\ \partial\sigma^{(1)}/\partial\Omega \ - \ H^{*(1)} \ = \ - J\ (\omega'/4\pi\omega)\ \sin\ (4\pi\Omega) \qquad (32)$$

from which

$$H^{*(1)} \ = \ 0 \qquad (33)$$

$$\sigma^{(1)} \ = \ J\ (\omega'/8\pi\omega^2)\ \cos\ (4\pi\Omega) \qquad (34)$$

Higher orders, derived by the use of (31), are

$$\left.\begin{aligned} H^{*(2)} &= - (J/16\pi)\ (\omega')^2/\omega^3 \\ \sigma^{(2)} &= - (J/64\pi)\ (\omega'/\omega^2)^2 \sin\ (8\pi\Omega)\ - \\ &\quad - (J/16\pi\omega)\ (\omega'/\omega^2)'\ \sin\ (4\pi\Omega) \end{aligned}\right\} (35)$$

$$H^{*(3)} \ = \ 0$$

$$\sigma^{(3)} \ = \ - (J/384\pi)(\omega'/\omega^2)^3 \cos\ (12\pi\Omega)$$

$$- (J/128\pi\omega) \left[(\omega'/\omega^2)^2\right]' \cos\ (8\pi\Omega) \ + \ (J/128\pi)(\omega'/\omega^2)^3 \cos\ (4\pi\Omega)$$

$$- (J/32\pi\omega) \left[(\omega'/\omega^2)'/\omega\right]' \cos\ (4\pi\Omega) \qquad (36)$$

Note that any term in an $O(\varepsilon^k)$ expression contains the dash operator exactly k times, corresponding to the factor ε^{-k} "hidden inside."

At this stage eqs. (12) and (13) could be used to express (J, Ω^*) in terms of (J^*, Ω), up to $O(\varepsilon^3)$. In fact, expressing the result in this manner, in terms of "mixed variables", is not too useful, and it pays to "invert" the result and express (J^*, Ω^*) in terms of (J, Ω), or vice versa. The shortest way to achieve this is by means of the <u>direct transformation</u> technique[17]. If

$$\underline{y} \equiv (\underline{p}, \underline{q})$$

are the "old" variables and

$$\underline{z} \equiv (\underline{P}, \underline{Q})$$

are the "new" ones, and if the relation between the two sets has the "direct" form

$$\underline{z} = \underline{y} + \sum_{k=1} \varepsilon^k \underline{\zeta}^{(k)}(\underline{y}) \tag{37}$$

then for this to be a <u>canonical</u> transformation, $\underline{\zeta}^{(k)}$ must have the form

$$\underline{\zeta}^{(k)} = \bar{\nabla}\chi^{(k)} + \underline{f}^{(k)} \tag{38}$$

where $\bar{\nabla}$ is a gradient operator in "conjugate phase space"

$$\bar{\underline{y}} \equiv (\underline{q}, -\underline{p})$$

the $\chi^{(k)}$ are arbitrary functions and $\underline{f}^{(k)}$ are prescribed expressions involving lower orders. In particular, if (37) is the "direct" form of the transformation generated by (10), one may choose

$$\chi^{(k)}(\underline{y}) = -\sigma^{(k)}(\underline{p}, \underline{q}) = -\sigma^{(k)}(\underline{y}) \tag{39}$$

(i.e. $\underline{P}$ is everywhere replaced by $\underline{p}$). The corresponding $\underline{f}^{(k)}$ is

$$\underline{f}^{(k)} = -\sum_{m=1}^{k-1} U^{(m)}{}_{*}\, \bar{\nabla}\sigma^{(k-m)}(\underline{y}) \tag{40}$$

with $U^{(m)}$ expansion operators depending only on the momentum-like components $\underline{\pi}^{(m)}$ of $\underline{\zeta}^{(m)}$

$$\underline{\pi}^{(m)} = (\zeta_1^{(m)}, \dots \zeta_N^{(m)}, 0, \dots 0)$$

with

$$\begin{aligned} U^{(1)} &= \underline{\pi}^{(1)} \cdot \nabla \\ U^{(2)} &= \underline{\pi}^{(2)} \cdot \nabla + \tfrac{1}{2} \underline{\pi}^{(1)} \underline{\pi}^{(1)} : \nabla\nabla \end{aligned} \tag{41}$$

and so on.

Of particular interest is the derivation of the adiabatic invariant

$$z_1 = J^* = \sum \varepsilon^k J^{*(k)} \tag{42}$$

which will now be outlined.

To obtain $\sigma^{(k)}(\underline{y})$ one simply uses the expressions (34)-(36) <u>without</u> restoring the asterisk superscript (as was originally planned). To derive (41), note that only one component of canonical momentum enters the calculation, so that

$$\begin{aligned} U^{(1)} &= J^{*(1)} (\partial/\partial J) \\ U^{(2)} &= J^{*(2)} (\partial/\partial J) + \tfrac{1}{2} (J^{*(1)})^2 \partial^2/\partial J^2 \end{aligned} \tag{43}$$

The second-derivative terms may be safely ignored, since all orders of $\sigma^{(k)}$ used here are found to be linear in J. Finally, the components of the conjugate gradient $\bar{\nabla}$ contributing to z_1 are simply

$$\partial\sigma^{(k)}/\partial\bar{y}_1 = \partial\sigma^{(k)}/\partial\Omega$$

For the first order, $\underline{f}^{(1)}$ vanishes and one obtains

$$J^{*(1)} = - \partial\sigma^{(1)}/\partial\Omega$$

$$= J\,(\omega'/2\,\omega^2)\,\sin\,(4\pi\Omega) \qquad (44)$$

The next terms are

$$J^{*(2)} = (J/8)\,(\omega'/\omega^2)^2 + (J/4\omega)\,(\omega'/\omega^2)'\,\cos\,(4\pi\Omega)$$

$$(45)$$

$$J^{*(3)} = (J/16)\,(\omega'/\omega^2)^3\,\sin\,(4\pi\Omega) - (J/8\omega)\,\left[(\omega'/\omega^2)'/\omega\right]'\sin(4\pi\Omega)$$

THE "OLD" NOTION OF ADIABATIC INVARIANCE

In some texts of mechanics[5] and in the older literature, the definition of adiabatic invariance differs somewhat from the one given here. The alternative definition is usually applied to one-dimensional systems (though generalizations for several dimensions exist) and is as follows:

> "Given a slowly perturbed periodic motion, consider the action integral
>
> $$J = \oint p\,dq \qquad (46)$$
>
> evaluated over one period of the unperturbed system. As the system is perturbed, an "instantaneous" J may be evaluated at any time by "freezing" slowly varying quantities. Then J has the property of adiabatic invariance: if the system undergoes a finite perturbation -- e.g., a finite change of the Hamiltonian from H_1 to H_2 -- the corresponding change in J may be made arbitrarily small by stretching out the perturbation over a sufficiently long time."

The action variable J of (46) is the same as the zero-order action variable with which the previously-developed perturbation scheme begins, but its "adiabatic invariance" differs in two respects from what was earlier defined as adiabatic invariance. First, there exists here no hierarchy of invariants each of which is conserved to some specified order and secondly, the definition concerns itself with the cumulative change in J over a long period in time. In fact, this property does not follow automatically from the definition of adiabatic invariance used earlier. It is nevertheless an extremely useful property, since it allows one to derive, using only the unperturbed variables, a quantity with long-term invariance properties, without even specifying the perturbation.

Since J is the zero-order part of J^*, we may use (12) to obtain (compare also eq. 44)

$$\begin{aligned} J^* &= J - \varepsilon\, \partial\sigma^{(1)}/\partial\Omega + O(\varepsilon^2) \\ &= J + \varepsilon J^{*(1)} + O(\varepsilon^2) \end{aligned} \tag{47}$$

As in (12), $\sigma^{(1)}$ means $\sigma^{(1)}(J^*, \Omega, t)$; since J^* is a constant of the motion, only Ω and the slow direct dependence on t contribute to the variation of the first order correction $J^{*(1)}$. The basic reason for the "long-term adiabatic invariance" of J, stated earlier, is that by the arguments of eq. (24) $J^{*(1)}$ is purely periodic in Ω, and therefore "nearly" purely periodic in t. Over long time intervals, its variation is therefore bounded, causing the long-term conservation of J to be better than might otherwise be expected.

To demonstrate this, expand (47) to

$$
\begin{aligned}
J &= J^{*} - \varepsilon J^{*(1)}(J^{*},\Omega,\ t) + O(\varepsilon^{2}) \\
&= J^{*} - \varepsilon J^{*(1)}(J^{*},\Omega,\ 0) - \varepsilon^{2} t\ \partial J^{(1)}/\partial(\varepsilon t) + \ldots + O(\varepsilon^{2})
\end{aligned}
\tag{48}
$$

Let a time $T = O(\varepsilon^{-1})$ pass. The first term on the right is conserved, while the second one will vary only through the variation of Ω. Since the dependence of this term on Ω is periodic, the resulting contribution is bounded and due to the factor preceding it, of order ε. The next term is also $O(\varepsilon)$ and the same holds for higher terms in the expansion of the slow direct time dependence of $J^{*(1)}$. The $O(\varepsilon^{2})$ terms may contribute to dJ/dt a term of form $\varepsilon^{2}\psi$, but its contribution to the total change of J will again be of order ε:

$$\varepsilon^{2}\int_{0}^{T}\psi\ dt = O(\varepsilon^{2}T) = O(\varepsilon)$$

Hence the long-term variation of J is $O(\varepsilon)$.

The variation of other dynamical quantities, on the other hand, will be finite. For instance, for H

$$
\begin{aligned}
\Delta H &= \int_{0}^{T}(dH/dt)\ dt = \int_{0}^{T}(\partial H/\partial t)\ dt \\
&= T\,(\partial H/\partial t)_{\text{aver.}} = (\varepsilon T)\,(\partial H/\partial(\varepsilon t))_{\text{aver}}
\end{aligned}
\tag{49}
$$

and each factor here is $O(1)$. Thus by making ε arbitrarily small, but keeping $T = O(\varepsilon^{-1})$, the variation of J may be made as small as is desired while that of H remains finite.

THE POINCARÉ - VON ZEIPEL METHOD FOR SLOW DEPENDENCE ON CANONICAL VARIABLES

Let a perturbed periodic motion be given, represented by a Hamiltonian

$$H = \sum \varepsilon^k H^{(k)}(\underline{p}, \underline{q}) \qquad (50)$$

with (p_1, q_1) the action-angle variables (J, Ω) of the unperturbed motion; since we have already derived methods dealing with slow time dependence, we will simplify matters by not including such a dependence here. The motion represented by $H^{(0)}$ alone is assumed to be periodic and soluble: we shall not require at this stage that $H^{(0)}$ has the form (9), but we note that it must be independent of Ω , since J is a constant of the unperturbed motion.

Instead, we shall assume that the canonical variables y_i fall into two groups: "normal" variables for which $\partial/\partial y_i$ maintains the same order in ε and "slow" ones for which it raises the order by one level. It is useful to define parameters that distinguish between the two groups: let γ_i equal 0 or 1 depending on whether q_i is normal or slow, and let δ_i play the same role for p_i . One can then define

$$\mathbb{Q}_i = \varepsilon^{\gamma_i} q_i \qquad (51)$$

$$\mathbb{P}_i = \varepsilon^{\delta_i} p_i \qquad (52)$$

so that (for example) $\partial H/\partial \mathbb{Q}_i$ and $\partial H/\partial \mathbb{P}_i$ are always of the same order as H itself.

As before, let a generating function

$$\sigma(\underline{P}, \underline{q}) = \sum_i P_i q_i + \sum_{k=1} \varepsilon^k \sigma^{(k)}(\underline{P}, \underline{q}) \qquad (53)$$

define a near-identity transformation to a new canonical set $(\underline{P}, \underline{Q})$, with the new Hamiltonian H^* independent of the transformed angle variable Ω^* (a term with $k = 0$ could be included, but since it may not depend on Ω^* it is not useful here). Again, the basic equation is

$$H^*(\underline{P}, \underline{Q}) = H(\underline{p}, \underline{q}) \tag{54}$$

and this again is expressed in powers of ε and expressed solely in terms of $(\underline{P}, \underline{q})$. Since $\underline{p}$ no longer appears, it helps to **redefine** $\mathbb{P}_i$ as

$$\mathbb{P}_i = \varepsilon^{\delta_i} P_i \tag{55}$$

and this will be the definition used in the remainder of this section. In analogy with (14) one finds

$$H^{*(k)}(\underline{P}, \underline{Q}) = \exp\left\{\sum_{m=1} \varepsilon^m \sum_i (\partial\sigma^{(m)}/\partial P_i)(\partial/\partial q_i)\right\} * H^{*(k)}(\underline{P}, \underline{q})$$

$$= \exp\left\{\sum_{i,m} \varepsilon^{m+\delta_i+\gamma_i} (\partial\sigma^{(m)}/\partial\mathbb{P}_i)(\partial/\partial\mathbb{Q}_i)\right\} * H^{*(k)}(\underline{P}, \underline{q})$$

$$= \exp \sum_m \varepsilon^m \sum_i (\partial\sigma^{(m-\delta_i-\gamma_i)}/\partial\mathbb{P}_i)(\partial/\partial\mathbb{Q}_i) * H^{*(k)}(\underline{P}, \underline{q})$$

$$= \sum_{m=0} \varepsilon^m v^{(m)} * H^{*(k)} \tag{56}$$

with $v^{(m)}$ suitable operators and $\sigma^{(m)}$ vanishing for all non-positive values of m. Expanding the exponential gives

$$V^{(0)} = 1$$

$$V^{(1)} = \sum_i (\partial\sigma^{(1-\delta_i-\gamma_i)}/\partial\mathbb{P}_i)(\partial/\partial\mathbb{Q}_i) \tag{57}$$

and so forth; because these operators are expressed solely in terms of $\mathbb{P}_i$ and $\mathbb{Q}_i$, their action on any function maintains the ordering in powers of $\mathcal{E}$.

Similarly

$$H^{(k)}(\underline{p}, \underline{q}) = \sum_{m=0} \varepsilon^m R^{(m)} * H^{(k)}(\underline{P}, \underline{q}) \tag{58}$$

with

$$R^{(0)} = 1$$

$$R^{(1)} = \sum_i (\partial\sigma^{(1-\delta_i-\gamma_i)}/\partial\mathbb{Q}_i)(\partial/\partial\mathbb{P}_i) \tag{59}$$

and so forth. Substituting these operators and collecting terms associated with ε^k then gives, in analogy with (18)

$$\sum_{m=0}^{k} \left\{ V^{(m)}{}_* H^{*(k-m)} - R^{(m)}{}_* H^{(k-m)} \right\} = 0 \tag{60}$$

Again, the terms with $m=0$ and $m=k$ are separated. For the latter terms one gets, in analogy with (19)

$$V^{(k)} = \sum_i (\partial\sigma^{(k-\delta_i-\delta_i)}/\partial \mathbb{P}_i)(\partial/\partial \mathbb{Q}_i) + M^{(k)}$$
(61)
$$R^{(k)} = \sum_i (\partial\sigma^{(k-\delta_i-\delta_i)}/\partial \mathbb{Q}_i)(\partial/\partial \mathbb{P}_i) + N^{(k)}$$

with $M^{(k)}$ and $N^{(k)}$ involving only lower orders. By taking $k=0$ in (60) one again finds

$$H^{*(0)} = H^{(0)} \tag{62}$$

so that (60) becomes, for the general case

$$H^{*(k)} + \sum_i \Big\{ (\partial\sigma^{(k-\delta_i-\delta_i)}/\partial \mathbb{P}_i)(\partial H^{(0)}/\partial \mathbb{Q}_i) - $$
(63)
$$- (\partial\sigma^{(k-\delta_i-\delta_i)}/\partial \mathbb{Q}_i)(\partial H^{(0)}/\partial \mathbb{P}_i) \Big\} = G^{(k)}$$

where

$$G^{(k)} = H^{(k)} + \sum_{m=1}^{k-1} \Big\{ R^{(m)}{}_* H^{(k-m)} - V^{(m)}{}_* H^{*(k-m)} \Big\} + (M^{(k)} - N^{(k)})_* H^{(0)} \tag{64}$$

involves only given functions and lower orders. For every k a relation of this type is obtained, constituting a k-th order recursion formula for the derivation of $H^{*(k)}$ and $\sigma^{(k)}$.

Now the action-angle variables associated with the zero-order periodicity (and used here in "mixed" form)

$$(J^*, \Omega) \equiv (P_1, q_1)$$

are assumed to be "normal", so that the left-hand side of (63) will include a term

$$(\partial \sigma^{(k)}/\partial \Omega)(\partial H^{(0)}/\partial J^*)$$

If $\sigma^{(k)}$ appears <u>nowhere else</u>, the equation assumes the form of (21) and is solved in the same manner.

On the other hand, if $\sigma^{(k)}$ appears anywhere else in (63), it may not be possible to derive it, for the equation then becomes a partial differential equation for $\sigma^{(k)}$. To prevent this from happening, it is required that for all (p_i, q_i) appearing in $H^{(0)}$ other than the action-angle pair

$$\delta_i + \gamma_i \geqslant 1 \tag{65}$$

Hence the recursion can be carried out <u>if</u>:

(1) $H^{(0)}$ has normal dependence on p_1 but does not depend on q_1

(2) $H^{(0)}$ may depend on any "slow" variable.

(3) $H^{(0)}$ may depend on any "normal" variable, provided its canonical conjugate is "slow".

Furthermore, it may be shown by extending the present calculation that

(4) H may include a term $H^{(-1)}$ of order ε^{-1}, provided it depends only on slow variables having slow conjugates. Such terms are then transformed intact to the new Hamiltonian.

As an example, the Hamiltonian of a charged particle in a time-independent electromagnetic field, in the regime of guiding-center motion, may be brought to the form (15)(18)

$$H \;=\; p_2{}^2/2m \;+\; p_1\omega/2\pi \;+\; \varepsilon^{-1}e\,\phi^{(0)} \;+\; O(\varepsilon) \qquad (66)$$

Here (p_1, q_1) are canonical variables associated with the rapid gyration, (p_2, q_2) represent the motion along field lines and (p_3, q_3) describe the identity of the guiding field line, which changes slowly with time (in the references, subscripts 1 and 3 have reversed meanings); the variables (p_1, q_1, p_2) are "normal" whereas the remaining ones are "slow". Furthermore, the gyration frequency ω and the lowest order $\phi^{(0)}$ of the electric potential are both functions of slow variables only. The last term is of order ε^{-1}, since its derivatives are proportional to the components of the lowest order of the electric field $\underline{E}$, which are of order unity.

Evidently H meets all the previously stated conditions except for one: if $\phi^{(0)}$ contains q_2, condition (4) is violated, since p_2 is not slow. One therefore must impose an additional requirement that $\partial\phi^{(0)}/\partial q_2$ vanishes: this reduces to the well-known restriction in guiding center theory that the electric field may have no zero-order component parallel to the magnetic field.

DIRECT CANONICAL TRANSFORMATIONS WITH SLOW VARIABLES

The generating function $\sigma^{(k)}$ gives the transformation equations as in (12)-(13), in "mixed" form. To bring them to the "direct" form (37) it is useful to generalize (38) for cases in which slow variables are present, and this will now be done.

Let

$$\underline{y} = (\underline{p}, \underline{q})$$

be a canonical set and

$$\underline{\bar{y}} = (\underline{q}, -\underline{p})$$

be its conjugate[17]. One may now define an index Γ_i equaling 0 or 1 depending on whether y_i is "normal" or "slow" , and an index Δ_i which has a similar relation to $\bar{y}_i$. With this notation it is possible to define vectors $\underline{Y}$ and $\underline{\bar{Y}}$ satisfying relations similar to (51) and (52)

$$Y_i = \varepsilon^{\Gamma_i} y_i \tag{67}$$

$$\bar{Y}_i = \varepsilon^{\Delta_i} \bar{y}_i \tag{68}$$

As with quantities defined in (51) and (52), $\partial/\partial Y_i$ and $\partial/\partial \bar{Y}_i$ are always $O(1)$.

We now seek the condition for a near-identity transformation (37) to be canonical. Actually, in what follows the recursion may still be carried out even if the transformation is not a near-identity one and $\zeta_i^{(0)}$ terms are included (satisfying appropriate conditions) but we shall not develop this possibility here. One then finds, as a condition for canonical behavior

$$[y_i, y_j] = [z_i, z_j]$$

$$= \left[y_i + \sum \varepsilon^k \zeta_i^{(k)} ,\; y_j + \sum \varepsilon^m \zeta_j^{(m)} \right] \qquad (69)$$

$$= [y_i, y_j] + \sum \varepsilon^k \left\{ [\zeta_i^{(k)}, y_j] - [\zeta_j^{(k)}, y_i] + \sum_{m=1}^{k-1} [\zeta_i^{(m)}, \zeta_j^{(k-m)}] \right\}$$

Expressing derivatives in terms of $\underline{Y}$ and $\overline{\underline{Y}}$ gives

$$[a, b] = \sum_s (\partial a/\partial \bar{y}_s)(\partial b/\partial y_s)$$

$$= \sum_s \varepsilon^{\Delta_s + \Gamma_s} (\partial a/\partial \bar{Y}_s)(\partial b/\partial Y_s) \qquad (70)$$

In particular

$$[a, y_i] = \varepsilon^{\Delta_i} (\partial a/\partial \bar{Y}_i) \qquad (71)$$

Thus

$$0 = \sum_{k=1} \left\{ \varepsilon^{k+\Delta_j} (\partial \zeta_i^{(k)}/\partial \bar{Y}_j) - \varepsilon^{k+\Delta_i} (\partial \zeta_j^{(k)}/\partial \bar{Y}_i) \right.$$

$$\left. + \sum_s \varepsilon^{k+\Delta_s+\Gamma_s} \sum_{m=1}^{k-1} (\partial \zeta_i^{(m)}/\partial \bar{Y}_s)(\partial \zeta_j^{(k-m)}/\partial Y_s) \right\} \qquad (72)$$

Dividing by $\varepsilon^{\Delta_i + \Delta_j}$

$$0 = \sum_{k=\Delta_i+\Delta_j+1} \Big\{ \varepsilon^{k-\Delta_i} (\partial \zeta_i^{(k)} / \partial \bar{Y}_j) - \varepsilon^{k-\Delta_j} (\partial \zeta_j^{(k)} / \partial \bar{Y}_i) +$$

$$+ \sum_s \varepsilon^{k+\Gamma_s+\Delta_s-\Delta_i-\Delta_j} \sum_{m=1}^{k-1} (\partial \zeta_i^{(m)} / \partial \bar{Y}_s)(\partial \zeta_j^{(k-m)} / \partial Y_s) \Big\} \tag{73}$$

It is useful at this point to redefine k for each term so that all powers of ε become ε^k and also to replace m by

$$M = m - \Delta_i \tag{74}$$

Because the exponent of ε differs for each term in (73), the new summation over k will begin at a different value for each term; this summation limit may, however, be uniformly set equal to 1 if it is assumed that $\zeta_i^{(u)}$ vanishes for non-positive values of u. With these changes, (73) gives

$$0 = \sum_{k=1} \varepsilon^k \Big\{ (\partial \zeta_i^{(k+\Delta_i)} / \partial \bar{Y}_j) - (\partial \zeta_j^{(k+\Delta_j)} / \partial \bar{Y}_i) +$$

$$+ \sum_{M=1-\Delta_i}^{k-\Delta_i-1} \sum_s (\partial \zeta_i^{(M+\Delta_i)} / \partial \bar{Y}_s)(\partial \zeta_j^{(k-M-\Gamma_s-\Delta_s+\Delta_j)} / \partial Y_s) \Big\} \tag{75}$$

This suggests the introduction of new "staggered" vectors

$$\underline{\mathfrak{z}}^{(k)} = \underline{\zeta}^{(k+\Delta_i)} \tag{76}$$

i.e.

$$\zeta_i^{(k)} = \zeta_i^{(k)} \quad \text{if} \quad \Delta_i = 0$$
$$\zeta_i^{(k)} = \zeta_i^{(k+1)} \quad \text{if} \quad \Delta_i = 1 \tag{77}$$

It is also useful (in analogy to what was done in ref. 17) to introduce a curl operator in $\underline{\bar{Y}}$ space. With this notation (75) may be rewritten as

$$(\bar{\nabla}_Y \times \underline{\zeta}^{(k)})_{ij} =$$
$$= \sum_s \sum_{M=1}^{k-1-\Gamma_s-\Delta_s+\Delta_j} (\partial \zeta_i^{(M)}/\partial \bar{Y}_s)(\partial \zeta_j^{(k-M-\Gamma_s-\Delta_s)}/\partial Y_s) \tag{78}$$

This equation may in principle be used to derive $\underline{\zeta}^{(k)}$ recursively, but this turns out to be a rather inconvenient approach. It is more useful in determining the degree of arbitrariness associated with a near-identity canonical transformation of the form (37). Let two such transformations be given, characterized by "staggered" vectors $\underline{\zeta}^{(m)}$ and $\underline{x}^{(m)}$ which are identical for orders up to and including the (k-1). For the k-th order one finds that the right-hand side of (78), which depends only on lower orders, is identical for both expansions, giving

$$\bar{\nabla}_Y \times (\underline{\zeta}^{(k)} - \underline{x}^{(k)}) = 0 \tag{79}$$

from which

$$\underline{\zeta}^{(k)} = \underline{x}^{(k)} + \bar{\nabla}_Y \chi^{(k)} \tag{80}$$

Thus the arbitrariness in specifying the canonical transformation at each level of $\underline{\zeta}^{(k)}$ is contained in the gradient in $\bar{Y}$ space of an arbitrary scalar $\chi^{(k)}$. The general form of $\underline{\zeta}^{(k)}$ for a canonical transformation may be written, in analogy with (38)

$$\underline{\zeta}^{(k)} = \bar{\nabla}_Y \chi^{(k)} + \underline{F}^{(k)} \tag{81}$$

where $\underline{F}^{(k)}$ is a vector involving orders of $\underline{\zeta}^{(m)}$ lower than the k-th and constitutes one particular solution of (78). In the following sections two such particular solutions will be derived, analogous to those found in ref. (17) for <u>small</u> perturbations.

DERIVATION BASED ON $\sigma(\underline{P}, \underline{q})$

Let a near-identity transformation of $n = 2N$ variables

$$\underline{y} = (\underline{p}, \underline{q}) \qquad \underline{z} = (\underline{P}, \underline{Q}) \tag{82}$$

be given by (37), and let a generating function (53) be assumed to produce the same transformations via the equations (12) and (13). In what follows the relation between (37) and (53) will be established in a way resembling what was done in ref. (17) for the case when no slow variables are present. As before, the calculation may be broadened somewhat beyond what is done here, since the method only requires that canonical <u>momenta</u> transform in near-identical fashion.

With the notation of (51) and (55) eqs. (12) and (13) give

$$
\begin{aligned}
P_i &= p_i - \sum \varepsilon^{k} (\partial \sigma^{(k)} / \partial q_i) \\
&= p_i - \sum \varepsilon^{k+\delta_i} (\partial \sigma^{(k)} / \quad _i) \qquad (83)
\end{aligned}
$$

$$
\begin{aligned}
Q_i &= q_i + \sum \varepsilon^{k} (\partial \sigma^{(k)} / \partial P_i) \\
&= q_i + \sum \varepsilon^{k+\delta_i} (\partial \sigma^{(k)} / \partial \mathbb{P}_i) \qquad (84)
\end{aligned}
$$

All functions on the right depend on mixed variables $(\underline{P}, \underline{q})$; to introduce a dependence on $\underline{y}$, it is useful to define "partial vectors" adding up to $\underline{\zeta}^{(k)}$

$$
\begin{aligned}
\underline{\pi}^{(k)} &= (\zeta_1^{(k)}, \ldots \zeta_{N-1}^{(k)}, 0, \ldots 0) \\
\underline{\theta}^{(k)} &= (0, \ldots 0, \zeta_N^{(k)}, \ldots \zeta_n^{(k)})
\end{aligned} \qquad (85)
$$

From this

$$
P_i = p_i + \sum \varepsilon^{k} \pi_i^{(k)}(\underline{y}) \qquad (86)
$$

$$
Q_i = q_i + \sum \varepsilon^{k} \theta_{N+i}^{(k)}(\underline{y}) \qquad (87)
$$

If the vectors $\underline{\pi}^{(k)}$ are known, they may also be used to expand any function of mixed variables $(\underline{P}, \underline{q})$ in terms of $\underline{y}$, e.g.

$$
\begin{aligned}
F(\underline{P}, \underline{q}) &= F(\underline{y} + \sum \varepsilon^k \underline{\pi}^{(k)}) \\
&= \exp\left\{\sum \varepsilon^k (\underline{\pi}^{(k)} \cdot \partial/\partial \underline{y})\right\} * F(\underline{y}) \\
&= \exp\left\{\sum \varepsilon^k \sum_s (\underline{\pi}^{(k-\Gamma_s)} \cdot \partial/\partial Y_s)\right\} * F(\underline{y}) \\
&= \sum_{k=0} \varepsilon^k L^{(k)} * F(\underline{y})
\end{aligned}
\tag{88}
$$

where, if we collectively denote all "slow" components of $\underline{y}$ by $\underline{R}$ and all "normal" ones by $\underline{r}$ and if $\underline{\pi}_R^{(k)}$ and $\underline{\pi}_r^{(k)}$ denote vectors composed of the corresponding components of $\underline{\pi}^{(k)}$,

$$
\begin{aligned}
L^{(0)} &= 1 \\
L^{(1)} &= \underline{\pi}_r^{(1)} \cdot \partial/\partial \underline{r} \\
L^{(2)} &= \underline{\pi}_r^{(2)} \cdot \partial/\partial \underline{r} + \underline{\pi}_R^{(1)} \cdot \partial/\partial \varepsilon \underline{R} \\
&\quad + \tfrac{1}{2} \underline{\pi}_r^{(1)} \underline{\pi}_r^{(1)} : (\partial/\partial \underline{r})(\partial/\partial \underline{r})
\end{aligned}
\tag{89}
$$

and so forth. Note that since ε is implicit, $\underline{\pi}^{(1)}$ and $\underline{\pi}^{(2)}$ should have factors ε^{-1} and ε^{-2} "hidden inside", since they are teamed up with the corresponding positive powers in (86). For the same reason $L^{(2)}$ should contain a factor ε^{-2} and indeed, inspection of the last equality in (89) shows that all terms there have such a factor.

Substitution in (83) yields

$$P_i = p_i - \sum \varepsilon^{k+\gamma_i} \sum_m \varepsilon^m L^{(m)} * (\partial\sigma^{(k)}/\partial Q_i)$$

$$= p_i - \sum \varepsilon^k \left\{ (\partial\sigma^{(k-\gamma_i)}/\partial Q_i) + \right.$$

$$\left. + \sum_{m=1}^{k-\gamma_i} L^{(m)} * (\partial\sigma^{(k-m-\gamma_i)}/\partial Q_i) \right\} \qquad (90)$$

where all terms of σ are viewed as functions of $\underline{y}$, i.e. with $\underline{p}$ replacing $\underline{P}$ wherever the latter originally appeared. This should be identical to (83) and therefore

$$\pi_i^{(k)} = -\partial\sigma^{(k-\gamma_i)}/\partial Q_i - \sum_{m=1}^{k-1} L^{(m)} * \partial\sigma^{(k-m-\gamma_i)}/\partial Q_i \qquad (91)$$

The highest order of $L^{(m)}$ appearing on the right is $k-1$ ($\sigma^{(0)}$ only appears if $\gamma_i = 1$, for if it depends on "normal" variables, the transformation is no longer one of near-identity: this is the reason for the change in summation limit) and this is therefore also the highest order of $\underline{\pi}^{(s)}$ appearing on the right. Thus (91) is a usable recursion relation for deriving $\underline{\pi}^{(k)}$.

Expanding (84) in a similar manner gives

$$\Theta_i^{(k)} = \partial\sigma^{(k-\gamma_i)}/\partial \mathbb{P}_i + \sum_{m=1}^{k-1} L^{(m)} * (\partial\sigma^{(k-m-\gamma_i)}/\partial \mathbb{P}_i) \qquad (92)$$

where the definition of $\mathbb{P}_i$ reverts to (52) .

Now if

$$y_i = q_j$$

then

$$\bar{Y}_i = -\mathbb{P}_j \quad ; \quad \Delta_i = \delta_j$$

and if

$$y_i = p_j$$

then

$$\bar{Y}_i = \mathbb{Q}_j \quad ; \quad \Delta_i = \gamma_j$$

Inspection then shows that (91) and (92) may be combined to one equation

$$z_i = y_i - \sum \varepsilon^k \left\{ \partial\sigma^{(k-\Delta_i)}/\partial \bar{Y}_i + \sum_{m=1}^{k-1} L^{(m)} * \partial\sigma^{(k-m-\Delta_i)}/\partial \bar{Y}_i \right\} \tag{93}$$

The dependence on Δ_i may be removed by introducing $\underline{\gamma}^{(k)}$ defined in (76). Then, using the gradient operator in $\underline{\bar{Y}}$ space, (93) becomes

$$\underline{\gamma}^{(k)} = -\bar{\nabla}_Y \sigma^{(k)} - \sum_{m=1}^{k-1} L^{(m)} * \bar{\nabla}_Y \sigma^{(k-m)} \tag{94}$$

Since it has already been established in (81) that $\underline{\gamma}^{(k)}$ is arbitrary within some gradient in $\underline{\bar{Y}}$ space, the summation term represents a particular solution of (78).

LIE TRANSFORMS WITH SLOW VARIABLES [16]

If L_W is the operator denoting Poisson bracketing with a function W of the canonical variables

$$L_W(f) = [f, W] \tag{95}$$

Then it may be shown[17][19][20][21] that the transformation from $y = (\underline{p}, \underline{q})$ to

$$\underline{z} = \exp(\varepsilon L_W) * \underline{y} \tag{96}$$

(with the exponential operator defined by its series expansion) is canonical. In what follows the form of the Lie transform in the presence of slow variables will be derived, again following closely the derivation for the simpler case when all y_i vary on the same scale[17]. Let W be expandable in ε

$$W = \sum \varepsilon^k W^{(k)}(\underline{y}) \tag{97}$$

and let operators $L_W^{(k)}$ be defined through

$$\varepsilon L_W = \sum \varepsilon^k L_W^{(k)} \tag{98}$$

If one defines (compare ref. 17, eq. 35)

$$\chi^{(k)} = - W^{(k-1)} \tag{99}$$

then

$$L_W = - \sum \varepsilon^{k+1} \sum_s (\partial W^{(k)}/\partial \bar{y}_s)(\partial/\partial y_s)$$

$$= \sum \varepsilon^k \sum_s (\partial \chi^{(k-\Gamma_s-\Delta_s)}/\partial \bar{Y}_s)(\partial/\partial Y_s) \qquad (100)$$

where the lower limit of k in the last summation may be chosen as 1 if quantities with negative or zero index are understood to be zero. This gives

$$L_W^{(k)} = \sum_s (\partial \chi^{(k-\Gamma_s-\Delta_s)}/\partial \bar{Y}_s)(\partial/\partial Y_s) \qquad (101)$$

Expanding a typical component of (96)

$$z_i = \left\{ 1 + (\sum \varepsilon^k L_W^{(k)}) + \tfrac{1}{2}(\sum \varepsilon^k L_W^{(k)})^2 + .. \right\} * y_i$$

$$= \sum \varepsilon^k M^{(k)} * (\varepsilon^{-\Gamma_i} Y_i)$$

$$= \sum \varepsilon^k M^{(k+\Gamma_i)} * Y_i \qquad (102)$$

where the $M^{(k)}$ all have the form

$$M^{(k)} = L_W^{(k)} + N^{(k)} \qquad (103)$$

with $N^{(k)}$ some operator involving lower orders. One then gets

$$\gamma_i^{(k)} = \zeta_i^{(k+\Delta_i)}$$

$$= M^{(k+\Gamma_i+\Delta_i)} * Y_i$$

$$= L_W^{(k+\Gamma_i+\Delta_i)} * Y_i + N^{(k+\Gamma_i+\Delta_i)} * Y_i$$

$$= \partial\chi^{(k)}/\partial\bar{Y}_i + N^{(k+\Gamma_i+\Delta_i)} * Y_i \qquad (104)$$

which again is the sum of a gradient in $\underline{\bar{Y}}$ space and an expression involving lower orders which (presumably) is a particular solution of (78).

THE KRYLOV - BOGOLIUBOV - KRUSKAL METHOD WITH SLOW VARIABLES

Krylov and Bogoliubov[9][10][11] investigated the solution of a set of n equations vectorially represented by

$$d\underline{y}/dt = \sum_{k=0} \varepsilon^k \underline{g}^{(k)}(\underline{y}) \qquad (105)$$

with

$$\underline{g}^{(0)} = (0, 0, \ldots 0, g_n^{(0)}) \qquad (106)$$

ensuring that in the "unperturbed" limit $\varepsilon \to 0$, y_n alone varies and all other components of $\underline{y}$ (to be collectively denoted by $\underline{\tilde{y}}$) are constant. It is further assumed that the unperturbed system is periodic and that y_n is an angle variable appearing only in the angle argument of periodic functions. The zero-order growth of y_n is then assumed to be linear, from which follows that $g_n^{(0)}$ may depend on $\underline{\tilde{y}}$ but not on y_n.

To eliminate the periodicity from this motion, Krylov and Bogoliubov used a near-identity transformation to new variables $\underline{z}$, given in a direct form inverse to that of (37)

$$\underline{y} = \underline{z} + \sum \varepsilon^k \underline{\eta}^{(k)}(\underline{z}) \tag{107}$$

The new variables, which can be derived by a suitable recursive method, have the property that the equations by which they evolve do not contain the transformed angle variable z_n on the right-hand side but have the form

$$d\underline{z}/dt = \sum \varepsilon^k \underline{h}^{(k)}(\underline{\tilde{z}}) \tag{108}$$

The first (n-1) equations of this set, representing $d\underline{\tilde{z}}/dt$, then form an autonomous set not involving z_n and can be solved independently.

If $\underline{y}$ represents a perturbed periodic canonical system with a Hamiltonian of the form (50), then the canonical equations of motion have the form (105) and the Krylov-Bogoliubov method can be used to eliminate the angle variable $y_n = \Omega$. Unfortunately, unless precautions are taken(22)(23) , the $\underline{z}$ variables will in general not be canonical, so that the transformed variable corresponding to the canonical conjugate of y_n will in general not be a constant of the perturbed motion, as is automatically achieved by the Poincaré-Von Zeipel method.

On the other hand, the Krylov-Bogoliubov method has a much wider validity and can be used in non-hamiltonian systems. A similar elimination procedure which derives the transformation in the form (37) has been devised by

Kruskal[8][12], who followed it by the derivation (for canonical systems only) of a constant J of the motion, obtained by an ingenious application of integral invariants (it is the same constant as is obtained by the Poincaré-Von Zeipel method).

Here the Krylov-Bogoliubov method will be generalized for the case when slow variables are present. As with the Poincaré-Von Zeipel method, this allows the restrictions on the form of the zero-order equations — embodied in $\underline{g}^{(0)}$ — to be eased. Specifically, some variables other than y_n are now allowed to have a zero-order variation and this variation (as in the canonical method) is passed intact to the "reduced" equations involving $\underline{z}$. The calculation will be done for the transformation (107) ; the treatment of Kruskal's method, using (37), follows practically identical steps and will therefore be omitted.

Following the notation of (89), let $\underline{R}$ and $\underline{r}$ denote the slow and normal components of $\underline{z}$, and let $\underline{\eta}_R^{(k)}$ and $\underline{\eta}_r^{(k)}$ be corresponding components of $\underline{\eta}^{(k)}$. Substituting (107) in the left-hand side of (105) gives, with the definitions (67)(68) extended to $\underline{z}$ variables

$$d\underline{y}/dt = d\underline{z}/dt + \sum_{k,\,s} \varepsilon^k (\partial \underline{\eta}^{(k)}/\partial z_s)(dz_s/dt)$$

$$= d\underline{z}/dt + \sum_{k,\,s} \varepsilon^k (\partial \underline{\eta}^{(k-r_s)}/\partial z_s) \sum_{m=0} \varepsilon^m h_s^{(m)}$$

$$= d\underline{z}/dt + \sum \varepsilon^k \sum_{m=0}^{k-1} (\partial \underline{\eta}^{(k-m-r_s)}/\partial z_s)\, h_s^{(m)} \qquad (109)$$

Expressing a typical term of the right-hand side of (105) in terms of $\underline{z}$

gives, in a manner similar to (88)

$$\begin{aligned} \underline{g}^{(k)}(\underline{y}) &= \underline{g}^{(k)}\left(\underline{z} + \sum \varepsilon^m \underline{\eta}^{(m)} \right) \\ &= \exp\left\{ \sum_{m,\, s} \varepsilon^m \left(\underline{\eta}_s^{(m)} \cdot \partial/\partial z_s \right) \right\} * \underline{g}^{(k)}(\underline{z}) \\ &= \exp\left\{ \sum_{m,\, s} \varepsilon^m \left(\underline{\eta}_s^{(m-r_s)} \cdot \partial/\partial z_s \right) \right\} * \underline{g}^{(k)}(\underline{z}) \\ &= \sum \varepsilon^m K^{(m)} * \underline{g}^{(k)}(\underline{z}) \end{aligned} \tag{110}$$

The operators $K^{(m)}$ resemble those of (89) but with $\underline{\eta}^{(m)}$ everywhere replacing $\underline{\Pi}^{(m)}$. Substituting preceding results in (105) gives

$$\begin{aligned} d\underline{z}/dt = \sum_{k=0} \varepsilon^k \Bigg[& \sum_{m=1}^{k-1} \left\{ K^{(m)} * \underline{g}^{(k-m)} - \sum_s \left(\partial \underline{\eta}^{(k-m-r_s)} / \partial z_s \right) \cdot h_s^{(m)} \right\} \\ & + \left(K^{(k)} * \underline{g}^{(0)} \right) - \sum_s \left(\partial \underline{\eta}^{(k-r_s)} / \partial z_s \right) \cdot g_s^{(0)} \\ & + \left(1 - \delta_{k0}\right) \underline{g}^{(k)} \Bigg] \end{aligned} \tag{111}$$

where the factor preceding the last term denotes that it be omitted for $k = 0$ (in that case it is already counted as the term involving $K^{(k)}$) and where in the summation preceding this term $\underline{h}^{(0)}$ has been replaced by $\underline{g}^{(0)}$, which equals it since in the limit $\varepsilon \to 0$ eqs. (105) and (108) coincide.

Comparison with (108) shows that the expression in the square brackets equals $\underline{h}^{(k)}(\tilde{\underline{z}})$, and this equality forms the basis of the recursive derivation of $\underline{h}^{(k)}$ and $\underline{\eta}^{(k)}$.

The situation now resembles that of (63): in order that the recursion be at all possible, unknown components of $\underline{\eta}^{(k)}$ must appear in (111) only <u>once</u>, otherwise the result is a partial differential equation and cannot be easily integrated. One term which <u>always</u> contains $\underline{\eta}^{(k)}$ is contributed by the last summation in (111) and equals

$$(\partial \underline{\eta}^{(k)}/\partial z_n) \quad g_n^{(0)}$$

since $g_n^{(0)}$ does not vanish and z_n , the transformed angle variable, is "normal". No other appearance is permitted, hence

$$g_s^{(0)} = 0 \qquad \text{for} \qquad \Gamma_s = 0 \tag{112}$$

or, stated in words, "Only slow variables and the principal angle variable are allowed to have a zero-order variation."

In addition, $\underline{\eta}^{(k)}$ could enter through the term containing $K^{(k)}$, which has the form (compare eq. 89)

$$K^{(k)}{}_* \underline{g}^{(0)} = \left\{ \underline{\eta}_r^{(k)} \cdot \partial/\partial \underline{x} + N^{(k)} \right\} * \underline{g}^{(0)} \tag{113}$$

with $N^{(k)}$ containing lower orders. No problem arises here <u>provided</u> $\underline{\eta}_r^{(k)}$ is derived first and $\underline{\eta}_R^{(k)}$ only afterwards: because of (112), this term is absent in the first part of the derivation, while in the second part those components of $\underline{\eta}^{(k)}$ that appear in it are already known. In either case one gets

$$(\partial \underline{\eta}^{(k)}/\partial z_n)\, g_n^{(0)} \quad + \quad \underline{h}^{(k)}(\underline{z}) \quad = \quad \underline{\lambda}^{(k)} \tag{114}$$

where $(k > 0)$

$$\underline{\lambda}^{(k)} = \sum_{m=1}^{k-1} \left\{ K^{(m)} * \underline{g}^{(k-m)} - \sum_{s} (\partial \underline{\eta}^{(k-m-r_s)}/\partial z_s)\, h_s^{(m)} \right.$$

$$+ \quad (K^{(k)} * \underline{g}^{(0)}) \quad + \quad \underline{g}^{(k)} \quad - \quad \sum_{s \neq n} (\partial \underline{\eta}^{(k-r_s)}/\partial z_s)\, g_s^{(0)} \tag{115}$$

depends only on lower orders. The solving of (114) then resembles that of (21)

CONCLUSION

In the preceding sections the main methods of classical perturbation theory have been extended to slowly (or adiabatically) perturbed systems. At the same time, the basic concepts associated with such systems (e.g. adiabatic invariance and implicit ε) were examined and clarified.

The restrictions on the forms of the zero-order equations for slow perturbations have been derived and are generally less severe than for small perturbations. The extension of the methods themselves is relatively straightforward, involving mainly the shifting of indices for quantities corresponding either to slow variables ot (as in the case of $\underline{\gamma}^{(k)}$) to variables with slow conjugates. With the use of expansion operators the treatment is only slightly more complicated than for small perturbations.

REFERENCES AND COMMENTS

(1) This term has occasionally been applied (even by this author, in ref. 12) to denote approximate constants of motion derived from $\oint \underline{p} \cdot d\underline{q}$. Actually such constants may belong to either of the types discussed here.

(2) G. Contopoulos, J. Math. Phys. 7, 788 (1966)

(3) H. Poincaré, Les Methodes Nouvelles de la Mecanique Celeste, vol. II, Gauthier-Villars, Paris 1893; reprinted by Dover Publications, New York, 1957; NASA translation TT F-451, 1967.

(4) H. Von Zeipel, Ark. Astr. Mat. Fys. 11, no. 1 (1916)

(5) H. C. Corben and P. Stehle, Classical Mechanics, 2nd Ed. ; John Wiley and Sons, 1960 ; Sect. 75.

(6) D. Ter Haar, Elements of Hamiltonian Mechanics, North Holland Publishing Co. 1961 ; Chapter 7 .

(7) P. Musen, J. of the Astronautical Sciences 12, 129 (1965)

(8) M. Kruskal, J. Math. Physics 3, 806 (1962)

(9) N. M. Krylov and N. N. Bogoliubov, Introduction to Non-Linear Mechanics, Kiev 1936.

(10) N. N. Bogoliubov and D. N. Zubarev, Ukr. Math. J. VII, 5 (1955) ; Translated by Burton Fried and published 1960 by Space Technology Laboratories.

(11) N.N. Bogoliubov and Y. A. Mitropolski, Asymptotic Methods in the Theory of Nonlinear Oscillations, Gordon and Breach, N.Y. 1961; Section 25.

(12) D. Stern, J. Math. Phys. 11, 2771 (1970)

(13) Ref. (6), Section 6-3.

(14) M. Born, The Mechanics of the Atom, 1925, 2nd. Edit. of English Translation publ. 1960 by Frederick Ungar Publ. Co. ; Section 10.

(15) C. Gardner, Phys. Rev. 115, 791 (1959)

(16) May be omitted at first reading.

(17) D. Stern, J. Math. Phys. 11, 2776 (1970)

(18) D. Stern, Euler Potentials and their Application in the Hamiltonian Formulation of Guiding Center Motion, 1969, Goddard Space Flight Center publication X-641-69-97

(19) S. Lie, Theorie der Transformationgruppen I, Teubner, Leipzig 1888

(20) G. Hori, Publications of the Astronomical Society of Japan 18, 287 (1966)

(21) A. Deprit, Celestial Mechanics 1, 12 (1969)

(22) E. L. Burshtein and L. S. Solov'ev, Soviet Physics (Doklady), English translation 6, 731 (1962)

(23) D. Stern, Canonization of Nice Variables, 1970, Goddard Space Flight Center publication X-641-70-316