in a very small neighborhood of $\mathbf{k} = 0$, corresponding to the fact that $\beta\omega_\mathbf{k}$ becomes very small in that neighborhood.[7] Then it follows from (A6) that $w(\mathbf{k})$ differs infinitesimally from $2\,|\epsilon|\,\sigma(\mathbf{k})$ in this neighborhood.[21] Thus by (A8) and (A9) one finds, upon replacing $e^{\beta\omega(\mathbf{k})}$ by the first two terms of its power-series expansion,

$$n_\mathbf{k} \approx 2\,|\epsilon|\,\sigma(\mathbf{k})\kappa T/\omega^2(\mathbf{k}),$$

$$\langle a_\mathbf{k} a_{-\mathbf{k}}\rangle \approx -2\epsilon^*\sigma(\mathbf{k})\kappa T/\omega^2(\mathbf{k}), \qquad (A11)$$

for $T < T_c$ and $\mathbf{k} \to 0$. More explicit results can be obtained by solving (A4) for $n_\mathbf{k}$ after the substitutions $\alpha = -|\alpha|$, $w(\mathbf{k}) \approx 2\,|\epsilon|\,\sigma(\mathbf{k})$ [the latter following from $\omega(\mathbf{k}) \approx 0$]. This gives

$$n_\mathbf{k} = \tfrac{1}{2}(k_s^2 - k^2)\Omega/4\pi\,|\alpha|, \quad T < T_c, \quad k < k_s, \quad (A12)$$

where

$$k_s = \{2[\mu + 4\pi\rho\,|\alpha| + 2\,|\epsilon|\,\sigma(\mathbf{k})]\}^{\frac{1}{2}}. \quad (A13)$$

---

[21] We assume that $\sigma(\mathbf{k}) \geq 0$.

Then, by (A11),

$$\langle a_\mathbf{k} a_{-\mathbf{k}}\rangle = -\tfrac{1}{2}e^{-i\vartheta}(k_s^2 - k^2)\Omega/4\pi\,|\alpha|,$$
$$T < T_c, \quad k < k_s, \quad (A14)$$

where $\epsilon = |\epsilon|\,e^{i\vartheta}$. Equations (5) and (6) then give

$$p_c = \rho_c = f\rho, \quad T < T_c,$$
$$p_c = \rho_c = 0, \quad T > T_c, \qquad (A15)$$

where[11]

$$(1 - f)\rho = 2.612(\kappa T/2\pi)^{\frac{3}{2}}, \quad T < T_c, \quad (A16)$$
$$\kappa T_c = 2\pi(\rho/2.612)^{\frac{2}{3}},$$

and

$$\lim_{\epsilon \to 0} k_s = [15(2\pi)^3 f\rho\,|\alpha|/\Omega]^{\frac{1}{5}}, \quad T < T_c. \quad (A17)$$

The infinite susceptibility of $\langle a_\mathbf{k} a_{-\mathbf{k}}\rangle$ to the symmetry-breaking perturbation $V_\epsilon$ is clear in this model; with $\epsilon = 0$ a grand canonical ensemble calculation gives $\langle a_\mathbf{k} a_{-\mathbf{k}}\rangle = 0$; on the other hand, if $\langle a_\mathbf{k} a_{-\mathbf{k}}\rangle$ is interpreted as the quasiaverage (3), it is not zero and in fact gets large like $\langle N_\mathbf{k}\rangle$ as $\mathbf{k} \to 0$, for arbitrarily small but nonzero $\epsilon$ and $T < T_c$.

---

# Nonlinear Perturbations*

Timothy P. Coffey† and G. W. Ford
*Department of Physics, The University of Michigan, Ann Arbor, Michigan*

(Received 12 June 1967)

The perturbation theory of Bogoliubov and Mitropolsky for systems having a single rapid phase is generalized to systems having several rapid phases. It is shown that one can avoid the classic problem of small divisors to all orders in the perturbation theory. The method has the advantage of providing a single approach to many problems conventionally treated by a variety of specialized techniques.

## 1. INTRODUCTION

The techniques of perturbation theory for nonlinear systems, initiated by Poincaré three quarters of a century ago, have been extended and developed by many workers. One such technique, the method of averaging, was introduced by Krylov and Bogoliubov thirty years ago.[1] The essential feature of this method is the separation of a given motion into a secular motion plus a rapidly fluctuating motion of small amplitude; the given motion is then expressed in terms of the solution of a system of differential equations which describe the secular motion alone.

A wide variety of physical problems may be handled by this method, e.g., Case in a recent publication has shown how the method can be applied to time-dependent perturbation theory in quantum mechanics.[2]

Bogoliubov and Mitropolsky have presented a form of the method of averaging, called the method of rapidly rotating phase, which is especially convenient for systems in which a single variable, called the phase, has a rapid secular motion.[3] Our purpose in this paper is to extend this method to systems with

---

[1] N. N. Krylov and N. N. Bogoliubov, *Introduction to Non-Linear Mechanics* (Academy of Sciences of the Ukranian S.S.R., Kiev, 1937), trans. by S. Lefschetz in *Annals of Mathematics Studies*, No. 11 (Princeton University Press, Princeton, N.J., 1947).

[2] K. M. Case, Suppl. Progr. Theoret. Phys. (Kyoto) 37, 1 (1966). See also R. Y. Y. Lee, "On a New Perturbation Method" Thesis, The University of Michigan, Ann Arbor, 1964.
[3] N. N. Bogoliubov and D. N. Zubarev, Ukrain. Mat. Zh. 7, 5 (1955); N. N. Bogoliubov and Y. A. Mitropolsky, *Asymptotic Methods in the Theory of Non-Linear Oscillations* (Hindustan Publishing Co., Delhi, India, 1961), Chaps. 5 and 6. See also N. Minorsky, *Nonlinear Oscillations* (D. van Nostrand Co., Princeton, N.J., 1962), and M. Kruskal, J. Math. Phys. 3, 806 (1962).

several rapid phases and to succinctly describe how the method works, first in the nondegenerate case, and then in the more interesting degenerate case. In a following paper one of us (T. P. C.) will use the method to discuss the perturbation by a weak, transverse, spatially periodic magnetic field of the motion of a charged particle gyrating in a uniform magnetic field.[4]

## 2. NONDEGENERATE PERTURBATION THEORY

We consider the following set of coupled differential equations[5]:

$$\dot{x}_i = \epsilon A_i(\mathbf{x}, \boldsymbol{\psi}), \qquad i = 1, 2, \cdots, r, \quad (2.1a)$$

$$\dot{\psi}_j = \omega_j(x) + \epsilon B_j(\mathbf{x}, \boldsymbol{\psi}), \quad j = 1, 2, \cdots, s, \quad (2.1b)$$

where $\epsilon$ is a small parameter, $\mathbf{x} = (x_1, \cdots, x_r)$, $\boldsymbol{\psi} = (\psi_1, \cdots, \psi_s)$, and the $A_i$'s and $B_j$'s are periodic functions of each of the $\psi_k$'s with period $2\pi$. The dot represents differentiation with respect to time.

When $\epsilon = 0$, the $x_i$'s will be constants and the $\psi_j$'s will be linear functions of time. When $\epsilon$ is small but finite, the $x_i$'s will experience a slow secular growth on which is superimposed small-amplitude rapid fluctuations. Similarly, the $\psi_j$'s will experience a rapid secular growth on which is superimposed small-amplitude rapid fluctuations. Our aim is to separate this secular motion from the rapid fluctuating motion. To do this we seek a solution in the form

$$x_i = y_i + \sum_{n=1}^{\infty} \epsilon^n F_i^{(n)}(\mathbf{y}, \boldsymbol{\phi}), \quad i = 1, 2, \cdots, r, \quad (2.2a)$$

$$\psi_j = \phi_j + \sum_{n=1}^{\infty} \epsilon^n G_j^{(n)}(\mathbf{y}, \boldsymbol{\phi}), \quad j = 1, 2, \cdots, s, \quad (2.2b)$$

where the $F_i^{(n)}$'s and $G_j^{(n)}$'s are periodic functions of each of the $\phi_k$'s with period $2\pi$. We further require that the new variables $y_i$ and $\phi_j$ satisfy the following differential equations:

$$\dot{y}_i = \sum_{n=1}^{\infty} \epsilon^n a_i^{(n)}(\mathbf{y}), \qquad i = 1, 2, \cdots, r, \quad (2.3a)$$

$$\dot{\phi}_j = \omega_j(\mathbf{y}) + \sum_{n=1}^{\infty} \epsilon^n b_j^{(n)}(\mathbf{y}), \quad j = 1, 2, \cdots, s, \quad (2.3b)$$

where the right-hand sides of Eq. (2.3) are required to be independent of the $\phi_k$'s. The idea here is that the $y_i$ and $\phi_j$ exhibit only secular motion, since they are solutions of a system of differential equations which are independent of the rapidly increasing (or decreasing) phases $\phi_j$. The rapid fluctuations of the $x_i$ and $\psi_j$ about the $y_i$ and $\phi_j$ are given by the terms in the series in (2.2). We must now show that we can

[4] T. P. Coffey, J. Math. Phys. **10**, 1362 (1969).
[5] The generalization to the general case where the right-hand sides of (2.1) are power series in $\epsilon$ is straightforward.

construct the function $F_i^{(n)}$, $G_j^{(n)}$, $a_i^{(n)}$, and $b_j^{(n)}$ so that (2.2) is indeed a solution of the set of differential equations (2.1).

If we insert Eq. (2.2) in Eq. (2.1) and then use Eq. (2.3) we find, upon equating equal powers of $\epsilon$,

$$a_i^{(1)}(\mathbf{y}) + \sum_{k=1}^{s} \omega_k \frac{\partial F_i^{(1)}(\mathbf{y}, \boldsymbol{\phi})}{\partial \phi_k} = A_i(\mathbf{y}, \boldsymbol{\phi}), \quad (2.4a)$$

$$b_j^{(1)}(\mathbf{y}) + \sum_{k=1}^{s} \omega_k \frac{\partial G_j^{(1)}(\mathbf{y}, \boldsymbol{\phi})}{\partial \phi_k}$$

$$= B_j(\mathbf{y}, \boldsymbol{\phi}) + \sum_{l=1}^{r} F_l^{(1)}(\mathbf{y}, \boldsymbol{\phi}) \frac{\partial \omega_j(\mathbf{y})}{\partial y_l}, \quad (2.4b)$$

from the first power of $\epsilon$, and

$$a_i^{(2)} + \sum_{k=1}^{s} \omega_k \frac{\partial F_i^{(2)}}{\partial \phi_k} = \sum_{k=1}^{s} G_k^{(1)} \frac{\partial A_i}{\partial \phi_k} + \sum_{l=1}^{r} F_l^{(1)} \frac{\partial A_i}{\partial y_l}$$

$$- \sum_{l=1}^{r} a_l^{(1)} \frac{\partial F_i^{(1)}}{\partial y_l} - \sum_{k=1}^{s} b_k^{(1)} \frac{\partial F^{(1)}}{\partial \phi_k}, \quad (2.5a)$$

$$b_j^{(2)} + \sum_{k=1}^{s} \omega_k \frac{\partial G_j^{(2)}}{\partial \phi_k}$$

$$= \frac{1}{2} \sum_{l=1}^{r} \sum_{m=1}^{r} F_l^{(1)} F_m^{(1)} \frac{\partial^2 \omega_j}{\partial y_l \partial y_m} + \sum_{l=1}^{r} F_l^{(2)} \frac{\partial \omega_j}{\partial y_l}$$

$$+ \sum_{k=1}^{s} G_k^{(1)} \frac{\partial B_j}{\partial \phi_k} + \sum_{l=1}^{r} F_l^{(1)} \frac{\partial B_j}{\partial y_l}$$

$$- \sum_{l=1}^{r} a_l^{(1)} \frac{\partial G_j^{(1)}}{\partial y_l} - \sum_{k=1}^{s} b_k^{(1)} \frac{\partial G^{(1)}}{\partial \phi_k}, \quad (2.5b)$$

from the second power of $\epsilon$, and so on. We thus obtain a sequence of equations for the determination of the unknown functions.

Each of these equations is of the general form

$$a(\mathbf{y}) + \sum_{k=1}^{s} \omega_k(\mathbf{y}) \frac{\partial F(\mathbf{y}, \boldsymbol{\phi})}{\partial \phi_k} = A(\mathbf{y}, \boldsymbol{\phi}), \quad (2.6)$$

where $a(\mathbf{y})$ and $F(\mathbf{y}, \boldsymbol{\phi})$ are to be determined and $A(\mathbf{y}, \boldsymbol{\phi})$ is a periodic function of the $\phi_k$ which is known in terms of the solutions of the previous equations. Note that the dependence upon $\mathbf{y}$ is trivial, the $y_i$ behaving as parameters in this equation, so we may suppress this dependence for the moment and write the equation in the form

$$\sum_{k=1}^{s} \omega_k \frac{\partial F(\boldsymbol{\phi})}{\partial \phi_k} = A(\boldsymbol{\phi}) - a. \quad (2.7)$$

This equation, viewed as an equation for determining $F(\boldsymbol{\phi})$, is a first order, linear, inhomogeneous partial differential equation with constant coefficients. Solutions of such an equation exist only if the inhomogeneous term is orthogonal to all solutions of the

homogeneous equation:

$$\sum_{k=1}^{s} \omega_k \frac{\partial F(\boldsymbol{\phi})}{\partial \phi_k} = 0. \qquad (2.8)$$

But the solutions of this equation are all of the form

$$F(\boldsymbol{\phi}) = \exp\left\{ i\sum_{k=1}^{s} p_k \phi_k \right\}, \qquad (2.9)$$

where, because $F(\boldsymbol{\phi})$ must be periodic in each of the $\phi_k$, the $p_k$ must be integers and, because (2.9) must be a solution of (2.8), these integers must satisfy the identity

$$\sum_{k=1}^{s} p_k \omega_k = 0. \qquad (2.10)$$

In the nondegenerate case we assume there are no sets of integers satisfying this identity except for the trivial set in which all the $p_k$ are zero, i.e., $F(\boldsymbol{\phi})$ is a constant. If there is a nontrivial set of integers satisfying (2.10) we say there is a degeneracy; we discuss this case in Sec. 4.

We see, therefore, that in the nondegenerate case $a$ must be chosen so there is no constant term on the right-hand side of (2.7); $F(\boldsymbol{\phi})$ is then the solution of the resulting differential equation. To exhibit this solution more explicitly, we return to Eq. (2.6) where the **y** dependence is indicated. The given function $A(\mathbf{y}, \boldsymbol{\phi})$, since it is periodic in the $\phi_k$, must be expressible in the form

$$A(\mathbf{y}, \boldsymbol{\phi}) = \sum_{\mathbf{p}} A_{\mathbf{p}}(\mathbf{y}) e^{i\mathbf{p}\cdot\boldsymbol{\phi}}, \qquad (2.11)$$

where the sum is over all sets of integers $\mathbf{p} = (p_1, p_2, \cdots, p_s)$ and

$$\mathbf{p} \cdot \boldsymbol{\phi} \equiv \sum_{k=1}^{s} p_k \phi_k. \qquad (2.12)$$

The function $a(\mathbf{y})$ must be chosen to cancel the terms corresponding to $\mathbf{p} = \mathbf{0}$, in which all the $p_k$ are zero:

$$a(\mathbf{y}) = A_0(\mathbf{y})$$
$$= \frac{1}{(2\pi)^s} \int_0^{2\pi} d\phi_1 \int_0^{2\pi} d\phi_2 \cdots \int_0^{2\pi} d\phi_s A(\mathbf{y}, \boldsymbol{\phi}). \quad (2.13)$$

The solution of (2.6) is then

$$F(\mathbf{y}, \boldsymbol{\phi}) = -i\sum_{\mathbf{p}}{}' \frac{A_{\mathbf{p}}(\mathbf{y})}{\mathbf{p}\cdot\boldsymbol{\omega}} e^{i\mathbf{p}\cdot\boldsymbol{\phi}} + f(\mathbf{y}), \qquad (2.14)$$

where the prime indicates that the term $\mathbf{p} = \mathbf{0}$ is absent from the sum and

$$\mathbf{p} \cdot \boldsymbol{\omega} \equiv \sum_{k=1}^{s} p_k \omega_k(\mathbf{y}). \qquad (2.15)$$

The function $f(\mathbf{y})$ in (2.14) is arbitrary; the solution of

an inhomogeneous, linear, partial differential equation is determined only up to an arbitrary solution of the homogeneous equation. We usually choose $f(\mathbf{y})$ to be zero.

Thus, we see how the two functions $a(\mathbf{y})$, given by (2.13), and $F(\mathbf{y}, \boldsymbol{\phi})$, given by (2.14), are determined from the single equation (2.6). Since each of the equations in the sequence for the determination of the functions $F_i^{(n)}(\mathbf{y}, \boldsymbol{\phi})$, $G_j^{(n)}(\mathbf{y}, \boldsymbol{\phi})$, $a_i^{(n)}(\mathbf{y})$, and $b_j^{(n)}(\mathbf{y})$ is of the form (2.6), we may, in the nondegenerate case, successively solve to determine these functions. To be more explicit, we first note that the given functions $A_i(\mathbf{x}, \boldsymbol{\psi})$ and $B_j(\mathbf{x}, \boldsymbol{\psi})$ in (2.1) are periodic in each of the $\psi_k$ and so may be expanded in the form

$$A_i(\mathbf{x}, \boldsymbol{\psi}) = \sum_{\mathbf{p}} A_{i,\mathbf{p}}(\mathbf{x}) e^{i\mathbf{p}\cdot\boldsymbol{\psi}}, \qquad (2.16a)$$

$$B_j(\mathbf{x}, \boldsymbol{\psi}) = \sum_{\mathbf{p}} B_{j,\mathbf{p}}(\mathbf{x}) e^{i\mathbf{p}\cdot\boldsymbol{\psi}}. \qquad (2.16b)$$

Then from (2.4a) we find

$$a_i^{(1)}(\mathbf{y}) = A_{i,0}(\mathbf{y}) \qquad (2.17)$$

and

$$F_i^{(1)}(\mathbf{y}, \boldsymbol{\phi}) = -i\sum_{\mathbf{p}}{}' \frac{A_{i,\mathbf{p}}(\mathbf{y})}{\mathbf{p}\cdot\boldsymbol{\omega}} e^{i\mathbf{p}\cdot\boldsymbol{\phi}}. \qquad (2.18)$$

Using this solution in (2.4b) we then find

$$b_j^{(1)}(\mathbf{y}) = B_{j,0}(\mathbf{y}) \qquad (2.19)$$

and

$$G_j^{(1)}(\mathbf{y}, \boldsymbol{\phi})$$
$$= -i\sum_{\mathbf{p}}{}' \left[ \frac{B_{j,\mathbf{p}}(\mathbf{y})}{\mathbf{p}\cdot\boldsymbol{\omega}} - i\frac{\sum_{l=1}^{r} \frac{\partial\omega_j(\mathbf{y})}{\partial y_l} A_{l,\mathbf{p}}(\mathbf{y})}{(\mathbf{p}\cdot\boldsymbol{\omega})^2} \right] e^{i\mathbf{p}\cdot\boldsymbol{\phi}}, \quad (2.20)$$

and so on. The expressions become increasingly cumbersome, but we can, in principle, solve to obtain explicit expressions for the $F_i^{(n)}$, $G_j^{(n)}$, $a_i^{(n)}$, and $b_j^{(n)}$ so (2.2) is a solution of the system of equations (2.1) to any desired order in $\epsilon$.

## 3. THE VAN DER POL EQUATION

As a simple example illustrating the working of the general method for the nondegenerate case, we consider the van der Pol equation

$$\ddot{z} + \epsilon(z^2 - 1)\dot{z} + z = 0. \qquad (3.1)$$

We cast this equation into the standard form (2.1) by introducing variables $x$ and $\psi$ through the substitution:

$$z = x^{\frac{1}{2}} \cos\psi,$$
$$\dot{z} = -x^{\frac{1}{2}} \sin\psi, \qquad (3.2)$$

that is,

$$x = z^2 + \dot{z}^2,$$
$$\psi = -\arctan \dot{z}/z. \qquad (3.3)$$

Forming the time derivative of both sides of this last pair of equations and using (3.1) and (3.2) on the right-hand sides, we find

$$\dot{x} = 2\epsilon x(1 - x\cos^2\psi)\sin^2\psi$$
$$= \epsilon x(1 - \tfrac{1}{4}x - \cos 2\psi + \tfrac{1}{4}x\cos 4\psi) \quad (3.4a)$$

$$\dot{\psi} = 1 + \epsilon(1 - x\cos^2\psi)\sin\psi\cos\psi$$
$$= 1 + \epsilon[(\tfrac{1}{2} - \tfrac{1}{4}x)\sin 2\psi - \tfrac{1}{8}x\sin 4\psi]. \quad (3.4b)$$

These equations are in the standard form (2.1) for applying the method of rapidly rotating phase, with $\epsilon$ a small parameter.

According to our general method, we seek a solution in the form

$$x = y + \epsilon F^{(1)}(y, \phi) + \epsilon^2 F^{(2)}(y, \phi) + \cdots, \quad (3.5a)$$

$$\psi = \phi + \epsilon G^{(1)}(y, \phi) + \epsilon^2 G^{(2)}(y, \phi) + \cdots, \quad (3.5b)$$

where

$$\dot{y} = \epsilon a^{(1)}(y) + \epsilon^2 a^{(2)}(y) + \cdots, \quad (3.6a)$$

$$\dot{\phi} = 1 + \epsilon b^{(1)}(y) + \epsilon^2 b^{(2)}(y) + \cdots. \quad (3.6b)$$

Inserting (3.5) in (3.4), using (3.6), and equating powers of $\epsilon$, we get the following sequence of equations.

$$a^{(1)} + \frac{\partial F^{(1)}}{\partial \phi} = y(1 - \tfrac{1}{4}y - \cos 2\phi + \tfrac{1}{4}y\cos 4\phi),$$
$$(3.7a)$$

$$b^{(1)} + \frac{\partial G^{(1)}}{\partial \phi} = (\tfrac{1}{2} - \tfrac{1}{4}y)\sin 2\phi - \tfrac{1}{8}y\sin 4\phi, \quad (3.7b)$$

from the first power of $\epsilon$, and

$$a^{(2)} + \frac{\partial F^{(2)}}{\partial \phi} = (1 - \tfrac{1}{2}y - \cos 2\phi + \tfrac{1}{2}y\cos 4\phi)F^{(1)}$$
$$+ (2y\sin 2\phi - y^2\sin 4\phi)G^{(1)}$$
$$- a^{(1)}\frac{\partial F^{(1)}}{\partial y} - b^{(1)}\frac{\partial F^{(1)}}{\partial \phi}, \quad (3.8a)$$

$$b^{(2)} + \frac{\partial G^{(2)}}{\partial \phi} = -(\tfrac{1}{4}\sin 2\phi + \tfrac{1}{8}\sin 4\phi)F^{(1)}$$
$$+ [(1 - \tfrac{1}{2}y)\cos 2\phi - \tfrac{1}{2}y\cos 4\phi]G^{(1)}$$
$$- b^{(1)}\frac{\partial G^{(1)}}{\partial \phi} - a^{(1)}\frac{\partial G^{(1)}}{\partial y}, \quad (3.8b)$$

from the second power of $\epsilon$, and so on.

We solve this sequence of equations as indicated in the previous section. From (3.7a) we find

$$a^{(1)}(y) = y(1 - \tfrac{1}{4}y), \quad (3.9a)$$

$$F^{(1)}(y, \phi) = y(-\tfrac{1}{2}\sin 2\phi + \tfrac{1}{16}y\sin 4\phi), \quad (3.9b)$$

while from (3.7b) we find

$$b^{(1)}(y) = 0, \quad (3.9c)$$

$$G^{(1)}(y) = -\tfrac{1}{4}(1 - \tfrac{1}{2}y)\cos 2\phi + \tfrac{1}{32}y\cos 4\phi. \quad (3.9d)$$

From (3.8a), using the solutions (3.9), we find

$$a^{(2)}(y) = 0, \quad (3.10a)$$

$$F^{(2)}(y, \phi)$$
$$= y^2(\tfrac{1}{32}(y - 5)\cos 2\phi - \tfrac{1}{64}\cos 4\phi + \tfrac{1}{96}y\cos 6\phi), \quad (3.10b)$$

while from (3.8b) we find

$$b^{(2)}(y) = -\frac{1}{8} + \frac{3y}{16} - \frac{11y^2}{256}, \quad (3.10c)$$

$$G^{(2)}(y, \phi)$$
$$= -\frac{y(1 + y)}{128}\sin 2\phi - \frac{16 - 4y + 3y^2}{512}\sin 4\phi$$
$$+ \frac{y(3 - 2y)}{384}\sin 6\phi - \frac{y^2}{2048}\sin 8\phi. \quad (3.10d)$$

These expressions, when inserted in (3.5) and (3.6), give the complete reduction of the problem through second order in $\epsilon$.

The method of rapidly rotating phase does not in general lead to an explicit solution of the original set of differential equations. Rather, it is a method for separating the secular motion from the rapid periodic fluctuations and reducing the problem to that of solving the differential equations for the secular motion alone. The solution of these equations, i.e., in the general case the equations (2.3), may be a very difficult problem, but in the case of the van der Pol equation it is quite simple. Using (3.9a) and (3.10a), the differential equation (3.6a) becomes

$$\dot{y} = \epsilon y(1 - \tfrac{1}{4}y) \quad (3.11)$$

through second order in $\epsilon$. The solution is

$$y(t) = \frac{4y(0)}{y(0) + [4 - y(0)]e^{-\epsilon t}}. \quad (3.12)$$

Here we see the well-known feature of the van der Pol equation: for long times the amplitude approaches a constant independent of the initial amplitude. Inserting this solution in (3.6b) we can integrate to find

$$\phi(t) = \phi(0) + \left(1 - \frac{\epsilon^2}{16}\right)t$$
$$+ \frac{\epsilon}{16}\log\frac{y(0) + [4 - y(0)]e^{-\epsilon t}}{4}$$
$$+ \frac{11\epsilon}{64}y^2(0)\frac{1 - e^{-\epsilon t}}{y(0) + [4 - y(0)]e^{-\epsilon t}}. \quad (3.13)$$

Here we see there is a shift in the frequency of the rapid phase together with a slow secular shift of the phase.

This discussion of the van der Pol equation is only intended to be illustrative of the method. We refer, for example, to a recent paper by Struble and Fletcher, who give a much more thorough discussion of the van der Pol equation using a somewhat different method.[6]

## 4. DEGENERATE PERTURBATION THEORY

In the degenerate case we must consider what changes must be made when there is a nontrivial set of integers satisfying (2.10). More generally, we must consider the situation when

$$|\mathbf{p} \cdot \boldsymbol{\omega}| < O(\epsilon). \qquad (4.1)$$

That is, the case when the factors in the denominators of our solutions, e.g., (2.13) or (2.16) or (2.18), are small of order $\epsilon$. When this occurs the successive terms in the series (2.2) are no longer small if $\epsilon$ is small; they no longer represent small amplitude fluctuations of the given motion about the mean motion. This so-called problem of small divisors is, of course, closely related to the degeneracy problem for which the divisors are zero.

The solution of this problem is already indicated by our discussion of Eq. (2.6) in the nondegenerate case. There we saw that the function $a(\mathbf{y})$ has to be chosen to cancel the terms in $A(\mathbf{y}, \boldsymbol{\phi})$ which correspond to solutions of the homogeneous equation (2.8). In the nondegenerate case, the only such term was the constant term, but in the degenerate or near degenerate case we must cancel all the terms corresponding to sets of integers satisfying (4.1). That is, we generalize to allow $a(\mathbf{y}, \boldsymbol{\phi})$ to depend upon those combinations of the $\phi_j$ which give rise to small divisors and then choose $a(\mathbf{y}, \boldsymbol{\phi})$ to cancel those terms in $A(\mathbf{y}, \boldsymbol{\phi})$.

Our procedure is formally similar to the nondegenerate case. We seek a solution of (2.1) in the form

$$x_i = y_i + \sum_{n=1}^{\infty} \epsilon^n F_i^{(n)}(\mathbf{y}, \boldsymbol{\phi}), \quad i = 1, 2, \cdots, r, \quad (4.2a)$$

$$\psi_j = \phi_j + \sum_{n=1}^{\infty} \epsilon^n G_j^{(n)}(\mathbf{y}, \boldsymbol{\phi}), \quad j = 1, 2, \cdots, s, \quad (4.2b)$$

where the $F_i^{(n)}(\mathbf{y}, \boldsymbol{\phi})$ and $G_j^{(n)}(\mathbf{y}, \boldsymbol{\phi})$ are periodic functions of each of the $\phi_k$. We further require that

$$\dot{y}_i = \sum_{n=1}^{\infty} \epsilon^n a_i^{(n)}(\mathbf{y}, \boldsymbol{\phi}), \quad i = 1, 2, \cdots, r, \quad (4.3a)$$

$$\dot{\phi}_j = \omega_j(\mathbf{y}) + \sum_{n=1}^{\infty} \epsilon^n b_j^{(n)}(\mathbf{y}, \boldsymbol{\phi}), \quad j = 1, 2, \cdots, s. \qquad (4.3b)$$

_____
[6] R. A. Struble and J. E. Fletcher, J. Math. Phys. **2**, 880 (1961). See also N. Minorsky, Ref. 2, pp. 219–224 and pp. 329–338.

Inserting (4.2) in (2.1) and using (4.3) we find, upon equating powers of $\epsilon$,

$$a_i^{(1)}(\mathbf{y}, \boldsymbol{\phi}) + \sum_{k=1}^{s} \omega_k(\mathbf{y}) \frac{\partial F_i^{(1)}(\mathbf{y}, \boldsymbol{\phi})}{\partial \phi_k} = A_i(\mathbf{y}, \boldsymbol{\phi}), \quad (4.4)$$

$$b_j^{(1)}(\mathbf{y}, \boldsymbol{\phi}) + \sum_{k=1}^{s} \omega_k(\mathbf{y}) \frac{\partial G_j^{(1)}(\mathbf{y}, \boldsymbol{\phi})}{\partial \phi_k}$$
$$= B_j(\mathbf{y}, \boldsymbol{\phi}) + \sum_{l=1}^{r} \frac{\partial \omega_j(\mathbf{y})}{\partial y_l} F_l^{(1)}(\mathbf{y}, \boldsymbol{\phi}), \quad (4.5)$$

and so on. The sequence of equations we obtain differs from that in the nondegenerate case only in that the $a_i^{(n)}(\mathbf{y}, \boldsymbol{\phi})$ and $b_j^{(n)}(\mathbf{y}, \boldsymbol{\phi})$ depend upon $\boldsymbol{\phi}$ as well as $\mathbf{y}$. The formal solution of these equations is straightforward. Using again the expansions (2.16), from (4.4) we obtain:

$$a_i^{(1)}(\mathbf{y}, \boldsymbol{\phi}) = \sum_{|\mathbf{p} \cdot \boldsymbol{\omega}| < O(\epsilon)} A_{i,\mathbf{p}}(\mathbf{y}) e^{i\mathbf{p} \cdot \boldsymbol{\phi}}, \qquad (4.6)$$

where the sum is over all sets of integers fulfilling (4.1), and

$$F_i^{(1)}(\mathbf{y}, \boldsymbol{\phi}) = -i \sum_{|\mathbf{p} \cdot \boldsymbol{\omega}| > O(\epsilon)} \frac{A_{i,\mathbf{p}}(\mathbf{y})}{\mathbf{p} \cdot \boldsymbol{\omega}} e^{i\mathbf{p} \cdot \boldsymbol{\phi}}, \qquad (4.7)$$

where the sum is over all sets of integers not contained in the sum in (4.6). Continuing, from (4.5) we obtain

$$b_j^{(1)}(\mathbf{y}, \boldsymbol{\phi}) = \sum_{|\mathbf{p} \cdot \boldsymbol{\omega}| < O(\epsilon)} B_{j,\mathbf{p}}(\mathbf{y}) e^{i\mathbf{p} \cdot \boldsymbol{\phi}} \qquad (4.8)$$

and

$$G_j^{(1)}(\mathbf{y}, \boldsymbol{\phi})$$
$$= -i \sum_{|\mathbf{p} \cdot \boldsymbol{\omega}| > O(\epsilon)} \left[ \frac{B_{j,\mathbf{p}}(\mathbf{y})}{\mathbf{p} \cdot \boldsymbol{\omega}} - i \frac{\sum_{l=1}^{r} \frac{\partial \omega_j(\mathbf{y})}{\partial y_l} A_{l,\mathbf{p}}(\mathbf{y})}{(\mathbf{p} \cdot \boldsymbol{\omega})^2} \right] e^{i\mathbf{p} \cdot \boldsymbol{\phi}}.$$
$$(4.9)$$

It should be clear that in this manner we can successively solve the equations for the determination of the $F_i^{(n)}(\mathbf{y}, \boldsymbol{\phi})$, $G_j^{(n)}(\mathbf{y}, \boldsymbol{\phi})$, $a_i^{(n)}(\mathbf{y}, \boldsymbol{\phi})$, and $b_j^{(n)}(\mathbf{y}, \boldsymbol{\phi})$ to obtain explicit expressions in which small divisors do not occur. Of course, the equations (4.3) for the determination of the secular motion are more complicated than the corresponding equations in the nondegenerate case; they explicitly involve certain combinations of the $\phi_k$. However, these equations still describe the slowly varying secular motion, since those combinations of the $\phi_k$ which do appear are themselves slowly varying. Thus, if the combination $(\mathbf{p} \cdot \boldsymbol{\phi})$ appears in (4.3), then

$$(\mathbf{p} \cdot \dot{\boldsymbol{\phi}}) \cong (\mathbf{p} \cdot \boldsymbol{\omega}) < O(\epsilon), \qquad (4.10)$$

i.e., this combination is slowly varying in exactly the

same sense that the $y_k$ are slowly varying. The basic idea of the expansion (4.2), or of (2.2), is the separation of the secular motion from the rapidly fluctuating motion, and this separation is preserved in the degenerate case.

There is, however, a serious difficulty in our general formulation of the degenerate perturbation problem. This is the question of deciding which combinations of the $\phi_k$ are to be included in the secular motion. For any particular set of $\omega_k$, we can always find a set of integers $p_k$ such that $(\mathbf{p} \cdot \boldsymbol{\omega})$ is as close as we please to any real number. That is, the values of $(\mathbf{p} \cdot \boldsymbol{\omega})$ are dense in the whole range

$$-\infty < \mathbf{p} \cdot \boldsymbol{\omega} < \infty. \qquad (4.11)$$

This means that we cannot, in general, make a sharp separation between the terms for which $|\mathbf{p} \cdot \boldsymbol{\omega}| < O(\epsilon)$, which we put into the secular motion, and the remaining terms, which we put into the fluctuating motion. We can do so only if the coefficients $A_{i,\mathbf{p}}(\mathbf{y})$ and $B_{j,\mathbf{p}}(\mathbf{y})$ vanish sufficiently rapidly for large values of $|\mathbf{p}| \equiv (p_1^2 + p_2^2 + \cdots + p_s^2)^{\frac{1}{2}}$. The point here is that, for the most general functions $A_i(\mathbf{x}, \boldsymbol{\phi})$ and $B_j(\mathbf{x}, \boldsymbol{\phi})$ in Eq. (2.1), it is not possible to sharply separate the secular motion from the periodic fluctuations; when these functions are such as to allow a sharp separation, the method we have outlined will work.

## 5. CONCLUSION

The method of rapidly rotating phase which we have presented here is applicable to a wide range of physical problems. On the one hand, it can be shown to be equivalent to classical perturbation theory of Hamiltonian systems, at least in the nondegenerate case.[7] On the other hand, Rayleigh–Schrödinger perturbation theory in quantum mechanics is also a special case. In both cases the treatment of degeneracy

or near-degeneracy is generally simplest by the method of rapidly rotating phase. Thus, the advantages of the method are its generality and its simplicity.

Of course, not all perturbation problems can be cast into the form of a set of coupled differential equations in the standard form (2.1) appropriate for the method. In general, we can say that the method is suited for the discussion of small perturbations of periodic or multiple-periodic motions, but we cannot precisely characterize such problems.

As we remarked earlier, an aspect of the method which may cause difficulty in applications is that the Eqs. (2.3) or (4.2), which describe the slow secular motion, may not be appreciably easier to solve than the original equations. (Here we are speaking of the finite versions of (2.3) or (4.2) which are obtained by truncating the series on the right.) The point is that the method is designed to separate the secular motion from the fluctuating motion; it gives no help in the discussion of the equations for the secular motion. This is a characteristic feature of all averaging methods.

We close with a few remarks about convergence. It should be clear that the method of rapidly rotating phase is asymptotic in the sense that the approximate solution is intended to be valid for long times, i.e., for times of order $\epsilon^{-1}$, the characteristic time of the secular motion. What can be proved is a typical asymptotic convergence theorem: With suitable restrictions on the perturbing functions, the approximate solutions obtained by solving the differential Eqs. (2.3) or (4.2), truncated at a finite order in $\epsilon$, and inserting the resulting solution in (2.2) or (4.1), also truncated, differ from the exact solution by an error which is small but which grows in time like $\exp\{c\epsilon t\}$, with $c$ a constant.[8] This is a rather weak theorem, but we have not been able to improve it in the general case. The question of the convergence of the infinite series in Eqs. (2.2), (2.3), (4.2), and (4.3) is still open.

---

[7] For a proof see T. P. Coffey, "Analytical Methods in the Theory of Non-Linear Oscillations," thesis, The University of Michigan, Ann Arbor, 1966.

[8] For a proof, see Ref. 7.