APPLICATION OF A NON LINEAR WKB METHOD TO THE
KORTEWEG-DEVRIES EQUATION*

ROBERT M. MIURA† AND MARTIN D. KRUSKAL‡

Abstract. The WKB method used in quantum mechanics for solving linear second order ordinary
differential equations is generalized to apply to nonlinear partial differential equations. In particular,
this nonlinear WKB method, which is similar to the averaging method due to Whitham, is used to
study nearly-periodic solutions of the Korteweg–deVries equation when the dispersion parameter
is small. The emphasis of this paper is on a detailed analysis of the leading-order problem arising from
the application of the nonlinear WKB method. An explicit representation of the leading-order solution
is obtained in terms of unknown functions whose qualitative properties are studied. These unknown
functions are governed by a first order system of nonlinear partial differential equations which is of
hyperbolic type.

1. Introduction. One effective (though pedestrian) method for studying general
properties of solutions of nonlinear partial differential equations is to use the
computer. This method is not only useful for obtaining quantitative properties
of the solution but it is also useful as a synergetic tool in the mathematical analyses
of such equations (see Zabusky [1]). One such equation which has been studied
by this method and which has been shown to describe a large number of different
physical systems which exhibit nonlinear dispersive wave phenomena [2]–[11]
is the Korteweg–deVries equation (KdV for short)

\[ u_t + uu_x + \delta^2 u_{xxx} = 0, \]  

where \( \delta^2 \ll 1 \) is a small (dispersion) parameter depending on various physical
quantities. There is, however, one practical problem in obtaining computer-
calculated solutions of this equation, as well as of other equations which describe
dispersive phenomena; this is that as \( \delta^2 \) takes on smaller and smaller values,
oscillations in the solutions become more and more closely spaced and it becomes
necessary to introduce additional space steps to accurately describe the solution.
Coupled with such factors as numerical accuracy and stability, this places a severe
restriction on the total evolution time of the solution that can be computed.

To overcome this difficulty, Kruskal and Zabusky [12] (Kruskal [5], Zabusky
[1]) used the fact that the solutions become nearly periodic for small \( \delta^2 \) and
developed a method for treating such nonlinear waves. By “nearly periodic” is
meant that the changes in amplitude and wavelength are \( O(\delta) \) over a few oscilla-
tions in the solutions. (This should not be confused with almost periodic functions.)
This method is an asymptotic one which arises as a natural nonlinear generaliza-
tion of the usual linear WKB method and is similar to the averaging method of
Whitham [13]–[15]. (This method has not been called an averaging method
here since although averaging plays a role, it is by no means all of it.)

* Received by the editors April 27, 1972. This work was supported by the U.S. Air Force Office
of Scientific Research under Contracts AFOSR-71-2053 and F44620-70-C-0033.
† Courant Institute of Mathematical Sciences, New York University, New York, New York.
Now at Department of Mathematics, Vanderbilt University, Nashville, Tennessee 37235.
‡ Princeton Plasma Physics Laboratory and Program in Applied Mathematics, Princeton
University, Princeton, New Jersey 08540.

376
As presented here, the application of the method to the KdV equation is mainly formal and a number of very basic problems remain open. The main emphasis of the present paper will be a detailed discussion of the leading-order problem.

In § 2, a short general summary of the method is given and the formulation for the KdV equation is presented in § 3. The leading-order problem is treated in detail in § 4 and the dependence of the leading-order solution on the angle-like variable is obtained explicitly. Sections 5 and 6 study qualitative and quantitative properties of two functions which arise in the leading-order solution. In § 7, we show that the system of partial differential equations governing the averaged variables introduced in § 4 is hyperbolic. This classification appears to be intimately connected with the requirement that the asymptotic solution as obtained by this method be periodic in the angle-like variable.

2. The nonlinear WKB method. As is well known, the linear WKB method used in quantum mechanics obtains approximations to “nearly periodic” solutions of the linear ordinary differential equation

\[ \delta^2 \psi_{xx} + V(x) \psi = 0, \quad \delta^2 \ll 1, \]

where \( V(x) \) varies little over a few oscillations in the solution \( \psi \). The WKB method assumes a solution in the form

\[ \psi \sim W(x) e^{i \theta}, \quad \theta \equiv B(x)/\delta, \]

where \( W \) and \( B \) are to be determined. Having introduced two functions \( W \) and \( B \) in place of the single function \( \psi \) we have the freedom to specify one arbitrary condition between them.

These ideas of the linear WKB method can now be generalized to treat nearly periodic solutions (in \( x \)) of nonlinear partial differential equations. Because of the nonlinearity, we cannot expect the simple exponential form of (2.2) to suffice; therefore, we assume the formal series representation

\[ u(x, t; \delta) \sim U(\theta, x, t; \delta) = U^{(0)}(\theta, x, t) + \delta U^{(1)}(\theta, x, t) + \cdots, \]

where \( \theta = \theta(x, t; \delta) \). Here the dependence of \( U \) explicitly on \( x \) and \( t \) is analogous to the dependence of \( W \) on \( x \) and the \( \theta \)-dependence generalizes the exponential in (2.2). One important unsolved problem is to establish in what sense the series expansion \( U(\theta, x, t; \delta) \) on the right approximates the actual solution \( u(x, t; \delta) \) on the left; hence the meaning of the symbol \( \sim \) in (2.3) is unclear.

The series representation of the solution given by (2.3) is so general that it does not appear to lead to any major simplifications. However, we retain two properties from the linear WKB method—periodicity and one arbitrary condition. We assume that \( U \) is strictly periodic in \( \theta \), i.e., that \( \theta \) is an “angle-like” variable, and, since both \( U \) and \( \theta \) have been introduced, we choose the one free condition such that the period in \( \theta \) is unity; thus

\[ U(\theta, x, t; \delta) = U(\theta + 1, x, t; \delta). \]

We have a final difficulty; whereas in the linear WKB method the dependence of \( \psi \) on \( \theta \) was specified, i.e., exponential, here the dependence of \( U \) on \( \theta \) must be
determined. Later we see how \( \theta \) depends on \( x \) and \( t \). This has been called the “method of extension” by Sandri [16] and is basically the method of multiple scales. To illustrate this method, consider Fig. 1, where we have sketched a solution surface \( U(\theta, x) \) in the two variables \( x \) and \( \theta \) (for simplicity in this illustration we suppress the dependence of \( U \) on \( t \)). One of the objectives of the method is to determine \( \theta(x) \) so that the solution \( u(x) \) to the original problem will correspond to the curve obtained by the intersection of the \( U(\theta, x) \) surface and the \( \theta(x) \) surface.

3. Application to the Korteweg-deVries equation. To apply the nonlinear WKB method to the KdV equation, we first select the dependence of \( \theta \) on \( \delta \) using the “principle of minimal simplification” proposed by Kruskal [17]. This principle states that the dependence of \( \theta \) on \( \delta \) should be chosen in such a way that to leading order there is minimal simplification in the complexity of the equation.

For the KdV equation all terms can be made equally important so that the nonlinear and dispersive terms must balance; thus we assume

\[
(3.1) \quad \theta = B(x, t; \delta)/\delta,
\]

where \( B = O(1) \) is a formal power series in \( \delta \). Hence \( \theta \) varies rapidly for finite changes in \( B \).

Our objectives, therefore, are to obtain \( U(\theta, x, t; \delta) \) and \( B(x, t; \delta) \) as formal power series in \( \delta \). Again we remark that we will only obtain formal power series for \( U \) and \( B \), and it remains unsettled whether or in what sense the series \( U(B/\delta, x, t; \delta) \) approximates the solution of the original problem.
3.1. Extended equation. We complete the formulation of the extended problem by deriving the governing equation for $U$. Since only partial derivatives of $B$ will appear in our equation, reflecting the arbitrariness in the phase shift of the angle variable $\theta$, we define

$$L \equiv B_t, \quad K \equiv B_x, \quad l \equiv \frac{B_t}{B_x} = \frac{L}{K},$$

which are functions dependent only on $x$ and $t$ (and $\delta$). The partial differential operators become

$$\frac{\partial}{\partial t} \rightarrow \frac{L}{\delta} \frac{\partial}{\partial \theta} + \frac{\partial}{\partial t},$$

$$\frac{\partial}{\partial x} \rightarrow \frac{K}{\delta} \frac{\partial}{\partial \theta} + \frac{\partial}{\partial x}.$$  

Using (3.3) in (1.1) yields the “extended KdV equation”

$$l U_\theta + U U_\theta + K^2 U_{\theta\theta\theta} + \delta \left[ \frac{1}{K} (U_t + U U_x) + 3(KU)_x \right]$$

$$+ \delta^2 \left\{ \frac{1}{K} \left[ K_{xx} U + 3(KU)_x \right] \right\}_\theta + \delta^3 \frac{1}{K} U_{xxx} = 0.$$  

Note that the leading-order part involves only differentiations with respect to $\theta$ with coefficients $l$ and $K$ which depend on $x$ and $t$.

3.2. Averaged equations. As noted above, if we substitute the series (2.3) into the extended equation (3.4) and keep only the leading-order terms, we obtain an ordinary differential equation in $\theta$ for $U^{(l)}$,

$$l^{(0)} U_\theta^{(0)} + U^{(0)} U_\theta^{(0)} + K^{(0)^2} U_{\theta\theta\theta}^{(0)} = 0,$$

where $l^{(0)}$ and $K^{(0)}$ are the leading-order terms in formal expansions of $l$ and $K$. (Note that the effect of the dispersive third derivative term is retained at leading order and is not lost as is the effect of the dissipative second derivative term in the Burgers equation when the viscosity approaches zero.) Integrating (3.5) yields

$$l^{(0)} U^{(0)} + \frac{1}{2} U^{(0)^2} + K^{(0)^2} U_{\theta\theta}^{(0)} = m,$$

and after multiplying by $U_\theta^{(0)}$, a second integration gives

$$\frac{1}{2} l^{(0)} U^{(0)^2} + \frac{1}{6} U^{(0)^3} + \frac{1}{2} K^{(0)^2} U_\theta^{(0)^2} = m U^{(0)} + n,$$

where $m$ and $n$ depend only on $x$ and $t$.

If $k \geq 1$, we obtain linear ordinary differential equations for $U^{(k)}$ with inhomogeneous terms involving lower order solutions and coefficients,

$$\mathcal{L} U^{(k)} = M(U^{(0)}, \ldots, U^{(k-1)}, l^{(0)}, \ldots, l^{(k-1)}, K^{(0)}, \ldots, K^{(k-1)}),$$

where

$$\mathcal{L} \equiv l^{(0)} \frac{\partial}{\partial \theta} + \frac{\partial}{\partial \theta} \left( U^{(0)} + K^{(0)^2} \frac{\partial^3}{\partial \theta^3} \right).$$

The dependence of $U^{(k)}$ on $x$ and $t$ is determined from compatibility conditions for the existence of solutions to the inhomogeneous equation (3.8). For $k \geq 1$,
these require the inhomogeneous term $M$ to be orthogonal to the null space of the adjoint operator $\mathcal{L}^*$ given by

\begin{equation}
\mathcal{L}^* \equiv I^{(0)} \frac{\partial}{\partial \theta} + U^{(0)} \frac{\partial}{\partial \theta} + K^{(0)} \frac{\partial^3}{\partial \theta^3}.
\end{equation}

Two functions that span the null space of periodic solutions are $1$ and $U^{(0)}$, which lead to two averaged solutions (see § 4.2). The variation of constants method yields a third linearly independent solution which is not periodic; dropping superscripts,

\begin{equation}
V(\theta) = U(\theta) \int_{\theta_0}^{\theta} \frac{d\theta'}{U_d(\theta')^2} - \int_{\theta_0}^{\theta} \frac{U(\theta') d\theta'}{U_d(\theta')^2},
\end{equation}

where $\theta_0$ is arbitrary. (The integrals over $U_d^{-2}$ through any zero of $U_d$ should be interpreted as the finite part since by extension into the complex plane the residue of $U_d^{-2}$ is zero wherever $U_d = 0$ because by (3.5), $U_{d00} = 0$ there. Hence the value of the integral is the same whether the path goes above or below the pole.) If $V(\theta)$ were periodic with period $p > 0$, then the periodicity conditions $0 = V(\theta_0 + p)$ and $0 = V_d(\theta_0) = V_d(\theta_0 + p)$ would require

\begin{equation}
\int_{\theta_0}^{\theta_0 + p} \frac{d\theta'}{U_d(\theta')^2} = 0, \quad \int_{\theta_0}^{\theta_0 + p} \frac{U(\theta') d\theta'}{U_d(\theta')^2} = 0,
\end{equation}

where we choose $\theta_0$ such that $U_d(\theta_0) \neq 0$. Multiplying (3.6) by $U_d^{-2}$ and integrating over a period requires

\begin{equation}
\int_{\theta_0}^{\theta_0 + p} \frac{U(\theta')^2 d\theta}{U_d(\theta')^2} = 0.
\end{equation}

Finally, multiplying (3.6) by $UU_d^{-2}$ and (3.7) by $U_d^{-2}$, integrating over a period, and then eliminating $\int_{\theta_0}^{\theta_0 + p} d\theta' U(\theta')^2 [U_d(\theta')]^{-2}$ between them yields $p = 0$, a contradiction.

An alternative approach to derive averaged equations is to average conservation laws. This is by no means as general since the nonlinear WKB method is applicable to equations which possess weak dissipation, e.g., which may have only one conservation law. Also, since the KdV equation possesses an infinite sequence of conservation laws (see Miura, Gardner, and Kruskal [18]), we should prove that the averaged equations lead to only a finite number of independent equations. This last point remains an open problem.

A conservation law has the form

\begin{equation}
T_t + X_x = 0,
\end{equation}

where for the KdV equation we consider only $T$ and $X$ which are polynomials in $u$ and its $x$ derivatives. In the extended variables, these conservation laws take the form

\begin{equation}
T_t + X_x + \frac{1}{\delta}(LT + KX)_\theta = 0,
\end{equation}

where $T$ and $X$ are polynomials in $U, K$, and their derivatives. Integrating over a period in $\theta$, we obtain the averaged equations

\begin{equation}
\langle T \rangle_t + \langle X \rangle_x = 0,
\end{equation}

where $\langle F \rangle \equiv \int_0^1 F(\theta) d\theta$. 
For the KdV equation, the conservation laws [18] are obtainable from

\begin{equation}
 w_t + \left( \frac{1}{2} w^2 + \frac{\varepsilon^2}{18} w^3 + \delta^2 w_{xx} \right)_x = 0,
\end{equation}

(3.14)

\begin{equation}
 w = u - i\varepsilon \delta w_x - \frac{\varepsilon^2}{6} w^2.
\end{equation}

(3.15)

Each conservation law is obtained by solving (3.15) for \( W \) as a formal power series in \( \varepsilon \), inserting the result into (3.14), and then setting the coefficient of some power of \( \varepsilon \) equal to zero. Introducing a new variable \( W(\theta, x, t; \delta; \varepsilon) \), the averaged equations (3.13) become

\begin{equation}
 \left< W \right>_t + \left< \frac{1}{2} W^2 + \frac{\varepsilon^2}{18} W^3 + \delta^2 W_{xx} \right>_x = 0,
\end{equation}

(3.16)

where \( W = U - i\varepsilon \delta [W_x +(1/\delta)(KW)_\theta] - \varepsilon^2 W^2/6 \). We shall find that, at least to leading order, three averaged equations yield a determined system.

The problem of going to higher order in \( \delta \) has not been studied here and hence will not be discussed. A discussion of this problem has been given by Luke [19] for the nonlinear Klein–Gordon equation and a general second order variational equation. The asymptoticity of the solutions has not been proved.

In the remainder of this paper, we concern ourselves with a detailed investigation of the leading-order solution. As we proceed there will be many changes of dependent variables; thus we shall attempt to avoid confusion by stating explicitly with which variables we are dealing. We started with \( u(x, t; \delta) \) and went to the extended variables \( U(\theta, x, t; \delta) \) and \( B(x, t; \delta) \). Now we have changed to the variables \( U(\theta, x, t; \delta), K(x, t; \delta) \), and \( l(x, t; \delta) \).


4.1. Leading-order solution in \( \theta \). In this section we shall derive an explicit solution for \( U^{(0)} \) as a function of \( \theta \). With the assumption that \( U \) can be expressed by (2.3) as a series in \( \delta \), we find that the \( \delta \)-dependence of \( U^{(0)} \) is governed by (3.5),

\begin{equation}
 lU_{\theta} +UU_{\theta} + K^2 U_{\theta\theta} = 0,
\end{equation}

(4.1)

where for convenience we have omitted all superscripts; therefore, \( K, l, \) and \( U \) in (4.1) are now all zero order quantities. Two successive integrations yield (3.6) and (3.7) which we rewrite without superscripts,

\begin{equation}
 lU + \frac{1}{2} U^2 + K^2 U_{\theta\theta} = m,
\end{equation}

(4.2)

\begin{equation}
 \frac{1}{2} lU^2 + \frac{1}{6} U^3 + \frac{1}{4} K^2 U^2_{\theta} = mU + n.
\end{equation}

(4.3)

Requiring \( U \) to be periodic in \( \theta \) imposes restrictions on \( K, l, m, \) and \( n \). One can see these restrictions intuitively in Fig. 2 which is a schematic graph of \( K^2 U_{\theta}^2 \) as a function of \( U \) obtained from (4.3). The first condition for periodic \( U \) is the existence of a local minimum and maximum, i.e., that the two roots of

\begin{equation}
 \frac{1}{2} U^2 + lU - m = 0,
\end{equation}

(4.4)

given by

\begin{equation}
 U_{1,2} = -l \mp \sqrt{l^2 + 2m},
\end{equation}

(4.5)
be real. This requires
(4.6) \[ A^2 \equiv t^2 + 2m \geq 0, \quad (A \geq 0). \]

A second (stronger) condition is that there exist three real roots of
(4.7) \[ -H(U) \equiv \frac{1}{6}U^3 + \frac{1}{2}tU^2 - mU - n = 0 \]
giving the inequality
(4.8) \[ H(U_1) \leq 0 \leq H(U_2), \]
where \( U_1 \) and \( U_2 \) given by (4.5) determine the local minimum and maximum respectively. Define the quantity \( \alpha \) by
(4.9) \[ A^3 \alpha \equiv -2(\frac{1}{6}t^3 + ml - n); \]
then (4.8) requires
(4.10) \[ -\frac{2}{3} \leq \alpha \leq \frac{2}{3}. \]
The final condition for periodic \( U \) is that \( U \) be restricted to lie between \( U_b \) and \( U_c \) which correspond to two roots of (4.7).

If we also define a new variable
(4.11) \[ z \equiv -(U + l)/A, \]
then (4.3) can be rewritten in the integrated form
(4.12) \[ \theta - \theta_0 = K \sqrt{\frac{3}{A}} \int_{a(x)}^{b(x)} \frac{dz'}{\sqrt{z^3 - 3z' + 3x}}, \quad a(x) \leq z \leq b(x), \]
where \( \theta_0 \) is an integration constant. The roots of the cubic under the square root are designated by \( a(x) \leq b(x) \leq c(x) \) and correspond to \( U_a, U_b, \) and \( U_c \). We can convert the right side of (4.12) into an elliptic integral of the first kind by defining
(4.13) \[ x = \operatorname{sn} \Theta = \sin \phi \equiv \left( \frac{z - a}{b - a} \right)^{1/2}, \]
where \( sn \) is a Jacobi elliptic function, with the result

\[
\Theta = \frac{1}{2K} \sqrt{\frac{(c - a)A}{3}} (\theta - \theta_0) = \int_0^\phi \frac{d\phi'}{\sqrt{1 - k^2 \sin^2 \phi'}}
\]

\[
= \int_0^{sn \phi} \frac{dx'}{(1 - x'^2)(1 - k^2 x'^2)},
\]

where

\[
k^2 \equiv \frac{b - a}{c - a}.
\]

Imposing the requirement that the period be unity, from (4.15) we obtain the condition

\[
K = \frac{1}{4} \sqrt{\frac{(c - a)A}{3}} [\mathcal{K}(k)]^{-1},
\]

where \( \mathcal{K} \) is the complete elliptic integral of the first kind. Then from (4.11), (4.13), (4.15) and (4.17), we easily obtain the solution

\[
U(\theta, x, t) = -l - aA - (b - a)A \text{sn}^2[2\mathcal{K}(k)(\theta - \theta_0)][k],
\]

where all quantities on the right are functions of \( l, m, \) and \( n \) which in turn depend only on \( x \) and \( t \). These \( (l, m, n) \) are our three new variables.

Since \( \alpha \) is given explicitly as a function of \( l, m, \) and \( n \) from (4.6) and (4.9), it remains to determine the dependence of \( a, b, \) and \( c \) on \( \alpha \). We can solve the cubic equation \( z^3 - 3z + 3\alpha = 0 \) by recognizing that \(-2 \leq a \leq b \leq c \leq 2\) and \(-2 \leq z \leq 2\); thus letting \( z = 2\sin \psi \), the cubic becomes

\[
4 \sin^3 \psi - 3 \sin \psi = -\frac{3}{2} \alpha,
\]

which is a trigonometric identity if

\[
\alpha = \frac{3}{2} \sin (3\psi).
\]

Restricting \( \psi \) to \(-\pi/6 \leq \psi \leq \pi/6\), the roots of the cubic are given by

\[
a = -2 \sin (\pi/3 + \psi), \quad -2 \leq a \leq -1,
\]

\[
b = 2 \sin \psi, \quad -1 \leq b \leq 1,
\]

\[
c = 2 \sin (\pi/3 - \psi), \quad 1 \leq c \leq 2,
\]

where

\[
\psi = \frac{1}{3} \sin^{-1}(\frac{3}{2} \alpha).
\]

We have completely obtained the \( \theta \)-dependence of the zero order solution. The coefficients in (4.18) depend only on \( l, m, \) and \( n \) and it remains to determine their dependence on \( x \) and \( t \). Note that \( K \) is known in terms of \( l, m, \) and \( n \) by virtue of the periodicity condition (4.17).

### 4.2. Leading-order averaged equations

Introduce the averaged variables

\[
P \equiv \langle U \rangle, \quad Q \equiv \langle U^2 \rangle, \quad R \equiv \langle U^3 \rangle.
\]
Then the two compatibility equations referred to after (3.9) and the third averaged conservation law become
\begin{align}
(4.26) & \quad P_t + \frac{1}{2} Q_x = 0, \\
(4.27) & \quad Q_t + (\frac{3}{2} R - 3 K^2 \langle U^2 \rangle_x) = 0, \\
(4.28) & \quad (R - 3 K^2 \langle U^2 \rangle_x) + \left(\frac{3}{8} \langle U^4 \rangle - 12 K^2 \langle U U^2 \rangle + 9 K^2 \langle U^2 \rangle_x \right)_x = 0.
\end{align}

We shall now show that $m$, $n$, and the additional unknown averaged quantities in (4.27) and (4.28) can be eliminated in favor of the variables $l$, $P$, $Q$, and $R$. Then we shall show that $l$ is a function of $P$, $Q$, and $R$, and hence, that the above three equations form a closed system for the new variables $P$, $Q$, and $R$.

To accomplish this we make extensive use of the leading-order equations (4.2) and (4.3). Averaging (4.2) gives
\begin{equation}
(4.29) \quad m = lP + \frac{1}{2} Q.
\end{equation}

To evaluate $K^2 \langle U^2 \rangle_x$, we multiply (4.2) by $U$, with $m$ replaced by (4.29), and average
\begin{equation}
(4.30) \quad K^2 \langle U^2 \rangle_x = \frac{1}{4} R - lP^2 + lQ - \frac{1}{2} PQ,
\end{equation}
where by periodicity $\langle U U_{th} \rangle = -\langle U^2 \rangle$. Using this in the average of (4.3) gives
\begin{equation}
(4.31) \quad n = \frac{3}{14} R + lQ - \frac{3}{2} lP^2 - \frac{1}{2} PQ.
\end{equation}

The quantities $\langle U^4 \rangle$ and $K^2 \langle U U^2 \rangle$ are determined by deriving two linearly independent equations for them. One equation is the average of (4.2) multiplied by $U^2$ and the other is the average of (4.3) multiplied by $U$. The results are
\begin{align}
(4.32) & \quad \langle U^4 \rangle = \frac{1}{4} (-18 lR + 54 lPQ + 15 Q^2 - 36 lP^3 - 18 P^2 Q + 10 PR), \\
(4.33) & \quad K^2 \langle U U^2 \rangle = \frac{1}{14} (-2 lR + 20 lPQ + 4 Q^2 - 18 lP^3 - 9 P^2 Q + 5 PR).
\end{align}

We obtain $K^4 \langle U^2 \rangle_x$ from the average of (4.2) multiplied by $U_{th}$:
\begin{equation}
(4.34) \quad K^4 \langle U^2 \rangle_x = -l^2 P^2 + \frac{1}{14} lPQ + l^2 Q + \frac{3}{14} lR + \frac{5}{14} Q^2
\end{equation}
\begin{equation}
- \frac{9}{4} lP^3 - \frac{9}{14} P^2 Q + \frac{5}{14} PR.
\end{equation}

Then (4.27) and (4.28) become
\begin{align}
(4.35) & \quad Q_t + \left(\frac{3}{2} R + 3 lP^2 + \frac{3}{2} PQ - 3 lQ\right)_x = 0, \\
(4.36) & \quad \left(-\frac{3}{8} R + lP^2 + \frac{1}{2} PQ - lQ\right)_t
\end{align}
\begin{equation}
+ \left(lR - lPQ + \frac{1}{2} Q^2 - 3 lP^2 + 3 l^2 Q\right)_x = 0.
\end{equation}

Note that $K$ has automatically been eliminated from these equations.

To give a detailed study of (4.26), (4.35), and (4.36), it is convenient at this point to define the new variables
\begin{align}
(4.37) & \quad q \equiv \langle (U - P)^2 \rangle = Q - P^2 \geq 0, \\
(4.38) & \quad r \equiv -\frac{1}{6} \langle (U - P)^3 \rangle - \frac{1}{3} \langle (U - P)^2 \rangle
\end{align}
\begin{equation}
= -\frac{1}{6} R - lQ - \frac{1}{2} PQ + lP^2 + \frac{3}{2} P^3,
\end{equation}
where
\begin{equation}
(4.39) \quad \lambda \equiv P + l.
\end{equation}
The variables \( q \) and \( r \) are defined relative to the mean so that when no oscillations are present they are both zero. The particular form for \( r \) was chosen to avoid any time derivatives of \( l \) resulting in \( 4.36 \), and \( \lambda \) is introduced since the combination \( P + l \) often appears. We have, therefore, gone from the set of three variables \( l, m, \) and \( n \) to the new set of four variables \( \lambda, P, q \) and \( r \). Now we shall show that in fact \( \lambda \) is a function of \( q \) and \( r \) and that it does not depend on \( P \). Then with these new variables we shall derive a system of three first order quasi-linear partial differential equations for \( P, q, \) and \( r \).

4.3. Implicit equations for \( \lambda \). In terms of \( \lambda, q, \) and \( r \) (note that \( P \) is not included) we rewrite \((4.6)\) and \((4.9)\) as

\[
A^2 = \lambda^2 + q,
\]

\[
A^3 \alpha = -(\frac{\lambda^3}{3} + 4q \lambda + 5r). \tag{4.41}
\]

Eliminating \( A \) gives an algebraic equation for determining \( \lambda \) as a function of \( q, r, \) and \( \alpha \). It remains to obtain one more such relation to eliminate \( \alpha \).

This third relationship is obtained by integrating \((4.18)\) over a period in \( \theta \). This gives

\[
\frac{\lambda}{A} = -a - \frac{b - a}{2k^2} \int_0^{2\pi} \text{sn}^2 u \, du = -a - \frac{b - a}{2k^2} \left[ u_0^2 - 2E(\text{am} u_0) \right]
\]

\[
= -c + (c - a) \frac{\epsilon}{\mu} \equiv s(\lambda), \tag{4.42}
\]

where \( E \) is an elliptic integral of the second kind and \( \epsilon \) is the corresponding complete elliptic integral of the second kind. We have used the fact that \( \epsilon = E(\text{am} \mu) \). Therefore, \((4.40), (4.41), \) and \((4.42)\) are three implicit equations for \( \lambda \) as a function of \( q \) and \( r \). In \S\ 5 we study properties of \( s(\lambda) \) and in \S\ 6 we prove there exists a unique \( \lambda \) for all \( q > 0 \) and any \( r \).

4.4. Determined system for \( P, q, \) and \( r \). Having established the dependence of \( \lambda \) on only \( q \) and \( r \), we can rewrite \((4.26), (4.35), \) and \((4.36)\) as the following determined system of first order quasi-linear partial differential equations in \( P, q, \) and \( r \):

\[
P_t + PP_x + \frac{1}{2}q_x = 0, \tag{4.43}
\]

\[
q_t + 2qP_x + (P + 2\lambda + 2q\lambda_q)q_x + (5 + 2q\lambda_s)r_x = 0, \tag{4.44}
\]

\[
r_t + (2q\lambda + 6r)P_x + [-3\lambda^2 - 6(q\lambda + r)\lambda_q]q_x + [P - 6\lambda - 6(q\lambda + r)\lambda_s]r_x = 0. \tag{4.45}
\]

The derivatives \( \lambda_q \) and \( \lambda_s \) can be eliminated in favor of \( \lambda, q, \) and \( r \) (see Appendix). This system of equations is invariant under change of sign of both \( x \) and \( t \) and under Galilean transformation. We prove that this system of equations is hyperbolic in \( \S \ 7 \).

It is clear that if we solve the implicit equations for \( \lambda(q, r) \) and the partial differential equations for \( P, q, \) and \( r \) subject to appropriate initial and boundary conditions, then we have the explicit zero order solution \( U^{(0)}(\theta, x, t) \). Also, we obtain \( U^{(0)}(x, t) \) from \( K \) and \( L \) which are given respectively by \((4.17)\) and

\[
L = K(\lambda - P). \tag{4.46}
\]
5. **Properties of** \( s(\alpha) \). Although (4.42) can also be used to obtain qualitative properties of \( s(\alpha) \), it is somewhat easier to obtain these properties by a study of its derivative. To this end we give a different derivation of (4.42) from which we can derive a first order nonlinear ordinary differential equation for \( s(\alpha) \).

Define the two integrals

\[
F(\alpha) \equiv \int_a^b \sqrt{\frac{1}{3} z^3 - z + \alpha} \, dz,
\]

\[
G(\alpha) \equiv \int_a^b z \sqrt{\frac{1}{3} z^3 - z + \alpha} \, dz.
\]

Then we obtain the relations

\[
F'(\alpha) = \frac{1}{2} \int_a^b \frac{dz}{\sqrt{\frac{1}{3} z^3 - z + \alpha}} = \frac{1}{4} \frac{\sqrt{A}}{K},
\]

\[
G'(\alpha) = \frac{1}{2} \int_a^b \frac{z \, dz}{\sqrt{\frac{1}{3} z^3 - z + \alpha}} = \frac{1}{4} \frac{\sqrt{\lambda}}{K \sqrt{A}},
\]

where the first relation is simply (4.17) and the second is obtained by use of (4.11). Thus we see that eliminating \( K \) yields (4.42):

\[
- \frac{G'(\alpha)}{F'(\alpha)} = \frac{\lambda}{A} = s(\alpha).
\]

To derive the desired differential equation for \( s(\alpha) \), we differentiate with respect to \( \alpha \):

\[
\frac{ds}{d\alpha} = \frac{G'F'' - F'G''}{(F')^2}.
\]

We now evaluate \( F'' \) and \( G'' \) in terms of \( \alpha, F', \) and \( G' \). From (5.1), multiplying and dividing by the square root and integrating by parts appropriately, we obtain

\[
3\alpha F' - 2G' = \frac{3}{2} F,
\]

where we have used (5.3) and (5.4). Similarly, from (5.2) we obtain

\[
3\alpha G' - 2F' = \frac{3}{2} G.
\]

Differentiating (5.7) and (5.8) and solving for \( F'' \) and \( G'' \) yields for (5.6),

\[
\frac{ds}{d\alpha} = \frac{1 + 3\alpha s + s^2}{9\alpha^2 - 4},
\]

after using (5.5).

A study of the integral curves of \( s(\alpha) \) requires a certain amount of detail which we omit. One such calculation is to show that

\[
\frac{ds}{d\alpha} (-\frac{3}{2}) = -\frac{1}{4}.
\]

The solution curve is shown in Fig. 3 and is indicated by the solid curve.

Finally, we prove that this curve is unique. Define

\[
z \equiv s + \frac{1}{3} \alpha.
\]

We omit an easy geometrical proof that \( z \geq 0 \). Then (5.9) becomes

\[
\frac{dz}{d\alpha} = \frac{5}{4} + \frac{z^2}{9\alpha^2 - 4}.
\]
If \( z_1 \) and \( z_2 \) are two distinct solution curves satisfying the end conditions and \( z_2 \geq z_1 \geq 0 \), then

\[
\frac{d(z_2 - z_1)}{dx} = \frac{z_2^2 - z_1^2}{9x^2 - 4}.
\]

For \(|x| \leq \frac{1}{3}\) we have \(9x^2 - 4 \leq 0\); thus the right side of (5.13) is nonpositive, which contradicts our original hypothesis that \( z_2 \geq z_1 \), hence \( z_1 = z_2 \).

6. Existence and uniqueness of \( \lambda(q, r) \). In this section we prove the existence and uniqueness of \( \lambda \) for all \( q > 0 \) and any \( r \). We showed in § 4.3 that \( \lambda \) is to be determined from the equations

\[
A^2 = \lambda^2 + q, \quad A > 0, \quad q > 0,
\]

\[
A^3 \dot{x} = -(\frac{1}{3} \lambda^3 + 4q \lambda + 5r),
\]

\[
s(\dot{x}) = \lambda/A,
\]

\[
\frac{ds}{d\lambda} = \frac{1 + 3xs + s^2}{9x^2 - 4}, \quad s\left(\pm \frac{2}{3}\right) = \pm 1, \quad \frac{ds}{d\lambda} \left| \frac{-2}{3}\right) = -\frac{1}{4},
\]

where for convenience we have replaced the explicit solution (4.42) by the differential equation.

6.1. Simple existence proof. First, we give a geometrical proof of the existence of \( \lambda \) for all \( q > 0 \) and any \( r \). This is based on the simple idea that \( \lambda \) exists if and only if the graphs of each side of (6.3) intersect in the domain of interest.
Specifically, consider \( q \) and \( r \) as parameters, and treat \( \lambda \) as an independent variable. Define the right side of (6.3) by

\[
\bar{s}(\lambda; q) \equiv \frac{\lambda}{\sqrt{\lambda^2 + q}}.
\]

For any fixed \( q > 0 \), \( \bar{s} \) is clearly a monotone increasing function of \( \lambda \) with \( \bar{s}(\pm \infty) = \pm 1 \). The graph of \( \bar{s} \) is shown in Fig. 4.

To prove \( \alpha(q, r) \) exists, we now show for fixed \( q > 0 \) and fixed \( r \) that the graph of \( s(\lambda; q, r) \) obtained from (6.1), (6.2), and (6.4) intersects the graph of \( \bar{s}(\lambda; q) \) at least once for \(-1 \leq s \leq 1\). Given \( q \) and \( r \), \( \alpha \) is a function of \( \lambda \) alone which we denote by

\[
\bar{\alpha}(\lambda; q, r) = -\frac{\frac{3}{2} \lambda^3 + 4q\lambda + 5r}{(\lambda^2 + q)^{3/2}}.
\]

The resulting \( \bar{\alpha}(\lambda; q, r) \) is shown in Fig. 5. In the domain of interest, \(-\frac{3}{2} \leq \alpha \leq \frac{3}{2}\), we see that \( \bar{\alpha} \) decreases monotonically; therefore, \( s(\bar{\alpha}) \) increases monotonically.
from $-1$ to $1$. Moreover, this increase occurs over the finite domain $\lambda_1 \leq \lambda \leq \lambda_2$. The resulting curve for $s(\tilde{\lambda})$ is shown in Fig. 6. Superimposing $s(\tilde{\lambda})$ onto Fig. 4 yields the obvious inequalities

$$-1 = s(\tilde{\lambda}(\lambda_1) = \frac{3}{2}) < \tilde{s}(\lambda_1) < s(\tilde{\lambda}(\lambda_2) = -\frac{3}{2}) = 1,$$

which by continuity of $\tilde{s}$ and $s$ proves that there exists a solution $\lambda = \lambda(q, r)$ for all $q > 0$ and any $r$.

### 6.2. Existence and uniqueness of $\lambda(q, r)$.

We shall now prove both existence and uniqueness of $\lambda$ for given $q > 0$ and $r$ using a different method. Define new variables

$$\bar{\lambda} \equiv \lambda/\sqrt{q}, \quad \bar{r} \equiv r/q^{3/2}.$$

After elimination of $A$, (6.2) and (6.3) become

$$\left(\bar{\lambda}^2 + 1\right)^{3/2} \bar{z} = -\frac{3}{2}\bar{\lambda}^3 - 4\bar{\lambda} - 5\bar{r},$$

$$s(\bar{z}) = \frac{\bar{\lambda}}{\sqrt{\bar{\lambda}^2 + 1}} = \bar{s}(\bar{\lambda})$$

which are independent of $q$ and involve only $\bar{\lambda}$, $\bar{r}$, and $x$. We now treat $\bar{r}$ as the independent variable which is equivalent to treating $q$ and $r$ as independent variables.

Our present goal is to derive a differential equation for $\bar{\lambda} = \bar{\lambda}(\bar{r})$. Differentiating $s$ with respect to $\bar{r}$ and solving for $d\bar{\lambda}/d\bar{r}$ gives

$$\frac{d\bar{\lambda}}{d\bar{r}} = \frac{1}{ds/d\lambda - (ds/dx)(\partial \lambda/\partial r)},$$

where $ds/dx$ and $ds/d\lambda$ are to be computed from (6.4) and (6.10) respectively. The indicated differentiations are easily carried out and after some algebra we arrive at the equation

$$\frac{d\bar{\lambda}}{d\bar{r}} = \frac{3\bar{\lambda}(3\bar{\lambda} + 5\bar{r}) - 1}{3(2\bar{\lambda} + 3\bar{r})(3\bar{\lambda} + 5\bar{r})}.$$

![Fig. 6](image-url)
We omit the details of the qualitative properties of the integral curves and simply summarize some of the results in Fig. 7. In the figure, we have designated various regions as $\lambda^2 + \tilde{r}^2 \to \infty$ by Roman numerals. Then the integral curves can be classified according to the regions in which they lie for large $\lambda^2 + \tilde{r}^2$. These five types of integral curves are:

1. I–III,
2. II–III,
3. V–VII,
4. VI–VII,
5. III–VII.

Fig. 7
The only integral curves which become multivalued with respect to \( \hat{\lambda}(\hat{r}) \) are of type 5. A simple study of the asymptotic behavior of the solution curve near \( \alpha = -\frac{1}{2} \) shows that it is of type 3.

Although all integral curves of type 3 are single-valued, it remains to prove that a unique integral curve corresponds to the solution curve. We prove this by showing that for \( \hat{\lambda} = 0 \), there corresponds a unique \( \hat{r} \). From (6.9) and (6.10) we have

\[
\alpha = -5\hat{r}, \quad s(\alpha) = 0.
\]

There corresponds a unique \( \alpha \) for \( s = 0 \) and hence we obtain a unique \( \hat{r} \). Then the one-to-one correspondence between the solution curve \( \hat{\lambda}(\hat{r}) \) and a unique integral curve in the \( (\hat{r}, \hat{\lambda}) \)-plane follows by the Cauchy–Lipschitz theorem for an autonomous system which states that through every point in the plane there passes one and only one integral curve.

7. Classification of the system of partial differential equations. Our main objective in this section will be to show that the system of first order quasi-linear partial differential equations (4.43), (4.44), and (4.45) is hyperbolic. Whitham [13] has derived an equivalent system of three averaged equations for the KdV equation and has shown it to be hyperbolic.

If we denote the characteristic roots of the system by \( \mu \) and define \( \gamma \equiv P - \mu \), then the roots are obtained from the vanishing of the characteristic determinant

\[
\begin{vmatrix}
\gamma & \frac{1}{2} & 0 \\
2q & \gamma + 2\lambda + 2q\lambda & 5 + 2q\lambda_r \\
2q\lambda + 6r & -3\lambda^2 - 6q\lambda + r \lambda & \gamma - 6\lambda - 6q\lambda + r \lambda_r
\end{vmatrix} = 0.
\]

In the Appendix we derive expressions for \( \lambda_q \) and \( \lambda_r \) in terms of \( \lambda, q, \) and \( r \) with the results

\[
\lambda_q = \frac{\lambda^2(3q\lambda + 5r) + \frac{1}{2}qr}{(2q\lambda + 3r)(3q\lambda + 5r)},
\]

(7.2)

\[
\lambda_r = \frac{3\lambda(3q\lambda + 5r) - q^2}{3(2q\lambda + 3r)(3q\lambda + 5r)}.
\]

(7.3)

We recognize (7.3) as the nonnormalized form of (6.12). Substituting these into (7.1) yields a cubic equation for \( \gamma \):

\[
\gamma^3 + \left( -6\lambda + \frac{q^2}{3q\lambda + 5r} \right) \gamma^2 + \left( 9\lambda^2 - \frac{2q^2\lambda}{3q\lambda + 5r} \right) \gamma
\]

\[
+ 15(q\lambda + r) - \frac{4q^3}{3(3q\lambda + 5r)} = 0.
\]

(7.4)

For simplicity, we rewrite this as

\[
\gamma^3 + 3a\gamma^2 + 3b\gamma + c = 0,
\]

where

\[
a \equiv -2\lambda + \frac{q^2}{3(3q\lambda + 5r)}.
\]

(7.5)
\( b \equiv 3\lambda^2 - \frac{2q^2\lambda}{3(3q\lambda + 5r)} \),
\( c \equiv 15(q\lambda + r) - \frac{4q^3}{3(3q\lambda + 5r)} \).

We transform this into standard form by defining
\( \rho \equiv \gamma + a, \)
\( d \equiv a^2 - b, \)
\( e \equiv 2a^3 - 3ab + c, \)
giving
\( \rho^3 - 3d\rho + e = 0. \)

We now specify inequalities such that the cubic equation (7.12) should possess three real roots. We then determine if these inequalities are satisfied by our functions \( \lambda, q, \) and \( r. \) The limiting cases are, of course, determined by the boundary separating the real roots from complex conjugate roots, i.e., where two real roots coalesce. This occurs for
\( \rho = \pm \sqrt{d} \equiv \pm f. \)
But then (7.12) yields
\( e = \pm 2f^3. \)

Therefore, in order that (7.12) have all real roots, the coefficients \( d \) and \( e \) must satisfy the inequalities (found from extreme cases)
\( d \geq 0, \)
\( \left| \frac{e}{2f^3} \right| \leq 1. \)

For notational simplicity, define
\( g \equiv \frac{q^2}{3(3q\lambda + 5r)}, \quad h \equiv 15(q\lambda + r). \)

Our analysis in § 6.2 and omitted calculations show that the solution curve in the \((\bar{r}, \bar{\lambda})\)-plane satisfies the inequalities
\( 3q\lambda + 5r < 0, \quad q\lambda + r < 0; \)
therefore,
\( g < 0, \quad h < 0. \)

We write (7.6), (7.7), and (7.8) as
\( a = -2\lambda + g, \quad b = 3\lambda^2 - 2\lambda g, \quad c = h - 4qg. \)

Thus (7.10) becomes
\( d = \lambda^2 - 2\lambda g + g^2 = (\lambda - g)^2 \geq 0, \)
which satisfies the inequality (7.15).
Using the definition of \( e \) given by (7.11), the ratio in (7.16) becomes

\[
\frac{e}{2a^{3/2}} = 1 + \frac{g(3\lambda - 2g)^2 + h - 4qg}{2(\lambda - g)^3},
\]

where without loss of generality, we have taken the positive root for \( f \). The inequality (7.16) then becomes

\[
2 \leq \frac{g(3\lambda - 2g)^2 + h - 4qg}{2(\lambda - g)^3} \leq 0.
\]

We now show that this inequality is satisfied for all \( \lambda, q, \) and \( r \). From the definitions (7.17) we can write \( \lambda \) in terms of \( q, g, \) and \( h \):

\[
\lambda = \frac{1}{6} \left( \frac{h}{q} - \frac{q}{g} \right).
\]

Then the numerator in (7.23) becomes

\[
g(3\lambda - 2g)^2 + h - 4qg = \frac{1}{q^2} \left( 2qq^2 - \frac{1}{2} q^2 - \frac{1}{2} hq \right)^2 \leq 0.
\]

The denominator in (7.23) is

\[
(\lambda - g)^3 = g^3 \left( \frac{\lambda}{g} - 1 \right)^3,
\]

and

\[
\frac{\lambda}{g} - 1 = \frac{1}{q^2} [3q(3q\lambda + 5r) - q^2],
\]

which from (7.3) is negative because the solution curve \( \lambda(q, r) \) satisfies \( \partial \lambda / \partial r < 0 \), as can be shown from (6.12), and \( (2q\lambda + 3r)(3q\lambda + 5r) > 0 \), by averaging (7.18). Then since \( g < 0 \), the right inequality in (7.23) is satisfied.

From the positivity of the denominator in (7.23), the left inequality is equivalent to

\[
4\lambda^3 - 3\lambda^2 g + h - 4qg \geq 0,
\]

or in terms of \( \lambda, q, \) and \( r \) with the constraint \( 3q\lambda + 5r < 0 \), this becomes

\[
36q\lambda^4 + 60r\lambda^3 + 132q^2\lambda^2 + 360qr\lambda + 225r^2 - 4q^3 \leq 0.
\]

To prove that this inequality is satisfied, we square both sides of (6.2), having eliminated \( A \) by (6.1), and use the periodicity condition (4.10), with the resulting inequality

\[
4(\lambda^2 + q)^3 \geq 9(\lambda^3 + 4q\lambda + 5r)^2,
\]

which is precisely (7.29). This concludes the proof of hyperbolicity of the system of three first order quasi-linear partial differential equations.

It is undoubtedly not merely a coincidence that the inequality (7.30) is the same for periodicity of the solution in \( \theta \) and for hyperbolicity of the system of averaged equations. However, it is unclear as to the precise connection and as to what the relationship would be for a general problem.
Appendix. Evaluation of $\lambda_q$ and $\lambda_r$. In proving the hyperbolicity of the system of equations (4.43), (4.44), and (4.45), it is convenient to eliminate the derivatives $\lambda_q$ and $\lambda_r$ in favor of the variables $\lambda$, $q$, and $r$. To accomplish this we use (6.1), (6.2), (6.3), and (6.4). Treat $q$ and $r$ as independent variables. Then differentiating (6.3) with respect to $q$ and $r$ yields for $\lambda_q$ and $\lambda_r$,

$$\lambda_q = -\frac{s_2 \alpha A^3 + \frac{1}{2} \lambda}{s_2 \alpha A^3 - q},$$

$$\lambda_r = -\frac{s_2 \alpha A^3}{s_2 \alpha A^3 - q},$$

where we have used (6.1).

For simplicity, introduce the notation

$$B \equiv 2q\lambda + 3r, \quad C \equiv 3q\lambda + 5r.$$  

We now evaluate partial derivatives of $\alpha$ with respect to $\lambda$, $q$, and $r$ using (6.2) where $A$ is eliminated using (6.1) with the results

$$\alpha_{\lambda} = \frac{-4qA^2 + 5\lambda B}{A^2},$$

$$\alpha_q = \frac{-3\lambda A^2 + \frac{5}{2} B}{A^2},$$

$$\alpha_r = -\frac{5}{A^2}.$$  

Eliminating $\alpha$ on the right side of (6.4) using (6.2) and (6.3) yields

$$s_2 = \frac{qA^4 - 5\lambda A^2 B}{D},$$

where $D \equiv -4qA^4 + 20\lambda A^2 B + 25B^2$. Thus we obtain

$$\lambda_q = \frac{-\lambda q^2 A^2 + \frac{1}{2} q B + 5\lambda^2 B}{3BC},$$

$$\lambda_r = -\frac{qA^2 + 5\lambda B}{3BC},$$

or written out in full,

$$\lambda_q = \frac{\lambda^2 (3q\lambda + 5r) + \frac{1}{2} qr}{(2q\lambda + 3r)(3q\lambda + 5r)},$$

$$\lambda_r = \frac{3\lambda (3q\lambda + 5r) - q^2}{5(2q\lambda + 3r)(3q\lambda + 5r)}.$$

Acknowledgment. One of the authors (RMM) wishes to thank J. B. Keller and J. K. Moser for helpful comments.
REFERENCES


