Institute of Mathematical Sciences

MAGNETO-FLUID DYNAMICS DIVISION

PHYSICS AND MATHEMATICS

NYO-7975

ON THE ASYMPTOTIC SERIES EXPANSION OF THE MOTION OF A CHARGED PARTICLE IN SLOWLY-VARYING FIELDS

by

J. Berkowitz and C. S. Gardner

December 23, 1957

NEW YORK UNIVERSITY

AEC Computing Facility
Institute of Mathematical Sciences
New York University

PHYSICS AND MATHEMATICS

NYO-7975

ON THE ASYMPTOTIC SERIES EXPANSION OF THE MOTION OF A CHARGED PARTICLE IN SLOWLY-VARYING FIELDS

by

J. Berkowitz and C. S. Gardner

December 23, 1957

Contract No. AT(30-1)-1480

NYO-7975
PHYSICS AND
MATHEMATICS

ABSTRACT

The asymptotic series expansion for the motion of a charged particle in a slowly-varying magnetic field has been determined by Kruskal. Kruskal has shown that the series *formally* satisfies the equations of motion. In this report it is proved that the series is in fact a valid asymptotic development of the exact solution.

TABLE OF CONTENTS

ON THE ASYMPTOTIC SERIES EXPANSION OF THE MOTION OF A CHARGED PARTICLE IN SLOWLY-VARYING FIELDS.

1. Introduction.

The equations of motion of a charged particle in an electromagnetic field may be written as follows:

$$\epsilon \ddot{\underline{R}} + \underline{B} \times \dot{\underline{R}} - \underline{E} = 0 \qquad (1.1)$$

Here $\underline{B}$, $\underline{E}$ are the magnetic and electric field vectors, and $\underline{R}$ is the position vector of the particle. The parameter ϵ is the ratio

$$\epsilon = m/e$$

of the mass m of the particle to its charge e. The dot signifies differentiation with respect to the time t.

Approximate formulas for the particle motion, i.e. for solutions of (1.1), were given by Alfvèn[1]. Alfvèn's formulas are valid if the field varies slowly. (The reference length and time used in Alfvèn's definition of what is meant by a "slow" variation are the radius and period of the circular gyration the particle would execute if the field were replaced

(1) H. Alfvèn, Cosmical Electrodynamics, Oxford, The Clarendon Press, 1953

by a constant field having the same instantaneous magnitude as the actual field. Thus Alfvèn's formulas are valid also if the field strength is large, or if the parameter ϵ is small.) Alfven's results have been derived in other ways and have been generalized, by a number of authors: Hellwig[2] gave the first few terms of an asymptotic series for the particle motion. Kruskal[3] has given the complete asymptotic series.

We assume that $\underline{B}$ and $\underline{E}$ are given as functions of the space coordinates $\underline{R}$, the time t, and the parameter ϵ[4]. We assume that $\underline{B}$, $\underline{E}$, and $\epsilon^{-1}(\underline{B} \cdot \underline{E})$ are continuous functions of $\underline{R}$, t, having continuous derivatives of all orders with respect to $\underline{R}$, t. We assume that any of the components of $\underline{B}$, $\underline{E}$, or $\epsilon^{-1}(\underline{B} \cdot \underline{E})$, or any of their derivatives with respect to $\underline{R}$, t is bounded in absolute value <u>uniformly</u> in $\underline{R}$, t, ϵ. Also, if $B = \sqrt{\underline{B} \cdot \underline{B}}$ is the magnetic field strength, we assume that B^{-1} is bounded uniformly in $\underline{R}$, t, ϵ. Thus B is assumed to have an absolute upper bound B_{max}, and also to have an absolute

---

(2) G. Hellwig, Über die Bewegung geladener Teilchen in schwach veränderlichen Magnetfeldern, Zeitschrift für Naturforschung <u>7</u>, 508-516, 1955.

(3) M. D. Kruskal, paper to be published in Nuovo Cimento.

(4) The notion of allowing $\underline{B}$, $\underline{E}$ to vary with ϵ is due to Kruskal

lower bound $B_{min} > 0$.

We are concerned with the behavior, as ϵ tends to zero, of the solution of the equations of motion (1.1) which satisfies the initial conditions

$$\text{for} \quad t = 0 \; ; \; \underline{R} = \overline{\underline{R}}, \; \underline{v} = \overline{\underline{v}} \tag{1.2}$$

Here $\overline{\underline{R}}$, $\overline{\underline{v}}$ are given <u>bounded</u> functions of ϵ.

Under the above conditions, Kruskal[3] has shown that a formal solution of the initial-value problem (1.1), (1.2) can be obtained by assuming for $\underline{R}(t;\epsilon)$ a series expansion of the following kind:

$$\underline{R} \sim \sum_{n=-\infty}^{\infty} \epsilon^{|n|} \, \underline{R}_n \, e^{in\,\theta/\epsilon} \tag{1.3}$$

Here $\theta(t,\epsilon)$ is determined (except for an additive constant) by

$$\dot{\theta} = B[\underline{R}_0, \; t; \; \epsilon] \tag{1.4}$$

The coefficients $\underline{R}_n$ are obtained as power series in ϵ:

$$\underline{R}_n(t,\epsilon) \sim \sum_{n=0}^{\infty} \underline{R}_{nk} \, \epsilon^k \tag{1.5}$$

The vectors $\underline{R}_{nk}$ may depend on ϵ as well as on t if $\underline{B}$ and $\underline{E}$ depend on ϵ. To first order in ϵ, the series

(1.3) represents the motion as a circular gyration, given by the terms $\epsilon \underline{R}_{+1} e^{i\theta/\epsilon}$, about a moving guiding center given by $\underline{R}_0$. The frequency of gyration is $\dot{\theta}/\epsilon$, which, by (1.4), is the cyclotron frequency evaluated at the guiding center.

The coefficients $\underline{R}_{nk}$ are obtained by inserting (1.3) into the differential equations (1.1) and initial conditions (1.2), computing as if the series did in fact converge, and as if $\underline{B}$ and $\underline{E}$ were analytic functions of R. The vector functions $\underline{B}$, $\underline{E}$ are expanded in Taylor series about $\underline{R}_0$. All terms multiplied by $e^{ik\theta/\epsilon}$ (for each k) are collected. The total coefficient of each exponential $e^{ik\theta/\epsilon}$ is set equal to zero. This procedure yields equations for the vectors $\underline{R}_n$. These equations are then solved by assuming the expansion (1.5) for each $\underline{R}_n$ and then solving the resulting equations -- some of which are finite equations, and some of which are differential equations. Kruskal[3] has shown that this procedure determines the coefficients $\underline{R}_{nk}$ uniquely, and in sequence. The $\underline{R}_{nk}$'s for small n, k are completely determined before the $\underline{R}_{nk}$ for larger values of n, k. In this sense, then, Kruskal has shown that (1.3) is a formal solution of the initial-value problem.

Our purpose in this paper is to show that (1.3) is in fact an asymptotic representation of the exact solution of the initial-value problem, valid as ϵ tends to zero.

2. Statement of the Results.

To state our results compactly it is convenient to use the notation "$O(\epsilon^n)$" -- to be read "a quantity of the order of magnitude of ϵ^n." To explain what this means, suppose that $f(t;\,\epsilon)$ is a quantity depending on t, ϵ, and we write

$$f = O(\epsilon^n)$$

Then we mean that given $t_1 > 0$ there is a positive number $A(t_1)$, depending on t_1, but not on ϵ, such that

$$\text{for} \quad 0 \leq t \leq t_1 \quad , \quad |f| < A\epsilon^n$$

It follows from Kruskal's analysis that if $N \geq 2$ is a positive integer, and $\underline{R}^{(N)}$ is the sum of a large enough, finite, number of terms of the series (1.3), that $\underline{R}^{(N)}$ approximately satisfies the equations of motion and initial conditions, in the following sense:

$$\epsilon\ddot{\underline{R}}^{(N)} + \underline{B}(\underline{R}^N,\, t;\, \epsilon)\times\dot{\underline{R}}^{(N)} - \underline{E}(\underline{R}^{(N)},\, t;\, \epsilon) = O(\epsilon^N) \qquad (2.1)$$

$$\left.\begin{aligned} \text{for} \quad t = 0, \quad \underline{R}^{(N)} &= \overline{\underline{R}} + O(\epsilon^N) \\ \dot{\underline{R}}^{(N)} &= \overline{\underline{v}} + \ (\epsilon^N) \end{aligned}\right\} \qquad (2.2)$$

The error estimates in (2.1), (2.2) may be obtained by using Taylor's theorem with remainder. We do not wish to repeat Kruskal's analysis here, so we shall assume that (2.1), (2.2) have been established. Assuming this, and assuming that

$\underline{B}$, $\underline{E}$, $\underline{\bar{R}}$, $\underline{\bar{v}}$ have the properties stated in the introduction, we shall prove the following theorems:

Theorem I: The initial-value problem stated by (1.1), (1.2) has exactly one solution, which is defined for all values of t.

Theorem II: This solution $\underline{R}$, and the approximation $\underline{R}^{(N)}$, and their derivatives, are bounded uniformly in ϵ. That is,

$$\underline{R} = O(1) \quad , \quad \dot{\underline{R}} = \underline{v} = O(1) \tag{2.4}$$

$$\underline{R}^{(N)} = O(1) \quad , \quad \dot{\underline{R}}^{(N)} = \underline{v}^{(N)} = O(1) \tag{2.5}$$

Our main result is to the effect that $\underline{R}^{(N)}$ not only approximately satisfies the differential equation (1.1) and initial conditions (1.2) but is indeed approximately equal to the exact solution $\underline{R}$. We state this result as follows:

Theorem III:

$$\underline{R} - \underline{R}^{(N)} = O(\epsilon^{N-1}) \tag{2.6}$$

$$\underline{v} - \underline{v}^{(N)} = O(\epsilon^{N-2}) \tag{2.7}$$

(when we say that a vector is $O(\epsilon^n)$ we mean that each component of the vector is $O(\epsilon^n)$.) Theorem III is a precise statement of what is meant by saying that Kruskal's series (1.3) is a valid asymptotic development of $\underline{R}(t; \epsilon)$ as ϵ tends to zero.

3. Proof of Theorems I, II.

Theorem I is not quite trivial, and it is necessary to be assured of its truth to know that it makes sense to talk about "the solution" of our initial-value problem.

We write (1.1) as a first-order system as follows:

$$\dot{\underline{R}} = \underline{v} \tag{3.1}$$

$$\dot{\underline{v}} = -\frac{1}{\epsilon}\,\underline{B}\times\underline{v} + \frac{1}{\epsilon}\,\underline{E} \tag{3.2}$$

The assumptions about the continuity and differentiability of $\underline{B}$, $\underline{E}$ guarantee that, for sufficiently small values of t, equations (3.1), (3.2) have a unique solution satisfying the initial conditions (1.2). Conveivably, however, this solution can fail to exist beyond a finite time. However, this can happen only if $\underline{R}$ or $\underline{v}$ become unbounded in a finite time. (That is, if $|\underline{R}|$ or $|\underline{v}|$ either tend to infinity, or oscillate with an amplitude which tends to infinity as t tends to some finite upper limit t_0). Thus, to show that the solution is defined for all t, it is sufficient to show that the solution is bounded in any finite time. We do this by proving Theorem II, according to which not only are $\underline{R}$, $\underline{v}$ bounded, but are bounded by bounds which do not depend on ϵ.

To prove Theorem II we transform (3.1), (3.2) by introducing in place of $\underline{v}$ the velocity $\underline{w}$ relative to a coordinate system moving with the "drift velocity" $\underline{v}_0$, given by

$$\underline{v}_0 = \frac{\underline{E} \times \underline{B}}{B^2} \tag{3.3}$$

Note that -- because, by assumption, B^{-1} is bounded uniformly in $\underline{R}$, t, ϵ -- $\underline{v}_0$ has all the continuity, differentiability, and boundedness properties which were assumed for $\underline{B}$, $\underline{E}$. We define $\underline{w}$, then, by

$$\underline{v} = \underline{v}_0 + \underline{w} \tag{3.4}$$

Then (3.2) becomes

$$\dot{\underline{w}} = -\frac{1}{\epsilon}\,\underline{B} \times w + \frac{1}{\epsilon}\,\frac{(\underline{B}\cdot\underline{E})}{B^2}\,\underline{B} - \dot{\underline{v}}_0 \tag{3.5}$$

The derivative of $\underline{v}_0$ can be expressed as a function of $\underline{w}$, $\underline{R}$, t (and ϵ) by

$$\dot{\underline{v}}_0 = [(\underline{v}_0 + \underline{w}) \cdot \underline{\nabla}]\,\underline{v}_0 + \frac{\partial}{\partial t}\,\underline{v}_0$$

where $\underline{\nabla}$ is the gradient operator with respect to $\underline{R}$. We now form the scalar product of $\underline{w}$ and (3.5), and obtain

$$\frac{d}{dt}\left(\frac{1}{2}\;w^2\right) = \frac{1}{\epsilon}\,\frac{(\underline{B}\cdot\underline{E})}{B^2}\,(\underline{w}\cdot\underline{B}) - \underline{w}\cdot\frac{\partial}{\partial t}\,\underline{v}_0 - \underline{w}\cdot\left\{[\underline{v}_0+\underline{w}]\cdot\underline{\nabla}\,\underline{v}_0\right\} \tag{3.6}$$

Here we have used d/dt, instead of the dot, to denote total time differentiation, and we have used the symbol w to denote the <u>length</u> of $\underline{w}$; that is,

$$w = \sqrt{\underline{w} \cdot \underline{w}}$$

Equation (3.6) is the equation for the rate of change of the kinetic energy of the particle, as defined in the moving coordinate system.

The important feature of (3.6) for our purposes is that the right-hand side, considered as a function of the components of $\underline{w}$, is a quadratic polynomial. The coefficients of this polynomial are not known, but they are known to be bounded. This follows from the assumptions stated in the introduction. (Here we make essential use of the assumption that $\epsilon^{-1}(\underline{E} \cdot \underline{B})$ is bounded.) Let $Q(\underline{w})$ denote the right-hand side of (3.6). Because $Q(\underline{w})$ is a quadratic polynomial with bounded coefficients, it follows that $Q(\underline{w})$ is bounded as follows:

$$|Q(\underline{w})| \leq A w^2 + C \tag{3.7}$$

where A, C are independent of $\underline{R}$, t, ϵ. Hence, by (3.6), we have

$$\frac{d}{dt}(w^2) \leq \left|\frac{d}{dt}(w^2)\right| \leq 2A\, w^2 + 2C$$

whence

$$\frac{d}{dt}\left(e^{-2At}\, w^2\right) \leq 2Ce^{-2At}$$

so that, on integrating, we obtain

$$w^2 \leq \overline{w^2} + \frac{C}{A}\,(e^{2At}-1) \tag{3.8}$$

where $\overline{w^2}$ means the initial value of w^2. But (3.8) implies

$$w^2 = O(1)$$

and thus

$$\underline{w} = O(1)$$

Hence

$$\underline{v} = \underline{v}_0 + \underline{w} = O(1) + O(1) = O(1) \tag{3.9}$$

and, since

$$\dot{\underline{R}} = \underline{v}$$

we have

$$\underline{R} = O(1) \tag{3.10}$$

The same arguments can be applied to the function $\underline{R}^{(N)}$ which satisfies (2.1), (2.2). The occurrence of an additional term which is $O(\epsilon^{N-1})$ $(N \geq 2)$ on the right-hand sides of some of the equations, does not affect the results. Thus we have

$$\underline{v}^{(N)} = O(1) \tag{3.11}$$

$$\underline{R}^{(N)} = O(1) \tag{3.12}$$

Having established (3.9) - (3.12) we have proved Theorem II, and at the same time, have completed the proof of Theorem I.

4. Some Properties of the Vector-Product Operator.

To prove Theorem III we shall have to make a further transformation of the equations of motion. We obtain this transformation by decomposing $\underline{w}$ into a sum of eigenvectors of the vector-product operator V defined by

$$V\underline{w} = i\underline{B} \times \underline{w} \tag{4.1}$$

(The factor i in (4.1) makes V hermitian). Therefore we insert here a discussion of some of the properties of V which we shall use.

By using (4.1), we can compute the iterates of V, as follows:

$$V^2\underline{w} = -\underline{B} \times (\underline{B} \times \underline{w}) = B^2\underline{w} - (\underline{B}\cdot\underline{w})\underline{B}$$

$$V^3\underline{w} = iB^2\underline{B} \times \underline{w} = B^2V\underline{w}$$

Thus V satisfies the characteristic equation

$$V^3 = B^2V \tag{4.2}$$

Thus the eigenvalues of V are 0, B, $-B$, or, for short, kB ($k = -1, 0, 1$). Any vector $\underline{y}$ can be written uniquely as a sum of three eigenvectors of V associated with the three eigenvalues kB. That is, for any $\underline{y}$ there is exactly one set of vectors $\underline{y}_k$ ($k = -1, 0, 1$) such that

$$\underline{y} = \underline{y}_1 + \underline{y}_0 + \underline{y}_{-1} = \sum_{k+ -1}^{1} \underline{y}_k \tag{4.3}$$

where

$$V \underline{y}_k = kB \underline{y}_k \tag{4.4}$$

The vector $\underline{y}_k$ is a linear function of y, given by

$$\underline{y}_k = P_k \underline{y} \tag{4.5}$$

The projection operators P_k are given by

$$P_0 = I - \frac{1}{B^2} V^2 \tag{4.6}$$

$$P_1 = \frac{1}{2B^2} (V^2+BV) \tag{4.7}$$

$$P_{-1} = \frac{1}{2B^2} (V^2-BV) \tag{4.8}$$

where I is the identity operator.

To prove these statements we note the following relations - easily verified by direct computation, if (4.2) is used to express V^3, V^4 in terms of V, V^2:

$$V\,P_k = P_k\,V = k\,B\,P_k \tag{4.9}$$

$$P_k P_\ell = 0 \quad \text{for} \quad k \neq \ell \tag{4.10}$$

$$P_0 + P_1 + P_{-1} = I \tag{4.11}$$

By (4.10), (4.11) we have

$$P_k{}^2 = P_k \sum_{\ell=-1}^{1} P_\ell = P_k\,I = P_k \tag{4.12}$$

Now (4.9), (4.11) show that

$$\underline{y} = \sum_{k=-1}^{1} P_k\,\underline{y} \tag{4.13}$$

is a decomposition of $\underline{y}$ of the type required. Conversely, the relations (4.4), (4.9) imply

$$0 = P_\ell[V\underline{y}_k - kB\underline{y}_k] = (\ell - k)BP_\ell\underline{y}_k$$

so that

$$P_\ell\underline{y}_k = 0 \quad \text{for} \quad \ell \neq k\ .$$

Consequently, it follows from (4.3), (4.11), that

$$P_k \underline{y} = P_k \underline{y}_k = \sum_{\ell=-1}^{1} P_\ell \underline{y}_k = \underline{y}_k$$

so that (4.13) is the unique decomposition of $\underline{y}$ into eigenvectors of V.

5. Transformation of the Equations of Motion.

We have seen that the vector $\underline{w}$ defined by (3.4) satisfies the differential equation

$$\dot{\underline{w}} = \frac{1}{\epsilon} V \underline{w} + \underline{f}(\underline{R}, t, \underline{w}; \epsilon) \tag{5.1}$$

The vector function $\underline{f}$ is, in detail, given by

$$\underline{f} = \frac{1}{\epsilon} \frac{(\underline{B} \cdot \underline{E})}{B^2} \underline{B} - \underline{w} \cdot \frac{\partial}{\partial t} \underline{v}_0 - (\underline{v}_0 + \underline{w}) \cdot \nabla \underline{v}_0 \tag{5.2}$$

We note that $\underline{f}$ is a continuous function of $\underline{R}$, t, $\underline{w}$ with continuous derivatives with respect to $\underline{R}$, t, $\underline{w}$. It follows, by Theorem II, that, for the values of $\underline{R}$, $\underline{w}$ defined by the solution of the initial-value problem, $\underline{f}$ and its partial derivatives with respect to $\underline{R}$, $\underline{t}$, $\underline{w}$ are all $O(1)$. The vector $\underline{R}$ satisfies the differential equation

$$\dot{\underline{R}} = \underline{v}_0(\underline{R}, t; \epsilon) + \underline{w} \tag{5.3}$$

The initial values of $\underline{R}$, $\underline{w}$ are given by

$$\text{for } t = 0, \quad \underline{R} = \overline{\underline{R}}$$

$$\underline{w} = \overline{\underline{w}} \qquad (5.4)$$

where

$$\overline{\underline{w}} = \overline{\underline{v}} - \frac{\underline{E} \times \underline{B}}{B^2}$$

where $\underline{E}$, $\underline{B}$ are evaluated at $t = 0$, $R = \overline{\underline{R}}$.

The equations of motion in the form (5.1), (5.3) are not well adapted to the study of the behaviour of the solution as ϵ tends to zero. The reason lies in the factor $^1/\epsilon$ occurring in the right-hand side of (5.1). A Lipschitz constant for (5.1) would be of the order of $^1/\epsilon$. Proceeding directly from (5.1) to estimate $\underline{R} - \underline{R}^{(N)}$ one would obtain an estimate which would tend to infinity as ϵ tended to zero, and which would hence be useless.

We avoid this difficulty by a transformation which begins by decomposing $\underline{w}$ into eigenvectors of V. We have

$$\underline{w} = \sum_{k=-1}^{1} \underline{w}_k \qquad (5.5)$$

where

$$\underline{w}_k = P_k \, \underline{w} \qquad (5.6)$$

$$V\underline{w}_k = kB\underline{w}_k \qquad (5.7)$$

Differentiation of (5.6) shows that

$$P_k \, \dot{\underline{w}} = \dot{\underline{w}}_k - \dot{P}_k \, \underline{w} \tag{5.8}$$

Thus if one operates on (5.1) with P_k one obtains

$$\dot{\underline{w}}_k - \frac{ikB}{\epsilon} \underline{w}_k = \underline{G}_k \tag{5.9}$$

where

$$\underline{G}_k = \dot{P}_k \, \underline{w} + P_k \, \underline{f}$$

We write G_k as a function of $\underline{w}_k$, $\underline{R}$, t, ϵ by writing

$$\dot{P}_k = (\underline{v}_0 + \underline{w}) \cdot \underline{\nabla} P_k + \frac{\partial}{\partial t} P_k$$

and then inserting $\sum\limits_{k=-1}^{1} \underline{w}_k$ for $\underline{w}$ everywhere.

The next step is to introduce new dependent variables $\underline{u}_k$ by the definition

$$\underline{u}_k = e^{-ik\Phi/\epsilon} \, \underline{w}_k \tag{5.10}$$

where the function $\Phi(t)$ is defined by

$$\Phi(t) = \int_0^t B(\underline{R}(t', \epsilon), t'; \epsilon) dt' \tag{5.11}$$

Then (5.9) becomes

$$\dot{u}_k = e^{-\frac{ik}{\epsilon}\Phi} \, \underline{G}_k \left(\sum \underline{u}_\ell \, e^{\frac{i\ell}{\epsilon}\Phi}, \underline{R}, t; \epsilon\right) \tag{5.12}$$

Now we make a further transformation. We introduce a new independent variable ϕ and a new dependent variable $T(\phi)$. The function $T(\phi)$ is defined as the <u>inverse function</u> of the function $\Phi(t)$ defined by (5.11). (This inverse function exists and is monotone increasing since because $B > B_{min}$, the function Φ in (5.11) is monotone increasing.) That is, $T(\phi)$ is defined by the relations

$$\begin{aligned} &\Phi\,(T(\phi)) = \phi \\ &T\,(\Phi(t)) = t \end{aligned} \tag{5.13}$$

Thus in (5.12) we can replace Φ by ϕ and t by T. By (5.11) we have

$$\frac{dT}{d\phi} = \frac{1}{B(R(T: \epsilon), T, \epsilon)} \tag{5.14}$$

By (5.12) we have, consequently,

$$\frac{d\underline{u}_k}{d\phi} = \frac{1}{B}\, e^{-ik\phi/\epsilon} \, \underline{G}_k \left(\sum \underline{u}_\ell e^{i\ell\phi/\epsilon}, \underline{R}, T; \epsilon\right) \tag{5.15}$$

and (5.3) becomes

$$\frac{d\underline{R}}{d\phi} = \frac{1}{B}\,\underline{v}_0(\underline{R},\ T;\ \epsilon) + \frac{1}{B}\sum \underline{u}_k e^{ik\phi/\epsilon} \tag{5.16}$$

Let $\underline{X}$ be a vector with a large number of components - (thirteen) - the components of $\underline{X}$ being the scalar T, the components of $\underline{R}$, and all the components of the vectors $\underline{u}_k$, arranged in some definite order. Then according to (5.14), (5.16), (5.15), the vector $\underline{X}$ satisfies a differential equation of the following form

$$\frac{d}{d\phi}\,\underline{X} = \underline{F}\,(\underline{X},\ \phi;\ \epsilon) \tag{5.17}$$

Note that the vector function $\underline{F}$ is a continuous function of the components of $\underline{X}$, having continuous partial derivatives with respect to the components of X. Furthermore, $\underline{F}$ and its partial derivatives are bounded <u>independently of ϵ</u> in the following sense: if $t_1 > 0$ there is a number $K(t_1)$ such that if

$$|\phi| < B\min t_1 \tag{5.18}$$

then

$$\left.\begin{aligned} |F| &< K \\ |F'| &< K \end{aligned}\right\} \tag{5.19}$$

where F' means any derivative of F with respect to a component of X, and the notation $|F|$ means the sum of the absolute values of the components of F. To see this, note that (5.14) and (5.18) imply

$$|T| < t,$$

so that (5.19) follows from Theorem II and the assumptions listed in the Introduction.

The vector $\underline{X}$ satisfies initial conditions, which can be inferred from (5.4), to the effect that

$$\text{for} \quad \phi = 0, \ \underline{X} = \overline{\underline{X}} \tag{5.20}$$

(including the condition

$$\text{for} \quad \phi = 0, \ T = 0.)$$

A similar transformation can be performed on the equations (2.1), (2.2) which govern the putative approximate solution $\underline{R}^{(N)}$. We define

$$\Phi^{(N)}(t) = \int_0^t B(\underline{R}^{(N)}(t', \epsilon); t'\epsilon)dt' \tag{5.21}$$

$$\underline{u}_k^{(N)} = e^{-ik\Phi^{(N)}/\epsilon} \, \underline{w}_k^{(N)} \tag{5.22}$$

$$\underline{w}_k^{(N)} = P_k \, \underline{w}^{(N)} \tag{5.23}$$

(Here P_k is regarded as a function of $\underline{B}$, which in (5.23) is a function of $\underline{R}^{(N)}$, t, ϵ.) Also, we define the function

$T^{(N)}(\phi)$ by

$$\Phi^{(N)}\ (T^{(N)}(\phi)) = \phi$$

$$T^{(N)}\ (\phi^{(N)}(t)) = t \tag{5.24}$$

We find that the corresponding vector $\underline{X}^{(N)}$ satisfies the differential equation

$$\frac{d}{d\phi}\,\underline{X}^{(N)} = \underline{F}\ (\underline{X}^{(N)},\ \phi;\ \epsilon) + \phi(\epsilon^{N-1}) \tag{5.25}$$

and the initial conditions

$$\text{For}\quad \phi = 0,\ \ \underline{X}^{(N)} = \overline{\underline{X}} + O(\epsilon^{N}) \tag{5.26}$$

The differential equations (5.17), (5.25) with the initial conditions (5.20), (5.26) will be used to complete the proof of Theorem III.

6. Proof of Theorem III.

It is clear that by the same reasons which were given for (5.19) that $K(t_1)$ can be found such that

$$\text{for}\quad |\phi| < B\min t_1 \tag{6.1}$$

we have

$$|\underline{F}| < K$$

$$|\underline{F}'| < K$$

if $\underline{F}$, $\underline{F}'$ are evaluated anywhere on the line segment joining $\underline{X}(\phi)$ and $\underline{X}^{(N)}(\phi)$. It follows, then, that

$$|\underline{F}(\underline{X}, \phi; \epsilon) - \underline{F}(\underline{X}^{(N)}, \phi; \epsilon)| \leq K |\underline{X} - \underline{X}^{(N)}| \qquad (6.2)$$

if (6.1) is satisfied. The inequality (6.2) states that $\underline{F}$ satisfies a Lipschitz condition uniformly in ϵ. It is now an easy matter to compare the vectors $\underline{X}(\phi)$, $\underline{X}^N(\phi)$.

To do this, we subtract the differential equation (5.25) for $\underline{X}^N$ from the equation (5.17) for $\underline{X}$; using (6.2) we obtain

$$\left|\frac{d}{d\phi}(\underline{X} - \underline{X}^{(N)})\right| \leq K |\underline{X} - \underline{X}^{(N)}| + 0(\epsilon^{N-1}) \qquad (6.3)$$

If we subtract the initial conditions (5.26) from the initial conditions (5.20), we obtain

$$\text{for } \phi = 0, \quad |\underline{X} - \underline{X}^{(N)}| = 0(\epsilon^N) \qquad (6.4)$$

Let us now integrate (6.3); we obtain

$$|\underline{X} - \underline{X}^N| - 0(\epsilon^N) \leq \left| \int_0^{\phi} \frac{d}{d\phi}(\underline{X} - \underline{X}^{(N)}|d\phi\right|$$

$$\leq \int_0^{\phi} \left|\frac{d}{d\phi}(\underline{X} - \underline{X}^{(N)})\right| d\phi \leq K \int_0^{\phi} |X - X^{(N)}| d\phi + 0(\epsilon^{N-1}) \qquad (6.5)$$

If H is defined by

$$H = \int_0^{\phi} |\underline{X} - \underline{X}^{(N)}| \, d\phi$$

then (6.5) implies

$$\frac{dH}{d\phi} \leq KH + O(\epsilon^{N-1}) \tag{6.6}$$

Noting that

for $\phi = 0$, $H = 0$

we conclude from (6.6), as in a similar calculation we have done before, that

$$H = O(\epsilon^{N-1})$$

and hence, by (6.5), we have

$$\underline{X} - \underline{X}^N = O(\epsilon^{N-1}) \tag{6.7}$$

Now in (6.7), the vectors $\underline{X}$, $\underline{X}^{(N)}$ are compared at the same value of ϕ. What we must do to obtain Theorem III is to compare R, $R^{(N)}$ and $\underline{u}$, $\underline{u}^{(N)}$ at the same values of $\underline{t}$. We now procede to make this comparison.

We have, if $\phi = \Phi^{(N)}(t)$,

$$|\underline{R}(t) - \underline{R}^{(N)}(t)| = |\underline{R}(T^{(N)}(\phi)) - \underline{R}^{(N)}(T^{(N)}(\phi))| \tag{6.8}$$

and

$$|\underline{R}(T^{(N)}(\emptyset)) - \underline{R}^{(N)}(T^{(N)}(\emptyset))| \leq |\underline{R}(T(\emptyset)) - \underline{R}^{(N)}(T^{(N)}(\emptyset))|$$

$$+ |\underline{R}(T(\emptyset)) - \underline{R}(T^{(N)}(\emptyset))| \qquad (6.9)$$

Now (6.7) implies, inter alia, that

$$|\underline{R}(T(\emptyset)) - \underline{R}^{(N)}(T^{(N)}(\emptyset))| = O(\epsilon^{N-1}) \qquad (6.10)$$

and

$$|T(\emptyset) - T^{(N)}(\emptyset)| = O(\epsilon^{N-1}) \qquad (6.11)$$

By (5.3) we see that

$$\dot{R} = O(1)$$

and hence, using (6.11), we see that

$$|\underline{R}(T(\emptyset)) - \underline{R}\ (T^{(N)}(\emptyset))| \leq O(1) \cdot O(\epsilon^{N-1})$$

so that

$$\underline{R}\ (T(\emptyset) - \underline{R}\ (T^{(N)}(\emptyset)) = O(\epsilon^{N-1}) \qquad (6.12)$$

Combining (6.8), (6.9), (6.10), (6.12) we conclude that

$$\underline{R} - \underline{R}^{(N)} = O(\epsilon^{N-1}) \qquad (6.13)$$

Likewise, noting by (5.12) that

$$\dot{\underline{u}}_k = O(1)$$

we obtain

$$\underline{u}_k - \underline{u}_k^{(N)} = O(\epsilon^{N-1}) \tag{6.14}$$

By (6.13) and the definitions (5.11), (5.21) of Φ, $\Phi^{(N)}$, we see that

$$\Phi - \Phi^{(N)} = O(\epsilon^{N-1})$$

Hence, by the definitions of $\underline{u}_k$, $\underline{u}_k^N$, we obtain (using Theorem II)

$$|\underline{w}_k - \underline{w}_k^{(N)}| \leq |e^{ik\Phi/\epsilon}| \quad |\underline{u}_k - \underline{u}_k^{(N)}|$$

$$+ |\underline{u}_k^N| \quad |e^{ik\Phi/\epsilon} - e^{ik\Phi^{(N)}/\epsilon}|$$

$$\leq O\ (\epsilon^{N-1}) + O(1) \cdot \frac{k}{\epsilon} \cdot O(\epsilon^{N-1}) = O(\epsilon^{N-2}) \ . \tag{6.15}$$

Adding the inequalities (6.15), we have, since

$$\underline{w} = \sum w_k, \ w^{(N)} = \sum w_k^{(N)}$$

the result

$$\underline{w} - \underline{w}^{(N)} = O(\epsilon^{N-2})$$

whence, using (6.13), we see that

$$\underline{v} - \underline{v}^{(N)} = O(\epsilon^{N-2}) \tag{6.16}$$

Now that (6.13) and (6.16) have been established, Theorem III is proved, and our task is finished.