Ricatti and soliton equations[1]

Michiel Hazewinkel
CWI
POBox 40799, 1009AB Amsterdam
The Netherlands

Abstract. In this lecture I note that a number of integrable systems can be viewed as rational quotients of linear dynamical systems. This holds for the KdV equation, the Toda lattices and the KP equations. In particular I point out that the KP equation is in fact an infinite dimensional Riccati equation. Conjecturally all integrable systems are rational quotients of linear ones.

AMS classification: 58F07

1. Introduction.

The main purpose of this lecture is to point out that many integrable systems can be viewed as rational quotients of linear dynamical systems, to argue that this is a good way to look at them, and to suggest that this rational quotient property might be a characteristic (and defining) property for integrable systems.

2. A micro course on the formal structure of the inverse scattering transform method for the KdV equation.

Consider the Korteweg–de Vries equation for shallow water waves in the form

$$u_t = 6uu_x - u_{xxx} \tag{2.1}$$

where $u(t,x)$ is a real valued function of x (place) and t (time), and subscripts denote partial derivatives. It is desired to solve the Cauchy problem (initial value problem) for (2.1) with initial data

$$u(0,x) = a(x) \tag{2.2}$$

where $a(x)$ is a given smooth function that decays sufficiently fast to zero as $|x| \to \infty$. For this formal outline of ISTM (inverse scattering transform method) the analytical details are of lesser importance. It suffices to know that the scheme works (and where to find the proofs, cf e.g. [1,2]).

[1] I dedicate this paper to Prof. Takeyuki Hida, at the occasion of his retirement in 1991 from Nagoya University, in admiration and gratitude for his creation of white noise analysis (infinite dimensional stochastic calculus). Without him the world of stochastics would not have been the same.

One considers an associated linear eigenvalue problem

$$(2.3)\qquad (-\frac{d^2}{dx^2}+u)\phi=\lambda\phi$$

where u is a given function of x (often called a potential) also decaying fast enough as $|x|\to 0$. Then it is known that the spectrum of the Hill operator (some say Schrödinger operator), $-\frac{d^2}{dx^2}+u$, consists of two parts:

(i) a discrete part consisting of a finite number of negative eigenvalues

$$(2.4)\qquad \lambda=-\kappa_n^2$$

(ii) a continuous part

$$\lambda=k^2.$$

The corresponding eigenfunctions are denoted $\phi_n(x)$ and $\psi_k(x)$ and these can be chosen such that

$$(2.5)\qquad \begin{aligned}&\phi_n(x)\sim c_n e^{\kappa_n x}\ as\ x\to\infty,\\ &\int\phi_n(x)^2dx=1,\ \phi_n(x)>0\,\text{as}\ x\to\infty\end{aligned}$$

$$(2.6)\qquad \psi_k(x)\sim e^{-ikx}+b(k)e^{ikx}\ as\ x\to\infty$$

$$(2.7)\qquad \psi_k(x)\sim a(k)e^{-ikx}\ as\ x\to-\infty$$

Thus one associates to a potential u a set of *spectral data* $(\kappa_n, c_n, b(k))$. The $a(k)$ and $b(k)$ are called transmission and reflection coefficients, respectively, and (2.6), (2.7) are viewed as presenting the scattering of an incident wave from ∞ by the potential u. The ISTM derives its name from this way of looking at things.

The crucial observation is now the following. If $u(x)$ evolves in time according to the KdV equation (2.1) then the associated spectral data to u evolve as follows

$$(2.8)\qquad \dot{\kappa}_n=0$$

$$(2.9)\qquad \dot{c}_n=4\kappa_n^3c_n$$

$$(2.10)\qquad \dot{b}(k)=8ik^2b(k)$$

Thus, given the spectral data at time zero $\kappa_n(0)$, $c_n(0)$, $b(k,0)$ it is a trivial matter to write down the spectral data at any time $t>0$. It remains to recover the potential u from its spectral data. This is done by means of the Gelfand–Levitan–Marchenko equation. The procedure is as follows. Given κ_n, c_n, $b(k)$ (as a function of t), form the kernel

$$(2.11)\qquad B(\xi)=\sum_{n=1}^{N}c_n^2e^{-\kappa_n^\xi}+\frac{1}{2\pi}\int_{-\infty}^{+\infty}b(k)e^{ik\xi}dk$$

Let

$$\bar{B}(x, y) = B(x + y) \tag{2.12}$$

Now consider the GLM equation for $K(x, y)$

$$K(x, y) + \bar{B}(z, y) + \int_x^\infty \bar{B}(z, y)K(x, z)dz = 0 \tag{2.13}$$

Then the potential u with the given scattering data κ_n, c_n, $b(k)$ is given by

$$u(x; t) = -2\frac{\partial}{\partial x}K(x, x; t) \tag{2.14}$$

Thus the procedure to solve the initial value problem for the KdV is as follows: find the initial scattering data (the direct scattering transform) from the initial potential $a(x) = u(x, 0)$; calculate the scattering data at time t by equations $(2.8) - (2.10)$ (a triviality); calculate the potential at time t from the scattering data at time t (the inverse scattering transform). It is worth pointing out that the first step, finding the spectrum of $u(x, 0)$, is the bottleneck for the calculation of actual concrete solutions.

Now let us take a very formal algebraic look at the inverse scattering calculation. To this end view the integral in (2.13) as a (convolution) product of $\bar{B}$ and K, so that (2.3) looks like

$$K + \bar{B} + K * \bar{B} = 0, \quad K = \bar{B} * (I + \bar{B})^{-1} \tag{2.15}$$

Thus the total picture of the transition $(\kappa_n, c_n^2, b(k) \mapsto u$ looks as follows $(\kappa_n, c_n^2,\ b(k) \overset{(1)}{\mapsto} \bar{B}(x, y) \overset{(2)}{\mapsto} K(x, y) \overset{(3)}{\mapsto} u(x))$ with (1) linear, (2) looks like taking a rational fraction, and (3) is linear again. So the KdV equation appears to be a rational quotient of the linear dynamical system

$$\dot{\kappa}_n = 0, \quad (c_n^2)^{\cdot} = 8\kappa_n^3 c_n^2, \quad \dot{b}(k) = 8ik^2 b(k).$$

I like to call this a rational covering linearization.

3. The matrix Riccati equation.

The presentation of the KdV equation as a rational quotient of a linear dynamical system in the preceding section is highly formal and from that it is far from clear that this can be made to work analytically. I have no doubt however, that the results of Segal, Wilson [5] can be seen to provide just such a picture (among other things).

Let me now turn to a covering linearization situation where the analytical details are trivial. Consider the matrix Riccati equation

$$\dot{P} = AP + PD + PBD + C \tag{3.1}$$

Here P is an unknown matrix of size $n \times m$ depending on time t and the A, B, C, D are known matrices of constants of sizes $n \times n$, $m \times n$, $n \times m$, $m \times m$, respectively.

Consider the linear dynamical system

$$\frac{d}{dt}\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} A & C \\ -B & -D \end{pmatrix}\begin{pmatrix} X \\ Y \end{pmatrix}$$

where X is an $n \times m$ matrix and Y an $m \times m$ matrix. Assume for the moment that $Y(t)^{-1}$ exists and set $P = XY^{-1}$. Then an easy calculation shows that P satisfies (3.2) if $(X, Y)^T$ satisfies (3.2). Thus (3.2) is a rational covering linearization of (3.1).

There is a small snag in the picture as presented just above in that in order to form the rational quotient XY^{-1}, the inverse, Y^{-1}, must be assumed to exist. This arises because the space of $n \times m$ matrices is not really the right space on which the Riccati equation (3.1) should be studied. It should be considered instead on the Grassmann manifold of m dimensional subspaces of $n+m$ space.

For definiteness sake let's work over the reals (the complex numbers would work just as well). Let $\mathbf{R}^{(n+m)\times m}$ be the space of all real matrices of size $(n+m) \times m$ and let $\mathbf{R}^{(n+m)\times m}_{reg}$ be the open subspace of all $(n+m) \times m$ matrices of maximal rank m. Then (3.2) is a dynamical system on this open subspace of $\mathbf{R}^{(n+m)\times m}$.

Let $Gr_m(\mathbf{R}^{n+m})$ be the Grassmann manifold of m dimensional subspaces of $\mathbf{R}^{n+m}$. To each $M \in \mathbf{R}^{(n+m)\times m}_{reg}$ associate the subspace of $\mathbf{R}^{n+m}$ spanned by the columns of M. This defines a rational quotient map

$$\pi : \mathbf{R}^{(n+m)\times m}_{reg} \rightarrow Gr_m(\mathbf{R}^{n+m}) \tag{3.3}$$

Observe that $\pi(MS) = \pi(M)$ for all $S \in GL(m; \mathbf{R})$, and inversely if $\pi(M) = \pi(M')$ then $M' = MS$ for some $S \in GL(m; \mathbf{R})$. It follows from this that π is compatible with the dynamical system (3.2); i.e. if $M(t)$, $M'(t)$ are solutions of (3.2) and $\pi(M(0)) = \pi(M'(0))$, then $\pi(M(t)) = \pi(M'(t))$ for all t. Thus (3.2) induces a dynamical system on $Gr_m(\mathbf{R}^{n+m})$, the *Riccati flow*. The space of $n \times m$ matrices, $\mathbf{R}^{n\times m}$, imbeds as an open dense subspace in $Gr_m(\mathbf{R}^{n+m})$ by

$$P \mapsto \pi\begin{pmatrix} P \\ I_m \end{pmatrix} \tag{3.4}$$

where I_m is the $m \times m$ identity matrix. The restriction of the vector field on $Gr_m(\mathbf{R}^{n+m})$ that describes the Riccati flow to the open dense subset $\mathbf{R}^{n+m}$ is the right hand side of (3.1).

The equation (3.1) has finite escape time phenomena and these are perfectly described by the Riccati flow on the compactification $Gr_m(\mathbf{R}^{n+m})$ of $\mathbf{R}^{n\times m}$.

4. The Toda lattices.

The N–particle (non-periodic) Toda lattice is given by the equations

$$\begin{aligned} \ddot{q}_n &= -e^{q_{n+1}-q_n} + e^{q_n - q_{n-1}} \\ q_0 &= \infty, \quad q_{N+1} = -\infty, \ n = 1, \dots, N \end{aligned} \tag{4.1}$$

That is, it is a Hamiltonian system with Hamiltonian

$$H = \frac{1}{2}\sum p_n^2 + \sum e^{q_{n+1}-q_n} \tag{4.2}$$

Let $a_k = -\frac{1}{2}p_k$, $b_k = \frac{1}{2}e^{\frac{1}{2}(q_k - q_{k+1})}$. Then, as is well known, the equations (4.1) transform into

$$\dot{A} = [A, B] \tag{4.3}$$

where A is the tridiagonal matrix with diagonal elements, $a_1, a_2, \ldots, a_N$; supradiagonal elements $b_1, \ldots, b_{N-1}$; and infradiagonal elements $b_1, \ldots, b_{N-1}$ and B is the tridiagonal matrix with zero diagonal elements, the same supradiagonal elements as A, and $-b_1, \ldots, -b_{N-1}$ as infradiagonal elements. I.e.

$$\begin{gathered}
A = \begin{pmatrix} a_1 & b_1 & & 0 \\ b_1 & a_2 & \ddots & \\ & \ddots & \ddots & b_{N-1} \\ 0 & & b_{N-1} & a_N \end{pmatrix}, \\
B = \begin{pmatrix} 0 & b_1 & & 0 \\ -b_1 & 0 & \ddots & \\ & \ddots & \ddots & b_{N-1} \\ 0 & & -b_{N-1} & 0 \end{pmatrix}
\end{gathered} \tag{4.4}$$

It is in this form (4.3) that I want to consider the Toda Lattice. In this setting there is the following theorem, [3,8]:

Consider the linear dynamical system

$$\dot{Y}(t) = A(0)Y, \quad Y(0) = I_N \tag{4.5}$$

so that $Y(t) = \exp(tA(0))$. Take a *QR decomposition* of $Y(t)$, i.e. write $Y(t)$ as a product

$$Y(t) = Z(t)X(t) \tag{4.6}$$

where $Z(t)$ is orthogonal and $X(t)$ is lower triangular with positive diagonal entries. Then

$$A(t) = Z(t)^{-1}A(0)Z(t) \tag{4.7}$$

solves (4.3) (with initial conditions $A(t) = A(0)$ at time zero). Note that $Y(t) \mapsto A(t)$ is again a fractional linear transformation.

The Toda lattices are one class of integrable systems that are obtained from semi-simple Lie algebras by what is known as the AKSRS (Adler Kostant Symes Reyman Semenov–Tian–Shansky) construction. In the general case much the same picture holds. The space on which the rational covering linearization lives is the socalled double of the underlying Lie algebra, [6].

5. Factorizing differential operators.

Consider a differential operator of order $n+m$ in x

$$Q = D^{n+m} + q_1 D^{n+m-1} + \cdots + q_{n+m-1} D + q_{n+m} \tag{5.1}$$

where D is short for $\frac{d}{dx}$ and the p_i are given functions of x. Let us consider the problem of factorizing Q into product

$$\begin{aligned} Q = R^*R, \quad R &= D^m + r_1 D^{m-1} + \cdots + r_m, \\ R^* &= D^n + r_1^* D^{n-1} + \cdots + r_n^* \end{aligned} \tag{5.2}$$

Here the $r_1, \dots, r_m; r_1^*, \dots, r_n^*$ are the unknowns. Writing this out one finds a set of differential equations for the r's and r^*'s involving the known functions $q_1, \dots, q_{n+m}$. In case $n = m = 1$ one finds the scalar Riccati equation, and, though, as we shall see, the problem has much to do with matrix Riccati equations, this little fact is a total red herring.

Let

$$V = \{u : Qu = 0\} \tag{5.3}$$

be the solution space of Q. This is a space of dimension $n+m$. The crucial observation of Mikio Sato is that factorizations (5.2) of Q correspond bijectively with m–dimensional subspaces of V. The correspondence is simple. Let $Q = R^*R$ be a factorization; then the solution space of R

$$W_R = \{u : Ru = 0\} \tag{5.4}$$

is a subspace of dimension m of V. Inversily, let W be an m–dimensional subspace of V. Choose a basis $u_1, \dots, u_m$ of W and define a differential operator R_W by the Wronskian formula

$$R_W u = \det \begin{pmatrix} u_1 & \dots & u_m \\ Du_1 & \dots & Du_m \\ \vdots & & \vdots \\ D^{m-1}u_1 & \dots & D^{m-1}u_m \end{pmatrix}^{-1} \begin{pmatrix} u_1 & \dots & u_m & u \\ Du_1 & \dots & Du_m & Du \\ \vdots & & \vdots & \vdots \\ D^m u_1 & \dots & D^m u_m & D^m u \end{pmatrix} \tag{5.5}$$

which defines a differential operator of the form $D^m + \dots$, and which clearly has W as its space of solutions.

6. Finite chunks of the KP equation.

The next step is to consider an operator A which takes the solution space V of Q into itself and to consider the linear flow induced by e^{tA} on $Gr_m(V)$. The easiest and most natural case is the one where Q has constant coefficients. Then the operators D^k all take V into itself. So we are interested in the flows

$$W \mapsto e^{t_k D^k} W \tag{6.1}$$

on $Gr_m(V)$. The question is: how will the operators R corresponding to W evolve if W evolves according to (6.1). The answer is the following. The factors R and R^* of Q will evolve according to the operator equations

$$\frac{\partial R}{\partial t_k} = P_k R - R D^k \tag{6.2}$$

$$\frac{\partial R^*}{\partial t_k} = - R^* P_k + D^k R^* \tag{6.3}$$

Here P_k is a differential operator of order k of the form $D^k +$ lower degree, and it is uniquely determined (by R) by the requirenment that (6.2) makes sense, i.e. that the right hand side of (6.2), which is a priori of degree $m + k - 1$, is in fact of degree $m - 1$. This determines the coefficients of $P_k = D^k + p_1 D^{k-1} + \cdots + p_k$ uniquely as differential expressions in the coefficients $r_1, \ldots, r_m$ of R, so that (6.2) becomes a set of partial differential equations (nonlinear) for $r_1, \ldots, r_m$.

Once one has guessed at (6.2) it is a triviality to prove that this equation does the job, i.e. that if u is in W_R then $e^{t_k D^k} u$ is in $W_{R(t_k)}$; simply calculate $\frac{d}{dt_k}(R(t_k)\exp(t_k D^k)(u))$.

All the above is due to M. Sato, [4]; I learned about it through some notes of K. Takasaki for which I have no proper reference.

7. The KP Hierarchy.

The next step is to observe what happens if the degree of R in (6.2) is shifted. So let R satisfy (6.2) and let $\bar{R} = RD$. Then one easily checks that

$$\frac{\partial \bar{R}}{\partial t_k} = P_k \bar{R} - \bar{R} D^k \tag{7.1}$$

(with exactly the same P_k). So there is independence (of a kind) of the degree m of R. The only natural thing to do is to multiply R with D^{-m} to get a pseudodifferential operator and this gives then the ∞ hierarchy of equations

$$R = 1 + r_1 D^{-1} + r_2 D^{-2} + \cdots \tag{7.2}$$

$$\frac{\partial R}{\partial t_k} = P_k R - R D^k \tag{7.3}$$

where calculating with the D^{-i} is done by the rule

$$D^{-1} u = u D^{-1} - u^{(1)} D^{-2} + u^{(2)} D^{-3} - u^{(3)} D^{-4} + \cdots \tag{7.4}$$

with $u^{(1)} = \frac{du}{dx}$, etc.

The equations (7.3) in fact are (more precisely, cover) the Kadomptsev–Petviashvili hierarchy of equations as shall now be shown.

The, by now, standard way of writing down the KP hierarchy is as follows. Consider a pseudo-differential operator S of the form

$$S = D + \sum_{i=1}^{\infty} s_i D^{-i} \tag{7.5}$$

For every pseudodifferential operator $U = \sum_{i=-n}^{\infty} u_i D^{-i}$ define $U_+ = \sum_{i=-n}^{0} u_i D^{-i}$ (its differential operator part). Then the KP hierarchy is

$$\frac{\partial S}{\partial t_k} = [(S^k)_+, S] \tag{7.6}$$

which is a set of partial differential equations for the coefficients $s_1, s_2, \ldots$ occurring in (7.5).

The relation between (7.3) and (7.6) is as follows. Let R satisfy (7.3); define

$$S = RDR^{-1} \tag{7.7}$$

Note that S is of the form (7.5). An easy calculation shows that

$$\frac{\partial S}{\partial t_k} = [P_k, S] \tag{7.8}$$

The final step is the simple lemma that if P_k is a differential operator of the form $P_k = D^k + p_1 D^{k-1} + \cdots + p_k$ such that $[P_k, S]_+ = 0$ (which is what is necessary for (7.8) to make sense) then P_k is of the form

$$P_k = (S^k)_+ + c_1(S^{k-1})_+ + \cdots + c_k \tag{7.9}$$

for certain constants $c_1, \ldots, c_k$. Thus up to a triangular constant coefficient transformation in the times $t_1, t_2, \ldots$ equations (7.8) are the same as (7.6). And the equations (7.8) are rationally covered by the equations (7.3) which in turn are infinite dimensional Riccati flows and hence rationally covered by a linear flow.

So the KP hierarchy is also a rational quotient of a linear flow.

Very many integrable systems arise as specializations of the KP hierarchy. However, there is nothing, a priori, that guarantees that such a specialization is compatible with the rational quotient structure. This would be the case if the specialization relations are given by linear relations at the level of the rational covering linearization. This is the case in the case of the KdV hierarchy as a specialization of the KP hierarchy. This specialization, though quite well known, is, for completeness, briefly recalled below in section 8.

This specialization problem is akin to (but different, though undoubtedly related) the specialization problem in the Zaharov–Shabat dressing method, where the question is which specializations are compatible with the dressing transformations, [9].

8. Obtaining the KdV hierarchy from the KP hierarchy.

In the case of the KdV the basic operator is the Hill operator

$$L = D^2 + u \tag{8.1}$$

Let S be the unique operator of the form (7.5) such that $S^2 = L$.

$$S = D + s_1 D^{-1} + s_2 D^{-2} \tag{8.2}$$

One readily finds $s_1 = \frac{1}{2}u$, $s_2 = -\frac{1}{4}u_x$, $s_3 = -\frac{1}{8}(u^2 + u_{xx})$, ... so that in this case S is a very special operator of this form.

The KdV hierarchy is now

$$\frac{\partial L}{\partial t_k} = [(S^k)_+, L] \tag{8.3}$$

which is a consequence of $\frac{\partial S}{\partial t_k} = [(S^k)_+, L]$. The KdV equation itself arises for the case $k = 3$. Indeed

$$(S^3)_+ = D^3 + 3s_1 D + 3s_{1x} + 3s_2 = D^3 + \frac{3}{2}uD + \frac{3}{4}u_x \tag{8.4}$$

so that we find (writing $t = t_3$)

$$u_t = \frac{\partial L}{\partial t} = [D^3 + \frac{3}{2}uD + \frac{3}{4}u_x, D^2 + u] = \frac{3}{2}uu_x + \frac{1}{4}u_{xxx} \tag{8.5}$$

one of the well known forms of the KdV equation.

9. Conjecture and conclusion.

We have seen that several (classes of) integrable dynamical systems are in fact rational quotients of linear systems. I would like to suggest here that this could always be the case and that this might be a good defining property. The word rational here is important. For of course integrable dynamical system, in any case in the finite degree of freedom case, can be canonically transformed into a linear one (action – angle coordinates). However this transformation is not (as a rule) rational.

The remarks made above raise a rather large number of questions which would be interesting (and probably very rewarding) to sort out. Some of these are the following. Is the very formal rational covering linearization of the KdV of section 1 compatible with the KdV as a specialization of the KP and the rational covering linearization of the KP? To what extent are rational covering linearizations unique? Can canonical transformations be lifted in some sense to a rational covering linearization? How should superposition principles, Bäcklund transformations, and dressing transformations be interpreted at the rational covering linearization level?

Finally, to touch on another aspect: the matrix Riccati equation has a well understood phase portrait, [7]. Can something similar be done for the KP equations as projective limits of Riccati equations?

References

1. P.G.Drazin, *Solitons*, Cambridge Univ. Pr., 1983.

2. W.Eckhaus, A. van Harten, *The inverse scattering transformation and the theory of solitons*, North–Holland, 1981.

3. A.G.Reyman, M.A.Semenov–Tian–Shansky, *Reduction of Hamiltonian systems, affine Lie algebras, and Lax equations* I, Inv. Math. **54** (1980), 81–100.

4. M.Sato, *Soliton equations as dynamical systems on infinite dimensional Grassmann manifolds*, RIMS Kokyuroku **439** (1981), 30–40.

5. G.B.Segal, G.Wilson, *Loop group and equations of KdV type*, Publ. Math. IHES **61** (1985), 5–65.

6. M.A.Semenov–Tian–Shansky, *What is a classical R–matrix*, Funk. Anal. i Pril. **17**: 4 (1983), 17–33. (English translation: Functional Analysis and Applications **17** (1984), 259–272).

7. M.A.Shayman, *Phase portrait of the matrix Riccati equation*, SIAM J. Control and Opt. **24** (1986), 1–65.

8. W.W.Symes, *Hamiltonian group actions and integrable systems*, Physica **1D** (1980), 339–374.

9. V.E.Zakharov, S.V.Manakov, *Soliton theory.* In: Physics Reviews Vol. 1, Harwood Acad. Publ., 1979, 133–190.