The Toda Hierarchy and the KdV Hierarchy

D. Gieseker

Department of Mathematics, University of California, Los Angeles, CA 90024, USA.
E-mail: dag@math.ucla.edu

Abstract: This paper considers the relation between the periodic KdV hierarchy and the limit of the periodic Toda hierarchies. By choosing the initial data of the Toda flows in a canonical way, the behavior of a certain Toda flow can mimic KdV flows. Conjecturally, a method of deforming the KdV hierarchy is given.

McKean and Trubowitz [2] showed that the theory of the KdV equation

$$\frac{\partial}{\partial t} g(x,t) = \frac{\partial^3}{\partial x^3} g(x,t) - 6g(x,t)\frac{\partial g}{\partial x}(x,t)$$

is intimately related to the geometry of a related hyperelliptic curve of infinite genus, the Bloch spectrum $\mathscr{B}_{g_t}$ of the operator

$$L_{g_t} : \psi \to \frac{d^2}{dx^2}\psi(x) + g(x,t)\psi(x) \,,$$

where $g_t = g(x,t)$. As was known classically, $\mathscr{B}_{g_t}$ is independent of t, when $g(x,t)$ evolves according to the KdV equation. Our purpose in this paper is to develop a theory of finite difference operators and their Bloch spectra and isospectral flows which mimics the KdV theory. The basic idea of this paper is to use the theory of the periodic Toda chain of length N. Here again, the periodic Toda chain can be understood in terms of a finite genus hyperelliptic curve and isospectral deformations, as van Moerbeke discovered. For instance, see [3]. So one would like to see what the relation of the Toda hierarchy is to the KdV hierarchy, how the conserved quantities are related and so forth. A start on these matters has been obtained by Toda in [4]. In this paper, the idea is that if we choose the initial data for the periodic Toda chain very carefully, then the evolution of this data under the various equations of the Toda hierarchy looks similar to the evolution of f under the KdV hierarchy. Given f, we will find a canonical choice of the initial data of the Toda equations so that the flow of this initial data under the Toda hierarchy looks like

the flow of f under the KdV hierarchy, at least to high accuracy. This choice will be given by an asymptotic series in N^{-1}. The main result of this paper is the formulation and partial verification of the conjecture given below. The method studied here also gives a way of producing analytically approximate solutions to the Toda chain hierarchy, at least conjecturally. My motivation is to use this case as a model for constructing finite genus models for the KP hierarchy, which I have studied in [1]. I also hope to use these methods to develop discrete models of the sine-Gordon equation, the non-linear Schrödinger equation and other infinite dimensional integrable systems that are related to isospectral flows. At the end of this paper, I discuss using the methods developed in this paper to construct a finite difference equation to numerically solve the KdV equation. This finite difference equation has many conserved quantities. A different approach to deforming the KdV hierarchy can be found in E. Frenkel, *Deformations of the KdV hierarchy and related soliton equations*, IMRN, 1996, 2, pp. 55–76.

We first review the KdV hierarchy from an isospectral point of view. Let T be the translation operator defined on real valued functions on $\mathbf{R}$ by

$$T(g)(x) = g(x+1) \, .$$

Let $\mathscr{C}$ be the set of functions g so that g is analytic on $\mathbf{R}$ and $T(g) = g$, i.e. g is periodic with period 1. If $g \in \mathscr{C}$, then we can define an operator

$$L_g(\psi) = -\psi'' + g(x)\psi(x)$$

on the space of infinitely differentiable functions ψ on $\mathbf{R}$. We define the Bloch spectrum $\mathscr{B}_g$ of L_g to be the set of $(\lambda, \alpha) \in \mathbf{C} \times \mathbf{C}^*$ so that there is a non-zero function ψ with $L_g(\psi) = \lambda\psi$ and $T(\psi) = \alpha\psi$. The KdV hierarchy is a sequence of non-linear differential operators $D_i : \mathscr{C} \to \mathscr{C}$ so that for any i if we have an analytic function $g(x,t)$ and we define $g_t(x)$ to be $g(x,t)$ and

$$D_i(g_t) = \frac{d}{dt}g_t \, ,$$

then $\mathscr{B}_{g_t}$ is independent of t.

The Toda hierarchy is quite similar to the KdV hierarchy. Let us fix a positive integer N. If $A : \mathbf{Z} \to \mathbf{C}$, then we can again define the translation operator T by $T(A)(n) = A(n+N)$. Let $\mathscr{C}_N$ be the set of (A, B) with $T(A) = A$ and $T(B) = B$. Given $(A, B) \in \mathscr{C}_N$, we can form an operator on the space of all functions $\psi : \mathbf{Z} \to \mathbf{C}$ by

$$L_{(A,B)}(\psi)(n) = (-\psi(n+1) + A(n)\psi(n) + B(n)\psi(n-1))N^2 \, .$$

We define the Bloch spectrum $\mathscr{B}_{(A,B)}$ of $L_{(A,B)}$ to be the set of $(\lambda, \alpha) \in \mathbf{C} \times \mathbf{C}^*$ so that there is a non-zero function ψ with $L_{(A,B)}(\psi) = \lambda\psi$ and $T(\psi) = \alpha\psi$. Then there is a sequence of non-linear difference operators $D_{1,k}(A,B)$ and $D_{2,k}(A,B)$ so that if A_t and B_t are in $\mathscr{C}_N$ and $A_t(n)$ and $B_t(n)$ are differentiable functions of t and we have

$$D_{1,k}(A_t, B_t) = \frac{d}{dt}A_t$$

and

$$D_{2,k}(A_t, B_t) = \frac{d}{dt}B_t \, ,$$

then $\mathscr{B}_{(A_t, B_t)}$ is independent of t.

The Toda hierarchy can be defined in the following way: We first inductively define a sequence of complex valued functions d_i on $\mathbf{Z}$ as follows: We take $d_i = 0$ for $i < 0$ and $d_0 = 1$. We also take $d_i(0) = 0$ and finally

$$d_i(n + 1) = d_i(n) + A(n)d_{i-1}(n) + B(n)d_{i-2}(n) .$$

The first Toda equation is defined by

$$D_{1,1}(A, B)(n) = (B(n) - B(n + 1))N$$

and

$$D_{2,1}(A, B)(n) = B(n)(A(n) - A(n - 1))N .$$

To define the kth member of the Toda hierarchy $(k > 1)$, we define a sequence of functions $a_{k-1}, \ldots, a_0$ by descending induction on p by

$$a_p(n) = d_{k-p}(n + k) - d_{k-p}(n) - \sum_{r=p+1}^{k-1} a_r(n)d_{r-p}(n + r) .$$

We define

$$D_{1,k}(A, B)(n) = (a_1(n)B(n + 1) - B(n)a_1(n - 1))N$$

and

$$D_{2,k}(A, B)(n) = B(n)(a_0(n) - a_0(n - 1))N .$$

The numbers $d_i(N)$ are conserved quantities of the flow. If necessary, we denote $d_i(N)$ by $d_i(N, A, B)$.[1]

Given $g \in \mathscr{C}$, our question is whether it is possible to find N and $(A, B) \in \mathscr{C}_N$ so that the Toda flows with initial data A and B look like the KdV flows for g and so that $\mathscr{B}_g$ looks like $\mathscr{B}_{A,B}$. Rather than define precisely what we mean by *looks*

---

[1]While it is not necessary for the rest of the paper, here is the algebro-geometric interpretation of the Toda lattice equations above. The curve $\mathscr{B}_{A,B}$ is a double cover of the λ line. It can be compactified by adding two points P and Q over $\lambda = \infty$ and $\lambda^{-1} = z$ is a parameter at these two points. Further, the function α has a pole of order N at P and a zero of order N at Q. Let $h(z) = \alpha z^N$, regarded as a function of z in a neighborhood of P. To simplify notation, let's assume that $h(0) = 1$. If A and B are generic with $h(0) = 1$, then given a generic effective divisor D of degree g, the genus of $\mathscr{B}_{A,B}$, we can find unique functions s_n in $L(D + n(P - Q))$ so that the top coefficient of the Laurent expansion of s_n is one. Thus $s_0 = 1$ and $s(N) = \alpha$. We define $d_i(n)$ to be the ith Taylor coefficient of $z^n s_n$ at P. Thus $d_i(N)$ is the ith Taylor coefficient of $h(z)$ expanded as a power series in z. In particular, $d_i(N)$ does not depend on D, but only on the curve $\mathscr{B}_{A,B}$. It is an easy consequence of Riemann–Roch that s_n, s_{n+1}, s_{n-1}, and λs_n satisfy a linear relation, which turns out to be of the form

$$\lambda s_n + s_{n+1} = A_D s_n + B_D s_{n-1} ,$$

where A_D and B_D are functions which depend only on the linear equivalence class of D. Further, there is a divisor D_0 so that A_{D_0} is the original A and B_{D_0} is the original B. It is easy to find a recursive formula for the $d_i(n)$ in terms of the A_D and B_D. We can map the curve $\mathscr{B}_{A,B}$ to the Jacobian J_g of $\mathscr{B}_{A,B}$ by $j(x) = D_0 + P - x$. The derivative of j at P is an element of the tangent space of J_g at D_0, which can be extended to a translation invariant vector field on J_g. The first Toda equation is just the derivatives of the functions A_D and B_D as functions on J_g with respect to this translation invariant vector field. The higher Toda equations come in a similar way. Instead of using the tangent vector to $j(\mathscr{B}_{A,B})$, one uses vectors in the higher osculating subspaces of $j(\mathscr{B}_{A,B})$ at P. The calculation of the Toda equations from this point of view is essentially worked out in [1, Sect. 3]. The Toda equations are the case $M = 1$ of that paper. The paper shows how to calculate the third Toda equation, but the other Toda equations are similarly worked out.

like, we will just try to solve the problem through a series of attempts. We first define

$$\varepsilon = \frac{1}{N} \ .$$

We will use this notation consistently throughout the paper. We first consider the problem of making $\mathscr{B}_{A,B}$ look like $\mathscr{B}_g$. We try the following:

$$A(n) = 2 + \varepsilon^2 g(\varepsilon n)$$

and

$$B(n) = -1 \ .$$

Suppose that $L_g(\psi) = \lambda\psi$ and $T(\psi) = \alpha\psi$ so that $(\lambda, \alpha) \in \mathscr{B}_g$. Define $\phi(n) = \psi(n\varepsilon)$. Plugging this A and B into

$$L_{(A,B)}(\phi)(n) = \frac{-\phi(n+1) + A(n)\phi(n) + B(n)\phi(n-1)}{\varepsilon^2}$$

and using Taylor's theorem to expand $\psi(\varepsilon(n+k))$ around εn, we see that

$$L_{A,B}(\phi)(n) = -\frac{d^2}{dx^2}\psi(\varepsilon n) + g(\varepsilon n)\psi(\varepsilon n) + O(\varepsilon) \ .$$

Since $L_g(\psi) = \lambda\psi$, we have

$$L_{A,B}(\phi)(n) = \lambda\phi(n) + O(\dot{\varepsilon})$$

and $T(\phi) = \alpha\phi$. Thus ϕ is an approximate eigenfunction of $L_{A,B}$ and T. It is at least plausible that there is a true eigenfunction ϕ_ε of $L_{A,B}$ and T with $L_{A,B}(\phi_\varepsilon) = \lambda_\varepsilon\phi_\varepsilon$ and $T(\phi_\varepsilon) = \alpha_\varepsilon\phi_\varepsilon$ and

$$\lim_{N\to\infty} (\lambda_\varepsilon, \alpha_\varepsilon) = (\lambda, \alpha) \ .$$

Conversely, if λ is fixed and N is large, it is reasonable to assume that if we have $(\lambda, \alpha) \in \mathscr{B}_N$, then there are points of $\mathscr{B}_g$ nearby. Thus $\mathscr{B}_g$ looks like $\mathscr{B}_{A,B}$ with this choice of (A, B). More precisely, we conjecture that if K is any constant and

$$S_K = \{(\lambda, \alpha) : |\lambda| < K\} \ ,$$

we can make the two sets $S_K \cap \mathscr{B}_g$ and $S_K \cap \mathscr{B}_{A,B}$ be as close together as we want by making N large.

The problem with this approach is that the above choice of A and B is somewhat arbitrary. If we let

$$A(n) = 2 + \varepsilon^2 \left(\frac{1}{2}g(\varepsilon n) + q(\varepsilon n)\right)$$

and

$$B(n) = -1 + \varepsilon^2 \left(\frac{1}{2}g(\varepsilon n) - q(\varepsilon n)\right),$$

the same argument shows that $\mathscr{B}_g$ looks like $\mathscr{B}_{A,B}$ with this choice of (A, B). However, the first conserved quantity

$$d_1(N) = \sum_{k=0}^{N-1} A(k) = 2N + \varepsilon \int_0^1 \frac{1}{2}g(x) + q(x)\,dx + O(\varepsilon^2)$$

depends on our choice of q. That is, we know what $A + B$ is to order ε^3, but we do not know how to define A and B. But the conserved quantities depend on A and B, not just on $A + B$. So given g, we want to have a canonical choice of A and B in terms of g. Then we can hope that the conserved quantities of the A and B will just depend on g. However, we want this canonical choice of A and B to be in some sense invariant under the various Toda flows.

We will first concentrate on the first Toda equation:

$$D_{1,1}(A,B)(n) = (B(n) - B(n+1))N$$

and

$$D_{2,1}(A,B)(n) = B(n)(A(n) - A(n-1))N \; .$$

Let us consider two functions $f_t(x) = f(x,t)$ and $h_t(x) = h(x,t)$. Set

$$A_t(n) = 2 + \varepsilon^2(f(\varepsilon n, t) + h(\varepsilon n, t))$$

and

$$B_t(n) = -1 + \varepsilon^2(f(\varepsilon n, t) - h(\varepsilon n, t))$$

and assume that $f(x,0) = g(x)/2$. With this choice of A and B, $\mathscr{B}_g$ still looks like the Bloch spectrum attached to A_0 and B_0. In general, $\mathscr{B}_{2f_t}$ looks like $\mathscr{B}_{A_t,B_t}$. Let's consider the equations

$$\frac{d}{dt}A_t(n) = D_{1,1}(A,B)(n) + O(\varepsilon^3)$$

and

$$\frac{d}{dt}B_t(n) = D_{2,1}(A,B)(n) + O(\varepsilon^3) \; .$$

That is we ask that $A_t(n)$ and $B_t(n)$ be approximate solutions to the Toda flow problem. This means that $\mathscr{B}_{A_t,B_t}$ presumably do not move much as t changes and hence $\mathscr{B}_{2f_t}$ also does not change much with t. Plugging in and using Taylor's theorem, we get

$$\frac{d}{dt}(f+h)(\varepsilon n) = -\frac{d}{dx}(f-h)(\varepsilon n) + O(\varepsilon)$$

and

$$\frac{d}{dt}(f-h)(\varepsilon n) = -\frac{d}{dx}(f+h)(\varepsilon n) + O(\varepsilon) \; .$$

So if f and h are going to produce solutions to the Toda flow that are even approximately correct, then we should ask that

$$\frac{d}{dt}(f+h)(x,t) = -\frac{d}{dx}(f-h)(x,t) + O(\varepsilon)$$

and

$$\frac{d}{dt}(f-h)(x,t) = -\frac{d}{dx}(f+h)(x,t) + O(\varepsilon) \; .$$

Obviously, we cannot achieve this by taking $f = h$, which corresponds to the case of $B = -1$. However, we can take $h = 0$ and then we do get an approximate solution

to the Toda flow with the choice

$$A_t(n) = 2 + \varepsilon^2(f(\varepsilon n, t)) \tag{1}$$

and

$$B_t(n) = -1 + \varepsilon^2(f(\varepsilon n, t)) \tag{2}$$

with

$$\frac{d}{dt} f(x, t) = -\frac{d}{dx} f(x, t) .$$

Equations (1) and (2) define a canonical choice for A and B in terms of $f = g/2$ modulo ε^3. Of course, this is a rather uninteresting solution.

Equations (1) and (2) only define A and B to order ε^3. We next ask if we can modify (1) and (2) to be canonical to order ε^4,

$$A_t(n) = 2 + \varepsilon^2 f(\varepsilon n, t) - \varepsilon^3 \Phi_3(f)(\varepsilon n, t) \tag{3}$$

and

$$B_t(n) = -1 + \varepsilon^2 f(\varepsilon n, t) + \varepsilon^3 \Phi_3(f)(\varepsilon n, t) , \tag{4}$$

where $\Phi_3(f)$ is an unknown polynomial in f and the derivatives of f. We want to choose $\Phi_3(f)$ in such a way that

$$\frac{d}{dt} A_t = (B(n) - B(n+1))N + O(\varepsilon^4)$$

and

$$\frac{d}{dt} B_t = B(n)(A(n) - A(n-1))N + O(\varepsilon^4) \tag{5}$$

for some choices at least of $f(x, t)$. It is not obvious that such a Φ_3 should exist. Using Taylor's theorem to evaluate $A(n+i)$ and $B(n+i)$, we should have

$$\frac{df}{dt} - \frac{d\Phi_3(f)}{dt} \varepsilon = -\frac{df}{dx} + \varepsilon \left(-\frac{d^2 f}{2dx^2} - \frac{d\Phi_3(f)}{dx} \right) \tag{6}$$

and

$$\frac{df}{dt} + \frac{d\Phi_3(f)}{dt} \varepsilon = -\frac{df}{dx} + \varepsilon \left(\frac{d^2 f}{2dx^2} + \frac{d\Phi_3(f)}{dx} \right) . \tag{7}$$

In particular, adding (6) and (7), we see that $f(x, t)$ should satisfy

$$\frac{d}{dt} f(x, t) = -\frac{d}{dx} f(x, t) + O(\varepsilon^2) . \tag{8}$$

Now since

$$\frac{d}{dt} f(x, t) = -\frac{d}{dx} f(x, t) + O(\varepsilon) ,$$

we have

$$\frac{d\Phi_3(f)}{dt} = -\frac{d\Phi_3(f)}{dx} + O(\varepsilon)$$

by using the chain rule to evaluate

$$\frac{d\Phi_3(f)}{dt} .$$

So assuming that (8) holds, then (6) will hold if

$$\frac{d\Phi_3(f)}{dx} = -\frac{1}{4}\frac{d^2 f}{dx^2} .$$

Thus we can take

$$\Phi_3(f) = -\frac{1}{4}\frac{df}{dx} .$$

Thus if (8) holds and

$$A_t(n) = 2 + \varepsilon^2(f(\varepsilon n, t)) - \varepsilon^3 \left(-\frac{1}{4}\frac{df}{dx}(\varepsilon n, t) \right)$$

and

$$B_t(n) = -1 + \varepsilon^2(f(\varepsilon n, t)) + \varepsilon^3 \left(-\frac{1}{4}\frac{df}{dx}(\varepsilon n, t) \right) ,$$

then A and B satisfy the Toda equations to order ε^4. Next we attempt to continue this process. That is, we try to find a function

$$\Phi_4(f)$$

of f and the derivatives of f so that if we define

$$A_t(n) = 2 + \varepsilon^2(f(\varepsilon n, t)) - \varepsilon^3 \left(-\frac{1}{4}\frac{df}{dx}(\varepsilon n, t) \right) - \varepsilon^4 \Phi_4(f)(\varepsilon n, t) \qquad (9)$$

and

$$B_t(n) = -1 + \varepsilon^2(f(\varepsilon n, t)) + \varepsilon^3 \left(-\frac{1}{4}\frac{df}{dx}(\varepsilon n, t) \right) + \varepsilon^4 \Phi_4(f)(\varepsilon n, t) , \qquad (10)$$

then A and B satisfy the Toda equations to order ε^5, provided that $f(x, t)$ satisfies some condition. To find this condition on f, we plug (9) and (10) into

$$\frac{d}{dt}A_t = (B(n) - B(n+1))N + O(\varepsilon^5) \qquad (11)$$

and

$$\frac{d}{dt}B_t = B(n)(A(n) - A(n-1))N + O(\varepsilon^5) . \qquad (12)$$

Adding (11) and (12), we see that f must satisfy:

$$\frac{d}{dt}f(x) = -\frac{d}{dx}f(x) + \left(-\frac{1}{24}\frac{d^3 f(x)}{dx^3} + \frac{1}{2}f(x)\frac{df(x)}{dx} \right)\varepsilon^2 + O(\varepsilon^3) . \qquad (13)$$

Assuming (13) and using the fact that

$$\frac{d\Phi_4(f)}{dt} = -\frac{d\Phi_4(f)}{dx} + O(\varepsilon) ,$$

some computation shows that

$$\frac{d\Phi_4(f)}{dx} = -\frac{1}{4}f\frac{df}{dx} .$$

So we see that if

$$\Phi_4(f) = -\frac{1}{8}f^2\,,$$

then both (11) and (12) will hold if (13) holds.

When we say that f satisfies (13), we mean that (13) is satisfied in the asymptotic sense. More generally, consider the equation

$$\frac{df}{dt} = c\frac{df}{dx} + \varepsilon\Psi_1(f) + \cdots + \varepsilon^n\Psi_n(f) + O(\varepsilon^{n+1})\,, \qquad (14)$$

where the Ψ_k are polynomials in f and the x-derivatives of f. Of course, it can be extremely difficult to solve

$$\frac{df}{dt} = c\frac{df}{dx} + \varepsilon\Psi_1(f) + \cdots + \varepsilon^n\Psi_n(f)\,. \qquad (15)$$

So instead we try an asymptotic solution of the form

$$f(x,t) = f_0(x,t) + f_1(x,t)\varepsilon + \cdots + f_n(x,t)\varepsilon^n\,. \qquad (16)$$

To solve our problem, substitute (15) into (16) and multiply everything out and denote the coefficient of ε^k by V_k. Then we can easily find f_k for k from 0 to n so that $V_k = 0$ for k from 0 to n. Indeed, all we have to do is to solve

$$\frac{df_k}{dt} = c\frac{df_k}{dx} + L_k(f_0,\ldots,f_{k-1})\,,$$

where the L_k are polynomials in the known $f_0,\ldots,f_{k-1}$ and their derivatives. Given initial conditions, we can then easily solve these equations. So an asymptotic solution in this sense to (13) will solve (11) and (12) to order ε^5 for t in any given bounded region. That is, for any given T, there is an integer N and a constant C so that if $n > N$, then

$$\left|\frac{d}{dt}A_t - (B(n) - B(n+1))N\right| < C\varepsilon^5$$

and

$$\left|\frac{d}{dt}B_t - B(n)(A(n) - A(n-1))N\right| < C\varepsilon^5$$

if $|t| < T$. Of course, these estimates may break down over long time intervals, i.e. C may depend on T. However, this kind of asymptotic analysis provides a tool for guessing the behavior of the equations of the Toda hierarchy, although rigorously connecting the asymptotic analysis and the behavior of the equations is probably extremely difficult.

I conjecture that this process of finding Φ_k can be continued indefinitely. At any rate suppose we define

$$
R(f) = -\frac{1}{4}\frac{df(x)}{dx} - \frac{1}{8}f(x)^2\varepsilon + \frac{1}{192}\frac{d^3f(x)}{dx^3}\varepsilon^2
$$

$$
+ \left(\frac{1}{64}f(x)\frac{d^2f(x)}{dx^2} + \frac{1}{64}\frac{df(x)}{dx}^2 - 1/32f(x)^3\right)\varepsilon^3
$$

$$
+ \left(-\frac{1}{7680}\frac{d^5f(x)}{dx^5} + \frac{1}{64}f(x)^2\frac{df(x)}{dx}\right)\varepsilon^4
$$

$$
+ \left(\frac{3}{256}\frac{df(x)}{dx}^2 f(x) + \frac{3}{512}f(x)^2\frac{d^2f(x)}{dx^2} - \frac{5}{512}f(x)^4\right.
$$

$$
\left. - \frac{1}{1536}f(x)\frac{d^4f(x)}{dx^4} - \frac{3}{2048}\frac{d^2f(x)}{dx^2}^2 - \frac{1}{384}\frac{df(x)}{dx}\frac{d^3f(x)}{dx^3}\right)\varepsilon^5
$$

and f satisfies

$$
\frac{df}{dt} = -\frac{df(x)}{dx} + \left(-\frac{1}{24}\frac{d^3f(x)}{dx^3} + \frac{1}{2}f(x)\frac{df(x)}{dx}\right)\varepsilon^2
$$

$$
+ \left(-\frac{1}{1920}\frac{d^5f(x)}{dx^5} + \frac{1}{16}f(x)^2\frac{df(x)}{dx} + \frac{1}{32}\frac{df(x)}{dx}\frac{d^2f(x)}{dx^2}\right.
$$

$$
\left. + \frac{1}{48}f(x)\frac{d^3f(x)}{dx^3}\right)\varepsilon^4 + \left(\frac{1}{128}\frac{d^2f(x)}{dx^2}^2 + \frac{1}{128}\frac{df(x)}{dx}\frac{d^3f(x)}{dx^3}\right)\varepsilon^5
$$

$$
+ \left(-\frac{1}{322560}\frac{d^7f(x)}{dx^7} + \frac{1}{128}\frac{df(x)}{dx}^3 - \frac{1}{1536}\frac{d^2f(x)}{dx^2}\frac{d^3f(x)}{dx^3}\right.
$$

$$
+ \frac{1}{384}f(x)^2\frac{d^3f(x)}{dx^3} + \frac{1}{64}\frac{df(x)}{dx}f(x)\frac{d^2f(x)}{dx^2} + \frac{1}{64}f(x)^3\frac{df(x)}{dx}
$$

$$
\left. + \frac{1}{3840}f(x)\frac{d^5f(x)}{dx^5} + \frac{1}{1536}\frac{df(x)}{dx}\frac{d^4f(x)}{dx^4}\right)\varepsilon^6 .
$$

If we define

$$
A_t(n) = 2 + \varepsilon^2(f(\varepsilon n, t)) - \varepsilon^3 R(f)(\varepsilon n, t) \tag{17}
$$

and

$$
B_t(n) = -1 + \varepsilon^2(f(\varepsilon n, t)) + \varepsilon^3 R(f)(\varepsilon n, t) , \tag{18}
$$

then

$$
\frac{d}{dt}A_t = -(B(n+1) - B(n))N + O(\varepsilon^9)
$$

and

$$
\frac{d}{dt}B_t = B(n)(A(n) - A(n-1))N + O(\varepsilon^9) .
$$

These computations were carried out using Maple, as is described in the Appendix.

Next let's consider the second pair of equations in the Toda hierarchy. Actually, the answer is more comprehensible if we try to solve the equations:

$$(D_{1,2} + 2D_{1,1})(A_t, B_t) = \frac{d}{dt}A_t + O(\varepsilon^9) \tag{19}$$

and

$$(D_{2,2} + 2D_{2,1})(A_t, B_t) = \frac{d}{dt}B_t + O(\varepsilon^9) . \tag{20}$$

Now define A_t and B_t by Eqs. (17) and (18). In order for (19) and (20) to hold, we must have that

$$\varepsilon^2 \frac{df}{dt} = (D_{1,2} + 2D_{1,1} + D_{2,2} + 2D_{2,1})(A_t, B_t) + O(\varepsilon^9) .$$

Plugging in our definition of A and B and using Taylor's theorem many times, we see we must have

$$
\begin{aligned}
\frac{df}{dt} =& \left(-\frac{1}{4}\frac{d^3 f(x)}{dx^3} + 3f(x)\frac{df(x)}{dx} \right) \varepsilon^2 \\
&+ \left(\frac{3}{8}f(x)\frac{d^3 f(x)}{dx^3} - \frac{9}{8}f(x)^2\frac{df(x)}{dx} - \frac{1}{64}\frac{d^5 f(x)}{dx^5} + \frac{11}{16}\frac{df(x)}{dx}\frac{d^2 f(x)}{dx^2} \right) \varepsilon^4 \\
&+ \left(\frac{3}{64}\frac{d^2 f(x)}{dx^2}^2 + \frac{3}{64}\frac{df(x)}{dx}\frac{d^3 f(x)}{dx^3} \right) \varepsilon^5 \\
&+ \left(-\frac{1}{2560}\frac{d^7 f(x)}{dx^7} - \frac{1}{64}\frac{df(x)}{dx}^3 + \frac{49}{768}\frac{d^2 f(x)}{dx^2}\frac{d^3 f(x)}{dx^3} \right. \\
&\quad - \frac{5}{64}f(x)^2\frac{d^3 f(x)}{dx^3} - \frac{7}{32}\frac{df(x)}{dx}f(x)\frac{d^2 f(x)}{dx^2} - \frac{9}{32}f(x)^3\frac{df(x)}{dx} \\
&\quad \left. + \frac{11}{640}f(x)\frac{d^5 f(x)}{dx^5} + \frac{35}{768}\frac{df(x)}{dx}\frac{d^4 f(x)}{dx^4} \right) \varepsilon^6 .
\end{aligned}
\tag{21}
$$

Conversely, if (21) is valid and we use (17) and (18) to define A and B, then (19) and (20) are valid.

Next we repeat this process for the third set of equations in the Toda hierarchy. Here the result is that if

$$
\begin{aligned}
\frac{df}{dt} =& \left(\frac{5}{4}f(x)\frac{d^3 f(x)}{dx^3} - \frac{15}{2}f(x)^2\frac{df(x)}{dx} - \frac{1}{16}\frac{d^5 f(x)}{dx^5} + \frac{5}{2}\frac{df(x)}{dx}\frac{d^2 f(x)}{dx^2} \right) \varepsilon^4 \\
&+ \left(-\frac{1}{192}\frac{d^7 f(x)}{dx^7} - \frac{5}{4}\frac{df(x)}{dx}^3 + \frac{155}{192}\frac{d^2 f(x)}{dx^2}\frac{d^3 f(x)}{dx^3} \right. \\
&\quad - \frac{45}{32}f(x)^2\frac{d^3 f(x)}{dx^3} - \frac{85}{16}\frac{df(x)}{dx}f(x)\frac{d^2 f(x)}{dx^2} + \frac{15}{8}f(x)^3\frac{df(x)}{dx} \\
&\quad \left. + \frac{11}{64}f(x)\frac{d^5 f(x)}{dx^5} + \frac{47}{96}\frac{df(x)}{dx}\frac{d^4 f(x)}{dx^4} \right) \varepsilon^6
\end{aligned}
$$

and we use (17) and (18) to define A and B, then

$$(D_{1,3} - 2D_{1,1} + 2D_{1,2})(A_t, B_t) = \frac{d}{dt}A_t + O(\varepsilon^9) \tag{19}$$

and

$$(D_{2,3} - 2D_{2,1} + 2D_{2,2})(A_t, B_t) = \frac{d}{dt}B_t + O(\varepsilon^9). \tag{20}$$

Finally, for the fourth Toda equation, we have that

$$(D_{1,4} + 4D_{1,1} - 2D_{1,2} + 2D_{1,3})(A_t, B_t) = \frac{d}{dt}A_t + O(\varepsilon^9) \tag{19}$$

and

$$(D_{2,4} + 4D_{2,1} - 2D_{2,2} + 2D_{2,3}))(A_t, B_t) = \frac{d}{dt}B_t + O(\varepsilon^9) \tag{20}$$

if

$$\frac{df}{dt} = \left(-\frac{1}{64}\frac{d^7 f(x)}{dx^7} - \frac{35}{8}\frac{df(x)^3}{dx} + \frac{35}{16}\frac{d^2 f(x)}{dx^2}\frac{d^3 f(x)}{dx^3} \right.$$

$$-\frac{35}{8}f(x)^2\frac{d^3 f(x)}{dx^3} - \frac{35}{2}\frac{df(x)}{dx}f(x)\frac{d^2 f(x)}{dx^2} + \frac{35}{2}f(x)^3\frac{df(x)}{dx}$$

$$\left. +\frac{7}{16}f(x)\frac{d^5 f(x)}{dx^5} + \frac{21}{16}\frac{df(x)}{dx}\frac{d^4 f(x)}{dx^4} \right)\varepsilon^6.$$

This suggests the following conjecture:

Conjecture. *There are polynomials $\Phi_k(f)$ and $\Psi_{k,j}(f)$ in f and the derivatives of f so that if*

$$R(f) = \sum_{k=1}^{L} \Phi_k(f)\varepsilon^k$$

and

$$Z_j(f) = \sum_{k=0}^{L+1} \Psi_{k,j}(f)\varepsilon^k$$

and we define

$$A_t(n) = 2 + \varepsilon^2(f(\varepsilon n, t)) - \varepsilon^3 R(f)(\varepsilon n, t)$$

and

$$B_t(n) = 2 + \varepsilon^2(f(\varepsilon n, t)) + \varepsilon^3 R(f)(\varepsilon n, t)$$

and f satisfies the equation

$$\frac{df}{dt} = Z_j(f) + O(\varepsilon^{L+1}),$$

then $A_t(n)$ and $B_t(n)$ satisfy the jth Toda equations to order ε^{L+4}. Further, by taking suitable linear combinations of the $Z_j(f)$, we can produce asymptotic series whose leading terms in ε are the KdV hierarchy if L is large enough.

For the rest of the paper, let us assume this conjecture is true. Then we can also calculate out the conserved quantities asymptotically. First, let's consider $d_1(N)$, the first conserved quantity. We have

$$d_1(N) = \sum_{n=0}^{N-1} A(n).$$

By the Euler–Maclaurin summation formula, we have that asymptotically

$$d_1(N) = N\left(\int_0^1 (2 + \varepsilon^2 f(x) - \varepsilon^3 R(f))\, dx\right).$$

So we can write out the first few terms of the expansion of $d_1(N)$:

$$d_1(N) = \frac{1}{\varepsilon}\left(2 + \varepsilon \int_0^1 f(x)\, dx + \frac{1}{8}\varepsilon^3 \int_0^1 f(x)^2\, dx + \frac{1}{32}\varepsilon^5 \int_0^1 f(x)^3\, dx + \cdots\right).$$

In particular, the following quantity is conserved asymptotically:

$$C1 = \int_0^1 f(x)\, dx + \frac{1}{8}\varepsilon^2 \int_0^1 f(x)^2\, dx + \frac{1}{32}\varepsilon^4 \int_0^1 f(x)^3\, dx + \cdots.$$

Next consider $d_2(N)$. We can write

$$d_2(N) = \left(\sum_{n=0}^{N-1} A(n)\right)^2 + \sum_{n=0}^{N-1} B(n) + \sum_{n=0}^{N-1} A(n)^2.$$

Again the Euler–Maclaurin summation formula allows us to calculate each of the terms as an integral. In order to make the formulas look nicer, we consider the conserved quantity

$$C2 = -\frac{4}{3}\left(d_2(N) - \frac{2}{\varepsilon^2} + \frac{3}{\varepsilon^2}\right) + \left(\frac{2-\varepsilon}{\varepsilon}C1\right) - \frac{1}{2}C1^2. \tag{22}$$

Then we have an expansion for $C2$,

$$\varepsilon^3 \int_0^1 f(x)^2\, dx + \varepsilon^5\left(\frac{1}{4}\int_0^1 f(x)^3\, dx - \frac{1}{24}\int_0^1 f(x)\frac{d^2 f(x)}{dx^2}\, dx\right) + o(\varepsilon^7).$$

Thus if we solve a given Toda equation asymptotically to high enough order, we can expect

$$\int_0^1 f(x)^2\, dx + \varepsilon^2\left(\frac{1}{4}\int_0^1 f(x)^3\, dx - \frac{1}{24}\int_0^1 f(x)\frac{d^2 f(x)}{dx^2}\, dx\right) + \cdots$$

to be conserved to order ε^4.

The third conserved quantity is $C3$ which has the expression:

$$d_3(N) - \frac{1}{12\varepsilon}(-16 + 72\varepsilon - 56\varepsilon^2 - 24C1\varepsilon + 60C1\varepsilon^2 - 36C1\varepsilon^3 + 12C1^2\varepsilon^2$$

$$+ 18C2\varepsilon^2 + 2C1^3\varepsilon^3 + 18C2\varepsilon^3 - 9C1C2\varepsilon^3).$$

We have the following expression for $C3$:

$$\varepsilon^5 \left(-\frac{7}{12} \int\limits_0^1 f(x)^3 \, dx + \frac{1}{8} \int\limits_0^1 f(x) \frac{d^2 f(x)}{dx^2} \, dx \right) + \cdots .$$

Let's use (19) and (20) to try to find a numerical method for solving KdV. We first define $a(n) = N^2(A(n) - 2)$ and $b(n) = N^2(B(n) + 1)$. Then we define

$$L_{a,b}(k) = 2b(k+1) - 2b(k) - a(k+1) + a(k-1)$$

and

$$M_{a,b}(k) = 2a(k) - 2a(k-1) - b(k+1) + b(k-1) ,$$

$$F_{a,b}(k) = b(k+1)a(k) + b(k+1)a(k+1) - b(k)a(k) - b(k)a(k-1)$$

and

$$G_{a,b}(k) = -2b(k)a(k) + 2b(k)a(k-1) + a(k)^2 - a(k-1)^2 + b(k)b(k+1)$$
$$-b(k)b(k-1) + (-b(k)a(k)^2 + b(k)a(k-1)^2)/N^2 .$$

Equations (19) and (20) become

$$\frac{da}{dt}(k) = N(L_{a,b}(k) + \varepsilon^2 F_{a,b}(k)) \tag{23}$$

and

$$\frac{db}{dt}(k) = N(M_{a,b}(k) + \varepsilon^2 G_{a,b}(k)) . \tag{24}$$

Let's suppose we have a function $f(x)$ we want to use as initial data for KdV where f is periodic with period 1. Define $c(n) = f(n/N)$. Then we can mimic (17) and (18) to define $a(n)$ and $b(n)$ to any desired degree of accuracy in ε. Thus let's define

$$\Delta(b)(k) = \frac{N(b(k+1) - b(k-1))}{2} .$$

So we can approximate the first derivative of f by

$$\Delta(c) - \frac{1}{6}\varepsilon^2 \Delta^3(c) .$$

For instance, we could set

$$a(n) = c(n) + \frac{1}{4}\varepsilon^2 \left(\Delta(c)(n) - \frac{1}{6}\varepsilon^2 \Delta^3(c) \right) + \frac{1}{8}\varepsilon^3 c(n)^2 - \frac{1}{192}\varepsilon^4 \Delta^3(c) \tag{25}$$

and

$$b(n) = c(n) - \frac{1}{4}\varepsilon^2 \left(\Delta(c)(n) - \frac{1}{6}\varepsilon^2 \Delta^3(c) \right) - \frac{\varepsilon^3}{8} c(n)^2 + \frac{1}{192}\varepsilon^4 \Delta^3(c)) , \tag{26}$$

and use these a and b as initial conditions for (23) and (24) to get functions $a(n,t)$ and $b(n,t)$. We can reasonably expect that

$$\frac{a(n,t) + b(n,t)}{2}$$

will remain close to $f(n/N, t)$, where $f(x, t)$ is a solution to

$$\frac{df}{dt} = \left(-\frac{1}{4}\frac{d^3 f(x)}{dx^3} + 3f(x)\frac{df(x)}{dx} \right) \varepsilon^2 \, .$$

In fact, this method should be accurate to order ε^4. Further, we can expect a whole series of conserved quantities similar to (22).

There are serious problems in discretizing Eqs. (23) and (24) in time efficiently. This problem can be seen by setting $a(n) = \varepsilon^2 v(n/N)$ and $b(n) = \varepsilon^2 w(n/N)$. Then to order ε, we have

$$\frac{dv}{dx} = 2\frac{dw}{dx} - 2\frac{dv}{dx} \tag{27}$$

and

$$\frac{dw}{dx} = 2\frac{dv}{dx} - 2\frac{dw}{dx} \, . \tag{28}$$

So while arbitrary solutions tend to move with speeds on the order of magnitude 1, the special solutions we have constructed with (25) and (26) tend to move with speed ε^2. Thus if one applies some explicit method such as Runge–Kutte, one is virtually guaranteed that the differencing in time will produce instability if the time step size is not small with respect to 1. One can intuitively understand the situation as follows: Eqs. (23) and (24) define a vector field on the space V of functions a and b. For most of the a and b coming from smooth data, the vectors in this field have order of magnitude one. But for the set W of those coming from (25) and (26), the order of the magnitude of these vectors is ε^2. Thus W is almost fixed by the flow. In fact, W tends to act as a neutral fixed set, so solutions near W tend to stay near W. A large step size will tend to push away from W, and as more steps are taken, the distance from W increases exponentially. In practice, the high frequencies are most increased and so instability results. I believe that this problem can be overcome using the Crank–Nicholson method. Given a and b we wish to find a'' and b'', the results of propagating a and b using (23) and (24) over a time step Δt. We try to solve the equations

$$a(k) + \frac{\Delta t}{2}(N^3 L_{a,b}(k) + NF_{a,b}(k)) = a'(k) - \frac{\Delta t}{2}(N^3 L_{a',b'}(k) + NF_{a',b'}(k))$$

and

$$b(k) + \frac{\Delta t}{2}(N^3 M_{a,b}(k) + NG_{a,b}(k)) = b'(k) - \frac{\Delta t}{2}(N^3 M_{a',b'}(k) + NG_{a',b'}(k))$$

to obtain a Crank–Nicholson approximation a' and b' to a'' and b''. If we take the step size small with respect of N^2, then

$$\frac{\Delta t}{2}NF_{a,b}(k)$$

is small. So in solving for a' and b' in terms of a and b, the non-linear terms are small and so one should be able to find an algorithm to solve these equations using some version of Newton's method. At any rate, I tried a rather *ad hoc* method which was stable on the examples I tried.

Appendix

The calculations in this paper were made using the computer algebra system Maple V, Release 2. Here is an explanation of the functions in this program, followed by the program:

- dt is a procedure designed to differentiate g by t, given that

$$\frac{df}{dt} = h \, .$$

  So dt acts as a derivation of sums, products and exponentials. Further,

$$\frac{d}{dt}\frac{d}{dx}$$

  is replaced by

$$\frac{d}{dx}\frac{d}{dt} \, .$$

- d(i, n) is $d_i(n)$. It implements the recursive relation defining the $d_i(n)$.
- translate produces the Taylor expansion of $f(x + n\varepsilon)$ to order N.
- R is defined in the main part of this paper.
- aa implements the recursive definition of the Toda equations.
- XZ[p] is $D_{1,k}$ and YZ[p] is $D_{2,k}$.
- AA[p] and XX are the derivative of f for that particular Toda equation.
- check[p] checks that the time derivative of B is the same as the derivative calculated using the pth Toda equation.
- The last three lines calculate the various linear combinations of AA[p] to get the KdV hierarchy.

```
dt := proc(g,f,h)
      local i,n,p;
          if type(g,'+') then
              p := 0;
              n := nops(g);
              for i to n do p := p+dt(op(i,g),f,h) od;
              RETURN(p)
          elif type(g,'*') then
              p := 0;
              for i to nops(g) do p := p+dt(op(i,g),f,h)/op(i,g) od;
              RETURN(expand(g*p))
          elif type(g,'**') then
              RETURN(op(2,g)*op(1,g)**(op(2,g)-1)*dt(op(1,g),f,h))
          elif type(g,function) and op(0,g) = f then RETURN(h)
          elif type(g,function) and op(0,g) = diff then
              RETURN(diff(dt(op(1,g),f,h),x))
          elif type(g,numeric) or g = epi then RETURN(0)
          fi
      end
  d :=
      proc(i,n)
```

```
        options remember;
            if i<0 then RETURN(0)
            elif i=0 then RETURN(1)
            elif n=0 then RETURN(0)
            elif n<0 then RETURN(d(i,n+1)-A(n)*d(i-1,n)-B(n)*d(i-2,n-1))
            elif 0<n then RETURN(d(i,n-1)+A(n-1)*d(i-1,n-1)+B(n-1)*d(i-2,n-2))
            fi
        end
        ;
translate :=
        proc(f,n)
        local i,y;
            y := f(x); for i to N do y := y+diff(f(x),x $ i)*(epi*n)**i/i! od
        end;
R :=
proc(x)
        -1/4*diff(f(x),x)-1/8*epi*f(x)**2+1/192*diff(f(x),x $ 3)*epi**2+
            epi**3*(1/64*f(x)*diff(f(x),x $ 2)+1/64*diff(f(x),x)**2-1/32*f(x)**3)+
            epi**4*(-1/7680*diff(f(x),x $ 5)+1/64*f(x)**2*diff(f(x),x))+epi**5*(
            3/256*diff(f(x),x)**2*f(x)+3/512*f(x)**2*diff(f(x),x $ 2)-5/512*
            f(x)**4-1/1536*f(x)*diff(f(x),x $ 4)-3/2048*diff(f(x),x $ 2)**2-1/384
            *diff(f(x),x)*diff(f(x),x $ 3))
end
;
N := 11;
F := proc(x) 2+epi**2*f(x)-epi**3*R(x) end;
G := proc(x) -1+epi**2*f(x)+epi**3*R(x) end;
XZ[1] := B(0)-B(1);
YZ[1] := B(0)*(A(0)-A(-1));
for k from 2 to 4 do
        aa := proc(p,n)
                local r,X;
                    X := d(k-p,n+k)-d(k-p,n);
                    for r from p+1 to k-1 do X := X-aa(r,n)*d(r-p,n+r) od;
                    X
                end;
        XZ[k] := expand(aa(1,0)*B(1)-aa(1,-1)*B(0));
        YZ[k] := expand(B(0)*(aa(0,0)-aa(0,-1)))
od;
        A := proc(n) options remember; translate(F,n) end;
        B := proc(n) options remember; translate(G,n) end;
        for p to 4 do
            XX := convert(taylor(1/2*(XZ[p]+YZ[p])/epi**3,epi,12),polynom);
            AA[p] := collect(expand(XX),epi);
            YY := taylor(dt(G(x),f,XX)-YZ[p]/epi,epi,9);
            YY := collect(expand(YY),epi);
            check[p] := YY;
            print(p,check[p])
        od;
```

```
printlevel := 1;
AA[1];
AA[2] := collect(expand(AA[2]+2*AA[1]),epi);
AA[3] := collect(expand(AA[3]-6*AA[1]+2*AA[2]),epi);
AA[4] := collect(expand(AA[4]+20*AA[1]-6*AA[2]+2*AA[3]),epi)
```

References

1. Gieseker, D.: A lattice version of the KP equation. Acta Math. **168**, 219–248 (1992)
2. McKean, H.P., Trubowitz, E.: Hill's operator and hyperelliptic function theory in the presence of infinitely many branch points. Comm. Pure Appl. Math. **29**, 143–226 (1976)
3. van Moerbeke, P., Mumford, D.: The spectrum of difference operators and algebraic curves. Acta Math. **143**, 93–154 (1979)
4. Toda, M.: Theory of nonlinear lattices. Berlin-Heidelberg-New York: Springer-Verlag, 1989

Communicated by S.-T. Yau