COMMUNICATIONS ON PURE AND APPLIED MATHEMATICS, VOL. XXVIII, 141–188 (1975)

Periodic Solutions of the KdV Equation*

PETER D. LAX

Dedicated to Arne Beurling

Abstract

In this paper we construct a large family of special solutions of the KdV equation which are periodic in x and almost periodic in t. These solutions lie on N-dimensional tori; very likely they are dense among all solutions.

The special solutions are characterized variationally; they minimize $F_N(u)$, subject to the constraints $F_j(u)=A_j$, $j=-1,\cdots,N-1$; here F_j denote the remarkable sequence of conserved functionals discovered by Kruskal and Zabusky. The above minimum problem was originally suggested by them. In exploring the manifold of solutions of this minimum problem we make essential use of Gardner's discovery that these functionals are in involution with respect to a suitable Poisson bracket.

Gardner, Greene, Kruskal and Miura have shown that the eigenvalues of the Schrödinger operator are conserved functionals if the potential is a function of t and satisfies the KdV equation. In Section 6 a new set of conserved quantities is constructed which serve as a link between the eigenvalues of the Schrödinger operator and the F_j.

Another result in Section 6 is a slight sharpening of an earlier result of the author and J. Moser: for the special solutions constructed above all but $2N+1$ eigenvalues of the Schrödinger operator are double.

The simplest class of special solutions, $N=1$, are cnoidal waves. In an appendix, M. Hyman describes the results of computing numerically the next simplest case, $N=2$. These calculations show that the shape of these solutions recurs exactly after a finite time, in a shifted position. The theory verifies this fact.

1. Introduction

The equation

$$u_t+uu_x+u_{xxx}=0 \tag{1.1}$$

is called after its discoverers the Korteweg–de Vries equation (see [12]), abbreviated as KdV; here u is a real-valued scalar variable, t is time and x a single spatial variable. The initial value problem is to relate solutions of (1.1) to their initial values, i.e., their values at $t=0$. Recently, the initial value problem has been studied extensively for sufficiently differentiable solutions which tend to zero rapidly as $|x|$ tends to ∞. It is easy to show (see

* Results obtained at the Courant Institute of Mathematical Sciences, New York University, under Contract AT(11-1)-3077 with the U.S. Atomic Energy Commission. Reproduction in whole or in part is permitted for any purpose of the United States Government.

e.g. [13], p. 467) that solutions in this class are uniquely determined by their initial values. It has been further shown (see e.g. [20] or [15]) that the initial values of u may be prescribed arbitrarily in this class.

In 1965, Kruskal and Zabusky [25] devised a numerical scheme which they used to construct and study solutions of the KdV equation. They observed that after the elapse of some time each solution contained a number of waves, each of which propagates with a positive speed and without altering its shape. They identified these waves as the *solitary waves* already discovered by Korteweg and deVries, i.e., waves which are zero at $x=\pm\infty$ and which propagate with constant speed and unaltered shape. The culmination of these investigations was [7], in which the initial value problem for KdV was reduced to the inverse problem in scattering theory for potentials vanishing at $x=\pm\infty$. This led to a procedure for constructing solutions of the KdV equation in terms of their initial data which involved solving the Gelfand–Levitan equation. The resulting expression turned out to be the perfect tool for studying the long-time behavior of solutions (see [1], [21], and [8]).

The initial value problem for solutions which are periodic in x turns out to be more difficult than for solutions which vanish at $x=\pm\infty$. Here too the initial steps were taken by Kruskal and Zabusky, who observed in their numerical calculations that solutions with sinusoidal initial values develop for a while into a rather complicated wave pattern, but that at certain times the waves recombine into a reasonable semblance of the initial shape; this led them to suspect that these *solutions of* KdV *are almost periodic in time.*

In this paper we construct a large class of solutions of KdV which are periodic in x and almost periodic in t. I believe that the initial values of these special solutions are dense among all periodic initial values, and I further believe that all spatial periodic solutions of KdV are almost periodic in time.

The last section contains new information on the relation of the KdV equation to the Schrödinger equation.

In their recent study, Henry McKean and Pierre van Moerbeke [17] have constructed explicit formulas for the special class of solutions constructed in this paper, and for the KdV flow on this class of solutions. Their formulas involve hyperelliptic functions.

I learned from Prof. Fomenko at the Vancouver International Congress of Mathematicians that S. Novikov and Dubrovin have done similar work on this subject (see [19]).

Acknowledgement. I acknowledge with pleasure many stimulating conversations with Jürgen Moser.

This paper is dedicated to Arne Beurling, in recognition of his influence in the development of inverse problems of spectral theory.

2. Conserved Functionals of Equations of Evolution

Consider a nonlinear evolution equation

$$u_t = K(u) . \tag{2.1}$$

We assume that solutions of (2.1) exist for all t, positive and negative, are uniquely determined by their initial values, and that their initial values can be prescribed arbitrarily within a certain set in a linear space. Solutions of (2.1) induce a *flow* on this set. Let $F(u)$ be a *functional* which is *conserved* under the flow (2.1), i.e., one for which $F(u(t))$ is independent of t when $u(t)$ is a solution of (2.1). If F is a conserved functional, the sets $F(u) = c$ are invariant sets for the flow. These sets have codimension one and so are rather large; we show now how to construct with the aid of F much smaller invariant sets.

THEOREM 2.1. *Let F be a conserved functional of* (2.1) *which is Frechet differentiable. Then the set of stationary points of F forms an invariant set for the flow* (2.1).

Remark 1. The result is intuitively clear for stationary points which are local extremals, i.e., which are maxima or minima, since the flow carries a local maximum into a local maximum.

Remark 2. Take the very special case of linear, unitary flows in a Hilbert space, i.e., one governed by an equation of the form

$$u_t = Au , \tag{2.1'}$$

A skew selfadjoint. In this case we can construct invariant *quadratic* functionals; they are of the form

$$F(u) = (u, Bu) , \tag{2.2}$$

where B is a linear operator which *commutes* with A. The stationary points of (2.2) are the null vectors of B; so in this case Theorem 2.1 says that the null space of a linear operator B which commutes with A is an invariant subspace for solutions of (2.1)'. This is well known, of course.

Proof of Theorem 2.1: Frechet differentiability means that for every u and v the derivative

$$\frac{d}{d\varepsilon} F(u + \varepsilon v)$$

exists at $\varepsilon=0$ and is a linear functional of v. We assume that the linear space in which solutions lie is equipped with an *inner product* (,). Then linear functionals can be expressed as an inner product; in particular,

$$\frac{d}{d\varepsilon} F(u+\varepsilon v)\big|_{\varepsilon=0}=(G_F(u), v)\,. \tag{2.3}$$

$G_F(u)$ is called the *gradient* of F at u; G_F is a nonlinear operator defined on the space. The stationary points of F form the *null set* of G_F. We want to show that if $u(t)$ is a solution of (2.1) whose initial value belongs to the null set of G_F, then all values $u(t)$ belong to the null set. To see this we imbed the solution $u(t)$ in a differentiable one-parameter family of solutions $u(t, \varepsilon)$ which for $\varepsilon=0$ reduces to $u(t)$. For example, we may define $u(t, \varepsilon)$ as that solution of (2.1) whose value at $t=T$ is

$$u(T, \varepsilon)=u(T)+\varepsilon w\,, \tag{2.4}$$

where T is an arbitrary but fixed time and w an arbitrary but fixed vector. We assume that there is a dense set of vectors w for which the dependence on ε of u thus defined is differentiable.

Since F is conserved, $F(u(t, \varepsilon))$ is independent of t, for every fixed ε. Then the same is true of its ε-derivative at $\varepsilon=0$. Using the definition of the gradient given in (2.3), we can write this derivative as

$$(G_F(u(t)), v(t))\,, \tag{2.5}$$

wnere

$$v(t)=\frac{d}{d\varepsilon} u(t, \varepsilon)\big|_{\varepsilon=0}\,. \tag{2.6}$$

By assumption, $u(0)$ belongs to the null set of G_F; therefore, (2.5) is zero at $t=0$. Since (2.5) is independent of t, it is zero for all t. In particular, it is zero at $t=T$:

$$(G_F(u(T)), v(T))=0\,. \tag{2.7}$$

Comparing (2.6) and (2.4) we see that

$$v(T)=w\,,$$

so that (2.7) says that $G_F(u(T))$ is orthogonal to w. Since w is an arbitrary element, it follows that $G_F(u(T))=0$, and since T is an arbitrary time, it follows that $G_F(u(t))=0$ for all t. This completes the proof of Theorem 2.1.

We derive now a further result along these lines which will be useful in Section 5.

THEOREM 2.2. *Let F be a differentiable conserved functional for the evolution equation* (2.1). *Denote by $M(u)$ the derivative of the nonlinear operator $K(u)$ appearing on the right in* (2.1):

$$Mv = \frac{d}{d\varepsilon} K(u+\varepsilon v)|_{\varepsilon=0}\,; \tag{2.8}$$

M is a linear operator. Let $u=u(t)$ be a solution of (2.1); *then $G_F = G_F(u(t))$ satisfies*

$$\left(\frac{\partial}{\partial t} + M^*\right) G_F = 0\,. \tag{2.9}$$

where M^* is the adjoint of M with respect to the inner product used to define the gradient.

Proof: Let $u=u(t,\varepsilon)$ be the one-parameter family of solutions of (2.1) introduced earlier. Differentiating (2.1) with respect to ε at $\varepsilon=0$, we get, using v as defined by (2.6) and M defined by (2.8), the following linear equation for v:

$$v_t = Mv\,. \tag{2.10}$$

Since (2.5) is independent of t, its t derivative is zero:

$$\left(\frac{\partial}{\partial t} G_F, v\right) + \left(G_F, \frac{\partial}{\partial t} v\right) = 0\,.$$

Substituting Mv for v_t and applying the adjointness relation we get the relation

$$\left(\left(\frac{\partial}{\partial t} + M^*\right) G_F, v\right) = 0$$

for all t. Since at any particular value of t the values of v can be prescribed arbitrarily, it follows that equation (2.9) is satisfied as asserted in Theorem 2.2.

THEOREM 2.3. *Denote by $G_F'(u)$ the second derivation of F; $G_F'(u)$ is an operator defined by*

$$\frac{d}{d\eta} G_F(u+\eta w)|_{\eta=0} = G_F'(u)w\,. \tag{2.11}$$

Assertion: G'_F *is a symmetric operator with respect to the parentheses* $(\ ,\ )$ *used to define* G_F.

Proof: Substituting the definition (2.3) of G_F into (2.11) we get

$$\frac{\partial}{\partial \eta}\frac{\partial}{\partial \varepsilon} F(u+v+\eta v)\Big|_{\substack{\varepsilon=0\\ \eta=0}} = (G'_F(u)w, v) .$$

The symmetry of $G'_F(u)$ follows from the equality of mixed partial derivatives $\partial^2/\partial\eta\,\partial\varepsilon$ and $\partial^2/\partial\varepsilon\,\partial\eta$.

3. Conserved Functionals for the KdV Equation

The usefulness of the results contained in Section 2 hinges on the existence of plenty of conserved functionals whose gradient has a nontrivial but not too large null set. The KdV equation (1.1) is rich in such functionals; three of them are classical:

$$\begin{aligned} F_{-1}(u) &= \int u\,dx ,\\ F_0(u) &= \int \tfrac{1}{2}u^2\,dx , \qquad (3.1)\\ F_1(u) &= \int (-\tfrac{1}{3}u^3 + u_x^2)\,dx . \end{aligned}$$

Kruskal and Zabusky made the remarkable discovery that there are further invariant functionals of which the first is

$$F_2(u) = \int (\tfrac{1}{4}u^4 - 3uu_x^2 + \tfrac{9}{5}u_{xx}^2)\,dx . \qquad (3.1)_2$$

Eventually, Gardner, Kruskal and Miura showed in [18] that these four are merely the first of an infinite sequence of conserved functionals; the n-th functional has the form

$$F_n(u) = \int P_n\,dx , \qquad (3.1)_n$$

where P_n is a polynomial in u and its derivatives up to order n.

The following characterization of the polynomials P_n is given in [18]: *Each term in* P_n *has weight* $n+2$, *where the weight of a term is defined as the* *sum of the weights of its factors, and the weight of* $u^{(k)}=\partial^k u$ *is defined as* $1+\frac{1}{2}k$.

It follows in particular that P_n is at most quadratic in $u^{(n)}$:

$$P_n = a_n u^{(n)2} + b_n u^{(n)} + c_n, \tag{3.2}$$

where a_n is a constant, and b_n and c_n are polynomials in $u^{(j)}$, $j<n$, of weight $1+\frac{1}{2}n$ and $2+n$. It was shown in [18] that a_n is non-zero; we choose it to be positive.

It further follows that, for $n>1$, P_n is quadratic in $u^{(n)}$ and $u^{(n-1)}$:

$$P_n = au^{(n)^2} + bu^{(n)}u^{(n-1)} + cu^{(n-1)^2} + du^{(n)} + eu^{(n-1)} + f, \tag{3.2$'$}$$

a as before and $b, \cdots, f$ polynomials in $u^{(j)}$, $j<n-1$.

THEOREM 3.1. (a) *Let* u *be a smooth periodic function,* n *some non-negative integer; the quantities*

$$\max\{|u(x)|, |u^{(1)}(x)|, \cdots, |u^{(n-1)}(x)|\}, \qquad \int u^{(n)2}\, dx \tag{3.3}$$

can be bounded in terms of $F_0(u), F_1(u), \cdots, F_n(u)$.

(b) *The functional*

$$F_n(u) \geqq \tfrac{1}{2}a_n \int u^{(n)^2} - f_{n-1}, \tag{3.4}$$

where the quantity f_{n-1} *depends on* $F_0(u), F_1(u), \cdots, F_{n-1}(u)$.

Proof: The case $n=0$ is trivial; to prove the result for $n=1$ we introduce the abbreviations

$$M = \max|u(x)|, \qquad S = \int u_x^2. \tag{3.5}$$

By the mean value theorem, there is a point x_0 where

$$u^2(x_0) = \frac{1}{p}\int u^2\, dx = \frac{F_0(u)}{p},$$

p the period. For any x,

$$u(x) = u(x_0) + \int_{x_0}^{x} u_x\, dx;$$

using the Schwarz inequality, we get

$$u^2(x) \leqq 2u^2(x_0) + 2p\int u_x^2\,dx\,,$$

which implies that

$$M^2 \leqq \frac{2F_0}{p} + 2pS\,. \tag{3.6}$$

By definition $(2.8)_1$ of F_1,

$$S = \int u_x^2\,dx = F_1 + \int \tfrac{1}{3}u^3\,dx\,.$$

Using the definition $(2.8)_0$ of F_0 we get the inequality

$$S \leqq F_1 + \frac{MF_0}{3}\,. \tag{3.7}$$

Substituting this into (3.6) we have

$$M^2 \leqq \frac{2F_0}{p} + 2pF_1 + \frac{2pMF_0}{3}\,;$$

from this we easily estimate M in terms of F_0, F_1, and p. From (3.7) we get a similar estimate for S. This completes the proof of part (a) for $n=1$. For $n>1$ we proceed by induction. Using the quadratic character (3.2)$'$ of P_n and the positivity of the coefficient a we can easily estimate $\int u^{(n)^2}\,dx$ in terms of $F_n(u)$ and quantities already estimated by the induction assumption; $\max|u^{(n-1)}|$ can then be estimated in terms of $\int u^{(n)^2}\,dx$ and F_0. This completes the inductive proof of part (a).

To prove part (b) for $n>1$ we use formula (3.2)$'$ for P_n; since the coefficients $b,\cdots,f$ depend on $u^{(j)}$, $j<n-1$, it follows from part (a) that they, as well as $\int u^{(n-1)^2}\,dx$, are bounded.

Using the Schwarz inequality one can show that the sum of terms beyond the first is less than $\frac{1}{2}a\int u^{(n)^2}\,dx+\text{const}$; this implies inequality (3.4).

The case $n=1$ can be handled as in part (a).

We consider now the following extremum problem, originally posed by Kruskal and Zabusky:

Minimize $F_N(u)$, *given the values of* $F_{-1}(u), \cdots, F_{N-1}(u)$:

$$F_j(u) = A_j\,, \qquad j = -1, \cdots, N-1\,. \tag{3.8}$$

The constants A_j have to be so chosen that the constraints (3.8) can be satisfied by some admissible function, and furthermore so that A_j is *not* an extremal value of $F_j(u)$ when the other constraints are imposed. Analytically this means that for any function u satisfying the constraints (3.8), the gradients $G_j(u)$, $j = -1, \cdots, N-1$, are *linearly independent*. We call such constraints *admissible*.

THEOREM 3.2. (a) *Suppose the constraints* (3.8) *are admissible; then there is a periodic function* u *which minimizes* $F_N(u)$ *among all functions satisfying the constraints.*

(b) *Every minimizing function* u *satisfies an Euler equation of the form*

$$G(u) = G_N(u) + \sum_{-1}^{N-1} a_j G_j(u) = 0\,, \tag{3.9}$$

where the a_j *are some numbers, and* G_j *abbreviates* G_{F_j}, *the gradient of* F_j *with respect to the* L_2 *scalar product.*

Proof: It follows from inequality (3.4) that $F_N(u)$ is bounded from below for all u satisfying the constraints. Let u_n be a minimizing sequence of functions satisfying the constraints. It follows from part (a) of Theorem 3.1 that $u^{(j)}(x)$ are bounded for $j < N-1$ uniformly for all n. From inequality (3.4) we know that also the sequence $\int |u_n^{(N)}|^2\,dx$ is bounded; therefore so is $u^{(N-1)}(x)$.

We can select a subsequence for which $u_n^{(N)}$ converges in the weak topology in L_2 to some limit $u^{(N)}$; since the L_2 norm is lower semicontinuous in the weak topology,

$$\int u^{(N)2}\,dx \leqq \liminf_{n\to\infty} \int u_n^{(N)2}\,dx\,.$$

It follows from the Rellich compactness theorem that for a further subsequence, $u_n^{(j)}(x)$ converges uniformly to some limit $u^{(j)}$ for all $j < N$. Clearly, $u^{(j)}$ is the j-th derivative of $u^{(0)}$ in the sense of distributions; equally clearly, $u^{(0)} = u$ satisfies the constraints, and minimizes F_N. This completes the proof of part (a) of Theorem 3.2.

To prove part (b) we consider one-parameter families of deformations of u of the form

$$u(\varepsilon)=u+\varepsilon v+\sum_{-1}^{N-1} a_k v_k\,, \tag{3.10}$$

where v is an arbitrary function and the v_k are N fixed functions so chosen that the matrix

$$(G_j(u),\, v_k) \tag{3.11}$$

is nonsingular. It is possible to choose such v_k since for admissible constraints the functions $G_j(u)$, $j=-1,\cdots,N-1$, are linearly independent. The coefficients $a_k=a_k(\varepsilon)$ are so chosen that the constraints

$$F_j(u+\varepsilon v+\sum a_k v_k)=A_j\,, \tag{3.12}$$

$j=-1,\cdots,N-1$, are satisfied. According to the implicit function theorem, we can choose such a_k for every sufficiently small ε, so that $a_k(0)=0$ and a_k depends differentiably on ε. Differentiating (3.12) at $\varepsilon=0$, we get

$$(G_j(u),\, v+\sum \dot a_k(0)v_k)=0\,.$$

Since (3.11) is nonsingular, this determines $\dot a_k(0)$ uniquely. In particular, if v satisfies

$$(G_j(u),\, v)=0\,, \qquad j=-1,\cdots,N-1\,, \tag{3.13}$$

then

$$\dot a_k(0)=0 \tag{3.14}$$

for all k.

Since $u=u(0)$ minimizes $F_N(u)$, it follows that

$$\frac{d}{d\varepsilon} F_N(u(\varepsilon))\big|_{\varepsilon=0}=0\,;$$

this can be expressed as

$$(G_N(u),\, v+\sum \dot a_k(0)v_k)=0\,.$$

In particular, if v satisfies (3.13), all $\dot a_k(0)=0$, so that then

$$(G_N(u),\, v)=0\,. \tag{3.15}$$

This can happen only when $G_N(u)$ is a linear combination of $G_j(u)$, $j=-1,\cdots,N-1$, as asserted in (3.9), completing the proof of Theorem 3.2.

Equation (3.9) states that u is a stationary point of

$$F(u)=F_N(u)+\sum a_jF_j(u)\,. \tag{3.16}$$

F, being the linear combination of conserved functionals, is itself conserved; we conclude therefore from Theorem 2.1:

THEOREM 3.3. *Solutions of* (3.9) *form an invariant set for the* KdV *flow.*

According to $(3.1)_n$, F_n is the integral of a polynomial of u and its derivatives up to order n. It follows that G_n is a differential operator of order $2n$. Based on formulas (3.1) we can easily compute the gradients of the first four conserved functionals:

$$\begin{aligned} G_{-1}(u)&=1\,,\\ G_0(u)&=u\,,\\ G_1(u)&=-2u_{xx}-u^2\,,\\ G_2(u)&=\tfrac{18}{5}u_{xxxx}+6uu_{xx}+6u_x^2+u^3\,. \end{aligned} \tag{3.17}$$

For $N=0$, equation (3.9) is

$$u+a_{-1}=0\,, \tag*{$(3.18)_0$}$$

whose solutions are constant, a very uninteresting case. For $N=1$, we have

$$-2u_{xx}-u^2+a_0u+a_{-1}=0\,. \tag*{$(3.18)_1$}$$

Multiplying this by u_x we get, after integration,

$$u_x^2+\tfrac{1}{3}u^3-a_0\tfrac{1}{2}u^2-a_{-1}u=\text{const}\,. \tag{3.19}$$

Solutions of this equation are *elliptic functions.*

THEOREM 3.4. *If* $u(x)$ *is a solution of* $(3.18)_1$, *then*

$$u(x,t)=u(x-a_0t) \tag{3.20}$$

is a solution of the KdV *equation.*

Proof: Differentiating $(3.18)_1$ with respect to x we get, after dividing by -2,

$$u_{xxx}+uu_x-a_0u_x=0\,. \tag{3.21}$$

On the other hand, substituting (3.20) into the KdV equation (1.1), we have

$$u_t+uu_x+u_{xxx}=-a_0u_x+uu_x+u_{xxx}$$

which according to (3.21) is zero.

Solutions of the form (3.20) travel with constant speed a_0, without altering their shape. The existence of such solutions, and the differential equation (3.21) governing them, was discovered by Korteweg and deVries; they called these solutions *cnoidal waves.*

In an appendix, M. Hyman presents a numerical study of the case $N=2$. Solutions of equation (3.9) and the KdV flow on this manifold of solutions are described.

In Section 5 we present a theoretical study of solutions of (3.9) for general N; we shall show that the set of solutions constitutes an N-dimensional torus, and that the KdV flow on this torus is quasi-periodic.

4. Hamiltonian Mechanics

In this section we give a brief review of some topics in classical Hamiltonian mechanics which will be used in Section 5. The Hamiltonian form of the equations of motion is

$$\frac{d}{dt}q_n=\frac{\partial H}{\partial p_n}, \qquad \frac{d}{dt}p_n=-\frac{\partial H}{\partial q_n}, \qquad n=1,\cdots,N, \tag{4.1$_H$}$$

where H, the Hamiltonian, is some function of the $2N$ variables p_n, q_n, $n=1,\cdots,N$. The *Poisson bracket* of any pair of functions F and K of the variables p, q is defined as follows:

$$[F,K]=\sum\frac{\partial F}{\partial q_n}\frac{\partial K}{\partial p_n}-\frac{\partial F}{\partial p_n}\frac{\partial K}{\partial q_n}\,. \tag{4.2}$$

A straightforward calculation shows that

(a) the Poisson bracket is a bilinear, alternating function of F and K,

(b) the Jacobi identity

$$[[F,H],K]+[[H,K],F]+[[K,F],H]=0 \tag{4.3}$$

is satisfied.

When p and q change according to $(4.1)_H$, the rate of change of any function F of p and q equals the Poisson bracket of F and H:

$$
\begin{aligned}
\frac{dF}{dt} &= \sum \frac{\partial F}{\partial q}\frac{dq}{dt}+\frac{\partial F}{\partial p}\frac{dp}{dt} \\
&= \sum \frac{\partial F}{\partial q}\frac{\partial H}{\partial p}-\frac{\partial F}{\partial p}\frac{\partial H}{\partial q}=[F, H]
\end{aligned} \tag{4.4}
$$

It follows from this that F *is constant along the trajectories of the flow* $(4.1)_H$ *if and only if* $[F, H]=0$.

More generally, let v be a function of p, q and t. Relation (4.4) implies that

$$
\begin{aligned}
\frac{dv}{dt} &= \frac{\partial v}{\partial t}+\frac{\partial v}{\partial q}\frac{dq}{dt}+\frac{\partial v}{\partial p}\frac{dp}{dt} \\
&= \frac{\partial v}{\partial t}+[v, H].
\end{aligned}
$$

Thus we conclude that v *is constant along trajectories of the flow* $(4.1)_H$ *if and only if* v *satisfies*

$$\frac{\partial v}{\partial t}+[v, H]=0. \tag{4.5_H}$$

For fixed H, $[v, H]$ is a first-order partial differential operator acting on v, called the *Liouville operator* and denoted by L_H:

$$L_H v=[v, H]. \tag{4.6}$$

Equation (4.5) is called the *Liouville equation.* Its solution can be written symbolically as

$$v(t)=e^{-tL_H}v(0)\,; \tag{4.7_H}$$

e^{-tL_H} is the *solution operator* for the Liouville equation $(4.5)_H$. Let L_H and L_K be two Liouville operators. It is an immediate consequence of the Jacobi identity (4.3) that

$$L_K L_H - L_H L_K = L_{[H,K]}. \tag{4.8}$$

In particular, *if* $[H, K]=0$, *the operators* L_K *and* L_H *commute.* It follows then that also their exponentials, $\exp\{-tL_H\}$ and $\exp\{-tL_K\}$, the solution operators of $(4.5)_H$ and $(4.5)_K$, commute. Since solutions of these equations are characterized as functions of p, q, t which are constant along trajectories of $(4.1)_H$ and $(4.1)_K$, respectively, we have established

THEOREM 4.1. *If the Poisson bracket of K and H is zero, the Hamiltonian flows $(4.1)_H$ and $(4.1)_K$ commute.*

5. The KdV Equation as a Hamiltonian System

In [6], Clifford Gardner has given a Hamiltonian formulation of the KdV equation in the periodic case. In [3], Faddeev and Zacharov have given a Hamiltonian formulation of the KdV equation on the whole real axis; they have shown, using the inverse method of Gardner, Green, Kruskal and Miura, that this Hamiltonian system is completely integrable.

In this section we show how to use Gardner's results to study the manifold of periodic solutions of (5.1) and the KdV flow on this manifold.

Gardner defines a Poisson bracket for the class of C^∞ functionals F, H of smooth periodic functions u as follows:

$$[F,H]=(G_F,\partial G_H)\,, \tag{5.1}$$

where the parenthesis (,) is the L_2 scalar product, G_F and G_H the gradients of F and H with respect to (,), and $\partial = d/dx$.

THEOREM 5.1. (a) *The Poisson bracket defined by* (5.1) *is a bilinear and alternating function of F and H.*

(b) *The Jacobi identity* (4.3) *holds.*

Proof: Bilinearity is obvious, and the alternative character follows from the antisymmetry of the operator ∂.

To prove the Jacobi identity we first compute the gradient of $[F,H]$:

$$\begin{aligned} G_{[F,H]} &= \frac{d}{d\varepsilon}[F(u+\varepsilon v),H(u+\varepsilon v)]\big|_{\varepsilon=0} \\ &= \frac{d}{d\varepsilon}(G_F(u+\varepsilon v),\partial G_H(u+\varepsilon v)) \\ &= (G_F'(u)v,\partial G_H(u))+(G_F(u),\partial G_H'(u)v)\,. \end{aligned} \tag{5.2}$$

According to Theorem 2.3, G_F' and G_H' are symmetric; using this and the antisymmetry of ∂ we can rewrite (5.2) as

$$(v,G_F'\partial G_H)-(G_H'\partial G_F,v).$$

This shows that

$$G_{[F,H]} = G'_F \partial G_H - G'_H \partial G_F . \tag{5.3}$$

Using (5.3) we can write

$$\begin{aligned}[][[F,H],K]+[[H,K],F]+[[K,F],H] \\ &= (G'_F \partial G_H - G'_H \partial G_F, \partial G_K) + (G'_H \partial G_K - G'_K \partial G_H, \partial G_F) \\ &\quad + (G'_K \partial G_F - G'_F \partial G_K, \partial G_H) .\end{aligned}$$

Using once more the symmetry of the operators G'_F, G'_H, G'_K we conclude that the above sum is zero. This proves the Jacobi identity.

We turn now to the KdV equation (1.1). Using formula (3.17) for the gradient G_1 of the conserved functional F_1 we can write the KdV equation in the following form:

$$u_t = \tfrac{1}{2} \partial G_1 .$$

Let F be any functional; by definition of gradient,

$$\frac{dF}{dt} = (G_F, u_t) . \tag{5.4}$$

Using (5.4) and the definition (5.1) of the Poisson bracket we can write this as

$$\frac{dF}{dt} = (G_F, \tfrac{1}{2} \partial G_1) = [F, \tfrac{1}{2} F_1] . \tag{5.5}$$

Comparing this with (4.4) we recognize *the* KdV *equation as a Hamiltonian flow for the Poisson bracket* (5.1) *and with Hamiltonian* $H = \frac{1}{2} G_1$.

It follows from (5.5) that F is conserved for the KdV flow if and only if $[F, F_1] = 0$; since the functionals F_n defined by (2.8) are conserved, we conclude that

$$[F_n, F_1] = 0 , \qquad n = -1, 0, \cdots . \tag{5.6}$$

Gardner [6] has proved the following remarkable result:

THEOREM 5.2 (Gardner):

$$[F_n, F_m] = 0 \quad \textit{for} \quad \textit{all} \quad n, m . \tag{5.7}$$

Remark 1. The proof of this result first appeared in [6]; I learned this result from Gardner in the Spring of 1967, and remarked on its possible use at the end of [13].

Remark 2. The case $m=-1$ is trivial. For $m=0$, (5.7) expresses the invariance of the functionals F_n under translation with respect to x. For $m=1$, (5.7) expresses the fact that the F_n are conserved by the KdV flow.

Remark 3. In Section 6 we shall present the proof of a related result from which Theorem 5.2 follows.

We define the m-th generalized KdV *equation as*

$$u_t=\partial G_m(u)\,. \tag{5.8$_m$}$$

THEOREM 5.3. (a) *All functionals* F_n *are conserved for each equation* $(5.8)_m$.

(b) *The flows defined by equations* $(5.8)_m$ *commute with each other.*

Proof: Part (a) follows from Theorem 5.2, and part (b) from Theorems 4.1 and 5.2.

Remark. In view of Theorem 3.1 (a), u and its derivatives are bounded in terms of the functionals $F_n(u)$. Therefore we can derive, using part (a) of Theorem 5.3, *a priori* bounds for solutions of the generalized KdV equations $(5.8)_m$ in terms of their initial data. It can be shown that the initial value problem is properly posed for these equations.

We shall denote by $S_m(t)$ the *solution operator* of the m-th KdV equation; that is, $S_m(t)$ *relates the initial values of solutions of* $(5.8)_m$ *to their values at time* t. Part (b) of Theorem 5.3 states that *the operators* $S_m(t)$ *commute.*

THEOREM 5.4. *Let* u_0 *be a solution of equation* (3.9). *Then so is*

$$\prod_0^{N-1} S_m(t_m)u_0=u(t_0,\cdots,t_{N-1}) \tag{5.9}$$

for every value of the parameters $t_0,\cdots,t_{N-1}$.

Proof: Equation (3.9) is

$$G(u)=0\,,$$

where G is the gradient of F:

$$F = F_N + \sum_{-1}^{N-1} a_j F_j .$$

According to part (a) of Theorem 5.3, F is a conserved quantity for the m-th generalized KdV flow; therefore it follows from Theorem 2.1 that $S_m(t)$ maps a solution of (3.9) into a solution of (3.9). This proves Theorem 5.4.

DEFINITION. We denote the n-parameter family of functions defined by (5.9) by S.

Since (3.9) is an ordinary differential equation of order $2N$, its solutions are uniquely determined by the Cauchy data $\{u(x_0), u^{(1)}(x_0), \cdots, u^{(2N-1)}(x_0)\}$ at an arbitrary point x_0. It is convenient to regard the set S as imbedded via its Cauchy data in $\mathbb{R}^{2N}$.

THEOREM 5.5. *The set S defined by (5.9) is a bounded subset of $\mathbb{R}^{2N}$.*

Proof: Since by Theorem 5.3 each F_n is a conserved functional for each flow $(5.8)_m$, $F_n(u) = F_n(u_0)$ for all u, $n = 0, 1, \cdots, 2N$. It follows then from part (a) of Theorem 3.1 that the Cauchy data of u are bounded uniformly for all u in S; this proves Theorem 5.5.

THEOREM 5.6. *Suppose u_0 is a solution of the variational problem discussed in Theorem 3.2; then S is an N-dimensional manifold immersed in $\mathbb{R}^{2N}$.*

Proof: It was assumed in Theorem 3.2 that the constraints (3.8) are admissible. This implies, as we observed just before stating Theorem 3.2, that the gradients $G_j(u)$, $j = -1, \cdots, N-1$, form a linearly independent set of functions for every u satisfying the constraints. Since by (3.17), $G_{-1}(u) = 1$, the functions ∂G_j, $j = 0, \cdots, N-1$, are linearly independent. According to (5.8), $\partial G_j(u)$ is the direction in which the j-th generalized KdV flow starts.

Let $u = u(t_0, \cdots, t_{N-1})$ be any point of S. As we have shown above, $F_n(u) = F_n(u_0)$ for all n; consequently u satisfies the same constraints as u_0. It follows from the above that the directions $\partial G_j(u)$, $j = 0, \cdots, N-1$, are linearly independent. Since, by part (b) of Theorem 5.3, the flows S_j commute,

$$u(t_0 + \varepsilon_0, \cdots, t_{N-1} + \varepsilon_{N-1}) = S_0(\varepsilon_0), \cdots, S_{N-1}(\varepsilon_{N-1})u .$$

The linear independence of $\partial G_j(u)$, $j = 0, \cdots, N-1$, implies that for $\varepsilon_0, \cdots, \varepsilon_{N-1}$ small enough these points form an open subset of a smoothly imbedded N-dimensional manifold. This completes the proof of Theorem 5.6.

Remark 1. The tangent space of S at u is spanned by $\partial G_0(u), \cdots, \partial G_{N-1}(u)$. We observe that, by continuity, $F_n(u) = F_n(u_0)$ not only at all points u of S but also for all points u of $\bar{S}$. Therefore we deduce:

COROLLARY 5.7. *Let u be a point of $\bar{S}$, the closure of S; then*

$$S_0(\varepsilon_0) \cdots S_{N-1}(\varepsilon_{N-1})u\,, \qquad |\varepsilon_0| < \varepsilon\,,$$

forms an open subset of a smoothly imbedded N-dimensional manifold.

The next 8 lemmas lead up to showing that S is an algebraic variety. The first one is a calculus lemma:

LEMMA 5.8. *Suppose that Q is a polynomial in $u^{(0)}, u^{(1)}, \cdots, u^{(j)}$ such that, for every periodic function u of period p,*

$$\int_0^p Q(u)\, dx = 0\,. \tag{5.10}$$

Then there exists a polynomial J in $u^{(0)}, u^{(1)}, \cdots, u^{(j-1)}$ such that

$$Q = \partial J\,. \tag{5.11}$$

Proof: Let y be an arbitrary point, $0 < y < p$. Let u be a C^∞ function such that

$$u(0) = u^{(1)}(0) = \cdots = u^{(j-1)}(0) = 0\,. \tag{5.12}$$

Define the function v as follows:

$$v(x) = \begin{cases} u(x) & \text{for} \quad 0 \leqq x \leqq y\,, \\ q(x) & \text{for} \quad y \leqq x \leqq y+\varepsilon\,, \\ 0 & \text{for} \quad y+\varepsilon \leqq x \leqq p\,. \end{cases} \tag{5.13}$$

Here $q(x)$ is defined as that polynomial of degree $2j-1$ which satisfies the $2j$ interpolation conditions

$$q(y) = u(y), \quad q^{(1)}(y) = u^{(1)}(y), \quad \cdots, \quad q^{j-1}(y) = u^{(j-1)}(y)\,. \tag{5.14$_1$}$$

and

$$(5.14)_2 \qquad q(y+\varepsilon)=q^{(1)}(y+\varepsilon)=\cdots=q^{(j-1)}(y+\varepsilon)=0 .$$

It follows from $(5.14)_1$ that v and its derivatives of order less than j are continuous, and it follows from $(5.14)_2$ and (5.12) that they are periodic. The j-th derivative of v may have jump discontinuities at $x=y$ and $x=0$. Identity (5.10) is true for such a function:

$$\int_0^p Q(v)\,dx=0 .$$

Using the definition of v we get by (5.13)

$$(5.15) \qquad \int_0^y Q(u)\,dx=-\int_y^{y+\varepsilon} Q(q)\,dx .$$

The right side is a polynomial in the coefficients of q and ε; q in turn is completely determined by the data $(5.14)_1$. On the other hand, the left side is independent of ε. So we conclude from (5.15) that

$$\int_0^y Q(u)\,dx=J$$

for all u satisfying (5.12), where J is a polynomial in $u^{(0)}(y), u^{(1)}(y), \cdots, u^{(j-1)}(y)$. Differentiating this relation with respect to y we obtain (5.11); this completes the proof of Lemma 5.8.

Remark. J is uniquely determined if its constant term is taken to be zero.

According to Theorem 5.3,

$$(5.16) \qquad [F_m, F_n]=\int G_m\,\partial G_n\,dx=0 .$$

Both G_m and G_n are polynomials in u. Applying Lemma 5.8 to

$$H=G_m\,\partial G_n$$

we have

LEMMA 5.9.

$$(5.17) \qquad G_m\,\partial G_n=\partial J_{mn} ,$$

where J_{mn} *is a polynomial of* $u^{(0)}, u^{(1)}, \cdots, u^{(l)}$,

$$l=\begin{cases}2n & \text{if} \quad m\leqq n\,, \\ 2m-1 & \text{if} \quad m>n\,.\end{cases} \tag{5.17$'$}$$

Remark 1. The following relations are obviously true:

$$J_{mn}+J_{nm}=G_mG_n\,.$$

Remark 2. The leading term in J_{mn} is $u^{(2m)}u^{(2n)}$ for $m\leqq n$, and $u^{(2m-1)}u^{(2n+1)}$ for $m>n$.

The second derivative $G'_m(u)$ of F_m is a linear differential operator whose coefficients depend on u and its derivatives. According to Lemma 2.3, G'_m is symmetric; this means that, for any periodic u and any pair of periodic functions v and w,

$$\int_0^p (vG'_m(u)w-wG'_m(u)v)\,dx=0\,.$$

This implies that the integrand is a perfect derivative and proves

LEMMA 5.10.

$$vG_mw-wG'_wv=\partial C_m(v,w)\,, \tag{5.18}$$

where C_m *is a bilinear differential operator, whose coefficients depend on* u *and derivatives of* u.

LEMMA 5.11. *Suppose* u *satisfies the generalized* KdV *equation* $(5.8)_m$; *then* J_{jk} *satisfies*

$$\frac{\partial}{\partial t}J_{jk}+C_m[\partial G_j,\partial G_k]-G_jG'_m\partial G_k=0\,. \tag{5.19}$$

Proof: We show first that the expression on the left in (5.19) is independent of x. The x derivative of the first term, $\partial\frac{\partial}{\partial t}J_{jk}$, can be expressed with the aid of (5.17) as

$$\left(\frac{\partial}{\partial t}G_j\right)(\partial G_k)+G_j\,\partial\frac{\partial}{\partial t}G_j\,. \tag{5.20$_1$}$$

The x derivative of the second term can be expressed from (5.18) with $v=\partial G_j$ and $w=\partial G_k$ as

$$(\partial G_j)G_m\,\partial G_k-(\partial G_k)G_m\,\partial G_j\,. \tag{5.20$_2$}$$

The x derivative of the third term is

$$-(\partial G_j)G'_m\,\partial G_k-G_j\,\partial G'_m\,\partial G_k\,. \tag{5.20$_3$}$$

Adding these three expressions we get

$$(G_j\,\partial)\left[\frac{\partial}{\partial t}-G'_m\,\partial\right]G_k+(\partial G_k)\left[\frac{\partial}{\partial t}-G'_m\,\partial\right]G_j\,. \tag{5.21}$$

Next we turn to Theorem 2.2 and take for K the generalized KdV operator

$$K=\partial G_m\,.$$

The derivative M of K is then

$$M=\partial G'_m\,.$$

Using the antisymmetry of ∂ and the symmetry of G'_m we have

$$M^*=-G'_m\,\partial\,. \tag{5.22}$$

For the conserved functional F occurring in Theorem 2.2 we take F_k and F_j; setting G_k and G_j for G_F in (2.9), and using the value (5.22) of M^*, we obtain

$$\left(\frac{\partial}{\partial t}-G'_m\,\partial\right)G_k=0$$

and

$$\left(\frac{\partial}{\partial t}-G'_m\,\partial\right)G_j=0\,.$$

Using these relations in (5.21) we conclude that (5.21) is zero; since (5.2) is the x-derivative of (5.19), we see that (5.19) is independent of x.

Relation (5.19) is a polynomial in u and its derivatives with respect to x and t. If u satisfies the generalized KdV equation $(5.8)_m$, the t derivatives can be expressed in terms of x derivatives, so that (5.19) becomes a polynomial in the x derivatives of u alone. The initial values of u are arbitrary, so this polynomial is independent of x for arbitrary u. But this can be true only if all non-constant terms in the polynomial are zero. On the other hand, the constant term in this polynomial is zero since the ingredients of (5.19), J_{kj}, G_k, and G_j, all have zero constant terms. This proves that (5.19) is zero, and completes the demonstration of Lemma 5.11.

Suppose u is a solution of (3.9):

$$G(u)=\sum_{-1}^{N} a_j G_j = 0\,, \qquad a_N = 1\,. \tag{5.23}$$

We multiply this by ∂G_k; using the definition of J_{kj} and relation (5.17), we can write

$$\partial \sum a_j J_{jk} = 0\,.$$

Introducing the abbreviation

$$J_k = \sum a_j J_{jk}\,, \tag{5.24}$$

we deduce from the result above the following:

LEMMA 5.12. *If u satisfies $G(u)=0$, $J_k(u)$ is independent of x.*

Suppose $u=u(x,t)$ satisfies the generalized KdV equation $(5.8)_m$; then according to Lemma 5.11 the relations (5.19) hold. Multiplying (5.19) by a_j and summing, we get a relation which, using the definition J_k in (5.24) and (5.23) of G, can be written as follows:

$$\frac{\partial}{\partial t} J_k + C_m[\partial G, \partial G_k] - G G_m' \partial G_k = 0\,. \tag{5.25}$$

LEMMA 5.13: *$J_k(u)$ has the same value for all u in S.*

Proof: According to Theorem 5.4, $G(u)=0$ for every u in S. By definition of S, every point of S can be reached from u_0 by generalized KdV flows. Along such a flow, (5.25) holds. Since $G(u)=0$ on S, (5.25) becomes $\partial J_k/\partial t = 0$; since according to Lemma 5.12, $J_k(u)$ is independent of x for every u on S, the conclusion of Lemma 5.13 follows.

We have noted in Remark 2 following Lemma 5.9 that for $m>n$ the leading term of J_{mn} is $u^{(2m-1)}u^{(2n+1)}$; hence the leading term of J_k is $u^{(2N-1)}u^{(2k+1)}$. From this one easily deduces

LEMMA 5.14. *The polynomials* J_k, $k=0, 1, \cdots, N-1$, *in the* $2N$ *variables* $u^{(0)}, \cdots, u^{(2N-1)}$ *are algebraically independent.*

Denote by c_k the constants $J_k(u_0)$, and denote by V the set in $\mathbb{R}^{2N}$ satisfying

$$J_k = c_k\,, \qquad k=0, \cdots, N-1\,. \tag{5.26}$$

According to Lemma 5.14, V is an N-dimensional algebraic variety; according to Lemma 5.13, V contains S. We shall show next that S is a component of V; to see this we look at $\bar{S}$, the closure of S.

LEMMA 5.15. (a) $\bar{S}$ *is invariant under the generalized* KdV *flows.*
(b) $\bar{S} \subset V$.
(c) $\bar{S}$ *is compact.*

Denote as before the solution operator of the m-th KdV flow by S_m. Let u be some point of $\bar{S}$; by definition,

$$u = \lim u_n\,, \qquad u_n \in S\,.$$

Since $S_m(t)$ is a continuous operator, $S_m(t)u_n$ converges to $S_m(t)u$; this proves part (a) of the lemma.

Part (b) follows from the continuity of the functions J_k. Finally, according to Theorem 5.5, S is bounded, proving part (c).

According to Corollary 5.7, for any point u of $\bar{S}$, the set $S(u, \varepsilon)$ defined by

$$S(u, \varepsilon) = \prod S_j(t_j)u\,, \qquad |t_j| < \varepsilon\,, \tag{5.27}$$

is an open subset of a smoothly imbedded N-dimensional manifold. We claim that $S(u, \varepsilon)$ contains a neighborhood of u in $\bar{S}$. For suppose this is not the case; then there would be a sequence of elements u_n of $\bar{S}$ such that

$$\begin{aligned} &\text{(i)} \quad u_n \to u\,, \\ &\text{(ii)} \quad u_n \notin S(u, \varepsilon)\,. \end{aligned} \tag{5.28}$$

According to part (a) of Lemma 5.15 the set $S(u_n, \varepsilon)$ belongs to $\bar{S}$. We claim that $S(u_n, \frac{1}{2}\varepsilon)$ and $S(u, \frac{1}{2}\varepsilon)$ do not intersect. Suppose they did:

$$\prod S_j(t_j)u = \prod S_j(t_j')u_n\,, \qquad |t_j| < \tfrac{1}{2}\varepsilon,\ |t_j'| < \tfrac{1}{2}\varepsilon\,.$$

Since the S_j commute, this implies

$$\prod S_j(t_j - t_j')u = u_n$$

which would make u_n belong to $S(u, \varepsilon)$, contrary to $(5.28)_{ii}$.

We construct now a subsequence of u_n, again denoted by u_n, so that u_{n+1} does not belong to any $S(u_j, \varepsilon)$, $j \leqq n$; since each $S(u_j, \varepsilon)$ has a positive distance from u, this will be the case if u_{n+1} is close enough to u. Arguing as above we see that the sets $S(u_n, \frac{1}{2}\varepsilon)$ are pairwise disjoint. By part (a) of Lemma 5.15 these sets belong to $\bar{S}$, and so by part (b) of Lemma 5.15 they belong to V. But an algebraic variety V cannot have such a strudel-like structure with infinitely many leaves; so we conclude that $S(u, \varepsilon)$ contains a neighborhood of u in $\bar{S}$.

By definition of $\bar{S}$, every neighborhood of u in $\bar{S}$ contains points of S; thus some points of $S(u, \varepsilon)$ belong to S, which means that they are of the form $\prod S_j(r_j)u_0$. But then, again using the commutativity of the S_j, u itself is of this form. This shows that $\bar{S} = S$. Combining Lemma 5.15, Theorem 5.6 and the definition of S we can state

THEOREM 5.16. *S is a compact, connected, open and closed subset of the algebraic variety V; every point of S is a regular point.*

Next we turn to showing that S is an N-dimensional torus; we need

LEMMA 5.17. *There is a number T such that*

$$S \subset S(u_0, T). \tag{5.29}$$

Proof: By part (c) of Lemma 5.15, $\bar{S}$ is compact. Therfore, there are a finite number of points, $u_1, \cdots, u_m$ such that $\{S(u_j, \varepsilon)\}$ cover S. Each pair of points in one neighborhood can be connected to the other by applying $\prod S_k(t_k)$, $|t_j| < 2\varepsilon$; since each point can be linked to the base point u_0 by a chain of at most m links, (5.29) follows with $T = 2m\varepsilon$.

Lemma 5.17 implies that there must exist among the operators S_m relations of the form

$$\prod_0^{N-1} S_m(t_m) = I. \tag{5.30}$$

Now consider the set of all N-tuples $(t_0, t_1, \cdots, t_{N-1})$ for which (5.30) holds. These form a *module* $\mathcal{M}$ in $\mathbb{R}^N$ over the integers. It follows from Theorem 5.6 that (5.30) does not hold for $|t_m| < \varepsilon$ if ε is small enough except for $t_m \equiv 0$; this shows that the module $\mathcal{M}$ is *discrete*. According to linear algebra, a discrete module is a *lattice*, i.e., can be represented as

$$\sum_{j=1}^{K} n_j \omega_j \; ; \tag{5.31}$$

as the n_j range over all integers, each point of $\mathcal{M}$ has exactly one representation of form (5.31). K is the *dimension* of $\mathcal{M}$, where $K \leqq N$.

By Lemma 5.15, $S = \mathbb{R}^N \pmod{\mathcal{M}}$ is compact; according to linear algebra this implies that the lattice $\mathcal{M}$ is N-dimensional. In this case the quotient $\mathbb{R}^N/\mathcal{M}$ is an N-dimensional torus. This proves

THEOREM 5.18.

$$\prod S_m(t_m) : \mathbb{R}^N/\mathcal{M} \to \mathbb{R}^{2N} \tag{5.32}$$

is a one-to-one mapping of the torus $\mathbb{R}^N/\mathcal{M}$ onto S. In particular, S is a torus; the KdV *flow, given by $S_1(t_1)$ on $\mathbb{R}^N/\mathcal{M}$, is quasi-periodic.*

The case $N=1$ is trivial from the point of view of Theorem 5.18; since, according to (3.17), $\partial G_0 = \partial$, $S_0(t)$ is translation across the x-axis. The module $\mathcal{M}$ is generated by a single relation which says that functions in S are periodic; the KdV flow also is translation along the x-axis.

The case $N=2$ is more interesting; here $\mathcal{M}$ has two generators, at least one of which, call it $\omega = (t_0, t_1)$, has second component $t_1 \neq 0$. From (5.30) we see that on S

$$S_1(t_1) = S_0(-t_0) \,. \tag{5.33}$$

Since S_0 is translation, (5.33) says that if an initial function belonging to S is subjected to the KdV flow, it resumes its initial shape after time t_1, translated by the amount t_0. The calculations of M. Hyman, presented in the appendix bear this out strikingly.

Actually, the calculations were carried out first and were helpful in pointing the theory in the right direction. In particular, Lemma 5.13 was suggested by numerical evidence.

6. The Spectrum of the Schrödinger Operator

Let L denote the Schrödinger operator

$$L = \partial^2 + v \,, \tag{6.1}$$

$v(x)$ being some periodic potential, acting on periodic functions. L has as discrete spectrum $\{\lambda_j\}$, $j = 1, 2, \cdots$. Whereas the eigenvalues are completely determined by the potential v, the converse is not true: operators L with different potentials v may very well have the same spectrum. In fact, Gardner, Kruskal and Miura have shown in [7] that if $u(x, t)$ is a solution of the KdV equation, then, for

$$v(x, t) = \tfrac{1}{6}u(x, t)\,, \tag{6.2}$$

the spectrum of the operator $L = L(t)$ defined by (6.1) is independent of t. In [13], I gave a new derivation of this fact and indicated how to construct an infinite sequence of differential equations whose solutions have the same property; Gardner has identified these equations as the generalized KdV equations discussed in Section 5. We start by reproducing very briefly the derivation of these results.

Let $L(t)$ be a one-parameter family of selfadjoint operators in Hilbert space. Clearly the spectrum of $L(t)$ is independent of t if the operators $L(t)$ are unitarily equivalent to each other, i.e., if there is a one-parameter family of unitary operators $U(t)$ such that, for all t,

$$U^*(t)L(t)U(t) = L(0)\,. \tag{6.3}$$

Assume that both $L(t)$ and $U(t)$ depend differentiably on t. Differentiating with respect to t the relation

$$U(t)U^*(t) = I\,,$$

we see that

$$U_t U^* = B \tag{6.4}$$

is *antisymmetric:*

$$B^* = -B\,. \tag{6.5}$$

From (6.4) and the relation $U^*U = I$ we deduce that

$$U_t = BU\,. \tag{6.6}$$

Conversely, given $B(t)$ as an antisymmetric operator-valued function of t, we can construct an operator $U(t)$ satisfying (6.6) by solving the initial value problem

$$h_t(t) = B(t)h(t)\,, \qquad h(t_0) = h_0\,, \tag{6.6$'$}$$

and setting $U(t)h(0)=h(t)$. Clearly the operator $U(t)$ defined in this way is isometric; if the initial value problem (6.6)′ can be solved for a dense set of h_0 for every t_0, then both $U(t)$ and its inverse are densely defined, so that the closure of $U(t)$ is unitary. Differentiate (6.3) with respect to t; using (6.6) and (6.5) we get

$$L_t = BL - LB. \tag{6.7}$$

This condition in turn implies the unitary equivalence (6.3); so we have proved

THEOREM 6.1. *Let $L(t)$ be a one-parameter family of selfadjoint operators, $B(t)$ a one-parameter family of anti-selfadjoint operators. Suppose that (6.7) holds, and that (6.6)′ can be solved for a dense set of h_0; then the operators $L(t)$ are unitarily equivalent.*

We apply this theorem to L given by (6.1) where v depends on t. In this case, $L_t = v_t$, and (6.7) becomes

$$v_t = BL - LB. \tag{6.7)$'$}$$

In [13] it is described how to satisfy this equation by differential operators B_m, with real coefficients and leading coefficient 1:

$$B_m = \partial^{2m+1} + \sum_0^{m-1} b_j^{(m)} \partial^{2j+1} + \partial^{2j+1} b_j^{(m)}. \tag{6.8$_m$}$$

Clearly, B_m is antisymmetric. Relation (6.7)′ requires the commutator of B_m and L to be multiplicative. Now the commutator of B_m and L is a differential operator of order $2m$; furthermore, this operator is symmetric. Therefore the requirement that this operator be actually of order zero imposes m conditions, which can be satisfied by appropriate choice for the m coefficients $b_j^{(m)}$ in (6.8)$_m$.

Explicit calculation for the cases $m=0$ gives

$$B_0 L - L B_0 = u_x;$$

so (6.7)$'_0$ is

$$v_t = v_x. \tag{6.9$_0$}$$

For $m=1$ we obtain, with $b = b'_0$,

$$B_1 L - L B_1 = \partial(3v_x - 4b_x)\partial + v_{xxx} - b_{xxx} + 2bv_x.$$

Setting $b=\frac{1}{3}v$, and $v=\frac{1}{6}u$ we get in $(6.7)_1'$:

$$(6.9)_1 \qquad 4u_t = uu_x + u_{xxx}\,.$$

This is, except for an inessential change of sign and scale in t, the KdV equation, and of course $(6.9)_0$ is the zero-th KdV equation. We introduce the notation

$$(6.10)_m \qquad B_m L - L B_m = K_m\,.$$

As already reported in [13], Gardner has observed that K_m is related to the m-th conserved functional of KdV:

THEOREM 6.2 (Gardner). *Let L be the operator*

$$(6.11) \qquad L=\partial^2+\tfrac{1}{6}u$$

and B_m the operator of form $(6.8)_m$ so chosen that $(6.10)_m$ is satisfied. Then

$$(6.12)_m \qquad K_m = c_m\,\partial G_m\,,$$

G_m being the gradient of the m-th conserved functional F_m, c_m some constant.

Thus equation

$$(6.9)_m \qquad u_t = K_m$$

is, except for a change of scale in time, the m-th generalized KdV equation $(5.8)_m$.

Using Theorem 6.1 and relation $(6.10)_m$ we see that if u satisfies $(6.9)_m$, the operators $L(t)$ are unitarily equivalent. The eigenvalues λ_j of L are functionals of the potential u appearing in L. We recognize this dependence of λ_j on u by writing

$$(6.13) \qquad \lambda_j(u)\,.$$

The invariance of the spectrum of $L(t)$ when u varies subject to $(6.9)_m$ can be expressed as follows:

THEOREM 6.3. *The eigenvalues (6.13) are conserved functionals under all generalized KdV flows.*

For $m=0$, the KdV flow is just translation along the x-axis; for $m=1$, this result was discovered by Gardner, Green, Kruskal and Miura (see [7]).

We have now two different kinds of conserved functionals for the generalized KdV equations: the functionals $(3.1)_n$ described in Section 3 and the eigenvalues (6.13) of the operator (6.11). Already Kruskal and Zabusky have observed that these functionals are not independent: the functionals F_m defined in $(3.1)_m$ can be related to the asymptotic behavior of the eigenvalues for large n. An elegant derivation of these relations is given by McKean and van Moerbeke in [17].

We turn now to the functional λ_j:

THEOREM 6.4. *The functionals $\lambda_j(w)$ are in involution; i.e.,*

$$[\lambda_j, \lambda_k]=0 \tag{6.14}$$

for $j \neq k$, where $[\ ,\ ]$ is the Poisson bracket (5.1).

Since the functionals F_n are functions of the λ_j, Theorem 6.4 gives another proof of Gardner's result contained in Theorem 5.2.

For the proof we need the following lemmas:

LEMMA 6.5. *Let λ be a simple eigenvalue of L, and denote by w the eigenfunction*

$$Lw=\lambda w \tag{6.15}$$

normalized by $(w, w)=1$. The gradient of λ is

$$G_\lambda=\tfrac{1}{6}w^2 . \tag{6.16}$$

Proof: We recall that the gradient of λ is defined by

$$\frac{d}{d\varepsilon}\lambda(u+\varepsilon v)|_{\varepsilon=0}=(G_\lambda(u), v) , \tag{6.17}$$

where $(\ ,\ )$ denotes the L_2 scalar product. Replace u in (6.15) by $u+\varepsilon v$, differentiate with respect to ε and set $\varepsilon=0$. Recalling the definition (6.11) of L we get

$$L\frac{dw}{d\varepsilon}+\tfrac{1}{6}vw=\lambda\frac{dw}{d\varepsilon}+\left(\frac{d\lambda}{d\varepsilon}\right)w .$$

Taking the scalar product with w, using the symmetry of L, the eigenvalue equation (6.15) and $(w, w)=1$, we obtain

$$\frac{1}{6}\int vw^2=\frac{d\lambda}{d\varepsilon} .$$

Comparing this with (6.17) we deduce (6.16).

Remark. If λ is a double eigenvalue at u, it cannot be defined as a single-valued functional for nearby u. However, $\lambda_1+\lambda_2$ is an unequivocally defined functional. It is not hard to show that

$$G_{\lambda_1+\lambda_2}=w_1^2+w_2^2\,, \tag{6.16$'$}$$

where w_1, w_2 is any orthonormal pair of eigenfunctions.

LEMMA 6.6. *Suppose* w *satisfies the differential equation* (6.15); *then* w^2 *satisfies the following third order differential equation:*

$$Hw^2=4\lambda\,\partial w^2\,, \tag{6.18}$$

where H *is the linear differential operator*

$$H=\partial^3+\tfrac{2}{3}u\,\partial+\tfrac{1}{3}u_x\,. \tag{6.19}$$

Remark. This lemma is well known.

Proof: We have

$$\partial^1 w^2=2ww_x\,,$$

$$\partial^2 w^2=2ww_{xx}+2w_x^2\,,$$

$$\partial^3 w^2=2ww_{xxx}+6w_xw_{xx}\,.$$

Substituting this into (6.19), we get

$$Hw^2=2ww_{xxx}+6w_xw_{xx}+\tfrac{4}{3}uww_x+\tfrac{1}{3}u_xw^2\,. \tag{6.20}$$

According to (6.15) and the definition of L in (6.11),

$$w_{xx}+\tfrac{1}{6}uw=\lambda w\,. \tag{6.21}$$

Differentiating with respect to x we have

$$w_{xxx}+\tfrac{1}{6}uw_x+\tfrac{1}{6}u_xw=\lambda w_x\,. \tag{6.21$'$}$$

Multiply (6.21) by $6w_x$ and (6.21)$'$ by $2w$ and add; using (6.20) we obtain (6.18).

Observe that *the operator* H *defined by* (6.19) *is antisymmetric:*

$$H^* = -H.$$

We are now ready to evaluate the Poisson bracket (6.14). By definition (5.1) of the bracket,

$$[\lambda_j, \lambda_k] = (G_{\lambda_j}, \partial G_{\lambda_k})$$

which by (6.16) equals

$$\tfrac{1}{36}(w_j^2, \partial w_k^2). \tag{6.22}$$

Using $(6.18)_k$, we can rewrite this as

$$\frac{1}{36 \times 4\lambda_k}(w_j^2, Hw_k^2).$$

Since H is antisymmetric, this equals

$$\frac{-1}{36 \times 4\lambda_k}(Hw_j^2, w_k^2).$$

Using $(6.18)_j$ and then the antisymmetry of ∂ we get

$$\frac{-\lambda_j}{36 \times \lambda_k}(\partial w_j^2, w_k^2) = \frac{\lambda_j}{36\lambda_k}(w_j^2, \partial w_k^2).$$

Comparing this with (6.22) we see that, for $\lambda_j \neq \lambda_k$,

$$(w_j^2, \partial w_k^2) = 0.$$

This completes the proof of Theorem 6.4.

The generalized KdV flows are not the only ones which leave the spectrum of L invariant and which commute with each other. In Theorem 6.10 we present a whole one-parameter family of such flows. First we need some auxiliary results.

Let α be any real number, and consider the equation

$$(L - \alpha)w = 0, \tag{6.23}$$

where L is given by (6.11). The Wronskian of any two solutions w_1 and w_2 of (6.23),

$$w_1\,\partial w_2 - w_2\,\partial w_1\,, \tag{6.24}$$

is independent of x. Hence, according to Floquet's classical theory, (6.23) has two particular solutions w_1 and w_2 which satisfy

$$w_1(x+p)=\kappa w_1(x)\,, \tag{6.25_1}$$

$$w_2(x+p)=\kappa^{-1} w_2(x)\,. \tag{6.25_2}$$

The parameter κ, called the *Floquet exponent*, is uniquely determined by the potential u and the parameter α which occurs in equation (6.23). If α lies in a stability interval, κ is complex of modulus 1. If α is an eigenvalue of L operating on functions with period p, $\kappa=1$; if α is an eigenvalue of L operating on functions of period $2p$, $\kappa=-1$. In these cases, and these cases only, w_1 and w_2 can coincide.

For fixed α, κ is a functional of u; we shall compute now the gradient of $\kappa(u)$. We normalize w_1, say, by $(w_1, w_1)=1$. Set $u+\varepsilon v$ in place of u in (6.23):

$$\partial^2 w + (\tfrac{1}{6}(u+\varepsilon v)-\alpha)w = 0\,,$$

and differentiate with respect to ε; we get at $\varepsilon=0$

$$(L-\alpha)\frac{dw_1}{d\varepsilon}+\tfrac{1}{6}vw_1=0\,. \tag{6.26}$$

Differentiating $(6.25)_1$ with respect to ε gives

$$\frac{dw_1}{d\varepsilon}(x+p)=\kappa\frac{dw_1}{d\varepsilon}(x)+\frac{d\kappa}{d\varepsilon}w_1(x)\,. \tag{6.27}$$

We multiply (6.26) by w_2 and integrate by parts:

$$\begin{aligned} 0 &= \int_0^p w_2(L-\alpha)\frac{dw_1}{d\varepsilon}+\tfrac{1}{6}vw_1w_2\,dx \\ &= \int (L-\alpha)w_2\frac{dw_1}{d\varepsilon}+\tfrac{1}{6}vw_1w_2\,dx + w_2\,\partial\frac{dw_1}{d\varepsilon}-\partial w_2\frac{dw_1}{d\varepsilon}\Big|_0^p\,. \end{aligned} \tag{6.28}$$

Since w_2 satisfies (6.23), $(L-\alpha)w_2=0$; the boundary terms in (6.28) can be evaluated with the aid of the relations (6.27) and $(6.25)_2$ and we get

$$W\frac{1}{\kappa}\frac{d\kappa}{d\varepsilon}=\frac{1}{6}\int vw_1w_2\,dx\,, \tag{6.29}$$

where W is the Wronskian (6.24).

We recall the definition of the gradient:

$$\left.\frac{d\kappa(u+\varepsilon v)}{d\varepsilon}\right|_{\varepsilon=0}=(G_\kappa(u),v)\,,$$

and conclude from (6.29) that the following lemma holds.

LEMMA 6.7. *Let κ be the Floquet exponent of equation* (6.23) *as defined by relation* (6.25). *The gradient of κ with respect to u is*

$$G_\kappa=\tfrac{1}{6}\kappa w_1w_2=\tfrac{1}{6}\kappa n\,, \tag{6.30}$$

where w_1, w_2 are solutions of (6.23) *satisfying* (6.25) *and so normalized that their Wronskian equals* 1. *The symbol n is an abbreviation:*

$$n=w_1w_2\,. \tag{6.31}$$

Remark. Although w_1 and w_2 separately are not periodic, it follows easily from (6.25) that their product n is periodic.

LEMMA 6.8. *The product n of any two solutions of* (6.23) *satisfies the equation*

$$Hn=4\alpha\,\partial n\,, \tag{6.32}$$

where H is the linear operator (6.19).

Proof: This lemma is a corollary of Lemma 6.6, according to which squares of solutions of (6.23) satisfy (6.32). For, the product $n=w_1w_2$ can be written as a linear combination of $(w_1+w_2)^2$, w_1^2 and w_2^2.

THEOREM 6.9. (a) *The functionals $\kappa(u,\alpha)$ are in involution, i.e.,*

$$[\kappa(\alpha),\kappa(\beta)]=0 \tag{6.33}$$

for $\alpha\neq\beta$, neither α nor β being equal to an eigenvalue of L.

(b) *The functionals* $\kappa(u, \alpha)$ *and* $\lambda_j(u)$ *are in involution, i.e.,*

$$[\kappa(\alpha), \lambda_j] = 0 . \tag{6.33$'$}$$

Here $[\ ,\ ]$ *is the Poisson bracket defined by* (5.1).

Proof: By definition of the Poisson bracket in (5.1),

$$[\kappa(\alpha), \kappa(\beta)] = (G_{\kappa(\alpha)}, \partial G_{\kappa(\beta)})$$

and

$$[\kappa, \lambda] = (G_\kappa, \partial G_\lambda^*) .$$

Using (6.30) and (6.16) we can write the above expressions as

$$\tfrac{1}{36}\kappa(\alpha)\kappa(\beta)(n(\alpha), \partial n(\beta))$$

and

$$\tfrac{1}{36}\kappa(n, \partial w^2) .$$

According to Lemma 6.6,

$$Hw^2 = 4\lambda\, \partial w^2 ,$$

and according to Lemma 6.8,

$$Hn = 4\alpha\, \partial n .$$

Using these relations, and the antisymmetry of ∂ and H, we can complete the proof of Theorem 6.9 along the lines of the proof of Theorem 6.4.

For the next step we need the following observation, easily deduced using the standard tools of the theory of differential equations:

w_1 *and* w_2, *measured in the* C^1 *norm, depend Lipschitz continuously on u, measured in the* C^0 *norm.* From this it follows by the Banach space version of the Picard existence theorem that the differential equation

$$u_t = \partial n(\alpha, u) \tag{6.34}$$

has a unique solution, local in time, for arbitrary continuous initial function u_0.

THEOREM 6.10. (a) *The flows* (6.34) *corresponding to different values of* α *commute.*

(b) *The spectrum of* L *does not change under the flows* (6.34).

(c) *For* C^∞ *initial data, the functionals* $F_m(u)$ *introduced in Section* 3 *are conserved under the flows* (6.34).

(d) *For* C^1 *initial data, the flows* (6.34) *exist for all time.*

Proof: Parts (a) and (b) follow from parts (a) and (b) of Theorem 6.9, using Theorem 4.1 of Hamiltonian theory. Part (c) follows from the fact mentioned earlier that for C^∞ potentials the functionals $F_m(u)$ are completely determined by the spectrum of L.

To prove (d) we use Theorem 3.1, according to which the maximum of $u(x)$ can be estimated in terms of $F_0(u)$ and $F_1(u)$. According to part (c) above, $F_0(u(t))$ and $F_1(u(t))$ are independent of t; therefore, we have an estimate for $|u(x,t)|$ which is uniform for all t. This implies that the Lipschitz constant occurring in the dependence of w_1 and w_2 on u is uniform; under these conditions, equation (6.34) has a solution for all t.

It is easy to show, on the basis of perturbation theory, that n defined in Lemma 6.8 depends analytically on α:

$$n = \sum_0^\infty n_j \alpha^j . \tag{6.35}$$

Substituting this into (6.32) gives the following relation among the n_j:

$$Hn_j = 4\partial n_{j-1} ; \tag{6.36$_j$}$$

here we make the convention that $n_{-1} = 1$.

These equations can be easily solved recursively; since the operator H has a one-dimensional null space spanned by n_0, its range is orthogonal to n_0. It is easy to verify that the right sides of $(6.36)_j$ are orthogonal to n_0. Making use of the antisymmetry of ∂, then of equation $(6.36)_1$, then of the antisymmetry of H, and finally of equation $(6.36)_{j-1}$, we can write

$$\begin{aligned}(4\partial n_{j-1}, n_0) &= -(n_{j-1}, 4\partial n_0)\\ &= -(n_{j-1}, Hn_1) = (Hn_{j-1}, n_1) = (4\partial n_{j-2}, n_1) .\end{aligned}$$

Repeating this a number of times we obtain an expression of one of two forms:

$$(\partial n_k, n_k) \qquad \text{or} \qquad (Hn_k, n_k) ;$$

since both ∂ and H are antisymmetric and since H and n_k are real, the above scalar products are zero.

Since by Lemma 6.7, $n(\alpha)$ is the gradient of $6 \log \kappa(\alpha)$, it follows that n_j is the gradient of $(6/j!)(d/d\alpha)^j \log \kappa$. Therefore the flow

$$(6.37)_j \qquad u_t = \partial n_j(u)\,, \qquad j = 0, 1, \cdots,$$

is a Hamiltonian flow.

THEOREM 6.11. (a) *The flows* $(6.37)_j$ *commute with each other.*
(b) *The flows* $(6.37)_j$ *commute with the generalized* KdV *flows.*
(c) *The spectrum of* L *is invariant under the flows* $(6.37)_j$.
(d) *For* C^1 *initial data, the flows* $(6.37)_j$ *exist for all* t.

Proof: Denote the flow governed by the differential equation (6.34) by $S(\alpha, t)$. According to part (a) of Theorem 6.10,

$$S(\alpha, t)S(\beta, r) = S(\beta, r)S(\alpha, t)\,;$$

differentiating this j times with respect to α, and k times with respect to β gives part (a) of Theorem 6.11.

Part (c) of Theorem 6.11 can be deduced by a similar argument from part (b) of Theorem 6.10. Since, as mentioned earlier, the functionals F_m are determined by the spectrum of L, it follows that $F_m(u)$ are conserved along the flows $(6.37)_j$. This implies that F_m and n_j are in involution, which yields part (b) of Theorem 6.11. Part (d) follows by the argument used to prove part (d) of Theorem 6.10.

According to a result of A. Lenard (see equation (5.11) of [8]) the gradients G_m of the functionals F_m satisfy the recursion formula

$$(6.38)_m \qquad HG_m = 4\partial G_{m+1}\,.$$

(One has to replace u by $-6u$ in Lenard's formula to account for a different form of L used in [8].) This is exactly the recursion relation (6.36) satisfied by the functionals n_j, except for a change in sign in the index, showing that the sequence n_j can be regarded as a continuation of G_m to negative indices through the relation

$$(6.39) \qquad G_{-m} = n_{m-2}\,, \qquad m = 2, \cdots.$$

Therefore, the sequence of conserved functionals F_m too can be extended to negative values of the index

$$F_{-m} = \frac{6}{(m-2)!}\left(\frac{d}{d\alpha}\right)^{m-2} \log \kappa(\alpha)\Big|_{\alpha=0}. \tag{6.40}$$

We conclude these observations with some remarks on $\log \kappa$ as function of complex values of α:

$\log \kappa$ is a 2-valued function of α, with branchpoints of order 2 at the simple eigenvalues of L for functions of period p and $2p$. Furthermore, the growth of κ is exponential:

$$|\log \kappa(\alpha)| < \text{const.} + p\,|\alpha|^{1/2}. \tag{6.41}$$

Inequality (6.41) can be deduced from standard estimates for the theory of ordinary differential equations.

In his paper [23], Peter Ungar employs the function $\log \kappa$ to give a simple proof of the following theorem of Borg [2]:

If all but one of the eigenvalues of $L = \partial^2 + v$, acting on functions with period $2p$, are double, then v is constant.

In his proof, Ungar eliminates the branchpoint at the lowest eigenvalue λ_0 by introducing the function

$$f(z) = \log \kappa(\lambda_0 + z^2).$$

It follows from (6.41) that $f(z)$ has a simple pole at ∞. If λ_0 is the only simple eigenvalue, $f(z)$ is a linear function.

If there are in addition to λ_0 a finite number $2N$ of simple eigenvalues, then $f(z)$ is a double-valued analytic function with branch points of order 2 at the simple eigenvalues, and a simple pole at ∞; the theory of such functions plays an important role in the work of McKean and van Moerbeke.

The last 3 results of this section deal with the spectrum of L when u is one of the special potentials satisfying an equation of the form (3.9).

The next theorem is a slight sharpening of one due to J. Moser and the author given in [14]:

THEOREM 6.12. *Let u be a solution of* (3.9):

$$G(u) = \sum_{-1}^{N} a_j G_j(u) = 0, \qquad a_N = 1. \tag{6.42}$$

Then all but $2N+1$ eigenvalues of the operator L given by (6.11) *are double.*

Proof: Multiply equation $(6.10)_j$ by a_j/c_j and sum. We introduce the abbreviation

$$\sum \frac{a_j}{c_j} B_j = B ; \tag{6.43}$$

using (6.12), we can write the resulting relation as

$$BL - LB = \partial \sum a_j G_j . \tag{6.44}$$

If u satisfies (6.42),

$$BL - LB = 0 , \tag{6.45}$$

i.e., B commutes with L. Now suppose that λ is a simple eigenvalue of L, with the corresponding eigenvector w. It follows from (6.45) that

$$LBw = BLw = \lambda Bw ,$$

i.e., that Bw is also an eigenfunction of L. Since λ is assumed simple,

$$Bw = \kappa w . \tag{6.46}$$

B, being a linear combination of antisymmetric operators, is itself antisymmetric; so its spectrum is pure imaginary. Therefore, κ in (6.39), which is real, is zero:

$$Bw = 0 . \tag{6.47}$$

Thus all eigenfunctions with simple eigenvalues satisfy the *same* equation (6.47). Now eigenfunctions corresponding to distinct eigenvalues are linearly independent; on the other hand, B being a linear differential operator of order $2N+1$ cannot have more than $2N+1$ linearly independent solutions. This proves Theorem 6.12.

Remark. I suspect, but cannot prove, that the $2N+1$ simple eigenvalues are the lowest ones. Since the lowest eigenvalue is always simple, this is the case for $N=0$. In [14], an argument was given to show that for $N=1$ the lowest 3 eigenvalues are indeed simple. The argument is based on the observation that the eigenfunction w, in addition to satisfying the equation (6.47) of order $2N+1$, also satisfies the second-order equation (6.21). This equation can be used to express all derivatives of w of any order as a linear combination of w and w_x. Using this in (6.47) yields a first order equation for w of the form

$$aw_x + bw = 0\,, \tag{6.48}$$

where $a = a(\lambda, x)$ is a polynomial in λ of degree N whose leading coefficient is 1.

All but the lowest eigenfunctions w have zeros; in fact the $(2m-1)$-st and $2m$-th have $2m$ zeros. Let x_0 be such a zero: $w(x_0) = 0$. Since w satisfies $Lw = \lambda w$, and is not equal to 0, $w_x(x_0) \neq 0$; it follows then from (6.48) that

$$a(\lambda, x_0) = 0\,. \tag{6.49}$$

Since a is a polynomial in λ, we can get from this relation an upper bound for the simple eigenvalues λ in terms of estimates for u and its derivatives up to order $2N-1$.

We return now to Theorem 6.12. We have seen in Section 3, equation (3.19), that solutions of (3.9) are elliptic functions; it follows therefore from Theorem 6.12 that if the potential u occurring in L of the form (6.11) is such an elliptic function, then all but the first 3 eigenvalues of L are double. This fact is known (see Magnus and Winkler [16]). In a remarkable paper [11], H. Hochstadt has raised and answered the converse proposition: i.e., he has shown that if all but the first 3 eigenvalues of L are double, then the potential u is an elliptic function. Furthermore, he has shown that if L has only a finite number of simple eigenvalues, then u is C^∞.

In [14], after proving that if u satisfies (3.9), then L has only a finite number of simple eigenvalues, I raised the question whether the converse might be true. This has been answered affirmatively by W. Goldberg [9], using the method of Hochstadt. Another proof has been given by H. Flaschka [4]. Here is yet another proof of this fact by McKean and van Moerbeke, based on the following observation:

If L_0 has only $2N+1$ simple eigenvalues, the set of potentials u for which L has the same spectrum as L_0 forms an N-parameter family.

The proof of this relies on Borg's uniqueness theorem in [2].

THEOREM 6.13. *Suppose L_0 has only $2N+1$ simple eigenvalues; then u_0 satisfies an equation of the form* (6.42).

Proof: According to the theorem of Hochstadt quoted above, such a u_0 is C^∞; therefore u_0 belongs to the domain of each $S_m(t)$, the solution operator of the m-th generalized KdV flow. It follows from Theorem 6.3 that if we take u to be

$$u = u(t_0, \cdots, t_N) = \prod_0^N S_m(t_m)u_0 , \tag{6.50}$$

then L is unitarily equivalent with L_0. According to the observation quoted above, the set of such u forms an N-parameter family. The parameters are continuous functions of the $N+1$ parameters appearing in (6.50); hence the parameters appearing in (6.50) cannot all be independent. This is so if and only if the directions in which S_m start are linearly dependent; these directions are $G_m(u_0)$, and therefore there is a linear relation among the $G_m(u_0)$:

$$\sum_0^N a_m G_m(u_0) = 0 .$$

This is an equation of the form (6.42). We note that a_N cannot be zero; otherwise, according to Theorem 6.12, the operator L_0 would have only $2J+1$ simple eigenvalues, where J is the index of the highest nonzero a_m. This completes the proof of Theorem 6.13.

Theorem 6.12 shows that if u satisfies equation (6.42), the eigenfunctions of L corresponding to simple eigenvalues satisfy a differential equation (6.47). It is remarkable that also the squares of eigenfunctions satisfy an equation of the same kind.

THEOREM 6.14. *Suppose u is a solution of* (6.42):

$$G(u) = 0 .$$

L denotes the operator (6.11).

(a) *Let w be an eigenfunction of L corresponding to a simple eigenvalue. Then*

$$G'(u)\,\partial w^2 = 0 . \tag{6.51$_1$}$$

(b) *Let w_1 and w_2 be a pair of orthonormal eigenfunctions of L corresponding to a double eigenvalue. Then*

$$G'(u)\,\partial(w_1^2 + w_2^2) = 0 . \tag{6.51$_2$}$$

Proof: We appeal to Theorem 2.2 about time dependent solutions of nonlinear equations of the form

$$u_t = K(u)\ ;$$

according to Theorem 2.2, if $F(u)$ is a conserved quantity for this equation, its gradient $G_F(u)$ satisfies the linear equation

(6.52) $$\left(\frac{\partial}{\partial t} + M^*\right) G_F = 0\ ,$$

where M^* is the adjoint of M, the derivative of κ defined in (2.8). In our present situation where u satisfies $G(u) = 0$, we regard u as time independent solution of the evolution equation

(6.53) $$u_t = \partial G(u)\ .$$

For this equation, $K = \partial G(u)$ whose derivative is $M = \partial G(u)$. Since by Theorem 2.3 G' is selfadjoint,

$$M^* = -G' \partial\ .$$

So it follows from (6.52) that if F is a conserved functional,

(6.54) $$G'(u)\, \partial G_F = 0\ .$$

According to Theorem 6.3, the eigenvalues of L are conserved functionals for (6.53); according to Lemma 6.5, if λ is differentiable, $G_\lambda = \frac{1}{6} w^2$. Setting this into (6.54) we obtain $(6.51)_1$. According to the remark following Lemma 6.5, for a double eigenvalue, $\lambda_1 + \lambda_2$ is differentiable, and

$$G_{\lambda_1 + \lambda_2} = \tfrac{1}{6}(w_1^2 + w_2^2)\ .$$

Setting this into (6.47) we obtain $(6.51)_2$; this completes the proof of Theorem 6.14.

For G of the form (6.42), G' is an operator of order $2N+1$; therefore its null space is at most $(2N+1)$-dimensional. It follows therefore that there is a linear relation between $w_0^2, \cdots, w_{2N}^2$, the squares of the simple eigenfunctions, and $w_{(1)}^2 + w_{(2)}^2$ for any double eigenfunction. I suspect that the latter can be expressed as a linear combination of the former. This is certainly so in the trivial case $N = 0$, when $w_0^2 = 1$, and $w_{(1)} = \sin nx$, $w_{(2)} = \cos nx$.

Appendix

James M. Hyman

In this appendix we describe how the construction of special solutions of the KdV equation which minimize F_2 subject to the constraint $F_i = A_i$ was implemented numerically. We saw in Section 3 that a minimizing function satisfies the Euler equation

$$G_2 + \sum_{-1}^{1} a_i G_i = 0 \,. \tag{A.1}$$

A solution of (A.1) is an extremal for

$$F = F_2 + \sum_{-1}^{1} a_i F_i \,, \tag{A.2}$$

and presumably can be obtained by minimizing that functional. This is indeed what we did. We chose the constants a_{-1}, a_0, a_1, then discretized the functional (A.2) by specifying u at N equidistant points and expressed the first and second derivatives of u in (A.2) by difference quotients. The resulting function of N variables was minimized using A. Jameson's version[1] of the Fletcher-Powell-Davidov algorithm [5]. The resulting discretized function is a somewhat crude approximation to the function, u, we are looking for. To get u more accurate we proceeded as follows:

We are looking for a periodic solution of (A.1); since this is a fourth order equation, its solutions are parametrized by their four Cauchy data at, say, $x = 0$. Periodicity requires the Cauchy data at $x = p$ to be equal to the Cauchy data at $x = 0$. We sought to satisfy this requirement by a sequence of approximations constructed by "shooting". The Cauchy data of u_{n+1} at $x = 0$ were chosen by applying the rule of false position to $u_n(p) - u_n(0)$, $u_n'(p) - u_n'(0)$, $u_{n-1}(p) - u_{n-1}(0)$ and $u_{n-1}'(p) - u_{n-1}'(0)$. Then $u_{n+1}(p)$ and $u_{n+1}'(p)$ were computed numerically. For this purpose we used the ODE package developed by A. Hindmarsh [10].

The following observations were helpful:

(a) Instead of matching all four Cauchy data for the fourth order equation (A.1), it sufficed to match only two, since according to Lemma 5.12 of Section 5 there exist two polynomials J_0 and J_1 of the Cauchy data which are independent of x.

(b) The initial guesses for the Cauchy data come from the approximate

[1] My thanks are due to A. Jameson for acquainting me with his FPD package.

solution obtained by the variational procedure. Without such a good initial guess we were unable to construct a periodic solution of (A.1).

The periodic solution of (A.1) constructed above was then used as initial value data. We solved numerically the initial value problem for the KdV equation using Fred Tappert's[2] method [22] and code. The accuracy of the solution was monitored by checking the constancy of the functionals F_0, F_1, F_2, J_0 and J_1, and the extent to which the solution satisfied the ODE (A.1). The solution constructed by Tappert's method passed these tests of accuracy reasonably well. Solutions constructed using earlier methods of Kruskal and Zabusky [25] and Vliegenthart [24] were less accurate and were not used in this study. The finest vindication of Tappert's KdV solver was that after a finite elapse of time the solution resumed its initial shape, in a shifted position. During the intervening time the shape of the solution underwent considerable gyrations.

It was proved rigorously at the end of Section 5 that the shape of the solution recurs exactly after a finite time. Here is another proof, based on an idea of Lax in [13]:

According to Lemma 5.13 a solution $u(x, t)$ of KdV whose initial values satisfy equation (A.1) also satisfies

$$J_0 = c_0 , \qquad J_1 = c_1 , \tag{A.3}$$

where c_0 and c_1 are constants and J_0 and J_1 are polynomials in u and its x derivatives up to order 3. J_0 and J_1 can be calculated explicitly from formulas (3.17), (5.17) and (5.24).

Equations (A.3) can be solved to express u_{xx} and u_{xxx} in terms of u and u_x. These expressions become particularly simple when $u_x = 0$; they are of the form

$$u_{xx} = P^{1/2}(u) , \tag{A.4}$$

and

$$u_{xxx} = R^{1/2}(u) , \tag{A.5}$$

where P is a polynomial in u and R is a polynomial in u and $P^{1/2}$.

Let $m = m(t)$ denote the value of the maximum of $u = u(x, t)$ with respect to x. At the point $y = y(t)$ where the maximum is achieved,

$$u_x(y(t), t) = 0 . \tag{A.6}$$

Therefore formulas (A.4) and (A.5) are valid at this point. Note that since at a maximum $u_{xx} < 0$, the square root in (A.4) has to be taken as the negative root.

[2] My thanks are due to Fred Tappert for acquainting me with his KdV solver.

Differentiating

$$m(t) = u(y(t), t)\,,$$

we get

$$m_t = u_x y_t + u_t\,.$$

From the KdV equation we have

$$u_t = -uu_x - u_{xxx}\,.$$

Combining these two and using (A.6) we obtain

$$m_t = -u_{xxx}\,.$$

Using (A.5) we deduce that

$$m_t = -R^{1/2}(m)\,. \tag{A.7}$$

Solutions of (A.7) behave as follows:

If at $t = 0$ the value of u_{xxx} is greater than 0 at $y(0)$, then $m(t)$ decreases until it reaches the nearest zero of $R(m)$. After that u_{xxx} changes sign and $m(t)$ starts increasing until it reaches the next zero of $R(m)$. After that u_{xxx} changes sign again and $m(t)$ starts decreasing until it reaches the value $m(0)$. At this time T, and at the point of $y(T)$, the Cauchy data u, u_x, u_{xx} and u_{xxx} have the same value as at $t = 0$ and at $y(0)$. Since both $u(x, 0)$ and $u(x, T)$ are solutions of (A.1), it follows that

$$u(x, T) = u(x - L, 0)\,, \tag{A.8}$$

where

$$L = y(T) - y(0)\,. \tag{A.9}$$

Since equation (A.7) can be integrated by quadrature, T can be expressed as

$$T = 2\int_a^b \frac{dm}{R^{1/2}(m)}\,, \tag{A.10}$$

where

$$R(a) = R(b) = 0\,.$$

Specifically, a and b are those zeros of $R(m)$ between which the maximum of $u(x, 0)$ is located.

The value of $L(T)$ can also be determined by quadrature. Differentiating (A.6) with respect to t gives

$$u_{xx} y_t + u_{xt} = 0 ,$$

so that

$$y_t = -\frac{u_{xt}}{u_{xx}} . \tag{A.11}$$

Differentiating KdV with respect to x gives

$$u_{tx} = -u_x^2 - u u_{xx} - u_{xxxx} .$$

Substituting this into (A.11) we get, using $u_x = 0$,

$$y_t = u + \frac{u_{xxxx}}{u_{xx}} . \tag{A.12}$$

Using (A.1) we can express u_{xxxx} in terms of derivatives of u of lower order. The values of these at $x = y(t)$ have already been expressed in terms of m; so from (A.12) we can get a relation of the form

$$y_t = Y(m) , \tag{A.13}$$

where $Y(m)$ is an explicitly computable function of m. Integrating (A.13) gives, using (A.7),

$$L = y(T) - y(0) = \int_0^T y_t \, dt = \int_0^t Y(m) \, dt$$

$$= \int Y(m) \frac{dm}{dt} \, dt = \int Y(m) R^{1/2}(m) \, dm .$$

Hence, altogether,

$$L = 2 \int_a^b Y(m) R^{1/2}(m) \, dm . \tag{A.14}$$

We conclude by presenting some numerical data:

Figures 1 and 2 give the graphs of functions minimizing $F(u)$ of the form

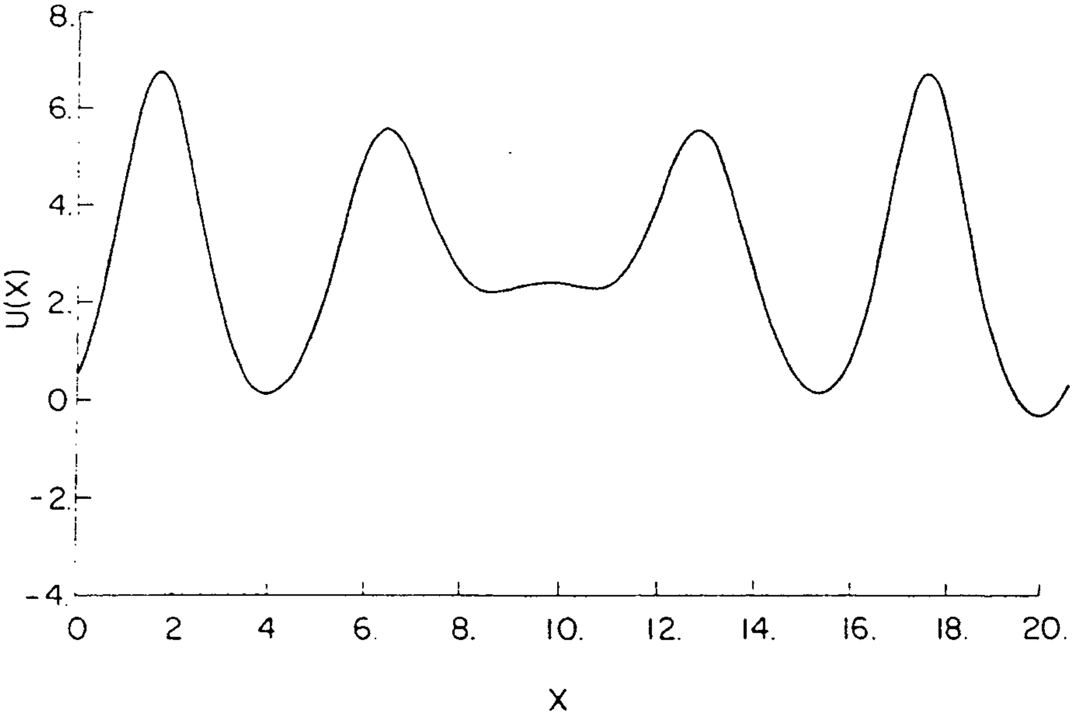

Figure 1. A periodic solution to equation (A.1) with $a_{-1}=0$, $a_0=-5$, $a_1=-2$.

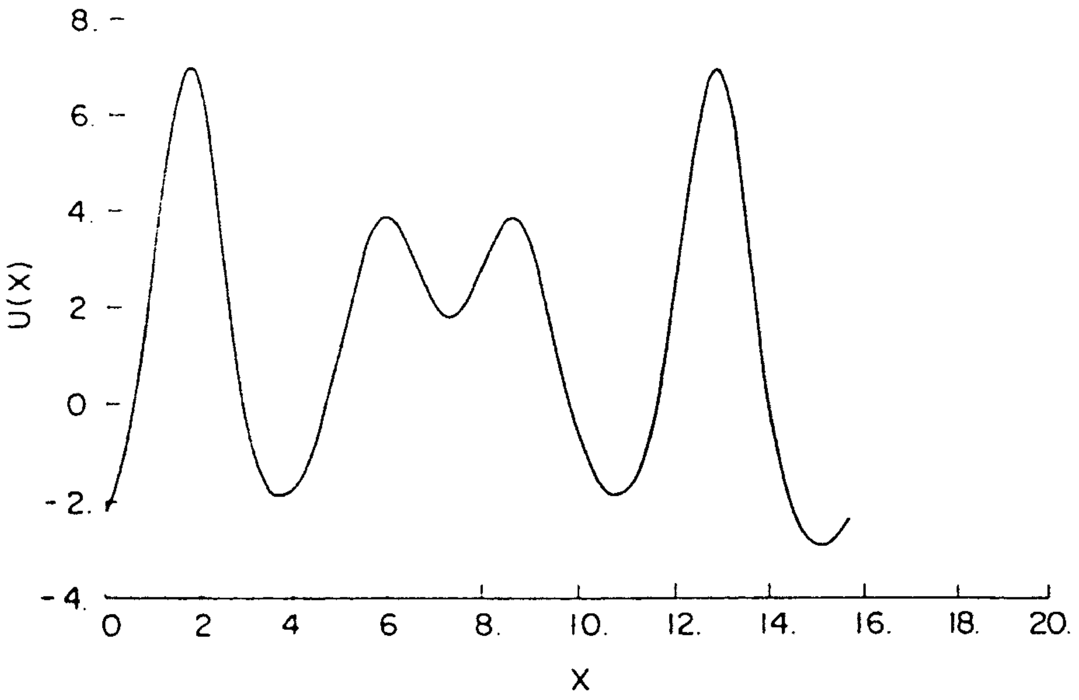

Figure 2. A periodic solution to equation (A.2) with $a_{-1}=0$, $a_0=-8$, $a_2=2$.

(A.2), with the following choice of parameters, time periods and shift:

Case 1: $p=20.7$, $a_{-1}=0$, $a_0=-5$, $a_1=-2$,
$T=1.11$, $L=-3.3$,

Case 2: $p=15.7$, $a_{-1}=0$, $a_0=-8$, $a_2=2$,
$T=1.11$, $L=-5.1$.

In a forthcoming publication we shall describe some further details of the time history of $u(x,t)$.

Bibliography

[1] Ablowitz, M. J., and Newell, A. C., *The decay of the continuous spectrum for solutions of the Korteweg-de Vries equation*, J. Math. Phys., 14, 1973, pp. 1277–1284.

[2] Borg, G., *Eine Umkehrung der Sturm-Liouvilleschen Eigenwertaufgabe*, Acta Math., Vol. 78, 1946, pp. 1–96.

[3] Faddeev, L., and Zakharov, V. E., *Korteweg-de Vries equation as completely integrable Hamiltonian system*, Funktsional. Anal. i Prilozhen. 5, 1971, pp. 18–27. (In Russian.)

[4] Flaschka, H., *Integrability of the Toda lattice*, Theoretical Physics, Vol. 51, 1974, p. 703.

[5] Fletcher, R., and Powell, M. J. D., *A rapidly convergent descent method for minimization*, Comput. J., Vol. 6, 1963/64, pp. 163–168.

[6] Gardner, C. S., *Korteweg-de Vries equation and generalizations. IV. The Korteweg-de Vries equation as a Hamiltonian system*, J. Mathematical Phys. 12, 1971, pp. 1548–1551.

[7] Gardner, C. S., Greene, J. M., Kruskal, M. D., and Miura, R. M., *Method for solving the Korteweg-de Vries equation*, Phys. Rev. Letters, 19, 1967, pp. 1095–1097.

[8] Gardner, C. S., Greene, J. M., Kruskal, M. D., and Miura, R. M., *Korteweg-de Vries equation and generalizations. VI. Methods for exact solutions*, Comm. Pure Appl. Math., Vol. XXVII, 1974, pp. 97–133.

[9] Goldberg, Wallace, *On the determination of Hill's equation from its spectrum*, Bull. Amer. Math. Soc., 80, 1974, pp. 1111–1112.

[10] Hindmarsh, A., *GEAR: Ordinary differential equation system solver*, Lawrence Livermore Laboratory, UCID-30001, 1972.

[11] Hochstadt, H., *On the characterization of a Hill equation via its spectrum*, Arch. Rational Mech. Anal., 19, 1965, pp. 353–362.

[12] Korteweg, D. J., and de Vries, G., *On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves*, Phil. Mag., 39, 1895, pp. 422–443.

[13] Lax, P. D., *Integrals of nonlinear equations of evolution and solitary waves*, Comm. Pure Appl. Math., 21, 1968, pp. 467–490.

[14] Lax, P. D., *Periodic Solutions of the KdV Equation*, Lectures in Applied Math., Vol. 15, AMS, Providence, Rhode Island, 1974.

[15] Lions, J. L., and Lattès, R., *The Method of Quasi-Reversibility*, American Elsevier, 1969.

[16] Magnus, W., and Winkler, S., *Hill's Equation*, Interscience, Tracts in Pure and Appl. Math., No. 20, Wiley, New York, 1966.

[17] McKean, H., and van Moerbeke, P., *The spectrum of Hill's equation*, to appear in Invenciones Mat.

[18] Miura, R., Gardner, C. S., and Kruskal, M. D., *Korteweg-de Vries equation and generalizations. II. Existence of conservation laws and constants of motion*, J. Math. Phys., 9, 1968, pp. 1204–1209.

[19] Novikov, S. P., *The periodic problem for the Korteweg-de Vries equation, I*, Funk. Anal. Prilozh., 8, No. 3, 1974, pp. 54–66.

[20] Sjoberg, A., *On the Korteweg-de Vries equation*, J. Math. Anal. Appl., 29, 1970, pp. 569–579.

[21] Tanaka, S., *Korteweg-de Vries equation: Asymptotic behavior of solutions*, to appear.

[22] Tappert, F., *Numerical Solution of the KdV Equation and its Generalizations by the Split-Step Fourier Method*, Lectures Appl. Math., Vol. 15, AMS, Providence, Rhode Island, 1974.

[23] Ungar, P., *Stable Hill equation*, Comm. Pure Appl. Math., Vol. XIV, 1961, pp. 707–710.

[24] Vliegenthart, A. C., *On finite difference methods for the Korteweg-de Vries equation*, J. of Engineering Math., Vol. 5, 1971, pp. 137–155.

[25] Zabusky, N. J., and Kruskal, M. D., *Interaction of "solitons" in a collisionless plasma and the recurrence of initial states*, Phys. Rev. Letters, 15, 1965, pp. 240–243.

Received November, 1974.