Korteweg-deVries Equation and Generalizations. V. Uniqueness and Nonexistence of Polynomial Conservation Laws

Martin D. Kruskal, Robert M. Miura, and Clifford S. Gardner
Plasma Physics Laboratory, Princeton University, Princeton, New Jersey

and

Norman J. Zabusky
Bell Telephone Laboratories, Inc., Whippany, New Jersey

(Received 5 September 1969)

The conservation laws derived in an earlier paper for the Korteweg-deVries equation are proved to be the only ones of polynomial form. An algebraic operator formalism is developed to obtain explicit formulas for them.

1. INTRODUCTION

Research in recent years exhibits an increase of interest in nonlinear dispersive wave phenomena. One of the simplest nonlinear dispersive wave equations is the Korteweg-deVries (KdV) equation

\[ u_t + uu_x + u_{xxx} = 0, \tag{1} \]

where subscripts denote partial differentiations. This equation is conservative and dispersive, in sharp contrast to the much studied Burgers’ equation which would obtain if \( u_{xxx} \) were replaced by \( -uu_x \), \( \nu > 0 \). Burgers’ equation can describe a constant pressure, incompressible, viscous fluid. In most applications (see Papers I and III for references to physical applications of the KdV equation), a dispersion parameter appears as a coefficient of the \( u_{xxx} \) term in the KdV equation. However, there is no loss of generality (for present purposes) in confining ourselves to (1), because of the scaling properties discussed in Appendix A.

In Paper II we proved that the KdV equation possesses an infinite sequence of polynomial conservation laws in the form

\[ T_t + X_x = 0, \]

where \( T \), the conserved density, and \( X \), the flux of \( T \), are polynomials (not explicitly dependent on \( x \) or \( t \)) in \( u \) and its derivatives. In this paper, we give a detailed discussion of these conservation laws (also called conservation equations). We frequently use the abbreviations c.d., p.c.d., and c.l. for conserved density, polynomial conserved density, and conservation law.

Two c.l. for the KdV equation are obtained immediately,

\[ u_t + (\frac{1}{4}u^3 + u_{xxx}) = 0, \]
\[ (\frac{1}{4}u^3)_{t} + (\frac{1}{4}u^3 + uu_{xx} - \frac{1}{4}u^3) = 0. \]

The first is simply the equation itself rewritten, and the second is obtained after multiplying by \( u \) and rewriting.

Whitham\(^{3}\) needed a third c.l. to study nonlinear dispersive waves by a method of averaging, hence he sought (and found) one, with a term \( u^3 \) in its p.c.d. A fourth c.l., with a term \( u^4 \) in its p.c.d., was found by Kruskal and Zabusky.\(^{3}\) In an effort to understand the results of certain numerical computations, they developed an asymptotic theory for the KdV equation when a small parameter \( \delta^2 \) multiplies the \( u_{xxx} \) term. If \( \delta^2 \) is set equal to zero, then for a large class of (smooth) initial data, the solution becomes discontinuous after a finite time due to the nonlinearity. However, if \( \delta^2 \) is small but not zero, then, as the solution tends to become discontinuous, the \( u_{xxx} \) term becomes important in spite of its small coefficient and keeps the solution smooth. Moreover, small-wavelength finite-amplitude oscillations develop, initially occupying a small region at the near-discontinuity but then spreading indefinitely. (For Burgers’ equation with \( \nu \to 0 \) no such oscillations develop, but rather a conventional hydrodynamic shock structure.)

The asymptotic theory of the KdV equation developed by Kruskal and Zabusky in this region of oscillations is a nonlinear generalization of the WKB method. They obtained a system of equations for the evolution of the solution on the “macroscopic” scale (together with a detailed description of the oscillations on the “microscopic” scale). Solutions of the limit system \( (\delta^2 \to 0) \) exhibit discontinuities of an unusual type which may be considered as “reversible (nondissipative) shock waves.” A count of the appropriate number of jump conditions across such shock waves led to the search for and discovery of the fourth conservation equation and to the investigation presented here.

Definitions and operators are introduced in Sec. 2 along with a number of commutation formulas. In Sec. 3 we prove that the KdV equation has no p.c.d. other than those established in Paper II.

952
Further, we outline the method of undetermined coefficients used to establish them there. However, this method does not yield an explicit formula for all p.c.d.; in Sec. 4 we find such formulas by other means.

We had originally intended to include a section proving uniqueness and nonexistence theorems for some generalizations of the KdV equation. However, the material has accumulated to such an extent as to justify our collecting it into an eventual separate Paper VII in this series.

2. DEFINITIONS AND OPERATORS

In this paper we deal extensively with polynomials and eventually formal infinite power series in \( u \) and its \( x \) derivatives. In fact, the words polynomial and series are always to be understood in this sense. Here we introduce some definitions to be used throughout.

We denote the \( t \) and \( x \) differentiation operators by

\[
\frac{\partial}{\partial t} \quad \text{and} \quad \frac{\partial}{\partial x},
\]

and the \( i \)th-order \( x \) derivative of \( u \) by \( u_i \). In much of the following we treat the \( u_i \) as independent variables.

Any term in a polynomial (or series) has the form \( c_a u_i u_j \cdots u_l \) where \( a_i \geq 0 \); \( i \) is the order of the term. (Throughout this paper \( c \) and \( c_i \), \( i = 0, 1, \cdots \), denote arbitrary constants.) Of any two distinct terms we call that one dominant which has larger \( l \), or the same \( l \) but larger \( a_i \), or the same \( l \) and \( a_i \) but larger \( a_{i-1} \), etc. Obviously dominance is an ordering relation (it is transitive and irreflexive) and is complete (of any two distinct terms, one dominates the other).

The total number of factors in a term is its degree

\[
m \equiv \sum_{i=0}^{l} a_i.
\]

The total number of differentiations will be called the derivative index (or simply index)

\[
n \equiv \sum_{i=1}^{l} ia_i.
\]

The rank of a term (introduced in Paper II) is defined by

\[
r \equiv m + \frac{1}{2} n = \sum_{i=0}^{l} (1 + \frac{1}{2} i) a_i;
\]

t a term scales as the \((-2r)\)th power of \( c \) under the scale transformation (see Appendix A). For example, the last two terms in the KdV equation are of equal rank \( \frac{1}{2} \). From the above definitions it is clear that acting on a term with the operator \( \mathcal{D} \) leaves the degree of that term unchanged but raises its index by 1 and its rank by \( \frac{1}{2} \), whereas acting on a term with \( \partial_i \) raises the rank by \( \frac{1}{2} \) but has a nonuniform effect on the degree and on the index. A polynomial of degree \( m \) has \( m \) as the l.u.b. of the degrees of all its terms, and similarly for the order, index, and rank.

Associated with the three "labelings" degree, index, and rank are the linear differential operators

\[
\mathcal{M} \equiv \sum_{i} u_i \partial_i,
\]
\[
\mathcal{N} \equiv \sum_{i} iu_i \partial_i,
\]
\[
\mathcal{R} \equiv \mathcal{M} + \frac{1}{2} \mathcal{N} = \sum_{i} (1 + \frac{1}{2} i) u_i \partial_i,
\]

where \( \sum_{i} \) denotes \( \sum_{i=0}^{\infty} \) and \( \partial_i \equiv \partial/\partial u_i \). (For \( i \) negative, we shall interpret \( \partial_i \) as zero.) The effect of applying \( \mathcal{M} \), \( \mathcal{N} \), or \( \mathcal{R} \) to any term is to multiply that term by its degree, index, or rank. Thus any term, indeed any polynomial of uniform degree (i.e., with all terms of the same degree), is an eigenpolynomial of \( \mathcal{M} \) with eigenvalue \( m \), and similarly for \( \mathcal{N} \) and \( \mathcal{R} \). For each of these labelings we define slice operators, i.e., projection operators that select terms of a certain degree, index, or rank. Thus, for example, the \( m \)-degree slice operator \( \mathcal{M}_{m} \), acting on a polynomial, leaves unaffected all terms of degree \( m \) and annihilates all terms of degree different from \( m \). (Because of the eigenproperty, such an \( m \)-degree slice operator could be defined formally as \( \mathcal{M}_{m} \equiv \partial_{\mathcal{M},m} \), where the right side is the Kronecker delta function, here equal to unity if the eigenvalue of \( \mathcal{M} \) operating on a term is \( m \), and otherwise zero.)

Besides \( \mathcal{R} \), we have occasion to consider other operators linear in \( \mathcal{M} \) and \( \mathcal{N} \). The operator \( c_1 \mathcal{M} + c_2 \mathcal{N} + c_3 \) has the effect of multiplying any term by the number \( c_1 m + c_2 n + c_3 \). We define \( (c_1 \mathcal{M} + c_2 \mathcal{N} + c_3)^{-1} \) to be the operator whose effect is to divide any term by \( c_1 m + c_2 n + c_3 \), or annihilate any term for which this number vanishes. If there are no annihilated terms we have a true inverse, otherwise a pseudo-inverse; in any case it commutes with the original operator.

In Sec. 4 we use series (of nonnegative powers) extensively. Many of the properties (e.g., vanishing and equality) and operations (e.g., addition, multiplication, differentiation, integration, exponentiation, \( \mathcal{M}, \mathcal{N}, \mathcal{R} \), and the slicings) generalize naturally and obviously from polynomials to series. A series is considered to have a property if arbitrarily high polynomial truncates of it do; in particular, it vanishes only if all its coefficients do. The result of an operation on a series consists of all terms common to the results of the same operation on arbitrarily high truncates;
for all the parenthetically mentioned operations, except multiplication and exponentiation, this amounts to operating term by term. The definitions of order, degree, index, and rank for polynomials have been so expressed as to apply also to series, now allowing infinity as a value. (A series of finite rank is necessarily a polynomial.)

Expressed as a differential operator in the space of functions of the \( u_i \), the \( x \) differentiation operator is

\[
\mathcal{D} = \sum_i u_{i+1} \partial_i. \tag{8}
\]

**Lemma 1:** A polynomial \( P \) of degree \( m \) and order \( l \) is either (a) a constant \( (m = 0) \), whereupon \( \mathcal{D}P = 0 \), or (b) not a constant \( (m > 0) \), whereupon \( \mathcal{D}P \neq 0 \) is of order \( l + 1 \) and is linear in \( u_{l+1} \).

**Proof:** Case (a) is obvious. Case (b) follows immediately from (8), the dominant term in \( \mathcal{D}P \) coming from \( i = l \).

Using the KdV equation (1), the \( t \) differentiation operator can be written

\[
\partial_t = \sum_i (\partial_i u_i) \partial_i = -\sum_i [\mathcal{D}(u_0 u_i + u_2) \partial_i]
= -(\mathcal{B} + \mathcal{D}_0), \tag{9}
\]

where

\[
\mathcal{B} \equiv \frac{1}{2} \sum_{i+j=0}^{i+j+1} \binom{i+j}{i} u_j u_{i-j+1} \partial_i, \tag{10}
\]

\((i+j)^{+1}\) being the binomial coefficient, and \( \mathcal{D}_0 \) is a special case of

\[
\mathcal{D}_0 \equiv \sum_i u_{i+1} \partial_i. \tag{11}
\]

Other special cases of (11) are

\[
\mathcal{D}_0 = \mathcal{M}, \quad \mathcal{D}_1 = \mathcal{D}.
\]

We present here a number of results not needed until Sec. 4. First, a number of commutation formulas all obvious or straightforwardly derived:

\[
\partial_i \mathcal{D}_j = \mathcal{D}_j \partial_i + \partial_{i-j}, \tag{12}
\]

\[
\mathcal{D}_i \mathcal{D}_j = \mathcal{D}_j \mathcal{D}_i, \tag{13}
\]

\[
\partial_i \mathcal{N} = (\mathcal{N} + i) \partial_i, \tag{14}
\]

\[
\mathcal{D}_i \mathcal{N} = (\mathcal{N} - i) \mathcal{D}_i, \tag{15}
\]

\[
\partial_0 \mathcal{B} = \mathcal{B} \partial_0 + \mathcal{D}, \tag{16}
\]

\[
\partial_i \mathcal{B} = \partial_i \mathcal{B} + \mathcal{M} + \mathcal{N}, \tag{17}
\]

\[
\mathcal{D}_0 \mathcal{B} = \mathcal{B} \mathcal{D}_0, \tag{18}
\]

\[
\mathcal{D} \partial_i = \partial_i \mathcal{D}. \tag{19}
\]

The commutativity of \( x \) and \( t \) differentiations expressed by (19) has already been used in (9).

We next introduce

\[
\mathcal{Y} \equiv \sum_i (\mathcal{D})^i \partial_i, \tag{20}
\]

the Euler operator of the calculus of variations. [For a function \( F(u_0, \cdots, u_l) \), \( \mathcal{Y}F \) is the functional derivative of \( \int F \, dx \), aside from boundary terms.] Since we are dealing here with formal algebraic expressions, we give a purely algebraic proof of the familiar

**Lemma 2:** If \( P \) is a series [polynomial] with no constant term, then \( P \) is the derivative of some series [polynomial] if and only if \( \mathcal{Y}P = 0 \).

**Proof:** If \( P = \mathcal{D}Q \), then (12) for \( j = 1 \) gives

\[
\mathcal{Y} \mathcal{D}Q = \sum_i (\mathcal{D})^i \partial_i \mathcal{D}Q = \sum_i (\mathcal{D})^i (\mathcal{D} \partial_i + \partial_{i-1})Q
\]

which is seen to vanish by shifting indices (since \( \partial_{-1} = 0 \)).

Conversely, if \( \mathcal{Y}P = 0 \), we multiply by \( u_0 \) to restore the degree and integrate by parts repeatedly, obtaining

\[
0 = u_0 \mathcal{Y}P = u_0 \sum_i (\mathcal{D})^i \partial_i P
= u_0 \partial_0 P - \mathcal{D} \left( u_0 \sum_{i=1}^\infty (\mathcal{D})^{i-1} \partial_i P \right)
+ \cdots + \mathcal{M} P - \mathcal{D} \sum_j u_j \sum_{i=j+1}^\infty (\mathcal{D})^{i-j-1} \partial_i P. \tag{21}
\]

We apply \( \mathcal{M}^{-1} \) and note that \( \mathcal{M}^{-1} \mathcal{M} P = P \) because only constant terms are annihilated by \( \mathcal{M} \), and \( P \) has none. Furthermore \( \mathcal{M}^{-1} \mathcal{M} \mathcal{M}^{-1} = \mathcal{M}^{-1} (\mathcal{D}, \mathcal{M}) \mathcal{M}^{-1} \),

while \( \mathcal{M}^{-1} \mathcal{M} = \mathcal{M} \mathcal{M}^{-1} = \mathcal{I} \) is the identity operator except for constant terms, which, by (8), do not occur in derivatives and are annihilated by \( \mathcal{D} \). Thus we have

\[
P = \mathcal{D} \left( \mathcal{M}^{-1} \sum_j u_j \sum_{i=j+1}^\infty (\mathcal{D})^{i-j-1} \partial_i P \right). \tag{22}
\]

Note that the infinite summations are (formally) well defined since we could have worked with derivative index slices, in each instance of which the summations would obviously be finite. For the same reason, if \( P \) is a polynomial, so is the expression in brackets.

We will want operators which when applied to powers of \( \mathcal{D} \) produce something simpler. We seek
such operators in the form $\partial \equiv \sum c_i (-\partial)^{-r} \partial_i$
(which reduces to $\partial$ for $c_i = 1$). This gives

$$
\partial \mathcal{D} = \sum_{i} c_i (-\partial)^{-r} (\partial \partial_i + \partial_{i-1})
$$

$$
= \sum_{i} (c_{i+1} - c_i) (-\partial)^{i+1-r} \partial_i,
$$

which suggests choosing $c_i = \binom{i}{r}$ and defining the operators

$$
\mathcal{Y}_i \equiv \sum_{i=r}^{\infty} \binom{i}{r} (-\partial)^{i-r} \partial_i.
$$

(23)

Then $\mathcal{Y}_r \mathcal{D} = \mathcal{Y}_{r-1}$ so that

$$
\mathcal{Y}_r \mathcal{D} = \mathcal{Y}_0 \equiv \mathcal{Y}.
$$

(24)

Incidentally, a basic property of the operators $\mathcal{Y}_i$ is stated in the following generalization of Lemma 2. However, we make no use of it.

**Lemma 3:** If $P$ is a series [polynomial] with no constant term, then $P$ is the $(r + 1)$-order derivative of some series [polynomial] if and only if $\mathcal{Y}_i P = 0$ for $s = 0, 1, \ldots, r$.

**Proof:** Straightforward, inductive, and omitted.

**3. POLYNOMIAL CONSERVED DENSITIES**

**A. Preliminaries**

Our general objective is to find conserved densities. The existence of an infinite sequence of p.c.d.s., one for each integral rank (here always meaning uniform rank), was proved in Paper II.

Since differentiations with respect to $x$ and $t$ commute, any $x$-derivative term is a trivial c.d. Given any c.d. we can immediately obtain any number of others simply by adding $x$-derivative terms. It is therefore natural to call two polynomials $P_1$ and $P_2$ equivalent if their difference is an $x$-derivative; we denote this by

$$
P_1 \Leftrightarrow P_2.
$$

(25)

(The terminology is justified by the obvious properties of reflexivity, symmetry, and transitivity.)

Among the many equivalents of a polynomial we seek a canonical representative. Let us integrate by parts to reduce order whenever possible. For instance $u_0 u_4 u_5 \Leftrightarrow -u_0 u_4 u_5 - u_0 u_4 \Leftrightarrow -\frac{1}{4} u_0^2 - u_4 u_5^2$. A polynomial is irreducible if no term has its highest-derivative factor $u_i$ occurring linearly, i.e., if each of its terms has order $l = 0$ or has $l \geq 1$ and $a_i \geq 2$, since then we cannot integrate by parts appropriately.

Our first two theorems state that the irreducible polynomials can serve as canonical representatives.

**Theorem 1 (Existence):** Every polynomial has an irreducible equivalent.

**Proof:** Appropriate integrations by parts can be applied successively only a finite number of times, since the order always decreases.

**Theorem 2 (Uniqueness):** Any two equivalent irreducible polynomials are equal.

**Proof:** Their difference, an irreducible derivative, vanishes by Lemma 1.

We also have occasion to speak of equivalent operators, meaning operators whose difference can be written in the form $DA$ for some operator $A$. We keep the same notation, so that, e.g., $U_0 \Leftrightarrow \partial_0$ by (20).

**B. Uniqueness of Conserved Densities of Rank $r$**

In this section we prove for each $r$ the uniqueness theorem that, up to multiplication by a constant, there is at most one irreducible p.c.d. $T_r$ of uniform rank $r$. We give two proofs, one of which shows further that the “extreme” term $u_0^r$ necessarily appears (i.e., with nonzero coefficient), and the other that so does the opposite extreme term $u_0^{r-1}$. From either of these, we deduce the nonexistence theorem that for half-integral $r$ there are no nontrivial p.c.d. at all.

We rely upon the structural

**Lemma 4:** The dominant term of the irreducible equivalent of the result of applying $B$ to an irreducible term $u_0^r \cdots u_0^i$ may be obtained by multiplying that term by $[-1 + a_1 + 3a_2 + 4a_3 + \cdots + (l + 1)a_l] u_1$, unless $l = 0$, in which case $Bu_0^r \Leftrightarrow 0$.

**Proof:** The result is obtained by straightforward calculation followed by the appropriate integration by parts to reduce order. It is unnecessary to give the details; in any case a slightly more general lemma will be presented in Paper VII.

A basic result of this section is

**Theorem 3:** If $T_r$ is any nonzero irreducible p.c.d. of uniform rank $r$ and $T_r^{(m)}$ its highest degree slice, then the dominant term $Q$ of $T_r^{(m)}$ is $cu_0^r$ ($c \neq 0$).

**Proof:** The highest degree slice of $(B + DA)T_r \Leftrightarrow 0$ is $3T_r^{(m)} \Leftrightarrow 0$. By Lemma 4, the dominant term of the irreducible equivalent of $3T_r^{(m)}$ would be just the dominant term of the irreducible equivalent of $3Q$, if the latter did not vanish. Therefore $3Q \Leftrightarrow 0$. Then by Lemma 4 again, $Q = cu_0^r$.
As immediate corollaries we have two of our main theorems.

**Theorem 4 (Uniqueness):** The KdV equation has only one linearly independent irreducible p.c.d. of rank $r$.

**Proof:** Given any two, we form a linear combination of them such that the $u^r_i$ terms cancel. This is then a p.c.d. with no $u^r_i$ term and hence vanishes, so the two are linearly dependent.

**Theorem 5 (Nonexistence):** The only p.c.d. of half-integral rank (e.g., containing terms such as $u^a_i u^b_i$ where $a_i$ is odd) for the KdV equation are trivial.

**Proof:** There is no polynomial term $u^r_i$ for $r$ a half-integer.

An alternative proof of these results relies on the structural

**Lemma 5:** The dominant term of the irreducible equivalent of the result of applying $D_u$ to an irreducible term $Q$ may be obtained by multiplying $Q$ by a nonzero constant (depending on $Q$) and replacing each of the three highest factors $u_i$, $u_{i+1}$, and $u_k$ ($k \leq l$) by $u_{i-1}$, $u_{i+1}$, and $u_{k+1}$, unless $Q$ has degree less than three, in which case $D_uQ = 0$.

**Proof:** Same comments apply as for Lemma 4.

Analogous to Theorem 3 is

**Theorem 6:** The dominant term of any nonzero irreducible p.c.d. of uniform rank $r$ is $tu^2_i$ if $r \geq 2$, $cu_i$ if $r = 1$, and $c$ if $r = 0$ ($c \neq 0$).

**Proof:** Analogous to that of Theorem 3, but, if anything, simpler because we need not take any kind of slice before selecting the dominant term.

It is perhaps worth remarking, as a corollary of the preceding results, that the only irreducible p.c.d. of the “truncated” KdV equation $u_i + u_{i-2} = 0$ are polynomials in $u_0$ alone. Similarly, the only irreducible p.c.d. of the alternatively truncated (i.e., linearized) KdV equation $u_i + u_{i-2} = 0$ are linear combinations of $1$, $u_0$, and $u^2_i$, $i = 0, 1, 2, \ldots$.

**C. Method of Undetermined Coefficients**

In Paper II we exhibited ten irreducible p.c.d. for the KdV equation and seven of the associated fluxes. (An eleventh is given in Appendix B.) Here we outline the method of undetermined coefficients (MUC) used to obtain those p.c.d. and fluxes.

We write $T_r$ as a general linear combination of all possible irreducible terms of rank $r$. Then we seek to determine the coefficients such that $\partial_x T_r$ will be equivalent to zero, i.e., equal to an $x$ derivative. By Theorem 3 $T_r$ contains the term $u^r_i$, the coefficient of which we choose to be $1/r$ for computational convenience. (Later we will choose it to be $1/r!$ instead. Alternatively, we could have fixed the coefficient of $u^r_{i-2}$, by Theorem 6.)

We next work out $\partial_x T_r$, use the KdV equation to eliminate explicit $t$ derivatives, and convert the result to its irreducible equivalent by the integration-by-parts algorithm described in Sec. 3A. We demand that the final expression vanish identically, and obtain thereby conditions on the original coefficients.

To illustrate MUC consider the case $r = 6$. All possible irreducible terms are represented in

$$T_6 = \frac{1}{6} u^6_0 + b_1 u^2_1 u^4_0 + b_2 u^2_0 u^4_1 + b_3 u^4_0 + b_4 u^6_0,$$

$$+ b_5 u^4_2 + b_6 u^6_2 + b_7 u^8_0.$$

The first term, for instance, gives

$$\partial_x (\frac{1}{6} u^6_0) = u^6_0 (-u_0 u_1 - u_2)$$

$$= -D(u^6_0 + u^2_1 u^4_0 - \frac{1}{2} u^4_1 u^2_0) - 10 u^2_0 u^4_1,$$

which contributes only $-10 u^2_0 u^4_1$ to the final irreducible expression which is to vanish. We eventually obtain an (overdetermined) inhomogeneous linear algebraic system of seven equations for the six $b_i$ which turns out to have a unique solution. The result is

$$T_6 = \frac{1}{6} u^6_0 - 10 u^2_0 u^4_1 + 18 u^2_0 u^2_1 - 5 u^4_1$$

$$- 18 u^2_1 u^2_0 + \frac{1}{2} u^4_0 + 3 u^2_0 u^2_2.$$

It is evident that in obtaining the p.c.d., we simultaneously obtain the corresponding fluxes, so long as we keep all the $\partial_x$ terms.

An interesting aspect of MUC for obtaining irreducible p.c.d. is that for $r \geq 6$, the number of equations exceeds the number of unknowns $b_i$. (Nevertheless, in addition to the uniqueness assured by Theorem 4, the theory in Paper II guarantees the existence of a solution.) Table I lists, for each rank $r$ up to 20, the number of undetermined coefficients (which is one less than the number of irreducible terms of that rank) and the number of possible equations (i.e., the number of irreducible terms of rank $r + \frac{3}{2}$). Up to $r = 11$, every possible $r + \frac{3}{2}$ term actually occurs and so yields an equation, and presumably this is true in general.
Table I. Number of irreducible terms less one, and number of possible equations, up to rank 20.

<table>
<thead>
<tr>
<th>Rank r</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of ( b )</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>9</td>
<td>13</td>
<td>21</td>
<td>31</td>
<td>44</td>
<td>66</td>
<td>94</td>
<td>133</td>
<td>191</td>
<td>268</td>
<td>372</td>
<td>520</td>
<td>717</td>
<td>982</td>
</tr>
<tr>
<td>No. of ( r + \frac{1}{2} ) terms</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>7</td>
<td>11</td>
<td>17</td>
<td>28</td>
<td>43</td>
<td>63</td>
<td>96</td>
<td>139</td>
<td>199</td>
<td>287</td>
<td>406</td>
<td>566</td>
<td>792</td>
<td>1092</td>
<td>1496</td>
</tr>
</tbody>
</table>

4. OPERATOR FORMULA FOR CONSERVED DENSITIES

In Paper II we proved that there exists an infinite sequence of p.c.d. for the KdV equation and obtained recursion formulas for generating them. Muc, as described in Sec. 3C, is straightforward for deriving individual p.c.d. but is not well suited for obtaining an explicit general formula. In this section our main objective is to derive an operator-product formula for \( T_r \).

In Paper VII we will show that the generalized equation

\[
    u_{q+1} + u_p u_1 + u_q = 0, \quad p \geq 1, \quad q \geq 2, \tag{26}
\]

for odd \( q \) has exactly three independent p.c.d. unless \( p = 1 \) or 2 and \( q = 3 \) (in which case it has infinitely many as proved in Paper II), and for even \( q \) has exactly one. The formalism to be developed in this section (for the KdV equation) depends essentially on Galilean invariance, requiring \( p = 1 \), but would go through for any \( q \). However, the results would be meaningless (except for \( q = 3 \)) because the supposed p.c.d. do not exist.

A. Exponential Series

A comparison of the p.c.d. in Paper II shows that consecutive \( T_r \) are related by

\[
    \partial_0 T_r = (r - 1) T_{r - 1}, \quad r \geq 2. \tag{27}
\]

As it stands, this relationship is of limited use even though it provides a complete prescription for going from \( T_r \) to \( T_{r - 1} \). We cannot go up in rank from \( T_{r - 1} \) to \( T_r \) because nonzero "constants" of integration (functions of \( u_1, u_2, \ldots \)) appear in \( T_r \).

Our remaining results depend heavily on

**Lemma 6**: If \( T \) and \( X \) are the c.d. and flux of a c.l. for the KdV equation, then so are

\[
    \tilde{T} \equiv \partial_0 T, \quad \tilde{X} \equiv \partial_0 X - T. \tag{28}
\]

**Proof**: Making the Galilean transformation \( x' = x + ct, t' = t \), and \( u_0' = u_0 + c \) (see Appendix A), \( \partial_0 T + \partial_0 X = 0 \) can be written

\[
    \partial_{x'} T(u_0 + c, u_1, \ldots) + \partial_{x'} X(u_0 + c, u_1, \ldots) = 0,
\]

where \( \partial_{x'} = \partial_t - c\partial_x \) if \( \partial' = \partial \). Expanding with respect to \( c \), the coefficient of the linear term is

\[
    \partial_0 (\partial_0 T) + \partial_0 (\partial_0 X - T) = 0. \tag{29}
\]

Alternatively, following Wiley,\(^8\) we can derive (29) by operating on \( \partial_1 T + \partial_1 X = 0 \) with \( \partial_0 \) and using the commutation formulas

\[
    \partial_0 \partial_1 = \partial_1 \partial_0 - \partial, \quad \partial_0 \partial_1 = \partial_1 \partial_0 - \partial,
\]

obtained from (12) and (16). (A posteriori, one might have been led to this by applying \( \Psi \) in order to annihilate \( \partial X \) and then have observed that \( \Psi \approx \partial_0 \).

We can eliminate the factor \( r - 1 \) in (27) by rescaling each \( T_r \) so that \( u_0' \) has coefficient \( 1/r! \) rather than \( 1/r \). Then instead of (27) we have

\[
    \partial_0 T_r = T_{r - 1}, \tag{30}
\]

valid for all \( r \geq 0 \) if we define \( T_0 = 1 \) and \( T_{-1} = 0 \).

We define the formal infinite sum

\[
    T = \sum_r T_r, \tag{31}
\]

a kind of generating function from which each \( T_r \) can be recovered by taking the \( r \)-rank slice. Then the formal sum of the infinite sequence of relations (30) can be written as the single relation

\[
    \partial_0 T = T, \tag{32}
\]

which formally integrates to

\[
    T = e^{u_0} A(u_1, u_2, \ldots). \tag{33}
\]

Series of this exponential form, namely \( e^{u_0} \) times a series in \( u_1, u_2, \ldots \), will be called e-series. For the most part in the following we can confine ourselves to the class of e-series because it is obviously closed under many of our operators, e.g., \( \partial, \partial_1, \text{ and } \partial' \).

However,

\[
    \partial_r e^{u_0} B(u_1, u_2, \ldots) = -e^{u_0}[(u_0 u_1 + u_0) B + (B + \partial_0) B] = -e^{u_0} u_0 (u_1 + B) B - e^{u_0} C(u_1, u_2, \ldots)
\]

is not purely exponential. The offending terms have been collected and written first; note that only the \( j = 0 \) and \( j = i + 1 \) terms from the inner summation in \( B \) have factors \( u_0 \). Nevertheless, since we only use the results of operating with \( \partial_i \) up to equivalence (i.e.,
modulo $x$ derivatives), we can write
\[
\partial_x(e^{ux}B) = -(\partial_x(e^{u_0(x_0 - 1)})B - e^{u_0}u_0B) - e^{u_0}C \\
\leq e^{u_0}(u_0 - 1)\partial_x B - e^{u_0}u_0\partial_x B - e^{u_0}C \\
= -e^{u_0}(\partial_x B + C),
\]
which is an $e$-series.

B. Canonical $e$-Series

Factors $u_1$ can always be successively eliminated from an $e$-series by integrating by parts with $e^{u_1}$, e.g.,
\[
e^{u_1}u_1^2 u_3 \geq -e^{u_1}(u_2 u_3 + u_1 u_4) \\
\geq e^{u_1}(-u_2 u_3 + u_4).
\]
This process terminates because the power of $u_1$ keeps decreasing. An $e$-series $E$ will be called canonical if it is independent of $u_1$,
\[\partial_x E = 0.\]
(We are abandoning irreducibility because it has no comparably algebraic expression.) The justification of the terminology is given by the above existence property and the following theorem.

\textbf{Theorem 7 (Uniqueness):} Any two equivalent canonical $e$-series are equal.

\textbf{Proof:} Their difference $E$ is equivalent to zero, so $E = \partial_x F$. Thus
\[0 = \partial_x E = \partial_x(\partial_x F - F),\]
since $\partial_x$ commutes with $\partial_x$. From Lemma 1
\[\partial_x F - F = c,\]
so that (absorbing $c$ into $F$) $F = e^{u_0}B(u_1, u_3, \cdots)$. But $E$ is canonical, so that
\[0 = \partial_x E = \partial_x F = \partial_x F + F \\
= e^{u_0}(u_1 \partial_1 B + \partial_1 B + B)
\]
using (12). Dropping the exponential and taking an arbitrary $n$-index slice gives
\[0 = u_1 \partial_1^{(n)} + \partial_1 B^{(n)} + B^{(n)}
\]
where $B^{(n)}$, $N^{(n)}$ is a polynomial. If $B^{(n)}$ did not vanish, its highest degree slice would satisfy the same equation without the middle term, which is impossible because the operator $u_1 \partial_1 + 1$ has the effect of multiplying each term by a positive constant. Hence, all $B^{(n)} = 0, F = 0$, and $E = 0$.

C. Operator Formula for Conserved Densities

From now on we use $T$ to denote the canonical equivalent of (31) rather than (31) itself. Next we derive an equality (not just equivalence) for $T$ which is the basis of the further analysis. To $T$ corresponds a flux $X$ such that
\[\beta + \partial_x T = \partial_x X.\]  
(34)

We make use of the $e$-series property ($\partial_x T = T$) by applying $\partial_x$ and subtracting (34) itself. In view of the commutation formulas (12) and (16) and Lemma 1, we obtain
\[T = \partial_x X - X,\]  
(35)

after absorbing into $X$ an otherwise free constant as in the proof of Theorem 7; (35) characterizes the $u_0$ dependence of $X$ as
\[X = u_0 T + Y\]  
(36)

where $Y$ is an $e$-series.

We make use of canonicity ($\partial_x T = 0$) by applying $\partial_x$ to (34). In view of the commutation relations (12) and (17) we obtain
\[(\mathcal{M} + \mathcal{N})T = \partial_x \partial_x X + \partial_x X.\]
Eliminating $\partial_x X$ by (35), this can be written
\[(\mathcal{K} - 1)T = \partial_x \partial_x X + X,\]  
(37)

where the effect of the operator
\[\mathcal{K} \equiv \mathcal{M} + \mathcal{N}\]
is to multiply each term by the number of its factors and differentiations combined.

Formally, solving (37) for the $X$ term alone, we can substitute the resulting formula into itself recursively. Since $\partial_x (\mathcal{K} - 1)T = 0$, this gives $X = (\mathcal{K} - 1)T + 0 + 0 + \cdots$, and indeed
\[X = (\mathcal{K} - 1)T\]  
(38)

is evidently a solution and is consistent with (35), and (36), by (12) and (14). To prove that it is the solution we need a uniqueness theorem. But we have already proved such a theorem for the $e$-series $F$ in the proof of Theorem 7, and it applies for $X$ of the form (36) since $Y$ is an $e$-series.

Putting (38) back into (34) gives
\[\beta + \partial_x T = \partial_x (\mathcal{K} - 1)T\]  
(39)
as an equation for $T$ alone.

This can be rank-sliced to give the recursion formula
\[\beta + \partial_x T_T = \partial_x (\mathcal{K} - 1)T_{r+1}.\]  
(40)

In using this to generate the $T_T$, it is most convenient to start iterating from $T_1 = u_0$.

The difficulty with solving this in general for $T_{r+1}$ in terms of $T_r$ is invert $\partial_x$. This operator has no inverse because only polynomials of very special form are derivatives, and unessentially because it annihilates
constants. (The difficulty does not occur in carrying out the recursion explicitly step by step, since at each stage the left side is guaranteed to be a derivative by the existence results of Paper 11.) However, \( \mathcal{D} \) has (many) left inverses (excluding constants), one such being

\[
\mathcal{D}^{-1} = \mathcal{M}^{-1} \sum_{i} \mu_{i} \sum_{j>1} (-\mathcal{D})^{i-j-1} \partial_{i}^{j-1}
\]

obtained from (22). (Other expressions for \( \mathcal{D}^{-1} \) and a detailed discussion of the algebra of these and other operators will be given in a forthcoming paper.) Applying \( \mathcal{D}^{-1} \) and then \( (\mathcal{K} - 1)^{-1} \), we iterate to obtain

\[
T_{r+1} = (\mathcal{K} - 1)^{-1}(B + \mathcal{D})u_{0}. \tag{41}
\]

This is an explicit representation of the p.c.d., as desired. Note that \((\mathcal{K} - 1)^{-1}\) annihilates only terms of the form \(cu_{0}\), which are of rank 1 and hence never occur.

Alternatively, in order to avoid \( \mathcal{D}^{-1} \) with its arbitrariness and (moderate) complexity, we apply \( \mathcal{D} \) to (39) and commute operators to obtain for \( \mathcal{D}^{T}T_{r+1} \) the recursion formula

\[
(\mathcal{K} - r - 1)^{-1}(B + \mathcal{D}) = \mathcal{D}^{r}T_{r+1}, \tag{42}
\]

There is no difficulty in inverting \((\mathcal{K} - r - 1)^{-1}\), since every term of \( T_{r+1} \) has degree \( m \geq 2 \) and \( \mathcal{D} \) raises derivative index \( r \) by \( r \), so that \( m + n - r - 1 > 0 \); doing so and iterating gives for \( \mathcal{D}^{r}T_{r+1} \) the explicit formula

\[
(\mathcal{K} - r - 1)^{-1}(\partial_{i})^{(\mathcal{K} - r - 1)} \times (-\partial_{i})^{2r-1}(\partial_{i}^{r})u_{0}. \tag{43}
\]

Fortunately, we can bypass the need to invert \( \mathcal{D} \) by applying \( \partial_{i} \), because \( \partial_{i}^{\mathcal{D}} = \partial_{i} \) by (12) and so \( \partial_{i}^{\mathcal{D}} = \partial_{i} \partial_{i}^{\mathcal{D}} \). Since \( \partial_{i}T_{r+1} = T_{r+1} \) by (30), we thus have an equivalent representation of the p.c.d.

\[
T_{r} = (\mathcal{K} - r - 1)^{-1}(\partial_{i})^{\mathcal{K} - r - 1} \times \partial_{i}(\mathcal{K} - r)^{-1}\partial_{i} \times \partial_{i}(\mathcal{K} - r - 1)^{-1}\partial_{i}u_{0}. \tag{44}
\]

Instead of \( \partial_{i} \), we could have applied \( \Psi_{i} \), by (24) obtaining \( \Psi_{i}T_{r+1} \) on the left. This in itself is no improvement, being equivalent to \( \partial_{i} \), but it can be used to remedy a slight disadvantage of formula (43): though \( \partial_{i} \) (and the associated inverse “numerical” operator involving \( \mathcal{K} \)) is applied \( r \) times on the right as in (41), we obtain a p.c.d. of rank merely \( r \), rather than \( r + 1 \). To restore the rank lowered by \( \Psi_{i} \), we multiply by \( u_{0} \) and observe that \( u_{0}\Psi_{i} \) is \( \mathcal{M}^{-1} \), as seen in (21). Since \( \mathcal{M}^{-\mathcal{D}} = \mathcal{M}^{-1} \), applying \( \mathcal{M}^{-1} \) yields

\[
T_{r+1} = (\mathcal{K} - r - 1)^{-1}(\mathcal{M}^{-1}u_{0})(\mathcal{K} - r + 1)^{-1} \times \partial_{i}(\mathcal{K} - r)^{-1}\partial_{i} \times \partial_{i}(\mathcal{K} - r - 1)^{-1}\partial_{i}u_{0}. \tag{45}
\]

ACKNOWLEDGMENTS

The work presented here was supported by the Air Force Office of Scientific Research under Contract AF 49(638)-1555. One author (R. M. M.) also acknowledges support from the U.S. Atomic Energy Commission, Contract AT(30-1)-1480.

APPENDIX A: TRANSFORMATION PROPERTIES OF THE KdV EQUATION

It is of interest to find transformations of the dependent and independent variables that leave the KdV equation invariant. It is obviously invariant to arbitrary translations of the independent variables (since they do not occur explicitly)

\[
x \rightarrow x + c_{1}, \quad t \rightarrow t + c_{2}.
\]

Similarly, it is invariant to the scale transformation

\[
x \rightarrow cx, \quad t \rightarrow c^{2}t, \quad u \rightarrow c^{-2}u.
\]

In particular, for \( c = -1 \) the signs of \( x \) and \( t \) are reversed and we note a kind of reversibility of the solution. Further the KdV equation is Galilean-invariant, i.e., invariant under the transformation

\[
x \rightarrow x + ct, \quad t \rightarrow t, \quad u \rightarrow u + c.
\]

This property is essential for our results in Sec. 4.

Aside from the invariance transformations, there are (real-valued) transformations that change the signs of individual terms in the KdV equation. The sign of the \( t \)-derivative term (relative to the other two) is changed by either \( t \rightarrow -t \) or \( x \rightarrow -x \), that of the nonlinear term by \( u \rightarrow -u \), and that of the dispersive term by the combination \( u \rightarrow -u, t \rightarrow -t \) (or \( u \rightarrow -u, x \rightarrow -x \)).

Incidentally, the KdV equation possesses a similarity solution, i.e., a solution \( U \) invariant under the above scale transformation; \( U \) satisfies the ordinary differential equation

\[
\frac{d^{2}U}{d\eta^{2}} + (U - \eta)\frac{dU}{d\eta} - 2U = 0,
\]

where

\[
u(x, t) \equiv (3t)^{-\frac{3}{2}}U(\eta), \quad \eta \equiv x(3t)^{-\frac{1}{2}}.
\]

APPENDIX B: CONSERVED DENSITY OF RANK \( r = 11 \)

\[
T_{11} = \frac{1}{11} u_{1}^{11} - 45u_{0}u_{2}^{5} + 216u_{0}u_{2}^{2} - 1260u_{0}u_{4} + 648u_{0}u_{4}^{3} + 4320u_{0}u_{4}^{2} + 22680u_{0}u_{4}u_{2}^{2} - 7560u_{0}u_{4}^{2} + 1296u_{0}u_{2}^{4} - 32400u_{0}u_{2}u_{4}^{2} - 38880u_{0}u_{2}u_{4}^{2} + 45360u_{0}u_{2}u_{4}^{2} + 238464u_{0}u_{2}u_{4}^{2}.
\]
Nonphysical Region Singularities of the $S$ Matrix

PETER GODDARD

Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, England

(Received 9 May 1969)

The work of a previous paper, which examined the relationship between the assumption of Cutkosky-type discontinuity formulas to the positive-$\alpha$ condition for the singularity of anomalous thresholds, is extended using more powerful methods. The method is first described in general terms. It is then justified using homological techniques, and specific expressions are calculated for distortion terms. Finally, some specific cases are considered and more elaboration versions of the cancellation processes previously found are shown to occur.

1. INTRODUCTION

In a previous paper\(^1\) (hereafter referred to as I), we considered the way that the positive-$\alpha$ condition for anomalous thresholds and the assumption of Cutkosky-type discontinuity formulas with "vanishing cycle" contours both lead to criteria for the singularity of connected $S$-matrix elements. This permitted the testing of the consistency of these conjectures on the basis of the usual $S$-matrix-theory assumptions. We remarked on the technical difficulty of performing finite continuations of unitarity integrals outside the physical region. The method we used to perform these continuations can be described as a metrical one in that we took a specific form for the unitarity integral as a repeated integral in the space of (complex) internal momenta, the order of the integrations being chosen with reference to the possible distortions. The different orders of integration result in different but "homologous" contours of integration, whose equivalence follows from Cauchy's theorem. The lack of symmetry in the way the integration contour had to be chosen and the increasing complexity of the integrals, taken with the delicate cancellation processes found in any but the simplest examples, make it clear that

\[ + 136080 u_1 u_2 u_3^2 - \frac{19440}{11} u_4 u_5 + \frac{544320}{11} u_2 u_4^2 \]

\[ + 38880 u_1^2 u_2 u_3^2 - \frac{19440}{11} u_1 u_6 + \frac{544320}{11} u_2 u_4^2 \]

\[ - \frac{499920}{11} u_2 u_3 u_4^2 - \frac{668736 u_1 u_2 u_3^2}{11} \]

\[ - 58320 u_1 u_2^2 + \frac{1288880}{11} u_1 u_6 \]

\[ + 233280 u_2 u_4^2 \]

\[ - \frac{8855160}{11} u_2 u_4 u_5^2 + \frac{233280}{11} u_2 u_4^2 \]

\[ + \frac{58348860}{11} u_1 u_2 u_3^2 + \frac{8855160}{11} u_1 u_2 u_3^2 \]

\[ + \frac{28094116}{11} u_1 u_2 u_3^2 \]

\[ - \frac{1288880}{11} u_1 u_2 u_3^2 + \frac{233280}{11} u_2 u_4^2 \]

\[ - \frac{1288880}{11} u_1 u_2 u_3^2 \]

\[ + \frac{8855160}{11} u_2 u_4 u_5^2 + \frac{233280}{11} u_2 u_4^2 \]

\[ - \frac{1288880}{11} u_1 u_2 u_3^2 \]

\[ + \frac{233280}{11} u_2 u_4^2 \]

\[ - \frac{1288880}{11} u_1 u_2 u_3^2 \]

\[ + \frac{233280}{11} u_2 u_4^2 \]

\[ - \frac{1288880}{11} u_1 u_2 u_3^2 \]

\[ + \frac{233280}{11} u_2 u_4^2 \]

\[ - \frac{1288880}{11} u_1 u_2 u_3^2 \]

\[ + \frac{233280}{11} u_2 u_4^2 \]