ALGEBRAIC ASPECTS OF NONLINEAR DIFFERENTIAL EQUATIONS

Yu. I. Manin UDC [512.88+517.946]:530.145.6

INTRODUCTION

0.1. This article is an attempt to systematize and display the basic structures in a vast, beautiful, and important mathematical literature which has accumulated during the last decade as a result of the boom surrounding the Korteweg–de Vries equation $u_t = 6uu_x - u_{xxx}$.

This equation was proposed and investigated in 1895 for describing long surface waves in a channel with rectangular walls. Korteweg and de Vries also obtained its first solutions: "cnoidal wave" — an infinite periodic wave train moving with constant speed without changing form — and its limiting case — the solitary wave or soliton (the terminology of Kruskal and Zabusky).

As early as August 1834 Scott Russell observed a soliton on water. He subsequently described his observations in the "Report on Waves" (see the reference in [47]) which is these days widely cited everywhere from the SIAM Review to the Young Technician. The lively image of him galloping along the banks of a channel in pursuit of a solitary wave evokes in the modern reader a mild and pleasant nostalgia harmonious with the general style of a bygone era.

It is not often that a single problem brings to life an entire theory. It seems that we are presenting its inception. The number of publications on this topic is now counted in the hundreds; what is more important, in the theory of nonlinear differential equations a clear shift of interests and a reevaluation of priorities are taking place.

For the fundamental equations of classical fields, special hopes rest on the existence of soliton solutions and their peculiar nonlinear superpositions. The elementary particles are possibly related to such solutions.

There are still no general notions of what soliton and multisoliton solutions really are, which classes of equations have them, and how they are related to the presence of an infinite sequence of conversion laws and the so-called Bäcklund transformations. There is only a rich experimental material which remains to be sorted out. The word "experimental" is here meant in the broad sense to include direct observations, calculations on the computer, and investigation of particular interesting classes of equations. It is clear that the deciphering of the structure of these phenomena must play a fundamental role in understanding them. A good and to considerable extent algebraic and geometric theory is needed, and this remains to be created.

Translated from Itogi Nauki i Tekhniki, Sovremennye Problemy Matematiki, Vol. 11, pp. 5-152, 1978.

0096-4104/79/1101-0001$07.50

We shall briefly describe certain properties of the Korteweg–de Vries equation which may serve to clarify the plan of the present article.

0.2. Derivation of the Equation. The usual linear one-dimensional wave equation is written in the form $u_{tt}-c^2u_{xx}=0$. Its general solution is a sum of two waves of arbitrary form $u=f(x+ct)+g(x-ct)$ one of which moves to the left and the other to the right with constant speed c. We consider the equation $u_t+cu_x=0$, which distinguishes waves moving to the right. Among its solutions are the harmonic waves $u=\exp i(\omega t-kx)$ where the frequency ω and the wave number k are related by $\omega=ck$, or for waves of both types by $\omega^2=c^2k^2$, where c is characteristic of the medium.

If the wave equation remains linear but includes derivatives of higher order, then the relation between the frequency and the wave number of a harmonic wave may have the more general form $\omega^2=f(k^2)$, where f is not necessarily a linear function. In the approximation of long waves, i.e., small k, we may restrict attention to the first two terms of the Taylor series for f and write $\omega^2\cong c^2k^2+\varepsilon k^4$ or $\omega\cong ck+\frac{1}{2}\frac{\varepsilon k^3}{c}$. Waves with this dispersion relation (dependence of the frequency on the wave number) are described by the equation $u_t+cu_x-\frac{\varepsilon}{2c}u_{xxx}=0$.

On the other hand, the simplest nonlinearity enters if it is assumed that the speed depends on the amplitude u. For waves of small amplitude it may be assumed that the dependence is linear, and the equation can be written in the form $u_t+(c+\alpha u)u_x=0$. The dependence of the speed on the amplitude for suitable sign of α may cause the crest of the wave to move faster than the trough, i.e., curling of the front occurs with subsequent formation of breakers and decay of the wave.

Simultaneous consideration of dispersion and nonlinearity, leads to the equation $u_t+cu_x-\alpha uu_x-\frac{\varepsilon}{2c}u_{xxx}=0$. If we go over to a system of coordinates moving to the right with speed c, the term cu_x drops out, and we arrive at the Korteweg–de Vries equation up to a normalization constant which can be changed by scaling u, x, t. (Making use of this fact, in the main text we shall often write this equation with different coefficients.)

This derivation is good in that it nowhere appeals to hydrodynamics and indicates the universal applicability of the Korteweg–de Vries equation to one-dimensional media where the essential features are only weak dispersion and weak nonlinearity.

0.3. Cnoidal Waves and the Soliton. We shall seek a solution of the Korteweg–de Vries equation $u_t=6uu'-u'''$ in the form of a traveling wave $u(x,t)=U(x-vt)$, where U is the wave form and v is a constant speed (we recall v is really the speed by which the traveling wave exceeds the wave speed in the simplest approximation $u_t+cu_x=0$).

For U we obtain the equation $-vU'=6UU'-U'''$. Integrating we obtain $-vU=3U^2-U''+a$, where a is a constant. Multiplying by U' and integrating again, we find $-v\frac{U^2}{2}=U^3-\frac{1}{2}U'^2+aU+b$, where b is a new constant, or $U'^2=2U^3+vU^2+aU+b$. Up to a normalization constant, the general solution of this equation is the Weierstrass function $U(x-vt)=c_1\wp(x-vt)+c_2$, the periods of which are the periods of the elliptic curve Γ: $Y^2=2X^3+vX^2+$

$aX+b$; here c_1, and c_2 are suitable constants. This is a cnoidal wave train if the discriminant of the curve is different from zero; the period of the wave train is the real period Γ, i.e., $\int_\gamma \omega$, where $\omega = dX(2X^3+vX^2+aX+b)^{-\frac{1}{2}}$, and γ is a real cycle on the Riemann surface of Γ.

The soliton is obtained for the curve with a double point at the origin: $Y^2=2X^3+vX^2$, $a=b=0$. The explicit formula for it has the form $U(x-vt) = -\frac{v}{2}\,\mathrm{ch}^{-2}\left(\frac{\sqrt{v}}{2}(x-vt)\right)$. This is the solitary wave (in the present normalization it is rather a "solitary well") with trough at the point $x=vt$. The depth of the well is proportional to its speed which may be arbitrary. The soliton is the limit of the cnoidal wave train when its period tends to infinity.

0.4. Superposition of Solitons and Quasiperiodic Solutions. Since solitons decrease at infinity and large solitons move faster than small ones, we may attempt to consider the solution of the Cauchy problem for which $u(x, 0)$ is the sum of two widely separated solitons of which the left is larger than the right and therefore begins to move almost independently of the right soliton and strives to overtake it. After this occurs a period of essentially nonlinear interaction ensues, and it is of interest to consider what form the solution may have at a later time. Contrary to usual expectations, numerical experiment showed that after a rather long time the solution is nearly the sum of the same two solitons of which the greater has already overtaken the smaller, and the result of the collision is found only in a shift of phase but does not affect their form or speed.

This provoked attempts to analytically prove the existence of the superposition of solitons. Lax in [42], which had a great effect on the subsequent development of the theory, established, in particular, the existence of the two-soliton solution, and almost simultaneously explicit formulas were found for the superposition of any number of N solitons.

These formulas have the form $u(x,t) = -2\frac{\partial^2}{\partial x^2}\log\det(E+A)$, where E is the $N\times N$ identity matrix and the element A of the matrix ij is $c_i c_j(\alpha_i+\alpha_j)^{-1}\exp\left[(\alpha_i^3+\alpha_j^3)t-(\alpha_i+\alpha_j)x\right]$, $\alpha_i>0$, $i=1,\ldots,N$, $\alpha_i\neq\alpha_j$ for $i\neq j$. Asymptotically for $|t|\to\infty$ this solution decays into a sum of N solitons arranged in order of decreasing (as $t\to-\infty$) or increasing (as $t\to\infty$) amplitudes and speeds.

After some time solutions were discovered which are related to the N-soliton solutions in the same way as cnoidal waves are related to the single soliton solution (for the history of this discovery see the survey of Dubrovin, Matveev, and Novikov [8]). They were found to be related to the Riemann theta function for hyperelliptic curves of genus N (with equation $Y^2=F(X)$, where F is a polynomial of degree $2N+1$). The explicit formulas have the form $u(x,t) = -2\frac{\partial^2}{\partial x^2}\log\theta(x\alpha+t\beta+\gamma)+\mathrm{const}$, where α, β, γ are certain N-dimensional complex vectors. (For the details see the review of Matveev [44] and Sec. 7, Chap. 4 of the present work.) The N-soliton solutions are obtained in the limit as the hyperelliptic curve degenerates to a rational curve with N double points. Partial degeneration (with reduction of the genus) leads to a "multisoliton solution on the background of a quasiperiodic solution."

0.5. The Conservation Laws. A conservation law for the evolution equation $u_t = K(u, u', \ldots, u^{(N)})$ $\left(u^{(i)} = \frac{\partial^i u}{\partial x^i}\right)$ is a relation of the form $T_t + X_x = 0$, where T and X are functions of $u^{(j)}, j \geqslant 0$, which follows formally from the equation. If u is a solution of the equation which is rapidly decreasing at infinity, then $\frac{\partial}{\partial t} \int_{-\infty}^{\infty} T dx = -\int_{-\infty}^{\infty} X_x dx = 0$, so that T is the density of a quantity conserved in time. The first three conservation laws for the Korteweg–de Vries equation are obtained without difficulty: they have the form

$$u_t + (-3u^2 + u_{xx})_x = 0;$$
$$(u^2)_t + (-4u^3 + 2uu_x - u_x^2)_x = 0;$$
$$\left(u^3 + \frac{1}{2} u_x^2\right)_t + \left(-\frac{9}{2} u^4 + 3u^2 u_{xx} - 6uu_x^2 + u_x u_{xx} - \frac{1}{2} u_{xx}^2\right)_x = 0.$$

They can be interpreted as the laws of conservation of mass, momentum, and energy. However, the Korteweg–de Vries equation has an infinite sequence of conservation laws which are polynomials in the $u^{(j)}$. They were first written out by the method of undetermined coefficients; this work was practicable up through the ninth law. According to Miura [47], "in the summer of 1966 the rumor circulated that only nine conservation laws exist." Miura killed this rumor by spending a week of summer vacation computing the tenth law; after this a machine program was written which computed the eleventh law consisting of 45 terms. (With the previous program all storage capacity was used already at the sixth conservation law.) Very transparent proofs of the theorem on the existence of an infinite sequence of laws and results on their structure were then obtained in connection with important theoretical progress: the discovery of the Lax representation and the applicability of the technique of the inverse scattering problem.

0.6. The Lax Representation and the Inverse Problem. Lax [42] observed that the Korteweg–de Vries equation can be written in the form $L_t = [P, L]$, where $L = -\frac{\partial^2}{\partial x^2} + u(x, t)$ and $P = -4\frac{\partial^3}{\partial x^3} + 3\left(u(x, t)\frac{\partial}{\partial x} + \frac{\partial}{\partial x} u(x, t)\right)$. Here $[P, L]$ is the **commutator** in the ring of linear differential operators; L_t is the coefficientwise derivative of L with respect to the "parameter" t. Equations of this form for flows in Lie algebras have been known for a long time; the best known of these is the Schrödinger equation in the Heisenberg representation where P is the energy operator of the system and L is any observable. Classical Hamiltonian equations in the Lie algebra of functions on phase space with the Poisson bracket and also the equations for the rotation of a solid body can be written similarly.

There is a simple formalism of conservation laws connected with such equations: if there is a linear representation φ of the Lie algebra and a "generalized trace" function Tr on it which is zero on the commutators, then $\mathrm{Tr}\varphi(L^n)$ for $n \geqslant 0$ is conserved in time, since $(\mathrm{Tr}\varphi(L^n))_t = \mathrm{Tr}\varphi([P, L^n]) = 0$.

It is, however, far from obvious how to carry through this formalism for a Lie algebra of differential operators. This development led to the formulation of the extremely important method of inverse scattering theory. In general outline it reduces to tracing the evolution in t of the space of solutions of the linear problem $L\psi=\lambda\psi$ (λ a constant) in terms of the "scattering data." The scattering data consist of the discrete part of the spectrum of L, the normalization constant of the eigenfunctions, and also the scattering matrix for the (rapidly decreasing) potential u from the continuous part of the spectrum. For further details see the survey of Faddeev [23], the extensive literature, and also Sec. 4 of Chap. 4.

0.7. The Variational Formalism and the Hamiltonian Property. The Korteweg–de Vries equation can also be written in the form $u_t=\frac{\partial}{\partial x}\frac{\delta}{\delta u}\left(u^3+\frac{u_x^2}{2}\right)$, where $\frac{\delta}{\delta u}=\sum_{i\geq 0}(-1)^i\partial^i\frac{\partial}{\partial u^{(i)}}$ is the Euler–Lagrange operator or the variational derivative. This is a somewhat unusual Hamiltonian form: the standard form of the equation of a Hamiltonian evolution for a vector-valued function $\bar{u}$ of dimension $2n$ is $\bar{u}_t=\begin{pmatrix}0 & E\\ -E & 0\end{pmatrix}\frac{\delta H}{\delta \bar{u}}$, where E is the $(n\times n)$ identity matrix, $H=H(u_i^{(j)})$ is the Hamiltonian, and $\frac{\delta H}{\delta \bar{u}}=\left(\frac{\delta H}{\delta u_1},\ldots,\frac{\delta H}{\delta u_n}\right)^t$.

Various aspects of the Hamiltonian property were investigated in the important work of Gardner [36], Lax [43], and Zakharov and Faddeev [12]; it was shown, in particular, that the conservation laws commute in the sense of the Hamiltonian formalism. Equations of the form $u_t=\frac{\partial}{\partial x}\frac{\delta T}{\delta u}$, where $T_t+X_x=0$ is some conservation law for $u_t=6uu_x-u_{xxx}$, have received the name of higher Korteweg–de Vries equations. The solutions described in part 0.4 were characterized invariantly as flows induced by the Korteweg–de Vries equation on the stationary manifolds of conserved quantities defined by the ordinary differential equation $\frac{\delta T}{\delta u}=0$.

There are other equations. Practically all the effects described above for the Korteweg–de Vries equation were subsequently found for a large number of physically interesting equations including the sine-Gordon equation, the nonlinear Schrödinger equation, the equations for self-induced transparency, etc. For other "suspect" equations part of the properties have been verified sometimes by numerical experiment.

0.8. The Plan of the Article. In the vast majority of studies pertaining to the Korteweg–de Vries equation and its analogues, a substantial role is played by a system of purely algebraic structures connected with these equations which do not depend on assumptions of analytic character, the choice of function spaces, existence and uniqueness theorems, etc. The principal aim of this article, as previously mentioned, consists in displaying and systematically expounding the origins of the theory of these structures. This objective has determined the choice of the material as well as the order in which it is introduced.

The first chapter is devoted to the foundations of the variational calculus with higher order derivatives which is necessary for the natural introduction of the conservation laws and the Hamiltonian structure. Here an attempt is made to follow the invariant interpretation of the variational calculus in terms of differential forms and vector fields on spaces of jets without which the formulas, which become more complex with increasing order of the derivatives, are hard to interpret and work with in a practical manner. Special attention is focussed on the basic facts of the Hamiltonian and Lagrangian formalisms.

In the second chapter a detailed study is made of the structure of general Lax equations as well as of an enigmatic system of wave equations of Benney which displays many of the features of the systems described above but so far does not fit into the general theory. We consider it an interesting object for future investigations.

The third chapter is devoted to the Lax equations of multisoliton and quasiperiodic type. Here we have also strived to display in the clearest possible way the mechanism of the appearance of the algebrogeometric structures in the theory of the equations without writing out explicit formulas for their solutions which is done in a number of other surveys and is briefly considered in Sec. 6 of Chap. 4. Exceptions are the "solitons of higher rank" which are here obtained by algebrogeometric methods for the first time. One common feature of all problems solved should be emphasized: the introduction of an auxiliary fiber bundle over an algebraic manifold and the interpretation of the equation as the problem of finding a connection in this bundle with certain additional properties. Recently the problem of "instanton" solutions of the Yang–Mills equations (more precisely, the duality equations) in Euclidean field theory with the group SU(2) has been reformulated and advanced in this manner (Atiyah following preliminary work of Penrose, t'Hooft, A. S. Shvarts, Polyakov, and others). It reduces to the classification of two-dimensional complex vector bundles over $P^3(\mathbf{C})$, which are trivial on a certain class of lines in P^3 .

Finally, facts regarding particular interesting constructions and methods which have not yet been sufficiently thought out or subjected to systematization are collected in the fourth chapter. The exposition here follows the sources cited in the corresponding sections; proofs for the most part are omitted. The only exception which deserves mention is the beginning of Sec. 3 of Chap. 4 where an attempt is made (not entirely successful) to invariantly define a very interesting Lie algebra introduced by Estabrook and Wahlquist in connection with their theory of "prolongation structures" and generalized conservation laws.

The reader should not take the scattered references and credit of authorship for various results too seriously. Many similar works were done almost simultaneously and almost independently; many approaches revealed a parallelism unknown to their authors; many ideas hung in the air and continued to hang in the air some time after formal first publications. The history of our question, if it deserves such, remains to be written.

In spite of the length of the article, many interesting facts have remained beyond its scope. First of all, the analytic theory of the method of inverse scattering has been omitted completely in spite of its importance and the fact that it motivated the inception of many of the purely algebraic constructs described here. Regarding this question, the reader may find abundant information in the literature cited. Secondly, very little attention is devoted to specific solutions of particular equations or to their physical interpretation. Third, we have left untouched the interesting parallel theory of discrete systems such as the "Toda lattice" and such of its principal applications as the explanation of the Fermi–Pasta–Ulam paradox. Fourth, interesting investigations of flows of Lax type in finite-dimensional Lie algebras and the many-particle problems related to such algebras have been omitted. The informed reader will probably discover still more omissions voluntary or involuntary. Among the results not contained in this work but naturally related to it mention should be made of the investigation of Bogoyavlenskii and Novikov [1] and of Gel'fand and Dikii [3] on restricting Hamiltonian flows to stationary manifolds of conservation laws. They are of basic importance for understanding the relation between solitons and conservation laws and merit generalization to the multidimensional case.

We have not endeavored to compile a complete bibliography. In place of this the bibliography includes surveys with large bibliographies and collections devoted to specific aspects of the theory [2, 8, 23, 26, 33, 44, 47, 50, 53]. Beyond this, papers having a direct relation to the question touched on here have been selected as have several works to which we do not refer but which, in our opinion, deserve special attention.

0.9. It is impossible to overestimate the role that the author's many conversations with I. M. Gel'fand and also the work of Gel'fand and Dikii [2-5] played in the design and plan of this paper.

The selection of material for the paper was made during a special course which the author gave in the mechanics and mathematics department of Moscow State University in 1975/1976 and from the introduction to a seminar in 1976/1977. The participants in the course and seminar provided the author with a great deal of material which directly or indirectly affected the content of the paper.

In particular, a large part of the new results of Chap. 1 belong to B. A. Kupershmidt; their presentation is based on his published papers and notes which were kindly given to the author before their publication. The investigation of Benney's equations in Chap. 2 was carried out jointly by B. A. Kupershmidt and the author. The algebraic reworking of the Gel'fand–Dikii theory in Chap. 2 derives from a report of M. S. Shubin in the seminar in which the simplicity of the formalism of pseudodifferential operators over a one-dimensional base was revealed. The role of bimodules and connections to which Chap. 3 is devoted was clarified by V. G. Drinfel'd. From the report of S. I. Gel'fand the author first understood the technique of Estabrook and Wahlquist, while the exposition of the results of Lax in Sec. 5, Chap. 4 is based on the notes of I. Ya. Dorfman. Finally, conversations with B. A. Kupershmidt, M. A. Shubin, V. G. Drinfel'd, V. E. Zakharov, I. Ya. Dorfman, and D. R. Lebedev were very useful to the author. I am happy to express my deepest gratitude to them all.

CHAPTER I

THE VARIATIONAL FORMALISM

1. Differential Equations: Three Languages

1.1. The Classical Language. In this language we first of all choose a notation for the independent variables, say, $x_1, \ldots, x_m$, and for the unknown functions, say $u_1, \ldots, u_n$. Let $k=(k_1, \ldots, k_m)$ where $k_i \geqslant 0$ are integers, and let $|k|=\sum_{i=1}^{m} k_i$. We denote by the symbol $u_i^{(k)}$ the derivative $\dfrac{\partial^{|k|} u_i}{\partial x_1^{k_1} \ldots \partial x_m^{k_m}}$. A system of differential equations relative to $\{u_i\}$ is a collection of relations of the form

$$F_j(x_1, \ldots, x_m; u_1, \ldots, u_n; u_1^{(k)}, \ldots, u_n^{(k)}) = 0, \tag{1}$$

where the F_j are some functions.

It is sometimes convenient to distinguish one of the variables, say, t (the "time" as opposed to the spatial coordinates $x_1, \ldots, x_m$), and to consider a system of evolution equations of the form

$$\frac{\partial u_j}{\partial t} = u_{j,t} = F_j(x_1, \ldots, x_m, t; u_1, \ldots, u_n; u_1^{(k)}, \ldots, u_n^{(k)}). \tag{2}$$

We point out that F_j does not depend on the derivatives of u_i with respect to t; $j = 1, \ldots, n$.

1.2. The Language of Differential Algebra. Let A be a ring, and let M be a left A-module. We recall that a differentiation of A into M is any additive mapping $\partial: A \to M$ with the property $\partial(ab) = a\partial b + b\partial a$ for all $a, b \in A$.

The algebraic analogue of the system of equation (1) is a structure consisting of a ring A, some Lie algebra D with a differentiation into itself, and an ideal $I \subset A$, for which $DI \subset I$. More precisely, we suppose that F_j in (1) is infinitely differentiable in all its arguments and set $A = \bigcup_{l=0}^{\infty} C^\infty(x_1, \ldots, x_m; u_j^{(k)} \mid 1 \leqslant j \leqslant n, |k| \leqslant l)$, where the $u_j^{(k)}$ are formal independent variables. Let further $D = \sum_{j=1}^{m} A\partial_j$, where ∂_j takes x_i into δ_{ij}, and $u_i^{(k)}$ into $u_i^{(k+\varepsilon_j)}$, $\varepsilon_j = (0\ldots1\ldots0)$ (1 sits at the j-th place), and let I be the ideal in A, generated by all $\partial_1^{k_1} \ldots \partial_n^{k_n} F_j$. The structure (A, D, I) corresponds to (1). Any smooth solution of (1) corresponds to a homomorphism $A \to C^\infty(x_1, \ldots, x_m) = K$, which is the identity on K, contains I in its kernel, and takes ∂_j into $\frac{\partial}{\partial x_j}$.

If the right sides of (2) do not depend on t explicitly, then the algebraic analogue of (2) in this context can be constructed by not introducing t explicitly: it consists of (A, D) and the additional differentiation of "evolution" $X: A \to A$, defined by the conditions $[X, \partial_j] = 0$ for all j, and $Xu_j = F_j$ (the right sides of (2) for $j = 1, \ldots, n$). Any smooth solution of (2) corresponds to a K-homomorphism $A \to C^\infty(x_1, \ldots, x_m, t)$, which takes ∂_j into $\partial/\partial x_j$, and X into $\partial/\partial t$.

Variations are possible in the definition of the ring A. For example, if the F_j in (1) do not depend explicitly on x_i and are polynomials in $u_i^{(k)}$, it is possible to set $A = \mathbf{R}[u_i^{(k)}]$ and consider the algebraic object modeling (1) to be a minimal ideal in A, which is D-closed and contains F_j. It is also possible to take A to consist of analytic functions, meromorphic functions, germs of functions, etc., depending on the properties of F_j and the solutions of interest to us.

1.3. The Geometric Language. Here we start with some locally trivial, smooth (i.e., class C^∞) fibration $\pi: N \to M$. The role of the independent variables $x_1, \ldots, x_m$ in (1) is played by a point on M, and the role of the unknown functions $u_1, \ldots, u_n$ is assumed by a smooth section $s: M \to N$ of the fibration π. In order to define a geometric object corresponding to the system (1) it is necessary to introduce the tower of jet spaces of the fibration π . We recall the corresponding definitions.

Let s_1 and s_2 be local sections of the fibration π . They are tangent to one another at a point $\xi \in N$ to order $k \geqslant 0$, if they pass through this point and their Taylor series in any local coordinate system for $\pi(\xi)$ coincide through order k . A k-jet at the point ξ is the equivalence class of local sections which are tangent to one another at the point ξ to order k . Let $J^k\pi$ be the set of all k-jets. It is equipped with a natural smooth manifold structure and there is a tower of smooth fibrations $\ldots J^k\pi \to J^{k-1}\pi \to \ldots \to J^0\pi = N \to M$. An analogue of system (1) is then a closed subset in a suitable story of the tower $J^k\pi$, prescribed by the vanishing of the right sides of (1).

More precisely, the connection of the classical notation (1) with geometry is realized by means of a choice of consistent local coordinate systems on all the spaces $J^k\pi$. We begin with a choice of a pair of neighborhoods $\xi \in V \subset N$ and $\pi(\xi) \in U = \pi(V) \subset M$, such that the restriction $\pi : V \to U$ is diffeomorphically equivalent to the projection $R^{m+n} \to R^m$. The choice of this diffeomorphism is a local chart for π; prescribing it is equivalent to prescribing a local coordinate system in N of the form $(u_1, \ldots, u_n;\ x_1, \ldots, x_m)$ such that the restrictions of (u_i) to each fiber of π give a local chart in this fiber, and the (x_j) are lifts from the base. With respect to this local chart a local chart is canonically constructed in each $J^k\pi$ with a coordinate system which is denoted by $(u_j^{(l)},\ x_i \,|\, |l| \leqslant k)$ and is uniquely defined in the following manner. Let s be some local section of $\pi|_V$, which is represented in a local chart of π by the functions $u_1^s, \ldots, u_n^s \in C^\infty(x_1, \ldots, x_n)$. Its k-jet is a point of $J^k\pi$. The value of the coordinate $u_j^{(l)}$ at this point is by definition $\dfrac{\partial^{|l|} u_j^s}{\partial x_1^{l_1} \ldots \partial x_m^{l_m}}$. Obviously the collection of k-jets of the section s at all its points is a smooth section of $J^k\pi$; it is called the lift of s to $J^k\pi$, and we denote it by the same letter s . The lifts of sections are compatible with the projections of the jet spaces.

It is now clear that the system (1) defines some subset Φ in $J^k\pi$, and its solutions are those sections of π, whose lifts lie in Φ.

The connection of the geometric language with the language of differential algebra is as follows: $A = \bigcup_{k=0}^{\infty} C^\infty(J^k\pi)$ (imbeddings with respect to the natural projections), $D = D_c$ is the Lie algebra of "horizontal" differentiations of A (for the precise definition see the following section), $I \subset A$ is the ideal of functions which vanish on the "lifts Φ" (the geometric definition of a lift is most simply introduced by the condition of differential closedness of the corresponding ideal).

1.4. A Comparison of the Languages. The three languages briefly described are not equivalent either mathematically or esthetically; each has its advantages and shortcomings.

The classical language makes the fewest explicit assumptions on the form of the functions F_j, the sense of the derivatives $u_j^{(k)}$, the domains of existence of solutions, etc. It affords the freedom of interpreting, e.g., the u_j as various types of generalized solutions (and therefore, possibly, not elements of any natural ring). It also permits when necessary accentuating other algebraic structures which are essential for investigating the system and its solutions, for example, linear topological structures. However, this language is not suitable for investigating global properties of solutions related to the topology of the manifold of independent variables and fibrations of the unknown functions. Further, this language may poorly express the invariant properties of Eqs. (1) and (2) and their consequences.

The language of differential algebra is better suited for expressing such properties and puts at the disposal of the investigator the extensive apparatus of commutative algebra, differential algebra, and algebraic geometry; this is especially true if the F_j in (1) and (2) are polynomials, and we are interested in special classes of solutions. The numerous "explicit formulas" for the solutions of the classical and newest differential equations have good interpretations in this language; the same may be said for conservation laws. However, the language of differential algebra which has been traditional since the work of Ritt does not contain the means for describing changes of the functions u_j and the variables x_i and for clarifying properties which are invariant under such changes. This is one of the main reasons for the embryonic state of the theory of so-called "Bäcklund transformations" in which there has been a recent surge of interest.

The geometric language is especially well suited for formulating and clarifying global and invariant properties of general systems of equations and for applying to them the theory of differential-geometric constructs and ideas. Its main drawback is its generality which, on the one hand, requires a rather lengthy development of foundations without concrete applications and, on the other hand, creates the risk of overlooking interesting properties of special classes of equations connected with fortuitous additional structures.

For these reasons the present paper is written in a broken and somewhat eclectic jargon in which modes of expression from all three languages are mixed in those proportions which to the author seemed most suitable for the object of study. Equivalent or comparable formulations of the same facts and constructions in different languages are often presented.

1.5. The Lagrange and Hamiltonian Equations. The classes of equations which will be of main interest to us in this work are engendered either by a Lagrangian or Hamiltonian formalism. Their choice and investigation is strongly motivated by the finite-dimensional case which corresponds to the projection $\pi: N \to$ (point). In this case the analogue of our "Lagrange" problem consists in choosing a smooth function $L: N \to \mathbf{R}$ and finding its stationary points grad $L=0$, i.e., from a general point of view it is not intrinsically a problem in differential equations. The "Hamiltonian" problem is obtained if a Hamiltonian structure is given on N which makes it possible for each Hamiltonian $H: N \to \mathbf{R}$ to construct the appropriate vector field X_H on N; it is required to investigate the properties of its trajectories.

If the base M is not a point then the role of the finite-dimensional configuration (or phase) is taken over by the infinite-dimensional space of sections $s: M \to N$ of the fibration π, and the first problem is to choose a suitable class of functionals on the sections from which the Lagrangian and Hamiltonian can be selected. Here we adopt the point of view of the classical variational calculus according to which the initial functionals have the form

$\tilde{\omega}(s) = \int_M \omega^s$, where ω in a local chart has the form $L(x_i; u_j^{(k)}) dx_1 \wedge \ldots \wedge dx_m$ and ω^s is the restriction of ω to the corresponding lift s. (If M is not compact $\tilde{\omega}(s)$ are defined on all ω with compact support or — for a particular ω — on sufficiently rapidly decreasing sections.)

In the Lagrange problem ω is called a Lagrangian density (or simply the Lagrangian), $\tilde{\omega}(s)$ is the action (for the section s), and the problem consists in finding those s, for which the action is stationary: $\delta \int_M \omega^s = 0$ in classical notation.

In the Hamiltonian problem it is, in addition, necessary to give a Hamiltonian structure: $\tilde{\omega} \mapsto X_{\tilde{\omega}}$, where $X_{\tilde{\omega}}$ is the differentiation of evolution corresponding to the Hamiltonian $\tilde{\omega}$. Not every such (linear) mapping defines a Hamiltonian structure. In analogy with the finite-dimensional case, we introduce on $\{\tilde{\omega}\}$ the Poisson bracket by the formula $\{\tilde{\omega}_1, \tilde{\omega}_2\} = X_{\tilde{\omega}_1} \tilde{\omega}_2$, and we require that the mapping $\tilde{\omega} \mapsto X_{\tilde{\omega}}$ be a Lie-algebra morphism: $X_{\{\tilde{\omega}_1, \tilde{\omega}_2\}} = [X_{\tilde{\omega}_1}, X_{\tilde{\omega}_2}]$ (the commutator of the evolution fields).

It is well known that by means of integration by parts the condition $\delta \int_M \omega^s = 0$ reduces to a system of differential equations (the vanishing of the variational derivatives or the Euler—Lagrange operators of the form ω). This same mechanism works in studying the Poisson bracket $X_{\tilde{\omega}_1} \tilde{\omega}_2$. The formalism of the variational calculus is based on the fact that essentially all computations are carried out with the integrands, i.e., in the algebra of certain operators on forms on the space of jets. Since for us Lagrangians and Hamiltonians with derivatives in x_i of arbitrarily high orders are essential, it is important to clarify the invariant meaning of the classical constructions and formulas; dealing with these objects becomes complicated as the order of the derivatives increase, while invariance under change of coordinates is almost not amenable to verification. This is the basic objective of the first chapter. In Sec. 8 we restrict our consideration to the case of a one-dimensional base and introduce the class of Hamiltonian structures which are important in the sequel.

In conclusion we remark that the choice of the class of functionals $\{\tilde{\omega}\}$ on the sections is not the only possible one nor even the most important. In field theory, for example, the class of functionals of basic importance are those generated by integral transformations of polynomials of the form $P(u^s(x^{(1)}), \ldots, u^s(x^{(N)}))$, where $(x^{(1)}, \ldots, x^{(N)}) \in M \times \ldots \times M = M^N$, and u^s are the coordinates of the section. They include, say, Fourier transformations of the coordinates, and, if generalized kernels are admitted, also all functionals of the form $\tilde{\omega}$. Moreover, they form a ring, while functionals of the form $\tilde{\omega}$ constitute only a linear space. A systematic development of the variational formalism in this class has apparently not been carried out in spite of the importance of this problem.

§2. Fields and Forms on the Space of Jets

2.1. The Basic Model. We consider a fibration $N \xrightarrow{\pi} M$, locally diffeomorphic to the projection $\mathbf{R}^{n+m} \to \mathbf{R}^m$, and we describe a number of concepts and constructions related to differential forms and vector fields on various stories of the tower of jets. The definitions and formulations of results are given in invariant terms which automatically ensures their consistency in local models (an exception is the Legendre operator for which see Sec. 4). For this reason computations may be carried out in local coordinates if desired which clearly indicates the connection with the classical formalism. We thus introduce a number of bundles and homomorphisms between them, but we carry out almost all computations in a single local chart ($\pi_V: V \to U$ and a diffeomorphism with projection $\mathbf{R}^{n+m} \to \mathbf{R}^m$), and we assume as in Sec. 1 that $K = C^\infty(U)$, $A_i = C^\infty(J^i\pi_V)$, $i \geqslant 0$, and $A_{-1} = K \subset A_0 \subset A_1 \subset \ldots$. A section $s: U \to V$ we identify with the induced homomorphism $A \to K: P \mapsto P^s$. In a local chart we have $A_i = C^\infty(x_1, \ldots, x_m;\ u_j^{(k)} \,||\, |k| \leqslant i)$; $u_j^{(k)s} = \frac{\partial^{|k|}}{\partial x_1^{k_1} \ldots \partial x_m^{k_m}} u_j^s$ for any s.

2.2. Vector Fields on Jets. If L is a smooth manifold, $B = C^\infty(L)$, we denote by $D(B)$ the Lie B-algebra of vector fields on L, considered as differentiations of B into itself. If $C \to B$ is a homomorphism induced by a smooth mapping we denote by $D(B/C)$ the subalgebra of fields of $D(B)$, which are trivial on the image of C.

For $k \geqslant i$ a field $X_k \in D(A_k)$ is called an extension or lift of a field $X_i \in D(A_i)$, if $X_k|_{A_i} = X_i$. Since $J^k\pi_V \cong J^i\pi_V \times R^{n_i}$, any field X_i extends to $D(A_k)$. We denote by $D(A_i, A_k) \subset D(A_k)$ the submodule generated over A_k by all elements of $D(A_i)$ in $D(A_k)$. Finally, we denote by $D(A)$ the set of all differentiations $X: A \to A$, such that for each i there exists a $k \geqslant i$, for which $X|_{A_i} \in D(A_i, A_k)$.

In a local chart $D(A_i)$ is a free A_i-module freely generated by the partial derivatives $\left(\frac{\partial}{\partial x_j}, \frac{\partial}{\partial u_k^{(l)}} \middle| |l| \leqslant i\right)$. Any element of $D(A)$ can we written uniquely as an infinite linear combination $X = \sum_j P_j \frac{\partial}{\partial x_j} + \sum_{k,l} Q_{k,l} \frac{\partial}{\partial u_k^{(l)}}$; $P_j, Q_{k,l} \in A$. The coefficients $P_j, Q_{k,l}$ are determined by successively restricting X to A_i with $i \to \infty$. Obviously, $P_j = Xx_j$, $Q_{k,l} = Xu_k^{(l)}$. It is easily verified that $D(A)$ forms a Lie A-algebra.

2.3. LEMMA. Let $P \in A$ be such that $P^s = 0$ for all sections s. Then $P = 0$.

Proof. If $P(x_j, u_k^{(l)}) \neq 0$, then there exist $\xi_j, \nu_k^{(l)} \in \mathbf{R}$, such that $P(\xi_j, \nu_k^{(l)}) \neq 0$. We look for functions $u_k^s \in K$, such that at the point (ξ_j) we have $(\partial^{|l|}/\partial x_1^{l_1} \ldots \partial x_m^{l_m}) u_k^s(\xi_j) = \nu_k^{(l)}$. They define a section s, for which $P^s \neq 0$.

2.4. LEMMA. There exists a unique mapping $D(K) \to D(A): X \mapsto \overline{X}$ such that for all sections $s: U \to V$ and for all $X \in D(K)$, $P \in A$ we have $(\overline{X}P)^s = XP^s$. It is an imbedding of Lie K-algebras.

Proof. $D(K)$ is freely generated over K by the fields $\partial/\partial x_i$. We set $\overline{\partial/\partial x_i} = \partial_i = \frac{\partial}{\partial x_i} + \sum_{j,k} u_j^{(k+\varepsilon_i)} \partial/\partial u_j^{(k)} \in D(A)$, where $\varepsilon_i = (0 \ldots 1 \ldots 0)$ (one sits at the i-th place) and we extend this definition by K-linearity to all of $D(K)$. Obviously, $\partial_i|_K=\partial/\partial x_i$, $(\partial_i u_j^{(k)})^s=(u_j^{(k+e_i)})^s=\frac{\partial}{\partial x_i}(u_j^{(k)s})$. From the fact that $P\mapsto P^s$ is the identity on K and from formal properties of derivatives it follows that $(\overline{X}P)^s=XP^s$ for all $X\in D(K)$ and $P\in A$.

The uniqueness of $\overline{X}$ follows from Lemma 2.3.

The fact that $X\mapsto\overline{X}$ is a Lie-algebra homomorphism follows from the computation

$$[X,Y]\,P^s=XYP^s-YXP^s=X(\overline{Y}P)^s-Y(\overline{X}P)^s=$$
$$=(\overline{X}\overline{Y}P)^s-(\overline{Y}\overline{X}P)^s=([\overline{X},\overline{Y}]\,P)^s$$

and the uniqueness of $[\overline{X,Y}]$.

Obviously, ∂_j is the "total derivative" with respect to x_j in the classical terminology. We denote by $D_c\subset D(A)$ the image of $D(K)$ under the imbedding $X\mapsto\overline{X}$. It is clear that AD_c forms a Lie subalgebra in $D(A)$. In a local chart we shall write $P^{(l)}=\partial_1^{l_1}\ldots\partial_m^{l_m}P$. This is consistent with the notation $u_k^{(l)}$ and the homorphisms $P\mapsto P^s$.

2.5. LEMMA. There exists a unique imbedding of $AD(A_0/K)$ into $D(A):Y\mapsto\overline{Y}$ such that $[\overline{Y},\overline{X}]=0$ for all $\overline{X}\in D_c$ and $\overline{Y}P=YP$ for all $P\in A_0$. Its image is the Lie subalgebra of all fields which commute with D_c and are trivial on K.

Proof. The action of $\overline{Y}$ on x_j and u_k coincides with the action of Y, and on $u_k^{(l)}$ for $|l|\geqslant 1$ it is defined by the condition $[\overline{Y},D_c]=0$: $\overline{Y}u_k^{(l)}=(Yu_k)^{(l)}$. Therefore, the only possible formula for $\overline{Y}$ has the form $\overline{Y}P=\sum(Yu_k)^{(l)}\frac{\partial P}{\partial u_k^{(l)}}$.

The required properties of $\overline{Y}$ are obvious on the generators of D_c and A_0 and are hence valid everywhere. Any field which commutes with D_c and is trivial on K, can be represented in the form $\sum Q_k^{(l)}\frac{\partial}{\partial u_k^{(l)}}$, and therefore has the form of $\overline{Y}$. Finally, such fields obviously generate a Lie algebra.

We denote by $D_{ev}\subset D(A)$ the image of $AD(A_0/K)$ in $D(A)$. Assigning to each $\overline{Y}\in D_{ev}$ the system of evolution equations $u_{j,t}=Yu_j$, we see that $\overline{Y}:A\to A$ is the "total time derivative" by virtue of this system. The condition $[D_c,\overline{Y}]=0$ means that the total time and space derivatives commute.

2.6. LEMMA. a) $[D_{ev},AD_c]\subset AD_c$. b) $D_{ev}+AD_c=D_{ev}\oplus AD_c$ (sum of spaces) is a Lie subalgebra in $D(A)$. The restriction of $D_{ev}+AD_c$ to A_0 defines an isomorphism of this space with $AD(A_0)$.

Proof. a) For $\overline{Y}\in D_{ev}$, $\overline{X}\in D_c$, $P\in A$ we have $[\overline{Y},P\overline{X}]=(\overline{Y}P)\,\overline{X}\in AD_c$, since $[\overline{Y},\overline{X}]=0$.

b) The restriction of $\sum Q_k^{(l)}\frac{\partial}{\partial u_k^{(l)}}+\sum R_j\partial_j\in D_{ev}+AD_c$ to A_0 coincides with $\sum\limits_k Q_k\frac{\partial}{\partial u_k}+\sum\limits_j R_j\left(\frac{\partial}{\partial x_j}+\sum\limits_k u_k^{(e_j)}\frac{\partial}{\partial u_k}\right)=\sum\limits_k\left(Q_k+\sum\limits_j R_j u_k^{(e_j)}\right)\frac{\partial}{\partial u_k}+\sum R_j\frac{\partial}{\partial x_j}\in AD(A_0)$. Therefore, Q_k and R_j are uniquely determined for any field of $AD(A_0)$, whence b).

For any element $X\in AD(A_0)$ we denote finally by $\bar{X}\in D(A)$ the corresponding element of $D_{ev}+AD_c$. As is evident from the proof of Lemma 2.6, on D_{ev} and D_c this canonical lift operator coincides with those defined in Lemmas 2.4 and 2.5 respectively.

Although $D_{ev}+AD_c$ far from exhausts $D(A)$, these differentiations completely suffice in order that interior products with them should uniquely determine the differential forms on the space of jets (see below). Indeed, we have the following result.

2.7. LEMMA. Fields of the form $\bar{X}+\bar{Y}$, where $X\in D(A_0/K)$, $Y\in D(K)$ form a basis of the tangent space at any point of $J^k\pi_V$.

Proof. Fields $\bar{X}|_{A_k}$ have the form $\sum_{|j|<k}(Q_i)^{(j)}\frac{\partial}{\partial u_i^{(j)}}$, and an argument analogous to that used in Lemma 2.3 shows that for any point $\xi\in J^k\pi_V$ it is possible to find $Q_i\in A$ with any prescribed values $Q_i^{(j)}$ at this point. Thus, $\bar{X}$ generates the spaces of vertical tangent vectors, while the spaces generated by $\bar{Y}$ for $Y\in D(K)$ project onto the entire tangent space at $\pi(\xi)$ and therefore give the lacking horizontal complement. (It should be mentioned that $\bar{Y}$ takes A_k onto A_{k+1}, so that, strictly speaking, $\bar{Y}$ does not define an ordinary tangent vector. However, we shall use this lemma to verify that each differential form on $J^k\pi$ is determined by its values on fields of the form $\bar{X}+\bar{Y}$, and for this purpose our argument suffices.)

2.8. The de Rham Complex. For $i\geqslant -1$ we denote by $\Omega A_i=\bigoplus_{k=0}^{\infty}\Omega^k A_i$ the exterior algebra of C^∞ differential forms on $J^i\pi$ with differential $d:\Omega^k A_i\to\Omega^{k+1}A_i$. The canonical imbedding $A_i\to A_k$ $(k\geqslant i)$ define a system of imbeddings $\Omega A_i\to\Omega A_k$ consistent with d . We set $\Omega A=\varinjlim\Omega A_i=\bigcup_{i=-1}^{\infty}\Omega A_i$, $\Omega^j A=\bigcup_{i=-1}^{\infty}\Omega^j A_i$.

In a local chart $\Omega^b A$ is freely generated over A by elements of the form $du_{i_1}^{(k_1)}\wedge\dots\wedge du_{i_a}^{(k_a)}\wedge dx_{i_{a+1}}\wedge\dots\wedge dx_{i_b}$, $i_1<\dots<i_a$, $i_{a+1}<\dots<i_b$. For any vector field $X\in D(A_i)$ and forms $\omega\in\Omega^j A_i$ the standard compositions $i_X\omega\in\Omega^{j-1}A_i$ (interior product) and $L_X\omega=(i_Xd+di_X)\omega\in\Omega^j A_i$ (Lie derivative) are defined. If $k\geqslant i$, $\tilde{X}\in D(A_k)$, is an extension of $X\in D(A_i)$, and $\omega\in\Omega(A_i)\subset\Omega(A_k)$, then $i_{\tilde{X}}\omega=i_X\omega$ and $L_{\tilde{X}}\omega=L_X\omega$. Therefore, i_X and L_X are defined for $X\in D(A)$, $\omega\in\Omega(A)$ and possess the usual properties. In particular, the L_X (respectively, i_X) are additive in X, and are differentiations (respectively, antidifferentiations) of the algebra ΩA . Moreover, $[L_X,d]=0$, $[L_X,L_Y]=L_{[X,Y]}$; $L_{PX}=PL_X+dP\wedge i_X$; $[L_X,i_Y]=i_{[X,Y]}$; $i_{PX}=Pi_X$. We shall sometimes write $X\omega$ in place of $L_X\omega$.

We now concern ourselves with the restriction of forms of ΩA to sections. Any section s defines an algebra homomorphism $\Omega A\to\Omega K$: $\omega\mapsto\omega^s$.

2.9. Proposition. a) There exists a unique operator $\tau:\Omega A\to\Omega A$ with $\tau\Omega A\subset A\Omega K$ such that $(\tau\omega)^s=\omega^s$ for all sections s.

b) $\tau^2=\tau$, and τ is a homomorphism of graded algebras with exterior multiplication.

c) $\operatorname{Ker}\tau$ is the ideal in ΩA, generated by $\operatorname{Ker}\tau\cap\Omega^1 A$.

d) $\tau L_{\bar{X}}=\tau L_{\bar{X}}\tau$ for all $\bar{X}\in D_{ev}+AD_c$; $\tau L_{\bar{X}}=L_{\bar{X}}\tau$ for all $X\in D_{ev}$.

e) $\tau d\tau = \tau d$ and $\tau d\tau d = 0$.

Proof. a), b). We define an A-homomorphism of algebras $\Omega A \to A\Omega K$ on the generators by the formulas $\tau(dx_j) = dx_j$, $\tau(du_k^{(l)}) = \sum_{l=1}^{m} u_k^{(l+e_l)} dx_l$. It is clear from the definitions that $\tau\Omega A \subseteq A\Omega K$, $(\tau\omega)^s = \omega^s$, and $\tau^2 = \tau$. Uniqueness follows from the fact that $A\Omega K$ is a free A-module with generators $dx_{i_1} \wedge \ldots \wedge dx_{i_k}$, and therefore any of its elements is uniquely determined by its s-images according to Lemma 2.3.

We remark that $\tau dP = \sum_{j=1}^{m} \partial_j P dx_j$ for all $P \in A$.

c) Since $\tau^2 = \tau$, we have $\operatorname{Ker}\tau = \operatorname{Im}(1-\tau)$. The identity $\omega\wedge\nu - \tau(\omega\wedge\nu) = (\omega - \tau\omega)\wedge\nu + \tau\omega\wedge(\nu - \tau\nu)$ and induction on i show that $\operatorname{Im}(1-\tau)\cap\Omega^i A \subset \Omega^{i-1}A(\operatorname{Im}(1-\tau)\cap\Omega^1 A)$. The reverse inclusion follows from the fact that $\operatorname{Ker}\tau$ is an ideal containing $\operatorname{Im}(1-\tau)\cap\Omega^1 A$.

d) We shall show first that $L_{\overline{X}}\operatorname{Ker}\tau \subset \operatorname{Ker}\tau$ for $\overline{X} \in D_{\mathrm{ev}} + AD_c$. Since $L_{\overline{X}}$ is a differentiation and $\operatorname{Ker}\tau$ is an ideal, it suffices to verify this inclusion on the generators of $\operatorname{Ker}\tau$, i.e., on $u_i^{(k)} - \sum_j u_i^{(k+e_j)} dx_j$. Because of the additivity of $L_{\overline{X}}$, with respect to $\overline{X}$, it is possible to consider $\overline{X}$ separately for $X \in AD(A_0/K)$ and $\overline{X} \in AD_c$; we have

$$L_{\overline{X}}\left(du_i^{(k)} - \sum_j u_i^{(k+e_j)} dx_j\right) =$$
$$= dL_{\overline{X}} u_i^{(k)} - \sum_j \left(L_{\overline{X}} u_i^{(k+e_j)} dx_j + u_i^{(k+e_j)} dL_{\overline{X}} x_j\right).$$

For $\overline{X} = \sum_{i,k} Q_i^{(k)} \frac{\partial}{\partial u_i^{(k)}}$ the right side is equal to $(1-\tau)dQ_i^{(k)} \in \operatorname{Ker}\tau$. For $\overline{X} = P\partial_l$ the right side is equal to $P(1-\tau)du_i^{(k+e_l)} \in \operatorname{Ker}\tau$.

Now $\tau L_{\overline{X}}\tau - \tau L_{\overline{X}} = \tau L_{\overline{X}}(\tau - 1)$ and $\operatorname{Im} L_{\overline{X}}(\tau-1) \subset \operatorname{Im}(\tau - 1) = \operatorname{Ker}\tau$, whence $\tau L_{\overline{X}} = \tau L_{\overline{X}}\tau$. If, moreover, $X \in AD(A_0/K)$, then $L_{\overline{X}}\operatorname{Im}\tau \subset \operatorname{Im}\tau$, since by the preceding calculation (for $\overline{X} = \sum_{i,k} Q_i^{(k)} \frac{\partial}{\partial u_i^{(k)}}$)

$L_{\overline{X}}(\tau du_i^{(k)}) = \tau dQ_i^{(k)}$. Therefore $L_{\overline{X}}\tau = \tau L_{\overline{X}}\tau = \tau L_{\overline{X}}$.

e) Obviously, $\tau d - \tau d\tau = \tau d(1-\tau)$. If $\omega^s = 0$ for all s , then $(d\omega)^s = d(\omega^s) = 0$ for all s, i.e., $d\operatorname{Ker}\tau \subset \operatorname{Ker}\tau$ and hence $\tau d(1-\tau) = 0$. Multiplying the equation $\tau d = \tau d\tau$ on the right by d , we find that $(\tau d)^2 = 0$.

3. Integration by Parts

3.1. Probably all the invariant information contained in the classical procedure of integrating the varied Lagrangian density by parts is contained in the next proposition. Here we do not even assume that the "form in the variations" $\omega \in \Omega^1 A\Omega^m K$ ($du_i^{(k)}$ is the "variation" of $u_i^{(k)}$) is obtained from a Lagrangian; this case will be dealt with in the next section.

3.2. Proposition. Let $\omega \in \Omega^1 A\Omega^m K \subset \Omega A$. There exist forms $\omega_1 \in A\Omega^1 A_0\Omega^m K$, $\omega_2 \in \Omega^1 A\Omega^m K$, $\omega_3 \in \Omega^1 A\Omega^{m-1} K$ with the following properties:

a) $\omega = \omega_1 + \omega_2$.

b) For all $\overline{X} \in D_{\mathrm{ev}}$ we have $i_{\overline{X}}\omega_2 = \tau d i_{\overline{X}}\omega_3$. The forms ω_1, ω_2 are determined uniquely. The form ω_3 can be normalized by the following additional requirement: we choose $\nu \in A\Omega^m K$ arbitrarily and impose on ω_3 the condition:

c) $\tau\omega_3 = \nu$.

Then ω_3 exists and is uniquely determined in the case $m=1$ and also in the case $\omega \in \Omega^1 A_1 \Omega^m K$, if we require, in addition, the inclusion $\omega_3 \in A\Omega^1 A_0 \Omega^{m-1} K$.

Proof. Existence. It suffices to verify the existence of ω_1, ω_2 with properties a), b) on the additive generators of the group $\Omega^1 A\Omega^m K$ which we take to be the forms $Pdu_i^{(k)} \wedge dx_1 \wedge \ldots \wedge dx_m$ in a local chart. We set $k(\omega) = |k|$ and carry out induction on $k(\omega)$. For $k(\omega)=0$ the forms $\omega_1 = \omega$, $\omega_2 = \omega_3 = 0$ satisfy a) and b). Suppose that $k(\omega) \geqslant 1$ and let $k_l \geqslant 1$. Then

$$\omega = Pdu_i^{(k)} \wedge dx_1 \wedge \ldots \wedge dx_m = -\partial_l P du_i^{(k-e_l)} \wedge dx_1 \wedge \ldots \wedge dx_m + \\ + \partial_l (Pdu_i^{(k-e_l)}) \wedge dx_1 \wedge \ldots \wedge dx_m = \omega' + \nu.$$

By the induction hypothesis it is possible to choose ω_1', ω_2' and ω_3' with $\omega' = \omega_1' + \omega_2'$, $i_{\overline{X}}\omega_2' = \tau d i_{\overline{X}}\omega_3'$ for all $\overline{X} \in D_{ev}$, and these lie in the described subgroups of ΩA . Further,

$$\begin{aligned} i_{\overline{X}}\nu &= i_{\overline{X}}(\partial_l P du_i^{(k-e_l)} \wedge dx_1 \wedge \ldots \wedge dx_m) + i_{\overline{X}}(Pdu_i^{(k)} \wedge dx_1 \wedge \ldots \wedge dx_m) = \\ &= \partial_l P (\overline{X}u_i)^{(k-e_l)} dx_1 \wedge \ldots \wedge dx_m + P(\overline{X}u_i)^{(k)} dx_1 \wedge \ldots \wedge dx_m = \\ &= \partial_l (P(\overline{X}u_i)^{(k-e_l)}) dx_1 \wedge \ldots \wedge dx_m = \\ &= (-1)^{l-1} \tau d i_{\overline{X}} (Pdu_i^{(k-e_l)} \wedge dx_1 \wedge \ldots \wedge dx_{l-1} \wedge dx_{l+1} \wedge \ldots \wedge dx_m = \tau d i_{\overline{X}} \nu'. \end{aligned}$$

It remains to set $\omega_1 = \omega_1'$, $\omega_2 = \omega_2' + \nu$, $\omega_3 = \omega_3' + \nu'$.

Without destroying properties a) and b) we may add to ω_3 any element $\nu'' \in \Omega^1 A\Omega^{m-1}K$ with the property $i_{\overline{X}}\nu'' = 0$ for all $\overline{X} \in D_{ev}$. In particular, it is possible to take $\omega_3 - \tau\omega_3 + \nu$ in place of ω_3 for any $\nu \in A\Omega^m K$, since $i_{\overline{X}}(A\Omega^m K) = \{0\}$. Then $\tau\omega_3$ is replaced by $\tau\nu = \nu$, which gives condition c).

Uniqueness. Let $\omega = \omega_1' + \omega_2'$ and $i_{\overline{X}} = \tau d i_{\overline{X}}\omega_3'$ for all $\overline{X}$, i.e., $\omega_1', \omega_2', \omega_3'$ is another triplet of forms which satisfies a), b), c). Then for all $\overline{X} \in D_{ev}$ we have, on setting $\overline{\omega}_i = \omega_i - \omega_i'$,

$$i_{\overline{X}}\overline{\omega}_1 = -i_{\overline{X}}\omega_2 = -\tau d i_{\overline{X}}\overline{\omega}_3. \tag{3}$$

From the proof of Lemma 2.7 it follows that any form $\Omega^1 A\Omega^m K$ is uniquely determined by the values of $i_{\overline{X}}\omega$ for all $X \in D(A_0/K)$, since such $\overline{X}$ generate a basis for the vertical tangent vectors at any point of the space of jets. Therefore, to establish the uniqueness of ω_1 and ω_2 it suffices to verify that $i_{\overline{X}}\omega_1 = 0$. For all sections s we have by (3) and the definition of τ :

$$(i_{\overline{X}}\overline{\omega}_1)^s = -d\,(i_{\overline{X}}\overline{\omega}_3)^s.$$

We choose any point $x \in M$, a small open neighborhood U of it with boundary ∂U, and we consider the fields $\overline{X}$ for X with support strictly inside $\pi^{-1}(U)$. By Stokes' formula $\int_M d\,(i_{\overline{X}}\overline{\omega}_3)^s = \int_{\partial U} (i_{\overline{X}}\omega_3)^s = 0$, for all such $\overline{X}$. If $\overline{\omega}_1 = \sum_{i=1}^m P_i du_i \wedge dx_1 \wedge \ldots \wedge dx_m$ and $\overline{X}u_i = Q_i$, then

$(\Sigma P_i Q_i)^s|_U=0$ for all s and $Q_i \in A_0$, whence $P_i^s|_U=0$ and therefore $\bar{\omega}_1|_U=0$. We remark that this conclusion remains in force if it is assumed that $\bar{\omega}_3$ depends on $\bar{X}$. Finally, $\bar{\omega}_1=\bar{\omega}_2=0$, and according to (3) $\tau d i_{\bar{X}} \bar{\omega}_3=0$ for all $\bar{X} \in D_{ev}$.

Even the normalization of ω_3 by condition c), however, does not guarantee the uniqueness of $\bar{\omega}_3$. We cannot describe the complete kernel $\bigcap_{\bar{X}\in D_{ev}} \operatorname{Ker} \tau d i_{\bar{X}}$, but it contains, e.g., $d\tau d\Omega^{m-2}K$. Indeed,

$$\tau d\left(i_{\bar{X}} d\tau d\right)=\tau\left(d i_{\bar{X}}\right) d\tau d=$$
$$=\tau\left(L_{\bar{X}}-i_{\bar{X}} d\right) d\tau d=\tau L_{\bar{X}} d\tau d=\tau L_{\bar{X}} \tau d\tau d=0,$$

according to Proposition 2.9 d) and e).

It remains to consider cases in which we are able to guarantee the uniqueness of ω_3.

The case $m=1$. Up to a term in $A\Omega^m K$, which is uniquely normalized by condition c) and lies in the kernel of all the $\tau d i_{\bar{X}}$, the element $\bar{\omega}_3$ can be represented in the form $\bar{\omega}_3=\sum_k P_{j,k} du_j^{(k)}$, $P_{j,k} \in A$. If $\tau d i_{\bar{X}} \bar{\omega}_3=0$, then $\tau d\left(\sum P_{jk} Q_j^{(k)}\right)=\sum \partial_1\left(P_{jk} Q_j^{(k)}\right) dx_1=0$ for all $Q_j=\bar{X} u_j \in A$, i.e., $\sum P_{j,k} Q_j^{(k)}=\text{const}$ for any Q_j. From this it easily follows that $P_{j,k}=0$.

The case $\omega \in \Omega^1 A_1 \Omega^m K$, $\omega_3 \in A\Omega^1 A_0 \Omega^{m-1} K$. The existence of ω_3 in this subgroup for $\omega \in \Omega^1 A_1 \Omega^m K$ is obvious from the construction of ω_3 at the beginning of the proof. Up to a term of $A\Omega^m K$ the element $\bar{\omega}_3 \in A\Omega^1 A_0 \Omega^{m-1} K$, the difference of two choices of ω_3 can be represented in the form $\bar{\omega}_3=\sum du_i \wedge \omega_i$, $\omega_i \in A\Omega^{m-1} K$. Further,

$$i_{\bar{X}} \bar{\omega}_3=\sum_i \bar{X} u_i \omega_i=\sum_i \bar{X} u_i\left(\sum P_{i,l} dx_1 \wedge \ldots \wedge \widehat{dx_l} \wedge \ldots \wedge dx_m\right),$$
$$\tau d i_{\bar{X}} \bar{\omega}_3=\sum_{i,l}(-1)^{l-1} \partial_l\left(\bar{X} u_i P_{i,l}\right) dx_1 \wedge \ldots \wedge dx_m.$$

If the last expression is identically zero, then, since $\bar{X} u_i=Q_i \in A$ may be chosen arbitrarily and independently, we find that for each i the sum $\sum_l (-1)^{l-1} \times \partial_l\left(Q P_{i,l}\right)$ does not depend on Q. It is easy to see that this is possible only for $P_{i,l}=0$, which completes the proof.

We denote the forms ω_1 and ω_3, constructed in Proposition 3.2 on the basis of the form $\omega \in \Omega^1 A\Omega^m K$, by $\hat{\delta}\omega$ and $\hat{S}_v \omega$, respectively. We shall indicate their explicit forms in coordinates.

We begin with the following classical lemma.

Let B be a (not necessarily commutative) ring, and let $\partial_1, \ldots, \partial_m$ be its differentiations into itself. For any sequence of numbers $i_1, i_2, i_3, \ldots$, $1 \leqslant i_j \leqslant m$ we set $\partial\left(i_1, \ldots, i_k\right)=\partial_{i_1} \ldots \partial_{i_k}$: $B \rightarrow B$; $\partial(\varnothing)=\mathrm{id}$.

<u>3.3. LEMMA.</u> For any $x, y \in B$ we have

$$\partial(i_1, \dots, i_k)xy - (-1)^k x\partial(i_k, \dots, i_1)y =$$
$$= \sum_{a=1}^{k} (-1)^{a-1} \partial_{i_a} (\partial(i_{a+1}, \dots, i_k) x \partial(i_{a-1}, \dots, i_1) y).$$

The proof is obvious.

3.4. COROLLARY. We assume that $[\partial_i, \partial_j] = 0$ for all $1 \leqslant i, \ j \leqslant m$. Then there exist numbers $C(j; k, l) \geqslant 0$, $1 \leqslant j \leqslant m$; $k, l \in N^m$, such that for all $s \in N^m$, $\partial^s = \partial_1^{s_1} \dots \partial_m^{s_m}$

$$\partial^s xy - (-1)^{|s|} x\partial^s y = \sum_{\substack{j=1 \\ \varepsilon_j + k + l = s}}^{m} (-1)^{|l|+1} C(j; k, l) \partial_j (\partial^k x \partial^l y).$$

Proof. We write out the formula of Lemma 3.3 for all sequences $(i_1, \dots, i_{|s|})$ with $\sum_{j=1}^{|s|} \varepsilon_{i_j} = s$, take their arithmetic mean, and collect like terms.

We remark that for $m = 1$ we have $C(j, k, l) = 1$.

3.5. We can now write down formulas for $\hat{\delta}\omega$ and $\hat{S}_\nu \omega$. Let $\omega = \sum P_{l,s} du_l^{(s)} dx_1 \wedge \dots \wedge dx_m$. We set

$$d^m x = dx_1 \wedge \dots \wedge dx_m, \ d_j^m x = dx_1 \wedge \dots \wedge dx_{j-1} \wedge dx_{j+1} \wedge \dots \wedge dx_m.$$

Then by 3.4 we find that

$$\hat{\delta}\omega = \sum_s (-1)^{|s|} \partial^s P_{l,s} du_l \wedge d^m x,$$
$$\omega - \hat{\delta}\omega = \sum_{k+l+\varepsilon_j = s} (-1)^{|l|+1} C(j, k, l) \partial_j (\partial^l P_{l,s} du_l^{(k)}) \wedge d^m x.$$

For any $\overline{X} \in D_{ev}$ we have $i_{\overline{X}} (\partial_j (\partial^l P_{l,s} du_l^{(k)} \wedge d^m x)) = \tau d i_{\overline{X}} \times (-1)^{j-1} (\partial^l P_{l,s} du_l^{(k)}) \wedge d_j^m x$. Therefore

$$i_{\overline{X}} (\omega - \hat{\delta}\omega) = \tau d i_{\overline{X}} \sum_{k+l+\varepsilon_j = s} (-1)^{|l|+j} C(j, k, l) \partial^l P_{l,s} du_l^{(k)} \wedge d_j^m x.$$

Replacing here $du_l^{(k)}$ by $(d - \tau d) u_l^{(k)}$, we do not change the right side. In place of $(d - \tau d) u_l^{(k)}$ it is also possible to take $du_l^{(k)} - u_l^{(k+\varepsilon_j)} dx_j$ in the term $d_j^m x$; the remainder terms give zero. After this change the form under the sign $\tau d i_{\overline{X}}$ is in the kernel of τ. Finally, adding ν to it in order to satisfy condition c) of Proposition 3.2, we obtain finally

$$\hat{S}_\nu \omega = \nu + \sum_{k+l+\varepsilon_j = s} (-1)^{|l|+j} C(j, k, l) \partial^l P_{l,s} (du_l^{(k)} - u_l^{(k+\varepsilon_j)} dx_j) \wedge d_j^m x.$$

4. The Euler–Lagrange Operators and the Legendre Transformation

4.1. THEOREM. Let ω be any Lagrangian density for the fibration $\pi: N \to M$, i.e., a form which belongs locally to $A\Omega^m K$. Then there exist forms $\delta\omega$ and $S\omega$, locally belonging to $A\Omega^1 A_0 \Omega^m K$ and $\Omega^1 A \Omega^{m-1} K$, respectively, such that for all $\overline{X} \in D_{ev} + AD_c$ we have (locally)

$$\tau L_{\bar{X}}\omega=\tau\,(i_{\bar{X}}d+di_{\bar{X}})\,\omega=\tau\,(i_{\bar{X}}\delta\omega+di_{\bar{X}}S\omega) \tag{4}$$

and, moreover, $\tau S\omega=\omega$.

The form $\delta\omega$ is always uniquely defined by these conditions, while the form $S\omega$ is uniquely defined in the case $m=1$ or in the case $\omega\in A_1\Omega^m K$ if it is additionally required that $S\omega\in A\Omega^1 A_0\Omega^{m-1}K$ (locally).

Here δ is called the Euler–Lagrange operator, and S is the Legendre transform.

<u>Proof. a) The local case.</u> We set $\delta\omega=\hat{\delta}d\omega$, $S\omega=\hat{S}_\omega d\omega$, where $\hat{\delta}$ and $\hat{S}$ are defined as in Proposition 3.2, where ν in condition c) is ω. We then obtain according to 3.2:

$$d\omega=\delta\omega+\omega_2;\quad i_{\bar{X}}\omega_2=\tau di_{\bar{X}}S\omega \ \text{ for } \ \bar{X}\in D_{\mathrm{ev}};\quad \tau S\omega=\omega. \tag{5}$$

From this we deduce the identity (4) for $\bar{X}\in D_{\mathrm{ev}}$ and $\bar{X}\in AD_c$ individually.

If $\bar{X}\in D_{\mathrm{ev}}$, then $i_{\bar{X}}\omega=0$, whence according to (5)

$$\tau L_{\bar{X}}\omega=\tau i_{\bar{X}}d\omega=\tau i_{\bar{X}}\,(\delta\omega+\omega_2)=\tau i_{\bar{X}}\delta\omega+\tau di_{\bar{X}}S\omega.$$

The uniqueness properties of $\delta\omega$ and $S\omega$ follow from those proved in Proposition 3.2 if it is noted that $d\omega$, $\delta\omega\in A\Omega^1 A_1\Omega^m K$, so that these forms are uniquely determined by all values of $i_{\bar{X}}d\omega$, $i_{\bar{X}}\delta\omega$ for $\bar{X}\in D_{\mathrm{ev}}$, and these values lie in $A\Omega^m K$; therefore, τ is the identity on them, and hence the decomposition $d\omega=\delta\omega+\omega_2$ with the properties postulated in Proposition 3.2 follows from (4) for $\bar{X}\in D_{\mathrm{ev}}$.

It thus remain to verify (4) for $\bar{X}\in AD_c$. We shall first establish that $\tau(i_{\bar{X}}\Omega^{m+1}A)=0$, i.e., $(i_{\bar{X}}\Omega^{m+1}A)^s=0$ for all sections s. Since $(\Omega^{m+1}A)^s=\{0\}$, it suffices to verify that $(i_{\bar{X}}\nu)^s=i_X\nu^s$ for $X\in D(K)$. Since $i_{\bar{X}}$ is a differentiation and restriction to s is a homomorphism of the algebra of forms, it suffices to verify it on the generators. On A both sides coincide. Further,

$$(i_{\bar{X}}dx_j)^s=(\bar{X}x_j)^s=(Xx_j)^s=i_X\,(dx_j)^s,$$

$$(i_{\bar{X}}du_i^{(k)})^s=(\bar{X}u_i^{(k)})^s=\Big(\sum_l u_i^{(k+e_l)}Xx_l\Big)^s=\sum u_i^{(k+e_l)^s}Xx_l,$$

$$i_X\,(du_i^{(k)})^s=i_X\Big(\sum_l u_i^{(k+e_l)}dx_l\Big)^s=\sum_l u_i^{(k+e_l)^s}Xx_l.$$

Using this, we see that it is necessary to verify the formula

$$\tau di_{\bar{X}}\omega=\tau di_{\bar{X}}S\omega.$$

We have

$$\tau di_{\bar{X}}S\omega=\tau L_{\bar{X}}S\omega-\tau(i_{\bar{X}}dS\omega)=\tau L_{\bar{X}}S\omega.$$

From Proposition 2.9 we have

$$\tau L_{\bar X} S\omega = \tau L_{\bar X} \tau S\omega = \tau\,[d i_{\bar X} + i_{\bar X} d]\, \tau S\omega = \tau d i_{\bar X} \tau S\omega.$$

But $\tau S\omega = \omega$, which gives the required result.

b) The Global Case. Since $\delta\omega$ is locally uniquely defined and the properties characterizing it are consistent with the operation of restricting to a subdomain, all local constructions automatically fit together. For $S\omega$ an analogous argument goes through in cases where uniqueness is guaranteed. Aside from these cases, as has been shown, it does not hold; e.g., it is possible to add locally to $S\omega$ any element of $d\tau d(A\Omega^{m-2}K)$ without changing the values of $\tau d i_{\bar X} S\omega$ for all $\bar X \in D_{ev} + AD_c$. Therefore, to prove the existence of a global form $S\omega$ a special treatment is necessary in order to establish the possibility of consistent local constructions. This has been done by B. A. Kupershmidt, but we shall not reproduce the proof, since it will not be needed below. It is essentially a question of the vanishing of a certain cohomological obstacle, and the problems which arise here merit special attention. The question as to what addition conditions (functorial property, for example) it is necessary to impose on $S\omega$ in order that $S\omega$ may be uniquely chosen has not been solved.

4.2. Classical Formulas. We shall apply the formulas of 3.5 to the case $\omega = P d^m x$, $d\omega = \sum \frac{\partial P}{\partial u_i^{(s)}} du_i^{(s)} \wedge d^m x$. We obtain

$$\delta\omega = \hat\delta d\omega = \sum_{i=1}^{m} \frac{\delta P}{\delta u_i} du_i \wedge d^m x, \quad \frac{\delta P}{\delta u_i} = \sum_s (-1)^{|s|} \left(\frac{\partial P}{\partial u_i^{(s)}}\right)^{(s)}.$$

Further,

$$S\omega = \omega + \sum_{i,j,k} \frac{\delta L}{\delta u_i^{(k+e_j)}} (du_i^{(k)} - u_i^{(k+e_j)} dx_j) \wedge d_j^m x,$$

where

$$\frac{\delta L}{\delta u_i^{(k+e_j)}} = \sum_l (-1)^{|l|+j} C(j, k, l) \left(\frac{\partial P}{\partial u_i^{(k+e_j+l)}}\right)^{(l)}$$

(it can be shown that $C(j; k, l)$ depends only on $k + e_j, l$).

5. The Variational Complex

5.1. Definition of the Complex. The beginning of the complex is locally as follows:

$$A \xrightarrow{\tau d} A\Omega^1 K \xrightarrow{\tau d} \ldots \xrightarrow{\tau d} A\Omega^m K \xrightarrow{\delta} A\Omega^1 A_0 \Omega^m K. \tag{6}$$

The fact that $(\tau d)^2 = 0$, is proved in Proposition 2.9 d). The equality $\delta\tau d = 0$ is established as follows. The operator $\tau L_{\bar X} d$ on $A\Omega^{m-1}K$ can be transformed in two ways. First of all,

$$\tau L_{\bar X} d = \tau (i_{\bar X} d + d i_{\bar X}) d = \tau d (i_{\bar X} d).$$

Secondly, by formula (4) and Proposition 2.9 d) for $\overline{X}\in D_{ev}+AD_c$ $\tau L_{\overline{X}}d=\tau L_{\overline{X}}\tau d=\tau i_{\overline{X}}\delta\tau d+\tau d(i_{\overline{X}}S\tau d)$. From the uniqueness of the Euler–Lagrange operator we therefore find that $\delta\tau d=0$.

We shall now indicate how to extend complex (6) to the right. The extension operator will also be essentially an Euler–Lagrange operator but in a larger bundle $\pi\times\mathrm{id}:N\times\mathbf{R}\to M\times\mathbf{R}$. Locally we denote by x_{m+1} the coordinate on the additional factor $\mathbf{R}$, and we set $K^{(1)}=C^\infty(x_1,\ldots,x_{m+1})$, $A^{(1)}=\bigcup_{l=0}^{\infty}C^\infty(x_j,\ u_i^{(k)}\,|\,i=1,\ldots,n;\ k\in N^{m+1},\ |k|\leqslant l)$. The canonical inbedding $A\subset A^{(1)}$, corresponding to the projection of $\pi\times\mathrm{id}$ onto π, in local coordinates can be identified with the imbedding $u_i^{(k)}\mapsto u_i^{(k,0)}$, which thus has an invariant meaning: the coordinates "along the fiber" u_i in the extension to $N\times\mathbf{R}$ are assumed to be independent of the new variable.

The imbedding $A\to A^{(1)}$ determines the imbedding $\Omega A\to\Omega A^{(1)}$ which also corresponds to the lift from π to $\pi\times\mathrm{id}$. Identifying $\Omega^1A\Omega^mK$ with its image in $\Omega^1A^{(1)}\Omega^mK^{(1)}$, we may consider the composition of the operators $\Omega^1A\Omega^mK\to\Omega^1A^{(1)}\Omega^mK^{(1)}\xrightarrow{\tau^{(1)}}A^{(1)}\Omega^{m+1}K^{(1)}\xrightarrow{\delta^{(1)}}\Omega^1A^{(1)}\Omega^{m+1}K^{(1)}$, which we denote simply by $\delta^{(1)}\tau^{(1)}$ ($\delta^{(1)}$ is the Euler–Lagrange operator, and $\tau^{(1)}$ is the operator τ for the fibration $\pi\times\mathrm{id}$). This construction can obviously be iterated which leads to the sequence

$$\ldots\xrightarrow{\tau d}A\Omega^mK\xrightarrow{\delta}A\Omega^1A_0\Omega^mK\xrightarrow{\delta^{(1)}\tau^{(1)}}A^{(1)}\Omega^1A_0^{(1)}\Omega^{m+1}K^{(1)}\xrightarrow{\delta^{(2)}\tau^{(2)}}A^{(2)}\Omega^1A_0^{(2)}\Omega^{m+2}K^{(2)}\to\ldots \tag{7}$$

5.2. THEOREM. Sequence (7) is a complex which we shall call the (local) variational complex of the fibration π. The global complex is a bundle complex the local sections of which have the form (7).

Proof. We first of all introduce a new operator $\tau^+:\Omega^{m+k}A\to\Omega^kA\Omega^mK$ by the condition: for all $X_1,\ldots,X_k\in D(A/K)$ and $\omega\in\Omega^{m+k}A$

$$i_{X_1}\ldots i_{X_k}\tau^+\omega=\tau(i_{X_1}\ldots i_{X_k}\omega). \tag{8}$$

The existence and uniqueness of the operator τ^+ are obvious, since a form in $\Omega^kA\Omega^mK$ can be considered a skew-symmetric, multilinear function of k fields in $D(A/K)$ with values in Ω^mK. It is clear that τ^+ is A-linear.

5.3. LEMMA. For any Lagrangian $\omega\in A\Omega^mK$ we have $\delta\omega=\tau^+dS\omega$.

Proof. It suffices to verify that $i_{\overline{X}}\delta\omega=\tau i_{\overline{X}}dS\omega$ for all $\overline{X}\in D_{ev}$, because of Lemma 2.7. According to formula (4) and Proposition 2.9 d), we have

$$L_{\overline{X}}\omega=\tau L_{\overline{X}}\omega=\tau di_{\overline{X}}S\omega+i_{\overline{X}}\delta\omega=\tau(L_{\overline{X}}-i_{\overline{X}}d)S\omega+i_{\overline{X}}\delta\omega=$$
$$=\tau L_{\overline{X}}\tau S\omega-\tau i_{\overline{X}}dS\omega+i_{\overline{X}}\delta\omega=\tau L_{\overline{X}}\omega-\tau i_{\overline{X}}dS\omega+i_{\overline{X}}\delta\omega.$$

Therefore, $\tau i_{\overline{X}}dS\omega=i_{\overline{X}}\delta\omega$, as required.

5.4. LEMMA. $\tau^{(1)}=\tau^{(1)}\tau^+$ on $\Omega^2A\Omega^{m-1}K$.

Proof. In view of the A-linearity of both operators, it suffices to verify this equality on generators of the form $du_i^{(k)}\wedge d^mx$ and $du_i^{(k)}\wedge du_j^{(l)}\wedge d_r^mx$. On forms of the first kind the operator τ^+ is the identity. Further, as is not hard to check,

$$\tau^+(du_i^{(k)}\wedge du_j^{(l)}\wedge d_r^m x)=(-1)^{r-1}(u_j^{(l+\varepsilon_r)}du_i^{(k)}-u_i^{(k+\varepsilon_r)}du_j^{(l)})\wedge d^m x,$$

whence it follows that the value of $\tau^{(1)}\tau^+$ on this form is equal to (even in $\Omega A^{(1)}$)

$$(-1)^{r-1}(u_j^{(l+\varepsilon_r)}u_i^{(k+\varepsilon_{m+1})}-u_i^{(k+\varepsilon_r)}u_j^{(l+\varepsilon_{m+1})})\,dx_{m+1}\wedge dx_1\wedge\ldots\wedge dx_m.$$

On the other hand, the value of $\tau^{(1)}$ on our form is equal to

$$(u_i^{(k+\varepsilon_r)}dx_r+u_i^{(k+\varepsilon_{m+1})}dx_{m+1})\wedge(u_j^{(l+\varepsilon_r)}dx_r+u_j^{(l+\varepsilon_{m+1})}dx_{m+1})\wedge d^m x.$$

The last two expressions obviously coincide.

5.5. Completion of the Proof of Theorem 5.2. According to Lemmas 5.3 and 5.4

$$\delta^{(1)}\tau^{(1)}\delta=\delta^{(1)}\tau^{(1)}\tau^+dS=\delta^{(1)}\tau^{(1)}dS.$$

But $\delta^{(1)}\tau^{(1)}d=0$, by the argument at the beginning of this section applied to the ring $A^{(1)}$. Thus, (7) is a complex at the term $\Omega^1 A\Omega^m K$ and hence at all remaining new terms: $\delta^{(a+1)}\tau^{(a+1)}\delta^{(a)}=0$ for all $a\geqslant 0$. The fact that it is a complex in the terms to the left of $\Omega^1 A\Omega^m K$ was verified earlier.

The exactness of complex (7) in the terms to the left of $A\Omega^m K$ (the "lift" of the de Rham complex to the jets) was recently proved by A. M. Vinogradov. The exactness of the global complex (7) in the remaining terms was established by B. V. Kupershmidt. We shall restrict ourselves to the proof of the classical part ("if the variational derivative of the Lagrangian is equal to zero, then it is a divergence").

5.6. THEOREM. $\operatorname{Ker}\delta=\operatorname{Im}\tau d$.

Proof. We consider a contractive homotopy $\varphi_t:V\to V$, $t\in[0,1]$ of the open set $V\subset N$, over which the complex (6) is defined:

$$\varphi_t:(x_i,u_j)\mapsto(tx_i,u_j\exp[1-(1-t)^{-2}]).$$

It is obvious that φ_t is a diffeomorphism of the fiber space $V\to U$ onto itself with $0\leqslant t<1$, $\varphi_0=\mathrm{id}$ and $\varphi_1:U\to$ (point). Let $X_t\in D(A_0)$ be the corresponding vector field, and let $\bar X_t$ be its canonical lift to $D(A)$ according to Lemma 2.6. Since $X_t\in D(A_0)$, it is obvious that $\tau L_{\bar X_t}\omega=L_{\bar X_t}\omega$ for all $\omega\in A\Omega^m K$. Since X_t is the derivative of φ_t, we have $X_t=(\varphi_t^*)^{-1}\frac{d}{dt}(\varphi_t^*)$ for $t<1$, and hence $\bar\varphi_t^* L_{\bar X_t}=\frac{d}{dt}\bar\varphi_t^*$, where $\bar\varphi_t$ is the lift of φ_t to the jets. Applying $\bar\varphi_t^*$ to formula (4), we find that

$$\begin{aligned}\frac{d}{dt}\bar\varphi_t^*\omega=\bar\varphi_t^*L_{\bar X_t}\omega=\bar\varphi_t^*\tau d i_{\bar X_t}S\omega+\bar\varphi_t^*\tau i_{\bar X_t}\delta\omega=\\=\tau d\bar\varphi_t^* i_{\bar X_t}S\omega+\tau\bar\varphi_t^* i_{\bar X_t}\delta\omega,\end{aligned}\tag{9}$$

since, as is easily seen, $\bar\varphi_t^*\tau=\tau\bar\varphi_t^*$ (the mapping $\bar\varphi_t^*$ is linear on x_j and $u_i^{(k)}$). Although the field $\bar X_1$ is not defined, its limit $\bar\varphi_t^* L_{\bar X_t}\omega$ exists as $t\to 1$ for all ω. Therefore, integrating (9) on t from 0 to 1, we find

$$-\omega=\int_0^1\frac{d}{dt}\bar\varphi_t^*\omega dt=\tau d\int_0^1\bar\varphi_t^* i_{\bar X_t}S\omega dt+\tau\int_0^1\bar\varphi_t^* i_{\bar X_t}\delta\omega dt=\tau d\tau\int_0^1\bar\varphi_t^* i_{\bar X_t}S\omega dt+\tau\int_0^1\bar\varphi_t^* i_{\bar X_t}\delta\omega dt\overset{\text{def}}{=}-\tau d\psi_1(\omega)-\psi_2(\delta\omega),$$

since $\tau d = \tau d\tau$. Thus,

$$\mathrm{id}|_{A\Omega^m K} = \tau d\psi_1 + \psi_2 \delta, \quad \psi_1(\omega) = -\tau \int_0^1 \bar{\varphi}_t^* i_{\bar{X}_t} S\omega dt,$$

and we find that if $\delta\omega = 0$, then $\omega = \tau d\psi_1(\omega)$ which proves the theorem.

We remark that if the form ω is a polynomial in $u_i^{(k)}$ (respectively, in $u_i^{(k)}$, x_j), then so are also $S\omega$, $\delta\omega$ polynomials by virtue of the formulas of 4.2 and with these also $\psi_1(\omega)$, $\psi_2(\delta\omega)$, since $\bar{\varphi}_t^*$ is linear in $u_i^{(k)}$, x_j.

<u>5.7. The Formula $\delta^{(1)}\tau^{(1)}\delta = 0$ in Coordinates.</u> We set $\omega = P d^m x$. Then $\delta\omega = \sum_{i=1}^{n} \frac{\delta P}{\delta u_i} du_i \wedge d^m x$. Further $\tau^{(1)} du_i = \sum_{l=1}^{m} u_i^{(e_l)} dx_l + u_i^{(e_{m+1})} dx_{m+1}$, whence

$$\tau^{(1)}\delta\omega = \sum_{i=1}^{n} \frac{\delta P}{\delta u_i} u_i^{(e_{m+1})} dx_{m+1} \wedge dx_1 \wedge \dots \wedge dx_m.$$

Therefore,

$$\delta^{(1)}\tau^{(1)}\delta\omega = \sum_{i,j} \frac{\delta}{\delta u_j}\left(\frac{\delta P}{\delta u_i} u_i^{(e_{m+1})}\right) du_j \wedge dx_{m+1} \wedge d^m x =$$

$$= \sum_{i,j}\left(\sum_{s\in N^m} (-1)^{|s|}\left(\frac{\partial}{\partial u_j^{(s)}} \frac{\delta P}{\delta u_i} \cdot u_i^{(e_{m+1})}\right)^{(s)} du_j \wedge dx_{m+1} \wedge d^m x - \sum_{i=1}^{n}\left(\frac{\delta P}{\delta u_i}\right)^{(e_{m+1})} du_i \wedge dx_{m+1} \wedge dx_1 \wedge \dots \wedge dx_m\right).$$

Further,

$$\left(\frac{\delta P}{\delta u_i}\right)^{(e_{m+1})} = \sum_{j,k} \frac{\partial}{\partial u_j^{(k)}} \frac{\delta P}{\delta u_i} \partial^k u_j^{(e_{m+1})}.$$

Therefore, after division by $dx_{m+1} \wedge d^m x$ the formula $\delta^{(1)}\tau^{(1)}\delta\omega = 0$ assumes finally the following form:

$$\sum_{i=1}^{n} \sum_{j,k\in N^m} \frac{\partial}{\partial u_j^{(k)}} \frac{\delta P}{\delta u_i} \partial^k u_j^{(e_{m+1})} du_i = \sum_{i,j,s\in N^m} (-1)^{|s|}\left(\frac{\partial}{\partial u_i^{(s)}} \frac{\delta P}{\delta u_j} u_j^{(e_{m+1})}\right)^{(s)} du_i.$$

We recall now that over A the variables $\partial^k u_j^{(e_{m+1})}$ are free. Therefore, for all i, j we obtain a family of equalities among differential operators which is equivalent to the formula $\delta^{(1)}\tau^{(1)}\delta = 0$.

<u>5.8. LEMMA.</u> For any $P \in A$ and $1 \leqslant i, j \leqslant n$ we have

$$\sum_{s\in N^m}\left(\frac{\partial}{\partial u_j^{(s)}} \frac{\delta P}{\delta u_i}\right)\partial^s = \sum_{s\in N^m} (-1)^{|s|}\partial^s \circ \left(\frac{\partial}{\partial u_i^{(s)}} \frac{\delta P}{\delta u_j}\right).$$

We shall need this lemma and its corollary in Sec. 7 to investigate the Hamiltonian structure.

<u>5.9. COROLLARY.</u> For any $\bar{X} \in D_{\mathrm{ev}}$

$$\bar{X} \frac{\delta P}{\delta u_i} = \sum_{j,s} (-1)^{|s|}\partial^s\left(\frac{\partial}{\partial u_i^{(s)}} \frac{\delta P}{\delta u_j} \cdot \bar{X} u_j\right).$$

Proof. We apply the identity of Lemma 5.8 to $\overline{X}u_j$ and sum on j.

6. Nöther's Theorem and Lagrangian Conservation Laws

6.1. Let $\omega = Pd^m x \in A\Omega^m K$ be some Lagrangian. It is evident from the formulas of 4.2 that the equation $(\delta\omega)^s = 0$ relative to unknown series s coincides with the classical system of Euler–Lagrange equations for the functions $u_1^s, \ldots, u_n^s$ with Lagrangian P. Its solutions are called extremals of the Lagrangian ω.

6.2. Integrals, Flows, and Symmetries: Analytic Version. A conservation law or an integral for the form ω is a form $\nu \in A\Omega^{m-1}K$ such that $d(\nu^s) = 0$ for any extremal s of the Lagrangian ω. For $m = 1$ this means that $\nu \in A$ is a functional of the space of jets which is constant along any extremal. In the general case for $\nu = \sum P_l d_l^m x$ the relation $d(\nu^s) = 0$ has the form $\left(\sum_{l=1}^{m} (-1)^{l-1} \partial_l P_l\right)^s = 0$ along any extremal.

A flow for ω is a field $\overline{Y} \in AD_c$, such that $i_{\overline{Y}}\omega$ is an integral. If $\omega = Pd^m x, \overline{Y} = \sum_{l=1}^{m} Q_l \partial_l$, then $i_{\overline{Y}}\omega = \sum_{l=1}^{m} (-1)^{l-1} PQ_l d_l^m x$. It is evident from this that if P does not vanish on its extremals, then for any conservation law ν there is a unique current $\overline{Y}$ such that $\nu = i_{\overline{Y}}\omega$.

The classical Nöther theorem affords the construction of conservation laws on the basis of a Lie group G, which acts on $\pi: N \to M$ and preserves the Lagrangian ω or even just the action $s \mapsto \int_M \omega^s$. If X is an element of the Lie algebra of the group G, considered as a field on N, then the G-invariance of ω implies that $L_{\overline{X}}\omega = 0$. More generally, let $X \in AD(A_0)$ be a field on the jets such that $(L_{\overline{X}}\omega)^s = 0$ for any extremal s of the Lagrangian ω. We call it a (formal) symmetry of the Lagrangian ω.

6.3. Integrals, Flows, and Symmetries: Algebraic Version. The definitions of the preceding section appeal to the set of all extremals of the Lagrangian ω, which may be empty or very complicated. In practice integrals, currents, and symmetries always satisfy the following algebraic version of the definition whose application does not require vanishing on the extremals. Let $I(\delta\omega)$ be the minimal D_c-closed ideal in A, generated by expressions of the form $i_X i_{Y_1} \ldots i_{Y_m} \delta\omega$, where $X \in D(A/K)$, $Y_1, \ldots, Y_m \in D(K)$. Let $J(\delta\omega) \subset \Omega^1 A\Omega^{m-1}K$ be the minimal D_c-closed ideal in the ring ΩA, generated by $I(\delta\omega)$.

We call ν an algebraic conservation law if $d\nu \in J(\delta\omega)$. We call $\overline{Y}$ an algebraic current if $i_{\overline{Y}}\omega$ is an algebraic conservation law. Finally, $X \in AD(A_0)$ is an algebraic symmetry if $L_{\overline{X}}\omega \in J(\delta\omega)$.

We point out that the use of this definition makes it possible to obtain additional information. For example, if ν is a conservation law, then the explicit representation of $d\nu$ as an element of $J(\delta\omega)$ determines important invariants; the theory of characteristics of Gel'fand and Dikii [2, 3] is based on this.

6.4. THEOREM. a) If X is a symmetry, then $\tau i_{\overline{X}} S\omega$ is a conservation law (the formal Nöther theorem).

b) For any flow $\bar{Y}$ the field Y is a formal symmetry.

Proof. We write the proof in the analytic version; the algebraic version may be considered analogously.

a) Since X is a symmetry, for any extremal s we have $(L_{\bar{X}}\omega)^s=0$. It therefore follows from (4) that $(i_{\bar{X}}\delta\omega)^s+d(i_{\bar{X}}S\omega)^s=0$ on the extremals, whence

$$d(i_{\bar{X}}S\omega)^s=d(\tau i_{\bar{X}}S\omega)^s=-(i_X\delta\omega)^s=0.$$

The last inequality follows from the fact that locally $\delta\omega=\sum\frac{\delta P}{\delta u_i}\wedge d^m x$, so that $i_{\bar{X}}\delta\omega$ is a linear combination of the $\frac{\delta P}{\delta u_i}$, which are zero on the extremals.

b) Since $\bar{Y}$ is a flow, we have $d(i_{\bar{Y}}\omega)^s=0$ on the extremals, so that

$$(L_{\bar{Y}}\omega)^s=(di_{\bar{Y}}\omega+i_{\bar{Y}}d\omega)^s=(i_{\bar{Y}}d\omega)^s.$$

But the last expression is zero as verified in the proof of Theorem 4.1 a).

7. The Hamiltonian Structure

7.1. We shall give a local definition of the Hamiltonian structure on $\pi:N\to M$; globalization goes through automatically.

7.2. Definition. Let Γ be an $\mathbf{R}$-linear operator $\Gamma:A\Omega^m K\to D_{ev}$. We define a bilinear composition law on $A\Omega^m K$, by setting

$$\{\omega_1,\omega_2\}_\Gamma=L_{\Gamma(\omega_1)}\omega_2. \tag{10}$$

The operator Γ gives a Hamiltonian structure if it takes $\{\,,\,\}_\Gamma$ into the commutator in the Lie algebra D_{ev},

$$\Gamma\{\omega_1,\omega_2\}_\Gamma=[\Gamma(\omega_1),\Gamma(\omega_2)], \tag{11}$$

and its kernel contains $\operatorname{Im}\tau d=\operatorname{Ker}\delta$.

7.3. Comments. The motivation for this definition was given in Sec. 1; it is parallel to one of the characterizations of a finite-dimensional Hamiltonian structure. In particular, here $\omega\in A\Omega^m K$ is considered as a representative of the functional $\tilde{\omega}(s)=\int_M\omega^s$ on sections. Since $\widetilde{\tau d\nu}=0$ (for a form $\nu\in A\Omega^{m-1}K$ with compact support or on rapidly decreasing sections), we require that $\operatorname{Ker}\Gamma\supseteq\operatorname{Im}\tau d$.

7.4. The following result is obvious from the definitions. (We sometimes omit Γ in the notation $\{\omega_1,\omega_2\}_\Gamma$).

7.5. THEOREM. If $\Gamma:A\Omega^m K\to D_{ev}$ defines a Hamiltonian structure, then for any $\omega_i\in A\Omega^m K$ we have

$$\{\omega_1,\omega_2\}+\{\omega_2,\omega_1\}\in\operatorname{Ker}\Gamma,$$
$$\{\omega_1,\{\omega_2,\omega_3\}\}+\{\omega_3,\{\omega_1,\omega_2\}\}+\{\omega_2,\{\omega_3,\omega_1\}\}\in\operatorname{Ker}\Gamma.$$

In particular, the operation $\{\ ,\ \}_\Gamma$ induces on $A\Omega^m K/\mathrm{Ker}\,\Gamma$ the structure of a Lie algebra, and Γ induces an imbedding of the Lie algebra $A\Omega^m K/\mathrm{Ker}\,\Gamma \to D_{ev}$.

7.6. Since $\mathrm{Ker}\,\delta \subseteq \mathrm{Ker}\,\Gamma$, we can always represent Γ as a composition $A\Omega^m K \xrightarrow{\delta} \mathrm{Im}\,\delta \xrightarrow{B} D_{ev}$. Indeed, $\mathrm{Im}\,\delta \subset A\Omega^1 A_0 \Omega^m K$, and our operator B in all concrete cases will be a differential operator defined on the whole module $A\Omega^1 A_0 \Omega^m K$. We call such an operator a Hamiltonian if $\Gamma = B \circ \delta$ defines a Hamiltonian structure.

7.7. LEMMA. For any additive operator $B: \mathrm{Im}\,\delta \to D_{ev}$ we have

$$B\delta\,\{\omega_1, \omega_2\}_{B\delta} = B\delta\,(i_{B\delta\omega_1}\delta\omega_2). \tag{12}$$

Proof. For any $\bar{X} \in D_{ev}$, by formula (4) and item 5.1 we have

$$L_{\bar{X}}\omega_2 \equiv i_{\bar{X}}\delta\omega_2 \bmod \mathrm{Ker}\,\delta.$$

Setting here $\bar{X} = B\delta\omega_1$ and applying (10) with $\Gamma = B\delta$, we obtain (12).

7.8. We rewrite the Hamiltonian condition (11) of the operator $\Gamma = B\delta$ in terms of structures related to the choice of a local coordinate system (x_j, u_i).

Let us first agree on the following notation. Let $\mathcal{P}$ be an A-module. $\mathcal{P}^n$ denotes the module of column vectors of height n with elements in $\mathcal{P}$. If $\bar{P} \in \mathcal{P}^n$, then $\bar{P}^t$ denotes the transposed row vector (t will generally denote the transpose of a matrix of any size). If $\bar{Q} \in A^n$, $\bar{P} \in \mathcal{P}^n$, then $\bar{Q}^t\bar{P} \in \mathcal{P}$ denotes the scalar product $\sum_{i=1}^{n} Q_i P_i$.

The operator $\frac{\delta}{\delta \bar{u}}: A \to A^n$ takes $P \in A$ into the vector $\frac{\delta P}{\delta \bar{u}} = \left(\frac{\delta P}{\delta u_i}\right)$ (cf. 4.2). In place of $\left(\frac{\delta P}{\delta \bar{u}}\right)^t$ we write $\frac{\delta P}{\delta \bar{u}^t}$.

Further, let $A[D_c]$ be the ring of differential operators over A, generated by total differentiations along the base. For any $P \in A$ we denote by $D_{u_i}(P) \in A[D_c]$ the formal partial Fréchet derivative

$$D_{u_i}(P) = \sum_k \frac{\partial P}{\partial u_i^{(k)}}\partial^k, \quad \partial^k = \partial_1^{k_1} \dots \partial_m^{k_m}, \quad \partial_j = \overline{\partial/\partial x_j}.$$

Finally, for any vector $\bar{P} \in A^n$ we define the "Fréchet Jacobian" $D(\bar{P}) \in M_n(A[D_c])$ as the $(n \times n)$ matrix (not the determinant) with the operator $D_{u_j}(P_i)$ at the site (ij). Its invariant interpretation will become clear in the next section.

In order to write the Hamiltonian condition (11) in this notation, we identify the A-modules $A\Omega^m K$, $A\Omega^1 A_0 \Omega^m K$ and D_{ev}, respectively with A, A^n, A^n by means of the following mappings: $P dx_1 \wedge \dots \wedge dx_m \mapsto P$, $\sum P_i du_i dx_1 \wedge \dots \wedge d^m x \mapsto (P_i)$, $\bar{X} \mapsto (Xu_i)$ for $X \in AD(A_0/K)$. Then the operator $B: A\Omega^1 A_0 \Omega^m K \to D_{ev}$ will be represented by an operator $B: A^n \to A^n$, which in all examples lies in $M_n(A[D_c])$.

7.9. Proposition. The operator $B: A^n \to A^n$ is Hamiltonian if and only if for any P, $Q \in A$ we have

$$B\frac{\delta}{\delta\bar{u}}\left(\frac{\delta P}{\delta\bar{u}^t}B\frac{\delta Q}{\delta\bar{u}}\right)=D\left(B\frac{\delta P}{\delta\bar{u}}\right)B\frac{\delta Q}{\delta\bar{u}}-D\left(B\frac{\delta Q}{\delta\bar{u}}\right)B\frac{\delta P}{\delta\bar{u}}. \tag{13}$$

Proof. We shall verify that the left and right sides of (13) are identical with the left and right sides of (11), respectively, for $\Gamma=B\delta$, $\omega_1=Qd^mx$, $\omega_2=Pd^mx$.

According to (12), the left side of (11) is $B\delta(i_{B\delta\omega_1}\delta\omega_2)$. According to 4.2, the form $\delta\omega_1$ is represented by the vector $\frac{\delta Q}{\delta\bar{u}}$; the field $B\delta\omega_1$ is therefore represented by the vector $B\frac{\delta Q}{\delta\bar{u}}$. From the formula $i_{\bar{X}}(\delta\omega_2)=\sum_{i=1}^{n}\frac{\delta P}{\delta u_i}Xu_i d^mx$ it follows that the form $i_{B\delta\omega_1}\delta\omega_2$ is represented by $\frac{\delta P}{\delta\bar{u}^t}B\frac{\delta Q}{\delta\bar{u}}$. Finally, the operator $B\delta=\Gamma$ is represented by the operator $B\frac{\delta}{\delta\bar{u}}$. This reduces the left side of (11) to the left side of (13).

In order to compute the right side of (11), we note first of all that the field $B\delta\omega_2$ is represented by the vector $B\frac{\delta P}{\delta\bar{u}}$. Further, for any $R\in A$ and $\bar{X}\in D_{\text{ev}}$ we have $\bar{X}R=\sum_{i,k}\frac{\partial R}{\partial u_i^{(k)}}(Xu_i)^{(k)}=\sum_{i=1}^{n}D_{u_i}(R)Xu_i$; this explains the meaning of the Fréchet derivatives. Thus, for any vector $\bar{R}\in A^n$ we have $\bar{X}\bar{R}=D(\bar{R})X\bar{u}$. Applying this to the case $\bar{X}=B\delta\omega_1$ and $\bar{R}=B\frac{\delta P}{\delta\bar{u}}$, we write $B\delta\omega_1 B\delta\omega_2\bar{u}$ in the form $D\left(B\frac{\delta P}{\delta\bar{u}}\right)B\frac{\delta Q}{\delta\bar{u}}$. Similarly, the second term of the commutator gives the second term of the right side of (13). This completes the proof.

7.10. We shall now concern ourselves with transformation of the criterion (13). We recall first the formalism of the adjoints of differential operators. In a local chart we have $A[D_c]=A[\partial_1,\dots,\partial_n]$, $[\partial_i,\partial_j]=0$. Let $L\in M_n(A[D_c])$ (the ring of $(n\times n)$ matrix operators which act in the natural way on $\mathcal{P}^n$, where $\mathcal{P}$ is any $A[D_c]$-module). We define an additive mapping $L\mapsto L^+$, by setting for any matrix $a\in M_n(A)$: $(a\partial^k)^+=(-1)^{|k|}\partial^k\circ a^t$, where a^t is the transpose of the matrix a.

The following result is classical (and can be deduced without difficulty from 3.3 and 3.4).

7.11. LEMMA. a) For any L, M we have $(L^+)^+=L$, $(LM)^+=M^+L^+$. b) For any $\bar{P}\in A^n$, $\bar{Q}\in\mathcal{P}^n$, where $\mathcal{P}$ is an $A[D_c]$-module we have

$$\bar{P}^tL\bar{Q}-(L^+\bar{P})^t\bar{Q}\in\partial_1\mathcal{P}+\dots+\partial_n\mathcal{P}.$$

The operator L is called formally symmetric (respectively, skew-symmetric) if $L^+=L$ (respectively, $L^+=-L$).

The next result is obtained immediately from Lemma 5.8 and the definition of the Fréchet derivative.

7.12. LEMMA. For any element $P\in A$ the operator $D\left(\frac{\delta P}{\delta\bar{u}}\right)$ is symmetric.

It is moreover clear from the considerations of 5.7 that the part of the variational complex $A\Omega^mK\xrightarrow{\delta}A\Omega^1A_0\Omega^mK\xrightarrow{\delta^{(1)}\tau^{(1)}}A^{(1)}\Omega^1A_0^{(1)}\Omega^{m+1}K^{(1)}$ with the identifications of 7.8 may be replaced by

$$A \xrightarrow{\delta/\delta\bar{u}} A^n \xrightarrow{D(\cdot)-D^+(\cdot)} M_n(A[D_c]),$$

and the theorem of B. A. Kupershmidt on the exactness of this complex means that $D^+(\overline{Q})=D(\overline{Q})$, if and only if there exists a $P\in A$, such that $\overline{Q}=\frac{\delta P}{\delta \bar{u}}$. There is thus an effective criterion for determining if a vector $\overline{Q}\in A^n$ is a vector of partial derivatives of some element of A.

Suppose now that $B\in M_n(A[D_c])$ is some differential operator.

7.13. THEOREM. a) If B is skew-symmetric and $B\in M_n(K[D_c])$, then B is Hamiltonian. b) If B is skew-symmetric and $B\in M_n(A_0[D_c])$, then in order that B be Hamiltonian it is necessary and sufficient that for all $P, Q\in A$ the following identity hold:

$$[D,B]\frac{\delta P}{\delta \bar{u}}B\frac{\delta Q}{\delta \bar{u}}-[D,B]\frac{\delta Q}{\delta \bar{u}}B\frac{\delta P}{\delta \bar{u}}=B\left(\frac{\delta P}{\delta \bar{u}^t}\frac{\partial B}{\partial \bar{u}}\frac{\delta Q}{\delta \bar{u}}\right). \tag{14}$$

We begin with the following lemma.

7.14. LEMMA. Suppose that $B\in M_n(A_0[D_c])$ is skew-symmetric. Then

$$\frac{\delta}{\delta \bar{u}}\left(\frac{\delta P}{\delta \bar{u}^t}B\frac{\delta Q}{\delta \bar{u}}\right)=D\left(\frac{\delta P}{\delta \bar{u}}\right)B\frac{\delta Q}{\delta \bar{u}}-D\left(\frac{\delta Q}{\delta \bar{u}}\right)B\frac{\delta P}{\delta \bar{u}}+\frac{\delta P}{\delta \bar{u}^t}\frac{\partial B}{\partial \bar{u}}\frac{\delta Q}{\delta \bar{u}}, \tag{15}$$

where the last term is a column vector with coordinate $\frac{\delta P}{\delta \bar{u}}\frac{\partial B}{\partial u_i}\frac{\delta Q}{\delta \bar{u}}$ at the i-th row, and $\frac{\partial B}{\partial u_i}$ is the result of applying $\frac{\partial}{\partial u_i}$ to the coefficients of the operator B.

Proof. By the definitions

$$\left[\frac{\delta}{\delta \bar{u}}\left(\frac{\delta P}{\delta \bar{u}^t}B\frac{\delta Q}{\delta \bar{u}}\right)\right]^t d\bar{u}\wedge d^m x=\hat{\delta} d\left(\frac{\delta P}{\delta \bar{u}^t}B\frac{\delta Q}{\delta \bar{u}}d^m x\right). \tag{16}$$

Further,

$$d\left(\frac{\delta P}{\delta \bar{u}^t}B\frac{\delta Q}{\delta \bar{u}}d^m x\right)=\left[d\,\frac{\delta P}{\delta \bar{u}^t}B\frac{\delta Q}{\delta \bar{u}}+\frac{\delta P}{\delta \bar{u}^t}Bd\frac{\delta Q}{\delta \bar{u}}+\frac{\delta P}{\delta \bar{u}^t}dB\frac{\delta Q}{\delta \bar{u}}\right]d^m x,$$

where $dB=\sum_{i=1}^{m} du_i\frac{\partial B}{\partial u_i}$ and ΩA is considered as an A-bimodule with multiplication $\omega P=P\omega$ for any $P\in A$, $\omega\in\overline{\Omega}A$.

Therefore, for any $\overline{X}\in D_{ev}$, we have

$$i_{\overline{X}}d\left(\frac{\delta P}{\delta \bar{u}^t}B\frac{\delta Q}{\delta \bar{u}}d^m x\right)=\left[\left(\overline{X}\frac{\delta P}{\delta \bar{u}^t}B\frac{\delta Q}{\delta \bar{u}}+\frac{\delta P}{\delta \bar{u}^t}B\left(\overline{X}\frac{\delta Q}{\delta \bar{u}}\right)+\frac{\delta P}{\delta \bar{u}}(\overline{X}B)\frac{\delta Q}{\delta \bar{u}}\right]d^m x. \tag{17}$$

Using the formula $\overline{XR}=D(\overline{R})X\bar{u}$, we rewrite the right side of (17) in the form

$$\left[\left(B\frac{\delta Q}{\delta \bar{u}}\right)^t D\left(\frac{\delta P}{\delta \bar{u}}\right)X\bar{u}+\frac{\delta P}{\delta \bar{u}^t}BD\left(\frac{\delta Q}{\delta \bar{u}}\right)X\bar{u}+\left(\frac{\delta P}{\delta \bar{u}^t}\frac{\partial B}{\partial \bar{u}}\frac{\delta Q}{\delta \bar{u}}\right)^t X\bar{u}\right]. \tag{18}$$

We now note that for any $R\in A$ we have $\partial_j R d^m x=(-1)^{j-1}\tau d(Rd_j^m x)$. Therefore, transposing the operators in (18) while replacing them by their adjoints, as in Lemma 7.11 a), we do not change the value of the right side modulo $\operatorname{Im} \tau d$.

In the first term of (18) we transposed the operator $D\left(\frac{\delta P}{\delta \bar{u}}\right)$, and in the second term the operator $BD\left(\frac{\delta Q}{\delta \bar{u}}\right)$. Recalling that B is skew-symmetric while $D\left(\frac{\delta P}{\delta \bar{u}}\right)$ and $D\left(\frac{\delta Q}{\delta \bar{u}}\right)$ are symmetric, we obtain finally

$$i_{\bar{X}} d\left(\frac{\delta P}{\delta \bar{u}^t} B \frac{\delta Q}{\delta \bar{u}} d^m x\right) \equiv i_{\bar{X}}\left[\left(D\left(\frac{\delta P}{\delta \bar{u}}\right) B \frac{\delta Q}{\delta \bar{u}}\right)^t d\bar{u} \wedge d^m x - \left(D\left(\frac{\delta Q}{\delta \bar{u}}\right) B \frac{\delta P}{\delta \bar{u}}\right)^t d\bar{u} \wedge d^m x + \left(\frac{\delta P}{\delta \bar{u}^t} \frac{\partial B}{\partial \bar{u}} \frac{\delta Q}{\delta \bar{u}}\right)^t d\bar{u} \wedge d^m x\right]. \quad (19)$$

From the uniqueness of the Euler–Lagrange operator it follows that the form following $i_{\bar{X}}$ on the right side of (19) is $\delta d\left(\frac{\delta P}{\delta \bar{u}^t} B \frac{\delta Q}{\delta \bar{u}} d^m x\right)$. This and (16) give the assertion of the lemma.

<u>7.15. Proof of Theorem 7.13.</u> By formula (15) the left side of (13) is equal to

$$BD\left(\frac{\delta P}{\delta \bar{u}}\right) B \frac{\delta Q}{\delta \bar{u}} - BD\left(\frac{\delta Q}{\delta \bar{u}}\right) B \frac{\delta P}{\delta \bar{u}} + B\left(\frac{\delta P}{\delta \bar{u}^t} \frac{\partial B}{\partial \bar{u}} \frac{\delta Q}{\delta \bar{u}}\right).$$

If $B \in M_n(K[D_c])$, then $\frac{\partial B}{\partial \bar{u}} = 0$. Moreover, in this case $BD\left(\frac{\delta P}{\delta \bar{u}}\right) = D\left(B \frac{\delta P}{\delta \bar{u}}\right)$, since when applied to the vector $X\bar{u} \in A^n$, $\bar{X} \in D_{ev}$, they give $B\bar{X} \frac{\delta P}{\delta \bar{u}}$ and $\bar{X} B \frac{\delta P}{\delta \bar{u}}$, respectively, and these expressions coincide, since $K[D_c]$ commutes with D_{ev}. Thus, the criterion (13) is satisfied in this case.

Using the information obtained, for $B \in A_0[D_c]$ formula (14) obviously coincides with criterion (13).

In Sec. 8 we shall apply the criterion (14) to prove that two special operators B over a one-dimensional base are Hamiltonian: the Gel'fand–Dikii operator and the Benney operator.

<u>7.16. Hamiltonian Conservation Laws.</u> Suppose Γ defines a Hamiltonian structure on $\pi: N \to M$. For any form $\omega = Q dx_1 \wedge \ldots \wedge dx_m \in A\Omega^m K$ to the field $\Gamma(\omega)$ there corresponds the system of evolution equations $\bar{u}_t = \Gamma(\omega) \bar{u} = B \frac{\delta Q}{\delta \bar{u}}$ in the notation of the preceding sections. We call ω or Q its Hamiltonian and B the corresponding Hamiltonian operator.

The form $\omega_1 = P dx_1 \wedge \ldots \wedge dx_m$, or the coefficient P, is called an integral or a conservation law for this system if $\{\omega, \omega_1\}_\Gamma \in \operatorname{Im} \tau d = \operatorname{Ker} \delta$. In the corresponding analytic formulation this means that for an evolution s, because of our system,

$$\frac{d}{dt} \int_M \omega_1^s = \int_M (L_{\Gamma(\omega)} \omega_1)^s = \int_M \{\omega, \omega_1\}^s = 0,$$

i.e., $\int_M \omega_1^s$ is a quantity conserved in time. In the notation of 7.8–7.9 P is an integral for the Hamiltonian Q, if

$$\frac{\delta}{\delta \bar{u}}\left(\frac{\delta P}{\delta \bar{u}^t} B \frac{\delta Q}{\delta \bar{u}}\right) = 0.$$

The integrals P_1 and P_2 are said to commute if $\{\omega_1, \omega_2\}_\Gamma \in \mathrm{Ker}\,\delta$, i.e.,

$$\frac{\delta}{\delta \bar{u}}\left(\frac{\delta P_1}{\delta \bar{u}^t} B \frac{\delta P_2}{\delta \bar{u}}\right)=0.$$

We note in conclusion that if $\{\omega, \ldots, \{\omega, \omega_1\} \ldots\} \in \mathrm{Ker}\,\delta$ ($k \geqslant 2$ factors ω), then on those sections s, which are uniquely included in the flow with Hamiltonian ω, it is possible to find $k-1$ quantities conserved in time as follows. We have $\left(\frac{d}{dt}\right)^k \int_M \omega_1^s = \int_M (L^k_{\Gamma(\omega)}\omega_1)^s = \int_M \{\omega, \ldots, \{\omega, \omega_1\} \ldots\} = 0$. Therefore $\int_M \omega_1^s$ is a polynomial in t of degree $k-1$ in general. Normalizing the initial time on a given trajectory of the flow so that the $(k-2)$-th coefficient of the polynomial is zero, we find that the remaining $k-1$ coefficients are invariantly defined, conserved quantities.

7.17. Stationary Manifolds of Integrals. Let ω_1 be a conservation law for a Hamiltonian system of evolution corresponding to a field $\Gamma(\omega)$. By definition $L_{\Gamma(\omega)}\omega_1 \in \mathrm{Im}\,\tau d$ or $i_{\Gamma(\omega)}\delta\omega_1 \in \mathrm{Im}\,\tau d$. The latter condition means that if s is an extremal of the Lagrangian ω_1, then within a short time in the linear approximation the evolution of s, under the field $\Gamma(\omega)$, will also be an extremal or that the "field $\Gamma(\omega)$ is tangent to the manifolds of extremals ω_1." It is possible to give a precise meaning to the last assertion and to prove it at least for $m=1$ and polynomial integrals ω_1, when the manifold of extremals W can be identified with a finite algebraic manifold by assigning to each extremal its value at $x=0$. Then $\Gamma(\omega)$ defines on it a flow which is also Hamiltonian with Hamiltonian computed on the basis of ω and ω_1. This assertion is nontrival because the natural class of functions on W is not obtained by restricting functionals W to $\tilde{\gamma}$; it consists, for example, of smooth functions on W in the natural structure of W. For a precise formulation and proof see the papers of Bogoyavlenskii and Novikov [1] and also Gel'fand and Dikii [3].

We shall make use of this remark in Chap. 3 where in place of the equations $\bar{u}_t=\Gamma(\omega)\bar{u}$ we will solve jointly the system $\bar{u}_t=\Gamma(\omega)\bar{u}$ and $\delta\omega_1=0$, where ω_1 is an integral. It is just this procedure which distinguishes in an invariant way the class of multisolution and finite-zone solutions of the Korteweg–de Vries equation as already mentioned in the introduction.

7.18. Examples. a) Let $n=2r$, and $B=\begin{pmatrix} 0 & -E \\ E & 0 \end{pmatrix}$, where E is the identity matrix of order r. The system of evolution $\bar{u}_t=B\frac{\delta P}{\delta \bar{u}}$ is traditionally called a Hamiltonian system with Hamiltonian P in "canonical coordinates". It is also Hamiltonian in our sense according to Theorem 7.13 a), since B is a differential operator of order zero and $B^+=-B$. Below we shall show how it is possible to define a "cotangent fibration" $\bar{\pi}: \bar{N} \to M$ to any fibration $\pi: N \to M$, a canonical Hamiltonian structure on it, and to each local coordinate system (x_i, u_j) on N a system of coordinates (x_i, u_j, v_j) on $\bar{N}$, such that in these coordinates the operator B corresponding to this structure has the form $\begin{pmatrix} 0 & E \\ -E & 0 \end{pmatrix}$.

b) Let $m=1$ and $B=\partial=\partial_1$. That equations of the form $u_t=\partial\frac{\delta P}{\delta u}$, which include the usual and higher Korteweg–de Vries equations, are Hamiltonian was established by Lax in the functional version and in terms of Fourier coefficients by Gardner. Our proof of Theorem 7.13 is a generalization of the proof of Lax as reworked by M. A. Shubin. A formalized version of Gardner's arguments will be given at the end of this section.

7.19. The Cotangent Fibration. The cotangent fibration to $\pi:N\to M$ we call the fibration $\bar\pi:\bar N\to M$, where $\bar N=T^*(N/M)\otimes\Lambda^m\pi^*T^*(M)$. Here $T(N/M)$ is the tangent bundle to N along the fiber π, $T(M)$ is the tangent bundle to M, and the asterisk denotes dualization; the tensor product is taken over N; $\bar\pi$ is the composition $\bar N\xrightarrow{\sigma}N\xrightarrow{\pi}M$, where $\sigma:\bar N\to N$ is the natural projection.

Sections of σ are naturally identified (locally) with forms in $\Omega^1A_0\Omega^mK$.

Let (x_i, u_j) be coordinates on N. On the basis of these it is natural to define coordinates (x_i, u_j, v_j) on $\bar N$ $(j=1,\dots,n)$ as follows. If $x\in\bar N$ lies on a section of σ, corresponding to the form $\sum_{j=1}^{m}P_j du_j\wedge d^m x$, then $v_j(x)=P_j(\sigma(x))$.

Let $\bar A_0$ be the ring of smooth functions on $\bar\pi^{-1}(U)$, where U is a neighborhood with coordinates (x_i). Then $\bar A_0=C^\infty(x_i; u_j, v_j)$, and the natural imbeddings $K\subset A_0\subset\bar A_0$ correspond to the projections π and σ.

We consider $\rho=\sum v_j du_j\wedge d^m x\in\bar A_0\Omega^1(A_0)\Omega^mK$. It possesses the following property: if the section s of the fibration σ corresponds to the form $\omega\in\Omega^1(A_0)\Omega^mK$, then $\rho^s=\omega$. It is easy to see that this property determines ρ uniquely. Thus, ρ is defined canonically and globally.

Let $\bar D_{ev}$ be the evolution differentiations for $\bar\pi$. If $\bar X\in\bar D_{ev}$, then $i_{\bar X}d\rho\in\Omega^1\bar A_0\Omega^mK$ and the formula

$$i_{\bar X}d\rho=\sum_{j=1}^{m}(Xv_j du_j-Xu_j dv_j)\,d_x^m$$

shows that this mapping $\bar D_{ev}\to\Omega^1\bar A_0\Omega^nK$ is an isomorphism of $\bar A_0$-modules. We denote by B the inverse operator, $B:\Omega^1\bar A_0\Omega^mK\to\bar D_{ev}$, and we set

$$\Gamma=-B\delta:\bar A_0\Omega^mK\to\bar D_{ev}.$$

7.20. THEOREM. In the coordinates (x_i, u_j, v_j) on $\bar N$ with the conventions of 7.8 (as applied to $\bar\pi$) the operator $-B$ is represented by the matrix $\begin{pmatrix}0 & -E\\ E & 0\end{pmatrix}$, the operator Γ is Hamiltonian, and the system of evolution corresponding to the Hamiltonian $Pd^m x\in\bar A_0\Omega^mK$ has the form

$$\bar u_t=\frac{\delta P}{\delta\bar v},\quad \bar v_t=-\frac{\delta P}{\delta\bar u}.$$

All this follows readily from the definitions and Theorem 7.13 a). We note further that the classical expression $\left(\frac{\delta P}{\delta v_j}\frac{\delta Q}{\delta u_j}-\frac{\delta P}{\delta u_j}\frac{\delta Q}{\delta v_j}\right)d^m x$ is equal to $i_{B\delta\omega_1}B\delta\omega_2$ where $\omega_1=Pd^m x$, $\omega_2=Qd^m x$.

7.21. The Hamiltonian Property in Formal Fourier Coefficients. In this and the following sections we briefly describe the formalism of variational calculus in terms of Fourier coefficients in the simplest case: M is the unit circle, $N=M\times\mathbf{R}$. We restrict ourselves to Hamiltonians which are polynomials in u and do not depend explicitly on x , i.e., we start from the algebra $A=C[u,u',\ldots]$. We shall first introduce formal Fourier coefficients $(v_n\,|\,n\in\mathbf{Z})$, assuming that the Fourier series of u is $F(u)=\sum_{n=-\infty}^{\infty} v_n e^{2\pi inx}$. As the Fourier series for $u^{(j)}$ it is then natural to take $F(u^{(j)})=\sum_n (2\pi in)^j v_n e^{2\pi inx}$, i.e., $F(u^{(j)})=(2\pi in)^j v_n$. The Fourier coefficients of the polynomials in $u^{(j)}$ will, however, be special series of infinitely many variables v_n . In order to introduce the corresponding ring, we first set $B=C[v_n]$. In this ring we introduce the order function $\operatorname{ord}\left(v_{n_1}^{a_1}\ldots v_{n_k}^{a_k}\right)=|n_1|+\ldots+|n_k|$ and $\operatorname{ord}\left(\sum c_{\overline{n}}^{\overline{a}} v_{\overline{n}}^{\overline{a}}\right)=\min\left(\operatorname{ord} v_{\overline{n}}^{\overline{a}}\,\middle|\,c_{\overline{n}}^{\overline{a}}\neq 0\right)$. If B_N is the space of polynomials purely of order N, then obviously $B=\bigoplus_{N=0}^{\infty} B_N$ (direct sum). We set $\hat{B}=\prod_{N=0}^{\infty} B_N$. The elements of the space $\hat{B}$ are infinite series $\sum_{N=0}^{\infty} f_N$, $f_N\in B_N$. Since $\operatorname{ord}(fg)=\operatorname{ord} f+\operatorname{ord} g$ for $f\in B_M$, $g\in B_N$, $\hat{B}$ has a natural ring structure: $\left(\sum f_N\right)\left(\sum g_M\right)=\sum h_K$, $h_K=\sum f_{K-M}g_M$.

We call the ring

$$\hat{B}((e^{2\pi ix}))=\left\{\sum_{n\in\mathbf{Z}} f_n e^{2\pi inx}\,\middle|\, f_n\in\hat{B},\ \exists c>0,\ n_0\geqslant 0,\ \forall\,|n|\geqslant n_0,\ \operatorname{ord} f_n\geqslant c\,|n|\right\}$$

the ring of formal Fourier series; it has the natural multiplication $\left(\sum f_m e^{2\pi imx}\right)\left(\sum g_n e^{2\pi inx}\right)=\sum h_p e^{2\pi ipx}$, where $h_p=\sum_{m+n=p} f_m g_n$. The convergence of the series for h_p in $\hat{B}$ is ensured by the fact that $\operatorname{ord}(f_m g_n)\geqslant c\,(|m|+|n|)\to\infty$ together with $|m|,|n|$. The product lies in $\hat{B}((e^{2\pi ix}))$, because $\operatorname{ord} h_p\geqslant \min_{m+n=p} c\,(|m|+|n|)\geqslant c\,|p|$.

We extend to $\hat{B}((e^{2\pi ix}))$ the differentiation $\frac{\partial}{\partial v_n}$ by continuity: we extend from B to $\hat{B}$ and then set $\frac{\partial}{\partial v_n}(e^{2\pi ikx})=0$. This is clearly possible, since $\operatorname{ord}\left(\frac{\partial f}{\partial v_n}\right)\geqslant\operatorname{ord} f-n$. We similarly define a differentiation ∂, which on B has the form $\partial v_n=2\pi iv_n$.

We introduce a ring homomorphism $F:A\to\hat{B}((e^{2\pi ix}))$, by defining $F(u)=\sum v_n e^{2\pi inx}$, $F(u^{(j)})=\partial^j F(u)$. It is not hard to see that F is an imbedding which commutes with ∂. Let $F(P)=\sum_m F_m(P)e^{2\pi imx}$ for any $P\in A$.

7.22. LEMMA. For any $P\in A$ we have $F_n\left(\frac{\delta P}{\delta u}\right)=\frac{\partial}{\partial v_{-n}}F_0(P)$.

Proof. We consider the two mappings of $A \to \hat{B}((e^{2\pi ix}))$:

$$P \mapsto \frac{\partial}{\partial v_n} F(P), \quad P \mapsto \sum_{j=0}^{\infty} \frac{\partial}{\partial v_n} F(u^{(j)}) F\left(\frac{\partial P}{\partial u^{(j)}}\right).$$

They are both differentiations of A into the A-module $\hat{B}((e^{2\pi ix}))$ (relative to F). Further, on the generators $u^{(j)}$ they coincide, taking $u^{(j)}$ into $(2\pi in)^j e^{2\pi inx}$. They therefore coincide everywhere, whence

$$\frac{\partial}{\partial v_n} F_0(P) = F_0\left(\frac{\partial}{\partial v_n} F(P)\right) = F_0\left(\sum_{j=0}^{\infty} \frac{\partial}{\partial v_n} F(u^{(j)}) F\left(\frac{\partial P}{\partial u^{(j)}}\right)\right) = F_0\left(\sum_{j=0}^{\infty} \partial^j \frac{\partial}{\partial v_n} F(u) F\left(\frac{\partial P}{\partial u^{(j)}}\right)\right).$$

We note now that $F_0 \circ \partial = 0$, so that we may "integrate by parts" under the sign of F_0, and the last expression is equal to

$$F_0\left(\sum_{j=0}^{\infty} e^{2\pi inx}(-1)^j \partial^j F\left(\frac{\partial P}{\partial u^{(j)}}\right)\right) = F_{-n}\left(\frac{\delta P}{\delta u^{(j)}}\right),$$

which proves the lemma.

7.23. Corollary. The equation $u_t = \partial \frac{\delta P}{\delta u}$ in terms of Fourier coefficients has the form

$$(v_0)_t = 0, \quad (v_n)_t = 2\pi in \frac{\partial}{\partial v_{-n}} F_0(P), \; |n| \geqslant 1.$$

Thus, setting $p_n = v_{-n}$, $q_n = \frac{v_n}{2\pi n}$, $H(p, q) = iF_0(P)$, we find that on the "hyperplanes" $v_0 = \text{const}$ our equation formally has the Hamiltonian form

$$\frac{\partial p_n}{\partial t} = -\frac{\partial H}{\partial q_n}, \quad \frac{\partial q_n}{\partial t} = \frac{\partial H}{\partial p_n}.$$

Further, the Poisson bracket may be written

$$F\left(\frac{\delta Q}{\delta u} \partial \frac{\delta P}{\delta u}\right) = \left(\sum_n \frac{\partial}{\partial v_{-n}} F_0(Q) e^{2\pi inx}\right) \times \left(\sum_m \frac{\partial}{\partial v_{-m}} F_0(P) 2\pi ime^{2\pi imx}\right) = \sum_p \left(\sum_{m+n=p} 2\pi im \frac{\partial}{\partial v_{-n}} F_0(Q) \frac{\partial}{\partial v_{-m}} F_0(P)\right) e^{2\pi ipx},$$

whence

$$F_0\{Q, P\} = \sum_{n=-\infty}^{\infty} (-2\pi in) \frac{\partial}{\partial v_{-n}} F_0(Q) \frac{\partial}{\partial v_n} F_0(P) =$$

$$= 2\pi i \sum_{n=1}^{\infty} \left(\frac{\partial}{\partial v_n} F_0(P) n \frac{\partial}{\partial v_{-n}} F_0(Q) - \frac{\partial}{\partial v_{-n}} F_0(P) n \frac{\partial}{\partial v_n} F_0(Q)\right) = i \sum_{n=1}^{\infty} \left(\frac{\partial F_0(P)}{\partial q_n} \frac{\partial F_0(Q)}{\partial p_n} - \frac{\partial F_0(P)}{\partial p_n} \frac{\partial F_0(Q)}{\partial q_n}\right)$$

in agreement with the classical formalism.

8. Special Hamiltonian Operators Over a One-Dimensional Base

8.1. In this section we set $m = 1$, $x_1 = x$, $\partial = \partial_1 = \overline{\partial/\partial x}$, $K = C^\infty(x)$, and we work with the sequence of rings $A_n = \bigcup_{i=0}^{\infty} C^\infty(x;\, u_j^{(k)} \,|\, 0 \leqslant j \leqslant n,\, 0 \leqslant k \leqslant i)$, and also with their union $A = \bigcup_{n=0}^{\infty} A_n$. We point out that the enumeration of the variables u_j begins with zero in contrast to the preceding conventions. In correspondence with this the enumeration of the elements in the rows and columns and also the enumeration of rows and columns in different matrices also begins with zero.

We define the operator $B_n \in M_{n+1}(A_n[\partial])$ by the formulas

$$B_n=\sum_{i=1}^{\infty}[B_{n,i}\partial^i-(-1)^i\partial^i\circ B_{n,i}^t],\quad B_{n,i}\in M_{n+1}(A_n),$$
$$(B_{n,i})_{ab}=\begin{cases}\binom{i+a}{a}u_{a+b+i+1} & \text{for } a+b+i+1\leqslant n,\\ 0 & \text{for } a+b+i+1>n,\ a+b+i+1\neq n+2,\\ \binom{i+a}{a} & \text{for } a+b+i+1=n+2.\end{cases} \tag{20}$$

These operators were introduced by Gel'fand and Dikii [4]. Following their work, we shall show in the next chapter that the Lax equations $L_t=[P,L]$, where $L=\partial^{n+2}+\sum_{i=0}^{n}u_i\partial^i$ are Hamiltonian in the structure defined by the operator B_n, with a suitable Hamiltonian depending on P.

We further introduce the operator $\bar{B}\in M_\infty(A[\partial])$ by the formulas

$$\bar{B}=B\partial+\partial\circ B^t,\quad B\in M_\infty(A),$$
$$(B)_{ab}=\begin{cases}au_{a+b-1} & \text{for } a+b\geqslant 1,\\ 0 & \text{for } a+b=0.\end{cases} \tag{21}$$

Here $M_\infty(A)$ denotes the group of matrices with an infinite number of rows and columns downward and to the right beginning with the zero (index) columns and rows. The elements of $M_\infty(A[\partial])$ are considered as left operators on the A-module A_f^∞ of finitely supported infinite columns with elements in A (finitely supported means that only a finite number of the coordinates are nonzero). The range of such an operator lies in the A-module A^∞ of all infinite columns. The scalar product $\bar{P}^t\bar{Q}$ is defined if at least one of the two vectors $\bar{P},\bar{Q}$ is finitely supported. All vectors $\frac{\delta P}{\delta \bar{u}}$ for $P\in A$ are finitely supported, and therefore the $\bar{B}\frac{\delta Q}{\delta \bar{u}}$ are defined but are not necessarily finitely supported. Each row of the Fréchet Jacobian $D(\bar{P})$ is finitely supported for any $\bar{P}\in A^\infty$, since $D_{u_j}(P_i)=0$ for sufficiently large j (depending on i; see the definition in 7.8). Therefore, $D(\bar{P})$ may be applied to any vector of A^∞. The left and right sides of equalities (13) and (14) of Sec. 7 are thus defined, and we take them as the definition that B be Hamiltonian in the ring A of functions on the space of jets of the corresponding "infinite-dimensional" fibration (the projective limit of the finite-dimensional fibrations). It is not hard to verify that the remaining constructs of Sec. 7 with appropriate modifications carry over to A.

The operator (21) was introduced in the work of Kupershmidt and the author [19] to study the system of equations for long waves with a free surface suggested by Benney. The infinite sequence of unknown functions $u_n,\ n\geqslant 0$, is, in fact, the sequence of moments of the horizontal component of the velocity, and the system of Benney's evolution equations for it is found to be Hamiltonian with operator $\bar{B}$ and a corresponding Hamiltonian. Details may be found in the next chapter.

The following theorem is the main result of this section.

8.2. THEOREM. The operators B_n and $\bar{B}$ are Hamiltonian. The proof follows by means of lengthy computations. Among the reasons for this is probably the fact that we do not know an invariant (coordinate-free) characterization of these operators.

We note that according to formula (20) there is also a formal "limit" B_∞ of the operators B_n, which is a matrix operator of infinite order. The operator B_n is obtained from B_∞, by setting $u_{n+2}=1$, $u_j=0$ for $j=n+1, j>n+2$. Although we are not in a position to assign B_∞ a substantial meaning, in the computations to verify the Hamiltonian property it is possible to work directly with B_∞, since the identities of interest to us will obviously be preserved under the substitution $u_j=0$ for $j>n$, $j\neq n+2$, $u_{n+2}=1$.

These identities are obtained in the following manner. In place of the columns $\frac{\delta P}{\delta \bar{u}}$ and $\frac{\delta Q}{\delta \bar{u}}$ we consider the columns $X=(X_0, X_1, \ldots)^t$ and $Y=(Y_0, Y_1, \ldots)^t$ with formal variables as coordinates; the action of ∂^j on them is interpreted as the conversion of X_i, Y_k into the formal variables $\partial^j X_i = X_i^{(j)}$, $\partial^j Y_k = Y_k^{(j)}$, independently of one another and of $u_i^{(l)}$. Each element of the left and right sides of (14) will then be a formal infinite sum of monomials in u, X, Y and their derivatives which is trilinear in u, X, Y; we simply verify that the coefficients of each such monomial at corresponding places on the left and right coincide. (Here X and Y are not to be confused with the previous notation for fields on jets!)

8.3. For the reformulation of the Hamiltonian criterion (14) which we shall use below, we introduce some further notation. Let $P, Q \in A$. Since D_{u_j} is a differentiation, we have $D_{u_j}(PQ)=PD_{u_j}(Q)+QD_{u_j}(P)$, or in vector form $D_{\bar{u}^t}(PQ)=PD_{\bar{u}^t}(Q)+QD_{\bar{u}^t}(P)$, where $D_{\bar{u}^t}=(D_{u_0}, D_{u_1}, \ldots)$. Now let $\bar{P}$, $\bar{Q}$ be columns. From the last equality it then follows immediately that $D_{\bar{u}^t}(\bar{P}^t\bar{Q})=\bar{P}^t D(\bar{Q})+\bar{Q}^t D(\bar{P})$. Finally, let C be a matrix and Z a column. Then CZ is a column, and

$$D(CZ)=CD(Z)+Z^t D(\tilde{C}), \tag{22}$$

where the expression $Z^t D(\tilde{C})$ is to be interpreted as follows: $\tilde{C}$ is a transfinite object — an infinite column which contains at the i-th place $\tilde{C}_i=(i\text{-th place } C)^t$, $D(\tilde{C})$ is an infinite column containing at the i-th place the matrix $D(\tilde{C}_i)$, and $Z^t D(\tilde{C})$ is an infinite matrix with i-th row $\sum_{k\geq 0} Z_k \cdot$ (k-th row of $D(\tilde{C}_i)) = \sum_{k\geq 0} Z_k D_{\bar{u}^t}(C_{ik})$ and hence with ij-th element the operator $\sum_{k\geq 0} Z_k D_{u_j}(C_{ik})$.

Since d commutes with ∂, the identity $d(Pdx)=\sum_i D_{u_i}(P)\,du_i\wedge dx$ implies that D_{u_i} and D commute with ∂. From (22) we therefore find that

$$D(C\partial^l Z)=CD(\partial^l Z)+(\partial^l Z)^t D(\tilde{C})=C\partial^l\circ D(Z)+Z^{(l)t}D(\tilde{C}).$$
$$D(\partial^l(CZ))=\partial^l\circ D(CZ)=\partial^l\circ CD(Z)+\partial^l\circ Z^t D(\tilde{C}),$$

where $Z^{(l)}=\partial^l Z$ and the circle following ∂^l on the right is inserted in order that, for example, the notation $\partial^l\circ CD(Z)$ not be interpreted as $(\partial^l C)D(Z)$.

In this notation the criterion (14) after obvious calculations may be rewritten in the following manner.

8.4. Proposition. The Hamiltonian property of the skew-symmetric operator $\sum_{j\geq 0}(B_j\partial^j - (-1)^j\partial^j\circ B_j^t)$ is equivalent to the following identities between the infinite columns:

$$B\left(\sum_{j\geq 0} X^t\frac{\partial B_j}{\partial \bar{u}}Y^{(j)} - X^{t(j)}\frac{\partial B_j^t}{\partial \bar{u}}Y\right) = \left(\sum_{j\geq 0} X^{(j)t}\frac{\partial \tilde{B}_j}{\partial \bar{u}^t} - (-1)^j\partial^j\circ X^t\frac{\partial \tilde{B}^t}{\partial \bar{u}^t}\right)BY - \text{(the same with } X\leftrightarrow Y). \quad (23)$$

Here the right and left sides of (23) correspond to the left and right sides of (14), respectively, and the following additional notation is used: $\partial\tilde{B}_j/\partial\bar{u}^t$ is the column of infinite matrices containing at the i-th place the matrix $\frac{\partial}{\partial \bar{u}}(i\text{-th row of } B_j^t)$; $X^{(j)t}\frac{\partial \tilde{B}_j}{\partial \bar{u}^t}$ is the matrix with (ik)-element $\sum_{l>0} X_l\frac{\partial}{\partial u_k}(B_{j,il})$; $\tilde{B}_j^t = (B_j^t)^{\sim}$.

8.5. We begin with the verification that the operator (21) is Hamiltonian. The identity (23) to be verified assumes the form ($X' = \partial X$, etc.)

$$(B\partial + \partial\circ B^t)\left(X^t\frac{\partial B}{\partial \bar{u}}Y' - X^{t'}\frac{\partial B^t}{\partial \bar{u}}Y\right) = \left(X^{t'}\frac{\partial \tilde{B}}{\partial \bar{u}^t} + \partial\circ X^t\frac{\partial \tilde{B}^t}{\partial \bar{u}^t}\right)(B\partial + \partial\circ B^t)Y - \text{(the same with } X\leftrightarrow Y). \quad (24)$$

We expand here all the brackets. On both sides there are sums of monomials each of which contains $X^{(i)}$ and $Y^{(j)}$ (possibly transposed). The ordered pair ij we call the type of the corresponding monomial. By collecting all terms of the same type, we bring (24) to the form

$$B^{t'}X^t\frac{\partial B}{\partial \bar{u}}Y' + (B+B^t)X^t\frac{\partial B}{\partial \bar{u}}Y'' - B^{t'}X^{t'}\frac{\partial B^t}{\partial \bar{u}}Y + (B+B^t)X^{t'}\frac{\partial B}{\partial \bar{u}}Y' - (B+B^t)X^{t'}\frac{\partial B^t}{\partial \bar{u}}Y' -$$
$$-(B+B^t)X^{t''}\frac{\partial B^t}{\partial \bar{u}}Y = X^t\frac{\partial \tilde{B}^t}{\partial \bar{u}^t}B^{t''}Y + X^t\frac{\partial \tilde{B}^t}{\partial \bar{u}^t}(B' + 2B^{t'})Y' + \quad (25)$$
$$+X^t\frac{\partial \tilde{B}^t}{\partial \bar{u}^t}(B+B^t)Y'' + X^{t'}\frac{\partial(\tilde{B}+\tilde{B}^t)}{\partial \bar{u}^t}B^{t'}Y + X^{t'}\frac{\partial(\tilde{B}+\tilde{B}^t)}{\partial \bar{u}^t}(B+B^t)Y' - \text{(the same with } X\leftrightarrow Y).$$

We shall prove that this is an identity by showing that the terms of each type on the left and right cancel. All calculations making use of the concrete form of the matrix B we reduce to the following lemma.

8.6. LEMMA. For any $i\geq 0$, and any X and Y the following assertions hold:

a) The expression $X^t\frac{\partial \tilde{B}^t}{\partial \bar{u}^t}B^{t(i)}Y$ is symmetric in X and Y .

b) $(B+B^t)^{(i)}X^t\frac{\partial B}{\partial \bar{u}}Y = X^t\frac{\partial \tilde{B}^t}{\partial \bar{u}^t}(B+B^t)^{(i)}Y$.

c) $B^{(i)}X^t\frac{\partial B}{\partial \bar{u}}Y = Y^t\frac{\partial \tilde{B}}{\partial \bar{u}^t}B^{t(i)}X$.

d) The expression $X^t\frac{\partial \tilde{B}}{\partial \bar{u}^t}(B+B^t)^{(i)}Y$ is symmetric in X and Y .

Proof. a) and d). The matrix $X^t\frac{\partial \tilde{B}^t}{\partial \bar{u}^t}$ has at the ad-th place the element $\sum_{b>0} X_b \frac{\partial}{\partial u_d}(B_{ab}^t) = \sum_{b>0} X_b b\delta_{d,a+b-1}$, where δ is the Kronecker symbol. Further, $(B^{t(i)})_{dc} = cu_{d+c-1}^{(i)}$. Therefore, the a-th element of the column $X^t\frac{\partial \tilde{B}^t}{\partial \bar{u}^t}Y$ is equal to

$$\sum_{\substack{b,c>0\\ d=a+b-1}} X_b bcu_{d+c-1}^{(i)}Y_c = \sum_{b,c>0} bcu_{a+b+c-2}^{(i)}X_bY_c.$$

It is obvious that this expression is symmetric in X, Y.

Similarly, the matrix $X^t \frac{\partial \tilde{B}}{\partial \bar{u}^t}$ has at the ad-th place $\sum_{b>0} X_b \frac{\partial}{\partial u_d}(B_{ab}) = \sum_{b>0} X_b a \delta_{d,a+b-1}$, while the matrix $(B+B^t)^{(i)}$ at the dc-th site has the element $(c+d)u^{(i)}_{c+d-1}$, so that the a-th element of the column $X^t \frac{\partial \tilde{B}}{\partial \bar{u}^t}(B+B^t)^{(i)} Y$ is

$$\sum_{\substack{b,c>0\\ d=a+b-1}} X_b a(c+d) u^{(i)}_{c+d-1} Y_c = \sum_{b,c>0} a(a+b+c-1) u^{(i)}_{a+b+c-2} X_b Y_c,$$

and this expression is symmetric in X, Y.

b) The element at the d-th place of the column $X^t \frac{\partial B}{\partial \bar{u}} Y$ is $X^t \frac{\partial B}{\partial u_d} Y = \sum_{b,c>0} X_b b \delta_{d,b+c-1} Y_c$. Further, $(B+B^t)^{(i)}_{ad} = (a+d) u^{(i)}_{a+d-1}$. Therefore, the a-th element of the column $(B+B^t)^{(i)} X^t \frac{\partial B}{\partial \bar{u}} Y$ is equal to $\sum_{b,c>0} b(a+b+c-1) u^{(i)}_{a+b+c-2} X_b Y_c$.

In analogy with the computation at the start of the proof, the a-th element of the column $X^t \frac{\partial \tilde{B}^t}{\partial \bar{u}^t}(B+B^t)^{(i)} Y$ is

$$\sum_{\substack{b,c>0\\ d=a+b-1}} X_b b(c+d) u^{(i)}_{c+d-1} Y_c = \sum_{b,c>0} b(a+b+c-1) u^{(i)}_{a+b+c-2} X_b Y_c.$$

c) As above, the a-th element of the column $B^{(i)} X^t \frac{\partial B}{\partial \bar{u}} Y$ is $\sum_{b,\, c>0} ab u^{(i)}_{a+b+c-2} X_b Y_c$.

On the other hand, the matrix $Y^t \frac{\partial \tilde{B}}{\partial \bar{u}^t}$ has at the ad-th place the element $\sum_{c>0} Y_c \frac{\partial}{\partial u_d}(B_{ac}) = \sum_{c>0} a \delta_{d,a+c-1} Y_c$, and $(B^t)^{(i)}_{db} = b u^{(i)}_{b+d-1}$. Therefore, the a-th element of $Y^t \frac{\partial \tilde{B}}{\partial \bar{u}^t} B^{t(i)} X$ is equal to $\sum_{b,c>0} Y_c ab u^{(i)}_{a+b+c-2} X_b$. The proof of the lemma is complete.

8.7. Verification of the Identity (25). Type 00. On the left there are no such terms, while on the right there is $X^t \frac{\partial \tilde{B}^t}{\partial \bar{u}^t} B^{t\prime} Y$ — (the same with $X \leftrightarrow Y$), i.e., zero by Lemma 8.6 a).

Type 01. We must verify that

$$B^{t\prime} X^t \frac{\partial B}{\partial \bar{u}} Y' = X^t \frac{\partial \tilde{B}^t}{\partial \bar{u}^t}(B' + 2B^{t\prime}) Y' - Y^{t\prime} \frac{\partial (\tilde{B}+\tilde{B}^t)}{\partial \bar{u}^t} B^{t\prime} X.$$

On the right there is the monomial $X^t \frac{\partial \tilde{B}^t}{\partial \bar{u}^t} B^{t\prime} Y'$ and a similar term with opposite sign and $X \leftrightarrow Y'$, which cancel by Lemma 8.6a). The remaining terms we transform in accordance with Lemma 8.6 b) and c):

$$X^t \frac{\partial \tilde{B}^t}{\partial \bar{u}^t}(B' + B^{t\prime}) Y' = (B' + B^{t\prime}) X^t \frac{\partial B}{\partial \bar{u}} Y',$$

$$-Y^{t\prime} \frac{\partial \tilde{B}}{\partial \bar{u}^t} B^{t\prime} X = -B' X^t \frac{\partial B}{\partial \bar{u}} Y'.$$

Their sum is obviously equal to the monomial on the left side.

Type 02. We have the equality at once and it follows from Lemma 8.6 b).

Type 10. We must verify that

$$-B^{t'}X^{t'}\frac{\partial B^t}{\partial \bar{u}}Y=X^{t'}\frac{\partial(\tilde{B}+\tilde{B}^t)}{\partial \bar{u}^t}B^{t'}Y-Y^t\frac{\partial \tilde{B}^t}{\partial \bar{u}^t}(B'+2B^{t'})X'.$$

By Lemma 8.6 a) the term $X^{t'}\frac{\partial \tilde{B}^t}{\partial \bar{u}^t}B^{t'}Y$ on the right cancels with the term symmetric to it. We transform the remaining terms according to Lemma 8.6 c) and b):

$$X^{t'}\frac{\partial \tilde{B}}{\partial \bar{u}^t}B^{t'}Y=B'Y^t\frac{\partial B}{\partial \bar{u}}X',$$

$$-Y^t\frac{\partial \tilde{B}}{\partial \bar{u}^t}(B'+B^{t'})=-(B'+B^{t'})Y^t\frac{\partial B}{\partial \bar{u}}X'.$$

The sum of the right sides is equal to $-B^{t'}Y^t\frac{\partial B}{\partial \bar{u}}X'$. The proof is completed by the remark that $Y^t\frac{\partial B}{\partial \bar{u}}X'=X^{t'}\frac{\partial B^t}{\partial \bar{u}}Y$, which is immediately evident from the definition of these expressions.

Type 11. We must verify that

$$(B+B^t)X^{t'}\frac{\partial B}{\partial \bar{u}}Y'-(B+B^t)X^{t'}\frac{\partial B^t}{\partial \bar{u}}Y'=X^{t'}\frac{\partial(\tilde{B}+\tilde{B}^t)}{\partial \bar{u}^t}(B+B^t)Y'-Y^{t'}\frac{\partial(\tilde{B}+\tilde{B}^t)}{\partial \bar{u}^t}(B+B^t)X'.$$

By transforming the first term on the left according to Lemma 8.6 b) and cancelling with the corresponding term on the right, we obtain the equivalent identity:

$$-(B+B^t)X^{t'}\frac{\partial B^t}{\partial \bar{u}}Y'=X^{t'}\frac{\partial \tilde{B}}{\partial \bar{u}^t}(B+B^t)Y'-Y^{t'}\frac{\partial(\tilde{B}+\tilde{B}^t)}{\partial \bar{u}^t}(B+B^t)X'.$$

On the left we interchange X and Y, replacing B^t by B, again apply Lemma 8.6 b), and cancel by Lemma 8.6 d):

$$0=X^{t'}\frac{\partial \tilde{B}}{\partial \bar{u}^t}(B+B^t)Y'-Y^{t'}\frac{\partial \tilde{B}}{\partial \bar{u}^t}(B+B^t)X'.$$

Type 20. We must verify that

$$-(B+B^t)X^{t''}\frac{\partial B^t}{\partial \bar{u}}Y=-Y^t\frac{\partial \tilde{B}^t}{\partial \bar{u}^t}(B+B^t)X''.$$

On the left we make the change $X''\leftrightarrow Y$ and $B^t\leftrightarrow B$; applying Lemma 8.6 b), we obtain the required result.

This completes the proof that the operator (21) is Hamiltonian.

We now proceed to verify that the operator (20) is Hamiltonian. In criterion (23) we put $B=B_\infty=\lim_n B_n$ (formal limit), $B_j=\lim_n B_{n,j}$; X and Y are infinite columns.

The next lemma is verified by straightforward computations using (20), and we limit ourselves to formulating it.

8.8. LEMMA. a) In the a-th element of the column $\sum_j BX^t \frac{\partial B_j}{\partial \bar{u}} Y^{(j)}$ only the coefficient of a monomial of the form $X_\alpha^{(\beta)} Y_\gamma^{(\delta)} u_{a+\alpha+\beta+\gamma+\delta+2}$, can be nonzero, and this coefficient is equal to

$$\sum_j \left[\frac{(a+\beta+\delta-j)!}{a!\,\beta!\,(\delta-j)!} \frac{(\alpha+j)!}{\alpha!\,j!} - \frac{(a+\beta+\delta-j)!}{a!\,\beta!\,(\delta-j)!} \frac{(\gamma+j)!}{\gamma!\,j!} \right]. \tag{26}$$

Here and below all indices are integers $\geqslant 0$; terms with $(\delta-j)!$ (respectively, $(\beta-j)!$) are considered equal to zero for $j>\delta$ (respectively, $j>\beta$).

b) In the a-th element of the column $\sum_j BX^{t(j)} \frac{\partial B_j^t}{\partial \bar{u}} Y$ only the coefficient of a monomial of the form $X_\alpha^{(\beta)} Y_\gamma^{(\delta)} u^{(\eta)}_{a+\alpha+\beta+\gamma+\delta+\eta+2}$, can be nonzero, and this coefficient is equal to

$$(-1)^{\beta+\delta+\eta+1} \sum_j \left[(-1)^j \frac{(\alpha+\beta+\gamma+\delta+\eta+1)!}{(\alpha+\gamma+j+1)!\,\beta!\,(\delta-j)!\,\eta!} \frac{(\alpha+j)!}{\alpha!\,j!} - (-1)^j \frac{(\alpha+\beta+\gamma+\delta+\eta+1)!}{(\alpha+\gamma+j+1)!\,\delta!\,(\beta-j)!\,\eta!} \frac{(\gamma+j)!}{\gamma!\,j!} \right]. \tag{27}$$

c) In the a-th element of the column $\sum_j X^{(j)t} \frac{\partial \tilde{B}_j}{\partial \bar{u}^t} BY$ only the coefficient of $X_\alpha^{(\beta)} Y_\gamma^{(\delta)} u^{(\eta)}_{a+\alpha+\beta+\gamma+\delta+\eta+2}$ can be nonzero. For $\eta=0$ this coefficient is equal to

$$\frac{(a+\beta)!}{a!\,\beta!} \left[\frac{(a+\alpha+\beta+\delta+1)!}{(a+\alpha+\beta+1)!\,\delta!} + (-1)^\delta \frac{(\gamma+\delta)!}{\gamma!\,\delta!} \right]. \tag{28}$$

For $\eta \neq 0$ this coefficient is equal to

$$(-1)^{\eta+\delta} \frac{(\gamma+\delta+\eta)!}{\gamma!\,\delta!\,\eta!} \frac{(a+\beta)!}{a!\,\beta!}. \tag{29}$$

d) In the a-th element of the column $\sum_j (-1)^{j+1} \partial^j \circ X^t \frac{\partial \tilde{B}_j^t}{\partial \bar{u}^t} BY$ only the coefficient of $X_\alpha^{(\beta)} Y_\gamma^{(\delta)} u^{(\eta)}_{a+\alpha+\beta+\gamma+\delta+\eta+2}$, can be nonzero, and this coefficient is equal to

$$\sum_{j<\delta} (-1)^{\beta+\delta+j+\eta+1} \frac{(\alpha+\beta+\delta-j+\eta)!\,(a+\alpha+\beta+\delta+\eta+1)!}{(a+\alpha+\beta+\delta-j+\eta+1)!\,\alpha!\,\beta!\,(\delta-j)!\,\eta!\,j!} + \sum_{k<\eta,\, l<\delta} (-1)^{\beta+\delta+\eta} \frac{(\alpha+\beta+k+l)!\,(\gamma+\delta+\eta-k-l)!}{\alpha!\,\beta!\,\gamma!\,(\delta-l)!\,(\eta-k)!\,k!\,l!}. \tag{30}$$

We remark that parts a) and b) of this lemma refer to the entire left side of (23), while parts c) and d) refer only to half of the right side, the remaining terms being considered by permuting $X \leftrightarrow Y$, i.e., the pair of indices $(\alpha, \beta) \leftrightarrow (\gamma, \delta)$.

We now proceed to describe the cancellations of similar terms.

8.9. LEMMA.

$$\sum_{j<\delta} \frac{(a+\beta+\delta-j)!}{a!\beta!\,(\delta-j)!} \frac{(\alpha+j)!}{\alpha! j!} = \frac{(a+\beta)!}{a!\beta!} \frac{(a+\alpha+\beta+\delta+1)!}{(a+\alpha+\beta+1)!\,\delta!}.$$

Proof. After division by $\frac{(a+\beta)!}{a!\beta!}$ the sum on the left is the coefficient of $X^{a+\beta} Y^\alpha T^\delta$ in the polynomial

$$\sum_{j<\delta} (X+T)^{a+\beta+\delta-j} (Y+T)^{\alpha+j} = (X+T)^{a+\beta} (Y+T)^\alpha \frac{(X+T)^{\delta+1} - (Y+T)^{\delta+1}}{X-Y}.$$

Separating from the two first factors the coefficient of $X^{a+\beta-r}Y^{\alpha-s}$, and from the third the coefficient of $X^rY^sT^j$ (after the binary expansion of $(X+T)^{\delta+1}$ and $(Y+T)^{\delta+1}$ and division by $X-Y$), we obtain $\sum_{r+s+j=\delta}\binom{\delta+1}{j}\binom{a+\beta}{r}\binom{\alpha}{s}$, which is the coefficient of T^δ in the polynomial $(1+T)^{\delta+1}(1+T)^{a+\beta}(1+T)^{\alpha}$, i.e., just $\frac{(a+\alpha+\beta+\delta+1)!}{(a+\alpha+\beta+1)!\,\delta!}$.

8.10. COROLLARY. All terms of the form (26) cancel with the first term of (28) and its transform under the substitution $(\alpha,\beta)\leftrightarrow(\gamma,\delta)$.

8.11. LEMMA.

$$\sum_{j=0}^{\delta}(-1)^j\frac{(\alpha+\gamma+\delta+1)!}{(\delta-j)!\,(\alpha+\gamma+j+1)!}\,\frac{(\alpha+j)!}{\alpha!\,j!}=(-1)^\delta\frac{(\gamma+\delta)!}{\gamma!\,\delta!}.$$

Proof. We make use of the identity

$$\frac{(\alpha+\gamma+\delta+1)!}{(\delta-j)!\,(\alpha+\gamma+j+1)!}=\frac{(\alpha+\gamma+\delta)!}{(\delta-j-1)!\,(\alpha+\gamma+j+1)!}+\frac{(\alpha+\gamma+\delta)!}{(\delta-j)!\,(\alpha+\gamma+j)!}.$$

8.12. COROLLARY. The expression (27) is equal to

$$(-1)^{\beta+\delta+\eta+1}\frac{(\alpha+\beta+\gamma+\delta+\eta+1)!}{\beta!\,\eta!\,(\alpha+\gamma+\delta+1)!}\,\frac{(\gamma+\delta)!}{\gamma!\,\delta!}-(\text{the same with }(\alpha,\beta)\leftrightarrow(\gamma,\delta)).$$

8.13. COROLLARY. The first sum in (30) is equal to

$$(-1)^{\beta+\eta+1}\frac{(\alpha+\beta+\eta)!}{\alpha!\,\beta!\,\eta!}\,\frac{(a+\delta)!}{a!\,\delta!}.$$

8.14. COROLLARY. The second sum in (30) is equal to

$$(-1)^{\beta+\delta+\eta}\frac{(\alpha+\beta+\gamma+\delta+\eta+1)!\,(\alpha+\beta)!}{\alpha!\,\beta!\,\delta!\,\eta!\,(\alpha+\beta+\gamma+1)!}.$$

Proof. We first sum on l for fixed k with the help of Lemma 8.9 (with different values of the parameters) and then sum on k using the same lemma.

8.15. According to 8.12 and 8.14, the terms of (27) cancel in the second sum of (30) and its transform under the substitution $(\alpha,\beta)\leftrightarrow(\gamma,\delta)$. The remaining term of (28) and the coefficient of (29) cancel with the transform of the first sum in (30) under the substitution $(\alpha,\beta)\leftrightarrow(\gamma,\delta)$ according to 8.13. Finally, the last remaining terms cancel by symmetry also by 8.13.

CHAPTER II

THE STRUCTURE OF THE BASIC EQUATIONS

1. Introduction

1.1. The principal purpose of this chapter is to describe the construction, algebraic structure, and conservation laws of some classes of nonlinear partial differential equations and also ordinary differential equations related to them. The introduction is devoted to a brief description of these classes.

1.2. Lax Equations. Let $\mathcal{B}$ be some associative algebra over a field of characteristic k, which lies in the center of $\mathcal{B}$; let $\partial:\mathcal{B}\to\mathcal{B}$ be a differentiation of $\mathcal{B}$ which is trivial on k. The algebra of differential operators $\mathcal{B}[\partial]$ consists of all expressions of the form $\sum_{i=0}^{N} b_i\partial^i,\ b_i\in\mathcal{B}$, with the commutation rule $\partial\circ b=\partial b+b\partial$. If $L=\sum_{i=0}^{N} b_i\partial^i,\ b_N\neq 0$, the number N is called the order of the operator L and is denoted by $\operatorname{ord} L$. It is well known that the left $\mathcal{B}$-module of operators of order $\leqslant N$ is freely generated by $1,\partial,\ldots,\partial^N$.

Basic examples:

a) $\mathcal{B}$ is the ring (of germs) of smooth, analytic or meromorphic functions of the variable x, of the variables x, t, or of the variables x, y, t; $k=\mathbf{R}$ or $\mathbf{C}$; $\partial=\partial/\partial x$.

b) $\mathcal{B}=k[u_i^{(j)}\,|\,i=0,\ldots,n;\ j\geqslant 0]$; $\partial: u_i^{(j)}\mapsto u_i^{(j+1)}$; $k=\mathbf{Q}$, $\mathbf{R}$ or $\mathbf{C}$.

c) $\mathcal{B}=M_l(\mathcal{B}_0)$ is the ring of $(l\times l)$ matrices over a ring $\mathcal{B}_0$ of the type described in a) or b); ∂ is the unique extension of ∂ from $\mathcal{B}_0$ to $M_l(\mathcal{B}_0)$.

We suppose further that in $\mathcal{B}$ there are defined two additional differentiations ∂_1 and ∂_2 such that $\partial, \partial_1, \partial_2$ are pairwise commutative. In example a) $\partial_1=\partial/\partial t,\ \partial_2=\partial/\partial y$; in example b) ∂_1, ∂_2 are some differentiations commuting with ∂, i.e., formal analogues of the evolution fields of Chap. I. We shall usually write ∂_t, ∂_y in place of ∂_1, ∂_2 in case b) as well.

For any $L=\sum b_i\partial^i\in\mathcal{B}[\partial]$ we set $\partial_1 L=\sum \partial_1 b_i\partial^i$ and similarly for ∂_2. The symbol $[P, L]$ denotes the commutator $PL-LP$ of the operators P and L.

A pair of operators $P, L\in\mathcal{B}[\partial]$ is called a solution of the stationary Lax equation (respectively, the Lax equation, the equation of Zakharov–Shabat) if $[P,L]=0$ (respectively, $\partial_1 P=[P, L]$, $\partial_1 P+\partial_2 L=[P, L]$). For brevity, we shall often refer to all these equations as Lax equations.

We remark that in case b) (and similarly in the matrix case) the "unknowns" in the Lax equations, in addition to P, L, include the fields ∂_1, ∂_2, and solutions of these equations in such rings correspond to what in algebraic geometry are called "generic points" of the algebraic manifolds represented by systems of algebraic equations. From the more traditional point of view, when working with these rings, we shall investigate the structure of the Lax equations themselves rather than their solutions in the sense of traditional analysis. The latter case corresponds to rings of type a) and is treated in Chap. III.

The main results pertaining to Lax equations are proved in Secs. 2-5. These include the following: an explicit description of operators P with the property $\operatorname{ord}[P, L]\leqslant\operatorname{ord} L-1$ or $\operatorname{ord} L-2$; establishing that the equations $L_t=[P, L]$ are Hamiltonian over commutative rings of type b); formalization of the method of Zakharov–Shabat. We make use of the technique of fractional powers of Gel'fand and Dikii in a considerably simplified and formalized version. Another version based on the study of the resolvent may be found in [5].

1.3. The Benney Equations. This name we give to the following system of equations for the two-dimensional motion of a nonviscous, incompressible fluid in a gravitational field in the long-wave approximation:

$$\begin{cases} u_t+uu_x-u_y\int_0^y u_x\Big|_{y=\eta} d\eta+h_x=0, \\ h_t+\left(\int_0^h u\,dy\right)_x=0. \end{cases}$$

Here $-\infty<x<\infty$ is the horizontal coordinate; $0\leqslant y$ is the vertical coordinate; t is the time; $u=u(x,y,t)$ is the horizontal component of the velocity at the point (x,y) at time t; $h(x,t)$ is the height of the free surface above the point $(x,0)$ at time t. The notation u_t is an abbreviation for $\frac{\partial}{\partial t}u(x,y,t)$, etc. The system of units is chosen so that the gravitational acceleration and the density are equal to one. Integrals of the type $\int_0^y u_x d\eta$ arise from the equation of continuity $u_x+v_y=0$ where v is the vertical component of the velocity and from the boundary condition $v=0$ at $y=0$; these relations make it possible to eliminate v, by expressing it in terms of u: $v(x,y,t)=-\int_0^y u_x d\eta$. For the remaining details of the derivation see the work of Benney [27] who first discovered an unexpected property of the system: the existence of an infinite sequence of conservation laws for it.

The Benney equations display a number of unusual properties. We do now know of a Lax pair for them; the conservation laws in the interpretation of the present work (in contrast to the formal derivation of Benney) are obtained from a nonlinear integral equation with a parameter. This equation enables us to obtain the conservation laws of Miura [46] for the Benney system. We further establish the Hamiltonian character of the "reduced system" (with the additional condition $u_y=0$) and the commutativity of the reduced integrals. These results are then generalized to the full system.

Our exposition is based on the work of Kupershmidt and the author [19].

The main results pertaining to the Benney equations are formulated in more detail in Sec. 6; Secs. 7-13 are devoted to their proofs.

2. The Commutator and Fractional Powers of Differential Operators

2.1. In this section we begin the study of the Lax equations. If $\partial_t L=L_t=[P,L]$ or $[P,L]=0$, then in any case $\operatorname{ord}[P,L]\leqslant\operatorname{ord}L$. We therefore first investigate the conditions under which the commutator of two differential operators has lower order.

We set $L=\sum_{n=0}^{N} u_n\partial^n,\ P=\sum_{m=0}^{M} v_m\partial^m.$

2.2. LEMMA. Setting $\partial^\alpha w = w^{(\alpha)}$ for any $w \in \mathcal{B}$, $\alpha \geqslant 0$, we have

$$[P, L] = \sum_{\alpha, m, n \geqslant 0} \left[\binom{m}{\alpha} v_m u_n^{(\alpha)} - \binom{n}{\alpha} u_n v_m^{(\alpha)} \right] \partial^{m+n-\alpha}.$$

Proof. According to Leibnitz's rule $\partial^j (uv) = \sum_{\alpha=0}^{j} \binom{j}{\alpha} u^{(\alpha)} v^{(j-\alpha)}$.

2.3. COROLLARY. The coefficient of ∂^{M+N} in $[P, L]$ is equal to $[v_M, u_N]$. In particular, $\operatorname{ord}[P, L] \leqslant M+N-1$, if $u_N \in$ is contained in the center of $\mathcal{B}$.

2.4. COROLLARY. For $v \in \mathcal{B}$ we have $[v, L] = [v, u_N]\partial^N + (vu'_N - Nu_N v)\partial^{N-1} +$ (terms of order $\leqslant N-2$). In particular, if the center of $\mathcal{B}$ is infinite-dimensional over k, then the linear space of those P, for which $\operatorname{ord}[P, L] \leqslant N-1$, is infinite-dimensional.

We shall see below that the condition $\operatorname{ord}[P, L] \leqslant N-2$ in typical examples already leads to a finite-dimensional space.

2.5. COROLLARY. The coefficient of ∂^{N+k}, $k \in Z$, in the commutator $[P, L]$ depends only on those coefficients v_j of the operator P and their derivatives for which $j \geqslant k$. The terms depending on v_k, v_{k+1} and their derivatives in this coefficient have the form

$$[v_k, u_N] + [v_{k+1}, u_{N-1}] + Mv_{k+1}u'_N - Nu_N v'_{k+1}. \tag{1}$$

Below we shall always assume that u_N is invertible in $\mathcal{B}$ and is a ∂-constant (i.e., $\partial u_N = 0$). The operator L is fixed, while P may vary.

2.6. THEOREM. If u_N, u_{N-1} lie in the center of $\mathcal{B}$ and d is the dimension over k of the space of ∂-constants in $\mathcal{B}$, then the dimension of the space of operator $P \in \mathcal{B}[\partial]$, for which $\operatorname{ord}[P, L] \leqslant N-2$, does not exceed $(M+1)d$. It is equal to $(M+1)d$, if $\partial : \mathcal{B} \to \mathcal{B}$ is a surjective mapping (it is possible to "integrate on x" in the ring $\mathcal{B}$).

Proof. For $k = M, M-1, \ldots, 0, -1$ we call the "k-th equation" (relative to v_j) the condition that the coefficient of ∂^{N+k} in $[P, L]$ vanish. It is evident from (1) that the M-th equation is trivially satisfied. Further, for fixed $v_M, \ldots, v_{k+2}$ the k-th equation, according to (1), has the form $v'_{k+1} = w$, where w is a polynomial in u_N^{-1}, $u_j^{(\alpha)}$, $v_l^{(\beta)}$, $M \geqslant l \geqslant k$. If it is solvable at all, then any solution is obtained from one solution by addition of any ∂-constant in $\mathcal{B}$. It is clearly solvable if $\partial : \mathcal{B} \to \mathcal{B}$ is surjective. This completes the proof.

2.7. COROLLARY. If the ∂-constants in $\mathcal{B}$ coincide with k (e.g., $\mathcal{B}$ is the ring of germs of functions of x or $\mathcal{B} = k[u_i^{(j)}]$), then the dimension of the space described in Theorem 2.6 does not exceed $M+1$. (It will be shown below that in the commutative case it is equal to $M+1$, even if $\partial : \mathcal{B} \to \mathcal{B}$ is not surjective.)

For matrix rings $\mathcal{B}$ it is possible to obtain an analogous estimate under less restrictive assumptions which we now axiomatize.

2.8. Let $\operatorname{ad} u_N : \mathcal{B} \to \mathcal{B}$ be the operator $\operatorname{ad}_{u_N}(b) = [u_N, b]$. We set $\mathcal{B}^+ = \operatorname{Ker} \operatorname{ad} u_N$, $\mathcal{B}^- = \operatorname{Im} \operatorname{ad} u_N$ and assume that the following conditions are satisfied:

a) $\mathcal{B}=\mathcal{B}^+\oplus\mathcal{B}^-$ (as a k-space); $\mathcal{B}^+\mathcal{B}^-\subseteq\mathcal{B}^-$, $\mathcal{B}^-\mathcal{B}^+\subseteq\mathcal{B}^-$.

b) The operator $\operatorname{ad} u_N:\mathcal{B}^-\to\mathcal{B}^-$ is bijective.

c) $\partial\mathcal{B}^+\subseteq\mathcal{B}^+$.

For any element $w\in\mathcal{B}$ we denote by w^+, w^- the components of w in $\mathcal{B}^+$ and $\mathcal{B}^-$, respectively. If for u_N the conditions a)-c) are satisfied then we call u_N semisimple. In a commutative ring $\mathcal{B}$ every element is semisimple: $\mathcal{B}^+=\mathcal{B}, \mathcal{B}^-=\{0\}$. In the general case our terminology is motivated by the following example.

<u>2.9. Example.</u> Let $\mathcal{B}_0$ be commutative, $\mathcal{B}=M_l(\mathcal{B}_0)$, and u_N be a diagonal matrix with nonzero elements $c_i\in k$ on the diagonal. Then $\mathcal{B}^+$ consists of matrices with zeros at those sites for which $c_i\neq c_j$, and $\mathcal{B}^-$ of matrices at the sites ij, for which $c_i=c_j$. It is easy to verify that all conditions 2.8 a), b), and c) are satisfied. This implies that any semisimple matrix of $M_l(k)$ is also semisimple in our sense of the word. As is known, the converse is also true.

In conditions 2.8 we denote by d the dimension of $\operatorname{Ker}\partial\cap\mathcal{B}^+$ over k. In Example 2.9 it is equal to the sum of the squares of the multiplicities of the eigenvalues of u_N.

<u>2.10. THEOREM.</u> If u_N is invertible, constant, and semisimple and $u_{N-1}\in\mathcal{B}^-$, then the dimension of the space of operators P with $\operatorname{ord}[P,L]\leqslant N-2$ does not exceed $d(M+1)$. It is precisely equal to $d(M+1)$ if $\partial:\mathcal{B}^+\to\mathcal{B}^+$ is surjective.

<u>Proof.</u> The proof is analogous to the proof of Theorem 2.6, but it is somewhat more complicated due to the fact that the k-th equation now contains v_{k+1} as well as v_k:

$$[v_k, u_N]+[v_{k+1}, u_{N-1}]-Nu_Nv'_{k+1}=w.$$

We do induction k downwards assuming that $v^-_{k+1}, v_{k+2},\dots,v_M$ are already defined from equations with indices greater than or equal to $k+1$. Then the condition for solvability of the k for v_k consists in (by 2.8 b))

$$[v_{k+1}, u_{N-1}]^+-(Nu_Nv'_{k+1})^+=w^+.$$

If it is satisfied, then v^-_k is uniquely determined. On the other hand, $[v^+_{k+1}, u_{N-1}]^+=0$, since $u_{N-1}\in\mathcal{B}^-$. Therefore, $[v_{k+1}, u_{N-1}]^+=[v^-_{k+1}, u_{N-1}]^+$ and v^-_{k+1} is already defined by the induction hypothesis. Hence v^+_{k+1} is determined up to an element of $\operatorname{Ker}\partial\cap\mathcal{B}^+$, if it exists at all, and it necessarily exists if $\partial:\mathcal{B}^+\to\mathcal{B}^+$ is surjective. This completes the induction step and the proof.

2.11. Using the method of Gel'fand and Dikii, below we construct explicitly for each order an operator P with the condition $\operatorname{ord}[P, L]<N-2$ as the differential part of an appropriate fractional power of L. To this end we introduce the formal ring of symbols.

We denote by ξ the free variable and consider the polynomial ring $\mathcal{B}[\xi]$ (ξ commutes with $\mathcal{B}$). The mapping $L\mapsto\tilde{L}:\sum b_i\partial^i\mapsto\sum b_i\xi^i$ induces an isomorphism of left $\mathcal{B}$-modules but not of rings. The transferal of multiplication of operators to $\mathcal{B}[\xi]$ we shall call composition and denote it by $\circ$: $\widetilde{PL}=\tilde{P}\circ\tilde{L}$. In order to describe composition in $\mathcal{B}[\xi]$ in terms of the inner structures, we introduce two differentiations in $\mathcal{B}[\xi]$: $\partial:\sum b_i\xi^i\mapsto\sum \partial b_i\xi^i$ and $\partial_\xi:\sum b_i\xi^i\mapsto\sum ib_i\xi^{i-1}$. The next lemma follows from the Leibniz formula.

2.12. LEMMA. $\widetilde{PL}=\sum_{\alpha\geq 0}\frac{1}{\alpha!}\partial_\xi^\alpha\tilde{P}\partial^\alpha\tilde{L}.$

The idea is now to use the well-known construct of extensions of commutative rings and extensions of differentiations to them in order to learn how to extend composition by the formula of Lemma 2.12 and thus to construct useful ring extensions for differential operators.

For our purposes a single extension suffices: $\mathcal{B}=\mathcal{B}((\xi^{-1}))=\left\{\sum_{i=-\infty}^{k} b_i\xi^i \mid b_i\in\mathcal{B}\right\}$ (the ring of formal Laurent series). Obviously ∂ and ∂_ξ extends to $\mathcal{B}((\xi^{-1}))$ by continuity with the same formulas.

We set $\operatorname{ord}\left(\sum b_i\xi^i\right)=\max\{k\mid b_k\neq 0\}$ and further $\|a\|=2^{\operatorname{ord}a}$ for $a\in\mathcal{B}((\xi^{-1}))$ (we assume that $\operatorname{ord}0=-\infty$). The norm $\|\ \|$ is non-Archimedean: $\|a+b\|\leq\max(\|a\|,\|b\|)$. Moreover, $\|ab\|\leq\|a\|\|b\|$, $\|a\|=1$ for $a\in\mathcal{B}\setminus\{0\}$ and $\|a\|=0$ if and only if $a=0$. The algebra $\mathcal{B}((\xi^{-1}))$ is complete in this norm.

For any additive operator $F:\mathcal{B}\to\mathcal{B}$ we set $\|F\|=\sup_{a\neq 0}\|Fa\|/\|a\|$. Obviously, $\|\partial\|\leq 1$, $\|\partial_\xi\|=\frac{1}{2}<1$. In analogy with Lemma 2.12, we introduce in $\mathcal{B}((\xi^{-1}))$ the composition $\circ$, by setting

$$a\circ b=\sum_{\alpha\geq 0}\frac{1}{\alpha!}\partial_\xi^\alpha a\partial^\alpha b. \tag{2}$$

2.13. LEMMA. Series (2) converges in norm for any a, $b\in\mathcal{B}((\xi^{-1}))$ and defines on $\mathcal{B}((\xi^{-1}))$ the structure of an associate k-algebra.

Proof. The convergence of (2) follows from $\left\|\frac{1}{\alpha!}\partial_\xi^\alpha a\partial^\alpha b\right\|\leq 2^{-\alpha}\|a\|\|b\|\to 0$ as $\alpha\to\infty$. Bilinearity in a and b is obvious. The identity is $1\in\mathcal{B}$. Associativity is verified as follows:

$$(a\circ b)\circ c=\left[\sum_{\alpha\geq 0}\frac{1}{\alpha!}\partial_\xi^\alpha a\partial^\alpha b\right]\circ c=\sum_{\alpha,\beta,\gamma<\beta}\frac{1}{\alpha!\beta!}\binom{\beta}{\gamma}\partial_\xi^{\alpha+\gamma}a\partial_\xi^{\beta-\gamma}\partial^\alpha b\partial^\beta c,$$

$$a\circ(b\circ c)=\sum_{A\geq 0}\frac{1}{A!}\partial_\xi^A a\circ\partial^A\left[\sum_{B\geq 0}\frac{1}{B!}\partial_\xi^B\partial^B c\right]=\sum_{A,B,\Gamma<B}\frac{1}{A!B!}\binom{A}{\Gamma}\partial_\xi^A a\partial_\xi^B\partial^\Gamma b\partial^{A+B-\Gamma}c.$$

It is not hard to see that the substitution $\alpha=\Gamma$, $\beta=A+B-\Gamma$, $\gamma=A-\Gamma$ defines a bijection of the terms of both series and corresponding terms coincide. The proof of the lemma is complete.

We now suppose that M is a left $\mathcal{B}$-module to which the action ∂: $\partial(bm)=\partial b\cdot m+b\partial m$ has been extended for $b\in\mathcal{B}$, $m\in M$. We denote by $M((\xi^{-1}))$ the left $\mathcal{B}((\xi^{-1}))$-module $\left\{\sum_i m_i\xi^i \mid m_i\in M,\ \exists i_0\ \forall i>i_0,\ m_i=0\right\}$ with the obvious action. We extend the action of ∂ to $M((\xi^{-1}))$ coefficient-wise. It is not hard to see that formula (2) enables us to define on $M((\xi^{-1}))$ a new structure of a left $\mathcal{B}((\xi^{-1}))$-module relative to composition if we take $a\in\mathcal{B}((\xi^{-1}))$, $b\in M((\xi^{-1}))$.

Right composition $M((\xi^{-1}))\times\mathcal{B}((\xi^{-1}))\to M((\xi^{-1}))$, is defined similarly if M is a right $\mathcal{B}$-module with action ∂. Finally, if M is a $\mathcal{B}$-bimodule, then $M((\xi^{-1}))$ with left and right composition becomes a $\mathcal{B}((\xi^{-1}))$-bimodule. The only nonobvious assertions here concern associativity, and its verification in the proof of Lemma 2.13 carries over automatically in all the required cases. We shall need this construction in Sec. 3.

An interpretation of the elements of the ring $\mathcal{B}((\xi^{-1}))$ with negative powers of ξ as symbols of integrodifferential operators is given in Sec. 5.

2.14. LEMMA. An element $\sum b_i\xi^i=b$ is invertible in $\mathcal{B}((\xi^{-1}))$ both with respect to multiplication and composition if and only if $b_n\in\mathcal{B}$ is invertible in $\mathcal{B}$, where $n=\operatorname{ord}\left(\sum b_i\xi^i\right)$.

Proof. Invertibility with respect to multiplication is well known. Further, $\left(\sum b_i\xi^i\right)\circ b_n^{-1}\xi^{-n}=1+c$, $\operatorname{ord} c\leqslant -1$, as is evident from (2). Therefore, $\sum_{i=0}^{\infty}(-1)^i c^{\circ i}$ exists in $\mathcal{B}((\xi^{-1}))$, and is inverse to $1+c$ in the sense of composition: $(1+c)\circ\sum_{i=0}^{\infty}(-1)^i c^{\circ i}=1$ (we write $c^{\circ i}=c\circ\ldots\circ c$, i times). Then $b_n\xi^{-n}\circ(1+c)^{\circ(-1)}$ is inverse to $\sum b_i\xi^i$ on the right. A left inverse is established similarly. Finally, if $b\circ x=1$ or $x\circ b=1$, then by (2) the leading term of x must be equal to $b_n^{-1}\xi^{-n}$, so that b_n is invertible.

We proceed to the extraction of roots. If the element $c=\sum_{n<N}u_n\xi^n$ is an N-th power of X, then obviously $X=w\xi+\sum_{i\geqslant 0}x_i\xi^{-i}$ and $w^N=u_N$. In order to establish the converse assertion, we introduce the following notation. We call an element $w\in\mathcal{B}$ N-admissible $(N>0)$, if the mapping $\mathcal{B}\mapsto\mathcal{B}: x\mapsto\sum_{i=0}^{N-1}w^i x w^{N-1-i}$ is a bijection. In a commutative ring $\mathcal{B}$ with unique division by N precisely the invertible elements are admissible. In a matrix ring the element $w=\operatorname{diag}(c_1,\ldots,c_l)$ is admissible if and only if all elements $\sum_{i=0}^{N-1}c_r^i c_s^{N-1-i}$ $(r,s=1,\ldots,l)$ are different from zero.

2.15. LEMMA. Let $c=\sum_{n<N}u_n\xi^n$, $N>0$. Then for each N-admissible root of an N-th power w of u_N there exists a unique element $X\in\mathcal{B}((\xi^{-1}))$, for which $X^N=c$ and $X=w\xi+O(1)$ ($O(\xi^i)$ denotes some series of order $\leqslant i$).

Proof. We apply the method of successive approximations. We set $X_{-1}=w\xi$. Obviously, $X_{-1}^N=u_N\xi^N+O(\xi^{N-1})$. Suppose that for some $r\geqslant -1$ we have already proven the existence of $X_r\in\mathcal{B}((\xi^{-1}))$, such that $X_r=w\xi+O(1)$, $X_r^N=c+O(\xi^{N-r-2})$ and its uniqueness up to $O(\xi^{-(r+1)})$. We seek X_{r+1} in the form $X_{r+1}=X_r+x_{r+1}\xi^{-(r+1)}$. Using the distributivity of composition, we have

$$(X_r+x_{r+1}\xi^{-(r+1)})^{\circ N}=X_r^{\circ N}+\sum_{i=0}^{N-1}x_r^{\circ i}\circ x_{r+1}\xi^{-(r+1)}\circ x_r^{\circ(N-i-1)}+\text{(remainder)}.$$

We compute the terms on the right up to $O(\xi^{N-r-3})$. The sum has the form $\sum_{i=0}^{N-1}w^i x_{r+1}\cdot w^{N-i-1}\xi^{N-r-2}+O(\xi^{N-r-3})$. The remainder consists of a sum of products of $j\leqslant N-2$ of elements X_r and $N-j$ elements $x_{r+1}\xi^{-(r+1)}$ in different order. Therefore, its order does not exceed $\max_{j\leqslant N-2}(-(r+1)(N-j)+j)=N-2r-4\leqslant N-r-3$ for all $r\geqslant -1$. Hence, from the admissibility of w it follows that x_{r+1} exists and is uniquely determined by the requirement $X_{r+1}^{\circ N}=c+O(\xi^{N-r-3})$. This completes the proof.

2.16. We now choose an operator $L=\sum_{n=0}^{N} u_n\partial^n \in \mathcal{B}[\partial]$ and an N-admissible root of an N-th power w of u_N (if it exists). We construct a root X of degree N of $\tilde{L}$, as in Lemma 2.15; $X=w\xi+O(1)$. We denote by $\tilde{L}^{\circ s}$ for any $s=pN^{-1}\in Q_N=ZN^{-1}$ the element $X^p\in\mathcal{B}((\xi^{-1}))$. Let

$$\tilde{L}^{\circ s}=\sum_{j<p} v_j(s;w,L)\,\xi^j.$$

Assuming that w and L are fixed, we shall often write below $v_j(s;w,L)=v_j(s)$. For $s\geqslant 0$ we set

$$\langle L^s \rangle=\sum_{j=0}^{p} v_j(s)\,\partial^j.$$

2.17. THEOREM. For any $s\geqslant 0$, $s\in Q_N$, we have $\operatorname{ord}[\langle L^s\rangle, L]\leqslant N-1$. If u_N lies in the center of $\mathcal{B}$, then even $\operatorname{ord}[\langle L^s\rangle, L]\leqslant N-2$.

2.18. COROLLARY. Suppose that u_N, u_{N-1} lie in the center of $\mathcal{B}$, that u_N is invertible and is an N-th power of an admissible element w , that the set of ∂-constants in $\mathcal{B}$ coincides with k . Then for any $M\geqslant 0$ the space of operators $P\in\mathcal{B}[\partial]$ with the property $\operatorname{ord}[P,L]\leqslant N-2$, $\operatorname{ord}P\leqslant M$ is freely generated by the operators $\langle L^s\rangle$, $0\leqslant s\leqslant MN^{-1}$.

Indeed, since the order of the operator $\langle L^s\rangle$ is exactly sN, and its leading coefficient w^{sN} is invertible, all $\langle L^s\rangle$ of order $\leqslant M$ generate a space of dimension $M+1$. It remains to use Corollary 2.7.

For matrix rings, however, the operators $\langle L^s\rangle$, in general, generate only a part of the space of interest to us.

2.19. Proof of Theorem 2.17. All powers X from 2.17 commute pairwise and, in particular, commute with $\tilde{L}=X^N$. Therefore, in the ring $\mathcal{B}((\xi^{-1}))$ with composition we have $[\langle L^s\rangle^{\sim}, \tilde{L}]=[\tilde{L}, \tilde{L}^{\circ s}-\langle L^s\rangle^{\sim}]$. But $\operatorname{ord}(\tilde{L}^{\circ s}-\langle L^s\rangle^{\sim})\leqslant -1$ by definition of $\langle L^s\rangle$. Therefore, $\operatorname{ord}[\langle L^s\rangle^{\sim}, \tilde{L}]\leqslant N-1$ and is even $\leqslant N-2$, if u_N lies in the center of $\mathcal{B}$ (use is made of Corollary 2.3 which remains valid for composition in $\mathcal{B}((\xi^{-1}))$, by formula (2)).

3. The Hamiltonian Property for the Nonstationary Lax Equations and Their Integrals

3.1. In this section we set $\mathcal{B}=k[u_i^{(j)}]$, where $i=0,\dots,N-2$ $(N\geqslant 2)$, $j\geqslant 0$; $\partial: u_i^{(j)}\mapsto u_i^{(j+1)}$, are algebraically independent variables. Let further $L=\partial^N+\sum_{i=0}^{N-2} u_i\partial^i$. We choose $w=1$ and for any $s\in Q_N$ we set, as in 2.16,

$$\tilde{L}^{\circ s}=\sum_{j\leqslant Ns} v_j(s)\,\xi^j,\quad v_j(s)\in\mathcal{B}.$$

According to Corollary 2.18, all solutions of the Lax equation $\partial_1 L=[P, L]$ in the ring $\mathcal{B}$ have the following form $P=\sum c_i \langle L^{s_i} \rangle$, $c_i \in k$, $s_i \in Q_N$, $s_i \geqslant 0$; $\partial_1=\partial_{1,P}:\mathcal{B}\to\mathcal{B}$ is an evolution differentiation (i.e., it commutes with ∂ and is trivial on k) uniquely defined by the conditions

$$\partial_1 u_k = \text{coefficient of } \partial^k \text{ in } \left[\sum c_j \langle L^{s_j} \rangle, L\right], \; k=0,\ldots,N-2.$$

We shall here prove the following basic theorem.

<u>3.2. THEOREM.</u> a) The evolution differentiation ∂_1, defined by the condition $\partial_1 L=\left[\sum c_j \langle L^{s_j} \rangle, L\right]$, is Hamiltonian with operator B_{N-2} (cf. Sec. 8, Chap. I) and Hamiltonian $\sum c_j \frac{v_{-1}(s_j+1)}{s_j+1}$.

b) All Hamiltonians $v_{-1}(r)$, $r \in Q_N$, $r \geqslant 0$, commute pairwise in this Hamiltonian structure and are therefore conservation laws for any of the Lax equations described.

We begin with a number of auxiliary assertions. We set $\operatorname{res}\left(\sum b_i \xi^i\right)=b_{-1}$.

<u>3.3. LEMMA.</u> Let $a, b \in \mathcal{B}((\xi^{-1}))$. Then $\operatorname{res}(a\circ b - b\circ a) \in \partial\mathcal{B}$.

<u>Proof.</u> It suffices to verify this for $a=v\xi^m$, $b=u\xi^n$. It follows easily from formula (2) that

$$\operatorname{res}[a, b]=\binom{m}{m+n+1} vu^{(m+n+1)}-\binom{n}{m+n+1} uv^{(m+n+1)},$$

if $m+n+1>0$ and either $m>0$, or $n>0$, but $mn<0$, and $\operatorname{res}[a,b]=0$ in the remaining cases. Let us suppose that $m>0$, $n<0$; the second alternative is treated analogously. Then

$$\binom{n}{m+n+1}=\frac{n(n-1)\ldots(-m)}{(m+n+1)!}=(-1)^{m+n+1}\binom{m}{m+n+1}.$$

Therefore, $\operatorname{res}[a,b]$ is proportional to $vu^{(m+n+1)}-(-1)^{m+n+1}uv^{(m+n+1)}$ and is a total derivative by a lemma of Chap. I.

3.4. In order to verify that the Lax equations are Hamiltonian, we adapt the results of Chap. I on the characterization of variational derivatives $\frac{\delta}{\delta u_k}$ to our formal case.

To this end we denote by $\Omega^1(\mathcal{B})$ the universal module of differentials of the ring $\mathcal{B}/k$ in the algebraic sense of the word, and we let $\delta:\mathcal{B}\to\Omega^1(\mathcal{B})$ be the universal differentiation. It is well known that $\Omega^1(\mathcal{B})$ is freely generated over $\mathcal{B}$ by the elements $\delta u_i^{(j)}$ and $\delta P=\sum \frac{\partial P}{\partial u_i^{(j)}} \delta u_i^{(j)}$. We extend $\partial:\mathcal{B}\to\mathcal{B}$ to a differentiation $\partial:\Omega^1(\mathcal{B})\to\Omega^1(\mathcal{B})$, by setting $\partial(\delta u_i^{(j)})=\delta u_i^{(j+1)}$. This is a formal analogue of the operator $L_{\overline{\partial/\partial x_1}}$ of Chap. I where $\overline{\partial/\partial x_1}$ is the canonical lift of $\partial/\partial x_1$ to the jets. Obviously, $\delta\partial=\partial\delta$, since $[\delta, \partial]$ is trivial on the $u_i^{(j)}$. It is not hard to demonstrate the validity of the following formal version of Theorem 4.1 of Chap. I.

3.5. LEMMA. For any $P\in\mathcal{B}$ there exists a unique representation of δP in the form $\sum_{i=0}^{N-2} Q_i\delta u_i+\partial\omega$, $\omega\in\Omega^1(\mathcal{B})$; for this representation $Q_i^{i=0}=\frac{\delta P}{\delta u_i}$.

(All formulas of Sec. 4, Chap. I pertaining to the computation of variational derivatives are to be "divided by dx_1", $(m=1)$, and are to have $du_i^{(j)}$ replaced by $\delta u_i^{(j)}$).

3.6. We now pass to the ring $\mathcal{B}((\xi^{-1}))$. We define the $\mathcal{B}((\xi^{-1}))$-module $\Omega^1(\mathcal{B})((\xi^{-1}))$, consisting of series $\sum \omega_i\xi^i$, $\omega_i\in\Omega^1(\mathcal{B})$ with the usual rule of multiplication on the left and right. The differentiations ∂ and ∂_ξ extend to $\Omega^1(\mathcal{B})((\xi^{-1}))$: the first coefficient-wise and the second by the usual formula. This enables us to introduce another action of $\mathcal{B}((\xi^{-1}))$ on $\Omega^1(\mathcal{B})((\xi^{-1}))$ — the composition

$$b\circ\omega=\sum_{\alpha\geqslant 0}\frac{1}{\alpha!}\partial_\xi^\alpha b\,\partial^\alpha\omega.$$

According to the remark following Lemma 2.13, $\Omega^1(\mathcal{B})((\xi^{-1}))$ is converted into a left $(\mathcal{B}((\xi^{-1})),\circ)$-module. The composition $\omega\circ b$ defined by an analogous formula converts $\Omega^1(\mathcal{B})((\xi^{-1}))$ into a right $(\mathcal{B}((\xi^{-1})),\circ)$ -module. It in fact becomes a bimodule, since the formula of associativity $b\circ(\omega\circ c)=(b\circ\omega)\circ c$ is verified in the same way as for the associativity of multiplication in $\mathcal{B}((\xi^{-1}))$.

We extend the differential $\delta:\mathcal{B}\to\Omega^1(\mathcal{B})$ to $\delta:\mathcal{B}((\xi^{-1}))\to\Omega^1(\mathcal{B})((\xi^{-1}))$ coefficient-wise. The mapping $\mathrm{res}:\Omega^1(\mathcal{B})((\xi^{-1}))\to\Omega^1(\mathcal{B})$ picks out the coefficient of ξ^{-1}.

The easily verified compatibility properties of the structures introduced are collected in the following lemma.

3.7. LEMMA. a) The operators res, ∂, δ are pairwise commutative (in the sense of the commutativity of the corresponding diagrams).

b) $\delta(a\circ b)=\delta a\circ b+a\circ\delta b$ for any $a, b\in\mathcal{B}((\xi^{-1}))$.

The second assertion follows from the fact that δ also commutes with ∂_ξ .

We now apply this formalism to the computation of $\frac{\delta v_{-1}(s)}{\delta u_k}$.

3.8. LEMMA. $\delta v_{-1}(s)\equiv s\,\mathrm{res}\,\delta\tilde{L}\circ\tilde{L}^{\circ(s-1)}\bmod \mathrm{Im}\,\partial$ for $s\in Q_N$, $s>0$.

Proof. Let $X=\tilde{L}^{\circ(N^{-1})}$, $s=pN^{-1}$. Then $\tilde{L}=X^{\circ N}$ and $v_{-1}(s)=\mathrm{res}\,X^{\circ p}$.

We wish to establish that

$$\delta\,\mathrm{res}\,X^{\circ p}\equiv\frac{p}{N}\,\mathrm{res}\,\delta\,(X^{\circ N})\circ X^{\circ(p-N)}\bmod \mathrm{Im}\,\partial$$

and

$$N\delta\,\mathrm{res}\,X^{\circ p}\equiv p\delta\,(X^{\circ N})\circ X^{\circ(p-N)}\bmod \mathrm{Im}\,\partial.$$

According to Lemma 3.7,

$$N\delta\,\mathrm{res}\,X^{\circ p}=N\,\mathrm{res}\,\delta\,(X^{\circ p})=N\,\mathrm{res}\sum_{i=0}^{p-1}X^{\circ i}\circ\delta X\circ X^{\circ(p-i-1)}.$$

Further, $X^{\circ i}\circ\delta X\circ X^{\circ(p-i-1)}=\delta X\circ X^{\circ(p-1)}-[\delta X\circ X^{\circ(p-i-1)}, X^{\circ i}]$. It is shown by the same argument as in Lemma 3.3 that the residue of the commutator on the right lies in $\mathrm{Im}\,\partial$. Therefore,

$$N\delta \operatorname{res} X^{\circ p} \equiv Np \operatorname{res} \delta X \circ X^{\circ (p-1)} \bmod \operatorname{Im} \partial.$$

On the other hand,

$$p\delta (X^{\circ N}) \circ X^{p-N} = p \sum_{i=0}^{N-1} x^{i} \circ \delta X \circ X^{\circ (p-i-1)}$$

and since

$$X^{i} \circ \delta X \circ X^{p-i-1} = \delta X \circ X^{p-1} - [\delta X \circ X^{\circ (p-i-1)}, X^{\circ i}],$$

the same argument shows that

$$p \operatorname{res} \delta (X^{\circ N}) \circ X^{\circ (p-N)} \equiv Np \operatorname{res} \delta X \circ X^{\circ (p-1)} \bmod \partial.$$

This completes the proof of the lemma.

3.9. COROLLARY. For $s \in Q_N$, $s > 0$ and $0 \leqslant i \leqslant N-2$ we have

$$\frac{\delta v_{-1}(s)}{\delta u_i} = s \sum_{k=0}^{i} \binom{i}{k} \partial^k v_{k-i-1}(s-1). \tag{3}$$

Proof. According to the preceding lemma,

$$\delta v_{-1}(s) \equiv s \operatorname{res} \left(\left(\sum_{i=0}^{N-2} \delta u_i \xi^i \right) \circ \left(\sum_{j=-\infty}^{N(s-1)} v_j (s-1) \xi^j \right) \right) \equiv$$

$$\equiv s \operatorname{res} \left(\sum_{\alpha, i, j} \binom{i}{\alpha} \delta u_i v_j^{(\alpha)} (s-1) \xi^{i+j-\alpha} \right) \equiv s \sum_{i,j} \binom{i}{i+j+1} v_j^{(i+j+1)} (s-1) \delta u_i \bmod \operatorname{Im} \partial,$$

whence the required result follows in view of the characterization of variational derivatives in Lemma 3.5.

3.10. Proposition. For all $s > 0$, $s \in Q_N$ and $0 \leqslant i \leqslant N-2$ we have

$$v_{-i-1}(s-1) = \frac{1}{s} \sum_{j=0}^{i} \binom{i}{j} (-\partial)^j \frac{\delta}{\delta u_{i-j}} v_{-1}(s). \tag{4}$$

Proof. Relations (3) can be considered a system of equations for $v_{-1}(s-1), \ldots, v_{-N+1}(s-1)$, which has triangular form and can therefore be solved by induction. The closed formulas (4) are most easily obtained by writing (3) in operator form

$$\frac{\delta v_{-1}(s)}{\delta u_i} = s (1 + \partial \circ T)^i v_{-i-1}(s-1), \tag{5}$$

where T is the operator for increasing the index by one: $T(v_{-k}^{(\alpha)}) = v_{-k+1}^{(\alpha)}$. Equation (4) then becomes

$$v_{-i-1}(s-1) = \frac{1}{s} (1 - \partial \circ T^{-1})^i \frac{\delta v_{-1}(s)}{\delta u_i} \tag{6}$$

and the fact that (6) is the inversion of (5) is obtained by induction on i.

3.11. Proof of Theorem 3.2 a). It obviously suffices to consider the equation $\partial_1 L=[\langle L^s\rangle, L]$. We rewrite it in the algebra $(\mathscr{B}((\xi^{-1})),\circ)$, replacing L by $\tilde{L}$ and $\langle L^s\rangle$ by $\langle L^s\rangle^{\sim}-\tilde{L}^{\circ s}$ (it is recalled that $[\tilde{L}^{\circ s},\tilde{L}]=0$). We find on setting $u_{N-1}=0,\ u_N=1$:

$$\sum_{k=0}^{N-2}\partial_1 u_k\xi^k=\left[\sum_{\alpha=0}^{N}u_\alpha\xi^\alpha,\ \sum_{\beta=1}^{\infty}v_{-\beta}(s)\,\xi^{-\beta}\right].$$

This implies the two identities

$$\sum_{k=0}^{N-2}\partial_1 u_k\xi^k=\left[\sum_{\alpha=0}^{N}u_\alpha\xi^\alpha,\ \sum_{\beta=1}^{N-1}v_{-\beta}(s)\,\xi^{-\beta}\right],$$

$$0=\left[\sum_{\alpha=0}^{N}u_\alpha\xi^\alpha,\ \sum_{\beta=N}^{\infty}v_{-\beta}(s)\,\xi^{-\beta}\right]$$

(the second follows from the fact that the order of the commutator on the right is $\leqslant -1$). The commutator in the first identity is equal to

$$\sum_{\substack{\alpha<N,\beta<N-1\\ \gamma>0}}\left[\binom{\alpha}{\gamma}u_\alpha v_{-\beta}^{(\gamma)}(s)-\binom{-\beta}{\gamma}v_{-\beta}(s)\,u_\alpha^{(\gamma)}\right]\xi^{\alpha-\beta-\gamma}. \tag{7}$$

We sustitute in (7) the result of Proposition 3.10:

$$v_\beta^{(\gamma)}(s)=\sum_{\delta=0}^{\beta-1}\binom{\beta-1}{\delta}(-1)^\delta\partial^{\gamma+\delta}\left(\frac{1}{s+1}\,\frac{\delta}{\delta u_{\beta-1-\delta}}\,v_{-1}(s+1)\right). \tag{8}$$

We wish to represent the formula obtained in the form (cf. Sec. 8, Chap. I):

$$\partial_1 u_k=\sum_{j,l}(B_{j,kl})\,\partial^j\frac{\delta}{\delta u_l}\,\frac{v_{-1}(s+1)}{s+1}-(-\partial)^j\left(B_{j,lk}\frac{\delta}{\delta u_l}\left(\frac{v_{-1}(s+1)}{s+1}\right)\right) \tag{9}$$

with appropriate $B_{j,kl}$. In order to compare (7) and (9), we make in (7) and (8) the change of indices: $\alpha-\beta-\gamma=k,\ \gamma+\delta=j,\ \beta-1-\delta=l$. We then find that the sum of the first terms of (7) leads to the first terms of (9) if we set

$$B_{j,kl}=\begin{cases}\displaystyle\sum_{\delta=0}^{j}(-1)^\delta\binom{k+l+j+1}{j-\delta}\binom{l+\delta}{l}u_{j+k+l+1}=\binom{j+k}{k}u_{j+k+l+1} & \text{for } j+k+l+1\leqslant N\\ 0 \quad \text{for } j+k+l+1>N\end{cases} \tag{10}$$

(we have used one of the identities for the binomial coefficients proved in Sec. 8, Chap. I).

Similarly, the part of the second sum of (7) corresponding to given j,k,l, can be written in the form (for $k+l+j+1\leqslant N$ and zero otherwise):

$$-\sum_{\delta=0}^{j}(-1)^j\binom{j+l}{j-\delta}\binom{l+\delta}{\delta}\partial^\delta\frac{\delta}{\delta u_l}\,\frac{v_{-1}(s+1)}{s+1}\cdot\partial^{j-\delta}u_{k+l+j+1}=-\binom{j+l}{l}(-\partial)^j\left(u_{k+l+j+1}\frac{\delta}{\delta u_l}\,\frac{v_{-1}(s+1)}{s+1}\right).$$

This agrees with (9) and (10) and completes the proof of Theorem 3.2 a). The fact that the operator $\sum B_j\partial^j-(-\partial)^j\circ B_j^t$ is Hamiltonian is proved in Sec. 8, Chap. I.

3.12. Proof of Theorem 3.2 b). We fix r and s and consider the evolution of $v_{-1}(s)$, due to the system $\partial_1 L=[\langle L^r\rangle, L]$. Repeating word-for-word the proof of Lemma 3.8 with δ replaced by ∂_1 (it is only important that $[\partial_1,\partial]=[\partial_1,\partial_\xi]=0$) with the coefficient-wise action $\partial_1:\mathscr{B}((\xi^{-1}))\to\mathscr{B}((\xi^{-1}))$, we find

$$\partial_1 v_{-1}(s)\equiv s\operatorname{res}(\partial_1\tilde{L}\circ\tilde{L}^{\circ(s-1)}) \bmod \operatorname{Im}\partial.$$

Further,

$$\partial_1\tilde{L}\circ\tilde{L}^{\circ(s-1)}=[\langle L^r\rangle\tilde{,}\tilde{L}]\circ\tilde{L}^{\circ(s-1)}=\langle L^r\rangle\tilde{\circ}\tilde{L}^{\circ s}-\tilde{L}\circ\langle L^r\rangle\tilde{\circ}\tilde{L}^{\circ(s-1)}=$$

$$=\langle L^r\rangle\tilde{\circ}\tilde{L}^{\circ s}-[\tilde{L}\circ\langle L^r\rangle\tilde{,}\tilde{L}^{\circ(s-1)}]+\tilde{L}^{\circ s}\circ\langle L^r\rangle^{\sim}=[\langle L^r\rangle\tilde{,}L^{\circ s}]-[\tilde{L}\circ\langle L^r\rangle\tilde{,}L^{\circ(s-1)}].$$

Since the residues of the commutators lie in $\operatorname{Im}\partial$ by Lemma 3.3, we finally obtain $\partial_1 v_{-1}(s)\in\partial\mathscr{B}$. According to the definitions of Sec. 7, Chap. I, this means that $v_{-1}(r)$ and $v_{-1}(s)$ commute.

We shall now clarify the algebraic significance of commutativity. Let $\bar{u}=(u_0,\ldots,u_{N-2})$, let B be some Hamiltonian operator, and let $Q,P\in\mathscr{B}=k[u_i^{(j)}]$. Returning to the conventions of Chap. I, we denote by X_Q the evolution field $X_Q\bar{u}=B\frac{\delta Q}{\delta\bar{u}}$.

3.13. Proposition. If Q, P commute in the Hamiltonian structure with operator B, then X_Q takes the minimal ∂-closed ideal J_P, generated by the components of $B\frac{\delta P}{\delta\bar{u}}$ into itself. In other words, the X_Q flow is tangent to the finite-dimensional manifold of solutions of the system of ordinary differential equations $B\frac{\delta P}{\delta\bar{u}}=0$.

Proof. Since $[\partial,X_Q]=0$, it suffices to verify that $X_Q\frac{\delta P}{\delta u_i}\in J_P$. But $X_Q\frac{\delta P}{\delta\bar{u}}=D\left(B\frac{\delta P}{\delta\bar{u}}\right)B\frac{\delta Q}{\delta\bar{u}}$, and by the Hamiltonian criterion for B the latter expression is equal to $B\frac{\delta}{\delta\bar{u}}\left(\frac{\delta P}{\delta\bar{u}^t}B\frac{\delta Q}{\delta\bar{u}}\right)+D\left(B\frac{\delta Q}{\delta\bar{u}^t}\right)B\frac{\delta P}{\delta\bar{u}}$. The first term is equal to zero, since the commutativity of P, Q implies that $\frac{\delta P}{\delta\bar{u}^t}B\frac{\delta Q}{\delta\bar{u}}\in\operatorname{Im}\partial\subset\operatorname{Ker}\frac{\delta}{\delta\bar{u}}$. The second term (more precisely, its components) lies in J_P, since the Fréchet Jacobian belongs to $M_{n+1}(\mathscr{B}[\partial])$.

This proposition motivates the search for solution (in function rings) of the equations $\bar{u}_t=B\frac{\delta Q}{\delta\bar{u}}$, which remain for all time on the manifolds $B\frac{\delta P}{\delta\bar{u}}=0$, where P is an integral. Actually, instead of the equations $\frac{\delta P}{\delta\bar{u}}\in\operatorname{Ker}B$ the equations $\frac{\delta P}{\delta\bar{u}}=0$ — the extremals of the Lagrangian P — and the flows induced on them are usually investigated (cf. Gel'fand and Dikii [3]). Both formulations of the problem are equivalent if $\operatorname{Ker}B=k^{N-1}$ and in place of the equation $\frac{\delta P}{\delta\bar{u}}=\bar{c}\in k^{N-1}$ it is possible to consider the equation $\frac{\delta}{\delta\bar{u}}(P-\bar{c}^t\bar{u})=0$, noting that $P-\bar{c}^t\bar{u}$ is in this case an integral together with P. In the general case the question merits special investigation.

Regarding the Hamiltonian property for the induced flows see the work of Bogoyavlenskii and Novikov [1] and Gel'fand and Dikii [3].

3.14. We shall apply these considerations to the Lax equations $L_t=\left[\sum c_i\langle L^{s_i}\rangle, L\right]$ and their integrals $\sum d_i\frac{v_{-1}(r_i+1)}{r_i+1}=Q$. By Theorem 3.2 the condition $B\frac{\delta Q}{\delta\bar{u}}=0$, implies that

$[\sum d_i \langle L^{r_i} \rangle, L]=0$. Thus, the search for solutions of the Lax equation lying on the "quasi-extremals" of their integrals $\left(B\frac{\delta P}{\delta u}=0\right.$ in place of $\left.\frac{\delta P}{\delta u}=0\right)$, reduces to the joint solution of the system of equations $L_t=[P, L]$, $[Q, L]=0$. The condition $[Q, L]=0$ is called the auxiliary stationary problem (for the nonstationary problem $L_t=[P, L]$). Chapter III is mainly devoted to the description of such solutions.

3.15. Example: The Korteweg–de Vries Equation. We set $N=2$, $L=\partial^2+u$. The equation $L_t=u_t=[\langle L^{3/2} \rangle, L]$ is called the Korteweg–de Vries equation, while the equations $L_t=[\sum \langle L^{s_i} \rangle c_i, L]$ are its higher analogues. As $s_i \in Q_2$ it suffices to take half integers, since $\langle L^s \rangle = L^s$ for integral s (this remark applies to general N).

We observe that Corollary 3.9 assumes the form $\frac{\delta v_{-1}(s)}{\delta u}=s v_{-1}(s-1)$, so that increasing s by 1 corresponds to "variational integration."

3.16. The Matrix Case. We shall now briefly describe the changes to be made of the definitions and results of 3.1-3.12 in order to extend them to the matrix case. We set $L=\sum_{k=0}^{N} U_k \partial^k$, where $U_N=\operatorname{diag}(c_1, \ldots, c_l)$ is a semisimple matrix of $M_l(k)$, which is diagonal for simplicity and $U_0, \ldots, U_{N-1}$ are matrices with independent variables as elements: $U_i=(u_{i,\alpha\beta})$; $u_{N-1,\alpha\beta}=0$ for α, β with $c_\alpha=c_\beta$. Having chosen an admissible root of an N-th power W of U_N, we can construct the fractional powers $\tilde{L}_W^{\circ s}$, as in Sec. 2 for $s \in Q_N$. We shall further investigate equations of the form $\partial_1 L=\sum c_{i,W} \langle L_W^{s_i} \rangle$. We recall that they now do not exhaust all Lax equations over the ring $\mathcal{B}[\partial]$, $\mathcal{B}=M_l(\mathcal{B}_0)$, $\mathcal{B}_0=k[u_{i,\alpha\beta}^{(j)}]$.

We set $\tilde{L}_W^{\circ s}=\sum V_{-\beta}(s, W)\xi^{-\beta}$. A representation analogous to (7) holds as before but only up to a commutator in the term $\partial_1 U_0$, which dropped out for $l=1$:

$$\sum_{k=0}^{N-1} \partial_1 U_k \xi^k = \left[\sum_{\alpha=0}^{N} U_\alpha \xi^\alpha, \sum_{\beta=1}^{N-1} \sum_i c_{i,W} V_{-\beta}(s_i, W)\xi^{-\beta}\right] + \left[U_N, \sum_i c_{i,W} V_{-N}(s_i, W)\right].$$

The contribution from this commutator will vanish if we go over to equations for the matrix traces of the U_k in place of the U_k themselves. This becomes even more necessary if we wish to carry over to the matrix case at least a part of the results regarding the Hamiltonian property. Indeed, the key Lemma 3.3 ceases to hold, since the matrix $VU^{(m+n+1)}-(-1)^{m+n+1}\cdot UV^{(m+n+1)}$ no longer need be a total derivative. However, this expression differs from a total derivative by the commutator $[U, V^{(m+n+1)}]$, and therefore after going over to traces we again obtain $\operatorname{Tr}\operatorname{res}[a, b] \in \partial\mathcal{B}$. In the considerations of 3.4 and thereafter in place of $\Omega^1(\mathcal{B})$ it is now necessary to take the $\mathcal{B}[\partial]$-module $M_l(\Omega^1(\mathcal{B}_0))$; $\delta: \mathcal{B} \to M_l(\Omega^1(\mathcal{B}_0))$ is defined coefficient-wise, and the subsequent constructions are modified in an obvious manner. In Lemma 3.5 we must represent δP in the form $\sum Q_{i,\alpha\beta}\delta u_{i,\alpha\beta}+\partial\omega$, where $Q_{i,\alpha\beta}=\frac{\delta P}{\delta u_{i,\alpha\beta}}$ (coefficient-wise). The proof of the analogue of Lemma 3.8 leads to the result $\operatorname{Tr}\delta V_{-1}(s, W) \equiv s \operatorname{Tr}\operatorname{res}\delta\tilde{L}\circ\tilde{L}^{\circ(s-1)} \bmod \operatorname{Im}\partial$.

Analogues of Corollary 3.9 and Proposition 3.10 now connect $\frac{\delta}{\delta u_{l,\alpha\beta}} \operatorname{Tr} V_{-1}(s, W)$ with $\operatorname{Tr} V_j(s-1, W)$. Substitution of expressions analogous to (4) into the equations for $\operatorname{Tr}\partial_1 U_k$ leads to an assertion that these equations be "quasi-Hamiltonian" which is described in detail in the work of Gel'fand and Dikii [5]. We refer the reader to that work for further details.

4. The Stationary Lax Equations

4.1. In this section the first general result on stationary equations will be proved: if $L, P \in \mathcal{B}[\partial]$ are such that $[L, P]=0$, then (under weak additional conditions) L and P are connected by a polynomial relation with constant coefficients. This reduces the problem of solving the stationary Lax equations to the problem of imbedding one-dimensional commutative rings (i.e., rings of functions on affine algebraic curves) in rings of differential operators. The fruitfulness of such a reduction will be demonstrated in Chap. III.

4.2. We shall actually prove a more general theorem pertaining to rings of operators generated by several differentiations $\partial_1, \ldots, \partial_n$.

Let $\mathcal{B}$ be a not necessarily commutative Q-algebra which is free of finite rank as a module over its center $\mathcal{B}_0$. Let $\partial_1, \ldots, \partial_n : \mathcal{B} \to \mathcal{B}$ be differentiations taking $\mathcal{B}_0$ into itself and commuting pairwise. We write $i=(i_1, \ldots, i_n)$, $|i|=\sum_{j=1}^{n} i_j$, $\partial^i=\partial_1^{i_1} \ldots \partial_n^{i_n}$, $\mathcal{B}[\partial]=\left\{\sum b_i\partial^i \mid b_i \in \mathcal{B}\right\}$ with the usual rules of multiplication.

Let $L_1, \ldots, L_r \in \mathcal{B}[\partial]$ be a finite family of operators. We call it independent if the sum of $\mathcal{B}$-modules $\sum \mathcal{B}L_1^{p_1} \ldots L_r^{p_r}$ is direct. We write $L^p=L_1^{p_1} \ldots L_r^{p_r}$ and $|p|=\sum_{j=1}^{r} p_j$.

4.3. LEMMA. We assume that in the category of free $\mathcal{B}$-modules the concept of rank is well defined (i.e., $\mathcal{B}^r \cong \mathcal{B}^s \Rightarrow r=s$), and the rank of a submodule does not exceed the rank of the module. If $L_1, \ldots, L_r$ is an independent family of operators, then $r \leqslant n$.

Proof. There exists a constant l, depending on the degrees of the operators $L_1, \ldots, L_r$, such that for all $m \geqslant 0$ there is the following imbedding of free $\mathcal{B}$-modules: $\sum_{|p|<m} \mathcal{B}L^p \subset \sum_{|i|<lm} \mathcal{B}\partial^i$. But the $\mathcal{B}$-rank of the module on the left grows asymptotically like $c_1 m^r$ and that of the module on the right like $c_2 (lm)^n$. Therefore, $r \leqslant n$.

4.4. We shall call a family of operators $L_1, \ldots, L_n \in \mathcal{B}[\partial]$ maximal if it is independent and there exists a free $\mathcal{B}$-module of finite rank $M \subset \mathcal{B}[\partial]$, such that $\mathcal{B}[\partial]=\sum_i ML^i$, and the sum on the right is direct.

Example: in the case $n=1$ the family consisting of a single operator $b_N\partial^N + \ldots + b_0$, is maximal if the coefficient b_N is invertible in $\mathcal{B}$. Indeed, we set $M=\sum_{i=0}^{N-1} \mathcal{B}\partial^i$. Since $\mathcal{B}b_N^p\partial^{pN+i}=\mathcal{B}\partial^{pN+i}$, it follows easily that $\mathcal{B}[\partial]=\sum_{p=0}^{\infty}\left(\sum_{i=0}^{N-1} \mathcal{B}\partial^i\right) L^p$ and the sum on the right is direct.

4.5. THEOREM. Let $L_1, \ldots, L_n$ be a maximal family of operators in $\mathcal{B}[\partial]$ and let $P \in \mathcal{B}[\partial]$ be an operator such that $[P, L_1]= \ldots =[P, L_n]=0$. Then there exist ∂-constants $a_{i,q}$, $i=(i_1, \ldots, i_n)$, $q \geqslant 0$, lying in the center of $\mathcal{B}$ and not all zero such that $\sum a_{i,q} L^i P^q=0$.

We first prove the following auxiliary assertion. Let $\lambda_1, \ldots, \lambda_n$ be independent variables over $\mathcal{B}[\partial]$ which commute with one another and with $\mathcal{B}[\partial]$. We choose a module $M \subset \mathcal{B}[\partial]$, as in 4.4 for the family $L_1, \ldots, L_n$, and set

$$M[\lambda_1, \ldots, \lambda_n] = \left\{ \sum m_i \lambda^i \mid m_i \in M \right\} \subset \mathcal{B}[\partial][\lambda_1, \ldots, \lambda_n].$$

4.6. LEMMA. There is the direct sum decomposition

$$\mathcal{B}[\partial](\lambda_1, \ldots, \lambda_n) = M[\lambda_1, \ldots, \lambda_n] \oplus \sum_{i=1}^{n} \mathcal{B}[\partial][\lambda_1, \ldots, \lambda_n](L_i - \lambda_i), \tag{11}$$

where the last expression on the right is the left $\mathcal{B}[\partial][\lambda_1, \ldots, \lambda_n]$-ideal generated by $L_i - \lambda_i$.

Proof. a) We show first that the left side is the sum on the right. Since $\mathcal{B}[\partial] = \sum_j ML^j$, it suffices to verify that for any $m \in M$ and j the element mL^j is contained in the sum on the right. But $mL^j = m\lambda^j + m(L^j - \lambda^j)$, so that it remains to check that $L^j - \lambda^j$ lies in the left ideal generated by $L_i - \lambda_i$. Let $j = (j_1, \ldots, j_n)$ and let k be the largest index for which $j_k \neq 0$. We set $j' = (j_1, \ldots, j_{k-1}, 0, \ldots, 0)$. Then

$$L^j - \lambda^j = L^{j'} L_k^{j_k} - \lambda^{j'} \lambda_k^{j_k} = L^{j'} L_k^{j_k} - L^{j'} \lambda_k^{j_k} + L^{j'} \lambda_k^{j_k} - \lambda^{j'} \lambda_k^{j_k} =$$
$$= L^{j'} (L_k^{j_k} - \lambda_k^{j_k}) + \lambda_k^{j_k} (L^{j'} - \lambda^{j'}).$$

The first term on the right is divisible by $L_k - \lambda_k$, while the second is analogous to $L^j - \lambda^j$, but the index j' has fewer nonzero components than j. Therefore, induction on the number of components gives the required results.

b) We now show that the sum on the right of (11) is direct. For this it suffices to verify that if $\sum_i m_i \lambda^i \in \sum_i \mathcal{B}[\partial][\lambda](L_i - \lambda_i)$, then $m_i = 0$ for all i. But if $\sum_i m_i \lambda^i \in \sum_i \mathcal{B}[\partial][\lambda](L_i - \lambda_i)$, then according to the first part of the proof also $\sum m_i L^i \in \sum \mathcal{B}[\partial][\lambda](L_i - \lambda_i)$. We consider the homomorphism of left $\mathcal{B}[\partial]$-modules $\mathcal{B}[\partial][\lambda] \to \mathcal{B}[\partial]$, which is the identity on $\mathcal{B}[\partial]$ and takes the free generator $\lambda_1^{i_1} \ldots \lambda_n^{i_n}$ into $L_1^{i_1} \ldots L_n^{i_n}$. The submodule $\sum \mathcal{B}[\partial][\lambda](L_i - \lambda_i)$, clearly lies in the kernel of this homomorphism. Therefore, it as well as the entire kernel has zero intersection with $\mathcal{B}[\partial]$.

4.7. Proof of Theorem 4.5. We consider the $\mathcal{B}[\lambda]$-module $\overline{M} = \mathcal{B}[\partial][\lambda] \big/ \sum_i \mathcal{B}[\partial][\lambda](L_i - \lambda_i)$. According to Lemma 4.6, it is free of finite rank over $\mathcal{B}[\lambda]$, since it is isomorphic to $M[\lambda]$. It is therefore free of finite rank over $\mathcal{B}_0[\lambda]$, where $\mathcal{B}_0$ is the center of the ring $\mathcal{B}$. Since the operator $P \in \mathcal{B}[\partial]$ commutes with all the L_i, multiplication by P on the right induces a $\mathcal{B}_0[\lambda]$-endomorphism of the module $\overline{M}$. Therefore, the ring $\mathcal{B}_0[\lambda_1, \ldots, \lambda_n, P] \subset \mathcal{B}[\partial]$ acts naturally on $\overline{M}$. On representatives this action can be written as follows:

$$\overline{m} \circ b \lambda^i P^j = b \overline{m} \lambda^i P^j, \ b \in B_0, \ \overline{m} \in \overline{M}.$$

Since $\overline{M}$ is free over $\mathcal{B}_0[\lambda_1, \ldots, \lambda_n]$ there is a nonzero polynomial $F(\lambda_1, \ldots, \lambda_n, P) \in \mathcal{B}_0[\lambda_1, \ldots, \lambda_n, P]$ such that $\overline{M} \circ F = 0$. According to Lemma 4.6, this implies that $\mathcal{B}[\partial][\lambda] \circ F(\lambda_1, \ldots, \lambda_n, P) \subset$

$\sum_i \mathcal{B}[\partial][\lambda](L_i-\lambda_i)$. From the calculations in the first part of the proof of this lemma it follows that then $\mathcal{B}[\partial][\lambda]\circ F(L_1,\ldots,L_n,P)=0$ where the substitution of L_i in place of λ_i must be done in the reduced notation for F . Applying $F(L_1,\ldots,L_n,P)$ to $1\in\mathcal{B}[\partial][\lambda]$, we find that $F(L_1,\ldots,L_n,P)\in\mathcal{B}[\partial]\cap\sum_i\mathcal{B}[\partial][\lambda](L_i-\lambda_i)=\{0\}$, by the second part of the proof of Lemma 4.6.

We have thus obtained a polynomial relation between L_i and P but with coefficients in $\mathcal{B}_0$. We shall now show that for suitable F these coefficients are ∂-constants.

To this end we choose a free basis of the $\mathcal{B}_0[\lambda]$-module $\overline{M}$ and denote by D a differentiation $\mathcal{B}_0[\lambda]\to\mathcal{B}_0[\lambda]$ and its lift $\overline{M}\to\overline{M}$ with the following properties:

a) On $\mathcal{B}$ the differentiation D coincides with ∂.

b) $D\lambda_i=0,\ i=1,\ldots,n.$

c) The chosen basis of $\overline{M}$ is annihilated by D.

We denote by Λ the matrix of the endomorphism of multiplication by P in this basis and by F its characteristic polynomial.

In addition to D, there is the differentiation $\bar\partial:\overline{M}\to\overline{M}$, induced by multiplication on the left by ∂ in $\mathcal{B}[\partial][\lambda]$. It also extends ∂ . Therefore, $D-\bar\partial:\overline{M}\to\overline{M}$ is a $\mathcal{B}_0[\lambda]$ -linear operator:

$$(\bar\partial-D)(bm)=(\bar\partial-D)\,b\cdot m+b\,(\bar\partial-D)\,m=b\,(\bar\partial-D)\,m.$$

We set $U=\bar\partial-D$ and identify U with its matrix in the chosen basis. Since right multiplication by P commutes with left multiplication by ∂ , on calculating the action on this basis of the composition of ∂ and P in two different ways, we obtain $\partial\Lambda+\Lambda U=U\Lambda$ or $\partial\Lambda=[U,\Lambda]$. This implies that $\partial\Lambda^n=\sum_{i=0}^{n-1}\Lambda^i\partial\Lambda\Lambda^{n-i-1}=\sum_{i=0}^{n-1}\Lambda^i[U,\Lambda]\Lambda^{n-i-1}=[U,\Lambda^n]$. Therefore, the trace of $\partial\Lambda^n$ is zero for all $n\geqslant 0$. This means that the coefficients of the characteristic polynomial of Λ are ∂-constant by Newton's formula.

4.8. Integrals of the Stationary Lax Equations. We apply the preceding considerations to the case $\mathcal{B}=k\left[u_0^{(j)},\ldots,u_{N-2}^{(j)}\mid j\geqslant 0\right]$, $L=\partial^N+\sum_{i=0}^{N-2}u_i\partial^i$, $P=\sum c_i\langle L^{s_i}\rangle$. We denote by $\overline{\mathcal{B}}=\overline{\mathcal{B}}_{(c_i)}$ the factor ring of $\mathcal{B}$ by the ∂-closed ideal generated by $[L,P]$. Let $\bar\partial:\overline{\mathcal{B}}\to\overline{\mathcal{B}}$ be the induced differentiation where $\overline{L},\overline{P}\in\overline{\mathcal{B}}[\bar\partial]$ are the images of the operators L and P. It is obvious that $\overline{L},\overline{P}$ is a "general solution" of the stationary Lax equations with operators of the given degrees N,M, if the constants $c_0,\ldots,c_M$ are taken to be free variables and $s_i=i/N$.

Since the leading coefficient of $\overline{L}$ is equal to 1, $\overline{L}$ generates a maximal family, and $\overline{L}$, $\overline{P}$ are connected by the relation $\sum\bar d_{ij}L^iP^j=0$, where $\bar d_{ij}$ are $\bar\partial$-constants in $\overline{\mathcal{B}}$. There is a canonical relation: $\sum_j\bar d_{ij}\lambda^j$ are the coefficients of the characteristic polynomial of the endomorphism of multiplication by $\overline{P}$ for the $\overline{\mathcal{B}}[\lambda]$-module $\overline{\mathcal{B}}[\bar\partial][\lambda]/\overline{\mathcal{B}}[\bar\partial][\lambda](\overline{L}-\lambda)$. The lifts d_{ij} of the elements $\bar d_{ij}$ to $\mathcal{B}$ are integrals of the equation $[L,P]=0$.

We note now that by Proposition 3.13 we have another collection of integrals of this equation. Namely, for any $r \in Q_N$, $r \geqslant 0$ we have $\frac{\delta v_{-1}(r)}{\delta \bar{u}^t} B \frac{\delta}{\delta u} \sum c_i \frac{v_{-1}(s_i+1)}{s_i+1} = \partial I_r$, $I_r \in \mathcal{B}$ so that $\partial \bar{I}_r = 0$ in $\bar{\mathcal{B}}$.

The relation between the integrals d_{ij} and I_r is known for the higher Korteweg–de Vries equations; in particular, the ones are expressed in terms of the others.

5. The Zakharov–Shabat Formalism

5.1. In this section we present an initial version of the method of Zakharov–Shabat [13]. Our objective here is to clarify the algebraic side of their construction — the structure of "dressed" differential operators — while omitting functional-analytic considerations. We begin, however, by describing the basic functionals.

5.2. We consider the space of columns ψ of height N, the coordinates of which are functions of the variable x and possibly of the additional parameters t and z. Let $K(x, y; t, z)$ and $F(x, y; t, z)$ be two $(N \times N)$ matrix-valued functions. We assign to them the integral operators $\hat{K}, \hat{F}$:

$$(\hat{K}\psi)(x) = \int_x^\infty K(x, y)\psi(y)\,dy, \quad (\hat{F}\psi)(x) = \int_{-\infty}^\infty F(x, y)\psi(y)\,dy.$$

Similarly, for any operator L we understand by $(LK)\hat{}$ the operator with kernel LK and integration from x to ∞, and by $(LF)\hat{}$ the analogous operator with integration from $-\infty$ to ∞. The dependence on t, z is not indicated here explicitly. The functions K, F, and ψ are assumed such that all classical formulas used below for differentiation of integrals with respect to a parameter are valid (and, of course, the integrals themselves exist). The following sequence of lemmas on the commutation of various operators is preparatory to the formulation of the main theorem.

5.3 LEMMA. $[\partial_t, \hat{F}] = (\partial_t F)\hat{}$, $[\partial_t, \hat{K}] = (\partial_t K)\hat{}$, and similarly for ∂_z.

The proof is obvious.

Let $L_x = \Sigma l_i(x; t, z)\partial_x^i$, where the l_i are $(N \times N)$ matrices. We denote by the symbol FL_y^+ the kernel $\sum_i (-\partial_y)^i (F(x, y) l_i(x))$ and similarly for KL_y^+.

5.4. LEMMA. $[L_x, \hat{F}] = \{L_x F - FL_y^+\}\hat{}$.

Proof. Integrating by parts, we have

$$([L_x, \hat{F}]\psi)(x) = L_x \int_{-\infty}^\infty F(x, y)\psi(y)\,dy - \int_{-\infty}^\infty F(x, y) L_y \psi(y) =$$

$$= \int_{-\infty}^\infty L_x F(x, y)\psi(y)\,dy - \int_{-\infty}^\infty FL_y^+(x, y)\psi(y)\,dy = \{L_x F - FL_y^+\}\hat{\psi}.$$

5.5. LEMMA. If L_x satisfies the conditions of the preceding lemma, then

$$L_x \circ \hat{K} - (L_x K)^{\wedge} = \sigma_K(L_x),$$

where the differential operator $\sigma_K(L_x)$ is additive in L_x and for $l = l(x, t, z)$ is defined by

$$\sigma_K(l\partial_x^n)\psi = -\sum_{i=0}^{n-1} l\partial_x^i [(\partial_x^{n-1-i}K)(x, x)\psi].$$

Proof. It is clear that the left and right sides are linear on matrix-valued functions of x, t, z. It therefore suffices to compute $\sigma_K(\partial_x^n)$. For $n=0$ the formula is obvious. From n to $n+1$ the computation is as follows:

$$\sigma_K(\partial_x^{n+1})\psi = \partial_x^{n+1}\int_x^\infty K(x, y)\psi(y)\,dy - \int_x^\infty (\partial_x^{n+1}K)(x, y)\psi(y)\,dy =$$

$$= \partial_x\left[\int_x^\infty (\partial_x^n K)(x, y)\psi(y)\,dy + \sigma_K(\partial_x^n)\psi\right] - \int_x^\infty (\partial_x^{n+1}K)(x, y)\psi(y)\,dy = -(\partial_x^n K)(x, x)\psi(x) + \partial_x \sigma_K(\partial_x^n)\psi =$$

$$= -(\partial_x^n K)(x, x)\psi(x) - \sum_{i=0}^{n-1} \partial_x^{i+1}[(\partial_x^{n-1-i}K)(x, x)\psi(x)].$$

The last expression coincides with the formula for $\sigma_K(\partial_x^{n+1})\psi$, indicated in the lemma.

5.6. LEMMA. We assume that the coefficients of L_x do not depend on x. Then

$$\hat{K} \circ L_x - (K L_y^{+})^{\wedge} = \tau_K(L_x),$$

where τ_K is additive in L_x and

$$\tau_K(l\partial_x^n) = \sum_{i=0}^{n-1} (-1)^{n-i} (\partial_y^{n-i-1}K)(x, x)\, l\partial_x^i.$$

Proof. As above, for $n=0$ the assertion is obvious. We carry out the inductive step from n to $n+1$. Splitting the second integrals into parts, we have:

$$[\hat{K} \circ l\partial_x^{n+1} - K(l\partial_y^{n+1})^{+}]\psi(x) = \int_x^\infty K(x, y)\, l\partial_y^{n+1}\psi(y)\,dy -$$

$$-\int_x^\infty (-1)^{n+1}(\partial_y^{n+1}K)(x, y)\, l\psi(y)\,dy =$$

$$= \int_x^\infty K(x, y)\, l\partial_y^{n+1}\psi(y)\,dy + (-1)^{n+1}(\partial_y^n K)(x, x)\, l\psi(x) + \int_x^\infty (-1)^{n+1}(\partial_y^n K)(x, y)\, l\partial_y\psi(y)\,dy.$$

Here we have used the fact that $\partial_y(l\psi) = l\partial_y\psi$ since $l = l(t, z)$. The sum of the two integrals here is equal to

$$[\hat{K} \circ l\partial_x^n - K(l\partial_y^n)^{+}]\partial_x\psi = \tau_K(l\partial_x^n)\partial_x\psi$$

by the inductive hypothesis. The entire expression is therefore

$$(-1)^{n+1}(\partial_y^n K)(x, x)\, l\partial_x\psi + \tau_K(l\partial_x^n)\partial_x\psi = \tau_K(l\partial_x^{n+1})\psi.$$

5.7. We now consider the ring $\mathcal{B}[\partial_x]$ of matrix-valued operators in ∂_x, the coefficients of which are polynomials of the restrictions to the diagonal of the partial derivatives $\partial_x^\alpha K$, $\partial_y^\beta K$ and the derivatives of these restrictions along the diagonal. Let $L_x \in M_N(k)[\partial_x]$ (k consists of the constants, i.e., $\mathbf{R}$ or $\mathbf{C}$ in the analytic case). Then $(1+\tau_K)L_x$ has the same order as L_x, and its coefficients depend only on x, t, z. It is therefore permissible to apply σ_K and its powers to it while remaining in $\mathcal{B}[\partial_x]$. But the operator $\sigma_K:\mathcal{B}[\partial_x]\to\mathcal{B}[\partial_x]$, as is evident from Lemma 5.5, reduces the order. Therefore, the following expression is meaningful:

$$L_x^* = (1+\sigma_K)^{-1}\tau_K L_x = (1-\sigma_K+\sigma_K^2-\ldots+(-1)^M\sigma_K^M)\,\tau_K L_x \in \mathcal{B}[\partial_x],$$

where $M > \operatorname{ord} L_x$.

The operator L_x^* is called the "dressed" operator L_x, and the mapping $L_x \mapsto (1+\sigma_k)^{-1}\tau_K L_x$ is called the "raiment" of L_x.

The first part of the next theorem shows that L_x^* is roughly speaking the differential part of the operator obtained from L_x by means of conjugation with the operator $1+\hat{K}$:

5.8. THEOREM. a) $L_x^*(1+\hat{K})-(1+\hat{K})L_x = \{L_x^* K - K L_y^+\}^\wedge$. b) We assume that the functions K, F satisfy the equation

$$K(x,y)+F(x,y)+\int_x^\infty K(x,s)F(s,y)\,ds=0$$

or, more briefly, $K+F+K*F=0$. There is then the identity

$$(L_x^* K - K L_y^+)+(L_x F - F L_y^+)+(L_x^* K - K L_y^+)*F + K*(L_x F - F L_y^+)=0.$$

c) Under this same condition we have

$$\partial_t K + \partial_t F + \partial_t K * F + K * \partial_t F = 0$$

and similarly for ∂_z.

5.9. Application of Theorem 5.8. If $K+F+K*F=0$ and $RK+SF+RK*F+K*SF=0$ for some operators R, S, we shall say that the pair (R,S) "differentiates" (K,F). According to Theorem 5.8, the pair $(L_x^* - L_y^+ + \alpha\partial_t + \beta\partial_z,\, L_x - L_y^+ + \alpha\partial_t + \beta\partial_z)$ differentiates (K,F), if α, β are any constants.

We now suppose that the kernel F satisfies the system of linear differential equations

$$\begin{cases} L_x F - F L_y^+ + \alpha\partial_t F = 0, \\ L_x F - F L_y^+ + \beta\partial_z F = 0. \end{cases}$$

We find by Theorem 5.8 b) and c) that

$$\begin{cases} L_x^* K - K L_y^+ + \alpha\partial_t K + (L_x^* K - K L_y^+ + \alpha\partial_t K)*F = 0, \\ L_x^* K - K L_y^+ + \beta\partial_z K + (L_x^* K - K L_y^+ + \beta\partial_z K)*F = 0. \end{cases}$$

We suppose, moreover, that from this it can be concluded that

$$\begin{cases} L_x^* K - K L_y^+ + \alpha \partial_t K = 0, \\ L_x^* K - K L_y^+ + \beta \partial_z K = 0. \end{cases}$$

Using Theorem 5.8 a) and Lemma 5.3, we obtain, on the other hand,

$$(L_x^* + \alpha \partial_t)(1 + \hat{K}) - (1 + \hat{K})(L_x + \alpha \partial_t) = \{L_x^* K - K L_y^+ + \alpha \partial_t K\}^{\wedge} = 0,$$
$$(L_x^* + \beta \partial_z)(1 + \hat{K}) - (1 + \hat{K})(L_x + \beta \partial_z) = \{L_x^* K - K L_y^+ + \beta \partial_z K\}^{\wedge} = 0.$$

Assuming further that $1 + \hat{K}$ is invertible, we find that the differentiation operators $L_x^* + \alpha \partial_t$ and $L_x^* + \beta \partial_z$ are adjoint to $L_x + \alpha \partial_t$ and $L_x + \beta \partial_z$, respectively.

We now take the initial matrix operators $L_{1,x}$ and $L_{2,x}$ to have constant coefficients. We obtain the following result.

5.10. THEOREM. If $[L_{1x}, L_{2x}] = 0$ and F, K satisfy the conditions of 5.9, then L_{1x}^*, L_{2x}^* satisfy the equations of Zakharov–Shabat

$$[L_{1x}^* + \alpha \partial_t, L_{2x}^* + \beta \partial_z] = 0,$$

or

$$\alpha \partial_t L_{2x}^* - \beta \partial_z L_{1x}^* = [L_{2x}^*, L_{1x}^*].$$

We proceed to the proof of Theorem 5.8.

5.11. Proof of Theorem 5.8 a). According to Lemmas 5.5 and 5.6

$$L_x^* \circ \hat{K} = (L_x^* K)^{\wedge} + \sigma_K (L_x^*),$$
$$\hat{K} \circ L_x = (K L_y^+)^{\wedge} + \tau_K (L_x).$$

Therefore,

$$L_x^* (1 + \hat{K}) - (1 + \hat{K}) L_x = \{L_x^* K - K L_y^+\}^+ + L_x^* - L_x + \sigma_k (L_x^*) - \tau_K (L_x).$$

But $L_x^* = (1 + \sigma_k)^{-1} (1 + \tau) L_x$. Hence

$$L_x^* - L_x - \sigma_K L_x^* + \tau_K L_x = [(1 + \sigma_K)^{-1} (1 + \tau_K) - 1 + \sigma_K (1 + \sigma_K)^{-1} (1 + \tau_K) - \tau_K] L_x = 0.$$

5.12. Proof of Theorem 5.8 b). Applying the operators of Lemmas 5.5 and 5.6 to the identity $K + F + K * F = 0$, we obtain

$$L_x K + L_x F + (L_x K) * F + \sigma_K (L_x) F = 0, \tag{12}$$

$$K L_y^+ + F L_y^+ + K * F L_y^+ = 0. \tag{13}$$

By Lemma 5.6 applied to the columns of F, we have

$$K * F L_y^+ = -K * (L_x F - F L_y^+) + K * L_x F = -K * (L_x F - F L_y^+) + K L_y^+ * F + \tau_K (L_x) F.$$

Substituting this expression into (13) and subtracting the result from (12), we find

$$L_xK-KL_y^{+}+L_xF-FL_y^{+}+(L_xK-KL_y^{+})*F+K*(L_xF-FL_y^{+})+(\sigma_K-\tau_K)(L_x)F=0. \quad (14)$$

We set $L_x^{(0)}=L_x$, $L_x^{(i)}=\sigma_K^{i-1}(\sigma_K-\tau_K)L_x$ for $i\geqslant 1$ and apply (12) to $L_x^{(i)}$ in place of $L_x^{(0)}$. We obtain

$$L_x^{(i)}K+L_x^{(i)}F+(L_x^{(i)}K)*F+\sigma_K(L_x^{(i)})F=0. \quad (14)_i$$

We take the alternating sum of Eqs. $(14)_i$ with signs $(-1)^i$ over $i=0,\ldots,M$, $M>\operatorname{ord}L_x$. Noting that

$$\sum_{i=0}^{M}(-1)^iL_x^{(i)}=L_x-(\sigma_K-\tau_K)L_x+\sigma_K(\sigma_K-\tau_K)L_x-\ldots=L_x-(1+\sigma_K)^{-1}(\sigma_K-\tau_K)L_x=L_x^{*},$$

we obtain

$$L_K^{*}K-KL_y^{+}+L_xF-FL_y^{+}+(L_x^{*}K-KL_y^{+})*F+K*(L_xF-FL_y^{+})=0,$$

which completes the proof of the theorem.

5.13. Formal Analogues. Under appropriate analytic assumptions we have

$$K(x,y)=\sum\frac{1}{i!}(\partial_y^iK)(x,x)(y-x)^i$$

and

$$\hat{K}\psi=\sum(\partial_y^iK)(x,x)\int_x^{\infty}\frac{(y-x)^i}{i!}\psi(y)\,dy.$$

It is possible to eliminate the factor $\frac{(y-x)^i}{i!}$ under the integral by integrating by parts i times. Therefore, setting $(\partial_y^iK)(x,x)=u_i$, it is natural to assign to the Volterra operator $\hat{K}$ its symbol

$$\hat{K}\mapsto\sum_{i=1}^{\infty}(-1)^iu_i\xi^{-i}=\tilde{K}.$$

Here the u_i may be considered arbitrary elements of some differential ring $\mathcal{B}$, or matrices with independent coefficients as in 3.16.

In order to write formulas for the operators σ_K, τ_K, it is useful to introduce further the generator v_i, corresponding to $(\partial_x^iK)(x,x)$. It is not hard to see that they are linearly expressed in terms of the u_i and their derivatives (using the formula $((\partial_x+\partial_y)^jK)(x,x)=\partial_x^j(K(x,x)))$:

$$v_i=\sum_{j=0}^{i}\binom{j}{i}(-1)^{j-i}u_{j-i}^{(i)}.$$

Comparison with formulas (3) and (4) indicates that conversely the u_i are expressed in terms of the v_i . In these generators we obtain

$$\sigma_K:\mathcal{B}[\partial]\to\mathcal{B}[\partial]:l\partial^n\mapsto -\sum_{i=0}^{n-1} l\partial^i\circ v_{n-1-i},$$

$$\tau_K:\mathcal{B}[\partial]\to\mathcal{B}[\partial]:l\partial^n\mapsto \sum_{i=0}^{n-1}(-1)^{n-1-i}u_{n-1-i}l\partial^i.$$

Conjugation by means of $\tilde{K}=1+\sum_{i=1}^{\infty}(-1)^i u_i\xi^{-i}$ takes commuting operators in $(\mathcal{B}((\xi^{-1})),\circ)$ into commuting operators:

$$[L_1, L_2]=0\Rightarrow[\tilde{K}^{-1}L_1,\ \tilde{K},\ \tilde{K}^{-1}L_2\tilde{K}]=0$$

(we here identify L_1, and L_2 with their symbols). Generally speaking, the operators $\tilde{K}^{-1}L_i\tilde{K}$ are not purely differential operators. Considering only their differential parts $\langle \tilde{K}^{-1}L_i\tilde{K}\rangle$, we find

$$\operatorname{ord}[\langle \tilde{K}^{-1}L_1\tilde{K}\rangle,\ \langle \tilde{K}^{-1}L_2\tilde{K}\rangle]\leqslant \max\operatorname{ord}L_i-1.$$

Thus, this method leads to the construction of certain "general" solutions of the equations of Zakharov—Shabat.

In conclusion, we shall clarify which differential operators are conjugate by way of some $\tilde{K}$.

Now let $\sum_{i=0}^{N} u_i\xi^i,\ \sum_{i=0}^{N} w_i\xi^i\in\mathcal{B}[\xi]$ be two symbols (u_i, w_i are any elements). We shall assume that $\partial u_N=0$, u_N is semisimple in the sense of Sec. 2, and that $\partial:\mathcal{B}^+\to\mathcal{B}^+$ is surjective.

5.14. THEOREM. For the existence of a symbol $1+\sum_{k=1}^{\infty} v_{-k}\xi^{-k}$ with the property

$$\sum_{i=0}^{N} u_i\xi^i\circ\left(1+\sum_{k=1}^{\infty} v_{-k}\xi^{-k}\right)=\left(1+\sum_{k=1}^{\infty} v_{-k}\xi^{-k}\right)\circ\sum_{i=0}^{\infty} w_i\xi^i$$

it is necessary and sufficient that $u_N=w_N$ and $u_{N-1}-w_{N-1}\in\mathcal{B}^-$.

Proof. We rewrite the last equation in the form

$$\sum_{\substack{\alpha\geqslant 0,k\geqslant 1\\ i=0,\dots,N}}\left[\binom{l}{\alpha}u_i v_{-k}^{(\alpha)}-\binom{-k}{\alpha}v_{-k}w_i^{(\alpha)}\right]\xi^{l-k-\alpha}=\sum_{j=0}^{N}(w_j-u_j)\xi^j.$$

We call the k-th equation (for v) the equality for the coefficients of ξ^k on the left and right sides.

The N-th equation gives $u_N=w_N$; the $(N-1)$-st equation has the form $[u_N, v_{-1}]=w_{N-1}-u_{N-1}$. This shows the necessity of the condition of the theorem and uniquely determines v_{-1}^-. The rest of the argument is parallel to the proof of Theorem 2.10. We suppose that from the equations with indices $N,\dots,N_{-k+1}$ the quantities $v_{-1},\dots,v_{-k+1}^-$ have been determined. Then the $(N-k)$-th equation uniquely determines the element

$$u_{N-1}v_{-k+1}-v_{-k+1}w_{N-1}+Nu_Nv'_{-k+1}+[u_N, v_{-k}]=w.$$

Since $w_{N-1}=u_{N-1}+x$, $x\in\mathcal{B}^-$, the first two terms can be rewritten in the form $[u_{N-1}, v_{-k+1}]-xv_{-k+1}$. For determining v^{+}_{-k+1} we obtain the equation

$$Nu_N(v^{+}_{-k+1})'-(xv^{-}_{-k+1})=w^{+},$$

which is solvable if $\partial:\mathcal{B}^+\to\mathcal{B}^+$ is surjective.

6. The Benney Equations: Main Results

6.1. We recall that the system of Benney equations has the form

$$\begin{cases} u_t+uu_x-u_y\int_0^y u_x|_{y=\eta}d\eta+h_x=0, \\ h_t+\left(\int_0^h udy\right)_x=0. \end{cases} \tag{15}$$

The meaning of the notation is explained in Sec. 1 of this chapter. The formal investigation of the equations is based on the following lemma of Benney.

6.2. LEMMA. We define the moments $A_n(x, t)=\int_0^h u(x, y, t)^n dy$, $n\geqslant 0$. From system (15) there then follows an infinite system of equations for the moments:

$$A_{n,t}+A_{n+1,x}+nA_{n-1}A_{0,x}=0, \quad n\geqslant 0. \tag{16}$$

Proof. Multiplying the first equation of (15) by nu^{n-1} and regrouping terms, we obtain without difficulty

$$(u^n)_t+(u^{n+1})_x-\left(u^n\int_0^y u_x d\eta\right)_y+nu^{n-1}h_x=0.$$

We now integrate this relation on y from 0 to h. The fourth term becomes $nA_{n-1}A_{0,x}$, and the third [because of the second equation of (15)] becomes $u^n(h_t+uh_x)|_{y=h}$. Adding the first term of the last relation to $\int_0^h (u^n)_t dy$, and the second to $\int_0^h (u^{n+1})_x dy$, we obtain $A_{n,t}$ and $A_{n+1,x}$, respectively.

In Benney's work it is shown that there exist two sequences of polynomials $H_n\in Q[A_0, \ldots, A_n]$ and $F_n\in Q[A_0, \ldots, A_{k+1}]$ such that (16) implies local conservation laws for the system (15) of the form

$$H_{n,t}+F_{n,x}=0, \quad n>0. \tag{17}$$

In Miura's work [14] it is shown that there exist two further sequences of polynomials $\bar{H}_n, \bar{F}_n\in Q[u, A_0, \ldots, A_{n-1}]$ such that (16) implies local conservation laws of the form

$$\overline{H}_{n,t}+\overline{F}_{n,x}+(\overline{H}_n v)_y=0;\ n\geqslant 1,\quad v=-\int_0^h u_x d\eta. \tag{18}$$

(In Miura's table the coefficients of v do not coincide with $\overline{H}_n$ because of misprints and omissions.) As usual, under the assumptions of rapid decay of solutions at infinity, from this it is possible to obtain quantities conserved in time: these are $\int_{-\infty}^{\infty} H_n dx,\ n\geqslant 0.$

The next two sections are devoted to a new construction of relations (17) and (18). The description of the generating function of the system of polynomials H_n as solutions of a certain integral equation with parameter plays a central role in the construction.

6.3. THEOREM. We set $\Phi(\lambda)=\sum_{i=0}^{\infty}(-1)^i A_i\lambda^{-(i+1)}=\int_0^h(\lambda+u)^{-1}dy.$ Then there exists a unique solution $\mu(\lambda)$ of the equation

$$\mu(\lambda)+\Phi(\mu(\lambda))=\lambda \tag{19}$$

a) in the class of formal series of the form $\lambda+Q[A_0, A_1,\ldots][[\lambda^{-1}]]$; b) in the class of functions analytic in λ of the form $\lambda+o(\lambda^{-1})$ in a neighborhood of ∞ (depending on $u(x, y, t)$, $h(x, t)$) .

6.4. THEOREM. System (15) implies equations for $\mu(\lambda)$ of the form

$$\mu_t-(\mu^2/2+A_0)_x=0, \tag{20}$$

$$\left[\frac{\partial\mu}{\partial\lambda}(\mu+u)^{-1}\right]_t-\left[\mu\frac{\partial\mu}{\partial\lambda}(\mu+u)^{-1}\right]_x-\left[\frac{\partial\mu}{\partial\lambda}(\mu+u)^{-1}\int_0^y u_x d\eta\right]_y=0. \tag{21}$$

Theorems 6.3 and 6.4 are proved in Secs. 7-8. The Benney conservation laws (17) (up to constant factors) are obtained from (20) by setting $\mu=\lambda-\sum_{i=0}^{\infty}(-1)^i H_i\lambda^{-(i+1)}$, $\mu^2/2+A_0=\lambda^2/2+\sum_{i=0}^{\infty}(-1)^i F_i\lambda^{-(i+1)}$ and equating coefficients of $\lambda^{-(n+1)}$. Miura's conservation laws are deduced similarly from the relations (21) in which $(\mu+u)^{-1}$ is to be interpreted as $\sum_{i=0}^{\infty}(-1)^i u^i\mu^{-(i+1)}\in Q[u, A_i][[\mu^{-1}]]$.

6.5. The Reduced Equations. As already mentioned, the reduced system is obtained from Eqs. (15) by adding the condition $u_y=0$: the horizontal component of the velocity does not depend on height. The reduced system for the functions $u(x, t)$, $h(x, t)$ has the form (the classical equations for long waves):

$$\begin{cases} u_t=-\left(\frac{u^2}{2}+h\right)_x=\frac{\partial}{\partial x}\left(\frac{\partial H}{\partial h}\right), \\ h_t=-(uh)_x=\frac{\partial}{\partial x}\left(\frac{\partial H}{\partial u}\right),\ H=-\frac{h^2+hu^2}{2}. \end{cases} \tag{22}$$

This system is Hamiltonian with Hamiltonian operator $B=\begin{pmatrix}0 & \partial_x\\ \partial_x & 0\end{pmatrix}$, and Hamiltonian H.

The integrals H_n of the full system (15) are converted into integrals H_n^0 of the reduced system (22) after substituting $A_n \mapsto hu^n = A_n^0$: $H_n^0 \in Q[u, h]$. We explicitly compute these integrals in Sec. 9, we show there are no others, and we establish their commutativity.

6.6. THEOREM. a) Up to constant multiples we have $H_n^0 = \sum_{k=0}^{[n/2]} t_{n,k} u^{n-2k} h^{k+1}$, where

$$t_{n,k} = \frac{n!}{(n-2k)!\,k!\,(k+1)!} \tag{23}$$

and any polynomial H of u, h, for which $H_t \in \partial_x Q[u, h]$ is a linear combination of the H_n^0.

b) $[H_n^0, H_m^0] \in \partial_x Q[u, h]$.

Theorem 6.6 enables us to introduce the "higher reduced equations"

$$\begin{cases} u_t = \left(\dfrac{\partial H}{\partial h}\right)_x, \\ h_t = \left(\dfrac{\partial H}{\partial u}\right)_x \end{cases} \quad H = \sum_{n=0}^{N} c_n H_n^0,\ c_n \in \mathbb{R}. \tag{24}$$

From what has been said, they admit conservation laws of the form $H_{n,t}^0 + \tilde{F}_{n,x}^0 = 0$, where the $\tilde{F}_n^0$ depend on $c_0, \ldots, c_N$.

In Sec. 10 it is shown that the construction of commuting Hamiltonians $\{H_n{}^0\}$ can be considerably generalized.

In Sec. 11 we pass to the investigation of the full system (15) or, more precisely, of the system of equations for the moments (16). We begin by showing that they are Hamiltonian as a system of evolution for infinitely many unknown functions $A_n(x, t)$ of two variables by establishing the following facts.

We set $B = B_1\partial + \partial \circ B_1^t$, where $B_{1,ij} = iA_{i+j-1}$, $i, j \geqslant 0$. Then, as was shown in Sec. 8 of Chap. I, the operator B is Hamiltonian.

6.7. THEOREM. The Benney equations for the moments are Hamiltonian with operator B and Hamiltonian $-\frac{1}{2} H_2 = -\frac{1}{2}(A_2 + A_0^2)$.

(The Hamiltonian formalism is here applied to the ring $\mathcal{A} = k[A_i^{(j)}]$, $i, j \geqslant 0$, $A_i^{(j)}$ being the independent variables. It is also possible to work in the ring $\bigcup_{k=0}^{\infty} C^\infty(x, A_i^{(j)} \mid i + j \leqslant k)$ or in various intermediate rings.)

6.8. THEOREM. Let $H \in \mathcal{A}$ be any element. Then a system of equations for u, h of the form

$$\begin{cases} u_t = \left(\sum_{j=0}^{\infty} u^j H_{(j)}\right)_x - u_y \int_0^y \left(\sum_{j=0}^{\infty} j u^{j-1} H_{(j)}\right)_x \Big|_{y=\eta} d\eta, \\ h_t = \left(\sum_{j=0}^{\infty} j A_{j-1} H_{(j)}\right)_x, \end{cases} \tag{25}$$

where $H_{(j)} = \frac{\delta H}{\delta A_j}$, implies the system of equations $\bar{A}_t = B \frac{\delta H}{\delta \bar{A}}$ with Hamiltonian H .

This theorem is proved in Sec. 11. Theorem 6.7 is a special case of it for $H=-\frac{1}{2}H_2$, when the system (25) becomes the original system of equations (15) of part 1 and Theorem 6.7 becomes the Benney lemma. Theorem 6.8 indicates the rather surprising situation: inspite of the loss of information on passing from u, h to $\bar{A}$, the equations of evaluation for $\bar{A}$ with any Hamiltonian can be lifted to equations of evolution for u, h. Of course, the question of lifting solutions requires separate investigation.

6.9. THEOREM. Let the elements $H_i \in \mathcal{A}$ be defined from Benney's conservation laws (17). Then they commute relative to the Hamiltonian structure with operator B.

This theorem is proved in Sec. 12. It makes natural the consideration of systems of the form (25) with Hamiltonian $H=\sum c_i H_i$, which we call higher Benney equations in analogy with the higher Korteweg–de Vries equations. By the general formalism the higher equations possess conservation laws of Benney type $H_{i,t}+\tilde{F}_{i,x}=0$, where the $\tilde{F}_i$ depend on H. Conservation laws of Miura type hold for one of the higher equations.

6.10. THEOREM. There exist Miura conservation laws for the system (25) with Hamiltonian $H=cH_3$.

Unfortunately, we have been unable to determine if this fact holds for the other higher equations. Theorem 6.10 is proved in Sec. 13.

We began by considering the reduced higher Benney equations (with the additional condition $u_y=0$), for which $A_i=hu^i$. As was mentioned, they are Hamiltonian related to the operator $\begin{pmatrix} 0 & \partial \\ \partial & 0 \end{pmatrix}$ in the ring $\mathcal{A}^0=Q[u^{(j)}, h^{(j)} \mid j \geqslant 0]$.

6.11. THEOREM. The Hamiltonian structures described in the rings $\mathcal{A}$ and $\mathcal{A}^0$ are compatible.

A precise formulation and proof of this theorem is given in Sec. 14. That section also contains an explanation of the reasons for which the system (25) contains the operators

$$\sum_{j=0}^{\infty} u^j \frac{\delta}{\delta A_j}, \quad \sum_{j=0}^{\infty} jA_{j-1} \frac{\delta}{\delta A_j}.$$

7. The Function $\mu(\lambda)$

In this section Theorem 6.3 is proved, and further information regarding the function $\mu(\lambda)$ and its coefficients needed below is presented.

7.1. The Function $\mu(\lambda)$ as a Formal Series. We set $\mu(\lambda)=\lambda-\sum_{i=-1}^{\infty}(-1)^i H_i \lambda^{-(i+1)}$ and seek the coefficients H_i from Eq. (19) which we rewrite in the form

$$\sum_{i=-1}^{\infty}(-1)^i H_i \lambda^{-(i+1)}-\sum_{i=0}^{\infty}\lambda^{-(i+1)}(-1)^i A_i\left[1-\sum_{j=-1}^{\infty}(-1)^j H_j \lambda^{-(j+2)}\right]^{-(i+1)}=0. \quad (26)$$

This obviously implies that $H_{-1}=0$, $H_0=A_0$, $H_1=A_1$ and further $H_n=A_n+P_n$, where $P_n \in Z[A_0, \ldots, A_{n-2}, H_0, \ldots, H_{n-2}]$ for $n \geqslant 2$. Induction on n immediately shows that the $H_n \in A_n+$

$Z[A_0, \ldots, A_{n-2}]$ exist and are uniquely determined. The first H_n are: A_0, A_1, $A_2+A_0^2$, $A_3+3A_0A_1$, $A_4+4A_0A_2+2A_1^2+2A_0^3$.

7.2. The Function $\mu(\lambda)$ as an Analytic Function. For given differentiable $u(x, y, t)$, $h(x, t)$ the moments A_n become functions of x, t; it would be possible to attempt to estimate the $|H_n(x, t)|$ and show that they grow no faster than $C^n(x, t)$ for a suitable function $C(x, t)$, and this would establish the analyticity of $\mu(\lambda)$ in λ. However, it is inconvenient to obtain such an estimate directly, and instead of this we apply the method of iterations to the integral equation (19): $\mu(\lambda)+\int_0^h(\mu(\lambda)+u)^{-1}dy=0$. We set $\mu(\lambda)=\lambda+\varepsilon(\lambda)$. The uniqueness of a function μ analytic in λ with the property (19) and the estimate $\varepsilon=O(\lambda^{-1})$ follows from the uniqueness of the formal series. To prove existence we set $\varepsilon_0=0$, $\varepsilon_N=-\int_0^h(u+\lambda+\varepsilon_{N-1})^{-1}dy$ and show that $\varepsilon=\lim_{N\to\infty}\varepsilon_N$ exists (for given (x, t)) uniformly in λ, when $|\lambda|$ is so large that the following inequalities are satisfied:

$$\begin{cases} h(|\lambda|-U)^{-1}(1-4h(|\lambda|-U)^{-2})^{-1}\leqslant(|\lambda|-U)/2, \\ |\lambda|-U>2\sqrt{h};\ U=\sup\{u\,|\,0\leqslant y\leqslant h\}. \end{cases} \tag{27}$$

In particular, if u, h are bounded there is a domain of analyticity μ, not depending on x, t.

To this end we establish the following n inequalities by induction:

$$\begin{cases} |\varepsilon_n|\leqslant(|\lambda|-U)/2; \\ |\varepsilon_n-\varepsilon_{n-1}|\leqslant\theta\,|\varepsilon_{n-1}-\varepsilon_{n-2}|. \end{cases} \quad \left(\theta=\frac{4h}{(|\lambda|-U)^2}<1, \text{ by } (27)\right). \tag{28}$$

Indeed, $|\varepsilon_1|=\left|\int_0^h\frac{dy}{u+\lambda}\right|\leqslant h\sup|u+\lambda|^{-1}\leqslant h(2\sqrt{h})^{-1}<\frac{|\lambda|-U}{4}<\frac{|\lambda|-U}{2}$. Further, for any $N\geqslant 1$ we have $|\varepsilon_{N+1}-\varepsilon_N|\leqslant h\sup|(u+\lambda+\varepsilon_N)(u+\lambda+\varepsilon_{N-1})|^{-1}|\varepsilon_N-\varepsilon_{N-1}|$. Setting here $N=1$ and using (28) for $N=1$, we find by (27) $|u+\lambda|^{-1}\leqslant(|\lambda|-U)^{-1}$, $|u+\lambda+\varepsilon_1|^{-1}\leqslant 2(|\lambda|-U)^{-1}$, whence $|\varepsilon_2-\varepsilon_1|\leqslant 2h(|\lambda|-U)^{-2}|\varepsilon_1|<4h(|\lambda|-U)^{-2}|\varepsilon_1|$. This provides the basis for induction. Suppose now that the inequalities of (28) are satisfied for all $n\leqslant N$. Then

$$|\varepsilon_{N+1}-\varepsilon_N|\leqslant h\sup|(u+\lambda+\varepsilon_N)(u+\lambda+\varepsilon_{N-1})|^{-1}|\varepsilon_N-\varepsilon_{N-1}|\leqslant\theta\,|\varepsilon_N-\varepsilon_{N-1}|,$$

by (28) for $n=N-1, N$. Hence

$$|\varepsilon_{N+1}|\leqslant\sum_{n=1}^{N+1}|\varepsilon_n-\varepsilon_{n-1}|<(1-\theta)^{-1}|\varepsilon_1|\leqslant(1-\theta)^{-1}h(|\lambda|-U)^{-1}\leqslant(|\lambda|-U)/2,$$

by (27). This completes the proof of Theorem 6.3.

Below we shall mainly use $\mu(\lambda)$ as a formal series. The validity of the corresponding computations in the analytic version can be trivially verified. We shall establish several curious algebraic properties of the series for $\mu(\lambda)$ and its coefficients H_i.

7.3. Homogeneity. We call the number A_n the weight of the variable $n+2$, and we introduce the corresponding graded ring $Z[A_n]$. In this gradation the polynomials H_n are homogeneous of weight $n+2$. This fact is obtained by an obvious induction on n from the identity (26).

7.4. The Symmetry $A_i \leftrightarrow -H_i$, $\lambda \leftrightarrow \mu$. We define a homomorphism σ of the ring $\mathbf{Z}[A_i][[\lambda^{-1}]]$ into itself by the conditions $\sigma(A_i) = -H_i$, $\sigma(\lambda) = \mu$. It is obviously an automorphism. We shall show that it is an involution: $\sigma^2 = \mathrm{id}$.

Applying σ to (19), we obtain $\sigma(\mu) - \sum_{i=0}^{\infty} H_i (\sigma(\mu))^{-(i+1)} = \mu$. On the other hand, by the definition of μ, $\lambda - \sum_{i=0}^{\infty} (-1)^i H_i \lambda^{-(i+1)} = \mu$. But from the first equation, as in 7.2, the element $\sigma(\mu) \in \mu + \mathbf{Z}[A_i][[\mu^{-1}]] = \lambda + \mathbf{Z}[A_i][[\lambda^{-1}]]$ is uniquely determined. Hence $\sigma(\mu) = \lambda$. Applying σ again to the series for μ, we find $\mu - \sum_{i=0}^{\infty} (-1)^i \sigma(H_i) \mu^{-(i+1)} = \lambda$; since $\mu + \sum_{i=0}^{\infty} (-1)^i A_i \mu^{-(i+1)} = \lambda$ by (19), in view of the uniqueness of the expression for λ in terms of μ, we obtain $\sigma(H_i) = -A_i$, i.e., $\sigma(-H_i) = A_i$.

7.5. Summation of the Benney–Miura Series. In the work of Benney [27] and Miura [46] the generating series for the conservation laws were constructed in a different way. Benney started from the series $\Gamma(z) = \sum_{n=0}^{\infty} A_n z^n$ and showed that $\sum_{n=0}^{\infty} H_n z^n = \sum_{n=0}^{\infty} \left(z^2 \frac{\partial}{\partial z}\right)^n \frac{\Gamma^{n+1} z^{n+1}}{(n+1)!}$; similar formulas are given in [27, 46] for other series from (20), (21). After the substitution $z = -\lambda^{-1}$ and passage to $\Phi(\lambda) = \lambda^{-1}\Gamma(-\lambda^{-1})$ all the series of Benney and Miura can be represented in the form $\sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{\partial}{\partial \lambda}\right)^n [\varphi^n(\lambda)\psi(\lambda)]$, where $\varphi = -\Phi$, $\psi \in \mathbf{Z}[A_i]((\lambda^{-1}))$. Our basic integral equation (19) was obtained by applying to these series the following identity which is of independent interest.

7.6. Proposition. Let k be some Q-algebra, and let $\varphi \in k[[\lambda^{-1}]]$, $\psi \in k((\lambda^{-1}))$. Then $\sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{\partial}{\partial \lambda}\right)^n [\varphi^n \psi] = \partial\mu/\partial\lambda \cdot \psi(\lambda)$, where $\mu \in \lambda + k[[\lambda^{-1}]]$ is the unique root of the equation $\mu = \varphi(\mu) + \lambda$.

Proof. We note first of all that the series for $\left(\frac{\partial}{\partial \lambda}\right)^n [\varphi^n \psi]$ begins at least with the terms $\lambda^{-2n+\mathrm{const}}$, so that the sum converges λ^{-1}-adically. We introduce the auxiliary variable C and denote by μ_c the root of the equation $\mu_c = c\varphi(\mu_c) + \lambda$ in $\lambda + k[c][[\lambda^{-1}]]$. Its existence and uniqueness follow immediately from Theorem 6.3a). We shall show that $\sum_{n=0}^{\infty} \frac{c^n}{n!} \left(\frac{\partial}{\partial \lambda}\right)^n [\varphi^n \psi] = \frac{\partial \mu_c}{\partial \lambda} \psi(\mu_c)$ in the ring $k[c][[\lambda^{-1}]]$. For $c = 1$ this implies the result of the proposition.

This identity is equivalent to the sequence of identities $\left(\frac{\partial}{\partial \lambda}\right)^n [\varphi^n \psi] = \left(\frac{\partial}{\partial c}\right)^n \left[\frac{\partial \mu_c}{\partial \lambda} \psi(\mu_c)\right]\Big|_{c=0}$, which we establish by induction on n. For $n = 0$ the result is obvious, since $\mu_c|_{c=0} = \lambda$, $\frac{\partial \mu_c}{\partial \lambda}\Big|_{c=0} = 1$. Suppose that it is true for $n < N$ and all φ, ψ. In order to make the inductive step, we use the following remark. Differentiating the relation $\mu_c = c\varphi(\mu_c) + \lambda$ with respect to c and λ, we find $\partial\mu_c/\partial c = [1 - c\varphi'(\mu_c)]^{-1}\varphi(\mu_c)$, $\partial\mu_c/\partial\lambda = [1 - c\varphi'(\mu_c)]^{-1}$, whence $\partial\mu_c/\partial c = \partial\mu_c/\partial\lambda \cdot \varphi(\mu_c)$ and further $\partial^2\mu_c/\partial\lambda\partial c = \partial^2\mu_c/\partial\lambda^2 \cdot \varphi(\mu_c) + (\partial\mu_c/\partial\lambda)^2 \varphi'(\mu_c)$.

We now have

$$\left(\frac{\partial}{\partial\lambda}\right)^{N+1}[\varphi^{N+1}\psi]=\frac{\partial}{\partial\lambda}\Big[\left(\frac{\partial}{\partial\lambda}\right)^{N}(\varphi^{N}\cdot\varphi\psi)\Big]=\frac{\partial}{\partial\lambda}\Big\{\left(\frac{\partial}{\partial c}\right)^{N}\Big[\frac{\partial\mu_c}{\partial\lambda}\cdot(\varphi\psi)(\mu_c)\Big]\Big|_{c=0}\Big\}=$$
$$=\left(\frac{\partial}{\partial c}\right)^{N}\Big\{\frac{\partial^2\mu_c}{\partial\lambda^2}(\varphi\psi)(\mu_c)+\left(\frac{\partial\mu_c}{\partial\lambda}\right)^{2}[\varphi'(\mu_c)\psi(\mu_c)+\varphi(\mu_c)\psi'(\mu_c)]\Big\}\Big|_{c=0}=$$
$$=\left(\frac{\partial}{\partial c}\right)^{N}\Big[\frac{\partial^2\mu_c}{\partial\lambda\partial c}\psi(\mu_c)+\psi'(\mu_c)\frac{\partial\mu_c}{\partial\lambda}\frac{\partial\mu_c}{\partial c}\Big]\Big|_{c=0}=\left(\frac{\partial}{\partial c}\right)^{N+1}\Big[\frac{\partial\mu_c}{\partial\lambda}\psi(\mu_c)\Big]\Big|_{c=0}.$$

This completes the proof.

8. The Conservation Laws

In this section we shall prove Theorem 6.4. In order to establish a formal version of it, we introduce the independent variables $A_i^{(j)}$, $A_i^{(0)}=A_i$; $i, j\geqslant 0$, and we consider the following differentiations of the ring $Z[A_i^{(j)}][[\lambda^{-1}]]$ into itself. The differentiation $\partial/\partial A_i$ takes A_i into 1, and it takes $A_k^{(j)}$ for $k\neq i$, $A_i^{(j)}$ for $j\geqslant 1$ and λ^{-1} into zero; $\partial/\partial x$ takes $A_i^{(j)}$ into $A_i^{(j+1)}$ and λ^{-1} into zero; $\partial/\partial t$ commutes with $\partial/\partial x$, and takes A_i into $-A_{i+1}^{(1)}-iA_{i-1}A_0^{(1)}$ [cf. (16)] and λ^{-1} into zero. Moereover, all these differentiations are continuous in the λ^{-1} -adic topology.

8.1. LEMMA. Let D be one of the differentiations described above. Then $(1+\Phi'(\mu))D\mu=D\lambda-\Phi^D(\mu)$, where, by definition, $\Phi'(\cdot)=\sum_{i=0}^{\infty}(-1)^{i+1}A_i(i+1)(\cdot)^{-(i+2)}$, $\Phi^D(\cdot)=\sum_{i=0}^{\infty}(-1)^iDA_i(\cdot)^{-(i+1)}$.

COROLLARY. a) $(1+\Phi'(\mu))\partial\mu/\partial\lambda=1$; b) $\partial\mu/\partial x=-\partial\Phi/\partial x(\mu)\frac{\partial\mu}{\partial\lambda}$; c) $\partial\mu/\partial t=-\frac{\partial\Phi}{\partial t}(\mu)\frac{\partial\mu}{\partial\lambda}$; d) $\frac{\partial\mu}{\partial A_l}=(-1)^{l+1}\mu^{-(l+1)}\frac{\partial\mu}{\partial\lambda}$.

Proof. We apply D to the relation (19); we obtain $D\mu+\Phi^D(\mu)+\Phi'(\mu)D\mu=D\lambda$, i.e., Lemma 8.1. Setting here $D=\partial/\partial\lambda$, we obtain Corollary a). Finally, putting successively $D=\partial/\partial x$, $\partial/\partial t$, $\partial/\partial A_i$, and using Corollary a) and the fact that $\Phi^{\partial/\partial A_l}(\cdot)=(-1)^l(\cdot)^{-(l+1)}$, we obtain Corollaries b), c), and d).

8.2. Proof of Formula (20). We multiply relation (16) by $(-1)^n\mu^{-(n+1)}$ and sum on n from 0 to ∞. We obtain $\partial\Phi/\partial t(\mu)-\mu(\partial\Phi/\partial x(\mu)-A_{0,x}\mu^{-1})+A_{0,x}\Phi'(\mu)=0$, or $\partial\Phi/\partial t(\mu)-\mu\partial\Phi/\partial x(\mu)+A_{0,x}(1+\Phi'(\mu))=0$. Multiplying the last equality by $\partial\mu/\partial\lambda$ and using Corollaries a), b), c), we obtain (20).

8.3. Proof of Formula (21). We carry out the proof by analogous formal computations in the extended ring $Z[A_i^{(j)}, u^{(k,l)}, v_{(n)}^{(m)}][[\lambda^{-1}]]$, where $k, l, m, n\geqslant 0$; $u^{(0,0)}=u$; $v_{(0)}^{(0)}=v$ and corresponds to $-\int_0^y u_x d\eta$; all the new variables are independent of one another and of the previous variables. The differentiation $\partial/\partial x$ takes $u^{(k,l)}$ into $u^{(k+1,l)}$, $v_{(n)}^{(m)}$ into $v_{(n)}^{(m+1)}$, $A_i^{(j)}$ into $A_i^{(j+1)}$, and λ^{-1} into zero; the differentiation $\partial/\partial\lambda$ is trivial on the new variables; the differentiation $\partial/\partial t$ commutes with $\partial/\partial x$, $\partial/\partial\lambda$ and takes $v_{(n)}^{(m)}$ into $v_{(n+1)}^{(m)}$, A_0 into $-A_1$, u into $-uu^{(1,0)}-u^{(0,1)}v-A_0^{(1)}$; and λ^{-1} into zero; the differentiation $\partial/\partial y$ commutes with $\partial/\partial x$, $\partial/\partial t$ and takes $u^{(k,l)}$ into $u^{(k,l+1)}$, $v_{(n)}^{(m)}$ into $-(\partial/\partial t)^n u^{(m+1,0)}$ and is trivial on $A_i^{(j)}$ and λ^{-1} (the differentiations are chosen in correspondence with the first and second equations of (15)).

We now calculate (21) using the prime to denote differentiation with respect to λ:

$$(\mu'(\mu+u)^{-1})_t=[(\mu+u)\mu'_t-\mu'(\mu_t+u_t)](\mu+u)^{-2},$$

$$-(\mu\mu'(\mu+u)^{-1})_x=-[(\mu+u)(\mu_x\mu'+\mu\mu'_x)-\mu\mu'(\mu_x+u_x)](\mu+u)^{-2},$$

$$(v\mu'(\mu+u)^{-1})_y=[(\mu+u)(v_y\mu'+v\mu'_y)-v\mu'(\mu_y+v_y)](\mu+u)^{-2}.$$

Substituting into the numerators of these expressions the right sides of the relations $\mu_y=\mu'_y=0$, $v_y=-u_x$, $\mu_t=\mu\mu_x+A_{0,x}$ [formula (20)], $\mu'_t=\mu'\mu_x+\mu\mu'_x$ (the derivative of (20) with respect to λ), $u_t=-uu_x-u_yv-A_{0,x}$ and collecting like terms, we find that the sum of all numerators is equal to zero. This completes the proof of Theorem 6.4.

There are a number of relations between the "densities" of conserved quantities H_l, $\bar{H}_l$ and the "local currents" F_l, $\bar{F}_l$ which we now describe. We introduce two further differentiations $\mathcal{D}=\frac{\partial}{\partial x}+\sum_{j=1}^{\infty} jA_{j-1}\frac{\partial}{\partial A_j}$ and $D=\sum_{j=0}^{\infty} u^j\frac{\partial}{\partial A_j}$.

8.4. LEMMA. a) $\mathcal{D}\mu=\partial\mu/\partial\lambda-1$; b) $D\mu=-\partial\mu/\partial\lambda\,(\mu+u)^{-1}$.

Proof. a) Using Corollary 8.1 d), we find

$$\mathcal{D}\mu=\sum_{j=1}^{\infty} jA_{j-1}(-1)^{j+1}\mu^{-(j+1)}\partial\mu/\partial\lambda=-\Phi'(\mu)\,\partial\mu/\partial\lambda=\partial\mu/\partial\lambda-1$$

(the last equality follows from 8.1 a)). b) Similarly,

$$D\mu=\sum_{j=0}^{\infty} u^j(-1)^{j+1}\mu^{-(j+1)}\partial\mu/\partial\lambda=-\partial\mu/\partial\lambda\,(\mu+u)^{-1}.$$

8.5. Proposition. The following relations hold:

$$\text{a) } \mathcal{D}H_n=nH_{n-1};\ \text{b) } H_n=\frac{\partial}{\partial A_0}\frac{F_{n+1}}{n+1};\ \text{c) } \frac{\partial}{\partial A_l}H_n=\frac{\partial}{\partial A_{l+1}}F_n;$$
$$\text{d) } DH_n=-\bar{H}_n,\ DF_n=-\bar{F}_n;\ \text{e) } \mathcal{D}\bar{H}_n=n\bar{H}_{n-1};$$
$$\text{f) } \bar{H}_n=\frac{\partial}{\partial A_0}\frac{\bar{F}_{n+1}}{n+1};\ \text{g) } \frac{\partial}{\partial A_l}\bar{H}_n=\frac{\partial}{\partial A_{l+1}}\bar{F}_n.$$

Proof. Relation a) follows immediately from Lemma 8.4 a); d) is obtained from 8.4 b) and (21). Further, according to 8.1 d), $\partial/\partial A_0(\mu^2/2+A_0)=-\partial\mu/\partial\lambda+1$; this and (20) imply b). Relation c) is obtained similarly: $\partial/\partial A_{l+1}(\mu^2/2+A_0)=(-1)^l\mu^{-(l+1)}\partial\mu/\partial\lambda=-\partial\mu/\partial A_l$. The relations e), f), and g) follow from a), b), and c) respectively if to the latter we apply D and note that $[D,\mathcal{D}]=[D,\partial/\partial A_i]=0$.

8.6. Homogeneity. Setting the weight of u equal to 1, we find from 7.3 and Eqs. (20), (21) that the F_n are homogeneous of weight $n+3$; $\bar{H}_n$ and $\bar{F}_n$ are homogeneous of weights n and $n+1$, respectively.

9. Integrals of the Reduced System

Theorem 6.6 is proved in this section.

9.1. Proof of Theorem 6.6 a). The derivative of u^kh^l, by virtue of the system (22) is equal to $(u^kh^l)_t=-(k+l)u^kh^lu_x-(ku^{k-1}h^l+lu^{k+1}h^{l-1})h_x$. Therefore, for any polynomial $P\in Q[u,h]$, represented as the sum of terms homogeneous in u, h, $P=\sum P_i$ of degree $\deg P_i$, we have $P_t = -\sum \deg P_i \times P_i u_x - \left(u\frac{\partial}{\partial h} + \frac{\partial}{\partial u}\right) P \cdot h_x$. The Poincaré lemma implies that the condition $P_t \in \frac{\partial}{\partial x} Q[h, u]$ is equivalent to the requirement that

$$\Delta(P) \overset{\text{def}}{=} \partial/\partial h \left(\sum \deg P_i \cdot P_i\right) - \frac{\partial}{\partial u} \circ (u\partial/\partial h + \partial/\partial u) P = 0.$$

In $Q[u, h]$ we introduce the gradation with weight w, by setting $w(u)=1$, $w(h)=2$ (cf. 7.3). Obviously, Δ is a homogeneous operator of weight -2. Therefore, the kernel of Δ is the direct sum of its homogeneous components, and it suffices to compute it on a component of weight n, i.e., to solve the equation $\Delta\left(\sum_{k=0}^{[n/2]} c_{n,k} u^{n-2k} h^k\right) = 0$. Somewhat complicated computations show that $\Delta = 0$ on $Q + Qu$, while for $n > 1$ we have $c_{n,0} = 0$ and $c_{n,k+1} = (n-2k) \cdot (n-2k-1) k^{-1} (k+1)^{-1} c_{n,k}$ for $k \geqslant 1$. This implies that all homogeneous components of the kernel of Δ are one-dimensional and are generated by polynomials H_n^0 of the form (23).

Remark. The explicit form of $H_n{}^0$ could also be obtained by computing the coefficients of the function $\mu^0(\lambda)$, which in the reduced case satisfies the equation $\mu^0(\lambda) + h(\mu^0(\lambda) + u)^{-1} = \lambda$. It cannot, however, be proved that the integrals found form a complete system. On the other hand, by generalizing our arguments to the full system (15) it is possible to show that the Benney integrals also exhaust the space of integrals which are polynomials in A_i.

9.2. Proof of Thecrem 6.6 b). According to the general formalism of Chap. I, the Poisson bracket $[H_m{}^0, H_n{}^0]$ is the derivative with respect to t of $H_n{}^0$, by virtue of the system $u_t = (\partial H_m{}^0/\partial h)_x$, $h_t = (\partial H_m{}^0/\partial u)_x$.

In order to cover all pairs (n, m) simultaneously, we introduce the new formal generating function $\nu(\lambda) = \sum_{n=0}^{\infty} H_n^0 \lambda^n/n!$ and in the ring $Q[u, h][[\lambda, \lambda_1]]$ we compute the derivative with respect to t of $\nu(\lambda)$ on the basis of the system $u_t = \nu(\lambda_1)_{hx}$, $h_t = \nu(\lambda_1)_{ux}$. More precisely, setting $\nu = \nu(\lambda)$, $\nu_1 = \nu(\lambda_1)$, we show that $\nu_t \in \partial/\partial x Q[u, h][[\lambda, \lambda_1]]$. The conventions regarding the differentiation are analogous to those described in Sec. 8.

We note, first of all, that because of (23)

$$\nu = \sum_{n=0}^{\infty} \sum_{k=0}^{[n/2]} t_{n,k} u^{n-2k} h^{k+1} \lambda^n / n! = \sum_{n,k} \frac{u^{n-2k}}{(n-2k)!} \frac{h^k \lambda^n}{k!(k+1)!} = h \sum_{k=0}^{\infty} \frac{(h\lambda^2)^k}{k!(k+1)!} \sum_{m=0}^{\infty} \frac{(u\lambda)^m}{m!} = \omega(h\lambda^2) e^{\lambda u},$$

where $\omega(\cdot) = \lambda^{-2} \sum_{k=0}^{\infty} \frac{(\cdot)^{k+1}}{k!(k+1)!}$. It is important below that $[\omega(h\lambda^2)]_{hh} = \lambda^2 h^{-1} \omega(h\lambda^2)$, which is obvious from the definition of ω, and similarly $\nu_{hh} = \lambda^2 h^{-1} \nu$. Moreover, $\nu_u = \lambda \nu$.

Now the evolution of ν, due to the Hamiltonian ν_1, is determined by the relations $\nu_t = \nu_u u_t + \nu_h h_t = \nu_u \nu_{1hx} + \nu_h \nu_{1ux} = \nu_u (\nu_{1hh} h_x + \nu_{1hu} u_x) + \nu_h (\nu_{1uh} h_x + \nu_{1uu} u_x) = (\nu_u \nu_{1hh} + \nu_h \nu_{1uh}) h_x + (\nu_u \nu_{1hu} + \nu_h \nu_{1uu}) u_x$. The inclusion $\nu_t \in \frac{\partial}{\partial x} Q[u, h][[\lambda, \lambda_1]]$ is equivalent to the identity $(\nu_u \nu_{1hh} + \nu_h \nu_{1uh})_u = (\nu_u \nu_{1hu} + \nu_h \nu_{1uu})_h$, which is checked directly by means of the relations $\nu_{hh} = \lambda^2 h^{-1} \nu$, $\nu_u = \lambda \nu$ and analogous equalities for ν_1.

10. Other Spaces of Commuting Hamiltonians

The sequence of Hamiltonians $\{H_n^0\}$ constructed in Sec. 9 is an example of a general construction which makes it possible to form infinite-dimensional spaces of commuting Hamiltonians in the ring $R[u, h]$. We now develop this construction .

10.1. Notation. $C^2(u)$ denotes the ring of functions of u; with two continuous derivatives; $C^2(h)$ and $C^2(u, h)$ have an analogous meaning. The letters λ, λ_1 denote independent formal parameters; the derivatives of them with respect to u, h are equal to zero. We consider two formal Laurent series with a finite number of negative powers $\eta^u(\lambda)\in C^2(u)((\lambda))$, $\eta^h(\lambda)\in C^2(h)((\lambda))$. We set $\eta^u(\lambda)\eta^h(\lambda)=\sum\eta_i(u, h)\lambda^i$. We write $\eta^u=\eta^u(\lambda)$, $\eta_1^u=\eta^u(\lambda_1)$; η^h and η_1^h have an analogous meaning. We suppose that η^u, η_{uu}^u, η^h, η_{hh}^h vanish only at isolated points.

10.2. THEOREM. The Hamiltonians $\{\eta_i(u, h)\}$ commute pairwise relative to the Hamiltonian operator $\begin{pmatrix}0 & \partial_x\\ \partial_x & 0\end{pmatrix}$, if and only if there exist functions $V_1(u)$, $V_2(h)$ which are continuous outside a set of isolated points and a series $a\in\mathbf{R}((\lambda))$, such that η^u, η^h satisfy the equations

$$\eta_{uu}^u=a(\lambda)V_1(u)\eta^u,\quad \eta_{hh}^h=a(\lambda)V_2(h)\eta^h. \tag{29}$$

Example. The Hamiltonians $\{H_i^0\}$ are obtained by setting $V_1=1$, $V_2=h^{-1}$, $a(\lambda)=\lambda^2$ (cf. 9.2).

Proof. As in 9.2 it suffices to verify that condition (29) is equivalent to the condition $\eta_t\in\partial/\partial x C^2(u, h)((\lambda))$, where the derivative with respect to t is taken on the basis of equations with Hamiltonian η_1: $u_t=\eta_{1hx}$, $h_t=\eta_{1ux}$. Computing as in 9.2, we find

$$\eta_t=\eta_u u_t+\eta_h h_t=\eta_u\eta_{1hx}+\eta_h\eta_{1ux}=(\eta_u\eta_{1hh}+\eta_h\eta_{1hh})h_x+(\eta_u\eta_{1hu}+\eta_h\eta_{1uu})u_x.$$

Hence the condition $\eta_t\in\partial/\partial x C^2(u, h)((\lambda))$ is equivalent to the identity

$$(\eta_{uu}\eta_{1hh}+\eta_h\eta_{1uh})_u=(\eta_u\eta_{1hu}+\eta_h\eta_{1uu})_h,$$

which after cancelling like terms and a suitable division acquires the form $(\eta_{uu}^u/\eta^u)(\eta_{1uu}^u/\eta_1^u)^{-1}=(\eta_{hh}^h/\eta^h)(\eta_{1hh}^h/\eta_1^h)^{-1}$. From this it follows that both sides of the equation are independent of u, h, and hence have the form $a(\lambda)a(\lambda_1)^{-1}$, where $a(\lambda)\in\mathbf{R}((\lambda))$. Further, this implies that $\eta_{uu}^u/\eta^u=a(\lambda)V_1(u)$, $\eta_{hh}^h/\eta^h=a(\lambda)V_2(h)$, where V_1 and V_2 are continuous away from the zeros of η^u, η^h, respectively. Obviously, the converse is also true, and this completes the proof of Theorem 9.2.

11. Lifts of Equations of Evolution

Theorem 6.8 is proved in this section.

11.1. Writing out explicitly the equation for A_i with Hamiltonian H for the operator B, introduced in Theorem 6.7, we obtain

$$A_{i,t}=\sum_{j\geqslant 0} iA_{i+j-1}H_{(j),x}+(jA_{i+j-1}H_{(j)})_x,\quad H_{(j)}=\delta H/\delta A_j. \tag{30}$$

In order to derive this system from (15), we consider separately the evolution of A_i on the basis of (15).

11.2. Using the expressions for u_t and h_t of (15), we find

$$A_{i,t}=\frac{\partial}{\partial t}\int_0^h u^i dy=\int_0^h iu^{i-1}u_t dy+u^i|_h h_t=\int_0^h iu^{i-1}\Big(\sum_{j\geqslant 0}u^j H_{(j)}\Big)_x dy-$$

$$-\int_0^h iu^{i-1}u_y dy\int_0^y\Big(\sum_{j\geqslant 0}ju^{j-1}H_{(j)}\Big)_x d\eta+u^j|_h\Big(\sum_{j\geqslant 0}jA_{j-1}H_{(j)}\Big)_x. \qquad (31)$$

We first transform the second (double) integral in this expression. It has the form

$$\int_0^h\left\{\left[u^i\int_0^y\Big(\sum_{j>0}ju^{j-1}H_{(j)}\Big)_x d\eta\right]_y-u^i\Big(\sum_{j\geqslant 0}ju^{j-1}H_{(j)}\Big)_x\right\}dy=u^i|_h\int_0^h\Big(\sum_{j\geqslant 0}ju_j^{j-1}H_{(j)}\Big)_x dy-\int_0^h u^i\Big(\sum_{j\geqslant 0}ju^{j-1}H_{(j)}\Big)_x dy=$$

$$=u^i|_h\left\{\frac{\partial}{\partial x}\int_0^h\sum_{j\geqslant 0}ju^{j-1}H_{(j)}dy-\sum_{j>0}ju^{j-1}|_h H_{(j)}h_x\right\}-\int_0^h u^i\Big(\sum_{j\geqslant 0}ju^{j-1}H_{(j)}\Big)_x dy=$$

$$=u^i|_h\Big(\sum_{j\geqslant 0}jA_{j-1}H_{(j)}\Big)_x-\frac{\partial}{\partial x}\int_0^h\sum_{j>0}ju^{i+j-1}H_{(j)}dy+\int_0^h\sum_{j>0}ju^{j-1}H_{(j)}(u^i)_x dy.$$

Substituting the last expression into (31) and cancelling like terms, we obtain finally

$$A_{i,t}=\int_0^h\sum_{j>0}iu^{i+j-1}H_{(j),x}dy+\Big(\int_0^h\sum_{j\geqslant 1}ju^{i+j-1}H_{(j)}dy\Big)_x=\sum_{j>0}(iA_{i+j-1}H_{(j),x}+(jA_{i+j-1}H_{(j)})_x),$$

since the $H_{(j)}$ do not depend on y. This coincides with (30).

The derivation of Theorem 6.7 from Theorem 6.8 coincides with that of the Benney lemma 6.2 if it is noted that Eqs. (30) for $H=-\frac{1}{2}(A_2+A_0^2)$ coincide with the equations in the Benney lemma, while Eqs. (25) coincide with the origin system (15).

12. The Benney Integrals Commute

12.1. Theorem 6.9 is proved in this section. We recall that it means the following: $X_{H_i}H_j\in\partial_x\mathcal{A}$. This assertion follows from a stronger fact which we now formulate. As was shown in Sec. 7 setting $\mu(\lambda)=\lambda-\sum_{i=0}^{\infty}(-1)^i H_i\lambda^{-(i+1)}$, we have $\mu(\lambda)+\Phi(\mu(\lambda))=\lambda$, where $\Phi(\lambda)=\sum_{i=0}^{\infty}(-1)^i A_i\lambda^{-(i+1)}$; the formal series for $\mu(\lambda)$ is uniquely determined by this equation. We choose a variable λ not depending on λ_1 and set $\mu=\mu(\lambda)$, $\mu_1=\mu(\lambda_1)$. Exactly as in Sec. 8 it is possible to define the expressions $X_H\mu$ and $X_{\mu_1}\mu$ (on λ the differentiations act trivially).

12.2. THEOREM. a) For any $H=\sum_{j<N}c_jH_j$ we have

$$X_H\mu=\Big(\sum_{j\geqslant 0}\frac{\partial H}{\partial A_j}(-1)^{j+1}\mu^j\Big)_x;$$

$$\text{b) } X_{\mu_1}\mu=\Big(\frac{\partial\mu_1}{\partial\lambda_1}\frac{1}{\mu_1-\mu}\Big)_x.$$

<u>Proof.</u> Let $H=\sum_{j<N} c_j H_j$ or $H=\mu_1$. As above, we shall write $H_{(j)}=\frac{\partial H}{\partial A_j}=\frac{\delta H}{\delta A_j}$. We multiply formula (30) by $(-1)^i\mu^{-(i+1)}$ and sum (replacing $A_{i,t}$ by $X_H A_i$):

$$\Phi_t(\mu)=\sum_{i=0}^{\infty}(-1)^i\mu^{-(i+1)}\sum_{j=0}^{\infty}(iA_{i+j-1}H_{(j),x}+j(A_{i+j-1}H_{(j)})_x).$$

Taking the summation on j to the outside, we obtain

$$\Phi_t(\mu)=\sum_{j=0}^{\infty}H_{(j),x}\sum_{i=0}^{\infty}(-1)^i(i+j)A_{i+j-1}\mu^{-(i+1)}+\sum_{j=0}^{\infty}jH_{(j)}\sum_{i=0}^{\infty}A_{i+j-1,x}\mu^{-(i+1)}. \tag{32}$$

We transform the two inner sums on the right:

$$\sum_{i=0}^{\infty}(-1)^i(i+j)A_{i+j-1}\mu^{-(i+1)}=(-1)^j\mu^j\frac{\partial}{\partial\mu}\left(\sum_{i=0}^{\infty}(-1)^{i+j-1}A_{i+j-1}\mu^{-(i+j)}\right)=$$

$$=(-1)^j\mu^j\frac{\partial}{\partial\mu}\left[\Phi(\mu)-\sum_{k=0}^{j-2}(-1)^kA_k\mu^{-(k+1)}\right];$$

$$\sum_{i=0}^{\infty}(-1)^iA_{i+j-1,x}\mu^{-(i+1)}=(-1)^{j-1}\mu^{j-1}\sum_{i=0}^{\infty}(-1)^{i+j-1}A_{i+j-1,x}\mu^{-(i+j)}\Big)=$$

$$=(-1)^{j-1}\mu^{j-1}\left[\Phi_x(\mu)-\sum_{k=0}^{j-2}(-1)^kA_{k,x}\mu^{-(k+1)}\right].$$

We now substitute these expressions into (32), multiply the formula obtained by $\partial\mu/\partial\lambda$, and use the following identities (Lemma 8.1): $\Phi_t(\mu)\partial\mu/\partial\lambda=-\mu_t=-X_H\mu$, $\Phi'(\mu)\partial\mu/\partial\lambda=1-\partial\mu/\partial\lambda$, $\Phi_x(\mu)\partial\mu/\partial\lambda=-\mu_x$. We obtain

$$-X_H\mu=\sum_{j=0}^{\infty}H_{(j),x}\left[(-1)^j\mu^j\left(1-\frac{\partial\mu}{\partial\lambda}\right)+\sum_{k=0}^{j-2}(-1)^{k+j}(k+1)A_k\mu^{j-k-2}\frac{\partial\mu}{\partial\lambda}\right]+$$

$$+\sum_{j=0}^{\infty}H_{(j)}\left[(-1)^j(\mu^j)_x+\sum_{k=0}^{j-2}(-1)^{k+j}jA_{k,x}\mu^{j-k-2}\frac{\partial\mu}{\partial\lambda}\right].$$

The first terms in the inner sums give a total derivative with respect to x of the form $\left(\sum_{j=0}^{\infty}H_{(j)}(-1)^j\mu^j\right)_x$. We shall prove that the remainder is zero. After this, part a) of Theorem 6.9 is rapidly obtained, while part b) follows from the formula $\mu_{1(j)}=\frac{\partial\mu_1}{\partial A_j}=(-1)^{j+1}\mu_1^{-(j+1)}\frac{\partial\mu_1}{\partial\lambda_1}$ (Lemma 8.1) which gives

$$-\sum_{j=0}^{\infty}\mu_{1(j)}(-1)^j\mu^j=\sum_{j=0}^{\infty}\mu_1^{-(j+1)}\mu^j\frac{\partial\mu_1}{\partial\lambda_1}=\frac{\partial\mu_1}{\partial\lambda_1}\frac{1}{\mu_1-\mu}.$$

We prove, finally, that the remainder vanishes. It suffices to do this for $H=\mu_1$. Indeed, the coefficient of the remainder for $\lambda_1^{-(j+1)}$, $j\geqslant 0$, then vanishes which corresponds to the remainder for $H=(-1)^jH_j$; this implies the required assertion for any linear combinations of the H_j as well by linearity.

The remainder has the form

$$\sum_{j=0}^{\infty}\mu_{1(j),x}\left[(-1)^{j+1}\mu^{j}\frac{\partial\mu}{\partial\lambda}+\sum_{k=0}^{j-2}(-1)^{k+j}(k+1)A_{k}\mu^{j-k-2}\frac{\partial\mu}{\partial\lambda}\right]+\sum_{j=0}^{\infty}\mu_{1(j)}\sum_{k=0}^{j-2}(-1)^{k+j}jA_{k,x}\mu^{j-k+2}\frac{\partial\mu}{\partial\lambda}.$$

We fix $s\geqslant 0$ and show that in this expression the coefficient of $\mu^{s}\frac{\partial\mu}{\partial\lambda}$ vanishes. This coefficient is equal to

$$\mu_{1(s),x}(-1)^{s+1}+\sum_{\substack{j,k\geqslant 0\\ j-k-2=s}}(-1)^{k+j}(k+1)A_{k}\mu_{1(j),x}+\sum_{\substack{j,k\geqslant 0\\ j-k-2=s}}(-1)^{k+j}jA_{k,x}\mu_{1(j)}.$$

Substituting here $\mu_{1(j)}=(-1)^{j+1}\mu_{1}^{-(j+1)}\frac{\partial\mu_{1}}{\partial\lambda_{1}}$, we obtain

$$\left(\mu_{1}^{-(s+1)}\frac{\partial\mu_{1}}{\partial\lambda_{1}}\right)_{x}+\sum_{j-k-2=s}(-1)^{k+1}(k+1)A_{k}\left(\mu_{1}^{-(j+1)}\frac{\partial\mu_{1}}{\partial\lambda_{1}}\right)_{x}+(-1)^{k+1}jA_{k,x}\mu_{1}^{-(j+1)}\frac{\partial\mu_{1}}{\partial\lambda_{1}}=\left(\mu_{1}^{-(s+1)}\frac{\partial\mu_{1}}{\partial\lambda_{1}}\right)_{x}+$$

$$+\left(\sum_{j-k-2=s}(-1)^{k+1}(k+1)A_{k}\mu_{1}^{-(j+1)}\frac{\partial\mu_{1}}{\partial\lambda_{1}}\right)_{x}+(s+1)\sum_{j-k-2=s}(-1)^{k+1}A_{k,x}\mu_{1}^{-(j+1)}\frac{\partial\mu_{1}}{\partial\lambda_{1}}=$$

$$=\left(\mu_{1}^{-(s+1)}\frac{\partial\mu_{1}}{\partial\lambda_{1}}\right)_{x}+\left(\mu_{1}^{-(s+1)}\Phi'(\mu_{1})\frac{\partial\mu_{1}}{\partial\lambda_{1}}\right)_{x}-(s+1)\mu_{1}^{-(s+2)}\Phi_{x}(\mu_{1})\frac{\partial\mu_{1}}{\partial\lambda_{1}}=$$

$$=\left(\mu_{1}^{-(s+1)}\frac{\partial\mu_{1}}{\partial\lambda_{1}}\right)_{x}+\left[\mu_{1}^{-(s+1)}\left(1-\frac{\partial\mu_{1}}{\partial\lambda_{1}}\right)\right]_{x}+(s+1)\mu_{1}^{-(s+2)}\mu_{1x}=0.$$

In passing to the last expression we have again used Lemma 8.1.

13. Miura's Conservation Laws

13.1. The evolution of u on the basis of system (25) with Hamiltonian $H=\frac{1}{3}H_{3}=\frac{A_{3}}{3}+A_{0}A_{1}$, has the form

$$u_{t}=\left(A_{1}+A_{0}u+\frac{u^{3}}{3}\right)_{x}-u_{y}\int_{0}^{y}(A_{0}+u^{2})_{x}d\eta=A_{1,x}+A_{0,x}u+A_{0}u_{x}+u^{2}u_{x}-yu_{y}A_{0,x}+u_{y}w, \tag{33}$$

where $w=-\int_{0}^{y}(u^{2})_{x}d\eta$. (In place of $\frac{1}{3}H_{3}$ it is possible to take cH_{3} with any constant c; $\frac{1}{3}$ was chosen to simplify the coefficients.) The evolution of μ, under the same system has, according to Theorem 3.2 a), the form

$$\mu_{t}=\left(-A_{1}+A_{0}\mu+\frac{\mu^{3}}{3}\right)_{x}=-A_{1x}+A_{0x}\mu+A_{0}\mu_{x}+\mu^{2}\mu_{x}. \tag{34}$$

The conservation laws of Miura type of this system are obtained from the following relations.

13.2 THEOREM. $[\mu'(\mu+u)^{-1}]_{t}-[(A_{0}+\mu^{2})\mu'(\mu+u)^{-1}]_{x}-[(w-A_{0,x}y)\mu'(\mu+u)^{-1}-v\mu']_{y}=0$, where $v=-\int_{0}^{y}u_{x}d\eta$, $\mu'=\partial\mu/\partial\lambda$ [cf. formula (21)].

COROLLARY. There exist constants $\tilde{H}_{n}, \tilde{F}_{n}, \tilde{G}_{n,0}, \tilde{G}_{n,1}, \tilde{G}_{n,2}$, such that under the evolution of u, h according to system (25) with Hamiltonian $\frac{1}{3}H_{3}$, there are local conservation laws of the form

$$\tilde{H}_{n,t}+\tilde{F}_{n,x}+(\tilde{G}_{n0}y+\tilde{G}_{n1}v+\tilde{G}_{n2}w)_{y}=0.$$

To derive the corollary from the theorem it is necessary to expand the expressions under the derivative sign in Theorem 13.2 in powers of λ^{-1}, using the formula $(\mu+u)^{-1}=\sum_{i=0}^{\infty}(-1)^i u^i \mu^{-(i+1)}$, and equate to zero the coefficient of $\lambda^{-(n+1)}$ in the sum thus obtained.

13.3. Proof of the Theorem. For the proof it is necessary to differentiate all three terms of the expression by the usual rules and multiply the result by $(\mu+u)^2$ to eliminate the denominator. In the expression thus obtained the right sides of the formulas for u_t (33), μ_t (34), $\mu_t'=2\mu\mu'\mu_x+\mu^2\mu_x'+A_{0x}\mu_x'$, and $v_y=-u_x$, $w_y=-(u^2)_x$ are then substituted. An algebraic sum of forty-six monomials is obtained, all of which happily cancel if it is noted that $\mu_y=\mu_y'=0$.

14. Compatibility of the Hamiltonian Structures

14.1. Theorem 6.7 contains a Hamiltonian structure on the ring $\mathcal{A}=Q[A_i^{(j)}]$ with operator B. In 6.5 a Hamiltonian structure on the ring $\mathcal{A}^0=Q[u^{(j)}, h^{(j)}]$ with operator $B^0=\begin{pmatrix}0 & \partial\\ \partial & 0\end{pmatrix}$ was considered. The rings $\mathcal{A}$ and $\mathcal{A}^0$ are connected by a homomorphism $\mathcal{A}\to\mathcal{A}^0: P\mapsto P^0$, where $\mathcal{A}_i^0=hu^i$, which commutes with the structural differentiations in $\mathcal{A}$ and $\mathcal{A}^0$; these we denote by the same letter $\partial: A_i^{(j)}\mapsto A_i^{(j+1)}$, $h^{(j)}\mapsto h^{(j+1)}$, $u^{(j)}\mapsto u^{(j+1)}$. Let $H\in\mathcal{A}$ be any Hamiltonian. With respect to it the evolution differentiations $X_H:\mathcal{A}\to\mathcal{A}$ and $X_{H^0}:\mathcal{A}^0\to\mathcal{A}^0$ can be defined. The precise formulation of the compatibility theorem 6.11 is as follows.

14.2. THEOREM. For any $P, Q\in\mathcal{A}$ we have $(X_Q P)^0=X_{Q^0}P^0$.

For the proof we need the following lemma.

14.3. LEMMA. For any $Q\in\mathcal{A}$ we have

$$\frac{\delta Q^0}{\delta u}=\sum_{j\geqslant 0}\left(jA_{j-1}\frac{\delta Q}{\delta A_j}\right)^0,\quad \frac{\delta Q^0}{\delta h}=\sum_{j\geqslant 0}u^j\left(\frac{\delta Q}{\delta A_j}\right)^0.$$

Proof. Let $\Omega^1\mathcal{A}$, $\Omega^1\mathcal{A}^0$ be the modules of differentials of the rings $\mathcal{A}$ and $\mathcal{A}^0$, respectively. They are freely generated over $\mathcal{A}$, $\mathcal{A}^0$, respectively, by the differentials $\delta A_i^{(j)}$, $\delta u^{(j)}$, $\delta h^{(j)}$. The differentiation ∂ extends to these modules, and the variational derivatives are uniquely determined by the following conditions:

$$\delta Q=\sum_{j\geqslant 0}\frac{\delta Q}{\delta A_j}\delta A_j+\partial\omega_j,\quad \delta Q^0=\frac{\delta Q^0}{\delta u}\delta u+\frac{\delta Q^0}{\delta h}\delta h+\partial\omega^0;$$
$$\omega\in\Omega^1 A,\quad \omega^0\in\Omega^1 A^0.$$

Further, the homomorphism $\mathcal{A}\to\mathcal{A}^0$ induces a module homomorphism $\Omega^1\mathcal{A}\to\Omega^1\mathcal{A}^0$ commuting with ∂ for which $(\delta Q)^0=\delta Q^0$. In particular, $(\delta A_j)^0=\delta A_j^0=jhu^{j-1}\delta u+u^j\delta h$, whence

$$(\delta Q)^0=\delta Q^0=\sum_{j\geqslant 0}\left(\frac{\delta Q}{\delta A_j}\right)^0(jhu^{j-1}\delta u+u^j\delta h)+\partial\omega^0.$$

Recalling that $hu^{j-1}=A_{j-1}^0$ and the characterization of $\delta Q^0/\delta u$, $\delta Q^0/\delta h$ indicated above, we obtain the required result.

<u>14.4. Proof of Theorem 14.2.</u> If the identity $(X_Q P)^0 = X_{Q^0} P^0$ holds for a given Q and for P_1, P_2 in place of P, then it also holds for this Q and $P_1 + P_2$, $P_1 P_2$, and ∂P_1 in place of P. This follows from the basic properties of differentiations, the commutativity of X_Q, X_{Q^0}, and the fact that the homomorphism reduces ∂. It therefore suffices to verify the identity for $P = A_i$, $i \geqslant 0$.

According to formula (30), we have

$$X_Q A_i = \sum_{j \geqslant 0} [(i+j) A_{i+j-1} (\delta Q / \delta A_j)_x + j A_{i+j-1,x} \delta Q / \delta A_j].$$

Therefore, according to Lemma 14.3,

$$(X_Q A_i)^0 = \sum_{j \geqslant 0} [(i+j) h u^{i+j-1} (\delta Q / \delta A_j)^0_x + j h^{(1)} u^{i+j-1} (\delta Q / \delta A_j)^0 + j(i+j-1) h u^{i+j-2} u^{(1)} (\delta Q / \delta A_j)^0] =$$

$$= i h u^{i-1} \left[\sum_{j \geqslant 0} u^j (\delta Q / \delta A_j)^0 \right]_x + u^i \left[\sum_{j \geqslant 0} j h u^{j-1} (\delta Q / \delta A_j)^0 \right]_x = i h u^{i-1} (\delta Q^0 / \delta h)_x + u^i (\delta Q^0 / \delta u)_x.$$

On the other hand,

$$X_{Q^0} A_i^0 = X_{Q^0} (h u^i) = i h u^{i-1} X_{Q^0} u + u^i X_{Q^0} h = i h u^{i-1} (\delta Q^0 / \delta h)_x + u^i (\delta Q^0 / \delta u)_x.$$

This completes the proof.

CHAPTER III
SOLUTIONS OF ALGEBRAIC TYPE

1. Introduction

1.1. This chapter is devoted mainly to a description of algebraic structures at the basis of explicit formulas for certain classes of solutions of Lax equations. These explicit formulas include both solutions of multisoliton type as well as solutions of quasiperiodic character written in terms of theta functions and also solutions of mixed type. An invariant definition of this class of solutions was given in 3.14 of Chap. II: for nonstationary solutions of Lax equations and the equations of Zakharov–Shabat these are solutions of a compatible system obtained by adding to the initial equation $L_t = [P, L]$ or $L_t + P_y = [P, L]$ the auxiliary stationary equation $[Q, L] = 0$. We shall call them solutions of algebraic type.

1.2. With this in mind the first object of study are the stationary equations $[Q, L] = 0$. According to Sec. 4 of Chap. II, commuting operators are connected by a polynomial relation with constant coefficients. This relation defines an affine algebraic curve C_0. Its most important characteristics are the following: the genus g of its nonsingular projective model C; points of C, lying at infinity (relative to C_0) and singular points of the curve C_0. The singular points of C_0, at least when they have the simplest form, i.e., are double points with separated tangents, are responsible for the multisoliton component of the solutions of the corresponding nonstationary equations: to each double point there corresponds one soliton. In the case of genus $g = 0$ this is the entire solution; in the case $g > 0$ it has a quasiperiodic component related to a motion on the complex torus — the Jacobian variety of the curve C. Points of C lying at infinity control the imbedding of the ring of functions $\mathcal{O}$ on C_0 into the ring of differential operators (in the simplest case the image of this imbedding is generated by Q and L, where $[Q, L]=0$). In particular, they make it possible to relate the order of operators to the orders of poles of functions at infinity.

One further fundamental invariant of a solution, its rank, is determined in terms of the imbedding of the ring $\mathcal{O}$ in $\mathcal{B}[\partial]$. The simplest description of rank is as follows: under weak assumptions regarding $\mathcal{O}$ and $\mathcal{B}[\partial]$ the order of operators in the image of $\mathcal{O}$, greater than some constant consists of all multiples of some integer $r \geqslant 1$; this r is the rank. Solutions of rank 1 have now been described much more completely than solutions of higher ranks. For the Korteweg–de Vries equation all algebraic solutions have rank 1, since for it $L=\partial^2+u$, and Q has a representation as a linear combination $\sum c_i \langle L^{s_i/2} \rangle$, where s_i are odd integers, so that the degree of Q is odd, and the monomials $L^m Q^n$ beginning from some place onward can have any integral order.

1.3. In order to clarify the mechanism of the difficulties related to the rank, we consider the curve C_0: $\Phi(\varkappa, \lambda)=0$ ($\Phi(Q, L)=0$ is the relation coupling Q and L) and for each point $c \in C_0$ we write the system of linear differential equations $Q\psi=\varkappa(c)\psi$, $L\psi=\lambda(c)\psi$. If $\mathcal{B}$ consists of (l, l) matrices of functions, then ψ is a column of functions of height l. It is found that for almost all points $c \in C_0$ the solutions of this system form a free (right) module F_c over the constants in $\mathcal{B}$ of the same rank which coincides with the number r introduced above. The system of linear spaces $\{F_c\}$ for each value x can be equipped with the structure of a vector bundle over C_0, which actually extends to infinity. This is proved (for the case $l=1$ and a nonsingular curve C_0) in the work of Drinfel'd [7]. It is probably true in general. In any event, all solutions considered in the literature possess this property. The functions ψ with respect to c can then be interpreted as sections of this bundle; the dependence on x is determined by the variation of the bundle. For $r=1$ the generating module of solutions described in a suitable trivialization of the bundle is called the Akhiezer function (see, e.g., the work of Krichever [16] and Matveev [44]). If a trivialization is not fixed following Drinfel'd [7], then the action of ∂_x on sections is determined by an appropriate connection ∇_x on the bundle described which extends $\frac{\partial}{\partial x}$ and is trivial along C.

Since the space of modules of one-dimensional vibrations over the curve C essentially coincides with its Jacobian J_C, finding the connection ∇_x reduces to constructing a suitable vector field on this Jacobian and verifying its integrability. The integral curves of the field are found to be rectilinear coverings of the torus J_C. On the other hand, the sections of the corresponding bundle lifted to the universal covering of J_C, are represented by the classical theta functions. This explains their occurrence in the explicit formulas for the unknown in Lax equations (see the derivation of these formulas in the paper of Matveev [44]).

From this the difficulties related to rank $r \geqslant 2$ become clear: spaces of modules of bundles with rank $\geqslant 2$ over algebraic curves as well as their analytic uniformization have been much less studied.

1.4. In this paper we have tried to get by with a minimum of machinery from algebraic geometry and analysis. Therefore, for us the basic objects will be special algebraic structures: the bimodules of Krichever–Drinfel'd which are described in Sec. 2 and their standard realization in Sec. 3. The bimodule technique makes it possible to reduce the problem of solving Lax equations to a problem of the variation with respect to x, t, y of certain algebraic functions on a curve C, the number of which is exactly the rank. For $r=1$ this problem, just as in other versions, is solved almost to the very end, and the basic conclusion regarding the rectilinearity of the motion on the Jacobian is attained in a very economical manner. This is the topic of Sec. 4.

In the case of rank $r \geqslant 2$ our technique makes it possible to construct at least some solutions of multisoliton type which we call matrix solitons. The justification for this name is that explicit formulas for them contain exponents of the form $\exp(K_1 x + K_2 y + \Omega t)$, where K_1, K_2, and Ω are matrices of rank r, rather than scalars as in earlier known formulas.

From the bundle point of view our construction is motivated from the fact that if a curve of genus zero with singularities is considered as C_0, then after lifting the bundle to a smooth model of C_0 it becomes invariable (a not very complicated theorem of Grothendieck). Therefore, variations of the bundles over C arise only due to the joining of fibers at those points of the smooth model C, which coalesce on C_0. The space of modules of bundles essential reduces to a product of linear groups over k, and the problem of describing suitable vector fields on them and connections can be handled.

It has been observed repeatedly in the literature that soliton solutions correspond to degenerations of conditionally periodic solutions. Nevertheless, a detailed algebrogeometric investigation of such solutions can provide useful information even in the case $r=1$, since the invariants of complex degeneration are well revealed in the language of the structure of singularities of curves but rather poorly in the language of theta functions (see the work of Matveev [44] where the simplest degenerations of hyperelliptic curves are treated analytically). This is even more applicable in the case $r \geqslant 2$, where the corresponding "matrix theta functions" are unknown.

In Sec. 5 we construct bimodules of arbitrary rank, and in Sec. 6 we investigate multisoliton solutions of rank 2. We shall here describe the simplest case of a single soliton. The ring $\mathcal{B}$ consists of meromorphic functions of x, t.

1.5. The simplest Lax equation having soliton solutions of rank two has the form $\partial_t L = [P, L]$, where $L = \partial_x^4 + v\partial_x^2 + w\partial_x + z$, $P = \omega\partial_x^2 + c\partial_x + u$; here $\omega, c \in \mathrm{R}$ are constants, and u, v, w, and z are unknown functions of x, t. Writing out the coefficients and eliminating u, we obtain the equivalent system of equations

$$-\omega^{-1}v_t = 2v_{xx} - 2w_x - c\omega^{-1}v_x,$$
$$-\omega^{-1}w_t = -w_{xx} + 2v_{xxx} + vv_x - c\omega^{-1}w_x - 2z_x,$$
$$-\omega^{-1}z_t = \frac{1}{2}v_{xxxx} + \frac{1}{2}vv_x - z_{xx} + \frac{1}{2}v_x w - c\omega^{-1}z_x.$$

According to Sec. 3 of Chap. II, this system can be represented in Hamiltonian form. Beginning with an operator L of just fourth order an interesting feature of the corresponding Hamiltonian of the differential operator appears for the first time: its coefficients depend explicitly on the unknown functions (in the present case only on v). The single-soliton solution of this system has the following form. We set $\xi = x + ct$ and choose an arbitrary real constant $a \in \mathbb{R}$, $a \neq 0$. We introduce the auxiliary functions

$$\mu_1 = 2a^2 \frac{\operatorname{ch} 2a\xi - \cos 2a\xi + 2\sin 4a^2\omega t}{\operatorname{ch} 2a\xi + \cos 2a\xi + 2\cos 4a^2\omega t},$$
$$\mu_2 = -2a \frac{\operatorname{sh} 2a\xi - \sin 2a\xi}{\operatorname{ch} 2a\xi + \cos 2a\xi + 2\cos 4a^2\omega t}.$$

Then

$$v = -4\mu_{2x}; \quad w = -6\mu_{2xx} - 4\mu_{1x} + 4\mu_2\mu_{2x};$$
$$z = -4\mu_{2xxx} - 6\mu_{1xx} + 8(\mu_{1x})^2 + 4(\mu_1\mu_2)_x + 6\mu_2\mu_{2xx} - 4\mu_{1x}\mu_2^2.$$

A special feature of the behavior of this soliton consists in the following: it moves as a whole with speed $-c$, but changes form with period $\pi(2a^2\omega)^{-1}$. At infinity (with respect to ξ) the amplitude of these variations decays rapidly, but at times when $\cos 4a^2\omega t = -1$, infinite ejection (in the functions μ_1, w, z) periodically occurs at its center $\xi = 0$. Thus, our solution displays a geyserlike behavior and deserves the name "geyseron" or "shooting soliton". We remark that the vibration of the solution with frequency $4a^2\omega$ formally occurs due to the same mechanism as the vibration of a free relativistic electron in the Schrödinger solution of the Dirac equation.

Another interesting feature of the behavior of the soliton is that the speed $-c$ is completely determined by the equations while the amplitude (measured by the factor a) may vary arbitrarily. In Sec. 6 where this example is considered on the basis of the general theory it is shown that $-c$ is also the common speed of the multisoliton solutions which thus represent a "coupled system" of solitons in contrast, e.g., to the multisoliton solutions of the Korteweg–de Vries equations.

1.6. Finally, the last section of this chapter is devoted to the description of a special class of solutions of the Benney equations and its analogues. In the notation of 1.3 of Chap. II these solutions are obtained by adding to the Benney equations the conditions $u_y = 0$, $h = \frac{(u+c)^2}{4}$, where c is any constant after which the problem reduces to a well known problem. The character of these invariant manifolds is not altogether clear in contrast to the auxiliary stationary problems for Lax equations. This hinders, in particular, the elimination of the condition $u_y = 0$. We remark that since the conservation laws H of

Benney do not depend on the derivatives of the moments A_n, it is here meaningless to apply the technique of restricting to extremals of conservation laws, since by the results of Chap. II the set $\frac{\delta H}{\delta A}=0$ is defined algebraically rather than by differential equations. Questions hereto related merit further investigation.

2. The Bimodules of Krichever and Drinfel'd

2.1. Notation. In this section k is any field of characteristic 0, $\mathcal{B}$ is a central k-algebra which is not necessarily commutative, $\partial, \partial_1, \partial_2: \mathcal{B}\to\mathcal{B}$ are three pairwise commuting k-differentiations, and $\mathcal{B}[\partial]$ is the ring of ∂-differential operators with coefficients in $\mathcal{B}$, and the commutation rule $\partial\circ b - b\partial = \partial b$. The differentiations ∂_1 and ∂_2 act coefficient-wise on $\mathcal{B}[\partial]$.

We fix some ring $\mathcal{O}$ of an affine curve over k. Suppose there is given an imbedding of k-algebra $i:\mathcal{O}\to\mathcal{B}[\partial]$, $i(\varphi)=L_\varphi$ for any $\varphi\in\mathcal{O}$. Obviously, $[L_\varphi, L_\psi]=0$ for all $\varphi, \psi\in\mathcal{O}$ so that prescribing i is equivalent to giving an entire class of solutions of the stationary Lax equations. We start by associating with the imbedding i a certain k-linear space $\mathcal{M}$ having a series of additional structures: the $(\mathcal{B},\mathcal{O})$-bimodules of Krichever and Drinfel'd. We then show that the imbedding i is recovered on the basis of the given bimodule, and solutions of the Lax equations and the equation of Zakharov—Shabat are also constructed. In the next sections we investigate and construct the bimodules themselves.

2.2. Construction of a Bimodule on the Basis of the Imbedding i. We set $\mathcal{M}=\mathcal{B}[\partial]$ and consider the following structures on $\mathcal{M}$.

a) $\mathcal{B}$ acts on $\mathcal{M}$ by multiplication on the left and $\mathcal{O}=i(\mathcal{O})$ by multiplication on the right.

b) In $\mathcal{M}$ there is a distinguished element 1 — the identity operator.

c) In $\mathcal{M}$ there is an increasing filtration

$$\mathcal{M}_i=\left\{\sum_{j<i} b_j\partial^j \mid b_j\in\mathcal{B}\right\}.$$

d) The k-linear operator $\nabla:\mathcal{M}\to\mathcal{M}$, $\nabla m=\partial\circ m$ acts on $\mathcal{M}$. These structures satisfy the following axioms which are trivially verified:

e) The actions of $\mathcal{B}$ and $\mathcal{O}$ commute so that $\mathcal{M}$ is a $(\mathcal{B},\mathcal{O})$-bimodule.

f) $\mathcal{M}_{-1}=\{0\}$; $\mathcal{B}\mathcal{M}_i\subset\mathcal{M}_i$ for all i.

g) For each $i\geqslant -1$ the factor $\mathcal{M}_{i+1}/\mathcal{M}_i$ is a free $\mathcal{B}$-module of rank 1; 1 is the free generator of M_0.

h) For all $b\in\mathcal{B}$, $\varphi\in\mathcal{O}$ we have $\nabla(bm)=\partial b\cdot m+b\nabla m$, $\nabla(mL_\varphi)=(\nabla m)L_\varphi$. In other words, ∇ is a ∂-connection on the $\mathcal{B}$-module $\mathcal{M}$.

i) For each $i\geqslant -1$, $\nabla\mathcal{M}_i\subset\mathcal{M}_{i+1}$ and ∇ induces an isomorphism of $\mathcal{B}$-modules $\mathcal{M}_i/\mathcal{M}_{i-1}\to\mathcal{M}_{i+1}/\mathcal{M}_i$.

A connection of the action of $\mathcal{O}$ with the filtration $\{M_i\}$ is not postulated, but it can sometimes be described explicitly on the basis of the following Lemma due to Drinfel'd.

2.3. LEMMA. We assume that $\mathcal{O}$ has no zero divisors, $i(\mathcal{O})\not\subset k$ and that for any $\varphi\in\mathcal{O}$ the coefficient of the leading power of ∂ of the operator L_φ is not a zero divisior in $\mathcal{B}$. We denote by C the smooth projective k-model of the affine curve $\operatorname{Spec}\mathcal{O}$. Then on C there exists a unique closed k-point ∞, the image of which does not lie in $\operatorname{Spec}\mathcal{O}$, and such that for all $\varphi\in\mathcal{O}$ the order $\operatorname{ord}_\partial L_\varphi$ of the operator L_φ is equal to $r\operatorname{ord}_\infty\varphi$. The integer $r\geqslant 1$ does not depend on φ and is called the rank of the imbedding i.

COROLLARY. Under the hypotheses of the lemma $\mathcal{M}_i L_\varphi\subset\mathcal{M}_{i+r\operatorname{ord}\varphi}$. If, moreover, the leading coefficient of L_φ is invertible, then multiplication by L_φ induces an isomorphism $\mathcal{M}_i/\mathcal{M}_{i-1}\to\mathcal{M}_{i+r\operatorname{ord}\varphi}/\mathcal{M}_{i+r\operatorname{ord}\varphi-1}$.

Proof. The mapping $v:\mathcal{O}\to\mathbf{Z}$, $v(\varphi)=-\operatorname{ord}_\partial L_\varphi$, possesses the following properties: $v(\mathcal{O})\neq\{0\}$, $v(k)=\{0\}$; $v(\varphi\psi)=v(\varphi)+v(\psi)$; $v(\varphi+\psi)\geqslant\min(v(\varphi),v(\psi))$. It is easy to see that v extends to the quotient field of $\mathcal{O}$ by the formula $v(\varphi\psi^{-1})=v(\varphi)-v(\psi)$, and all properties described are preserved. Therefore, v defines a k-valuation on the field of functions on C. Let ∞ be the k-point corresponding to this valuation. It does not lie on $\operatorname{Spec}\mathcal{O}$ since $v(\varphi)<0$ for some $\varphi\in\mathcal{O}$. The group of values of $\operatorname{ord}_\infty$ coincides with $\mathbf{Z}$, while the group of values of v is $r\mathbf{Z}$ for suitable integral r; this is the rank of i.

This lemma is usually employed in the following manner. Suppose that we are interested in solutions of Lax equations $[Q,L]=0$ with an operator L of low order, e.g., $L=\partial^2+u$ for the Korteweg–de Vries equations. In order to ensure the existence in $\mathcal{O}$ of a function φ with $\operatorname{ord}_\partial L_\varphi=2$, we must have on C a function with its only pole at ∞ of second order. In the case $\infty\in C(k)$ and the closure of $\operatorname{Spec}\mathcal{O}$ has no singularities at infinity this requirement means that C is a hyperelliptic curve, possibly degenerate, and $r=1$.

We shall now indicate how solutions of Lax equations are constructed on the basis of the bimodule $\mathcal{M}$.

2.4. THEOREM. Suppose that a bimodule $\mathcal{M}$ is given with structure 2.2 a)-d) and axioms 2.2 e)-i). Then there exists an imbedding $i:\mathcal{O}\to\mathcal{B}[\partial]$ and an isomorphism of bimodules $\mathcal{B}[\partial]\simeq\mathcal{M}$, which preserves all these structures.

Proof. We denote by $\mathcal{B}[\nabla]$ the ring of k-endomorphisms of the space $\mathcal{M}$ generated by multiplications by elements $\mathcal{B}$ and the operator ∇. The canonical mapping $\mathcal{B}\to\mathcal{B}[\nabla]$ is an imbedding by 2.2 g). Any element of $\mathcal{B}[\nabla]$ can be represented in the form $\sum_{i=0}^{N} b_i\nabla^i$ using the commutation rule $\nabla b-b\nabla=\partial b$ [2.2 h)]. This representation is unique by 2.2 i). There is therefore a unique ring isomorphism $\mathcal{B}[\nabla]\simeq\mathcal{B}[\partial]$, $\nabla\mapsto\partial$, which is the identity on $\mathcal{B}$. It follows from 2.2i) that $\mathcal{M}$ is a free $\mathcal{B}[\nabla]$-module of rank 1, generated by 1, and that the isomorphism $\mathcal{B}[\nabla]\simeq\mathcal{B}[\partial]$ extends to an isomorphism of the modules $\mathcal{M}=\mathcal{B}[\nabla]1\simeq\mathcal{B}[\partial]$. We point out that in this argument the $\mathcal{O}$-module structure on $\mathcal{M}$ was nowhere used. This is important for Theorem 2.6 below.

For each $\varphi \in \mathcal{O}$ we now define an operator $L_\varphi \in \mathcal{B}[\partial]$ by the formula $1\varphi = L_\varphi 1$ using the right action of $\mathcal{O}$ on $\mathcal{M}$. By the foregoing, L_φ exists and is unique. The mapping $\varphi \mapsto L_\varphi$ is k-linear, has trivial kernel, and is multiplicative by 2.2 e) which completes the proof.

The next two theorems concern nonstationary equations.

2.5. THEOREM. Suppose that a bimodule $\mathcal{M}$ is given with the following additional structures: a differentiation $\partial_1 : \mathcal{B} \to \mathcal{B}$ which extends to a ∂_1-connection $\nabla_1 : \mathcal{M} \to \mathcal{M}$, with $[\nabla, \nabla_1] = 0$ and $(\nabla, 1)\varphi = \nabla_1(1\varphi)$ for all $\varphi \in \mathcal{O}$. For $\varphi \in \mathcal{O}$ we construct the operator L_φ, as in 2.4, and we define $P \in \mathcal{B}[\partial]$ from the equality $\nabla_1 1 = P1$. Then $\partial_1 L_\varphi = [P, L_\varphi]$.

Proof. We have $(L_\varphi - \mathrm{Id}\cdot\varphi)1 = 0$ and $(\nabla_1 - P)1 = 0$. Hence $[\nabla_1 - P, L_\varphi - \mathrm{Id}\cdot\varphi]1 = 0$. But $[\nabla_1, \mathrm{Id}\cdot\varphi]1 = 0$ and $[P, \mathrm{Id}\cdot\varphi]1 = 0$, while $[\nabla_1, L_\varphi]1 = \partial_1 L_\varphi 1$, since $[\nabla_1, \nabla] = 0$. Therefore, $(\partial_1 L_\varphi - [P, L_\varphi])1 = 0$, and the operator on the left is zero in $\mathcal{B}[\partial]$, since 1 is the free generator of $\mathcal{M}$ over $\mathcal{B}[\partial]$, by the argument of the preceding section.

2.6. THEOREM. Suppose that a $\mathcal{B}$-module $\mathcal{M}$ is given with the structures 2.2 a)–i) in the description of which all mention of $\mathcal{O}$ is omitted. Suppose, moreover, the following additional structures are defined on $\mathcal{M}$: $\mathcal{B}$-connections $\nabla_j : \mathcal{M} \to \mathcal{M}$, extending ∂_j, $j = 1, 2$, where ∇, ∇_1, ∇_2 commute pairwise. We define operators $L, P \in \mathcal{B}[\partial]$ from the conditions $\nabla_1 1 = L1$, $\nabla_2 1 = P1$. Then $\partial_1 P + \partial_2 L = [P, L]$.

Proof. From the relation $[\nabla_1 - L, \nabla_2 - P]1 = 0$ we obtain, as above, the required assertion by noting that $[\nabla_1, \nabla_2] = 0$, $[\nabla_1, P] = \partial_1 P$, $[\nabla_2, L] = \partial_2 L$.

Remark. Although as is evident from the formulation of the theorem, the $\mathcal{O}$-module structure on $\mathcal{M}$ is not essential, those modules which are constructed by the method of Krichever and Drinfel'd carry this structure with some $\mathcal{O}$ by the very construction.

2.7. We remark in conclusion that in the ring of symbols $\mathcal{B}((\xi^{-1}))$ it is possible to obtain additional information on commutative subrings using Theorem 5.14 of Chap. II. Let $\mathcal{O}_N \subset \mathcal{B}((\xi^{-1}))$ be such a subring and suppose it contains an element $L = \sum_{i=0}^{N} u_i \xi^i$, for which u_N is a ∂-constant, is semisimple, and $u_{N-1} \in [u_N, \mathcal{B}] = \mathcal{B}^-$. Theorem 5.14, assuming surjectivity of $\partial : \mathcal{B}^+ \to \mathcal{B}^+$, implies the existence of the symbol $1 + \sum_{k=1}^{\infty} w_{-k}\xi^{-k} = Q$ with the property $Q^{-1} \circ L \circ Q = u_N \xi^N$. Therefore, the entire ring $Q^{-1} \circ \mathcal{O} \circ Q$ consists of symbols which commute with $u_N \xi^N$.

Now it is evident from the results of Sec. 1 of Chap. II that the space of symbols commuting with $u_N \xi^N$, consists precisely of symbols of the form $\sum v_i \xi^i$, where $v_i \in \mathrm{Ker}\,\partial \cap \mathcal{B}^+$, since all such symbols commute with $u_N \xi^N$, and the subfactors of this space, considering terms in the given interval of orders, have the required dimension.

This implies several conclusions.

a) If the leading term u_N of the operator L is a ∂-constant and is semisimple, if $u_{N-1}\in\mathscr{B}^-$, and $\mathscr{B}^+$ is commutative, then any two operators commuting with L, commute with one another.

Indeed, then the ring $(\operatorname{Ker}\partial\cap\mathscr{B}^+)((\xi^{-1}))$ is commutative.

This result is applicable to the case $\mathscr{B}=M_l(\mathscr{B}_0)$, $u_N=\operatorname{diag}(c_1,\ldots,c_l)$, $c_i\neq c_j$ for $i\neq j$. In this case the entire ring of symbols commuting with $u_N\xi^N$, consists of diagonal matrices of symbols with constant coefficients. Therefore, l "partial" order functions are defined on it: the orders of the symbols at the site (ii), $1\leqslant i\leqslant l$. The argument of Lemma 2.3 can be applied to each of these separately, so we obtain the following result.

b) If in the commutative subring $\mathcal{O}\subset\mathscr{B}[\partial]$ with $\mathscr{B}=M_l(\mathscr{B}_0)$ there is an element L with invertible $u_N=\operatorname{diag}(c_1,\ldots,c_l)$, $c_i\neq c_j$, $u_{N-1}\in\mathscr{B}^-$, then on the smooth projective model of $\operatorname{Spec}\mathcal{O}$ there are $\leqslant l$ infinitely distant points the orders of the poles at which are determined by the partial order functions of the symbols of the corresponding operators.

In the centralizer $\tilde{L}$ in $\mathscr{B}((\xi^{-1}))$ the partial orders do not depend on one another, but in $\mathcal{O}$ they may be related and even completely determined by one of them. Therefore, among the points at infinity there are not necessarily l distinct points.

3. The Standard Realization of a Bimodule Over a Field

3.1. In this section we assume that the basic differential ring $\mathscr{B}$ is a field. This does not reduce appreciably the generality of the results, since for almost all initial conditions the solutions of the stationary and nonstationary Lax equations are locally analytic, and we can seek them in the field of germs of meromorphic functions. The base field k in applications is $\mathbf{R}$ or $\mathbf{C}$; it is assumed to be algebraically closed in $\mathscr{B}$ and to coincide with the ∂-constants in $\mathscr{B}$.

Let $\mathcal{O}\subset\mathscr{B}[\partial]$ be a commutative subring containing k. It contains no zero divisors. The orders of its elements form a semigroup; hence, there exists an integer n_0, such that all those orders greater than n_0, form an arithmetic progression with difference r. From the proof of Lemma 2.3 it is clear that r coincides with the rank of the $(\mathscr{B},\mathcal{O})$-bimodule $\mathscr{M}=\mathscr{B}[\partial]$. The semigroup of orders of $\mathcal{O}$ is finitely generated; choosing its generators and operators of the corresponding orders we find that the ring $\mathcal{O}$ is finitely generated over k. It is thus the ring of an absolutely irreducible curve. Let ∞ be the point of its smooth model defining the orders of operators of $\mathcal{O}$ as in Lemma 2.3. We shall assume that the leading coefficient of at least one element of $\mathcal{O}$ of nonzero order is a constant. From formula (1) of Chap. II it then follows that all leading coefficients of elements of $\mathcal{O}$ are constants.

In this section we consider $\mathscr{M}$ as a left $\mathscr{B}\otimes_k\mathcal{O}=\mathscr{B}\mathcal{O}$-module with action $(bf)m=bmf$, $b\in\mathscr{B}$, $m\in\mathscr{M}$, $f\in\mathcal{O}$, and we realize $\mathscr{M}$ as a submodule of vectors in the r-dimensional coordinate space K^r over the quotient field K of the ring $\mathscr{B}\mathcal{O}$. This realization is called the standard realization.

In order to describe it and carry over to its language the basic structures related to $\mathcal{M}$, it is convenient to use the elements of the theory of algebraic curves. The field K is the field of functions on an algebraic curve over $\mathcal{B}$ — the smooth complete model of Spec $\mathcal{O}$ with field of constants extended to $\mathcal{B}$. The points of the field are local rings in K containing $\mathcal{B}$. Those field points which come from points of the complete smooth model of Spec $\mathcal{O}$, are called constants and the remaining points are variables. For example, ∞ is a constant point. The principal part of an element $f \in K$ at a point P is called the class of f modulo the maximal ideal corresponding to P.

Any k-differentiation of $\mathcal{B}$, in particular ∂, extends uniquely to K by the condition $\partial \mathcal{O} = \{0\}$. The same letter ∂ is used to denote the coordinate-wise action on K^r.

We can now formulate a theorem on the standard realization.

3.2. THEOREM. The bimodule $\mathcal{M} = \mathcal{B}[\partial]$ is canonically isomorphic to the $\mathcal{B}\mathcal{O}$-submodule $\hat{\mathcal{M}} \subset K^r$ with the following properties:

a) $(\mathcal{B}\mathcal{O})^r \subset \hat{\mathcal{M}}$; the factor space $\hat{\mathcal{M}}/(\mathcal{B}\mathcal{O})^r$ is finite-dimensional over $\mathcal{B}$.

b) The element $1 \in \mathcal{M}$ is represented by the vector $\hat{1} = (1, 0, \ldots, 0)^t \in \hat{\mathcal{M}}$.

c) Let $\hat{\mathcal{M}}_{(i)}$ denote the subset of elements of $\hat{\mathcal{M}}$, the j-th coordinates of which have poles at ∞ of order not exceeding $\frac{i-j}{r}$, $j = 0, \ldots, r-1$. Then $\hat{\mathcal{M}}_{(i)} \subset (\mathcal{M}_i)\hat{}$ for all i and $\hat{\mathcal{M}}_{(i)} = (\mathcal{M}_i)\hat{}$ for all $i \geqslant i_0$, where i_0 is a suitable constant.

d) The connection $\nabla : \mathcal{M} \to \mathcal{M}$ is induced by a connection on K^r of the form $\hat{\nabla} = \partial + \Delta$, where $\Delta \in M_r(K)$ is a matrix of the form (for $r \geqslant 2$)

$$\Delta = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ . & . & . & . & . \\ \Delta_0 & \Delta_1 & \ldots & & \Delta_{r-1} \end{pmatrix}, \quad \Delta_i \in K.$$

Proof. We note first of all that the $\mathcal{B}\mathcal{O}$-module $\mathcal{M}$ has no torsion. Indeed, suppose Γ is the semigroup of the orders of elements of $\mathcal{O}$. For each $n \in \Gamma$ we choose an element $f_n \in \mathcal{O}$ with leading ∂^n. Clearly, the elements $\{f_n\}$ form a k-basis of $\mathcal{O}$. They therefore form a $\mathcal{B}$-basis of $\mathcal{B}\mathcal{O}$. If $\sum b_n f_n \neq 0$ is any element of $\mathcal{B}\mathcal{O}$ and $m \in \mathcal{M}$, $m \neq 0$, then $\left(\sum b_n f_n\right) m = \sum b_n m f_n$. Choosing the greatest n_0 with $b_{n_0} \neq 0$, we find that the leading term of $\left(\sum b_n f_n\right) m$ is equal to the leading term $b_{n_0} m f_{n_0}$, i.e., is different from zero. This means that $\mathcal{M}$ has no nontrivial torsion. This implies that the canonical mapping $\mathcal{M} \to K \otimes_{\mathcal{B}\mathcal{O}} \mathcal{M}$: $m \mapsto 1 \otimes m$ is an imbedding.

We establish the isomorphism $K \otimes_{\mathcal{B}\mathcal{O}} \mathcal{M} = K^r$. We set $M' = \mathcal{B}1\mathcal{O} + \ldots + \mathcal{B}\nabla^{r-1}1\mathcal{O}$. Since all elements of Γ are divisible by r, the orders of the elements of $\mathcal{B}\nabla^j 1\mathcal{O}$ are all congruent to $j \bmod r$, and therefore the sum of the $\mathcal{B}\nabla^j 1\mathcal{O}$ is direct. Further, the orders of elements of $\mathcal{M}'$ include all integers greater than some n_0, and therefore $\mathcal{M} = \mathcal{M}' + \mathcal{M}_{n_0}$. Since the space $\mathcal{M}_{n_0}$ is finite-dimensional over $\mathcal{B}$, the $\mathcal{B}\mathcal{O}$-module $\mathcal{M}/\mathcal{M}'$ is a torsion module. Indeed, if $f \in \mathcal{O}$ is any element of nonzero order and $m \in \mathcal{M}$, then there is always a nontrivial relation of the form $\sum b_i m f^i \in \mathcal{M}'$, i.e., $\sum b_i f^i$ is annihilated by $m \bmod \mathcal{M}'$.

Thus, the imbedding $\mathcal{M}' \to \mathcal{M}$ induces an isomorphism $K \otimes \mathcal{M}' \to K \otimes \mathcal{M}$. In view of the above, $K \otimes \mathcal{M}'$ has the canonical basis $\{\nabla^j 1 \mid 0 \leqslant j \leqslant r-1\}$ (more precisely, $\{1_K \otimes \nabla^j 1\}$). We identify $K \otimes \mathcal{M}'$ and $K \otimes \mathcal{M}$ with the space K^r of columns of height r with coordinates in K by means of this basis. This defines the canonical imbedding $\mathcal{M} \to K^r$. We denote its image by $\hat{\mathcal{M}}$. The inclusion $(\mathcal{B}\mathcal{O})^r \subset \hat{\mathcal{M}}$ and the finite-dimensionality of $\hat{\mathcal{M}}/(\mathcal{B}\mathcal{O})^r$ over $\mathcal{B}$ have actually already been proved: $(\mathcal{B}\mathcal{O})^r$ is the image of $\mathcal{M}'$, while $\hat{\mathcal{M}}/(\mathcal{B}\mathcal{O})^r$ is isomorphic to the factor space $\mathcal{M}_{n_0}/\mathcal{M}' \cap \mathcal{M}_{n_0}$. The element $1 \in \mathcal{M}'$ is represented by the vector $(1, 0, \ldots, 0)^t$.

The filtration with respect to order is easily described on $(\mathcal{B}\mathcal{O})^r$: to the vector $(b_0 f_0, \ldots, b_{r-1} f_{r-1})^t$, $(b_j \in \mathcal{B}, f_j \in \mathcal{O})$ there corresponds the operator $\sum b_j \nabla^j 1 f_j$, the order of which is equal to $\max(j + \operatorname{ord} L_{f_j} \mid b_j \neq 0) = \max(j - r \operatorname{ord}_\infty f_j \mid b_j \neq 0)$ by Lemma 2.3. Therefore, $(b_0 f_0, \ldots, b_{r-1} f_{r-1})^t \in (\mathcal{M}_i)^\wedge$, if and only if $j - r \operatorname{ord}_\infty f_j \leqslant i$, i.e., the order of the pole of f_j at ∞ does not exceed $\frac{i-j}{r}$ for $b_j \neq 0$. Taking the order of the pole of the zero element to be $-\infty$, we may remove the condition $b_j \neq 0$.

Thus $\hat{\mathcal{M}}_{(i)} \subset (\mathcal{M}_i)^\wedge$. In $\hat{\mathcal{M}}$ we now choose a finite-dimensional subspace over $\mathcal{B}$ complementary to $(\mathcal{B}\mathcal{O})^r$. The orders of the poles at ∞ of all the coordinates of elements of this subspace are uniformly bounded. They are therefore contained in $\hat{\mathcal{M}}_{(i)}$ for the sufficiently large i. This implies that $\hat{\mathcal{M}}_{(i)} = (\mathcal{M}_i)^\wedge$ for sufficiently large i. We have now verified assertions a)-c) of Theorem 3.2.

In order to establish d) we note first of all that the connection ∇ extends uniquely to a connection $\hat{\nabla}: K \otimes \mathcal{M} \to K \otimes \mathcal{M}$ with $\nabla(f \otimes m) = \partial f \otimes m + f \otimes \nabla m$ for all $f \in K$, $m \in \mathcal{M}$ (this is a standard fact regarding the extension of connections on a localization).

With the identification of $K \otimes \mathcal{M}$ with K^r we obtain a connection $\hat{\nabla}: K^r \to K^r$. The difference $\hat{\nabla} - \partial$ is a K-linear mapping $K^r \to K^r$. Let Δ be the matrix of this mapping. Since $\nabla(\nabla^j 1) = \nabla^{j+1} 1$, we have on setting $e_j = (0 \ldots 0 \underset{j}{1} 0 \ldots 0)^t$ a representative of $\nabla^j 1$ in K^r,

$$(\hat{\nabla} - \partial) e_j = \begin{cases} e_{j+1} & \text{for } j \leqslant r-2, \\ \sum_{j=0}^{r-1} \Delta_j e_j \; (\Delta_j \in K) & \text{for } j = r-1. \end{cases}$$

Therefore, Δ has the form indicated in the theorem. This completes the proof.

3.3. Remark. The standard realization of $\mathcal{M}$ shows that the matrix Δ, or its last row $(\Delta_0, \ldots, \Delta_{r-1})$ is essentially the unique invariant of $\mathcal{M}$ (for given $\mathcal{B}$ and $\mathcal{O}$): $\hat{\mathcal{M}}$ is recovered from it as $\sum_{j=0}^{\infty} \mathcal{B}(\partial + \Delta)^j (1, 0, \ldots, 0)^t$, together with the filtration, the action of $\mathcal{B}$ and the connection $\hat{\nabla}$. However, $(\Delta_0, \ldots, \Delta_{r-1})$ cannot be chosen arbitrarily, because the conditions that $\hat{\mathcal{M}}$ be invariant with respect to multiplication by $\mathcal{O}$, have finite type over $\mathcal{O}$, and that the filtration of $\hat{\mathcal{M}}$ be described in terms of the behavior at ∞, as in Theorem 3.2, impose strong and nontrivial restrictions on Δ. We shall occupy ourselves with them in the next sections. Here we note only that part of these restrictions have the character of differential equations: Δ is to be considered as a collection of functions on the model of $\operatorname{Spec}\mathcal{O}$, depending on parameters x, (x, t), (x, y, t) and satisfying certain differential equations in these parameters.

4. Bimodules of Rank 1

4.1. In this section we give a classification of bimodules of rank 1 which satisfy two additional conditions. We can formulate the first one immediately: it is that the ring $\mathcal{B}\mathcal{O}$ consist precisely of all functions of the field K, having a pole only at the point ∞. This condition is equivalent to the condition that $\operatorname{Spec}\mathcal{O}$ be obtained from its smooth complete model by excising the point at ∞. The second condition will be formulated later. We fix $\mathcal{M}$ and its standard realization $\hat{\mathcal{M}} \in K$.

Let $\Delta \in K$ be defined as in Theorem 3.2.

We call a pole of $\hat{\mathcal{M}}$ any point P of the field K, for which there exists an element $f \in \hat{\mathcal{M}}$, having a pole of order $\geqslant 1$ at P. We call the order of the pole P the greatest order of the pole of $f \in \hat{\mathcal{M}}$ at P (∞ if there is no greatest order).

4.2. THEOREM. a) There exists a finite number of nonconstant points $P_1, \ldots, P_s$ of the field K and positive integers $a_1, \ldots, a_s$ such that $\hat{\mathcal{M}}$ consists of all functions of the field having poles of order $\leqslant a_i$ at $P_i\,(1 \leqslant i \leqslant s)$ and a pole of any order at ∞.

b) The degree of the divisor $D = \sum a_i P_i$ is equal to the genus g of the field K.

c) Let z be an element of the field of fractions of the ring $\mathcal{O}$, which is finite at all the points $P_1, \ldots, P_s$, ∞ and such that $z - z(P_i)$, $z - z(\infty)$ are local parameters at the points P_i, ∞, respectively. Then the principal part of Δ at the points P_i, ∞, have the respective forms $-\frac{a_i \partial z(P_i)}{z - z(P_i)} + c_i$, $\frac{b}{z - z(\infty)} + c_\infty$, where c_i, c_∞ line in the fields of the residue classes of the points P_i, ∞. The divisor of the poles of Δ is precisely $\infty + \sum_{i=1}^{s} P_i$.

Proof. Let P be any point of the field K, and let $\mathcal{O}_P$ be its local ring. It is known that ∂ takes $\mathcal{O}_P$ into itself (cf. [20, Lemma 2]). Since $\hat{\mathcal{M}} = \sum_{j=0}^{\infty} \mathcal{B}(\partial + \Delta)^j 1$, all poles of $\hat{\mathcal{M}}$ must be contained among the poles of Δ, so that $\hat{\mathcal{M}}$ has a finite number of poles.

We shall first show that at ∞ $\hat{\mathcal{M}}$ has a pole of exactly first order. If Δ had no pole at ∞ then neither would $\hat{\mathcal{M}}$, and this would contradict the inclusion $\mathcal{B}\mathcal{O} \subset \hat{\mathcal{M}}$. If the pole of Δ at ∞ were of order $a \geqslant 2$, then the order of the pole of $(\partial + \Delta)^j 1$ would be precisely ja (an easy induction on j using the fact that ∂ does not increase the order of a pole at a constant point). But $(\partial + \Delta)^j 1$ corresponds to the operator $\partial^j \in \mathcal{B}[\partial] = \mathcal{M}$, and by Theorem 3.2 c) the order of the pole at ∞ of the function $(\partial + \Delta)^j 1$ for sufficiently large j must coincide with j. Therefore, the case $a \geqslant 2$ is impossible.

Suppose now that $P \neq \infty$ is a constant point. We shall show that it cannot be a pole of Δ and hence of $\hat{\mathcal{M}}$. Indeed, otherwise, the same argument as in the preceding paragraph shows that $\hat{\mathcal{M}}$ contains functions with a pole at P of arbitrarily higher order. But $\hat{\mathcal{M}}$ is finitely generated over $\mathcal{B}\mathcal{O}$, while elements of $\mathcal{B}\mathcal{O}$ have no poles at P at all; this is thus impossible.

We now consider a nonconstant pole P_i of the module $\hat{\mathcal{M}}$. Its order is finite, for otherwise, as above, $\hat{\mathcal{M}}$ could not be finitely generated over $\mathcal{B}\mathcal{O}$, since the elements of $\mathcal{B}\mathcal{O}$ have no poles at P_i. Suppose that $a_i \geq 1$ is the order of this pole. We choose z in the field of fractions at $\mathcal{O}$, as in part b) of the formulation of Theorem 4.2 and an element $f \in \hat{\mathcal{M}}$ with a pole of order a_i at P_i. It may be assumed that $f = (z - z(P_i))^{-a_i} + O((z - z(P_i))^{-a_i+1})$. Since $(\partial + \Delta) f \in \hat{\mathcal{M}}$ must also have a pole at P_i of order no greater than a_i, and

$$(\partial + \Delta) f = \Delta f + \frac{a_i \partial z(P_i)}{(z - z(P_i))^{a_i+1}} + O((z - z(P_i))^{-a_{i+1}},$$

we find that the expansion of Δ at P_i must begin with $-\frac{a_i \partial z(P_i)}{z - z(P_i)}$.

We set $D = \sum_{i=1}^{s} a_i P_i$ and denote by $\mathcal{L}(D + j\infty)$ in the linear space over $\mathcal{B}$ of all functions of K, the divisor of the poles of which does not exceed $D + j\infty$. According to **Theorem 3.2 a) and c), for sufficiently large j we have $\mathcal{L}(j\infty) \subset \hat{\mathcal{M}}_j \subset \mathcal{L}(D + j\infty)$.** Therefore, $\hat{\mathcal{M}}_j = \mathcal{L}(D + j\infty)$ for $j \gg 0$, since in $\hat{\mathcal{M}}_j$ there are elements with principal parts at P_i of order exactly a_i, and taking linear combinations of them with suitable coefficients in $\mathcal{B}\mathcal{O}$, we can make these principal parts whatever we wish.

Now the dimension of $\hat{\mathcal{M}}_j$ over $\mathcal{B}$ is equal to $j+1$, while the dimension of $\mathcal{L}(D + j\infty)$ for $j \gg 0$ by the Riemann-Roch theorem is equal to $\deg D - g + 1 + j \deg \infty$. Therefore, ∞ is a point of first degree and $\deg D = g$. This completes the proof.

4.3. We now formulate the second condition imposed on our bimodule $\mathcal{M}$. It is that the divisor $P_1 + \ldots + P_s$ defined in the preceding section be nonspecial, i.e., there exists no differential of first kind that vanishes at $P_1 + \ldots + P_s$. But then from the existence of a nonconstant function Δ with a divisor of poles $P_1 + \ldots + P_s + \infty$ it follows that $\deg(\Sigma P_i) = g$, $a_i = 1$; the function Δ is uniquely defined up to an additive constant.

The divisor $P_1 + \ldots + P_s$ is clearly nonspecial if this is so at the initial point $x = x_0$, and the set of such initial conditions is dense in the g-fold symmetric product of spec $\mathcal{O}$.

We shall now show that the conditions on Δ become a condition on the linear variation of the Jacobian coordinates of $P_1 + \cdots + P_s$ with respect to x in the functional case. We shall assume that $\mathcal{B}$ is the ring of germs of meromorphic functions of x and represent the P_i as germs of holomorphic paths on the Riemann surface of Spec $\mathcal{O}$, parametrized by x. It may be assumed that they are distinct, i.e., $s = g$.

We choose a basis of differentials of first kind $\omega_1, \ldots, \omega_g$ in the field of fractions of $\mathcal{O}$; a function z as in Theorem 4.2 c), and set $\omega_i = u_i dz$. The functions u_i are finite at the points $P_1, \ldots, P_g$. We set $J_i = J_i(x) = \sum_{j=1}^{g} \int_{\infty}^{P_j} \omega_i$. We define the element $b \in \mathcal{B}$ from the condition $\Delta = \frac{b}{z - z(\infty)} + O(1)$ near ∞.

4.4. THEOREM. a) $\partial_x J_i = b u_i(\infty)$ for all $i = 1, \ldots, g$. Hence the Jacobian coordinates of the divisor $P_1 + \ldots + P_g$ move in a constant direction $(u_i(\infty))$ with speed $b(x)$ at the point x.

b) Conversely, suppose $D = P_1 + \ldots + P_g$ is a nonspecial divisor of the field K, the Jacobian coordinates of which vary with x as in part a) with some function $b \in \mathcal{B}$. Then there exists a unique (up to a constant of $\mathcal{B}$) function Δ with principal part $\frac{b}{z - z(\infty)}$ at ∞ and divisor of poles $D + \infty$, and $\sum_{j=0}^{\infty} \mathcal{B}(\partial + \Delta)\partial 1$ is a $(\mathcal{B}, \mathcal{O})$-bimodule.

Proof. a) For any differential ω of the field K we have $\sum_Q \operatorname{res}_Q(\omega\Delta) = 0$. There can be poles of $\omega_i\Delta$ only at the points $P_1, \ldots, P_g, \infty$, so that $\sum_{j=1}^{g} \operatorname{res}_{P_j}(\omega_i\Delta) = -\operatorname{res}_\infty(\omega_i\Delta)$. Therefore,

$$\partial_x J_i = \sum_{j=1}^{g} u_i(P_j)\,\partial_x z(P_j) = \sum_{j=1}^{g} \operatorname{res}_{P_j} \frac{u_i(z)\,\partial_x z(P_j)}{z - z(P_j)}\,dz = -\sum_{j=1}^{g} \operatorname{res}_{P_j}(\omega_i\Delta) = \operatorname{res}_\infty(\omega_i\Delta) = b u_i(\infty).$$

b) Conversely, suppose that $P_1 + \ldots + P_g$ varies as in part a). Then the calculation of $\partial_x J_i$ read in the reverse order shows that at the points P_j the principal part of Δ begins with $-\frac{\partial_x z(P_j)}{z - z(P_j)}$, since the matrix $(u_i(P_j))$ is nondegenerate because $P_1 + \ldots + P_g$ is nonspecial. Arguments analogous to those given in the proof of Theorem 4.2 show that $\sum_{j\geq 0} \mathcal{B}(\partial + \Delta)^j 1 = \bigcup_{j\geqslant 0} \mathcal{L}(P_1 + \ldots + P_g + j\infty)$, and all the axioms of a bimodule for this space are verified without difficulty.

4.5. Theorem 4.4 gives almost a complete classification of bimodules of rank 1 over a field. In order to apply it to find solutions of nonstationary equations, it is useful to have in mind the following situations.

a) Let us assume that $\mathcal{B}$ is a ring of germs of functions of x, t. To solve nonstationary Lax equations we must extend ∂_t to $\hat{\mathcal{M}}$ by the condition $[\hat{\nabla}_x, \hat{\nabla}_t] = 0$. This extension of $\hat{\nabla}_t$ must have the form $\partial_t + \Delta_t$ where $\Delta_t \in K$. Since $(\partial_t + \Delta_t)1 \in \hat{\mathcal{M}}$, we necessarily have $\Delta_t \in \hat{\mathcal{M}}$. The commutation condition has the form $\partial_x \Delta_t = -\partial_t \Delta$. Since each element of $\hat{\mathcal{M}}$ is uniquely determined by its principal part at ∞ (because $P_1 + \ldots + P_g$ is nonspecial), Δ_t can only have the form $-\int_{x_0}^{x} \partial_t \Delta dx +$ (an element with a ∂_x-constant principal part at ∞). On the other hand, since the connection $\hat{\nabla}_t$ must take $\hat{\mathcal{M}}$ into itself, the behavior of Δ_t near P_j is determined by the same sort of conditions as in Theorem 4.2 c). The argument used in the proof of Theorem 4.4 shows that the motion of $P_1 + \ldots + P_g$ with respect to t also becomes rectilinear in Jacobian coordinates; its direction and speed are determined by the principal part of Δ_t at ∞. The same is true for the motion in y for solution of the Zakharov–Shabat equations.

b) We set $\psi=e^{\int_{x_0}^{x}\Delta dx}$. Then for any $m\in\hat{\mathcal{M}}$ we have $[(\partial_x+\Delta)m]\psi=\partial_x(m\psi)$. Therefore, $\mathcal{M}$ can also be realized as the space $\hat{\mathcal{M}}\psi$, to which all structures carry over in an obvious way, while the connection $\hat{\nabla}_x$ goes over in to ordinary differentiation with respect to x. The function ψ is called the (stationary) Akhiezer function. Prescribing this function is equivalent to prescribing Δ and hence the bimodule $\mathcal{M}$ in the standard realization. In the papers of Krichever [16] and Matveev [44] this function is taken as the initial object. If Δ does not depend on t nor Δ_t on x, then in the nonstationary case the function $\psi(x, t)=\exp\left(\int_{x_0}^{x}\Delta dx+\int_{t_0}^{t}\Delta_t dt\right)$. On $\hat{\mathcal{M}}\psi$ the connections $\hat{\nabla}_x$ and $\hat{\nabla}_t$ become ∂_x and ∂_t, respectively.

c) Since Δ, Δ_t are determined by their divisors and principal parts at ∞, they can be written out explicitly in terms of the classical Riemann theta functions. The coefficients of all operators which enter in the solution of Lax equations are expressed in terms of Δ, Δ_t and their derivatives. This leads to explicit formulas for the solutions. For further details we refer the reader to the papers [16], [44], and Sec. 6 of Chap. IV.

5. Bimodules of Higher Rank over a Rational Curve with Double Points

In this section a class of bimodules is constructed which may have arbitrary rank $r\geqslant 1$. For the motivation see 1.4 and 2.2. The realizations given here are close to the standard realizations but do not coincide with them.

5.1. The Initial Objects. We set $k=\mathbf{R}$ or $\mathbf{C}$, and we let $\mathcal{B}_0$ be a ring of germs of k-analytic or infinitely differentiable functions of the variable x (for stationary Lax equations), of (x, t) (for nonstationary Lax equations), or of (x, t, y) (for the equations of Zakharov–Shabat); $\mathcal{B}=M_l(\mathcal{B}_0)$ (the matrix algebra of order l over $\mathcal{B}_0$); $\partial=\partial_x=\partial/\partial x$, $\partial_1=\partial_t=\partial/\partial t$, $\partial_2=\partial_y=\partial/\partial y$.

To construct $\mathcal{O}$ we choose $2N$ distinct numbers $\alpha_i, \beta_i\in\mathbf{C}$, $i=1,\ldots,N$ with the condition $\alpha_i=\bar{\beta}_i$ if $k=\mathbf{R}$ and we set

$$\mathcal{O}=\{f(\lambda)\in k[\lambda]\,|\,\forall i=1,\ldots,N,\ f(\alpha_i)=f(\beta_i)\}. \tag{1}$$

It is obvious that $\mathcal{O}$ is a ring of functions on the affine line with N pairs of identified points realized as a subring of the functions on the line itself, i.e., $k[\lambda]$.

Similarly, we construct $\mathcal{M}$ as a submodule of the trivial $(\mathcal{B}, k[\lambda])$-bimodule $\mathcal{N}=\mathcal{B}[\lambda]^r$ of rank r with identification conditions at the points (α_i, β_i).

We shall write elements of $\mathcal{B}^r$ and $\mathcal{N}=\mathcal{B}[\lambda]^r$ as columns of height r with coordinates in $\mathcal{B}$, $\mathcal{B}[\lambda]$, respectively (λ commutes with $\mathcal{B}$). We introduce on $\mathcal{N}$ the left action of $\mathcal{B}$ by the formula

$$b\begin{pmatrix}b_1\\ \vdots\\ b_r\end{pmatrix}=\begin{pmatrix}b_1b^t\\ \vdots\\ b_rb^t\end{pmatrix},$$

where t is the transpose in $\mathcal{B}=M_l(\mathcal{B}_0)$. We define the right action of $k[\lambda]$ on $\mathcal{N}$ as coordinate-wise multiplication on the right by $k[\lambda]E_r$, where E_r is the identity matrix in $\mathrm{GL}(r,\mathcal{B})$. Finally, we introduce on $\mathcal{N}$ the natural left action of the group $\mathrm{GL}(r,\mathcal{B})$. Obviously, the actions of $\mathcal{B}$ and $k[\lambda]$ on $\mathcal{N}$ are effective and commute with one another and with the action of $\mathrm{GL}(r,\mathcal{B})$.

The identification rules defining $\mathcal{M}$ are described by a collection of N matrices $g_i \in \mathrm{GL}(r, \mathbf{C}\otimes\mathcal{B})$, on which additional conditions will later be imposed. Having chosen this collection, we set

$$\mathcal{M}=\{m(\lambda)\in\mathcal{B}[\lambda]^r=\mathcal{N} \mid \forall i=1,\ldots,N,\ m(\alpha_i)=g_i m(\beta_i)\}. \tag{2}$$

If $k=\mathbf{R}$ and $\beta_i=\bar{\alpha}_i$, this implies the condition $\bar{g}_i=g_i^{-1}$.

The remainder of this section is devoted to describing those conditions on (g_i) and (α_i,β_i), which enable us to introduce on $\mathcal{M}$ the bimodule structures with the axioms of 2.2.

5.2. The Actions of $\mathcal{B}$ and $\mathcal{O}$. We have described above the actions of $k[\lambda]$ and $\mathcal{B}$ on $\mathcal{N}$. It is obvious from definitions (1) and (2) that $\mathcal{B}\mathcal{M}\subset\mathcal{M}$, $\mathcal{M}\mathcal{O}\subset\mathcal{M}$ and the actions of $\mathcal{B}$ and $\mathcal{O}$ on $\mathcal{M}$ commute. That these actions are effective will become clear below following construction of the element $1\in\mathcal{M}$.

5.3. The Filtration. We first describe an auxiliary filtration on $\mathcal{N}$ by setting $\mathcal{N}_{-1}=\{0\}$, and for any $a\geqslant 0$, $0\leqslant k\leqslant r$;

$$\mathcal{N}_{ar+k}=\left\{\begin{pmatrix} b_1\\ \vdots\\ b_r\end{pmatrix}\in\mathcal{B}[\lambda]^r,\ b_1,\ldots,b_k \text{ are polynomials in } \lambda \text{ of degree } \leqslant a \ ; \right. \\ \left. b_{k+1},\ldots,b_r \text{ are polynomials in } \lambda \text{ of degree } \leqslant a-1\right\}. \tag{3}$$

(A polynomial of degree $\leqslant -1$ is zero.) It is obvious that $\mathcal{N}\subset\mathcal{N}_{l+1}$ and $\mathcal{N}=\bigcup_{l=0}^{\infty}\mathcal{N}_l$. Further, $\mathcal{N}_{l+1}=\mathcal{N}_l\oplus\mathcal{B}e_{l+1}$, where e_{l+1} has $\lambda^a E_r$ at the site $k+1$ for $l=ar+k$, $a\geqslant 0$, $0\leqslant k\leqslant r-1$ and zeros elsewhere. Therefore, $\mathcal{N}_{l+1}/\mathcal{N}_l$ is free of rank 1.

The filtration we need on $\mathcal{M}$ will be induced by this filtration on $\mathcal{N}$ and the following shift. We temporarily set $\mathcal{M}_{(k)}=\mathcal{N}_k\cap\mathcal{M}$. It is obvious that $\mathcal{M}=\bigcup_{k=0}^{\infty}\mathcal{M}_{(k)}$ and $\mathcal{B}\mathcal{M}_{(k)}\subset\mathcal{M}_{(k)}$. We impose the following conditions on the collections (g_i), (α_i,β_i).

5.4. The Nondegeneracy Condition. The block matrix in $M_{rN}(\mathcal{B})$: $G=(\alpha_i^j E_r-\beta_i^j g_i)$, $1\leqslant i\leqslant N$, $0\leqslant j\leqslant N-1$ is nondegenerate, i.e., belongs to $\mathrm{GL}(rN,\mathcal{B})$.

5.5. Lemma. If the condition of nondegeneracy 5.4 is satisfied, then $\mathcal{M}_{(i)}=\{0\}$ for $i\leqslant rN$, and the natural mapping $\mathcal{M}_{l+1}/\mathcal{M}_l\to\mathcal{N}_{l+1}/\mathcal{N}_l$ is an isomorphism of $\mathcal{B}$-modules for $l\geqslant rN+1$.

The filtration $\mathcal{M}_k=\mathcal{M}_{(k+rN+1)}$ for $k\geqslant -1$ therefore satisfies conditions 2.2e, f, and g where any $\mathcal{B}$-generator of the module $\mathcal{M}_{(rN+1)}$ may be taken as 1 .

Proof. According to (3), $\mathcal{N}_{rN}=\left\{\sum_{j\leqslant N-1} m_j\lambda^j \mid m_j\in\mathcal{B}^r\right\}$. Therefore, by (2)

$$\mathcal{M}_{(rN)}=\left\{\sum_{j\leqslant N-1} m_j\lambda^j \mid \forall i=1,\ldots,N,\ \sum_{j=0}^{N-1}(\alpha_i^j E_r-\beta_i^j g_i)\, m_j=0\right\}.$$

The condition in braces may be written

$$G\begin{pmatrix} m_0\\ \vdots \\ m_{N-1}\end{pmatrix}=0.$$

Since G is nondegenerate by hypothesis, it follows that $\mathcal{M}_{(rN)}=\{0\}$ (and hence $\mathcal{M}_{(i)}=0$ for $i\leqslant rN$) .

In order to verify that $\mathcal{M}_{l+1}/\mathcal{M}_l\to\mathcal{N}_{l+1}/\mathcal{N}_l$ is a $\mathcal{B}$-isomorphism for $l\geqslant rN$, it suffices to show that in $\mathcal{M}_{l+1}$ there is an element $\bar{e}_{l+1}$ of the form $\bar{e}_{l+1}\equiv e_l \bmod \mathcal{N}_l$, where e_{l+1} is described in 5.3.

To find $\bar{e}_{l+1}=\sum m_j\lambda^j$ we must write down the system of linear equations for the m_j, corresponding to the conditions (2). After displaying them over the commutative ring $\mathcal{B}_0$ we find in the system a nondegenerate minor of maximum possible rank rlN corresponding to the matrix G. The system is therefore solvable.

5.6. The Explicit Form of 1. From the proof of the preceding lemma it is clear that we may set

$$1=\sum_{j=0}^{N-1} m_j\lambda^j+\begin{pmatrix}\lambda^N\\ 0\\ \vdots\\ 0\end{pmatrix},\quad m_j\in\mathcal{B}^r, \tag{4}$$

where the m_j form a solution of the system $(i=1,\ldots,N)$:

$$\sum_{j=0}^{N-1}(\alpha_i^j E_r-\beta_i^j g_i)\, m_j+(\alpha_i^N E_r-\beta_i^N g_i)\begin{pmatrix}1\\ 0\\ \vdots\\ 0\end{pmatrix}=0. \tag{5}$$

5.7. The Connection $\nabla=\nabla_x$. The following results with obvious modifications will also be applied to the construction of the connections $\nabla_1=\nabla_t$, $\nabla_2=\nabla_y$.

We first continue $\partial_x:\mathcal{B}\to\mathcal{B}$ coordinate-wise to $\mathcal{B}^r$ and then to $\mathcal{B}^r[\lambda]$ so that $\partial_x\lambda=0$. Inasmuch as $\partial_x(b_1 b_2^t)=\partial_x(b_1)\,b_2^t+b_1(\partial_x b_2)^t$ for $b_1, b_2\in\mathcal{B}$, this action is a ∂_x-connection on $\mathcal{N}$. Any other connection extending ∂_x and trivial on λ, has the form ∂_x+d_x, where $d_x\in M_r(\mathcal{B}[\lambda])=\operatorname{End}\mathcal{B}[\lambda]^r$. (Triviality on λ gives the condition $\nabla_x(m\varphi)=(\nabla_x m)\varphi$ for $m\in\mathcal{N}$, $\varphi\in k[\lambda]$) . We use the freedom in the choice of d_x to ensure the condition $\nabla_x\mathcal{M}\subset\mathcal{M}$.

5.8. LEMMA. Let $\nabla_x=\partial_x+d_x(\lambda)$. Then $\nabla_x\mathcal{M}\subset\mathcal{M}$, if

$$\forall i=1,\ldots,N,\quad \partial_x g_i=g_i d_x(\beta_i)-d_x(\alpha_i)\, g_i. \tag{6}$$

Proof. According to (2), $m(\lambda)\in\mathcal{M}$, if and only if $m(\alpha_i)=g_i m(\beta_i)$ for all $i=1,\ldots,N$, whence

$$(\partial_x m)(\alpha_i)=\partial_x g_i m(\beta_i)+g_i\partial_x m(\beta_i).$$

Therefore,

$$[(\partial_x+d_x)m](\alpha_i)=\partial_x g_i m(\beta_i)+g_i\partial_x m(\beta_i)+d_x(\alpha_i)g_i m(\beta_i).$$

On the other hand,

$$g_i[(\partial_x+d_x)m](\beta_i)=g_i\partial_x m(\beta_i)+g_i d_x(\beta_i)m(\beta_i).$$

From (6) it is obvious that the right sides of these expressions coincide.

5.9. Solutions of Eqs. (6). We assume that $d_x(\alpha_i,x)$ and $d_x(\beta_i,x)$ commute respectively with $\exp\left(\int_0^x d_x(\alpha_i,\xi)\,d\xi\right)$ and $\exp\left(\int_0^x d_x(\beta_i,\xi)\,d\xi\right)$ (this is so if $[d_x(\alpha_i,x),d_x(\alpha_i,y)]=[d_x(\beta_i,x),d_x(\beta_i,y)]=0$). Then an explicit solution of (6) can be written in the form

$$g_i(x)=\exp\left(-\int_0^x d_x(\alpha_i,\xi)\,d\xi\right)g_i(0)\exp\left(\int_0^x d_x(\beta_i,\xi)\,d\xi\right). \tag{7}$$

In the examples we shall take $d_x(\lambda)\in M_r(M_l(k[\lambda]))$, and then

$$g_i(x)=\exp(-xd_x(\alpha_i))\,g_i(0)\exp(xd_x(\beta_i)). \tag{8}$$

5.10. The Connection of ∇_x with the Filtration in $\mathcal{M}$. LEMMA. Let $\nabla_x\mathcal{M}\subset\mathcal{M}$, $\nabla_x=\partial_x+d_x$. Then $\nabla\mathcal{M}_l\subset\mathcal{M}_{l+1}$ for all $l\geqslant-1$ and ∇_x induces an isomorphism $\mathcal{M}_l/\mathcal{M}_{l-1}\overset{\sim}{\rightarrow}\mathcal{M}_{l+1}/\mathcal{M}_l$, if the following conditions are satisfied:

$$\text{for } r=1:\ d_x=d_{0x}+d_{1x}\lambda,\ d_{1x}\in\mathcal{B}^*=\mathrm{GL}(1,\mathcal{B});$$
$$\text{for } r\geqslant 2:\ d_x=d_{0x}+d_{1x}\lambda,$$

where $d_{0x}\in M_r(\mathcal{F})$ has zeros on the main diagonal and above and invertible elements (in $\mathcal{B}^*$) along the diagonals below the main diagonal, while $d_{1x}\in M_r(\mathcal{B})$ has an invertible element in the right upper corner and zeros elsewhere.

Proof. Since $\mathcal{M}_l=\mathcal{M}\cap\mathcal{N}_{l+rN+1}$ and $\mathcal{M}_{l+1}/\mathcal{M}_l\simeq\mathcal{N}_{l+rN+2}/\mathcal{N}_{l+rN+1}$ for $l\geqslant rN$, it suffices to verify that under the hypotheses of the lemma $\nabla_x\mathcal{N}_l\subset\mathcal{N}_{l+1}$ and ∇_x induces an isomorphism $\mathcal{N}_{l+1}/\mathcal{N}_l\overset{\sim}{\rightarrow}\mathcal{N}_{l+2}/\mathcal{N}_{l+1}$. But the action of ∇_x on the last factor coincides with the action d_x, since $\partial_x\mathcal{N}_l\subset\mathcal{N}_l$, and the required result is checked by direct computation.

We have now completed the construction of the class of bimodules $\mathcal{M}$. The result of the section is devoted to some further remarks.

5.11. The Rank of $\mathcal{M}$. It is equal to r in the sense described in the introduction and also in the sense of Lemma 2.3 and its corollary. Indeed, if $\varphi\in\mathcal{O}$ is a polynomial of degree n, then $\mathcal{M}_l\varphi\subset\mathcal{M}_{l+rn}$, and multiplication by φ induces the isomorphism $\mathcal{M}_l/\mathcal{M}_{l-1}\overset{\sim}{\rightarrow}$

$\mathcal{M}_{l+rn}/\mathcal{M}_{l+rn-1}$. It suffices to verify this on the corresponding filtration submodules of $\mathcal{A}$ where all is obvious from the definitions.

5.12. The Connections ∇_t, ∇_y . We set $\nabla_t=\partial_t+d_t,\ \nabla_y=\partial_y+d_y$. For solving nonstationary equations these connections must satisfy conditions analogous to (6) but not Lemma 5.10. In place of it we need the commutation conditions $[\nabla_x, \nabla_t]=[\nabla_x, \nabla_y]=[\nabla_t, \nabla_y]=0$ (cf. Theorem 2.5, 2.6). They can be written in the form

$$\partial_x d_t-\partial_t d_x+[d_x, d_t]=0 \tag{9}$$

and similarly for the remaining pairs. In the case $d_x, d_t \in M_r(M_l(k[\lambda]))$ they become simply $[d_x, d_t]=0$ etc. If these conditions are satisfied, then the matrices $g_i(x, t)$ for the Lax equations have the form

$$g_i(x, t)=\exp(-xd_x(\alpha_i)-td_t(\alpha_i))\times g_i(0,0)\exp(xd_x(\beta_i)+td_t(\beta_i)). \tag{10}$$

Conditions (9) with constant d_x, d_t, d_y are most simply satisfied by taking d_t, d_y to be polynomials in d_x .

5.13. The Order of the Operators. According to 2.3 and 5.11, the order of L_φ, is equal to $r\deg\varphi$. It is not hard to see that the order of P, found from the condition $\nabla_t 1=P1$, does not exceed $r\deg d_t$ (degree in λ) and similarly for ∇_y. These considerations determine the choice of $\mathcal{O}$, ∇_t in the next section.

The case where $\mathcal{O}$ contains polynomials of degree 2 and $r=1$, leads to multisoliton solutions of the Korteweg–de Vries equation and its higher analogues and has been treated in the literature in various ways (by the method of inverse scattering, the method of Zakharov-Shabat, Hirota's method). The next most difficult case is the case $r=2$.

6. Example: Solitons of Rank 2

6.1. Parameters of an N-Soliton Solution. We shall construct an N-soliton solution of a nonstationary Lax equation of the form $L_t=[P, L]$, where the order of L is equal to four and the order of P is two. The parameters of the N-soliton solution are constants which go into the construction of the appropriate bimodule $\mathcal{M}$.

We take $k=\mathbb{R}$, $\mathcal{B}=\mathcal{B}_0$ be the germs of meromorphic functions of x, t over $\mathbf{R}$; the number of solitons N is the number of pairs (α_k, β_k) of dual points.

In order that $\mathcal{O}$ contain a polynomial of degree 2, it is necessary and sufficient that $\alpha_k+\beta_k$ not depend on k . Replacing λ by $\lambda+\text{const}$, we may assume that $\alpha_k+\beta_k=0$ for all k; then $\lambda^2\in\mathcal{O}$. Real α_k lead to nonlocalized solutions. We therefore take α_k to be pure imaginary. For convenience of subsequent computations we set $\alpha_k=2ia_k^2b$, $\beta_k=-2ia_k^2b$, $a_k, b\in\mathbf{R}$, where b is chosen as follows.

According to Lemma 5.10 and (8), the connection ∇_x has the form ∂_x+d_x, where $d_x=\begin{pmatrix}0 & a\lambda\\ b & 0\end{pmatrix}$, $a, b\in\mathbf{R}^*$. Replacing λ by $a^{-1}\lambda$ and modifying the value of a_k, correspondingly, we may assume that $a=1$; b remains free.

We subject the choice of ∇_t to the condition that the operator P, for which $\nabla_t 1 = P1$, have order 2. According to 5.12, it suffices that for $\nabla_t = \partial_t + d_t$ the matrix d_t depend on λ linearly; moreover, for constant coefficients of d_x and d_t the condition that ∇_x and ∇_t commute leads to the condition $d_t = \omega\lambda E_2 + c d_x$, where $\omega \in \mathbf{R}$, $c \in \mathbf{R}^*$ (the *a priori* admissible term $\omega_0 E_2$, $\omega_0 \in \mathbf{R}$, does not lead to alteration of the solution).

To construct the identification matrix $g_k(x, t)$, which finally defines the bimodule $\mathcal{M}$, it is necessary to further choose initial conditions $g_k(0,0) \in GL(2, \mathbf{C})$. According to 5.1, we must have $\overline{g_k(0,0)} = g_k(0,0)^{-1}$. We represent $g_k(0,0)$ in the form $g_k(0,0) = \gamma_k^{-1}\bar{\gamma}_k$, $\gamma_k \in GL(2, \mathbf{C})$.

We further set $\xi = b(x+ct)$, $\tau = \omega b t$. Then formula (10) becomes

$$g_k(x, t) = G_k(x, t)^{-1}\overline{G_k(x, t)}, \tag{11}$$

where

$$G_k(x, t) = e^{2ia_k^2\tau}\gamma_k \exp\begin{pmatrix} 0 & 2ia_k^2 \\ 1 & 0 \end{pmatrix}\xi. \tag{12}$$

Thus, $-c$ has the interpretation of a certain "group velocity" of the N-soliton solution; ω is the common frequency of oscillation, and b is a scale factor; a_k and γ_k determine the shape of the k-th soliton.

6.2. The Element 1. The coefficients m_j of the element 1, defined according to formula (4) are found from the system of equations (5) by transforming with the use of relations (11) and (12) $(j = 1, \dots, N)$:

$$\sum_{k=0}^{N-1}[(2ia_j^2 b)^k G_j - (-2ia_j^2 b)^k \bar{G}_j)]\, m_k = [(2ia_j^2 b)^N G_j - (-2ia_j^2 b)^N \bar{G}_j]\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{13}$$

6.3. The Operators L and P. We set $L = \partial_x^4 + u\partial_x^3 + v\partial_x^2 + w\partial_x + z$. To find u, v, w, z we solve in $\mathcal{M}$ the equation

$$(\nabla_x^4 + u\nabla_x^3 + v\nabla_x^2 + w\nabla_x + z)\,1 = 1\lambda^2 b^2. \tag{14}$$

The coefficients of L are uniquely expressed in terms of the components of the coefficients $m_{N-1} = \begin{pmatrix}\mu_1 \\ \mu_2\end{pmatrix}$ and $m_{N-2} = \begin{pmatrix}\mu_3 \\ \mu_4\end{pmatrix}$ by the formulas

$$u = 0;\quad v = -4\partial_x\mu_2;\quad w = -6\partial_x^2\mu_2 - 4b\partial_x\mu_1 + 4\mu_2\partial_x\mu_2;$$

$$z = -4\partial_x^3\mu_2 - 6b\partial_x^2\mu_2 - 4b\partial_x\mu_1 + 8(\partial_x\mu_2)^2 + 4b\mu_1\partial_x\mu_2 + 6\mu_2\partial_x^2\mu_2 + 4b\mu_2\partial_x\mu_1 - 4\mu_2^2\partial_x\mu_1. \tag{15}$$

To obtain them it suffices to equate the coefficients of λ^{N+2}, λ^{N+1}, λ^N on the left and right sides of (14).

Similarly, for the equation $\nabla_t 1 = P1$, we find

$$P = \omega b^{-1}\partial_x^2 + c\partial_x - 2\omega b^{-1}\partial_x\mu_2. \tag{16}$$

The condition $L_t=[P, L]$ has the form of a system of equations for v, w, z:

$$\begin{aligned}-b\omega^{-1}v_t&=2v_{xx}-2w_x-bc\omega^{-1}v_x;\\ -b\omega^{-1}w_t&=-w_{xx}+2v_{xxx}+vv_x-bc\omega^{-1}w_x-2z_x;\\ -b\omega^{-1}z_t&=\frac{1}{2}v_{xxxx}+\frac{1}{2}vv_{xx}-z_{xx}+\frac{1}{2}v_xw-bc\omega^{-1}z_x.\end{aligned} \tag{17}$$

Thus, the constants $b\omega^{-1}$ and the "speed" of the solution c determine the form of the equations; the remaining constants of N-soliton solutions can vary.

6.4. A Single Soliton. We solve Eq. (13) for the case $N=1, a_1=a, \gamma_1=iE_2$. The system acquires the form

$$(iG_0+i\bar{G}_0)\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}=(2a^2G_0-2a^2\bar{G}_0)\begin{pmatrix}1\\ 0\end{pmatrix},$$

whence

$$\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}=2a^2(\operatorname{Re}G_0)^{-1}\operatorname{Im}G_0\begin{pmatrix}1\\ 0\end{pmatrix}, \tag{18}$$

where

$$G_0=e^{2ia^2\tau}\exp\begin{pmatrix}0 & 2ia^2\\ 1 & 0\end{pmatrix}\xi. \tag{19}$$

Since $\begin{pmatrix}0 & 2ia^2\\ 1 & 0\end{pmatrix}^2=\begin{pmatrix}2ia^2 & 0\\ 0 & 2ia^2\end{pmatrix}=\begin{pmatrix}(1+i)a & 0\\ 0 & (1+i)a\end{pmatrix}^2$, by separating even and odd powers in the series for exp we obtain without difficulty

$$G_0=e^{2ia^2\tau}\begin{pmatrix}\operatorname{ch}(1+i)a\xi, & (1+i)a\operatorname{sh}(1+i)a\xi\\ (1+i)^{-1}a^{-1}\operatorname{sh}(1+i)a\xi, & \operatorname{ch}(1+i)a\xi\end{pmatrix}. \tag{20}$$

To simplify the notation we set $\operatorname{ch}=\operatorname{ch}a\xi$, $\operatorname{sh}=\operatorname{sh}a\xi$, $c=\cos a\xi$, $s=\sin a\xi$. $\cos=\cos 2a^2\tau$, $\sin=\sin 2a^2\tau$. Computing (20) explicitly, we obtain

$$\begin{aligned}\operatorname{Re}G_0&=\begin{pmatrix}\operatorname{ch}c\cos-\operatorname{sh}s\sin, & a(\operatorname{sh}c\cos-\operatorname{ch}s\cos-\operatorname{sh}c\sin-\operatorname{ch}s\sin)\\ \frac{1}{2a}(\operatorname{sh}c\cos+\operatorname{ch}s\cos-\operatorname{ch}s\sin+\operatorname{sh}c\sin), & \operatorname{ch}c\cos-\operatorname{sh}s\sin\end{pmatrix},\\ \operatorname{Im}G_0&=\begin{pmatrix}\operatorname{ch}c\sin+\operatorname{sh}s\cos, & a(\operatorname{sh}c\sin+\operatorname{sh}c\cos+\operatorname{sh}s\cos-\operatorname{ch}s\sin)\\ \frac{1}{2a}(\operatorname{sh}c\sin+\operatorname{ch}s\sin+\operatorname{ch}s\cos-\operatorname{sh}c\cos), & \operatorname{ch}c\sin+\operatorname{sh}s\cos\end{pmatrix}.\end{aligned} \tag{21}$$

After rather lengthy but straightforward transformations, we find from this

$$\det\operatorname{Re}G_0=\frac{1}{4}(\operatorname{ch}2a\xi+\cos 2a\xi+2\cos 4a^2\tau) \tag{22}$$

and then, using (18), (21), and (22),

$$\begin{aligned}\mu_1&=2a^2\frac{\operatorname{ch}2a\xi-\cos 2a\xi+2\sin 4a^2\tau}{\operatorname{ch}2a\xi+\cos 2a\xi+2\cos 4a^2\tau},\\ \mu_2&=-2a\frac{\operatorname{sh}2a\xi-\sin 2a\xi}{\operatorname{ch}2a\xi+\cos 2a\xi+2\cos 4a^2\tau}.\end{aligned}$$

The qualitative behavior of the corresponding (v, w, z) was described in the introduction where, for simplicity, we set $b=1$.

7. Solutions of the Reduced Benney Equations and Their Analogues

7.1. The reduced Benney equations are Eqs. (22) of Chap. II. In attempting to find solutions of them for which $h=\varphi(u)$, we find that $h=(u+c)^2/4$, $c\in\mathbf{R}$ is a constant from the compatibility condition. We have, moreover, the following result.

7.2. THEOREM. Let $c\in\mathbf{R}$, $h=(u+c)^2/4$, and let u be a solution of the equation $u_t=-(3u^2/4+cu/2)_x$. Then the pair (u, h) is a solution of system (22) of Chap. II.

Proof. Under the hypotheses of the theorem we have

$$h_t=(u+c)u_t/2=-(u+c)(3u+c)u_x/4=-(u(u+c)^2/4)_x=-(uh)_x,$$
$$u_t=-(3u^2/4+uc/2)_x=-(u^2/2+(u+c)^2/4)_x=-(u^2/2+h)_x.$$

We now consider the equation $u_t+(3u^2/4+cu/2)_x$. The following method is classical.

7.3. Proposition. Let ψ be differentiable, let φ be twice differentiable, and let $u=u(x, t)$ be a smooth solution of the functional equation $u=\psi(x-\varphi'(u)t)$ in the range of the variables (x, t). Then in this range $[u_t+\varphi(u)_x][1+\varphi''(u)t\cdot\psi'(x-\varphi'(u)t)]=0.$

COROLLARY. Under the hypotheses of the proposition $u(x, t)$ is a solution of the equation $u_t+\varphi(u)_x=0$ at points not in the set where $1+\varphi''(u)t\cdot\psi'(x-\varphi'(u)t)=0$.

Proof. Differentiating the relation $u=\psi(x-\varphi'(u)t)$ with respect to x and t, we obtain $u_x=\psi'(x-\varphi'(u)t)[1-\varphi''(u)u_xt]$, $u_t=\psi'(x-\varphi'(u)t)[-\varphi'(u)-\varphi''(u)tu_t]$. We multiply the first equation by $\varphi'(u)$, replace $\varphi'(u)u_x$ by $\varphi(u)_x$, twice, add the relations obtained, and take all terms to the left. We obtain assertion 7.3.

If the Cauchy problem is posed for the equation $u_t+\varphi(u)_x=0$, then ψ has an obvious physical interpretation: $\psi(x)=u(x, 0)$. We shall investigate for what initial conditions ψ the reduced system has a unique solution for all $x\in\mathbf{R}$ and $t\geqslant 0$. We must put $\varphi(u)=3u^2/4+cu/2$ (obviously, any constant can be taken in place of 3/4).

7.4. Proposition. The equation $u=\psi(x-(3u/2+ct/2))$ has a unique solution for all $x\in\mathbf{R}$ and $t\geqslant 0$, if and only if the smooth function ψ is nondecreasing everywhere.

Proof. For given x and $t\geqslant 0$ the graph of $\psi(x-3u/2-ct/2)$ as a function of u is obtained from the graph of ψ by reflection in the vertical axis (call this the graph of $\bar{\psi}$), compressing $\bar{\psi}$ horizontally 3/2 times, and translating to the left by $x-ct/2$. In order that after any compression with positive coefficient and any translation the graph of $\bar{\psi}$ have a unique intersection with the diagonal, it is necessary and sufficient that $\bar{\psi}$ be nonincreasing. Indeed, a small neighborhood of any local maximum or minimum of $\bar{\psi}$ under compression and translation can be made to intersect the diagonal twice. Hence $\bar{\psi}$ must be monotone. If it is nondecreasing everywhere, then it cannot be a nonhorizontal line, since a suitable compression and translation will take this line into the diagonal. If it is not a horizontal line, then there is a point u_0 with $\bar{\psi}'(u_0)>0$, $\bar{\psi}''(u_0)\neq 0$; $\bar{\psi}$ then lies locally on both sides of the tangent at the point u_0 . A small translation of this tangent will intersect $\bar{\psi}$ twice. A suitable compression and translation takes this secant into the diagonal.

Remarks. a) The initial condition $\psi(x)=x$ on some segment leads to the classical solution $u=\frac{2x-c^2t}{2+ct}$, $h=\frac{(x+c)^2}{(2+ct)^2}$, which is called the "ruptured dike" solution.

b) If the velocity profile ψ at the initial time has an inflection point with a horizontal tangent, then by an arbitrarily small deformation it is possible to obtain from it a profile with a maximum and minimum which leads to nonuniqueness of the solution in finite time.

7.5. We now present a method for finding invariant manifolds of the form $h=\varphi(u)$ for the Hamiltonian system described in Sec. 10 of Chap. II. Our calculations will show that the manifold $h=\frac{(u+c)^2}{4}$ is invariant also for the higher reduced Benney equations. It would be of interest to find an analogue of it for the unreduced equations.

In the notation of Sec. 10 of Chap. II we seek solutions of the equations $u_t=\eta_{1hx}$, $h_t=\eta_{1ux}$, subject to a relation of the form $S(u, h)=c\in\mathbb{R}$, where S is a suitable differentiable function. We must ensure the consistency of the system

$$0=S_t=S_u u_t+S_h h_t=S_u(\eta_{1u}^u\eta_{1h}^h u_x+\eta_1^u\eta_{1hh}^h h_x)+S_h(\eta_{1uu}^u\eta_1^h u_x+\eta_{1u}^u\eta_{1h}^h h_x), \tag{23}$$

$$0=S_x=S_u u_x+S_h h_x. \tag{24}$$

We multiply (24) by $-\eta_{1u}^u\eta_{1h}^h$ and add to (23); into the result we substitute the relations $\eta_{1uu}^u=a(\lambda_1)V_1\eta_1^u$ and $\eta_{1hh}^h=a(\lambda_1)V_2\eta_1^h$ and divide by $a(\lambda_1)\eta_1^h\eta_1^u$. We obtain the equation

$$S_h V_1 u_x+S_u V_2 h_x=0. \tag{25}$$

In order that the system of equations (24) and (25) have nontrivial solutions (u_x, h_x) it is necessary that $S_h^2 V_1-S_u^2 V_2=0$, i.e., $S_u/\sqrt{V_1(u)}\pm S_h/\sqrt{V_2(h)}=0$. This is satisfied if S is a function of $\int^u\sqrt{V_1}du\pm\int^h\sqrt{V_2}dh$. Since we are interested in the relation $S=\text{const}$, it is possible to set simply

$$S\equiv\int^u\sqrt{V_1}du\pm\int^h\sqrt{V_2}dh=\text{const}. \tag{26}$$

For the higher Benney equations the conditions (26) reduce to $u\pm 2\sqrt{h}=\text{const}$, i.e., $h=(u+c)^2/4$, as above.

We express h in terms of u from Eq. (26): $h=h(u)$ (when this is impossible we express u in terms of h and argue analogously).

Under condition (26) the equation $u_t=\eta_{1hx}$ becomes an equation $u_t=(\eta_{1h}|_{h=h(u)})_x$ of the type considered in 7.3. The equation $h_t=\eta_{1ux}$ is automatically satisfied. Indeed, from Eq. (23) and then (25) we find (always with $h=h(u)$):

$$h_t = -\frac{S_u}{S_h} u_t = -\frac{S_u}{S_h}(\eta_{1h})_x = -\frac{S_u}{S_h}(\eta_{1hh}h_x + \eta_{1hu}u_x) =$$

$$= -\frac{S_u}{S_h}(aV_2\eta_{1}h_x + \eta_{1hu}u_x) = aV_1\eta_1 u_x + \eta_{1hu}h_x = \eta_{1uu}u_x + \eta_{1hu}h_x = \eta_{1ux}.$$

From linearity considerations it is clear that the same formalism remains in force if the formal series η_1 is replaced by any linear combination of its coefficients (as the Hamiltonian). This shows that the manifold $\int^u \sqrt{V_1}\,du \pm \int_0^h \sqrt{V_2}\,dh = \text{const}$ is invariant with respect to equations with any Hamiltonian from the corresponding space.

CHAPTER IV

INDIVIDUAL RESULTS

1. The Hirota Formalism

1.1. In this section we describe the semiheuristic method of Hirota for solving nonlinear equations. The method consists of two steps: reduction of the equation to so-called bilinear form by finding a suitable substitution and then solving the bilinear equation by means of some version of perturbation theory or a lucky guess. The Hirota formalism has a great deal in common with the Zakharov–Habat formalism described in Sec. 5 of Chap. II, but their exact relationship has not been clarified. Our presentation is based on the papers [38, 39]; see also the literature cited in these works.

1.2. The Hirota Operators. Let $f(x, t)$, $g(x, t)$ be two functions of two variables which are differentiable an appropriate number of times. For a pair of integers m, n we define the expression

$$D_t^n D_x^m f\cdot g = \left(\frac{\partial}{\partial t} - \frac{\partial}{\partial t'}\right)^n \left(\frac{\partial}{\partial x} - \frac{\partial}{\partial x'}\right)^m f(t, x)\, g(t', x')\Big|_{t=t', x=x'}.$$

A system of equations in bilinear form is obtained by equating to zero some system of linear combinations of such expressions for the unknown functions f and g (the coefficients of the linear combinations in all examples are constants).

1.3. The Hirota Substitutions. We shall present a sequence of substitutions which reduce well-known nonlinear equations to bilinear form.

a) The Korteweg–de Vries equation $u_t + 6uu_x + u_{xxx} = 0$. Substitution: $u = 2(\log f)_{xx}$. Bilinear form:

$$D_x(D_t + D_x^3) f\cdot f = 0.$$

b) The equation $u_t + 45u^2u_x + 15(u_xu_{xx} + u_{xxx}u) + u_{xxxxx} = 0$. Substitution: $u = 2(\log f)_{xx}$. Bilinear form:

$$D_x(D_t + D_x^5) f\cdot f = 0.$$

This equation is similar to the second of the higher Korteweg–de Vries equations (corresponding to the operator $\langle (\partial_x^2+u)^{5/2} \rangle$ in the notation of Chap. 2) which has the form $u_t+30u^2u_x+10(2u_xu_{xx}+uu_{xxx})+u_{xxxxx}=0$. However, as Hirota observes, the latter does not reduce to 1.3 b) by linear transformations.

c) The equation for waves on shallow water:

$$u_t-u_{xxt}-3uu_t+3u_x\int_x^\infty u_t d\xi+u_x=0.$$

Substitution: $u=2(\log f)_{xx}$. Bilinear form:

$$D_x(D_t-D_x^2D_t+D_x)f\cdot f=0.$$

d) The Boussinesq equation: $u_{tt}-u_{xx}-3(u^2)_{xx}-u_{xxxx}=0$. Substitution: $u=2(\log f)_{xx}$. Bilinear form

$$(D_t^2-D_x^2-D_x^4)f\cdot f=0.$$

e) The two-dimensional Korteweg-de Vries equation: $u_{tx}\pm u_{yy}+6(uu_x)_x+u_{xxxx}=0$. Substitution: $u=2(\log f)_{xx}$. Bilinear form:

$$(D_tD_x\pm D_y^2+D_y^4)f\cdot f=0.$$

f) The modified Korteweg-de Vries equation: $v_t+6v^2v_x+v_{xxx}=0$. Substitution:

$$v=-i\left(\frac{\log(f+ig)}{f-ig}\right)_x.$$

Bilinear form:

$$(D_t+D_x^3)(f+ig)\cdot(f-ig)=0,$$
$$D_x^2(f+ig)\cdot(f-ig)=0.$$

g) The sine-Gordon equation: $v_{xt}=\sin v$. Substitution: $v=-2i\log(f+ig)/(f-ig)$. Bilinear form:

$$D_xD_tg\cdot f=gf,\ D_xD_t(f\cdot f-g\cdot g)=0.$$

h) The two-dimensional sine-Gordon equation: $v_{xx}+v_{yy}-v_{tt}=\sin v$. Substitution: $v=-2i\log(f+ig)/(f-ig)$.. Bilinear form:

$$(D_x^2+D_y^2-D_t^2)g\cdot f=gf,\ (D_x^2+D_y^2-D_t^2)(f\cdot f-g\cdot g)=0.$$

i) The equation $i\psi_t+3i\alpha|\psi|^2\psi_x+\beta\psi_{xx}+i\gamma\psi_{xxx}+\delta|\psi|^2\psi=0$, α, β, γ, $\delta\in\mathbf{R}$, $\alpha\beta=\gamma\delta$. Substitution: $\psi=G/F$, F a real function. Bilinear form:

$$(iD_t+D_x^2-\lambda)G\cdot F=0,\ (D_x^2-\lambda)F\cdot F=-2|G|^2.$$

j) The equation $i\psi_t+\psi_{xx}-2|\psi|^2\psi=0$. Substitution $\psi=G/F$, F a real function. Bilinear form

$$(iD_t+D_x^2-\lambda)G\cdot F=0,\quad (D_x^2-\lambda)F\cdot F=-2|G|^2.$$

The constant λ is determined by the behavior of ψ as $|x|\to\infty$.

k) The equations of two waves

$$\begin{cases}\varphi_{1t}+v_1\varphi_{1x}=-\varphi_1\varphi_2,\\ \varphi_{2t}+v_2\varphi_{2x}=\varphi_1\varphi_2.\end{cases}$$

Substitution: $\varphi_i=G_i/F$. Bilinear form:

$$(D_t+v_iD_x)G_i\cdot F=0,\quad i=1,2.$$

l) The equations of three waves:

$$\varphi_{it}+v_i\varphi_{ix}=q_i\varphi_j\varphi_k,\quad \{i,j,k\}=\{1,2,3\};\quad q_i=\pm 1.$$

Substitution: $\varphi_i=G_i/F$. Bilinear form:

$$(D_t+v_iD_x)G_i\cdot F=q_iG_jG_k,\quad i=1,2,3.$$

1.4. Sample Solutions. We present two examples of solutions of the bilinear equations and refer the reader to the papers cited of Hirota for other solutions.

a) The interaction of three waves (Eqs. 1.31). We expand F, G_i in powers of an auxiliary small parameter ε:

$$F=1+\sum_{n\geqslant 1}\varepsilon^{2n}f_{2n},\quad G_1=g_{10}+\sum_{n\geqslant 1}\varepsilon^{2n}g_{1n},$$
$$G_j=\sum_{n\geqslant 1}\varepsilon^{2n+1}g_{j,2n+1},\quad j=1,2.$$

We substitute these expansions into the bilinear equations 1.31 and equate coefficients of like powers of ε. Writing out the conditions for thevanishing of the leading coefficients, we find two types of solutions:

The solution with $g_{10}=0$. Here

$$\varphi_1=\frac{a_{22}\exp(\eta_2+\eta_3)}{1+a_{02}\exp 2\eta_2+b_{02}\exp 2\eta_3},$$
$$\varphi_2=\frac{a_{21}\exp\eta_2}{1+a_{02}\exp 2\eta_2+b_{02}\exp 2\eta_3},$$
$$\varphi_3=\frac{a_{31}\exp\eta_3}{1+a_{02}\exp 2\eta_2+b_{02}\exp 2\eta_3},$$

where

$$\eta_2=p_2(x-v_2t),\quad \eta_3=p_3(x-v_3t),$$
$$a_{22}=q_1\frac{a_{21}a_{31}}{2(v_1-v_2)p_2},$$
$$a_{02}=q_1q_3\frac{(a_{21})^2}{4(v_1-v_2)(v_2-v_3)p_2^2},$$
$$b_{02}=-q_1q_2\frac{(a_{31})^2}{4(v_1-v_3)(v_2-v_3)p_3^2}.$$

The letters a_{21} and a_{31} denote arbitrary constants, while p_2 and p_3 obey the conditions

$$(v_1-v_2)p_2=(v_1-v_3)p_3.$$

This solution was obtained by V. E. Zakharov and S. V. Manakov by methods of inverse scattering theory.

The solution with $g_{10} \neq 0$. Here

$$\varphi_1 = g_{10} \frac{1 - \exp 2\eta_1}{1 + \exp 2\eta_1},$$
$$\varphi_2 = \frac{a_1 \exp \eta_1}{1 + \exp 2\eta_1},$$
$$\varphi_3 = \frac{b_1 \exp \eta_1}{1 + \exp 2\eta_1},$$

where

$$\eta_1 = p_1 x - \Omega_1 t,$$
$$(\Omega_1 - v_2 p_1)(\Omega_1 - v_3 p_1) = q_2 q_3 g_{10}^2,$$
$$a_1^2 = -q_1 q_3 \cdot 4 (\Omega_1 - v_1 p_1)(\Omega_1 - v_3 p_1),$$
$$b_1^2 = -q_1 q_2 \cdot 4 (\Omega_1 - v_1 p_1)(\Omega_1 - v_2 p_1).$$

c) The two-dimensional sine-Gordon equation (Eq. 1.3 h). For this equation Hirota found solutions of three-soliton type. It is interesting that there are apparently no known solutions with a different number of solitons. The Hirota solution has the form

$$f + ig = \sum_{\mu = 0,1} \exp \left\{ \sum_{3 > i > j \geq 1} A_{ij} \mu_i \mu_j + \sum_{i=1}^{3} (\eta_i + i\pi/2) \mu_i \right\},$$

where

$$\eta_i = p_i x + q_i y - \Omega_i t,$$

and the constants obey the following restrictions:

$$p_i^2 + q_i^2 - \Omega_i^2 = 1,$$
$$\exp A_{ij} = -\frac{(p_i - p_j)^2 + (q_i - q_j)^2 - (\Omega_i - \Omega_j)^2}{(p_i + p_j)^2 + (q_i + q_j)^2 - (\Omega_i + \Omega_j)^2} = \frac{d_{ij} - 1}{d_{ij} + 1},$$
$$d_{ij} = p_i p_j + q_i q_j - \Omega_{ij};$$
$$\det (d_{ij}) = 0 \Leftrightarrow \det \begin{pmatrix} p_1 & p_2 & p_3 \\ q_1 & q_2 & q_3 \\ \Omega_1 & \Omega_2 & \Omega_3 \end{pmatrix} = 0.$$

1.5. The Hirota Identities. In seeking suitable substitutions relating the equations in the usual form to bilinear equations, Hirota makes use of identities samples of which are given below. The reader will have no difficulty in extending this list by induction.

$$D_x^m a \cdot b = (-1)^m D_x^m b \cdot a;$$
$$D_x^m a \cdot b = D_x^{m-1} (a_x \cdot b - a \cdot b_x);$$
$$D_x^m \exp (p_1 x) \cdot \exp (p_2 x) = (p_1 - p_2)^m \exp (p_1 + p_2) x;$$

$$D_x ab \cdot c = \frac{\partial a}{\partial x} bc + a (D_x b \cdot c);$$
$$D_x^2 ab \cdot c = \frac{\partial^2 a}{\partial x^2} bc + 2 \frac{\partial a}{\partial x} D_x b \cdot c + a (D_x^2 b \cdot c);$$

$$\exp(\delta D_t)[\exp(2\varepsilon D_x)a\cdot b]\cdot cd =$$
$$= \exp(\varepsilon D_x)[\exp(\varepsilon D_x+\delta D_t)a\cdot d]\cdot[\exp(\varepsilon D_x-\delta D_t)c\cdot b];$$
$$\exp(\varepsilon\partial/\partial x)a/b=[\exp(\varepsilon D_x)a\cdot b]/\cos h(\varepsilon D_x)b\cdot b;$$
$$\frac{\partial}{\partial x}\frac{a}{b}=\frac{D_x a\cdot b}{b^2};$$
$$\frac{\partial^2}{\partial x^2}\left(\frac{a}{b}\right)=\frac{D_x^2 a\cdot b}{b^2}-\frac{a}{b}\frac{D_x^2 b\cdot b}{b^2};$$
$$\frac{\partial^3}{\partial x^3}\frac{a}{b}=\frac{D_x^3 a\cdot b}{b^2}-3\frac{D_x a\cdot b}{b^2}\frac{D_x^2 b\cdot b}{b^2};$$
$$2\cos h(\varepsilon\partial/\partial x)\log f=\log\cos h(\varepsilon D_x)f\cdot f;$$
$$2\frac{\partial^2}{\partial x^2}\log f=\frac{D_x^2 f\cdot f}{f^2}=u;$$
$$D_x^4 f\cdot f/f^2=u_{2x}+3u^2;$$
$$D_x^6 f\cdot f/f^2=u_{4x}+15u_{2x}+15u^3.$$

2. Poles of the Solutions

2.1. In [40] Kruskal suggested considering the process of soliton interaction by tracing the poles of the analytic continuation of a multisoliton solution in the complex domain. For the equation $v_t+v_x^2+v_{xxx}$, which is closely related to the Korteweg—de Vries equation, the single-soliton solution has the form $v(x,t)=3\sqrt{c}\,\mathrm{th}\frac{\sqrt{c}(x-ct)}{2}$. The expansion of $\mathrm{th}\,z$ in a sum of principal parts near the poles has the form $\mathrm{th}\,z=\sum_{n\equiv 1(2)}\left(z-\frac{n\pi i}{2}\right)^{-1}$, whence

$$v(x,t)=6\sum_{n\equiv 1(2)}(x-ct-n\pi i/\sqrt{c})^{-1}.$$

Therefore, according to Kruskal, the evolution of the soliton may be considered as a "parade of poles" moving in strict file with speed c.

The two-soliton solution can also be represented in the form of a sum of principal parts:

$$v(x,t)=6\sum_{n\equiv 1(2)}(x-x^{(1)}-\xi_n^{(1)}t)^{-1}+6\sum_{n\equiv 1(2)}(x-x^{(2)}-\xi_n^{(2)}t)^{-1},$$

where $x^{(j)}$, $j=1,2$ are constants, and the $\xi_n^{(j)}$ are functions of time which for large $|t|$ behave like $c^{(j)}t+n\pi i/\sqrt{c^{(j)}}$, where $c^{(j)}$ is the speed of the j-th soliton. The speed of a soliton is proportional to the square of the density of its poles. When a fast soliton overtakes one which is twice as slow, pairs of poles of the first soliton momentarily coalesce with the poles of the second. If the speeds are close part of the poles of the fast soliton spring into the gaps between poles of the slow soliton.

2.2. In the joint work of Airault, McKean, and Moser [25] Kruskal's remark was developed and led to the discovery that the poles of multisoliton solutions evolve in correspondence with known Hamiltonian equations of the type introduced earlier by Calogero and investigated by Moser.

One of the results of their work which is most simply formulated is the following. We suppose that the function

$$v(x, t)=2\sum_{j=1}^{n}(x-x_j(t))^{-2} \tag{1}$$

is a solution of the equation $v_t+3vv_x-\frac{1}{2}v_{xxx}=0$. Then the functions $x_j(t)$ satisfy the system

$$\dot{x}_j=6\sum_{k\neq j}(x_j-x_k)^{-2},\quad j=1,\ldots,n,$$

with the additional condition

$$\sum_{k=1}^{n}(x_j-x_k)^{-3}=0,\quad j=1,\ldots,n. \tag{2}$$

The set (2) is empty if it is restricted to real values. However, for $n=\frac{d(d+1)}{2}$, $d\geqslant 1$, the set of its complex points is isomorphic to a Zariski-open, dense part of d-dimensional complex space (after symmetrization in $(x_1,\ldots,x_n)$).

2.3. In [30], which was completed almost simultaneously with the work of Airault, KcKean, and Moser, D. V. Choodnovsky and G. V. Choodnovsky obtained analogous results for other equations. Especially interesting is their remark that the evolution of the poles of the Burgers–Hopf equation $u_t=2uu_x+u_{xx}$ is free from restrictions of type (2).

2.4. Proposition. Function $u(x, t)=\sum_j(x-x_j(t))^{-1}$ is a solution of the equation $u_t=2uu_x+u_{xx}$, if and only if the $x_j(t)$ satisfy the system

$$\dot{x}_j(t)=-2\sum_{k\neq j}(x_j(t)-x_k(t))^{-1}.$$

(For an infinite set of indices j the following computations are formal; convergence requires a separate investigation.)

Proof. If $u(x, t)=\sum_j(x-x_j(t))^{-1}$, then

$$2uu_x+u_{xx}=-2\sum_{j,k;\,j\neq k}(x-x_j(t))^{-1}(x-x_k(t))^{-2},$$

$$u_t=\sum_k\dot{x}_k(t)(x-x_k(t))^{-2}.$$

Comparing the poles of these expressions (assuming that the $x_j(t)$ do not pairwise coincide), we immediately obtained the required assertion.

2.5. For the modified Korteweg–de Vries equations algebraic restrictions on the motion of the poles again arise: the function $u(x, t)=\sum_j c_j(x-x_j(t))^{-1}$, $c_j=\pm 1$ satisfies the equation $u_t=6u^2u_x-u_{xxx}$, if and only if

$$\dot{x}_j=\sum_{k\neq j}(x_j-x_k)^{-2},$$

$$\sum_{k\neq j}c_k(x_j-x_k)^{-1}=0.$$

2.6. The results briefly considered here are close in spirit to classical expansions in the theory of elliptic functions and also to the approach for obtaining solutions of Lax equations of algebraic type described in Chap. III. We recall that for the bimodule of rank 1 introduced there the invariant Δ also evolves in such a way that the poles of its derivative cancel in a particular manner with the poles of the function Δ itself. It is true that the poles considered there refer to the auxiliary "spectral parameter" — the variable point on the Riemann surface of the curve $\operatorname{Spec}\mathcal{O}$ — rather than to x as a function of t. However, it appears likely that recalculation of the evolution of Δ in terms of the poles of solutions will lead to a generalization of the results of 2.2 and 2.6.

3. Pseudopotentials and Generalized Conservation Laws

3.1. This section gives an introduction to the interesting formalism proposed by Estabrook and Wahlquist [54, 32] and further investigated, in particular, in the work of Corones [31], Corones and Tests [26], and Morris [49].

The central feature of these papers is a certain generalization of the concept of a conservation law for evolution equations of the form $\bar{u}_t=K(\bar{u}, \bar{u}', \ldots)$ where $\bar{u}=(u_1, \ldots, u_n)$. As we have repeatedly mentioned, an ordinary ("algebraic" in the terminology of Chap. I) conservation law for such an equation is a relation of the form $F_t-G_x=0$, which is a formal consequence of the equation $\bar{u}_t=K$. More precisely, let $\mathcal{A}=k[u_i^{(j)}]$ or $\bigcup_{l=0}^{\infty} C^\infty(u_i^{(j)} \| j|\leqslant l)$, $K\in\mathcal{A}$, and let $\partial_t:\mathcal{A}\to\mathcal{A}$ be an evolution differentiation corresponding to $\bar{u}_t=K$; then $F, G\in\mathcal{A}$ with $F_t=G_x$ constitute a conservation law.

If $\bar{u}^s$ is a solution of the system, then in terms of it and a conservation law it is possible to construct a potential v : a solution of the system $v_x=F(\bar{u}^s)\ v_t=G(\bar{u}^s)$.

Suppose F, G are vector-valued functions of $u_i^{(j)}$ of height N and the auxiliary variables $\bar{v}=(v_1, \ldots, v_N)$. We assume that the following system of equations is consistent (the right sides do not contain derivatives of v_i):

$$\begin{gathered}\bar{u}_t=K(\bar{u}, \bar{u}', \ldots);\quad \bar{v}_x=F(\bar{u}, \bar{u}', \ldots; \bar{v});\\ \bar{v}_t=G(\bar{u}, \bar{u}', \ldots; \bar{v}).\end{gathered} \tag{3}$$

Then for any solution $(\bar{u}^s, \bar{v}^s)$ the relation $\frac{\partial}{\partial t}F(\bar{u}^s, \ldots, \bar{v}^s)=\frac{\partial}{\partial x}G(\bar{u}^s, \ldots, \bar{v}^s)$ is satisfied. It is called a generalized conservation law, and $\bar{v}^s$ is called the corresponding pseudopotential.

In all examples considered F and G lie in $\mathcal{A}\underset{k}{\otimes}\mathcal{B}^N$, where $k=\mathbf{R}, \mathbf{C}$ and $\mathcal{B}=C^\infty(v_1, \ldots, v_N)$. An algebraic model of this situation can be formed as follows.

3.2. We consider the ring $\mathcal{A}\underset{k}{\otimes}\mathcal{B}$. On the subring $A\otimes 1$ there act the differentiations $\partial_x: u_i^{(j)}\otimes 1\mapsto u_i^{(j+1)}\otimes 1$ and $\partial_t:\bar{u}\otimes 1\mapsto K\otimes 1$; $[\partial_x, \partial_t]=0$. The right sides of the two last equations of (3) can be considered as the extension of these differentiations to all of $\mathcal{A}\underset{k}{\otimes}\mathcal{B}$, which acts on the generators according to the formulas

$$\bar{\partial}_x: 1\otimes\bar{v}\mapsto F,\quad \bar{\partial}_t: 1\otimes\bar{v}\to G.$$

The formal consistency condition of the system $\bar{v}_x = F$ and $\bar{v}_t = G$ under the condition $\bar{u}_t = K$ means simply that the differentiations $\bar{\partial}_x$ and $\bar{\partial}_t$ commute. We may therefore pose the following problem.

3.3. Problem. Given a k-algebra $\mathcal{A}$ and two commuting k-differentiations $\partial_1, \partial_2 : \mathcal{A} \to \mathcal{A}$. For some k-algebra $\mathcal{B}$ describe the set of extensions $\bar{\partial}_1, \bar{\partial}_2 : \mathcal{A} \otimes \mathcal{B} \to \mathcal{A} \otimes \mathcal{B}$ with the condition $[\bar{\partial}_1, \bar{\partial}_2] = 0$.

3.4. Without practical loss of generality, we may assume that $\bar{\partial}_1, \bar{\partial}_2 \in \mathcal{A} \otimes_k D(\mathcal{B}) + D(\mathcal{A}) \otimes_k \mathcal{B}$. In order to solve Problem 3.3, under certain restrictions on $\mathcal{A}$ we construct a Lie algebra k over $\mathcal{L} = \mathcal{L}(\mathcal{A}, \partial_1, \partial_2)$, such that the set of all extensions $\mathcal{P}_{\mathcal{B}}$ is in bijective correspondence with some subset of Lie algebra morphisms of $\mathrm{Hom}(\mathcal{L}, D(\mathcal{B}))$, where $D(\mathcal{B})$ is the algebra of k differentations of $\mathcal{B}$ into itself. The restrictions on $\mathcal{A}$ are as follows.

3.5. We consider a basis (U_j), $j \in J$ of the algebra $\mathcal{A}$ as a linear space over k, which determines the set of structure constants by

$$\partial_1 U_j = \sum_{l \in J} c_{jl}^{(1)} U_l, \quad \partial_2 U_j = \sum_{l \in J} c_{jl}^{(2)} U_l; \qquad (4)$$
$$U_j U_k = \sum_{l \in J} d_{jkl} U_l,$$

where the right sides are finite linear combinations with coefficients in k.

We suppose, moreover, that these coefficients possess the following property: for each $l \in J$ there exists only a finite number of pairs $(j, k) \in J \times J$ (respectively, elements $j \in J$), such that $d_{jkl} \neq 0$ (respectively, $c_{jl}^{(1)} \neq 0$, $c_{jl}^{(2)} \neq 0$).

3.6. For the algebra $(k[u_i^{(j)}], \partial_x, \partial_t)$ the basis of monomials in $u_i^{(j)}$ obviously satisfies (4); moreover, $d_{jkl} \neq 0$ only if U_j, U_k, divide U_l, and these are of finite number. Further, we define the weight $w(U_j)$ by additivity, setting $w(u_i^{(j)}) = j$. Then ∂_x is homogeneous and increases the weight by one, and the space of polynomials of a given weight is finite-dimensional; therefore, U_l can enter in only a finite number of the derivatives $\partial_1 U_j$. Thus, only the condition on the $c_{jl}^{(2)}$ (for $\partial_2 = \partial_t$) can be nontrivial. For evolution according to Korteweg-de Vries ($u_t = uu' + u'''$) it can be satisfied as for ∂_x, by introducing the new weight $w_1(u^{(i)}) = i + 2$. Then the field $\frac{\partial}{\partial t} = \sum (uu' + u''')^{(i)} \frac{\partial}{\partial u^{(i)}}$ is homogeneous with respect to w_1 and increases this weight by 3, so that U_l can enter the expansion of $\partial_2 U_j$ only for $w_1(U_j) = w_1(U_l) - 3$; there are a finite number of such j . Similar considerations are probably applicable to other Lax equations.

3.7. We now fix the algebra $(\mathcal{A}, \partial_1, \partial_2)$ and basis (U_j), $j \in J$, satisfying the conditions of 3.5. We consider the free Lie algebra $\mathcal{L}$, generated by X_k, Y_j, $j, k \in J$, with relations

$$\sum_j c_{jl}^{(1)} Y_j - \sum_j c_{jl}^{(2)} X_j = \sum_{j,k} d_{jkl} [Y_j, X_k], \ l \in J. \qquad (5)$$

The condition of 3.5 ensures that all these linear combinations are finite.

3.8. Proposition. The set of commuting extensions $(\bar{\partial}_1, \bar{\partial}_2)$ of the pair (∂_1, ∂_2) to $\mathcal{A}\otimes D(\mathcal{B})+D(\mathcal{A})\otimes\mathcal{B}$ corresponds bijectively to the set of homomorphisms of the Lie algebra $\mathcal{L}$ to the Lie algebra $D(\mathcal{B})$, which vanish on all but a finite number of the X_k, Y_j.

Proof. Any extension has the form

$$\bar{\partial}_1=\partial_1\otimes 1+\sum_{j\in J} U_j\otimes\xi_j,\quad \xi_j\in D(\mathcal{B}),$$
$$\bar{\partial}_2=\partial_2\otimes 1+\sum_{j\in J} U_j\otimes\eta_j,\quad \eta_j\in D(\mathcal{B}),$$

where almost all the ξ_j, η_j are equal to zero.

The condition $[\bar{\partial}_1, \bar{\partial}_2]=0$ means that

$$\sum_j \partial_j U_j\otimes\eta_j-\sum_j \partial_2 U_j\otimes\xi_j=\sum_{j,k} U_j U_k\otimes[\eta_j, \xi_k], \tag{6}$$

where the commutators are evaluated in the Lie algebra $D(\mathcal{B})$. Substituting here the right sides of the identities (4), we find

$$\sum_{j,l} c_{jl}^{(1)}\otimes U_l\otimes\eta_j-\sum_{j,l} c_{jl}^{(2)} U_l\otimes\xi_j=\sum_{j,k,l} d_{jkl} U_l\otimes[\eta_j, \xi_k].$$

Equating coefficients of the U_l, we obtain the identities (5) with X_k, Y_j in place of ξ_k, η_j This completes the proof.

3.9. In 3.7 an almost invariant definition of the Lie algebra k over $\mathcal{L}$ with respect to the pair $(\mathcal{A}, \partial_1, \partial_2)$ was given. The dependence on the choice of basis (U_j) with conditions 3.5 is probably not essential in the sense that for a certain class of bases all Lie algebras $\mathcal{L}$ obtained are canonically isomorphic.

We call $\mathcal{L}$ the EW algebra for $(\mathcal{A}, \partial_1, \partial_2)$ (in honor of Estabrook and Wahlquist).

The search for pseudopotentials for the evolution equation represented by $(\mathcal{A}, \partial_1, \partial_2)$, naturally breaks into two steps: a) description of the EW Lie algebra $\mathcal{L}$; b) description of its representations in the Lie algebras of vector fields $D(C^\infty(v_1, \ldots, v_N))$.

Both problems are far from being completely solved even for the Korteweg–de Vries equation. Apparently, even a finite-dimensional EW algebra for it is unknown. In the papers cited at the beginning of the section a sequence of factors of EW algebras is constructed for some interesting triples $(\mathcal{A}, \partial_1, \partial_2)$, which are generated by a finite number of generators and relations, and some particular representations of them by vector fields are given.

Factors of a finite-type EW algebra $\mathcal{L}$ are most easily obtained by setting $X_j, Y_k=0$, if U_j, U_k contain derivatives of u_i of sufficiently high order. In examples "sufficiently high" usually means $\geqslant 3$, and the additional wonder is that then by virtue of the relations (5) all but a finite number of the X_j, Y_k automatically vanish.

If the factor obtained has a nontrivial, finite-dimensional factor Lie algebra $\mathscr{L}_0$, then it is not hard to construct its representation by invariant vector fields on the corresponding Lie group. We observe that representations with an Abelian image, as is evident from (6), lead to the usual conservation laws, so that non-Abelian representations and the corresponding pseudopotentials are of special interest.

Another class of representations which can be constructed is obtained in the Lie algebra of fields over a one-dimensional base. Corones suggested calling the corresponding pseudopotentials simple. This class is nice in that over a one-dimensional base $[\xi, \eta]=0$ implies that $\xi=c\eta$, c a constant, which makes it possible to avoid the many relations (5) and their consequences.

With these rudiments of a systematic theory, we now present several samples of computations from the works cited above.

3.10. The Korteweg–de Vries Equation: $u_t+u_{xxx}+12uu_x=0$. In the work of Wahlquist and Estabrook [54] (see also Corones and Testa [26]) a factor EW-algebra $\mathscr{L}$ is constructed which is given by the following generators and relations:

$$\begin{gathered}[X_1, X_3]=[X_2, X_3]=[X_1, X_4]=[X_2, X_6]=0; \\ [X_1, X_2]+X_7=0;\ [X_1, X_7]-X_5=0;\ [X_2, X_7]-X_6=0; \\ [X_1, X_5]+[X_2, X_4]=0;\ [X_3, X_4]+[X_1, X_6]+X_7=0.\end{gathered} \tag{7}$$

The generalized conservation law $F_t=G_x$ has the form

$$\begin{gathered}F=2X_1+3uX_2+3u^2X_3; \\ G=-2(u_{xx}+6u^2)X_2+3(u_x^2-8u^3-2uu_{xx})X_3+ \\ +8X_4+8uX_5+4u^2X_6+4u_xX_7.\end{gathered}$$

Any representation of (7) in the algebra of vector fields on an N-dimensional manifold gives N concrete generalized conservation laws (X_i, F, G are expanded in terms of components).

In order to obtain a nontrivial factor of the algebra (7), Wahlquist and Estabrook set $X_8=[X_4, X_3]$ and impose the new relation $[X_1, X_5]=\sum_{m=1}^{8} c_mX_m$. They then verify that this leads to the zero algebra in all cases except $c_1=\ldots=c_6=0$, $c_7=-c_8=\lambda$, $\lambda\neq 0$. We denote the corresponding factor by $\mathscr{L}_\lambda$.

According to computations of S. I. Gel'fand, $\mathscr{L}_\lambda$ has as a direct component the Lie algebra $\mathrm{sl}(2)$, engendered by the generators $2X_5+\lambda X_6$, X_6, X_7-X_8, and the representation of Wahlquist and Estabrook coincides (up to a formal diffeomorphism) with a representation of the left-invariant fields on $\mathrm{SL}(2)$. In their notation this is a representation in $D(C^\infty(y_2, y_3, y_8))$, which is described as follows:

$$2X_5+\lambda X_6\mapsto e^{-2y_3}\frac{\partial}{\partial y_2}-y_8\frac{\partial}{\partial y_3}-y_8^2\frac{\partial}{\partial y_8},\quad X_6\mapsto\frac{\partial}{\partial y_8},\quad X_7-X_8\mapsto\frac{1}{2}\frac{\partial}{\partial y_3}+y_8\frac{\partial}{\partial y_8}.$$

The generalized conservation law corresponding to $\frac{\partial}{\partial y_8}$, has the form

$$(2u+y_8^2-\lambda)_t+4\left[(u+\lambda)(2u+y_8^2-\lambda)+2u_{xx}-2u_x y_8\right]_x=0. \tag{8}$$

If u is a solution of the Korteweg–de Vries equation and y_8 is a solution of (8), then $\tilde{u}=-u-y_8^2+\lambda$ is a new solution which is obtained by "adding" to u a single soliton with speed λ. The reasons for the appearance of the "Bäcklund transformation" in such a context remain somewhat puzzling.

3.11. The Hirota Equation $u_t=-3\alpha u\bar{u}u_x-\beta u_{xxx}+i\gamma u_{xx}+i\varepsilon u^2\bar{u}$. Here u is a complex function and $\bar{u}$ is its conjugate, they are considered algebraically independent, and the Hirota equation corresponds to a pair of equations for $u, \bar{u}$: the second is obtained by formal complex conjugation from the first. In [31] Corones constructs the following factor EW algebra for it:

$$-\alpha X_1+\beta[X_1,[X_1,X_2]]=0;$$
$$\alpha X_2+\beta[X_2,[X_1,X_2]]=0;$$
$$[X_3,[X_1,X_2]]=0;$$
$$i\varepsilon_1 X_1+i\gamma[X_1,[X_2,X_1]]-\beta[X_1,[X_2,[X_3,X_1]]]-\frac{1}{2}\beta[X_2,[X_1,[X_3,X_1]]]-\alpha[X_3,X_1]=0;$$
$$-i\varepsilon X_2-i\gamma[X_2,[X_2,X_1]]-\beta[X_2,[X_2,[X_3,X_1]]]-\frac{1}{2}\beta[X_1,[X_2,[X_3,X_2]]]-\alpha[X_3,X_2]=0;$$
$$-\frac{1}{2}\beta[X_1,[X_1,[X_3,X_1]]]=0;$$
$$-\frac{1}{2}\beta[X_2,[X_2,[X_3,X_2]]]=0;$$
$$-i\gamma[X_1,[X_1,X_3]]-\beta[X_3,[X_1,[X_3,X_1]]]-\beta[X_1,[X_3,[X_3,X_4]]]=0;$$
$$-i\gamma[X_2,[X_2,X_3]]-\beta[X_2,[X_3,[X_3,X_2]]]-$$
$$-\frac{1}{2}\beta[X_3,[X_2,[X_3,X_2]]]=0;$$
$$-i\gamma[X_1,[X_2,X_3]]+i\gamma[X_2,[X_3,X_1]]+i\gamma[X_3,[X_2,X_1]]-$$
$$-\beta[X_1,[X_3,[X_3,X_2]]]-\beta[X_2,[X_3,[X_3,X_1]]]-$$
$$-\beta[X_3,[X_2,[X_3,X_1]]]+[X_5,X_4]=0;$$
$$-i\gamma[X_3,[X_1,X_3]]-\beta[X_3,[X_3,[X_3,X_1]]]+[X_1,X_4]=0;$$
$$i\gamma[X_3,[X_2,X_3]]-\beta[X_3,[X_3,[X_3,X_2]]]+[X_2,X_4]=0;$$
$$[X_3,X_5]=0.$$

The corresponding conservation law does not depend on $u^{(j)}, \bar{u}^{(j)}$ for $j\geqslant 3$.

Corones observes that it has non-Abelian representations by one-dimensional fields only in the case $\beta\varepsilon-\alpha\gamma=0$. It is interesting that this is precisely the condition obtained by Hirota for multisoliton solutions of his equation.

3.12. The Equation $3u_{tt}+u_{xxxx}+6(uu_x)_x=0$. This equation was investigated by Morris [49]; the substitution $v=u_t$ reduces it to the usual system of evolution.

The EW factor algebra constructed by Morris is defined by the relations

$$[X_1, X_2]=X_{10};\quad [X_1, X_3]=4X_8;\quad [X_1, X_5]=X_{11};$$
$$[X_1, X_8]=X_5;\quad [X_1, X_{10}]=\frac{3}{4}X_1+\frac{9}{4}\mu X_3;$$
$$[X_1, X_{11}]=\frac{3}{2}X_4-3\mu X_2;\quad [X_2, X_4]=-X_{11};$$
$$[X_2, X_5]=-\frac{9}{4}X_8;\quad [X_2, X_8]=-\frac{3}{8}X_3;$$
$$[X_2, X_{10}]=-\frac{3}{4}X_2;\quad [X_2, X_{11}]=-\frac{9}{4}X_8;$$
$$[X_2, X_8]=-\frac{3}{8}X_3;\quad [X_2, X_{10}]=-\frac{3}{4}X_2;$$
$$[X_2, X_{11}]=-\frac{9}{4}X_5;\quad [X_3, X_4]=\frac{4}{3}X_{10};\quad [X_3, X_{10}]=-\frac{3}{2}X_3;$$
$$[X_3, X_{11}]=X_2;\quad [X_4, X_5]=-3\mu X_8;\quad [X_4, X_8]=-\frac{3}{4}\mu X_3-\frac{1}{4}X_1;$$
$$[X_4, X_{10}]=\frac{3}{2}X_4-3X_2;\quad [X_4, X_{11}]=-3\mu X_5;$$
$$[X_5, X_8]=\frac{1}{4}X_2;\quad [X_5, X_{11}]=-\frac{9}{16}X_1-\frac{9}{16}X_3;$$
$$[X_8, X_{10}]=-\frac{3}{4}X_8;\quad [X_8, X_{11}]=-\frac{1}{4}X_{10};$$
$$[X_{10}, X_{11}]=-\frac{3}{4}X_{11}.$$

Morris constructs its representation in the algebra $D(C^\infty(x_1, x_2, x_3))$ by setting $\left(\partial_i=\frac{\partial}{\partial x_i}\right)$:

$$X_1=-\mu x_1\partial_3-x_2\partial_1-x_3\partial_2;$$
$$X_2=\frac{3}{4}(x_1\partial_2+x_2\partial_3);$$
$$X_3=x_1\partial_3;$$
$$X_4=x_3\partial_1+\mu(x_1\partial_2+x_2\partial_3);$$
$$X_5=\frac{1}{4}(x_1\partial_1+x_3\partial_3-2x_2\partial_2);$$
$$X_8=\frac{1}{4}(x_1\partial_2-x_2\partial_3);$$
$$X_{10}=\frac{3}{4}(x_1\partial_1-x_3\partial_3);$$
$$X_{11}=\frac{3}{4}(x_3\partial_2-x_2\partial_1).$$

From the equation for the generalized conservation law Morris deduces the relations

$$\begin{cases} x_1'''+\frac{3}{2}ux_1'+\left(\frac{3}{4}u'+w\right)x_1=\mu x_1, \\ x_{1,t}+3x_1''+ux_1=0, \end{cases} \tag{9}$$

where w is a trivial potential: $w_x=-\frac{3}{4}u_t$, $w_t=\frac{1}{4}u'''+\frac{3}{2}(uu')'$. Equations (9) make it possible to apply the technique of scattering theory to the original problem.

<u>3.13. The Nonlinear Schrödinger Equation:</u> $iu_t+u_{xx}-\frac{1}{2}\varepsilon\bar{u}u^2=0$, $\bar{u}$ is the formal complex conjugate. The following factor EW algebra for this equation was found in the work of Estabrook and Wahlquist [32]:

$$\begin{gathered}[X_1, X_2]=[X_1, Y_2]=[X_2, Y_1]=[X_2, Z_1]=[Z_1, Z_2]=0;\quad [X_1, Z_1]=Z_2,\quad [Z_1, \bar{Z}_1]=\frac{1}{2}Y_1; \\ \frac{1}{2}[X_2, Z_2]+[Y_1, Z_1]-\varepsilon Z_1=0;\quad [X_1, Z_2]+2[Y_2, Z_1]=0; \\ [X_1, Y_1]+[X_2, Y_2]=2[\bar{Z}_1, Z_2]-2[Z_1, \bar{Z}_2].\end{gathered} \tag{10}$$

(To these relations are to be added their formal complex conjugates.)

Denoting complex variables by y_1, y_2 and by $\partial_i=\frac{\partial}{\partial y_i}$, $\bar{\partial}_i=\frac{\partial}{\partial \bar{y}_i}$ the standard differentiations, we obtain the following representation (k is any complex constant):

$$\begin{gathered} X_1=2(ky_1\partial_1+\bar{k}\bar{y}_1\bar{\partial}_1); \\ Y_1=\varepsilon\left(y_1\partial_1-\bar{y}_1\bar{\partial}_1-\frac{1}{2}\partial_2+\frac{1}{2}\bar{\partial}_2\right); \\ Y_2=-2(k^2y_1\partial_1-\bar{k}^2\bar{y}_1\bar{\partial}_1); \\ Z_1=-\frac{1}{2}(y_1^2\partial_1-\varepsilon\bar{\partial}_1-y_2\partial_2); \\ Z_2=-(ky_1^2\partial_1+\varepsilon\bar{k}\bar{\partial}_1-ky_1\partial_2). \end{gathered} \tag{11}$$

Equations for solving by the inverse-scattering method can also be obtained from the corresponding psuedopotential.

According to computations of S. I. Gel'fand, algebra (10) has the factor $\mathfrak{sl}(2)\times\mathfrak{sl}(2)$, and representation (11) coincides with a representation by left-invariant fields.

4. Bäcklund Transformations

4.1. A Bäcklund transformation B, relating two systems of differential equations $E(\bar{u})=0$ and $F(\bar{v})=0$, is a system of differential equations $B(\bar{u},\bar{v})=0$ such that E and B formally imply F, and F and B formally imply E. In the language of differential algebra, suppose that E corresponds to an ideal $I_E\subset\mathcal{A}$, F to an ideal $I_F\subset\mathcal{B}$ and $\mathcal{B}$ to an ideal $I_B\subset\mathcal{A}\otimes\mathcal{B}$. Then $I_B\supset I_E\otimes 1+1\otimes I_F$. Thus, a Bäcklund transformation is an analogue of a "correspondence" in algebraic geometry. (In our case $\mathcal{A}$, $\mathcal{B}$ are differential rings over a common base differential ring K, over which all tensor products are taken; the ideals are assumed to be differentially closed and to be radical for analytic applications.)

There is no systematic theory of Bäcklund transformations, and we limit ourselves to describing a portion of the experimental material at hand taken from [26].

4.2. The Bäcklund Transformation for the sine-Gordon Equation. While investigating surfaces of constant negative curvature, Bäcklund in 1880 found the following transformation relating the sine-Gordon equation $u_{xy}=\sin u$ to itself:

$$B_a:\begin{cases}(u_1-u_0)_x=2a\sin\frac{u_1+u_0}{2}, \\ (u_1+u_0)_y=\frac{2}{a}\sin\frac{u_1-u_0}{2},\end{cases} \tag{12}$$

where a is any constant. This transformation makes it possible to obtain a sequence of solutions of the equation in quadratures starting from the solution $u_0=0$. Bianchi observed that the pair of transformations B_{a_1}, B_{a_2} commutes in the following sense: starting from the initial solution u_0, the compositions $B_{a_1}\circ B_{a_2}$ and $B_{a_2}\circ B_{a_1}$ generate a certain common solution u_3, for which

$$\operatorname{tg}\frac{u_3-u_0}{4}=\frac{a_1+a_2}{a_1-a_2}\operatorname{tg}\frac{u_1-u_2}{4},$$

where u_1 and u_2 are the solutions generated by B_{a_1} and B_{a_2} from u_0.

4.3. A Remark of Rund. In his work in [26] Rund put forth the following idea. We suppose that the equation E is the Euler–Lagrange equation with Lagrangian ω: $\delta\omega=0$ (cf. Chap. I), and we are interested in Bäcklund transformations of E into itself. One can then seek ideals I_B such that $\omega(u)-\omega(v)\in \operatorname{Im} \tau d+(I_B)$; on solutions of the system B the Lagrangians $\omega(u)$ and $\omega(v)$ differ by a divergence and therefore have zero first variation simultaneously. Rund shows that the transformation (12) belongs to this class. A certain modification of his idea is applicable to the Korteweg–de Vries equation.

4.4. Bäcklund Transformations for the Korteweg–de Vries Equation. We consider the spectral problem related to this equation defined by the Lax pair

$$\begin{cases}(\partial_x^2+2u)\psi=\lambda\psi;\\(\partial_t+4\partial_x^3+6u\partial_x+6\partial_x\circ u)\psi=0.\end{cases} \tag{13}$$

Setting $v=\psi_x/\psi$, we see that it is possible to eliminate u from (13). The result is the modified Korteweg–de Vries equation for v:

$$v_t-6v^2v_x+6\lambda v_x+v_{xxx}=0, \tag{14}$$

discovered by Miura. Equation (14) has the trivial transformation $v\mapsto -v$; then applying the transformation inverse to (13) we obtain a Bäcklund transformation of the Korteweg–de Vries equation into itself which we denote by B_λ .

Estabrook and Wahlquist found that if $u_i=B_{\lambda_i}(u_0)$, $i=1,2$, then $B_{\lambda_1}B_{\lambda_2}(u_0)$ and $B_{\lambda_2}B_{\lambda_1}(u_0)$ contain the solution

$$u_3=u_0+\frac{\lambda_1-\lambda_2}{u_1-u_2}:$$

an analogoue of Bianchi's formula for the sine-Gordon equation.

4.5. Flaschka and MacLaughlin in their paper contained in [26] studied how the transformation (12) acts on the spectrum of the Schrödinger operator. Changing the normalization slightly, we write the equation in the form $u_t-6uu_x+u_{xxx}=0$ and the Schrödinger operator in the form $\partial_x^2+u=l$. Setting $u=2w_x$, we obtain the new Schrödinger operator $\partial_x^2+U=L$, where $U=2W_x$, and W is found from the equations

$$\begin{cases}W_x=-w_x+(W-w)^2+\lambda;\\W_t=-w_t+4[\lambda W_x+w_x^2+w_x(W-w)^2+w_{xx}(W-w)].\end{cases}$$

Here $W-w=-\psi_x/\psi$, where $l\psi=\lambda\psi$.

If u is rapidly decreasing as $|x|\to\infty$, the scattering data for $l=-\partial_x^2+u$ are determined as follows. For any complex number k with $\operatorname{Im}k\geqslant 0$ there are solutions of the problem $Lf=k^2f$, distinguished by the asymptotic behavior at infinity:

$$f_1(x, k) \simeq \begin{cases} e^{ikx}, & x \to +\infty; \\ b(k) e^{-ikx} + a(k) e^{ikx}, & x \to -\infty; \end{cases}$$

$$f_2(x, k) \simeq \begin{cases} -\bar{b}(k) e^{ikx} + a(k) e^{-ikx}, & x \to +\infty; \\ e^{-ikx}, & x \to -\infty. \end{cases}$$

On the imaginary axis there may be a finite number of points of the discrete spectrum of $l: k^2 = -\eta_j^2$, $j = 1, \ldots, N$, which determine constants c_j, $j = 1, \ldots, N$, by conditions on their eigenfunctions: $\psi_j \sim \sqrt{c_j} \exp(-\eta_j x)$, if $\int_{-\infty}^{\infty} \psi_j^2(x) dx = 1$. The collection $(a(k), b(k), \eta_j, c_j)$ is called the scattering data for b.

If $\psi_0(x, \lambda_0)$ is any solution of $l\psi_0 = \lambda_0 \psi_0$, and $\psi(x, \lambda)$ is a solution of $l\psi = \lambda\psi$, then $\Psi(x, \lambda) = \psi'(x, \lambda) - \psi(x, \lambda) \frac{\psi_0'(x, \lambda_0)}{\psi_0(x, \lambda_0)}$ is a solution of the equation $L\Psi = \lambda\Psi$. In order that $\psi(x, \lambda_0)$ not vanish it suffices that λ_0 lie to the left of the spectrum of l.

We choose $\lambda_0 = -\eta^2$ and denote by $(A(k), B(k), H_j, C_j)$ the scattering data for L.

According to the analysis of Flaschka and MacLaughlin, there are the following facts.

a) If $\psi_0 = f_1(x, i\eta)$, then

$$A(k) = a(k), \quad B(k) = \frac{\eta - ik}{\eta + ik} b(k), \quad C_j = \frac{\eta - \eta_j}{\eta + \eta_j} c_j,$$

and the discrete spectra of l and L coincide.

b) If $\psi_0 = f_2(x, i\eta)$, then

$$A(k) = a(k), \quad B(k) = \frac{\eta + ik}{\eta - ik} b(k), \quad C_j = \frac{\eta + \eta_j}{\eta - \eta_j} c_j,$$

and the discrete spectra of l and L coincide.

c) If $\psi_0 = D_1 f_1(x, i\eta) + D_2 f_2(x, i\eta)$, D_1, D_2 nonzero constants, then

$$A(k) = \frac{ik + \eta}{ik - \eta} a(k), \quad B(k) = -b(k), \quad C_j = \frac{\eta + \eta_j}{\eta - \eta_j} c_j,$$

N points of the discrete spectra for l and L coincide and, moreover, L has an additional point $-\eta^2$, with normalization constant depending on D_1 and D_2.

This implies several curious conclusions.

First of all, there are Bäcklund transformations which add no solitons (these correspond to points of the discrete spectrum). They only shift the phase of solitons present in the old solution (cases a) and b): $C_j \neq c_j$).

Secondly, iteration of the transformations defined by $f_2(x, i\eta)$ and then $D_1 F_1(x, i\eta) + D_2 F_2(x, i\eta) : u \to U$ does not change $b(k)/a(k)$ and c_j, but adds the eigenvalues $-\eta^2$. For it we have

$$U(x) = u(x) - 2D f_2^2(x, i\eta) \left(1 + D \int_{-\infty}^{x} f_2^2(z, i\eta) dz\right)^{-1},$$

where D is a certain constant.

Finally, if $u(x,t)$ is a solution of the Korteweg–de Vries equation, then, as is known $\frac{d}{dt}b(k,t)=8ik^3b(k,t)$, $\frac{d}{dt}c_j(t)=8\eta_j^3c_j(t)$. This implies that Bäcklund transformation by means of $f_1(x, i\eta; t)$, changing $b(k,t)$ into $\frac{\eta-ik}{\eta+ik}b(k,t)$, takes the solution u into the new solution U, i.e., it commutes with the Korteweg–de Vries flow.

With this we conclude our brief discussion, referring the reader to the literature for further information.

5. Lax's Method for Generating the Algebra of Korteweg–de Vries Integrals

5.1. We consider the algebra $A=k[u_i^{(j)}]$ and some operator B which is Hamiltonian in the sense of Sec. 7 of Chap. I. It generates a Lie algebra structure on $A/\mathrm{Ker}\, B\frac{\delta}{\delta u}$. Explicit description of operators B, for which in this Lie algebra there exists large, e.g., infinite-dimensional, Abelian subalgebras is a very interesting question. Equations $\bar{u}_t=B\frac{\delta F}{\delta \bar{u}}$, where F lies in such a subalgebra have infinitely many conservation laws. In Chap. II we described such subalgebras for the Gel'fand–Dikii operators corresponding to Lax equations and for both the reduced and unreduced Benney operator.

In this section we present the recurrence method of Lax for the operator $B=\partial_x$, which leads to the Korteweg–de Vries algebra. The exposition is based on notes of I. Ya. Dorfman who has kindly permitted the author to use them.

5.2. Let $A=k[u^{(j)} \mid j\geqslant 0]$, $\partial: u^{(j)}\mapsto u^{(j+1)}$. We set $\tilde{A}=A/\partial A$ and for any $f\in A$ we denote by $\tilde{f}$ the image of f in $\tilde{A}$. We shall write $f\sim g$, if $\tilde{f}=\tilde{g}$. There is a Lie algebra structure on $\tilde{A}$: $\{\tilde{f},\tilde{g}\}=\left(\partial\frac{\delta f}{\delta u}\cdot\frac{\delta g}{\delta u}\right)^{\sim}$. Let $H: A\to A$ be a linear operator which is formally antisymmetric in the sense that $fHg+gHf\sim 0$ for all $f, g\in A$.

We consider a sequence of elements $f_{-1}, f_0, \ldots, f_n\in A$, $f_{-1}=cu$, $c\neq 0$. The following result is motivated by the remark in 3.15 of Chap. II.

<u>5.3. Proposition.</u> If $H\frac{\delta}{\delta u}f_i=\partial\frac{\delta}{\delta u}f_{i+1}$, $-1\leqslant i\leqslant n-1$, then

$$\text{a) } \{\tilde{f}_i,\tilde{f}_j\}=0 \text{ for } -1\leqslant i,j\leqslant n.$$
$$\text{б) } H\frac{\delta}{\delta u}f_n\sim 0.$$

<u>Proof.</u> From the antisymmetry of H and ∂ it follows that for $1\leqslant j<i\leqslant n$: $\partial\frac{\delta f_i}{\delta u}\cdot\frac{\delta f_j}{\delta u}=H\frac{\delta}{\delta u}f_{i-1}\cdot\frac{\delta f_j}{\delta u}\sim-\frac{\delta f_{i-1}}{\delta u}\cdot\partial\frac{\delta f_{j+1}}{\delta u}\sim\partial\frac{\delta f_{i-1}}{\delta u}\cdot\frac{\delta f_{j+1}}{\delta u}$. Iterating this argument, we obtain

$$\partial\frac{\delta f_i}{\delta u}\cdot\frac{\delta f_j}{\delta u}\sim\begin{cases}\partial\frac{\delta f_s}{\delta u}\cdot\frac{\delta f_s}{\delta u}, & \text{if } i-j=2s;\\ \partial\frac{\delta f_{s+1}}{\delta u}\cdot\frac{\delta f_s}{\delta u}=H\frac{\delta f_s}{\delta u}\cdot\frac{\delta f_s}{\delta u}, & \text{if } i-j=2s+1.\end{cases}$$

Both expressions on the right lie in $\mathrm{Im}\,\partial$ (the second by the antisymmetry of H) which proves assertion a). Further, $H\frac{\delta f_n}{\delta u}\sim -c^{-1}\frac{\delta f_n}{\delta u}\cdot H\frac{\delta f_{-1}}{\delta u}=-c^{-1}\frac{\delta f_n}{\delta u}\cdot\partial\frac{\delta f_0}{\delta u}\sim 0$. This completes the proof.

5.4. Proposition 5.3 makes it possible to continue the commuting collection of elements $f_{-1}, \ldots, f_n \in A$ to the commuting collection $f_{-1}, \ldots, f_n, f_{n+1}$, if $\partial^{-1} H \frac{\delta}{\delta u} f_n \in \operatorname{Im} \frac{\delta}{\delta u}$: as f_{n+1} we take any element of $\left(\frac{\delta}{\delta u}\right)^{-1} \partial^{-1} H \frac{\delta}{\delta u} f_n$.

In order to guarantee the solvability of the equation for f_{n+1}, we impose additional conditions on H. We recall that the Fréchet operator $D(g)$ for $g \in A$ has the form $D(g) = \sum_{i \geq 0} \frac{\partial g}{\partial u^{(i)}} \partial^i$. We define a mapping $[D, H]: A \to L(A, A)$, where $L(A, A)$ is the space of linear operators on A by the formula

$$[D, H] f = D(Hf) - H \circ D(f)$$

(cf. Sec. 7 of Chap. I). If $H = \sum h_i \partial^i$, then

$$[D, H] f = \sum \partial^i f \cdot D(h_i).$$

We call the antisymmetric operator H Laxitive if for any pair $f, g \in A$ with the condition $Hf = \partial g$ the operator $[D, H] g \circ \partial - [D, H] f \circ H$ is formally symmetric.

5.5 Proposition. If H is a Laxitive operator and $f_{-1}, \ldots, f_n$ satisfy the condition of Proposition 5.3, then the equation $H \frac{\delta}{\delta u} f_n = \partial \frac{\delta}{\delta u} f_{n+1}$ is solvable.

Proof. We set $g_i = \frac{\delta}{\delta u} f_i$, $i = -1, \ldots, n$, and let g_{n+1} be a solution of $Hg_n = \partial g_{n+1}$, which exists by 5.3 b). According to the results of Sec. 7 of Chap. I, it suffices to verify that the operator $D(g_{n+1})$ is symmetric: this implies that $g_{n+1} \in \frac{\delta}{\delta u} A$. This, in turn, is equivalent to the symmetry of the operator $\partial \circ D(g_{n+1}) \circ \partial$.

We begin with the case $n = -1$. The operator $[D, H] c \circ \partial$ is symmetric because of the Laxitivity of H (for the pair $f = 0$, $g = c$ in the definition of this property). This means that the following operator is symmetric:

$$\partial \circ D(g_0) \circ \partial = D(\partial g_0) \circ \partial = D(Hc) \circ \partial = [D, H] c \circ \partial.$$

Now let $n > -1$. Inasmuch as $Hg_n = \partial g_{n+1}$, we have, $D(Hg_n) = \partial \circ D(g_{n+1})$, i.e., $[D, H] g_n + H \circ D(g_n) = \partial \circ D(g_{n+1})$. Similarly, from $Hg_{n-1} = \partial g_n$ we find $[D, H] g_{n-1} + H \circ D(g_{n-1}) = \partial \circ D(g_n)$. Multiplying the first equation on the right by ∂, the second by $-H$ and adding, we obtain

$$\partial \circ D(g_{n+1}) \circ \partial = [D, H] g_n \circ \partial - [D, H] g_{n-1} \circ H + H \circ D(g_n) \circ \partial + \partial \circ D(g_n) \circ H - H \circ D(g_{n-1}) \circ H.$$

From the Laxitivity of H it follows that the sum of the first two terms is symmetric. The operators $D(g_n)$ and $D(g_{n-1})$ are symmetric, since g_n and g_{n-1} are variational derivatives. Finally, ∂ and H are antisymmetric. This completes the proof.

5.6. Thus, for any Laxitive operator H and $c \in k$ the recurrence formula

$$f_n = \left(\frac{\delta}{\delta u}\right)^{-1} (\partial^{-1} H)^{n+1} c$$

defines the generators of an Abelian Lie subalgebra in $\tilde{A}$.

5.7. Proposition. Let $c_0, c_1 \in k$. Then the operator

$$H = \partial_0^3 + 2(c_0 u + c_1)\partial_0 + c_0 u'$$

is Laxitive.

The proof is obtained by direct verification. Direct computations show that this is the general form of Laxitive operators of third order lying in $k[u, u'][\partial]$.

Setting $c = \frac{1}{2}$, $c_0 = -2$, $c_1 = 0$, we obtain by the formulas of 5.6 a sequence of integrals of the Korteweg–de Vries equation (in the form $u_t = 6uu' - u'''$). The first members have the form

$$\frac{u}{2};\ -\frac{u^2}{2};\ u^3 - \frac{uu''}{2};\ -\frac{5}{2}u^4 + \frac{5}{3}uu'^2 + \frac{10}{3}u^2u'' - \frac{1}{2}uu''''.$$

6. Solutions of Algebraic Type and Theta Functions

6.1. This section is to be considered as an appendix to Chap. II: the formulas presented here can be obtained from the facts proved there regarding the structure of bimodules over a field if they are augmented to include the classical results from the analytic theory of Riemann surfaces and Jacobian varieties. Our exposition is based on the survey of Matveev [44]; for the proofs the reader should see this survey, the literature cited there, and also the paper of Krichever [16].

6.2. The Topology of Riemann Surfaces. To each field K, which is finitely generated and one-dimensional over $\mathbf{C}$, there corresponds a one-dimensional compact, complex variety Γ, the Riemann surface of K. The field K is isomorphic to the field of meromorphic functions on Γ. The genus of K and Γ is

a) half the first Betti number of Γ;

b) the dimension of the space of holomorphic differential 1-forms on Γ: Abelian differentials of the first kind.

A Riemann surface of genus zero is the Riemann sphere or the set of points of $P^1(C)$ of the projective line over $\mathbf{C}$. The genus of the curve $\Gamma: y^2 = \prod_{j=1}^{2g+1}(z - E_j)$ is equal to g; the differentials of first kind on Γ have the form $f(z)y^{-1}dz$, where $f \in \mathbf{C}[z]$, $\deg f(z) \leqslant g-1$. Such curves are called hyperelliptic (elliptic for $g=1$). An invariant definition of a hyperelliptic curve is given by either of the following two conditions (provided the genus $\geqslant 1$).

a) On Γ there is a function with a single pole of second order.

b) The ratios of differentials of first kind generate a field of kind zero.

On a compact Riemann surface Γ of genus $g \geqslant 1$ it is possible to choose a basis of the homology group $H_1(\Gamma, Z)$ of the form $\{a_i, b_j\}$, $i, j = 1, \ldots, g$, where $(a_i, a_j) = (b_i, b_j) = 0$, $(a_i, b_j) = \delta_{ij}$.

A typical choice of $\{a_i, b_j\}$ on Γ: $y^2=\prod_{i=1}^{2g+1}(z-E_i)$; E_i real, is as follows. We take two copies of the Riemann sphere with cuts along the intervals $(E_1, E_2), (E_3, E_4), \ldots, (E_{2g+1}, \infty)$ and we glue them together crosswise along the edges of the cuts. A clockwise contour around the cut (E_{2j-1}, E_{2j}) is a_j ; a contour which on each of the sheet joins an interior point of (E_{2j-1}, E_{2j}) with an interior point of (E_{2g+1}, ∞) is b_j.

6.3. Differentials and Periods. Any meromorphic 1-differential ω on Γ is called Abelian. The numbers $A_j(\omega)=\int_{a_j}\omega$, $B_j(\omega)=\int_{b_j}\omega$ are called its periods (along a_j, b_j). The numbers $c_j=\frac{1}{2\pi i}\int\omega$ (over a circle around a logarithmic singularity) are its residues. If all the residues of ω are trivial ω is called a differential of second kind; all differentials of first kind are also differentials of second kind.

A basis of differentials of first kind (ω_k) is called normalized if $A_j(\omega_k)=\delta_{jk}$ (0 for $j\neq k$, and 1 for $j=k$).

Any Abelian differential is uniquely determined by its A-periods and principal parts at singular points. For differentials of second kind it is possible to uniquely define ω by the conditions $A_j(\omega)=0$ and any value of the principal parts. For general differentials the same is true if the sum of the residues of the prescribed principal parts is equal to zero.

6.4. The Riemann Theta Function. We consider g-dimensional complex space $\mathbf{C}^g$ and the lattice $\mathbf{Z}^g\in\mathbf{C}^g$. Let B be some complex $(g\times g)$ matrix for which there exists a constant $c>0$, such that $(\operatorname{Im} Bk, k)\geqslant c\sum_{i=1}^{g}k_i^2$ for all $k\in\mathbf{Z}^g$, $k=(k_1,\ldots,k_g)$. Corresponding to this matrix there is a theta function $\theta:\mathbf{C}^g\to\mathbf{C}$:

$$\theta(p)=\sum_{k\in\mathbf{Z}^g}\exp\{\pi i(Bk, k)+2\pi i(p, k)\},\ p\in\mathbf{C}^g.$$

If $B_{jk}=B_j(\omega_k)$ it is called the theta function of the Riemann surface Γ, referred to the basis $\{a_i, b_j\}$ of the group $H_1(\Gamma, \mathbf{Z})$ and the normalized basis of differentials of first kind (ω_j) on Γ . The matrix B is called the matrix of periods of Γ ; it is symmetric.

The following properties of the theta function are easily verified:

a) $\theta(-p)=\theta(p)$.

b) $\theta(p+v)=\theta(p)$ for all $v\in\mathbf{Z}^g$.

c) $\theta(p+B^j)=e^{-\pi i B_{jj}-2\pi i p_j}\theta(p)$, where B^j is the j-th column of B.

We now assign to each point $P\in\Gamma$ its Jacobian coordinates: $P\mapsto\omega(P)=\left(\int_{P_0}^{P}\omega_k\right)\in\mathbf{C}^g$. Here P_0 is a fixed point of Γ. The Jacobian coordinates of P are defined up to the lattice of periods of Γ spanned by the columns of the matrix B , since different paths of integration from P_0 to P differ by some cycle representing a homology class in $H_1(\Gamma, \mathbf{Z})$.

The multivalued function $\Gamma\to\mathbf{C}:P\mapsto\theta(\omega(P)-e)$, where $e\in\mathbf{C}^g$, is called the Riemann theta function. It is meromorphic and is either identically zero or has exactly g zeros on Γ.

The Jacobian coordinates of these zeros are $e-\varkappa$, where the vector $\varkappa$ is defined by the relations

$$\varkappa_m=\frac{1}{2}B_{mm}-\frac{1}{2}-\sum_{j\neq m}\int_{a_j}\omega_j(P)\int_{P_0}^{P}\omega_m$$

(the Riemann constants).

If the Riemann theta function is nonzero, then the divisor of its zeros is nonspecial, and any nonspecial divisor of degree g is a divisor of the zeros of a suitable theta function.

From properties b) and c) of the theta function it follows that $d\log\theta(\omega(P)-e)$ is a meromorphic differential on Γ, having as its divisor of poles the zeros of θ. All poles have first order.

6.5. Analytic Description of the Akhiezer Function. We fix the following data: a Riemann surface Γ of genus g, a point $P_0\in\Gamma$, a nonspecial divisor $D=\sum_{i=1}^{g}P_i$ on Γ, a local parameter k^{-1} in a neighborhood of P_0, and two polynomials R, $Q\in\mathbb{C}[k]$ of degrees n and m, respectively.

6.6. THEOREM. There exists a unique meromorphic function $\psi:\mathbb{C}^3\times(\Gamma\setminus P_0)\to\mathbb{C}$ with the following properties:

a) The divisor of the poles of ψ in P coincides with D for any $x, y, t\in\mathbb{C}^3$.

b) At the point P_0 the function ψ has an essential singularity of the form

$$\psi(x, y, t; P)=\exp[k(P)+R(k(P))+Q(k(P))]\cdot(1+O(k(P)^{-1})).$$

c) $\psi(0, 0, 0; P)=1$.

This function has the form

$$\psi(x, y, t; P)=\frac{\theta(\omega(P)-e(x, y, t))\,\theta(-e(0, 0, 0))}{\theta(\omega(P)-e(0, 0, 0))\,\theta(-e(x, y, t))}\exp\left(\int_{P_0}^{P}(\omega_1x+\omega_2y+\omega_3t)\right),$$

where ω_1, ω_2, ω_3 are normalized Abelian differentials of second kind with a pole solely at P_0 and principal parts $d(k+c\log k)$, $dR(k)$, $dQ(k)$, respectively, and

$$e(x, y, t)=\frac{1}{2\pi i}(B(\omega_1)x+B(\omega_2)y+B(\omega_3)t)+\sum_{j=1}^{g}\omega(P_j)+\varkappa,$$

$$B(\omega_i)=\left(\int_{b_j}\omega_i\right),\quad \varkappa=(\varkappa_m)\quad \text{is the Riemann vector.}$$

6.7. The connection of this construction with the results of Sec. 3 of Chap. III is as follows. Let $\mathcal{B}$ be a field of meromorphic functions in x, y, t, and let $\mathcal{O}$ denote the ring of meromorphic functions on Γ with pole at P_0. Then on the space of functions $\mathcal{B}[\partial_x]\psi$ it is possible to introduce the structure of a $(\mathcal{B},\mathcal{O})$-bimodule with ψ as 1, in which the filtration is induced by the filtration of $\mathscr{B}[\partial_x]$, and $\nabla_x, \nabla_y, \nabla_t$ coincide with $\partial_x, \partial_y, \partial_t$, respectively; multiplication by $\mathcal{O}$ is the natural multiplication.

Further calculations along the lines of Chap. III but using the analytic information now at hand lead to the following explicit formulas.

6.8. The Korteweg–de Vries Equation $u_t = 6uu_x - u_{xxx}$. As was explained in Chap. III, in order that under the imbedding $\mathcal{O} \to \mathscr{B}[\partial_x]$ the image of $\mathcal{O}$ contain a Schrödinger operator of second order $-\partial_x^2 + u$, it is necessary that Γ be hyperelliptic. Let Γ be the Riemann surface of the curve $y^2 = \prod_{j=1}^{2g+1} (z - E_j)$, $E_j \neq E_k$. We choose a normalized basis of differentials of first kind

$$\omega_j = \left(\sum_{k=1}^{g} c_{jk} z^{g-k} \right) y^{-1} dz,$$

where the c_{jk} are found from the system of equations $\int_{a_k} \omega_j = \delta_{jk}$, and the a_k are defined as in 6.2. Then for $P_0 = \infty$ and a nonspecial divisor $D = P_1 + \ldots + P_g$ we have

$$u(x, t) = -2\partial_x^2 \log \theta (xg + tv + l) + c,$$

where $g, v, l \in \mathbf{C}^g$, $c \in \mathbf{C}$ and

$$g_j = 2ic_{j1}; \quad v_j = 8i \left(\frac{1}{2} c_{j1} \left(\sum_{m=1}^{2g+1} E_m \right) + c_{j2} \right);$$

$$l_j = -\sum_{k=1}^{g} \int_{\infty}^{P_k} \omega_j + \frac{j}{2} - \frac{1}{2} \sum_{m=1}^{g} B_{mj};$$

$$c = \sum_{j=1}^{2g+1} E_j - 2 \sum_{j,m=1}^{g} \int_{a_m} y\omega_j.$$

(The nonuniqueness in the choice of path of integration in the formulas for l_j does not affect $u(x, t)$).

The Akhiezer function

$$\psi(x, t, P) = \frac{\theta(\omega(P) + xg + tv + l)\,\theta(tv + l)}{\theta(\omega(P) + tv + l)\,\theta(xg + tv + l)} \exp\left(i \int_{\infty}^{P} x\nu \right)$$

is a solution of the Schrödinger equation $(-\partial_x^2 + u(x, t))\psi = y(P)\psi$. Here $P \in \Gamma$ and ν is a normalized differential of second kind with $\nu = d\sqrt{z} + O(z^{-3/2} dz)$ as $z \to \infty$, and having no other poles.

If all or part of the pairs (E_m, E_{m+1}) coalesce the limit solutions will be g-solitons or solitons on an almost-periodic background.

6.9. The Kadomtsev–Petviashvili Equations. These correspond to the Akhiezer function $\psi(x, y, t; P)$, defined in Theorem 6.6 for which Γ is any Riemann surface, P_0 is any point on it, $R(k) = k^2$, $Q(k) = k^3$. In the ring $\mathscr{B}[\partial_x]$ with the help of the bimodule $\mathscr{B}[\partial_x]\psi$ we seek operators

$$L=\partial_x^2+u(x,y,t),\ P=\partial_x^3+v_1(x,y,t)\partial_x+v_2(x,y,t),$$

satisfying the Zakharov–Shabat equation $\partial_y P-\partial_t L=[L,P]$. In terms of the coefficients u, v_1, v_2 this system can be rewritten in the form

$$\begin{cases} v_{2y}-u_t+v_{2xx}-u_{xx}-v_1u_x=0; \\ v_{1xx}+2v_{2x}-3u_{xx}+v_{1y}=0; \\ 2v_{1x}=3u_x. \end{cases}$$

Solutions constructed on the basis of the bimodule actually even satisfy the condition $2v_1=3u$. Eliminating v_1 and v_2 from this system, we obtain the Kadomtsev–Petviashvili equation

$$\frac{3}{4}u_{yy}+u_{xt}+\frac{1}{4}(u_{xxx}+6uu_x)_t=0,$$

which is also called the two-dimensional Korteweg–de Vries equation. Its solution can be written in the form $2\partial_x^2\log\theta(xg+tv+yh+l)$.

LITERATURE CITED

1. O. I. Bogoyavlenskii and S. P. Novikov, "On the connection of the Hamiltonian formalisms of stationary and nonstationary problems," Funkts. Analiz Prilozhen., 10, No. 1, 9-13 (1976).
2. I. M. Gel'fand and L. A. Dikii, "Asymptotic resolvents of Sturm–Liouville equations and the algebra of the Korteweg–de Vries equation," Usp.Mat. Nauk, 30, No. 5, 67-100 (1975).
3. I. M. Gel'fand and L. A. Dikii, "The structure of Lie algebras in the formal calculus of variations," Funkts. Analiz Prilozhen., 10, No. 1, 28-36 (1976).
4. I. M. Gel'fand and L. A. Dikii, "Fractional powers of operators and Hamiltonian systems," Funkts. Analiz Prilozhen., 10, No. 4, 13-29 (1976).
5. I. M. Gel'fand and L. A. Dikii, "Resolvents and Hamiltonian systems," Funkts. Analiz Prilozhen., 11, No. 2, 11-27 (1977).
6. I. M. Gel'fand, Yu. I. Manin, and M. A. Shubin, "Poisson brackets and the kernel of the variational derivative in the formal calculus of variations," Funkts. Analiz Prilozhen., 10, No. 4, 30-34 (1976).
7. V. G. Drinfel'd, "On commutative subrings of some noncommutative rings," Funkts. Analiz Prilozhen., 11, No. 1, 11-14 (1977).
8. B. A. Dubrovin, V. B. Matveev, and S. P. Novikov, "Nonlinear equations of Korteweg–de Vries type, finite-zone linear operators, and Abelian manifolds," Usp.Mat. Nauk, 31, No. 1, 55-136 (1976).
9. V. E. Zakharov, "A kinetic equation for solitons," Zh. Eksp. Teor. Fiz., 60, No. 3, 993-1000 (1971).
10. V. E. Zakharov and S. V. Manakov, "On complete integrability of the KdV equation and the nonlinear Schrödinger equation," Preprint IYaF, No. 68-73, Novosibirsk (1973).
11. V. E. Zakharov, L. A. Takhtadzhyan, and L. D. Faddeev, "A complete description of sine-Gordan solutions," Dokl. Akad. Nauk SSSR, 219, No. 6, 1334-1337 (1974).
12. V. E. Zakharov and L. D. Faddeev, "The Korteweg–de Vries equation — a completely integrable Hamiltonian system," Funkts.Analiz Prilozhen., 5, No. 4, 18-27 (1971).
13. V. E. Zakharov and A. B. Shabat, "A scheme for integrating nonlinear equations of mathematical physics by the method of inverse scattering," Funkts. Analiz Prilozhen., 8, No. 3, 43-53 (1974).
14. A. P. Its and V. B. Matveev, "On a class of solutions of the Korteweg–de Vries equations," in: Probl. Mat. Fiz., No. 8, Leningr. Univ., Leningrad (1976), pp. 70-92.
15. I. M. Krichever, "Potentials with zero reflection coefficients on a background of finite-zone potentials," Funkts. Analiz Prilozhen., 9, No. 2, 77-78 (1975).
16. I. M. Krichever, "Integration of nonlinear equations by methods of algebraic geometry," Funkts. Analiz Prilozhen., 11, No. 1, 15-31 (1977).

17. E. A. Kuznetsov and A. V. Mikhailov, "On the complete integrability of the classical two-dimensional Thirring model," Teor. Mat. Fiz., 30, No. 3, 303-314 (1977).
18. B. A. Kupershmidt, "The Lagrangian formalism in the calculus of variation," Funkts. Analiz Prilozhen., 10, No. 2, 77-78 (1976).
19. B. A. Kupershmidt and Yu. I. Manin, "Equations of long waves with a free surface. I. Conservation laws and solutions. II. The Hamiltonian structure and higher equations," Funkts. Analiz Prilozhen., 11, No. 3, 31-42 (1977).
20. Yu. I. Manin, "Rational points of algebraic curves over function fields," Izv. Akad. Nauk SSSR, Ser. Mat., 27, No. 6, 1397-1442 (1963).
21. S. P. Novikov, "The periodic problem for the Korteweg-de Vries equation. I," Funkts. Analiz Prilozhen., 8, No. 3, 54-66 (1974).
22. L. A. Takhtadzhyan and L. D. Faddeev, "An essentially nonlinear one-dimensional model of classical field theory," Teor. Mat. Fiz., 21, No. 2, 160-174 (1974).
23. L. D. Faddeev, "The inverse problem of quantum scattering theory. II," in: Sovr. Probl. Mat. (Itogi Nauki i Tekhn. VINITI AN SSSR), Moscow (1974), pp. 93-180.
24. L. D. Faddeev, "In search of multidimensional solitons," in: Nonlocal, Nonlinear, and Nonrenormalizable Field Theories," JINR (1976), pp. 207-223.
25. H. Airault, H. P. McKean, and J. J. Moser, "Rational and elliptic solutions of the Korteweg-de Vries equation and a related many-body problem," Preprint (1976).
26. R. M. Miura (editor), Bäcklund Transformations, the Inverse Scattering Method, Solitons, and Their Applications, NSF Research Workshop on Contact Transformations, Lecture Notes in Mathematics, No. 515, Springer-Verlag, Berlin (1976).
27. F. Calogero, "Exactly solvable one-dimensional many-body problems," Ist. di Fisica G. Marconi, Univ. di Roma, Preprint (1975).
28. F. Calogero and A. Degasperis, "Nonlinear equations solvable by the inverse spectral transform method. II," Ist. di Fisica G. Marconi, Univ. di Roma, Preprint No. 20 (1976).
29. F. Calogero and A. Degasperis, "Bäcklund transformations, nonlinear superposition principles, multisoliton solutions, and conserved quantities for the boomeron nonlinear evolution equation," Lettere Nuovo Cimento, 16, No. 14, 434-438 (1976).
30. D. V. Choodnovsky and G. V. Choodnovsky, "Pole expansion of nonlinear partial differential equations," Preprint (1977).
31. J. Corones, "Solitons and simple pseudopotentials," J. Math. Phys., 17, No. 5, 1867-1872 (1976).
32. F. B. Estabrook and H. D. Wahlquist, "Prolongation structures of nonlinear evolution equations," J. Math. Phys., 17, No. 7, 1293-1297 (1976).
33. Exact Treatment of Nonlinear Lattice Waves, Progr. Theor. Phys., 59, (1976).
34. L. D. Faddeev, "Quantization of solitons," Preprint, Inst. for Adv. Study, Princeton (1976).
35. L. D. Faddeev, "Some comments on many-dimensional solitons," Preprint, CERN, TH 2188 (1976).
36. C. S. Gardner, "The Korteweg-de Vries equation and generalizations. IV. The Korteweg-de Vries equation as a Hamiltonian system," J. Math. Phys., 12, No. 8, 1548-1551 (1971).
37. P. Hasenfratz and D. A. Ross, "Are quarks short-range solitons?" Phys. Lett., 64B, No. 1, 78-80 (1976).
38. R. Hirota, "A direct method of finding exact solutions of nonlinear evolution equations, Lecture Notes in Mathematics, No. 515, Springer-Verlag, Berlin (1976), pp. 40-68.
39. R. Hirota and J. Satsuma, "A variety of nonlinear equations generated from the Bäcklund transformation for the Toda lattice," Progr. Theor. Phys., No. 59, 64-100 (1976).
40. M. D. Kruskal, "The Korteweg-de Vries equation and related evolution equations," Lect. Appl. Math., 15, 61-83 (1974).
41. M. D. Kruskal, R. M. Miura, and C. S. Gardner, "Korteweg-de Vries equation and generalizations. V. Uniqueness and nonexistence of polynomial conservation laws," J. Math. Phys., 11, 952-960 (1970).
42. P. D. Lax, "Integrals of nonlinear equations of evolution and solitary waves," Comm. Pure Appl. Math., 21, No. 5, 467-490 (1968).
43. P. D. Lax, "Periodic solutions of the KdV equations," Comm. Pure Appl. Math., 28, No. 1, 141-188 (1975).
44. V. B. Matveev, "Abelian functions and solitons," Inst. Theor. Phys., Univ. Wroclaw, Preprint No. 373 (1976).
45. H. P. McKean and E. Trubowitz, "Hill's operator and hyperelliptic function theory in the presence of infinitely many branch points," Comm. Pure Appl. Math., 29, No. 2, 143-226 (1976).

46. R. M. Miura, "Conservation laws for the fully nonlinear long wave equations," Stud. Appl. Math., 53, No. 1, 45-56 (1974).
47. R. Miura, "The Korteweg–de Vries equation: a survey of results," SIAM Rev., 18, No. 3, 412-459 (1976).
48. R. M. Miura, C. S. Gardner, and M. D. Kruskal, "Korteweg–de Vries equations and generalizations. II. Existence of conservation laws and constants of the motion," J. Math. Phys., 9, No. 8, 1204-1209 (1968).
49. H. C. Morris, "Prolongation structures and a generalized inverse scattering problem," J. Math. Phys., 17, No. 10, 1867-1869 (1976).
50. Nonlinear Waves, Lect. Appl. Math. 1974, AMS, Providence (1974).
51. A. Patani, M. Schlinwein, and Q. Shafi, "Topological charges in the field theory," J. Phys. A: Math. and Gen., 9, 1513-1520 (1976).
52. K. Pohlmeyer, "Integrable Hamiltonian systems and interaction through quadratic constraints," Comm. Math. Phys., 46, 207-221 (1976).
53. A. C. Scott, F. Y. F. Chu, and D. W. MacLaughlin, "The soliton: a new concept in applied science," Proc. IEEE, 61, 1443-1483 (1973).
54. H. D. Wahlquist and F. B. Estabrook, Prolongation structures of nonlinear evolution equations. I," J. Math. Phys., 16, No. 1, 1-7 (1975).