

# TODA MAPS, COCYCLES, AND CANONICAL SYSTEMS

CHRISTIAN REMLING

ABSTRACT. I present a discussion of the hierarchy of Toda flows that gives center stage to the associated cocycles and the maps they induce on the  $m$  functions. In the second part, these ideas are then applied to canonical systems; an important feature of this discussion will be my proposal that the role of the shift on Jacobi matrices should now be taken over by the more general class of twisted shifts.

## 1. INTRODUCTION

The first part of this paper will present and advertise a certain new view of Toda flows. In the second part, I then discuss flows on canonical systems from this point of view, or rather make some suggestions; this is a completely new topic that does not seem to have received much attention yet.

The basic idea will be to take the shift map on Jacobi matrices and its transfer matrix cocycle as the starting point and then extend this to a cocycle for the action of a larger group (see the next section please for the precise definitions). This structure displays a remarkable amount of rigidity in great generality: any such joint cocycle will update the Titchmarsh-Weyl  $m$  functions along the action (Theorem 2.3), and this in turn will imply that the group action has to preserve the spectral properties and the reflection coefficients of the Jacobi matrices it acts on (Theorem 2.4). Furthermore, only finite gap type operators can have periodic orbits (Theorem 2.5). Moreover, with small modifications, all this works for the action of any group that contains an infinite cyclic group (this is the part that will be acting by shifts) as a normal subgroup; in this paper and for the discussion of the classical hierarchies, though, we will only need  $G = \mathbb{R}^N \times \mathbb{Z}$ , with  $1 \leq N \leq \infty$ .

Toda flows and more general integrable hierarchies are a classical subject that has been studied extensively; see, for example, [6, 8, 9, 27]

---

*Date:* November 30, 2017; revised January 12, 2018.

*Key words and phrases.* Jacobi matrix, Toda flow, cocycle, canonical system.  
*2010 Mathematics Subject Classification.* Primary 34L40 37K10 47B36 81Q10.

for textbook style treatments and [2, 4] for recent work and further references. Since we will develop it more or less from scratch here, it is clear that not everything in this paper can be completely new, and indeed, there will be considerable overlap with what has been done in the literature. However, since I approach the whole subject from a point of view that is not the usual one, I hope that even those parts that discuss very well known facts (such as the statement that any two Toda flows commute, or that they act by unitary conjugation) will be of some interest.

The following section continues this introduction: it gives a non-technical overview of the general ideas; further details as well as the proofs will be discussed in Sections 3–5. Finally, in Section 6, I use these ideas as a guideline to attempt a similar approach to evolutions of *general* canonical systems, which is a new topic that (strange as that may sound) seems to have received hardly any attention yet. I have emphasized the word *general* here: if only a subclass of special canonical systems is evolved, then one is back on familiar territory and in fact the Toda, KdV, AKNS hierarchies are all of this type. My main point in Section 6 will probably be the proposal to replace the shift by a more general family of flows and make the flows to be constructed commute with these rather than the shift itself.

*Acknowledgment:* I'd like to thank Injo Hur and Darren Ong for inspiring discussions on these topics and the anonymous referee for a very careful reading of the paper and helpful comments.

## 2. TODA FLOWS, COCYCLES, AND TODA MAPS

A *Jacobi matrix* is a difference operator of the form

$$(2.1) \quad (Ju)_n = a_n u_{n+1} + a_{n-1} u_{n-1} + b_n u_n.$$

Here, we assume that  $a_n > 0$  and  $b_n \in \mathbb{R}$  are bounded sequences, and then  $J$  is a bounded self-adjoint operator on  $\ell^2(\mathbb{Z})$ .

It is convenient to have the separate notation  $\tau u$  available when the difference expression from (2.1) is applied to an arbitrary sequence  $u$ , not necessarily in  $\ell^2$ .

We denote by  $\mathcal{J}$  the space of all such (bounded) Jacobi matrices, and we endow it with the metric

$$(2.2) \quad d(J, J') = \sum_{n \in \mathbb{Z}} 2^{-|n|} (|a_n - a'_n| + |b_n - b'_n|).$$

This metric is often much more useful than the operator norm because it tends to make sets compact more easily, and it interacts well with

other quantities of interest in spectral theory such as spectral measures and  $m$  functions.

*Toda flows* are global flows on  $\mathcal{J}$ . Alternatively, and this is my preferred point of view, we can think of the Toda hierarchy as an action of the (abelian) group of polynomials  $\mathcal{P} = \mathbb{R}[x]$  (with pointwise addition as the group operation) on the space  $\mathcal{J}$  of Jacobi matrices. Before we discuss the approach via cocycles in detail, let me quickly review what is probably the most common way to construct the Toda hierarchy. Suppose that  $p \in \mathcal{P}$  is given. Then the *Lax equation*

$$(2.3) \quad \dot{J} = [p(J)_a, J]$$

defines a global flow [27, Theorem 12.6], and we can then define  $p \cdot J$  as the time one map of this flow. Here, the anti-symmetric part  $p(J)_a$  of  $p(J)$  is defined via its matrix representation in terms of the standard unit vectors  $\delta_n \in \ell^2$ : if we write  $X_{jk} = \langle \delta_j, X \delta_k \rangle$  for a bounded self-adjoint operator  $X$ , then  $(X_a)_{jk} = X_{jk}$  for  $j < k$  and  $= -X_{jk}$  if  $j > k$ , and  $(X_a)_{jj} = 0$ .

One can then establish (with a fair amount of effort, that is) the following properties: First of all, these flows all commute with each other, and this we already anticipated by saying that the abelian group  $\mathcal{P}$  acts on  $\mathcal{J}$ . (This commutativity does then ensure that we have a group action because the right-hand side of (2.3) is linear in  $p$ .) Moreover, they also commute with the shift  $S$  that sends a Jacobi matrix  $J$  to  $SJ$ , which has shifted coefficients  $(a_{n+1}, b_{n+1})$ . So we have actually obtained an action of the larger abelian group  $G = \mathcal{P} \times \mathbb{Z}$ , with the second factor acting by shifts. Finally,  $g \cdot J = U^* J U$  for some unitary  $U = U(g, J)$ . (All these properties will be given alternative explanations later in this section and the next as we develop the material.)

The maps  $J \mapsto g \cdot J$  are continuous with respect to  $d$ , in the sense spelled out below; this property will be important for us, and it is usually not mentioned in the standard treatments, so I'll state it separately and also provide a proof later, in Section 3.

**Theorem 2.1.** *For any  $R > 0$  and  $p \in \mathcal{P}$ , the map  $J \mapsto p \cdot J$  is a homeomorphism of  $(\mathcal{J}_R, d)$ , where  $\mathcal{J}_R = \{J \in \mathcal{J} : \|J\| \leq R\}$ .*

What is usually discussed is the continuity with respect to the operator norm, and a similar argument will establish Theorem 2.1. Observe also that since  $p \cdot J$  is unitarily equivalent to  $J$ , the operator norm is preserved and  $\mathcal{J}_R$  is indeed invariant under the action.

The central objects of this paper are cocycles (in the sense the word is used in dynamical systems). Let me review the definition. A cocycle takes values in a group, and for us, a special group of matrix functions

will be of great importance, so let's introduce the name

$$\mathcal{SL} = \{T : \mathbb{C} \rightarrow SL(2, \mathbb{C}) : T(z) \text{ entire, } T(x) \in SL(2, \mathbb{R}) \text{ for } x \in \mathbb{R}\}$$

for this; also, note that  $\mathcal{SL}$  indeed is a group with the obvious operation of pointwise matrix multiplication.

**Definition 2.1.** An  $\mathcal{SL}$ -cocycle associated with an action of a group  $G$  on a space  $X$  is a map  $T : G \times X \rightarrow \mathcal{SL}$  satisfying the *cocycle identity*

$$(2.4) \quad T(gh, x) = T(g, h \cdot x)T(h, x).$$

We will occasionally encounter cocycles taking values in other groups such as  $SL(2, \mathbb{C})$ ,  $SL(2, \mathbb{R})$ ,  $\mathbb{C}^\times$ , and of course these are defined in the same way. The reader interested in cocycles in general may consult [7, 15, 16] for more information, but I should also issue a warning that while this material makes for interesting reading, it will be largely irrelevant to what we do here.

For our purposes, the most fundamental example of an  $\mathcal{SL}$ -cocycle is given by the transfer matrix  $T(n; J)$ . In this case,  $G = \mathbb{Z}$ , and this acts by shifts on  $X = \mathcal{J}$ , that is,  $n \cdot J = S^n J$ . Now  $T$  is defined as

$$(2.5) \quad T(1; J) = A(J), \quad A(J) \equiv \begin{pmatrix} z-b_1 & \frac{1}{a_1} \\ a_1 & 0 \end{pmatrix},$$

and then  $T(n; J) = A((n-1) \cdot J)A((n-2) \cdot J) \cdots A(J)$  and  $T(-n; J) = T(n; (-n) \cdot J)^{-1}$  for general  $n \geq 1$ , and  $T(0; J) = 1$ . These latter definitions, having chosen  $T(1; J)$ , are of course forced on us by the cocycle identity.

It is then easily verified that  $T$  is an  $\mathcal{SL}$ -cocycle, and it is also clear that, conversely, any cocycle for an action of  $G = \mathbb{Z}$  will be of this form, for some matrix function  $A : \mathcal{J} \rightarrow \mathcal{SL}$ , which we can recover from the cocycle as  $A(J) = T(1; J)$ .

Occasionally, we will consider  $T$  for a fixed  $z \in \mathbb{C}$ , and then it becomes an  $SL(2, \mathbb{C})$ - or, if  $z \in \mathbb{R}$ ,  $SL(2, \mathbb{R})$ -cocycle.

Of course, this is a rather highbrow view of the transfer matrix  $T$ ; often it is perfectly appropriate to just think of  $T$  as the matrix that updates solution vectors, as follows: if  $u$  solves the difference equation  $\tau u = zu$  and  $Y(n) = (u_{n+1}, -a_n u_n)^t$ , then  $T(n)Y(0) = Y(n)$ .

The *Titchmarsh-Weyl  $m$ -functions*  $m_\pm$ , one for each half line  $\mathbb{Z}_\pm$ , are defined for  $z \in \mathbb{C}^+ = \{z \in \mathbb{C} : \text{Im } z > 0\}$  by

$$m_\pm(z) = \mp \frac{f_\pm(1, z)}{a_0 f_\pm(0, z)},$$

where  $f_\pm$  solve  $\tau f = zf$  and  $f_\pm \in \ell^2(\mathbb{Z}_\pm)$ . Notice that if we identify a vector  $v \in \mathbb{C}^2$ ,  $v \neq 0$ , with the point  $v_1/v_2$  on the Riemann sphere

$\mathbb{C}_\infty = \mathbb{C} \cup \{\infty\}$ , then we can say that  $m_\pm(z) = \pm F_\pm(0, z)$ , where, as above,  $F(n) = (f_{n+1}, -a_n f_n)^t$ . In fact, this is the reason we defined  $F$  and  $T$  in this way.

This has the important consequence that  $T$  is not just a cocycle of arbitrarily evolving matrices; rather, the cocycle updates  $m_\pm$  correctly along the action of  $\mathbb{Z}$  in the sense that for  $z \in \mathbb{C}^+$ ,

$$(2.6) \quad \pm m_\pm(z; n \cdot J) = T(n; J)(\pm m_\pm(z; J)).$$

Here, a matrix  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  acts on the Riemann sphere as a linear fractional transformation:  $Aw = \frac{aw+b}{cw+d}$ . Throughout this paper, an expression of this type (that is, a matrix followed by a number) will always refer to this action. In other contexts, group actions will usually be indicated by the dot notation that we have already started employing above.

One more trivial but important point is worth making here: namely, the action of an  $A \in \mathrm{SL}(2, \mathbb{C})$  on  $\mathbb{C}_\infty$  is consistent with its action on vectors  $v \in \mathbb{C}^2$  in the sense that if we apply  $A$  to such a  $v \neq 0$  to obtain  $Av \in \mathbb{C}^2$ , but then change our mind and would rather let  $A$  act as a linear fractional transformation on the point from  $\mathbb{C}_\infty$  that  $v$  represents, then it suffices to also reinterpret  $Av \in \mathbb{C}^2$  as a point from  $\mathbb{C}_\infty$ .

The Toda flows similarly come with associated cocycles that have the same basic properties. Let me report quickly on this; the construction is discussed in detail in [27, Section 12.4]. Note, however, that this reference uses slightly different conventions (on what  $T$  is, for example) and a completely different perspective; in particular, cocycles are not mentioned in [27].

Consider a fixed flow of the Toda hierarchy; here, we really want to think of this as an action of  $G = \mathbb{R}$  on  $\mathcal{J}$ . Then there is a matrix function  $B = B(z, J)$ , entire in  $z$  and continuous in  $J$ , and real on the real line, with  $\mathrm{tr} B = 0$ , such that the following holds: If  $T = T(t; J)$  is defined as the solution of

$$(2.7) \quad \dot{T} = B(z, t \cdot J)T, \quad T(0) = 1,$$

then  $T(t; J)$  is an  $\mathcal{SL}$ -cocycle for the action of  $G = \mathbb{R}$  on  $\mathcal{J}$ , and again

$$(2.8) \quad \pm m_\pm(z; t \cdot J) = T(t; J)(\pm m_\pm(z; J)).$$

The fact that  $T$  is a cocycle follows just from the form of (2.7), and this would in fact work for arbitrary  $B$ . Conversely, a differentiable cocycle for a given flow will satisfy (2.7), with  $B(J) = (d/dt)T(t; J)|_{t=0}$ . In other words, (2.7), for general  $B$ , can be thought of as a rewriting of

the cocycle property, and then we claim that an appropriate choice of  $B$  will also give (2.8).

We have explicit formulae for these matrix functions  $B$ , which I will review in Section 5; see especially (5.1) below. To just give one simple concrete example for now, consider the classical Toda flow itself ( $p(x) = x$  in (2.3)): then

$$(2.9) \quad B(J) = \begin{pmatrix} z - b_1 & 2 \\ -2a_0^2 & b_1 - z \end{pmatrix}.$$

While we will not provide a derivation of the general formula (5.1) (see [27, Section 12.4] and [18, Theorem A.1] for this), we will indicate in Section 3 how (2.9) can be obtained in the framework proposed in this paper.

We then have such a matrix function  $B = B_p$  for any choice of the polynomial  $p$ , and thus we obtain a cocycle for every flow from the hierarchy, and of course if we consider just one fixed flow, then  $G = \mathbb{R}$  is the group that is acting on the Jacobi matrices. Our next result says that we in fact obtain a cocycle  $T$  with respect to the action of the whole (and much larger) group  $G = \mathcal{P} \times \mathbb{Z}$ ; recall that  $\mathbb{Z}$  acts by shifts here, so  $(p, n) \cdot J = p \cdot S^n J = S^n(p \cdot J)$ . Let me make the definition of  $T(g; J) \in \mathcal{SL}$  completely explicit: if  $g = (p, n)$ , then  $T(g; J) = T(p; n \cdot J)T(n; J)$ , where  $T(n; J)$  was defined in (2.5) and  $T(p; J)$  is constructed as described above, by solving (2.7) for  $B = B_p(tp \cdot J)$  and then evaluating at  $t = 1$ . The alternative definition  $T(g; J) = T(n; p \cdot J)T(p; J)$  would have worked, too, and in fact this is part of what the theorem says. (Of course, contrary to what I promised, this is not really a complete definition of  $T$  yet; we also need the explicit formula (5.1) for  $B$  from Section 5.)

**Theorem 2.2.** *The matrices  $T(g; J)$  form an  $\mathcal{SL}$ -cocycle for the action of  $G = \mathcal{P} \times \mathbb{Z}$ .*

This is known if we just act by either the shift or an individual Toda flow or a combination of these, but the theorem makes the much stronger claim that the cocycle identity (2.4) also holds if  $g, h$  refer to different flows from the Toda hierarchy (possibly combined with shifts also).

The existence of a *joint* cocycle in this sense, even if the acting group is just  $G = \mathbb{R} \times \mathbb{Z}$ , is such a strong condition that the Toda hierarchy can be recovered from this property, plus the additional requirement that the matrices  $B$  from (2.7) are polynomials in  $z$ . This will be discussed in concrete style in Section 3.

From a more philosophical point of view, I believe the next result expresses the main reason why this structure (a joint cocycle extending the shift cocycle) displays so much rigidity. It will be easy to prove, but I find it mildly surprising nevertheless.

We will give it in a rather general form and consider an action of an abelian group  $G$  on  $\mathcal{J}$ , and our only assumption on  $G$  is that it contains an element of infinite order or, equivalently, a subgroup isomorphic to  $\mathbb{Z}$ . It will be notationally convenient to just denote this subgroup as  $\mathbb{Z}$  itself.

This subgroup  $\mathbb{Z}$  will act by shifts, and the application we have in mind would be to  $G = \mathbb{R} \times \mathbb{Z}$ , so in addition to the shift we have a flow that commutes with it.

We could in fact consider general, possibly non-abelian groups  $G$ , and we would then assume that  $\mathbb{Z} \trianglelefteq G$  is contained in  $G$  as a *normal* subgroup. Essentially the same proof as the one we are going to give still works, with the only modification that  $m_{\pm}$  could also get swapped as part of the updating done by the cocycle. For example, imagine a group element acting by a reflection of the coefficients of  $J$ ; in fact, the corresponding group action, which is generated by the shift and a reflection, provides probably the most basic example of such an action of a non-abelian group with a cocycle that extends the shift cocycle. In this case, the acting group is the infinite dihedral group.

We don't have any use for this non-abelian version of Theorem 2.3 in this paper, so I won't make it explicit, but I do think that it suggests new directions and I intend to pursue these topics in a future project.

**Theorem 2.3.** *Let  $G$  be an abelian group with  $\mathbb{Z} \leq G$ , and consider an action of  $G$  on  $\mathcal{J}$ , with  $\mathbb{Z}$  acting by shifts  $n \cdot J = S^n J$ . Fix  $z \in \mathbb{C}^+$ , and suppose that  $T(g; J)$  is an  $\mathrm{SL}(2, \mathbb{C})$ -cocycle that extends the shift cocycle (for this fixed  $z$ ) in the sense that  $T(n; J)$  is given by (2.5). Then, if  $g \in G$ ,  $J \in \mathcal{J}$  satisfy*

$$(2.10) \quad \liminf_{n \rightarrow \pm\infty} \|T(g; n \cdot J)\| < \infty,$$

*then  $\pm m_{\pm}(z, g \cdot J) = T(g; J)(\pm m_{\pm}(z, J))$ .*

The additional assumption (2.10) is aesthetically displeasing, but it seems necessary and it should be extremely easy to verify in all cases of interest, and let me perhaps also make a few comments on why this should be the case: Typically, we expect  $T(g; n \cdot J)$  to be bounded, uniformly in  $n \in \mathbb{Z}$  and also in  $z$ , varying over a compact subset of  $\mathbb{C}$ , and this should follow from the orbit  $\{n \cdot J\}$  having compact closure in  $\mathcal{J}$ . Then continuity of  $T(g; J)$  in  $J$  would be enough to produce the desired conclusion. This argument does not literally work in complete

generality because to make  $\mathcal{J}_R$  compact, we would have to include Jacobi matrices with  $a_n = 0$  in the definition of  $\mathcal{J}$ , which is in fact often convenient and I have done this before on other occasions, but here it has the serious drawback that then the shift cocycle is undefined on those  $J$ . The Toda cocycles, on the other hand, do have continuous extensions to Jacobi matrices with  $a_n = 0$  (and these sets are invariant under the action of  $\mathcal{P}$ ). Also, it would not be completely ridiculous to restrict the whole treatment to Jacobi matrices satisfying a uniform bound  $a_n \geq \delta > 0$  (I did this in [24]), and then these technical problems disappear entirely.

Also, note that since we're dealing with holomorphic functions, in the case of  $\mathcal{SL}$ -cocycles (which is really all we're interested in here), it suffices to establish the hypotheses of Theorem 2.3 for set of  $z$ 's with an accumulation point in  $\mathbb{C}^+$  to have the statement available for all  $z \in \mathbb{C}^+$ .

Let's now discuss the statement (rather than the assumptions) of Theorem 2.3. I believe that several remarkable things have happened here. Originally, it is natural to think of a cocycle  $T(g; x)$  as something that evolves on the side, always keeping an eye on the base dynamics  $x \mapsto g \cdot x$ . Now in the situation of Theorem 2.3, it turns out that the cocycle itself already tells the full story:  $T$  updates  $\pm m_\pm$ , but these determine  $J$ , so we recover the base dynamics from  $T$ . The *zero curvature equation* (3.5) will make a very explicit statement about this.

Next, the mere fact that the updating is done by a matrix function  $T(z) \in \mathcal{SL}$  is a very strong restriction, so only very special group actions can ever have such an associated joint cocycle that extends the shift cocycle. Let's look at this in more detail. I defined such maps,

$$\pm m_\pm(z) \mapsto T(z)(\pm m_\pm(z)),$$

earlier in [25], where I called them *transformations of type TM*, which, by an unlikely coincidence, is consistent with the name *Toda maps*, which I would now like to give them; my original idea in [25] was to let TM stand for *transfer matrix*. Examples of Toda maps are given by the matrices from a Toda or shift cocycle, but of course there are many others.

Toda maps are most naturally considered for general Herglotz functions, not necessarily coming from a Jacobi matrix; equivalently, by the one-to-one correspondence between Herglotz functions and (trace normed) canonical systems [5, 22, 28], we can think of  $T \in \mathcal{SL}$  as inducing a map on whole line canonical systems  $Ju' = -zHu$ ,  $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ . I'll use these two points of view interchangeably; in particular, a canonical system (or just its coefficient function  $H(x)$ ) in a situation where



none was introduced before will be understood to be associated with a given pair of Herglotz functions. For now, we don't really need to understand anything about canonical systems; we can just view  $H(x)$  as a convenient short-hand notation for a pair  $m_{\pm}$  of Herglotz functions.

As a final general remark, notice that Toda maps have domains, possibly quite small (or empty) ones, since there is no guarantee that letting  $T$  act on a Herglotz function will produce another Herglotz function.

**Theorem 2.4.** *Let  $T \in \mathcal{SL}$ , and suppose that  $m_{\pm}^{(j)}$  are related by the Toda map*

$$\pm m_{\pm}^{(2)}(z) = T(z)(\pm m_{\pm}^{(1)}(z)).$$

*Then  $H_1$  and  $H_2$  are unitarily equivalent, and the reflection coefficients satisfy  $|R_1(x)| = |R_2(x)|$  for almost every  $x \in \mathbb{R}$ . In particular, for any Borel set  $A \subseteq \mathbb{R}$ ,  $H_1$  will be reflectionless on  $A$  if and only if  $H_2$  is.*

Some clarifying comments are in order:

(1) The general reflection coefficients, which are defined for any canonical system, not necessarily of classical scattering type, are given by

$$R_+(z) = \frac{\overline{m_+(z)} + m_-(z)}{m_+(z) + \overline{m_-(z)}}, \quad R_-(z) = \frac{m_+(z) + \overline{m_-(z)}}{m_+(z) + m_-(z)};$$

since  $|R_+(x)| = |R_-(x)|$  at those  $x \in \mathbb{R}$  at which both  $R_{\pm}(x) \equiv \lim_{y \rightarrow 0^+} R_{\pm}(x + iy)$  exist, it doesn't matter here which reflection coefficient we take. Please see [25] for more on reflection coefficients and [10, 26] for earlier uses of them in somewhat different situations.

(2) We call  $H$  *reflectionless* on a Borel set  $A \subseteq \mathbb{R}$  if  $m_+(x) = -\overline{m_-(x)}$  for almost every  $x \in A$ . This is equivalent to  $R(x) = 0$  almost everywhere on  $A$ , and thus the final statement of Theorem 2.4 is an immediate consequence of the earlier claim that  $|R_1(x)| = |R_2(x)|$ . Please see [3, 12, 20, 24] for (much) more on reflectionless operators and why they are important.

(3) Recall that for a whole line operator (think of specifically a Jacobi matrix or a Schrödinger operator perhaps), there is a standard way to construct a specific spectral representation from the half line  $m$  functions  $m_{\pm}$ . Namely, define the matrix valued Herglotz function  $M$  as

$$(2.11) \quad M(z) = \frac{1}{m_+(z) + m_-(z)} \begin{pmatrix} -1 & \frac{1}{2}(m_+(z) - m_-(z)) \\ \frac{1}{2}(m_+(z) - m_-(z)) & m_+(z)m_-(z) \end{pmatrix};$$

its Herglotz representation gives us a matrix valued measure  $\rho$ , and the operator is then unitarily equivalent to multiplication by the variable in  $L^2(\mathbb{R}, d\rho)$ . See, for example, [27, Section 2.5] for the Jacobi case.

In our general situation, we don't have an operator (construction of these is actually more involved for canonical systems, and we don't really need them here), so we take this as our definition of the spectral properties. In particular, the claim of Theorem 2.4 that  $H_1, H_2$  are unitarily equivalent really means that the operators of multiplication by the variable in  $L^2(\mathbb{R}, d\rho_j)$  are unitarily equivalent.

We write  $\mathcal{R}(A)$  for the set of Jacobi matrices that are reflectionless on  $A$ , and then, for a non-empty compact set  $K \subseteq \mathbb{R}$ , we define

$$(2.12) \quad \mathcal{R}_0(K) = \{J \in \mathcal{J} : J \in \mathcal{R}(K), \sigma(J) \subseteq K\}.$$

We can make the same definitions for canonical systems, and we will write  $\mathcal{R}^C(A)$  and  $\mathcal{R}_0^C(K)$  for these larger spaces.

$\mathcal{R}_0(K)$  is a compact subset of  $\mathcal{J}$  if  $K$  has positive Lebesgue measure; if  $K$  is essentially closed, then  $\sigma(J) = K$  for all  $J \in \mathcal{R}_0(K)$ . If  $K$  is a finite gap set (a disjoint union of finitely many closed intervals of positive finite length), then  $\mathcal{R}_0(K)$  is the set of finite gap Jacobi matrices with spectrum  $K$  and, in particular, is homeomorphic to a torus of dimension equal to the number of (bounded) gaps of  $K$ .

Theorem 2.4 implies that, for arbitrary compact  $K$ , the space  $\mathcal{R}_0(K)$  is invariant under maps induced by a  $T(z) \in \mathcal{SL}$ , provided that this maps Jacobi matrices to Jacobi matrices again. In particular, the  $\mathcal{R}_0(K)$  are invariant under all flows from the Toda hierarchy. The collection of these sets is dense in  $\mathcal{J}$ : in fact, the Jacobi matrices with periodic coefficients are already dense, and such a  $J$  lies in  $\mathcal{R}_0(K)$  with  $K = \sigma(J)$ . Therefore, what happens to a general initial point under a Toda flow is determined by what happens on the  $\mathcal{R}_0(K)$ . As we will see in a moment, this leads to a very easy and transparent explanation of the fact that any two flows from the Toda hierarchy commute with each other.

It is also natural to ask what the fixed points of a Toda map are. In the special case of a Toda flow, it is well known that these are exactly the  $J \in \mathcal{R}_0(K)$  for a finite gap set  $K$  (where  $K$  depends on the flow under consideration). The general case has essentially the same answer:

**Theorem 2.5.** *Suppose that  $\pm m_{\pm}(z) = T(z)(\pm m_{\pm}(z))$  for a  $T \in \mathcal{SL}$ ,  $T \not\equiv \pm 1$ . Then  $H \in \mathcal{R}_0^C(E)$ , and here  $E$  is a union of disjoint closed intervals (possibly unbounded or consisting of single points) whose endpoints do not accumulate anywhere. More specifically,*

$$E = \{x \in \mathbb{R} : -2 \leq \operatorname{tr} T(x) \leq 2\}.$$

Such an  $H$  automatically has empty singular continuous spectrum, the absolutely continuous spectrum is essentially supported by  $E$  and of multiplicity 2 locally, and point spectrum is only possible at the isolated points of  $E$ , if any. These statements follow from the fact that  $H \in \mathcal{R}_0^C(E)$ , if we use general results about reflectionless Herglotz functions; see [20] for more on this. However, they are also very easy to verify directly here, and we will do this in Section 4, when we prove Theorem 2.5.

We can now give very transparent explanations of the well known basic properties of Toda flows: By combining Theorem 2.3 with Theorem 2.4, we see that  $p \cdot J$  is unitarily equivalent to  $J$  and the absolute value of the reflection coefficient is preserved. In particular,  $p \cdot J \in \mathcal{R}(A)$  precisely if  $J \in \mathcal{R}(A)$  (a different and much more technical proof of this fact was earlier given in [21], and see also [26] in this context). So the sets  $\mathcal{R}_0(K)$  are invariant under the action of  $G = \mathcal{P} \times \mathbb{Z}$ . The fixed points are finite gap Jacobi matrices by Theorem 2.5. Finally, by their construction (if the Toda hierarchy is indeed constructed following the suggestions above and as outlined more explicitly in Section 3), Toda flows commute with the shift.

What is missing from this list of basic properties is the fact that any two Toda flows also commute with each other. This would in principle follow from corresponding properties of the matrices  $B = B(z, J)$  from the cocycle equations. However, a much more intuitive and less technical explanation is also possible. This depends on the following easy general fact from topological dynamics.

**Lemma 2.6.** *Let  $S$  be a homeomorphism on a compact metric space  $X$ , and suppose that  $(X, S)$  is an equicontinuous dynamical system with a dense orbit. Then any two continuous maps on  $X$  that commute with  $S$  also commute with each other.*

The equicontinuity assumption refers to the family of maps  $S^n, n \in \mathbb{Z}$ .

*Proof.* Denote the two maps by  $F$  and  $G$ , respectively. Suppose that  $\{S^n x\}$  is dense in  $X$ , and let  $y \in X$  be an arbitrary point. Then  $y = \lim S^{k_j} x$  for a suitable sequence  $k_j \in \mathbb{Z}$ , and similarly  $Fy = \lim S^{m_j} x$ ,  $Gy = \lim S^{n_j} x$ . Then

$$Gy = G \lim S^{k_j} x = \lim GS^{k_j} x = \lim_{j \rightarrow \infty} S^{k_j} \lim_{p \rightarrow \infty} S^{n_p} x,$$

and now the equicontinuity of  $S^{k_j}$  implies that also  $Gy = \lim S^{k_j+n_j} x$ . By repeating this argument, we find that  $FGy = \lim S^{k_j+m_j+n_j} x$ , and then that  $GFGy$  equals the same expression.  $\square$

**Corollary 2.7.** *Any two Toda flows commute.*

*Sketch of proof.* We apply the Lemma to  $X = \mathcal{R}_0(K)$ , with  $K$  being a finite gap set. Then  $(X, S)$  is equicontinuous; in fact, it is well known that  $(X, S)$  can be conjugated to become a translation  $\tau_a$  on a finite-dimensional torus [27, Section 9.1]. The additional assumption of Lemma 2.6, on dense orbits, will be satisfied if  $a$  has suitable properties. The claim of the Corollary now follows from Theorem 2.1 and the fact that the collection of these  $\mathcal{R}_0(K)$  is still dense in  $\mathcal{J}$ .  $\square$

### 3. RECONSTRUCTION OF THE TODA HIERARCHY FROM COCYCLES

Before we turn to the topic announced in the title of this section, let me give the proofs of Theorems 2.1, 2.3.

*Proof of Theorem 2.1.* For a given  $p \in \mathcal{P}$ , write  $X(J) = [p(J)_a, J]$  for the right-hand side of (2.3). This is Lipschitz continuous with respect to the operator norm and thus the standard Picard iteration technique may be employed (see [27, Section 12.2] for more details): the solution  $J(t)$  of (2.3) on  $0 \leq t \leq T$  with the initial value  $J$  may be obtained as  $J(t) = \lim J_n(t)$ , with

$$(3.1) \quad J_0(t) = J, \quad J_{n+1}(t) = J + \int_0^t X(J_n(s)) ds,$$

and the convergence is in operator norm, and it is uniform in  $t \in [0, T]$ . The length  $T$  of this interval only depends on the Lipschitz constant.

As a by-product, one also obtains the continuity of  $J \mapsto p \cdot J$  with respect to the operator norm from this method, but we would like to use  $d$  instead of  $\|\cdot\|$ , so it's now natural to wonder if we also have a Lipschitz condition with respect to  $d$ , and indeed this works: we have that

$$(3.2) \quad d(X(J_1), X(J_2)) \leq Ld(J_1, J_2)$$

for some  $L$  that only depends on the polynomial  $p$  and the bound  $R$  on  $\|J\|$ . To see this, observe that the matrix  $p(J)_a$  has finite band width in the sense that  $\langle \delta_j, p(J)_a \delta_k \rangle = 0$  for  $|j - k| > \deg p$ . So when we compute a matrix element  $X(J)_{nk}$ , then only those coefficients  $a_j, b_j$  of  $J$  at at most a certain distance from  $n$  or  $k$  are involved, and of course  $X(J)_{nk}$  is a smooth function of those coefficients, with uniformly bounded derivatives on  $J \in \mathcal{J}_R$ . From this it follows that there are constants  $C$  and  $D$ , depending only on  $R$  and  $p$ , such that

$$(3.3) \quad |X(J)_{nk} - X(J')_{nk}| \leq C \sum_{|j-n| \leq D} (|a_j - a'_j| + |b_j - b'_j|);$$

also recall that only  $k = n, n \pm 1$  gives non-zero matrix elements here. By multiplying (3.3) by  $2^{-|n|}$  and summing over  $n, k$ , we now indeed obtain (3.2).

The rest is routine: the argument alluded to above can simply be repeated with (3.2) as the key ingredient. To spell this out more explicitly, consider two initial values  $J, J'$ , and observe that the metric is of the form  $d(J_1, J_2) = \|J_1 - J_2\|_w$ , for a certain weighted  $\ell^1$  norm  $\|\cdot\|_w$ . Thus an inductive argument using (3.1) shows that

$$d(J_n(t), J'_n(t)) \leq d(J, J') \left( 1 + Lt + \dots + \frac{(Lt)^n}{n!} \right),$$

so  $d(J(t), J'(t)) \leq e^{Lt}d(J, J')$ .  $\square$

*Proof of Theorem 2.3.* I'll discuss the claim about  $m_+$ ; of course, a similar argument will work for  $m_-$ . Recall that  $M = m_+(z)$  may be characterized as the unique number that makes (the components of)  $T(n; J)(M, 1)^t$  square summable over  $n \geq 1$ . Moreover, a solution vector  $F_n = (f_{n+1}, -a_n f_n)^t$ , with  $\tau f = zf$ , will be square summable already if  $\liminf \|F_n\| = 0$ . This follows because the Wronskian  $F_n^t J G_n$ ,  $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ , of two solutions  $F, G$  is constant and the  $\ell^2$  solution goes to zero, so all other solutions must become large.

Consider now the solution

$$F_n = T(n; g \cdot J)T(g; J) \begin{pmatrix} M \\ 1 \end{pmatrix} = T(g; n \cdot J)T(n; J) \begin{pmatrix} M \\ 1 \end{pmatrix}.$$

Since, by assumption,  $T(g; n \cdot J)$  stays bounded on a subsequence  $n_k \rightarrow \infty$ , we see that  $\liminf \|F_n\| = 0$ . By our preliminary remarks, this identifies  $F_n$  as the  $\ell^2$  solution of the Jacobi matrix  $g \cdot J$ . Hence its  $m$  function  $m(g \cdot J)$  is the number represented by  $T(g; J)(M, 1)^t$ , but this is  $T(g; J)M$ , as claimed.  $\square$

We now indicate how the Toda hierarchy could be constructed starting from the requirement that we wish our evolutions to have an associated  $\mathcal{SL}$ -cocycle when combined with the shift, rather than from the Lax equation (2.3), as is usually done. We do not make rigorous claims in this part and will freely use formal calculations.

We consider a single flow, unknown at this point, which we'll denote by  $t \cdot J$  (in our notation from above, this would become  $tp \cdot J$ , for a fixed polynomial  $p$  and  $t \in \mathbb{R}$ ). So the acting group now is  $G = \mathbb{R} \times \mathbb{Z}$ , with  $n \cdot J$  acting by shifts and  $t \cdot J$  is what we're trying to construct. More careful notation for these group elements would have been  $(0, n)$  and  $(t, 0)$ , respectively, but our abbreviated version is more pleasing to look at, and we in fact already used similar conventions earlier.

Following our general plan, we now make the crucial (and strong) additional assumption that there is an associated  $\mathcal{SL}$ -cocycle  $T = T(t, n; J)$  whose shift part is given by (2.5). As in (2.7), the  $t$  part is described by a matrix function  $B = B(z, J)$ ,  $\text{tr } B = 0$ . We differentiate both sides of the identity

$$(3.4) \quad T(t; 1 \cdot J)T(1; J) = T(1; t \cdot J)T(t; J)$$

with respect to  $t$  at  $t = 0$ . Obviously, (3.4) is a special case of the cocycle identity (2.4); however, it is also true and easily verified that if we have individual cocycles  $T_0(t; J) = T(t, 0; J)$  and  $S(n; J) = T(0, n; J)$ , for the actions of  $\mathbb{R}$  and  $\mathbb{Z}$ , respectively, and the compatibility condition (3.4) holds, then  $T(t, n; J) := T_0(t; n \cdot J)S(n; J)$  defines a cocycle for the action of  $\mathbb{R} \times \mathbb{Z}$ .

Next, recall that  $T(1; J) = A(J)$ , with the  $A$  from (2.5), so, formally at least, we obtain that

$$(3.5) \quad B(1 \cdot J)A(J) = \dot{A}(J) + A(J)B(J),$$

where  $\dot{A}(J)$  is short-hand for  $(d/dt)A(t \cdot J)|_{t=0}$ . This equation is known as the *zero curvature equation*. Our derivation of it here gives it a rather transparent interpretation: It is the cocycle property for  $T(g; J)$  in differential form. More precisely, it is a compatibility condition that will ensure that the individual cocycles form a joint cocycle when glued together. Please see also Section 3 and especially Theorem 3.1 of [18], where these remarks are made more precise.

To obtain the Toda hierarchy from (3.5), we make the additional assumption that  $B(z, J) = \sum_{n=0}^N z^n B_n(J)$  is a polynomial in  $z$ . Since  $A(z, J) = A_0(J) + zA_1(J)$  also is of this type, we can then compare coefficients in (3.5).

Let me take a quick look at the cases  $N = 0$  and  $N = 1$ , as an illustration, without attempting to give a general treatment. Such a general treatment, however, is possible and was given by Ong [17].

If  $N = 0$ , then comparing coefficients of  $z^1$  in (3.5) shows that  $B_{12} = B_{21} = 0$ . Since  $\text{tr } B = 0$ , this means that  $B$  would have to be of the form  $B(J) = \alpha(J) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ , but now we can already see that only  $\alpha \equiv 0$  works here since this  $B$  leads to

$$T(t; J) = \begin{pmatrix} e^{\omega(t; J)} & 0 \\ 0 & e^{-\omega(t; J)} \end{pmatrix}, \quad \omega(t; J) \equiv \int_0^t \alpha(s \cdot J) ds,$$

so  $T(t)m = e^{2\omega}m$ , but half line Jacobi  $m$  functions satisfy  $m_+(z) = -1/z + O(z^{-2})$  for large  $z$ , so Theorem 2.3 now shows that we must have  $\omega = 0$ . The same conclusion could have been obtained by looking at (3.5) more closely. So we are not getting a non-trivial flow for  $N = 0$ .

Moving on to the case  $N = 1$  then, we again start out by comparing coefficients for the highest power  $z^2$ . This shows that  $B_1(1 \cdot J)A_1(J) - A_1(J)B_1(J) = 0$ , and since  $A_1 = (1/a_1) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ , it follows that

$$B_1(J) = \alpha_1 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

and here  $\alpha_1(1 \cdot J) = \alpha_1(J)$ . Since there are Jacobi matrices  $J$  whose orbits under the shift map are dense, this forces us to take  $\alpha_1$  as a constant, independent of  $J$ , at least if we want a continuous  $B_1(J)$ . In fact, since multiplying  $B$  by a constant amounts to rescaling time, we might as well set  $\alpha_1 = 1$ .

By comparing the coefficients of  $z$ , we obtain that

$$(3.6) \quad \dot{A}_1 = B_1 A_0(J) - A_0(J) B_1 + B_0(1 \cdot J) A_1(J) - A_1(J) B_0(J).$$

Since only the  $(1, 1)$  entry of  $A_1$  is non-zero, this also gives restrictions on the possible choices for  $B_0$ . More specifically, we find that

$$(3.7) \quad B_0(J) = \begin{pmatrix} \alpha_0(J) & 2 \\ -2a_0^2 & -\alpha_0(J) \end{pmatrix}.$$

Finally, we compare the coefficients of  $z^0$ ; this gives that

$$(3.8) \quad \dot{A}_0 = B_0(1 \cdot J) A_0(J) - A_0(J) B_0(J).$$

We work out the matrix elements, using the  $B_0$  from (3.7). From the  $(1, 2)$  element, we find that  $\dot{a}_1/a_1 = -2b_1 - \alpha_0(J) - \alpha_0(1 \cdot J)$ . On the other hand, comparing the  $(1, 1)$  elements of (3.6) yields  $\dot{a}_1/a_1 = \alpha_0(J) - \alpha_0(1 \cdot J)$ . So we must take  $\alpha_0(J) = -b_1$ . With these choices in place, (3.8) will now produce the familiar classical Toda equations

$$\dot{a}_n/a_n = b_{n+1} - b_n, \quad \dot{b}_n = 2(a_n^2 - a_{n-1}^2),$$

initially for  $n = 1$ , but then we also obtain the general case by considering  $n \cdot J$  instead of  $J$ . Moreover, and more importantly still for us perhaps, we have confirmed that  $B(J)$  for the Toda flow is indeed given by (2.9).

As I mentioned above, Ong [17] has carried out this whole analysis in a systematic fashion, and he proves that the whole Toda hierarchy can be reconstructed in this way.

#### 4. PROOF OF THEOREMS 2.4, 2.5

*Proof of Theorem 2.4.* I'll use the following notations in this proof:  $T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ , and I'll write  $m_{\pm}$  and  $M_{\pm}$  for  $m_{\pm}^{(1)}$  and  $m_{\pm}^{(2)}$ , respectively, and similarly for other quantities, to be introduced in a moment. Let  $h = m_+ + m_-$ ,  $g_0 = -1/h$ ,  $g_1 = m_+ m_- / h$ . Notice that  $g_0, g_1$  are the

diagonal elements of  $M$  from (2.11), so  $\rho = \rho_0 + \rho_1$  may be used as a spectral measure, where  $\rho_j$  denotes the measure associated with the Herglotz function  $g_j$ .

We thus want to show that  $\rho$  and  $P$  (“capital rho;” recall our general convention of using uppercase letters for the transformed quantities) are equivalent measures, and we also need to pay attention to possible spectral multiplicity, but this latter part will be easy since only the absolutely continuous part can have multiplicity greater than one.

By writing out the linear fractional transformation by which  $T$  acts, we obtain that

$$(4.1) \quad M_+ = \frac{am_+ + b}{cm_+ + d}, \quad M_- = -\frac{am_- - b}{cm_- - d},$$

and then that

$$(4.2) \quad G_0 = -(cm_+ + d)(cm_- - d)g_0, \quad G_1 = -(am_+ + b)(am_- - b)g_0.$$

We’ll treat the different parts of the spectrum separately, and we start with the absolutely continuous parts. These are easy to analyze: The corresponding half line spectral measures are given by  $d\rho_{\pm,ac}(t) = (1/\pi)\text{Im } m_{\pm}(t) dt$ , and we may look at the operator of multiplication by the variable in  $L^2(\mathbb{R}, d\rho_{+,ac}) \oplus L^2(\mathbb{R}, d\rho_{-,ac})$ ; the fact that such a direct reduction to the two half lines works is sometimes referred to as the *decomposition method*. Since  $a, b, c, d$  are entire functions, (4.1) will imply, after a quick calculation, that each of the two sets

$$\Sigma_{ac}(m_{\pm}) := \{t \in \mathbb{R} : \text{Im } m_{\pm}(t) > 0\}$$

differs from its counterpart for  $M_{\pm}$  by at most a set of Lebesgue measure zero. This proves the claim about the absolutely continuous parts, including multiplicity.

Since only the absolutely continuous part can have multiplicity greater than one, our discussion of the singular parts can focus on showing that  $\rho_s$  and  $P_s$  are equivalent measures. Now  $\rho_{0,s}$  is supported by the set where  $|g_0(t)| \equiv \lim_{y \rightarrow 0^+} |g_0(t + iy)| = \infty$ . Since  $m_{\pm}$  are Herglotz functions, this is equivalent to the three conditions

$$(4.3) \quad \text{Im } m_{\pm}(t) = 0, \quad \lim_{y \rightarrow 0^+} \text{Re } (m_+(t + iy) + m_-(t + iy)) = 0.$$

Now we are going to use Poltoratski’s Theorem [19] (see also [14]) on the comparison of the singular parts. This says that for  $\rho_{0,s}$ -almost every  $t$ , the limit

$$\lim_{y \rightarrow 0^+} \frac{G_0(t + iy)}{g_0(t + iy)} \equiv L(t)$$



exists, and  $L(t) = dP_0(t)/d\rho_{0,s}$ , the Radon-Nikodym derivative of that part of  $P_0$  (or  $P_{0,s}$ ) that is absolutely continuous with respect to  $\rho_{0,s}$ .

I now want to know how frequently  $L$  can be zero, to control how much of the measure  $\rho_{0,s}$  could potentially be lost after the transformation. We may then restrict our attention to those  $t$ 's that satisfy (4.3) to compute  $L$  because  $\rho_{0,s}$  gives zero weight to the set where these conditions fail. We now distinguish two cases:

(a)  $c(t) = 0$ : Then  $c(t+iy) = O(y)$ , and a Herglotz function  $F$  always satisfies  $\lim_{y \rightarrow 0^+} y \operatorname{Re} F(t+iy) = 0$ . Thus also  $\lim_{y \rightarrow 0^+} c(t+iy)m_+(t+iy) = 0$ , since we are currently only considering  $t$ 's with  $\operatorname{Im} m_+(t) = 0$ . Hence  $L(t) = d^2(t)$  and  $c, d$  cannot both be zero since  $\det T = 1$ , so we never have  $L = 0$  in this case.

(b)  $c(t) \neq 0$ : Recall that we are currently considering only  $t$ 's satisfying (4.3). Now if also  $L(t) = 0$  for such a  $t$ , then  $m_{\pm}(t) \equiv \lim_{y \rightarrow 0^+} m_{\pm}(t+iy)$  both exist and  $m_-(t) = -m_+(t) = d(t)/c(t)$ . It follows that

$$\frac{G_1(t+iy)}{g_0(t+iy)} = -(am_+ + b)(am_- - b) \rightarrow \left( -\frac{a(t)d(t)}{c(t)} + b(t) \right)^2 = \frac{1}{c^2(t)}.$$

Thus Poltoratski's theorem shows that if  $L(t) = 0$  on a set of positive  $\rho_{0,s}$ -measure, then  $dP_{1,s}/d\rho_{0,s} > 0$  for  $\rho_{0,s}$ -almost every such  $t$ .

Putting things together, we thus see that  $\rho_{0,s} \ll (P_0 + P_1)_s$ ; indeed, the argument we gave showed that when the density  $L(t)$  in the decomposition  $dP_{0,s} = Ld\rho_{0,s} + d\nu$  fails to be positive, then  $P_{1,s}$  comes to the rescue.

A similar analysis, with  $-1/m_{\pm}$  now taking over the roles of  $m_{\pm}$ , works for  $\rho_{1,s}$ , and thus it also follows that  $\rho_s \ll P_s$ . By symmetry, since  $T^{-1} \in \mathcal{SL}$  also, this then gives that  $P_s \ll \rho_s$  as well, so  $\rho_s$  and  $P_s$  are equivalent measures. This concludes the proof of the claim that  $H_1, H_2$  are unitarily equivalent.

It remains to show that  $|R_1| = |R_2|$  almost everywhere, but this is immediate from a calculation. Or, in flashier style, one can observe that

$$|R(z)| = \left| \frac{m_+(z) + \overline{m_-(z)}}{m_+(z) + m_-(z)} \right| = \tanh \left( \frac{1}{2} \gamma(m_+(z), -\overline{m_-(z)}) \right)$$

depends only on the hyperbolic distance  $\gamma$  of the numbers  $m_+(z), -\overline{m_-(z)} \in \mathbb{C}^+$ , which is preserved by the automorphism  $T(x) \in \operatorname{SL}(2, \mathbb{R})$ . Strictly speaking, this argument works only if  $m_+(x), -\overline{m_-(x)}$  actually lie in  $\mathbb{C}^+$ , but if at least one of them is real, then  $|R(x)| = 1$ , and

we have the same situation for the transformed operator by the unitary equivalence of the absolutely continuous parts, which we already proved.  $\square$

*Proof of Theorem 2.5.* As in the previous proof, we'll write  $T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ . Then, by writing out the assumption that  $\pm m_{\pm} = T(\pm m_{\pm})$ , we see that  $m_+(z)$  and  $-m_-(z)$  both solve the quadratic equation

$$(4.4) \quad c(z)x^2 + (d(z) - a(z))x - b(z) = 0.$$

Its solutions are

$$x = \frac{a-d}{2c} \pm \frac{1}{2c} \sqrt{D^2 - 4}, \quad D \equiv a+d,$$

at least if  $c(z) \neq 0$ . Observe that we cannot have  $c \equiv 0$  here: if  $z \in \mathbb{C}^+$ , then  $m_+(z) \neq -m_-(z)$ , so (4.4) must have two distinct solutions, and if  $c = 0$ , then this forces  $a = d = \pm 1$ ,  $b = 0$ , so  $T = \pm 1$ , but we explicitly assumed that  $T$  is not identically equal to the identity matrix or its negative.

Recall from the previous proof how we can extract the spectral properties of the whole line operator from  $m_{\pm}$ : we let  $h = m_+ + m_-$ ,  $g_0 = -1/h$ ,  $g_1 = m_+m_-/h$ , and then  $\rho = \rho_0 + \rho_1$  is a spectral measure of maximal type, where the measures  $\rho_0, \rho_1$  come from the Herglotz functions  $g_0, g_1$ .

In fact, this is needed only to study the singular part; the absolutely continuous part can be read off directly by identifying the sets where  $\text{Im} m_{\pm}(t) > 0$ . Here this happens for  $t \in \mathbb{R}$  precisely if  $|D(t)| < 2$ , so we see that the absolutely continuous spectrum is of (uniform) multiplicity 2 and is essentially supported by

$$\Sigma_{ac} = \sigma_{ac} = E = \{t \in \mathbb{R} : |D(t)| \leq 2\}.$$

We also obtain that  $m_+(t) = -\overline{m_-(t)}$  on this set, so  $H \in \mathcal{R}^C(E)$ , as claimed.

To complete the proof of the theorem, we would just have to show that there is no spectrum outside  $E$ ; we will instead, as promised, give a complete treatment of the spectral properties, since this is easy to do.

It is already clear, from what we reviewed above, that there is no singular continuous spectrum; it remains to prove that only the isolated points of  $E$  are possible eigenvalues of  $H$ . A calculation shows that

$$(4.5) \quad g_0 = -\frac{c}{\sqrt{D^2 - 4}}, \quad g_1 = \frac{b}{\sqrt{D^2 - 4}},$$

and eigenvalues occur when

$$(4.6) \quad \lim_{y \rightarrow 0^+} -iyg_j(t + iy) > 0$$

for at least one of  $j = 0, 1$ . This can happen only when  $D(t) = \pm 2$ , and let's consider the case where  $D(t) = 2$  and (4.6) holds for  $j = 0$  at this  $t$ . Use the Taylor expansions

$$c(t + iy) = c_k(iy)^k + O(y^{k+1}), \quad D(t + iy) = 2 + D_m(iy)^m + O(y^{m+1}),$$

with  $c_k, D_m \neq 0$  in (4.5). It follows that  $m = 2n \geq 2$  must be even, and since  $c_k, D_m \in \mathbb{R}$ , we also see that  $D_m > 0$ , so indeed

$$D(t + h) = 2 + D_m h^{2n} + O(h^{2n+1}) > 2$$

for  $x = t + h \in \mathbb{R}$  close to  $t$ , so  $t$  is an isolated point of  $E$ , as claimed. The other cases are similar.  $\square$

## 5. PROOF OF THEOREM 2.2

Before we can prove Theorem 2.2, we of course need to give the precise definition of the cocycle: The corresponding matrix function  $B$  is given by

$$(5.1) \quad B(J) = \begin{pmatrix} 2(z - b_1)G_1 - H_1 & 2G_1 \\ -2a_0^2 G_0 & -2(z - b_1)G_1 + H_1 \end{pmatrix}.$$

This formula needs some explanation. Also, recall that we have one such  $B = B_p$  for each polynomial  $p \in \mathcal{P}$ , which, as usual, is then used to produce  $T(p; J)$  by solving

$$\dot{T} = B_p(tp \cdot J)T, \quad T(0) = 1,$$

and evaluating at  $t = 1$ .

For a given polynomial  $p \in \mathcal{P}$ , the functions  $G, H$  from (5.1) depend on  $J \in \mathcal{J}$ ,  $n \in \mathbb{Z}$ , and  $z \in \mathbb{C}$ , and they are polynomials in  $z$ ; the indices in (5.1) refer to the  $n$  variable, so for example  $G_1 = G(n = 1)$ . The precise definitions are as follows: Suppose first of all that  $p(x) = x^k$ , and attach a superscript  $k$  to the corresponding functions  $G, H$ , as a reminder to ourselves that this is the  $p$  we have currently chosen. Then

$$G^{(k)}(z, J) = \sum_{j=0}^{k-1} z^{k-1-j} (J^j)_d,$$

$$H^{(k)}(z, J) = z^k - (J^k)_d + 2a \sum_{j=1}^{k-1} z^{k-1-j} (S J^j)_d;$$

here  $X_d$  refers to the diagonal part of a (let's say: bounded) operator, thought of as an infinite matrix. In other words,  $(X_d)_n = \langle \delta_n, X \delta_n \rangle$ . As

before,  $S$  denotes the shift, so  $(Sx)_n = x_{n+1}$  for  $x \in \ell^2$ , and  $a$  refers to the sequence  $(a_n)$ . So the dependence on  $n$  has not been made explicit in the notation; it is contained in the various diagonal parts and also in  $a = a_n$ .

We have now defined the functions  $G, H$  for  $p(x) = x^k$ ; for general  $p(x) = \sum c_k x^k$ , we put  $G = G_p = \sum c_k G^{(k)}$ , and similarly for  $H = H_p$ . In other words, we make  $G, H$  linear functions of  $p$ .

If  $p(x) = x$ , then we obtain that  $G = 1$ ,  $H = z - J_d = z - b$ , and thus we recover the  $B$  from (2.9) from (5.1).

We are now ready for the

*Proof of Theorem 2.2.* I will only discuss the cocycle identity for evolution along two distinct flows from the Toda hierarchy. The theorem also claims that the same property holds if we consider a flow together with the shift. This follows from the zero curvature equation, and the relevant calculations (though not the statement itself, at least not very explicitly) can be found in the standard literature, for example in [27, Section 12.2] or [9, Section 1.3]. The argument is presented in detail in the appendix of [18].

So we'll show that

$$(5.2) \quad T(p+q; J) = T(p; q \cdot J)T(q; J) \quad \text{for } p, q \in \mathcal{P}.$$

Rather than deal with this  $\mathcal{SL}$ -cocycle directly, we can reduce matters to the same question about a  $\mathbb{C}^\times$  valued cocycle, as follows. Introduce, for fixed  $z \in \mathbb{C}^+$ ,

$$V(J) = \frac{1}{\sqrt{m_+ + m_-}} \begin{pmatrix} m_+ & -m_- \\ 1 & 1 \end{pmatrix};$$

the choice of square root can be conveniently settled by requiring that  $\sqrt{m_+ + m_-} \in \mathbb{C}^+$ . Next, write

$$(5.3) \quad T(p; J) = V(p \cdot J)D(p; J)V(J)^{-1}, \quad D = \begin{pmatrix} \lambda(p; J) & 0 \\ 0 & \lambda(p; J)^{-1} \end{pmatrix}.$$

(This kind of transformation is usually described by referring to  $T$  and  $D$  as *cohomologous* cocycles; note, however, that  $T \in \mathcal{SL}$  while  $V$  is not an entire function of  $z$ . So the transformation, while convenient here, certainly has its drawbacks from a general point of view.) To interpret (5.3), let's first of all observe that  $V$  as a linear fractional transformation maps the standard unit vectors  $e_1, e_2$  (that is,  $\infty$  and  $0$ , as points on the Riemann sphere) to  $\pm m_\pm$ . This means that (5.3) simply records the general form of an  $\mathrm{SL}(2, \mathbb{C})$  matrix that updates  $\pm m_\pm(J)$  correctly to their new values  $\pm m_\pm(p \cdot J)$ . Next, the cocycle

property (5.2) for  $T$  is now equivalent to the same property for  $D$ , and since  $D$  is diagonal, (5.2) indeed simplifies to

$$(5.4) \quad \lambda(p+q; J) = \lambda(p; q \cdot J)\lambda(q; J),$$

and this is what we'll now establish.

Since we already have the cocycle property of  $D$  and  $\lambda$  with respect to the action of  $G = \mathbb{R}$ , for just a fixed individual flow, we know that, in analogy to (2.7), we can obtain  $\lambda$  by solving an ODE of the form

$$(5.5) \quad \dot{\lambda} = \omega_p(tp \cdot J)\lambda, \quad \lambda(0) = 1;$$

as usual,  $\lambda(p; J)$  then is the solution to this initial value problem, evaluated at  $t = 1$ .

We can find  $\omega = \omega_p(J)$  by comparing this with (2.7), (5.1). To do this, differentiate  $T = T(tp; J)$  with respect to  $t$  at  $t = 0$  to produce  $B_p(J)$ . Obviously, these manipulations assume that the various evolving quantities depend smoothly on  $t$ ; a few quick remarks about this issue can be found below, at the beginning of the proof of Proposition 5.1.

If the representation from the right-hand side of (5.3) is used in this calculation, then we obtain an alternative formula that relates  $B$  to  $\omega_p(J) = (d/dt)\lambda(tp; J)|_{t=0}$ . It is especially convenient to focus on the  $(2, 1)$  entry of  $B$  because the only contribution to this when differentiating (5.3) comes from  $V(J)\dot{D}V(J)^{-1}$ . In this way, we see that

$$(5.6) \quad \omega_p(J) = -a_0^2(m_+ + m_-)G_0 = (G/g)_0,$$

where  $g_0 = -1/(a_0^2(m_+ + m_-))$ . This is also the matrix element of the resolvent

$$(5.7) \quad g_0(z) = \langle \delta_0, (J - z)^{-1}\delta_0 \rangle,$$

which we already used in Section 4 (with the factor  $a_0^2$  removed). So alternatively, we could define  $g = (J - z)_d^{-1}$  and then interpret  $g_0$  exactly as above as the  $n = 0$  element of this sequence.

The choice of notation here was deliberate: the polynomial  $G$  is a truncated version of the Taylor series of  $g$  about  $z = \infty$ , plus a trivial shift of exponent. So the two functions  $g, G$  are closely related. There is an analog  $h$  of  $g$ , which bears the same relation to  $H$  as  $g$  does to  $G$  and will become important later. It is defined by

$$(5.8) \quad h(z) = (2aS(J - z)^{-1} - 1)_d;$$

here,  $S$  again denotes the shift operator  $(Sx)_n = x_{n+1}$ , and  $a$  must now be interpreted as the operator of multiplication by the sequence  $a_n$ .

We will now establish:

**Proposition 5.1.** *For fixed  $z \in \mathbb{C}^+$  and any two polynomials  $p, q \in \mathcal{P}$ ,  $\omega_p(tq \cdot J)$  is a smooth function of  $t \in \mathbb{R}$  and*

$$(5.9) \quad \frac{d}{dt} \omega_p(tq \cdot J)|_{t=0} = \frac{d}{dt} \omega_q(tp \cdot J)|_{t=0}.$$

Before proving this, let's discuss how we can obtain (5.4) (and thus finish the proof of Theorem 2.2) from Proposition 5.1. By applying this result to  $(sp + tq) \cdot J$  in place of  $J$ , we also obtain that

$$(5.10) \quad \frac{\partial}{\partial s} \omega_q((sp + tq) \cdot J) = \frac{\partial}{\partial t} \omega_p((sp + tq) \cdot J),$$

and then integration with respect to  $s$  from  $s = 0$  to  $s = t$  yields

$$(5.11) \quad \omega_q(t(p + q) \cdot J) - \omega_q(tq \cdot J) = \int_0^t \frac{\partial}{\partial t} \omega_p((sp + tq) \cdot J) ds.$$

We also have that  $\omega_{p+q} = \omega_p + \omega_q$ . To see this, recall how  $\omega$  was defined in (5.6): only  $G_0$  depends on  $p$  here, and we observed earlier that  $G$  indeed is a linear function of  $p$ . So we can rewrite (5.11) as

$$(5.12) \quad \begin{aligned} \omega_{p+q}(t(p + q) \cdot J) &= \omega_p(t(p + q) \cdot J) + \omega_q(tq \cdot J) \\ &\quad + \int_0^t \frac{\partial}{\partial t} \omega_p((sp + tq) \cdot J) ds. \end{aligned}$$

Next, integrate this from  $t = 0$  to  $t = 1$  and use Fubini-Tonelli on the right-hand side, which is justified since  $(sq + tp) \cdot J$  stays inside a compact set for  $s, t \in [0, 1]$ , and this implies that the integrand is bounded. We do need further information about the evolution under Toda flows for this step, which can be found in [27]; see also the discussion below, at the beginning of the proof of Proposition 5.1. Integration of the last term of (5.12) then produces

$$\begin{aligned} &\int_0^1 ds \int_s^1 dt \frac{\partial}{\partial t} \omega_p((sp + tq) \cdot J) = \\ &\int_0^1 (\omega_p(sp \cdot (q \cdot J)) - \omega_p(s(p + q) \cdot J)) ds \end{aligned}$$

and here the last term conveniently cancels the integral of the first term from the right-hand side of (5.12). So if we also recall (5.5) and exponentiate the integrated version of (5.12), then we indeed arrive at (5.4), as desired. This completes the proof of Theorem 2.2, assuming Proposition 5.1.  $\square$

*Proof of Proposition 5.1.* As for the smoothness claims, we essentially refer the reader to [27] and limit ourselves to a few remarks. See especially Theorem 12.6 there, but also Lemma 12.15. As a general

strategy, once the smoothness of (the coefficients of)  $tp \cdot J$  is known, everything else (such as the smoothness of  $m_{\pm}(tp \cdot J)$ ,  $G_0(tp \cdot J)$ ) will pretty much just fall into place by making use of the explicit formulae we have; the perhaps most challenging contributions here are  $m_{\pm}$ , but these, too, can be handled without much trouble by recalling that they are evolved by the cocycle and then using the explicit formulae for  $B$ . I'll leave the matter at that, and I'll focus on (5.9) now.

First of all, observe that it suffices to prove this for monomials  $p(x) = x^p$ ,  $q(x) = x^q$ . Indeed, suppose we had this already, and consider general polynomials  $p = \sum p_j x^j$ ,  $q = \sum q_j x^j$ . Then

$$\begin{aligned} \omega_p(tq \cdot J) &= F(t, t, \dots, t), \\ F(t_0, \dots, t_N) &\equiv \omega_p(t_0 q_0 \cdot t_1 q_1 x \cdot \dots \cdot t_N q_N x^N \cdot J), \end{aligned}$$

and since the group acting is abelian, the monomials  $t_j q_j x^j$  can be reshuffled at will here. Thus, by the chain rule,

$$\frac{d}{dt} \omega_p(tq \cdot J)|_{t=0} = \sum \frac{\partial F}{\partial t_j}(0, 0, \dots, 0) = \sum \frac{d}{dt} \omega_p(tq_j x^j \cdot J)|_{t=0}.$$

Moreover,  $\omega_p = \sum p_j \omega_{x^j}$ , so indeed (5.9) for monomials gives the general case also.

So it now suffices to take  $p(x) = x^p$ ,  $q(x) = x^q$ , and let's also assume that  $p > q$ . We'll write  $G^{(p)}$ ,  $G^{(q)}$ ,  $H^{(p)}$ ,  $H^{(q)}$  to indicate which polynomial (or, rather, monomial) is being used, and we now adopt the convention that evaluation at  $n = 0$  is understood if no such index is given: for example  $G^{(p)}$  will refer to  $(G^{(p)})_0$ .

We will also need the time evolution of  $g = g(tq \cdot J)$ , which is given by

$$(5.13) \quad \dot{g} = 2g(\omega_q h - H^{(q)});$$

here, as always, the dot notation is defined as  $\dot{X}(J) = (d/dt)X(tq \cdot J)|_{t=0}$ , and  $h$  was defined in (5.8). To prove (5.13), it is easiest to work with the explicit kernel of the Green function  $g$ ; please see [27, Section 12.4] for the details.

It now follows that

$$(5.14) \quad \frac{d}{dt} \omega_p(tq \cdot J)|_{t=0} = \frac{1}{g} \left( \dot{G}^{(p)} - 2\omega_q h G^{(p)} + 2H^{(q)} G^{(p)} \right).$$

To analyze this further, we expand  $g, h$  about  $z = \infty$ :

$$(5.15) \quad g(z) = - \sum_{n=0}^{\infty} c_n z^{-n-1}, \quad h(z) = - \sum_{n=-1}^{\infty} d_n z^{-n-1},$$

and here

$$\begin{aligned} c_0 &= 1, \quad c_n = \langle \delta_0, J^n \delta_0 \rangle \quad (n \geq 1); \\ d_{-1} &= 1, \quad d_0 = 0, \quad d_n = \langle \delta_0, 2aS J^n \delta_0 \rangle \quad (n \geq 1). \end{aligned}$$

These formulae follow by expanding the definitions (5.7), (5.8), and the series converge at least for  $|z| > R$ , with  $R = \|J\|$ . Note that this quantity does not change under any Toda flow because the evolved operator is unitarily equivalent to  $J$ .

It is again clear from [27, Theorem 12.6] that  $c_n, d_n \in C^\infty$  as a function of  $t$  along any flow. Moreover, we also obtain uniform bounds of the form  $\dot{c}_n, \dot{d}_n \lesssim R^n$ , so the expansions from (5.15) may be differentiated with respect to  $t$  term by term, and thus  $\dot{g}, \dot{h}$  have expansions similar to the ones from (5.15). This will become important in a moment because we will establish the desired identity by expanding and then comparing coefficients. Note that strictly speaking this will only give the identity for  $|z| > R$ , but of course that is good enough since both sides are holomorphic functions of  $z$ .

As I already pointed out, the functions  $g, h$  are closely related to  $G, H$ , and this can now be made more explicit: we have that

$$G^{(D)}(z) = z^D \sum_{n=0}^{D-1} c_n z^{-n-1}, \quad H^{(D)}(z) = z^D \sum_{n=-1}^{D-1} d_n z^{-n-1} - c_D, \quad D = p, q.$$

In other words, we have that (for example)

$$G^{(p)}(z) = [-z^p g(z)]_+,$$

where the notation  $[X]_+$  instructs us to expand  $X = \sum_{n \in \mathbb{Z}} X_n z^n$  into a Laurent series about  $z_0 = 0$  and then keep only the power series part  $[X]_+ = \sum_{n \geq 0} X_n z^n$  (which, in this paper, will always be a polynomial). We now use these formulae to compute the time derivatives. So, recalling (5.6) and (5.13), we can now say that

$$(5.16) \quad \dot{G}^{(p)} = [-2z^p (G^{(q)} h - H^{(q)} g)]_+.$$

In this step, we make use of our preliminary observations about term-by-term differentiation of (5.15). If (5.16) is combined with (5.14), then what we are trying to show can be rephrased as the claim that with

$$F(p, q) \equiv [-z^p (G^{(q)} h - H^{(q)} g)]_+ - \omega_q h G^{(p)} + H^{(q)} G^{(p)},$$

we have the identity  $F(p, q) = F(q, p)$ .



This symmetry property is clear for the second term if we recall (5.6), so we may drop this from  $F$  and focus on

$$F_1(p, q) \equiv [H^{(q)}G^{(p)} - z^p(G^{(q)}h - H^{(q)}g)]_+.$$

We can now finish the argument by a straightforward (if tedious) brute force calculation. We expand everything in powers of  $z$ . After a calculation, we find that

$$F_1(p, q) = \left[ z^{p+q} \left( \sum_{j=0}^{q-1} c_j z^{-j-1} \sum_{k \geq -1} d_k z^{-k-1} - \sum_{j \geq p} c_j z^{-j-1} \sum_{k=-1}^{q-1} d_k z^{-k-1} \right) \right]_+.$$

In the last sum, we may also sum over *all*  $k \geq -1$  since we will not get contributions to the  $[\dots]_+$  part from the  $k \geq q$ , and if written in this form, then  $F_1$  is easily seen to have the symmetry property we require.  $\square$

## 6. TWISTED SHIFTS AND FLOWS FOR CANONICAL SYSTEMS

As pointed out earlier, though not exceedingly popular at the moment, *canonical systems*

$$(6.1) \quad Ju'(x) = -zH(x)u(x), \quad x \in \mathbb{R},$$

are a very natural object from a mathematical point of view because when we normalize them by the requirement that  $\text{tr } H = 1$ , then they are in one-to-one correspondence to pairs of Herglotz functions  $m_{\pm} : \mathbb{C}^+ \rightarrow \overline{\mathbb{C}^+}$ . Here  $J$  denotes the matrix  $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ , and the basic assumptions on the coefficient function  $H$  are as follows: it takes values in  $\mathbb{R}^{2 \times 2}$ ,  $H(x) \geq 0$  almost everywhere, and  $H \in L^1_{\text{loc}}$  (this latter requirement follows automatically if the entries are measurable and we do normalize the trace). The half line  $m$  functions are defined as  $m_{\pm}(z) = \pm f_{\pm}(0, z)$ ,  $z \in \mathbb{C}^+$ , where  $f_{\pm}$  solves (6.1) and  $f_{\pm} \in L^2_H(\mathbb{R}_{\pm})$ , that is,  $\int_0^{\infty} f_{\pm}^*(x)H(x)f_{\pm}(x) dx < \infty$ , and similarly for  $f_-$ . Here and throughout this section, we assume limit point case at both endpoints  $\pm\infty$ : there is a unique, up to a multiplicative constant, square integrable solution at each endpoint. If  $\text{tr } H = 1$ , or, more generally,  $\text{tr } H \notin L^1(\mathbb{R}_{\pm})$ , then this follows automatically, by a Theorem of de Branges, see [5] and also [1]. Finally, when we write  $m_{\pm}(z) = \pm f_{\pm}(0, z)$  we have again used the convention that a non-zero vector  $v \in \mathbb{C}^2$  is identified with the point  $v_1/v_2$  on the Riemann sphere.

Since any Herglotz function is the  $m$  function of a (unique, when trace-normed) canonical system (on a half line), it must in particular be possible to rewrite Jacobi and Schrödinger equations

$$(6.2) \quad -y''(x) + V(x)y(x) = zy(x)$$

as canonical systems. This can be done explicitly, and these transformations are well known. Let me discuss the case of a Schrödinger equation (6.2), with potential  $V \in L^1_{\text{loc}}(\mathbb{R})$ , and in the limit point case at  $\pm\infty$ . Given a solution  $y$ , let  $Y = (y', y)^t$ , and observe that  $Y$  then solves

$$Y' = \begin{pmatrix} 0 & V - z \\ 1 & 0 \end{pmatrix} Y.$$

Let  $T_0(x) \in \text{SL}(2, \mathbb{R})$  be the matrix solution of this equation for  $z = 0$  and with the initial value  $T_0(0) = 1$ . Write

$$T_0(x) = \begin{pmatrix} p'(x) & q'(x) \\ p(x) & q(x) \end{pmatrix}.$$

Finally, introduce  $u(x)$  by writing  $Y = T_0 u$ . This is essentially variation of constants for (6.2) about  $z = 0$ , which seems a reasonable thing to try since we must get rid of the terms not involving  $z$  if we want to write (6.2) as a canonical system. A calculation then shows that: (1)  $u$  indeed solves (6.1), with

$$(6.3) \quad H(x) = \begin{pmatrix} p^2(x) & p(x)q(x) \\ p(x)q(x) & q^2(x) \end{pmatrix};$$

(2)  $y \in L^2(0, \infty)$  precisely if  $u \in L^2_H(0, \infty)$ , and since the  $m$  function of (6.2) with Dirichlet boundary conditions  $y(0) = 0$  is given by  $m_S(z) = F(0, z) = f'(0, z)/f(0, z)$ , with  $f \in L^2(0, \infty)$ , this in particular shows that, as intended,  $m_S = m_C$ , that is, the  $m$  functions of the Schrödinger equation and the canonical system agree. The same remarks apply to the left half line  $(-\infty, 0)$ .

Note that  $H$  from (6.3) will usually not be trace normed; we could pass to a trace normed version of this (or any)  $H$  by the change of variable  $t = \int_0^x \text{tr } H(s) ds$ , but we prefer not to do so here. One reason for this is that such a normalization is inconvenient here because it will typically not be preserved by the flows we are about to construct.

The classical hierarchy of evolution equations of Schrödinger operators is the KdV hierarchy. This can be constructed in the traditional way from a Lax equation, or, more in line with what we do here, in the same way as discussed in Section 3 for the Toda hierarchy:  $\mathbb{R}$  acts by shifts  $(s \cdot V)(x) = V(x + s)$  (it is now  $\mathbb{R}$  rather than  $\mathbb{Z}$  because we have moved from the discrete to the continuous setting), and we may

now look for new flows (= actions of  $\mathbb{R}$ ) that come with associated  $\mathcal{SL}$ -cocycles for the action of  $G = \mathbb{R} \times \mathbb{R}$ , where, as above, the extra copy of  $\mathbb{R}$  acts by shifts. There are definitely extra technical difficulties involved, compared to the Jacobi case, since we cannot expect global flows on very general initial conditions, but leaving that aside, one finds that at least the formal side of this works just like before.

Moving on to canonical systems then, this would suggest to just try the same thing here, but I will instead propose a modified and more general approach. I will not look for flows that commute with literally the shift  $H(x) \mapsto H(x+s)$ , but rather with a modified version of this. One strong immediate motivation for this comes from the observation that a shift  $V(x) \mapsto V(x+s)$  will not induce just a shift on the corresponding  $H$  from (6.3); in fact, since always  $H(0) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$  that clearly could not be true. Rather, the shift on  $V$  induces the action

$$(6.4) \quad (s \cdot H)(x) = T_0(s)^{-1t} H(x+s) T_0(s)^{-1}$$

on  $H$ . This can be checked by hand, by verifying that the right-hand side is still (6.3), but with  $T_0(x)$  replaced by  $T_0(x+s)T_0(s)^{-1}$  (this is the right way to do it because we need the solution that is the identity matrix at  $x=0$ ), or, alternatively, it could be checked that  $m_{\pm}$  get updated correctly, though this second argument would require a uniqueness result also.

Now we take this as our guideline for suitable flows on general canonical systems that might take over the role of the shift. Motivated by (6.4), I propose to consider  $\mathbb{R}$  actions of the form

$$(6.5) \quad (s \cdot H)(x) = M(s; H)^t H(x+s) M(s; H),$$

for functions  $M : \mathbb{R} \times \mathcal{C} \rightarrow \mathrm{SL}(2, \mathbb{R})$  (where I denoted the set of coefficient functions of limit point case canonical systems by  $\mathcal{C}$ ) satisfying the following condition.

**Proposition 6.1.** *Suppose that  $T := M^{-1}$  satisfies the cocycle identity*

$$T(s+t; H) = T(s; t \cdot H) T(t; H).$$

*Then (6.5) defines a flow on  $\mathcal{C}$ .*

*Proof.* This is trivial; check it by direct calculation. □

We will call such an action of  $G = \mathbb{R}$  on  $\mathcal{C}$  a *twisted shift*.

As we discussed earlier, (differentiable) cocycles correspond to matrix functions  $S = S(H)$ ,  $\mathrm{tr} S = 0$  (we used to denote these by  $B$ , but I now want to call them  $S$  as in *shift*): given such an  $S$ , the cocycle is then obtained from

$$(6.6) \quad \dot{T} = S(s \cdot H)T, \quad T(0) = 1.$$

Moreover, once we have the cocycle  $T$ , we then obtain a flow from (6.5). However, notice that this does not really give an explicit construction of group actions of  $\mathbb{R}$ , after having chosen a matrix function  $S$  because in order to be able to solve (6.6), we need to have that action already. In fact, a random choice of  $S$  seems quite unlikely to produce a group action via (6.5), (6.6), for reasons that will be much clearer in the discrete setting, so we postpone the more detailed discussion until the end of this paper. Please see (6.18), (6.19) below and the comments that follow.

This somewhat circular structure can be camouflaged if we pass to a differential formulation. (However, note that it becomes truly circular only if we start out at the wrong end, with a matrix function  $S$ ; if we are given  $M = M(s; H)$  instead, then this matrix function satisfies the condition of Proposition 6.1 or it doesn't, and when it does, then we do obtain an action and a cocycle. There is no circularity here because if  $M$  is given, then (6.5) gives a perfectly meaningful definition of  $s \cdot H$ , though, despite the notation, this map might fail to define a group action.) To do this, assume that everything on the right-hand side of (6.5) is differentiable and take the  $s$  derivative at  $s = 0$ . Since  $(d/ds)M = -T^{-1}(dT/ds)T^{-1}$  and  $T(0) = 1$ ,  $(dT/ds)(0) = S(H)$ , this yields

$$(6.7) \quad \frac{\partial H}{\partial s} - \frac{\partial H}{\partial x} = -S^t(H)H - HS(H),$$

where  $H = H(x, s) = (s \cdot H)(x)$ . More precisely, we initially obtain this equation for  $s = 0$ , but then the general case, at  $s = s_0$ , say, results by applying this to  $(s \cdot (s_0 \cdot H))(x)$ .

We could now make this equation our starting point. Having chosen a matrix function  $S = S(H) \in \mathbb{R}^{2 \times 2}$ ,  $\text{tr } S = 0$ , we hope to obtain an action of  $\mathbb{R}$  on (parts of)  $\mathcal{C}$  from (6.7). Of course, whether or not that will actually be the case would need further investigation in any given case; (6.7) is a non-linear, non-local (thanks to the arbitrary dependence of  $S$  on  $\{H(x) : x \in \mathbb{R}\}$ ) PDE, so in principle anything could happen, and, as I just pointed out, I'm not particularly optimistic about our prospects if  $S$  is just a randomly chosen matrix function.

If we make the simplest possible choice,  $S \equiv 0$ , then (6.7) becomes  $\partial_s H - \partial_x H = 0$ , so recovers the plain shift  $(s \cdot H)(x) = H(x + s)$ . A second important and more interesting choice of  $S$  is given by

$$S(H) = \begin{pmatrix} 0 & V(0) \\ 1 & 0 \end{pmatrix}, \quad V(x) \equiv \frac{1}{4} \det H''(x).$$

This extends the twisted shift that corresponds to the shift  $V(x) \mapsto V(x+s)$  on Schrödinger operators to general canonical systems (with a twice differentiable coefficient function  $H$ ). To see that this is the case, check by a computation that if  $H$  corresponds to a Schrödinger operator in the way explained above, so is of the form (6.3), then indeed  $\det H'' = 4V$ , with now  $V = p''/p = q''/q$  being the potential from the original Schrödinger operator.

In general, for an  $S$  of this general type, we *can* guarantee existence of global solutions to (6.7) and existence of a group action.

**Proposition 6.2.** *Suppose that  $S = S(H)$  is of the form*

$$(6.8) \quad S = F(\det H(0), \det H'(0), \dots, \det H^{(n)}(0)),$$

*for some function  $F \in C^1$ ,  $\text{tr } F = 0$ . Then (6.7) has a global classical solution for any initial value  $H(x, s=0) = H_0(x) \in C^n(\mathbb{R})$ . Moreover, (6.5), (6.6) yield an  $\mathbb{R}$  action on coefficient functions  $H \in C^n(\mathbb{R})$ .*

To motivate the proof, let me mention the following fact, which is also of some independent interest. It says that the determinant of  $H$  is preserved along the characteristics  $x+s=c$ .

**Proposition 6.3.** *Suppose that  $H(x, s)$  solves (6.7). Then  $\det H(x, s) = \det H(x+s, 0)$ .*

*Proof.* This is unsurprising since (6.7) was meant to be a rewriting of (6.5), and  $\det M = 1$ . We'll check it by a calculation. It is helpful to introduce the new variables  $u = s-x$ ,  $v = s+x$ . Then  $2\partial H/\partial u = -S^t H - HS$ , and we want to show that  $(\partial/\partial u) \det H = 0$ . Write

$$H = \begin{pmatrix} a & b \\ b & c \end{pmatrix}, \quad S = \begin{pmatrix} \alpha & \beta \\ \gamma & -\alpha \end{pmatrix}.$$

Then, with  $X' \equiv \partial X/\partial u$ ,

$$\begin{aligned} a' &= -\alpha a - \gamma b, \\ 2b' &= -\beta a - \gamma c, \\ c' &= \alpha c - \beta b, \end{aligned}$$

and plugging this into  $(\det H)' = a'c + ac' - 2bb'$  will show that indeed  $(\det H)' = 0$ , as claimed.  $\square$

*Proof of Proposition 6.2.* Motivated by Proposition 6.3, we make the following attempt: Given an initial value  $H_0(x)$  for  $H$ , define  $T(s)$  by solving

$$(6.9) \quad \dot{T}(s) = F(\det H_0(s), \det H_0'(s), \dots, \det H_0^{(n)}(s))T(s), \quad T(0) = 1.$$

If we compare this with (6.6), then we find that this is not exactly the “correct” way to compute  $T$ , since we’ve replaced the (unknown, at this point) action  $(s \cdot H_0)(x)$  by  $H_0(x + s)$ ; however, by Proposition 6.3, we expect that this will have no consequences as far as computation of the determinants is concerned. Indeed, if we now define

$$(6.10) \quad H(x, s) = T(s)^{-1t} H_0(x + s) T(s)^{-1},$$

then it is in fact obvious that  $\det H_0^{(j)}(x + s) = \det(\partial^j H / \partial x^j)(x, s)$ . Moreover,  $(d/ds)T^{-1} = -T^{-1}S$ , by (6.9), and now a computation shows that  $H(x, s)$  from (6.10) solves (6.7).

The *moreover* part has the same proof: define  $T$  by (6.9) and then check that this  $T$  works in (6.5), (6.6).  $\square$

The next result completes the analogy between plain and twisted shifts by providing an  $\mathcal{SL}$ -cocycle for twisted shifts also, which again updates  $m$  functions correctly. Notice that from the point of view of the cocycles, the twisted shift is a plain shift, followed by the action of an  $\mathrm{SL}(2, \mathbb{R})$  matrix (in other words, an automorphism of  $\mathbb{C}^+$ ).

**Theorem 6.4.** *Consider the action (6.5) of  $G = \mathbb{R}$ , and denote the associated  $\mathrm{SL}(2, \mathbb{R})$ -cocycle from Proposition 6.1 by  $T_0 = T_0(s; H) = M^{-1}$ . Let  $T_1 = T_1(x; H)$  be the  $\mathcal{SL}$ -cocycle of the shift:  $dT_1/dx = zJHT_1$ ,  $T_1(0) = 1$ . Then  $T(s; H) = T_0(s; H)T_1(s; H)$  defines an  $\mathcal{SL}$ -cocycle for this action, and*

$$(6.11) \quad \pm m_{\pm}(z; s \cdot H) = T(s; H)(\pm m_{\pm}(z; H)).$$

*Proof.* Let’s first recall the following well known fact: If  $A \in \mathrm{SL}(2, \mathbb{R})$  and  $H_A(x) := A^t H(x) A$ , then  $\pm m_{\pm}(z; H_A) = A^{-1}(\pm m_{\pm}(z; H))$ . This is easy to verify by looking at how solutions are changed under this transformation.

Since  $(s \cdot H)(x) = T_0^{-1t}(s; H)H(x + s)T_0^{-1}(s; H)$  and  $T_1$  updates  $m_{\pm}$  under the shift flow, this in particular implies (6.11). So we only need to verify that  $T$  is a cocycle. Since  $T_0, T_1$  are cocycles individually, we have that

$$\begin{aligned} T(s + t; H) &= T_0(s + t; H)T_1(s + t; H) \\ &= T_0(s; t \cdot H)T_0(t; H)T_1(s; S_t H)T_1(t; H) \end{aligned}$$

(writing  $(S_t H)(x) \equiv H(x + t)$ ), so this will follow if we can show that

$$(6.12) \quad T_0(t; H)T_1(s; S_t H) = T_1(s; t \cdot H)T_0(t; H).$$

Call the two sides of this equation  $A$  and  $B$ , respectively, and observe that  $A, B$  are absolutely continuous functions of  $s$  (even if  $T$  itself

isn't, which might happen if we choose a sufficiently irregular  $T_0$ ). We compute

$$\frac{dA}{ds} = zT_0(t; H)JH(s+t)T_1(s; S_tH) = zJ(t \cdot H)(s)A;$$

in the last step we have used the identity  $CJC^t = J$ , which is valid for  $C \in \text{SL}(2, \mathbb{R})$ . Similarly,

$$\frac{dB}{ds} = zJ(t \cdot H)(s)B,$$

so  $A, B$  solve the same linear ODE, and since also  $A = B$  at  $s = 0$ , we indeed obtain (6.12).  $\square$

Since  $T = T_0T_1$  is a cocycle, if it is sufficiently regular, it must then itself satisfy an equation of the type

$$\dot{T} = C(s \cdot H)T, \quad T(0) = 1,$$

and indeed this is obviously the case, with

$$(6.13) \quad C(H) = \frac{d}{ds} (T_0T_1)|_{s=0} = S(H) + zJH(0).$$

This concludes our discussion of twisted shifts.

Let me now take a quick look at what could be done with this, following the ideas from Sections 2, 3, without attempting a systematic discussion. Let's first try to give a version of Theorem 2.3 that applies to the new setting. Here, we will make essential use of the *Weyl disks*

$$\mathcal{D}(x, z; H) = T_1(x, z; H)^{-1}\overline{\mathbb{C}^+} \quad (x > 0, z \in \mathbb{C}^+);$$

as above,  $T_1$  denotes the cocycle for the shift, so  $T_1' = zJHT_1$ ,  $T_1(0) = 1$ , and the closure of  $\mathbb{C}^+$  is taken on the Riemann sphere, so includes  $\infty$ . This set  $\mathcal{D}(x, z; H)$  is indeed a disk; its boundary is the image of  $\mathbb{R}_\infty$  under the linear fractional transformation  $T_1^{-1}$ . It is in fact a disk that is contained in  $\overline{\mathbb{C}^+}$  itself. From the cocycle property it is obvious that these disks are nested as  $x$  increases:  $\mathcal{D}(x, z; H) \subseteq \mathcal{D}(y, z; H)$  if  $x \geq y$ . So as  $x \rightarrow \infty$ , the disks  $\mathcal{D}(x, z; H)$  shrink to a limiting object  $\bigcap_{x>0} \mathcal{D}(x, z; H)$ , and we are assuming limit point case here, which means that this will be a point. More precisely, the intersection of the Weyl disks will be the unique point  $m_+(z; H)$ .

It's also useful to observe that  $\mathcal{D}(x, z; H)$  can be directly understood as those values that  $m_+(z; H)$  can still take, given the values of  $H(t)$  on  $0 \leq t \leq x$ . This is an immediate consequence of the fact that  $T_1$  updates  $m_+$  when  $H$  is shifted, and if we don't know anything about a half line, then any (generalized) Herglotz function is possible as its  $m_+$ .

**Theorem 6.5.** Fix  $z \in \mathbb{C}^+$ , and consider an action of  $G = \mathbb{R} \times \mathbb{R}$  on canonical systems and then an orbit  $\mathcal{O} = \{g \cdot H_0\}$  of this action. Assume that there is an associated  $\mathrm{SL}(2, \mathbb{C})$ -cocycle  $T(t, x; H)$ . Suppose that the second copy of  $\mathbb{R}$  acts by twisted shifts, and the cocycle extends the one from Theorem 6.4 in the sense that  $T(0, x; H) = T_0(x; H)T_1(x; H)$ . As for the first copy of  $\mathbb{R}$ , we assume that  $T(t, 0; H)$  satisfies the usual equation  $\dot{T} = B(t \cdot H)T$ ,  $T(0) = 1$ , with  $B$  bounded on  $\mathcal{O}$ . Finally, assume that  $\{m_{\pm}(z; H) : H \in \mathcal{O}\}$  is contained in a compact subset of  $\mathbb{C}^+$ .

Then

$$\pm m_{\pm}(z; g \cdot H) = T(g; H)(\pm m_{\pm}(z; H))$$

for all  $g \in G$ ,  $H \in \mathcal{O}$ .

A few remarks on this perhaps:

(1) We now use the variable  $x$  (rather than  $s$ , as before) for the twisted shift part, which I hope is still rather suggestive notation.

(2) While the list of assumptions may seem lengthy, I would still expect them to be very easy to verify in situations of interest, by again referring to a combination of compactness and continuity properties; compare the comments I made earlier on Theorem 2.3. For this, we need a metric on canonical systems that can serve as a replacement for (2.2), and this is actually somewhat less straightforward here than in the Jacobi case. [23, Section 2] has a discussion of these issues for Schrödinger operators.

(3) By focusing on a single orbit we perhaps make the assumptions easier to verify, but really the main point here is to avoid the issues with existence and uniqueness for general initial conditions. In other words, we bypass this by *assuming* that we have an  $H_0$  that we can evolve, and then we focus on this orbit exclusively.

*Proof.* I'll present the argument for  $m = m_+$ , and what I'll actually show is that  $T(g; H)m(H) \in \mathcal{D}(x; g \cdot H)$  for all  $x > 0$ . This will give the claim, by Weyl theory, as discussed above. Also, I've dropped all reference to  $z \in \mathbb{C}^+$ , which was fixed at the beginning.

By the cocycle property and since we already know, from Theorem 6.4, that  $T(0, x; H)$  updates  $m$  correctly, it suffices to treat the case  $g = (t, 0)$ . We are given that  $T_1(x; H)m(H) \in \overline{\mathbb{C}^+}$  for all  $x > 0$ , and we want to show that, similarly,  $T_1(x; t \cdot H)T(t, 0; H)m(H) \in \overline{\mathbb{C}^+}$  for  $x > 0$ . Here  $T_1$  may be replaced by  $T$  since these only differ by a matrix  $T_0 \in \mathrm{SL}(2, \mathbb{R})$ , which acts by an automorphism of  $\mathbb{C}^+$  and thus has no effect on what we're trying to establish. Then the cocycle property lets



us rewrite the claim as

$$T(t, 0; x \cdot H)T(0, x; H)m(H) \in \overline{\mathbb{C}^+},$$

or, since we know that  $T(0, x; H)$  updates correctly,

$$(6.14) \quad T(t, 0; x \cdot H)m(x \cdot H) \in \overline{\mathbb{C}^+}.$$

Now by assumption  $m(x \cdot H) \in K$  for some compact subset  $K \subseteq \mathbb{C}^+$ . Moreover, for small  $t > 0$ , our assumption on  $B$  guarantees that  $T(t, 0; x \cdot H) = 1 + o(1)$  as  $t \rightarrow 0+$ , uniformly in  $x \geq 0$ . So (6.14) will indeed hold, at least for small  $0 \leq t \leq \delta$ . Notice also that  $\delta$  depends only on the compact subset  $K \subseteq \mathbb{C}^+$  and on the bound on  $B$ . So we have established the claim of the theorem for these  $t$ , but then we can repeat the whole argument and obtain it for  $\delta \leq t \leq 2\delta$  also, and so forth. Here it is important that we never leave the orbit  $\mathcal{O}$ , so neither  $K$  nor the bound on  $B$  change and thus the same  $\delta > 0$  works in all steps.  $\square$

Let's now try to look for such actions of  $G = \mathbb{R} \times \mathbb{R}$  that extend a twisted shift together with its cocycle. At a formal level (and this is all we will talk about here), things proceed much as before; differences arise only because we are dealing with an action of  $\mathbb{R}$  by (twisted) shifts rather than a shift *map* (that is, an action of  $\mathbb{Z}$ ).

As in Section 3 and Theorem 6.5 above, let's denote by  $B = B(H)$  the matrix function that describes the part of the cocycle corresponding to the unknown evolution we're trying to construct; compare (2.7). For the twisted shift part, the corresponding matrix is given by  $C$  from (6.13).

In the cocycle identity

$$T(t, 0; x \cdot H)T(0, x; H) = T(0, x; t \cdot H)T(t, 0; H),$$

we differentiate with respect to  $t$  and then  $x$  at  $t = 0$ ,  $x = 0$ . The first step yields

$$(6.15) \quad B(x \cdot H)T(0, x; H) = T(0, x; H)B(H) + \partial_t T(0, x; t \cdot H)|_{t=0},$$

and then from this we obtain that

$$(\partial_x B)(H) + B(H)C(H) = C(H)B(H) + (\partial_x X)|_{x=0},$$

with  $X$  denoting the second term on the right-hand side of (6.15). We now make the additional assumption that these two derivatives could have been taken in the reverse order, and this produces

$$\partial_t C - \partial_x B = [B, C];$$

as before, the derivatives must be interpreted as

$$(\partial_x B)(H) \equiv \frac{d}{dx} B(x \cdot H)|_{x=0},$$

and similarly for  $\partial_t C$ . Or, plugging  $C$  from (6.13) into this, we may rewrite the equation as

$$(6.16) \quad \partial_t S + zJ\partial_t H - \partial_x B = [B, S] + z[B, JH],$$

and this is the basic evolution equation (“zero curvature equation”) here.

As we discussed in Section 3, this can (and probably should) be thought of as the compatibility condition on the individual cocycles (twisted shift and the one of the unknown flow, which is described by  $B$ ) that guarantees that gluing these together will result in a joint cocycle.

We now have a very general framework; for every twisted shift, which will provide a matrix function  $S(H)$ , the zero curvature equation (6.16) potentially gives a hierarchy of evolutions. As in the Toda case, the cocycle that we started out with is not just a collection of arbitrarily evolving matrix functions; on the contrary, it again has the important property that it updates the  $m$  functions correctly (that is, assuming that we are able to verify the assumptions of Theorem 6.5, which we don’t even attempt here).

Unlike in the Toda case, there doesn’t seem to be an obvious way to make (6.16) spit out hierarchies of evolutions, and in fact some negative results have been obtained by Hur and Ong [13]. They discuss the case  $S = 0$  in (6.8), corresponding to the plain shift, so look at the simplified equation

$$(6.17) \quad zJ\partial_t H - \partial_x B = z[B, JH],$$

and then search for  $B$ ’s with polynomial dependence on  $z$  that satisfy this compatibility condition. For constant  $B$  (of trace zero), this recovers the flows

$$(t \cdot H)(x) = e^{-tB^t} H(x) e^{-tB}$$

that conjugate  $H$  by an  $\mathrm{SL}(2, \mathbb{R})$  matrix and that of course could have been defined directly. Note that these do preserve the property that  $H \geq 0$ . Whether or not that will be the case in more general situations is quite unclear. The cocycle is given by  $T(t, 0; H) = e^{tB}$ .

So everything is in perfect order in the degree zero case, but now Hur-Ong show that (6.17) will not yield evolutions with the desired properties beyond degree zero. This result also confirms that the added flexibility of twisted shifts may be crucial.

So a more productive choice of  $S$  would be the one from (6.8); we already know that we have the KdV hierarchy on the subclass of canonical systems that correspond to Schrödinger operators, and now our formalism will extend this evolution to general canonical systems. This works because we have found a way of expressing  $V = (1/4) \det H''$  in terms of  $H$ , and in such a way that this makes sense for arbitrary (smooth)  $H$ , not necessarily coming from a Schrödinger equation. The same procedure can now be applied to the  $B$ 's from the classical KdV hierarchy to extend everything to canonical systems.

I'll leave the matter at that, but clearly the general properties of the formalism need further investigation.

Let me close this paper by making a few very quick remarks on how to adapt this material to the case of evolutions commuting with a twisted shift *map* rather than a flow; in other words, we have an action of  $\mathbb{Z}$  rather than of  $\mathbb{R}$ , and so we return to the situation already familiar to us from the Jacobi case.

We proceed as above. As in (6.5), we consider maps induced by group elements  $n \in G = \mathbb{Z}$  of the form

$$(6.18) \quad (n \cdot H)(x) = M(n; H)^t H(x+n) M(n; H),$$

for some matrix function  $M : \mathbb{Z} \times \mathcal{C} \rightarrow \mathrm{SL}(2, \mathbb{R})$ . This will define a group action of  $\mathbb{Z}$  if  $T = M^{-1}$  has the cocycle property, and this is the same as saying that  $T$  must be of the form

$$(6.19) \quad T(n; H) = \begin{cases} A((n-1) \cdot H) \cdots A(H) & n \geq 0 \\ A^{-1}(n \cdot H) \cdots A^{-1}(-1 \cdot H) & n < 0 \end{cases}$$

for some matrix function  $A$  taking values in  $\mathrm{SL}(2, \mathbb{R})$ .

We can now finish the discussion that I started after stating (6.5), (6.6) and Proposition 6.1. In the current situation, this issue now takes the following form: Why can't I just, conversely, pick a matrix function  $A$  and obtain a cocycle and a  $\mathbb{Z}$  action from (6.18), (6.19)? Of course, as before, there is the obvious concern that (6.19) will only define a cocycle if I had the action already, but to define this via (6.18), I need the cocycle. We can make this more explicit: having chosen an  $A$ , (6.18), (6.19) will formally produce

$$(1 \cdot H)(x) = A(H)^{-1t} H(x+1) A(H)^{-1},$$

but this can be part of a group action only if the map  $H \mapsto 1 \cdot H$  is bijective, and it's easy to cook up  $A$ 's for which it fails to be injective.

In any event, if we do have a  $\mathbb{Z}$  action constructed in this way, then we will again call it a *twisted shift*. These twisted shift maps not only mimic what we did above, but they also again arise naturally if we want

to think of the Toda hierarchy in terms of canonical systems. Namely, if a Jacobi equation is rewritten as a canonical system, by a procedure that is similar to the one used above for Schrödinger equations, then the shift on  $J$  again induces a *twisted* shift on the corresponding canonical system. Everything is completely analogous to the Schrödinger case; in particular, the twisting is again done by the transfer matrix  $A = \begin{pmatrix} -b_1/a_1 & 1/a_1 \\ -a_1 & 0 \end{pmatrix}$  of the Jacobi matrix at  $z = 0$ .

Returning to the general situation then, we can next establish an analog of Theorem 6.4: twisted shift maps come with an  $\mathcal{SL}$ -cocycle, which, moreover, will update  $\pm m_{\pm}$  correctly.

Finally, suppose again that we have an action of  $G = \mathbb{R} \times \mathbb{Z}$  and an associated  $\mathcal{SL}$ -cocycle  $T(g; H)$ , with  $\mathbb{Z}$  acting by a twisted shift map and  $T$  extending the twisted shift cocycle from the analog of Theorem 6.4 that I just mentioned. Let's introduce some notation here and let's call this twisted shift cocycle  $S(n; H)$  (please do not confuse this with the  $S$  from above, which was also related to the twisted shift cocycle, but in a different way). So we are then assuming that  $T(0, n; H) = S(n; H)$ . Let  $B(H)$  again be the matrix that describes the part of the cocycle corresponding to the (unknown) flow. Then exactly the same (formal, without further clarification of the precise assumptions) calculation that led to (3.5) produces the zero curvature equation

$$\dot{S}(H) = S(1 \cdot H)B(H) - B(H)S(H);$$

here, I've used the (perhaps confusing) short-hand notation  $S(H) \equiv S(1; H)$ , and this also equals  $S(H) = A(H)T_1(1; H)$ , where  $T_1$  again denotes the transfer matrix of the plain shift,  $T_1' = zJHT_1$ ,  $T_1(0) = 1$ .

## REFERENCES

- [1] K. Acharya, An alternate proof of the de Branges theorem on canonical systems, *ISRN Math. Anal.* (2014), 7 pp.
- [2] I. Binder, D. Damanik, M. Lukic, and T. Vandenboom, Almost periodicity in time of solutions of the Toda lattice, arXiv preprint 1603.04905 (2016).
- [3] J. Breuer, E. Ryckman, and B. Simon, Equality of the spectral and dynamical definitions of reflection, *Comm. Math. Phys.* **295** (2010), 531–550.
- [4] D. Damanik and M. Goldstein, On the existence and uniqueness of global solutions for the KdV equation with quasi-periodic initial data, *J. Amer. Math. Soc.* **29** (2016), 825 – 856.
- [5] L. de Branges, *Hilbert Spaces of Entire Functions*, Prentice-Hall, Englewood Cliffs, 1968.
- [6] L. Dickey, *Soliton equations and Hamiltonian systems*, Advanced series in mathematical physics, 26, World Scientific Publishing, River Edge, 2003.
- [7] R. Feres, An introduction to cocycle super-rigidity, in: *Rigidity in dynamics and geometry*, 99 – 134, Springer, Berlin, 2002.

- [8] F. Gesztesy and H. Holden, Soliton equations and their algebro-geometric solutions, vol. I, (1+1)-dimensional continuous models, Cambridge Studies in Advanced Mathematics, 79, Cambridge University Press, Cambridge, 2003.
- [9] F. Gesztesy, H. Holden, J. Michor, and G. Teschl, Soliton equations and their algebro-geometric solutions, vol. II, (1+1)-dimensional discrete models, Cambridge Studies in Advanced Mathematics, 114, Cambridge University Press, Cambridge, 2008.
- [10] F. Gesztesy, R. Nowell, and W. Pötz, One-dimensional scattering theory for quantum systems with nontrivial spatial asymptotics, *Diff. Int. Equations* **10** (1997), 521–546.
- [11] F. Gesztesy and B. Simon, Inverse spectral analysis with partial information on the potential, I. The case of an a.c. component in the spectrum, *Helv. Phys. Acta* **70** (1997), 66–71.
- [12] I. Hur, M. McBride, and C. Remling, The Marchenko representation of reflectionless Jacobi and Schrödinger operators. *Trans. Amer. Math. Soc.* **368** (2016), 1251 – 1270.
- [13] I. Hur and D. Ong, in preparation (2017).
- [14] V. Jaksic and Y. Last, A new proof of Poltoratski’s theorem, *J. Funct. Anal.* **215** (2004), 103–110.
- [15] A. Katok, Combinatorial constructions in ergodic theory and dynamics, University Lecture Series 30, AMS, Providence, RI, 2003.
- [16] A. Katok and B. Hasselblatt, Principal structures, in: Handbook of dynamical systems IA, 1 – 203, Elsevier, 2002.
- [17] D. Ong, On a description of the Toda hierarchy using cocycle maps, arXiv preprint 1801.10021 (2018).
- [18] D. Ong and C. Remling, Generalized Toda flows, arXiv preprint 1801.03053 (2018).
- [19] A. Poltoratski, Boundary behavior of pseudocontinuable functions, *St. Petersburg Math. J.* **5** (1994), 389–406.
- [20] A. Poltoratski and C. Remling, Reflectionless Herglotz functions and Jacobi matrices, *Comm. Math. Phys.* **288** (2009), 1007–1021.
- [21] A. Poltoratski and C. Remling, Approximation results for reflectionless Jacobi matrices, *Int. Math. Res. Not.* **16** (2011), 3575 – 3617.
- [22] V.P. Potapov, The multiplicative structure of  $J$ -contractive matrix functions, *Amer. Math. Soc. Transl. (2)* **15** (1960), 131–243.
- [23] C. Remling, The absolutely continuous spectrum of one-dimensional Schrödinger operators, *Math. Phys. Anal. Geom.* **10** (2007), 359 – 373.
- [24] C. Remling, The absolutely continuous spectrum of Jacobi matrices, *Annals of Math. (2)* **174** (2011), 125–171.
- [25] C. Remling, Generalized reflection coefficients, *Comm. Math. Phys.* **337** (2015), 1011–1026.
- [26] A. Rybkin, On the evolution of a reflection coefficient under the Korteweg-de Vries flow, *J. Math. Phys.* **49** (2008), 15pp.
- [27] G. Teschl, Jacobi Operators and Completely Integrable Nonlinear Lattices, Mathematical Monographs and Surveys, 72, American Mathematical Society, Providence, 2000.
- [28] H. Winkler, The inverse spectral problem for canonical systems, *Int. Eq. Op. Theory* **22** (1995), 360–374.

MATHEMATICS DEPARTMENT, UNIVERSITY OF OKLAHOMA, NORMAN, OK  
73019

*E-mail address:* [christian.remling@ou.edu](mailto:christian.remling@ou.edu)

*URL:* [www.math.ou.edu/~cremling](http://www.math.ou.edu/~cremling)