On Hamiltonian flows whose orbits are straight lines

Hans Koch [1] and Héctor E. Lomelí [2]

Abstract. We consider real analytic Hamiltonians on $\mathbb{R}^n \times \mathbb{R}^n$ whose flow depends linearly on time. Trivial examples are Hamiltonians $H(q,p)$ that depend only on the coordinate $p \in \mathbb{R}^n$. By a theorem of Moser [5], every cubic Hamiltonian reduces to a Hamiltonian of this type via a linear symplectic change of variables. We show that the same does not hold for polynomials of degree ≥ 4. But we give a condition that implies linear-symplectic conjugacy to another simple class of Hamiltonians. The condition is shown to hold for all nondegenerate Hamiltonians that are homogeneous of degree 4.

1. Introduction and main results

Polynomial Hamiltonians and maps have been studied extensively and for a variety of different reasons. Among other things, they constitute local normal forms for more general Hamiltonians and maps, and they provide a convenient testing ground for new ideas in dynamical systems. The restriction to polynomials also adds interesting algebraic aspects to the problem. This includes the possibility of classifying polynomial maps with a given property, and of decomposing them into simpler ones.

The maps considered here are symplectic and arise from Hamiltonian flows. We use the standard symplectic form on $\mathbb{R}^n$, so a differentiable map $F : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ is symplectic if and only if

$$DF(x)^\top \mathrm{J} DF(x) = \mathrm{J}\,, \qquad \mathrm{J} = \begin{bmatrix} 0 & \mathrm{I} \\ -\mathrm{I} & 0 \end{bmatrix}, \qquad x \in \mathbb{R}^{2n}\,. \tag{1.1}$$

Here $DF(x)$ denotes the derivative of F at x, and $DF(x)^\top$ denotes its transpose (as a matrix). Let H be a smooth function on $\mathbb{R}^{2n}$. One of the basic facts from Hamiltonian mechanics is that the vector field $X = \mathrm{J}\nabla H$ defines a flow $\Phi : (t,x) \mapsto \Phi^t(x)$ whose time-t maps Φ^t are symplectic. We will mainly consider Hamiltonians of the following type:

Definition 1.1. *We say that a Hamiltonian H is affine-integrable if its flow Φ is linear in time:*

$$\Phi^t = \mathrm{I} + tX\,, \qquad X = \mathrm{J}\nabla H\,, \qquad t \in \mathbb{R}\,. \tag{1.2}$$

Equivalently, a Hamiltonian is affine-integrable if and only if the corresponding vector field X is constant along each orbit. In particular, $X \circ (\mathrm{I} + X) = X$. General (polynomial) maps $F = \mathrm{I} + X$ with this property are also called quasi-translations. They arise naturally in the study of singular Hessians [7]. The identity $X \circ (\mathrm{I} + tX) = X$ for $t = 1$ extends to $t \in \mathbb{Z}$ by induction, and to $t \in \mathbb{C}$ if X is a polynomial. Differentiating it with respect to t yields $(DX)X = 0$, or equivalently, $(X^\top\nabla)^2\ell = 0$ for all linear functions ℓ. This "local nilpotency" property is an alternative way of characterizing quasi-translations [10] and affine-integrable Hamiltonians [4,6].

[1] Department of Mathematics, University of Texas at Austin, Austin, TX 78712

[2] Department of Mathematics, Instituto Tecnológico Autónomo de México, México DF 01000, Mexico
Current address: Department of Mathematics, University of Texas at Austin, Austin, TX 78712

In numerical analysis and physics, symplectic quasi-translations are also called *jolt maps*. They constitute the basic building blocks in the so-called Dragt-Finn factorization [2] of more general symplectic maps. This factorization has proved to be very useful in symplectic numerical schemes, including the simulation of Hamiltonian flows in plasmas [4,9].

From a dynamical systems point of view, affine-integrable Hamiltonians are rather simple. Not only is the vector field X constant along each orbit, but its components X_j are Poisson-commuting invariants, as we will see later. So an affine-integrable Hamiltonian H is Liouville integrable, at least if it satisfies a suitable nondegeneracy condition. In addition, the geometry defined by the invariants X_j is quite restricted:

Theorem 1.2. *Let H be a real analytic affine-integrable Hamiltonian on $\mathbb{R}^{2n}$. Then H and its vector field X are constant on the affine subspaces $x + \mathrm{range}(DX(x))$. If $DX(x)$ has rank n then $x + \mathrm{range}(DX(x))$ is a local level set for X.*

The only affine-integrable Hamiltonians that we have been able to find in the literature are all linear-symplectically conjugate to Hamiltonians of the form $H(q, p) = K(p)$. The time-t map for such a Hamiltonian is a shear: $\Phi^t(q, p) = \big(q + t\nabla K(p), p\big)$.

Definition 1.3. *We call $H : \mathbb{R}^{2n} \to \mathbb{R}$ a shear Hamiltonian if there exists a linear symplectic change of variables $U : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ such that $(H \circ U)(q, p) = K(p)$ for some function $K : \mathbb{R}^n \to \mathbb{R}$.*

Remark 1. As we will describe later, the linear map U in the above definition can be chosen both symplectic and orthogonal (as a matrix).

One of our goals is to find an affine-integrable Hamiltonian that is not a shear, or to prove that there is no such Hamiltonian. Partial non-existence results can be obtained by restricting the class of Hamiltonians being considered. A trivial case: If $n = 1$ then line-orbits are necessarily parallel, so if $H : \mathbb{R}^2 \to \mathbb{R}$ is affine-integrable, then there exists a rotation U of $\mathbb{R}^2$, such that $(H \circ U)(q, p)$ is independent of q. In other cases the Definition 1.3 cannot be used directly. The following theorem gives a coordinate-independent characterization of shear Hamiltonians. It is a slight generalization of a result in [8] on Hamiltonians that are homogeneous polynomials of degree ≥ 3. Homogeneous vector fields (of positive degree) vanish at the origin; and if $X = \mathrm{J}\nabla H$ vanishes at some point, then H is regular in the following sense:

Definition 1.4. *We say that a Hamiltonian H is regular if there exists a point x where $X(x)$ belongs to the range of $DX(x)$.*

Theorem 1.5. *A regular real analytic function H on $\mathbb{R}^{2n}$ is a shear Hamiltonian if and only if $DX(x)DX(y) = 0$ holds for all $x, y \in \mathbb{R}^{2n}$.*

To continue our discussion of special cases, assume that H is regular and affine-integrable. Then $(DX)^2 = 0$, as we will see later. In particular, if H is quadratic then Theorem 1.5 implies that H is a shear Hamiltonian. The cubic case is covered by a result of Moser [5] on quadratic symplectic maps on $\mathbb{R}^{2n}$. It states that every such map F admits a decomposition $F = A \circ S \circ L$ into three simple symplectic maps: an affine map A, a shear $S(q,p) = (q + s(p), p)$, and a linear map L. It is not hard to show that this result implies – and is essentially equivalent to – the statement that every homogeneous affine-integrable Hamiltonian of degree 3 is a shear. A direct proof is given in Section 2.

This raises the question [8] whether every homogeneous affine-integrable Hamiltonian on $\mathbb{R}^{2n}$ is a shear. Locally, much more is true: It is well known that every smooth Hamiltonian H is of the form $H(q,p) = K(p)$ in some local symplectic chart, near any point where the vector field does not vanish. The local conjugacy (chart) is nonlinear in general. But if all orbits for H are straight lines with constant velocity, as is the case for affine-integrable Hamiltonians, and if X is constant on n-dimensional affine subspaces, then one might think that this conjugacy can be chosen to be linear. However, this is not true in general:

Theorem 1.6. *The following Hamiltonian on $\mathbb{R}^8$ is affine-integrable but not a shear:*

$$H(q,p) = q_1 p_3^3 + \sqrt{3} q_2 p_3^2 p_4 + p_1 p_4^3 - \sqrt{3} p_2 p_3 p_4^2 \,, \qquad q, p \in \mathbb{R}^4 \,. \tag{1.3}$$

In addition, H is nondegenerate in the sense defined below.

Definition 1.7. *A real analytic vector field X on $\mathbb{R}^{2n}$ is said to be nondegenerate if $DX(x)$ has rank $\geq n$ at some point $x \in \mathbb{R}^{2n}$. If $X = \mathrm{J}\nabla H$ then we also say that H is nondegenerate.*

We note that, if $X = \mathrm{J}\nabla H$, then the rank of $DX(x)$ can be no larger than n. And if X is analytic, then the rank is constant outside some analytic set of codimension one.

The example (1.3) belongs to a simple class of Hamiltonians that we shall now describe. Let $0 \leq d < n$. To simplify the description, we write $q = (Q, \bar{q})$ and $p = (P, \bar{p})$, where $Q, P \in \mathbb{R}^d$ and $\bar{q}, \bar{p} \in \mathbb{R}^{n-d}$. Consider a Hamiltonian of the form

$$H(q,p) = K(\bar{p}) + Q^\top V(\bar{p}) + P^\top W(\bar{p}) \,, \tag{1.4}$$

with $K : \mathbb{R}^{n-d} \to \mathbb{R}$ and $V, W : \mathbb{R}^{n-d} \to \mathbb{R}^d$ differentiable. Notice that H does not depend on $\bar{q}$, and thus $\bar{p}$ stays fixed under the flow. Furthermore, the coordinates Q and P evolve linearly (in time) under the flow. If $\bar{q}$ evolves linearly as well, then H is affine-integrable. As we will see later, this is the case if and only if

$$W(\bar{p})^\top DV(\bar{p}) - V(\bar{p})^\top DW(\bar{p}) = 0 \,. \tag{1.5}$$

If $d = 0$, then $P = Q = 0$ and $\bar{p}$ can be identified with p. In this case, (1.4) becomes $H(q,p) = K(p)$, so H is a shear Hamiltonian.

Remark 2. For reference below, we note that the Hamiltonian (1.4) can be written as the sum of $H_1 = Q^\top V(\bar{p})$ and $H_2 = K(\bar{p}) + P^\top W(\bar{p})$. What makes this decomposition interesting is that the Poisson bracket $\{H_1, H_2\} = (\nabla H_1)^\top \mathrm{J}(\nabla H_2)$ of H_1 and H_2 Poisson-commutes with both H_1 and H_2.

Theorem 1.8. *Let H be a nondegenerate real analytic affine-integrable Hamiltonian on $\mathbb{R}^{2n}$. Then H is linear-symplectically conjugate to a Hamiltonian of the form (1.4) if and only if $DX(x)DX(y)DX(z) = 0$ for all $x, y, z \in \mathbb{R}^{2n}$.*

Here, as in Theorem 1.2, a simple class of affine-integrable Hamiltonian is characterized by a nilpotency-type condition on the derivative of the vector field. This suggest there may be a natural hierarchy of such conditions, which characterize classes of increasingly complex affine-integrable Hamiltonians.

We will prove that the condition on DX in Theorem 1.8 holds if H is a homogeneous polynomials of degree 4. As a result we obtain

Theorem 1.9. *Let H be a nondegenerate affine-integrable Hamiltonian on $\mathbb{R}^{2n}$. If H is homogeneous of degree 4 then H is is linear-symplectically conjugate to a Hamiltonian of the form (1.4).*

Homogeneous affine-integrable Hamiltonians can be obtained from symplectic maps $F : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ for which $X = F - \mathrm{I}$ is a homogeneous polynomial for degree ≥ 2. Any map with this property is the time-one map of an affine-integrable Hamiltonian [8]. Combining this result with Theorem 1.9, and using the decomposition described in Remark 2, we obtain the following theorem.

Theorem 1.10. *Let F be a symplectic map on $\mathbb{R}^{2n}$ such that $X = F - \mathrm{I}$ is homogeneous of degree 3 and nondegenerate. Then F admits a decomposition $F = F_1 \circ F_4$, where F_1 and F_4 are the time-one maps of two shear Hamiltonians.*

For the proofs of Theorems 1.2, 1.5, 1.6, 1.8, 1.9, and 1.10, we refer to Sections 3, 2, 6, 4, 5, and 7, respectively. Some additional results can be found in Section 2.

2. Basic properties

In the remaining part of this paper we always assume that H is a real analytic Hamiltonian on $\mathbb{R}^{2n}$. Furthermore, by a "homogeneous" Hamiltonian we always mean a homogeneous polynomial of degree ≥ 3.

As is true generally, the Hamiltonian H is invariant under the flow that it generates, so $(DH)X = 0$. Assuming that H is affine-integrable, $X \circ (\mathrm{I} + tX) = X$, and thus $(DX)X = 0$. Furthermore, $\Phi^t = \mathrm{I} + tX$ is symplectic, which by (1.1) yields

$$\mathrm{J} + t\big[(DX)^\top \mathrm{J} + \mathrm{J}DX\big] + t^2\big[(DX)^\top \mathrm{J}DX\big] = \mathrm{J}\,. \tag{2.1}$$

Using that the terms of order t and t^2 have to vanish separately, we get $(DX)^2 = 0$. Differentiating the identity $X \circ (\mathrm{I} + tX) = X$ yields $[DX \circ (\mathrm{I} + tX)](\mathrm{I} + tDX) = DX$. Multiplying on the right by $(\mathrm{I} - tDX)$ and using that $(DX)^2 = 0$, we find in addition that $DX \circ (\mathrm{I} + tX) = DX$. In summary, we have the following

Lemma 2.1. *Let H be an affine-integrable Hamiltonian. Then the functions H and X and DX are constant along every orbit. Furthermore, $(DH)X = 0$ and $(DX)X = 0$ and $(DX)^2 = 0$.*

An equivalent formulation can be given in terms of Poisson brackets. Assume that H is affine-integrable. Let $\ell(x) = (\mathrm{J}u)^\top x$ for some vector $u \in \mathbb{R}^{2n}$. Then $\{\ell, H\}$ is the directional derivative of H in the direction u, which we denote by $\partial_u H$. Being linear in the coordinate x, ℓ evolves linearly in time, so $\{\ell, H\} = \partial_u H$ is invariant under the flow. This implies the first identity in

$$\{\partial_u H, H\} = 0\,, \qquad \{\partial_u \partial_v H, H\} = 0\,, \qquad \{\partial_u H, \partial_v H\} = 0\,. \tag{2.2}$$

The second and third identities are obtained from the first by applying a derivative ∂_v. This yields $\{\partial_u\partial_v H, H\} + \{\partial_u H, \partial_v H\} = 0$, and the two terms have to vanish separately since the first is symmetric in (u, v) and the second antisymmetric. If H is nondegenerate, then (2.2) shows that n of the vector field components X_j constitute a maximal set of Poisson-commuting invariants. So H is Liouville integrable, as mentioned earlier.

Let x be a fixed but arbitrary point in $\mathbb{R}^{2n}$. In the canonical splitting $\mathbb{R}^{2n} = \mathbb{R}^n \times \mathbb{R}^n$, we can represent the derivative of X and the Hessian $\mathbb{H}(x) = \nabla DH(x)$ as 2×2 matrices whose entries are $n \times n$ matrices,

$$DX(x) = \begin{bmatrix} \mathcal{Z}(x)^\top & \mathcal{A}(x) \\ -\mathcal{B}(x) & -\mathcal{Z}(x) \end{bmatrix}, \qquad \mathbb{H}(x) = \begin{bmatrix} \mathcal{B}(x) & \mathcal{Z}(x) \\ \mathcal{Z}(x)^\top & \mathcal{A}(x) \end{bmatrix}. \tag{2.3}$$

Given that $X = \mathrm{J}\nabla H$, we have $DX = \mathrm{J}\mathbb{H}$. Since $\mathbb{H}$ is symmetric, so are $\mathcal{A}$ and $\mathcal{B}$. In the case of an affine-integrable Hamiltonian, $\mathcal{A}\mathcal{Z}$ is symmetric as well, as a result of the identity $(DX)^2 = 0$.

Lemma 2.2. *A regular Hamiltonian is of the form $H(q, p) = K(p)$ if and only if $\mathcal{Z}(x) = 0$ and $\mathcal{B}(x) = 0$ for all x.*

Proof. The necessity of the conditions $\mathcal{Z} = 0$ and $\mathcal{B} = 0$ is obvious. Assume now that they are satisfied. Let $x_0 = (q_0, p_0)$ be a point where $X(x_0) = \bigl(\nabla_p H(x_0), -\nabla_q H(x_0)\bigr)$ belongs to the range of $DX(x_0)$. At this point we have $\nabla_q H(x_0) = 0$. Given that $D_q^2 H = 0$ by assumption, this implies that the function $q \mapsto H(q, p_0)$ is constant. Furthermore, $D_p H(q, p)$ does not depend on q, since $D_q D_p H = 0$. So $H(q, p)$ is independent of q as well, since $H(q, p) = H(q, p_0) + \int_0^1 D_p H(q, p_0 + sv)v\,ds$ with $v = p - p_0$. **QED**

Lemma 2.3. *Let H be an affine-integrable Hamiltonian. Given $x \in \mathbb{R}^{2n}$, there exists an orthogonal symplectic $2n \times 2n$ matrix U, and a diagonal $n \times n$ matrix A, such that*

$$U^{-1} DX(x) U = \begin{bmatrix} 0 & A \\ 0 & 0 \end{bmatrix}, \qquad U^\top \mathbb{H}(x) U = \begin{bmatrix} 0 & 0 \\ 0 & A \end{bmatrix}. \tag{2.4}$$

Proof. Let d be the rank of $M = \mathbb{H}(x)$. Let $(u_1, u_2, \ldots, u_d)$ be an orthonormal set of eigenvectors for the nonzero eigenvalues of M. Since $M\mathrm{J}M = 0$, the vectors $\mathrm{J}u_j$ are eigenvectors of M for the eigenvalue 0. Consider first the case $d = n$. Let U be the $2n \times 2n$ matrix whose columns vectors are $\mathrm{J}u_1, \ldots, \mathrm{J}u_n, u_1, \ldots, u_n$, in this order. Clearly, U is orthogonal and $U^\top M U$ diagonal. A simple computation shows that U is symplectic.

If $d < n$, consider the orthogonal projection P onto the span of $\mathrm{J}u_1, \ldots, \mathrm{J}u_d, u_1, \ldots, u_d$. Then P commutes with both M and J. So we can choose an orthonormal set of vectors $\mathrm{J}u_{d+1}, \ldots, \mathrm{J}u_n, u_{d+1}, \ldots, u_n$ in the null space of P and define U as above. **QED**

The same construction can be used to give a

Proof of Theorem 1.5. The necessity of the condition $DX(x)DX(y) = 0$ is obvious. Assume now that this condition holds, for all $x, y \in \mathbb{R}^{2n}$, and that H is regular.

First, we show that H is affine-integrable. By regularity, there exist $x_0, w \in \mathbb{R}^n$ such that $DX(x_0)w = X(x_0)$. Thus $X(x) = DX(x_0)w + \int_0^1 DX(x_0 + sv)v\,ds$, for any given $x \in \mathbb{R}^{2n}$, where $v = x - x_0$. This shows that $DX(x)X(x) = 0$ for all x, which implies that H is affine-integrable.

Let $(u_1, u_2, \ldots, u_d)$ be an orthonormal basis for the subspace spanned by all vectors $\mathbb{H}(y)z$ with $y, z \in \mathbb{R}^{2n}$. Then $\mathbb{H}(x)\mathrm{J}u_j = 0$ for all x and all j. Defining U as in the proof of Lemma 2.3, we obtain (2.4) simultaneously for all x. (The matrix A can depend on x and need not be diagonal.) So $(H \circ U)(q,p)$ is independent of q by Lemma 2.2, implying that H is a shear Hamiltonian. **QED**

As a corollary we obtain

Theorem 2.4. [5] *Every affine-integrable Hamiltonian H that is homogeneous of degree 3 is a shear.*

Proof. By (2.2) we have $\{\partial_u^k H, H\} = 0$ for $k \leq 2$. The same holds for $k \geq 3$ since H is of degree 3. It follows that $\{H(\,.\,+u), H\} = 0$ for all u. Or equivalently, $X(x)^\top \mathrm{J} X(y) = 0$ for all x and y. From this we get $DX(x)DX(y) = 0$ by differentiation, and the assertion follows from Theorem 1.5. **QED**

The matrix U described in Lemma 2.3 is both symplectic and orthogonal. This means that $U^\top \mathrm{J} U = \mathrm{J}$ and $U^\top U = \mathrm{I}$. As a result, we also have $\mathrm{J}U = U\mathrm{J}$. In fact, any two of the three properties imply the third. This is know as the 2-out-of-3 property of the unitary group $\mathrm{U}(n) = \mathrm{O}(n) \cap \mathrm{Sp}(2n, \mathbb{R}) \cap \mathrm{GL}(n, \mathbb{C})$. The complex structure here is given by the matrix J, and the equation $\mathrm{J}U = U\mathrm{J}$ simply says that U is "complex". Using the properties $U^\top U = \mathrm{I}$ and $\mathrm{J}U = U\mathrm{J}$, any matrix $U \in \mathrm{U}(n)$ can be written as

$$U = \begin{bmatrix} S & T \\ -T & S \end{bmatrix}, \qquad S^\top S + T^\top T = \mathrm{I}, \quad S^\top T = T^\top S\,. \tag{2.5}$$

We will refer to such a $2n \times 2n$ matrix as being *unitary*. The $n \times n$ submatrices S and T will be referred to as the real and imaginary parts of U, respectively.

Concerning the claim in Remark 1, we note that any symplectic matrix M can be written as a product $M = UAN$, where U is unitary, A positive diagonal, and N unipotent upper-triangular. This is the standard Iwasawa decomposition [3]. If H is a Hamiltonian such that $(H \circ M)(q,p)$ is independent q, then $(H \circ U)(q,p)$ is independent of q as well.

By Lemma 2.3, the Hessian $\mathbb{H}(x)$ of an affine-integrable Hamiltonian H is always of the form

$$\mathbb{H}(x) = \begin{bmatrix} S & T \\ -T & S \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & A \end{bmatrix} \begin{bmatrix} S^\top & -T^\top \\ T^\top & S^\top \end{bmatrix} = \begin{bmatrix} TAT^\top & TAS^\top \\ SAT^\top & SAS^\top \end{bmatrix}, \tag{2.6}$$

for any given $x \in \mathbb{R}^{2n}$, where the matrix A can be chosen to be diagonal. This representation is unique if $\mathbb{H}(x)$ has n distinct nonzero eigenvalues, and if the diagonal elements of A are required to be in some prescribed order.

If we do not require that the matrix A be diagonal, then we could replace S, T, and A by SV, TV and $V^{-1}AV$, respectively, where V can be any orthogonal $n \times n$ matrix. This fact is used in the lemma below.

Example 3. Let M be an $m \times n$ matrix of rank $m \le n$, and let $f : \mathbb{R}^m \to \mathbb{R}^n$ be real analytic. Assuming $Mf = 0$, the equation $\dot{q} = f(Mq)$ defines a flow on $\mathbb{R}^n$ that is linear in time: $q(t) = q_0 + f(Mq_0)$. This is similar to the flow considered in [11, Lemma 5]. As every flow on $\mathbb{R}^n$, it extends to a Hamiltonian flow on $\mathbb{R}^{2n}$. The Hamiltonian is $H(q,p) = p^\top f(Mq)$. Using Theorem 1.5, is is easy to check that H is a shear Hamiltonian. In fact, H can be trivialized explicitly: If we set $T = M^\top(MM^\top)^{-1}M$ and $S = \mathrm{I} - T$, then (2.5) defines a matrix $U \in \mathrm{U}(n)$, and we get $(H \circ U)(q,p) = H(p,p)$.

Definition 2.5. *We say that $\mathbb{H}(x)$ is in semi-normal form if $\mathcal{Z}(x) = 0$ and $\mathcal{B}(x) = 0$.*

Lemma 2.6. *Let H be an affine-integrable Hamiltonian and $x \in \mathbb{R}^{2n}$. If $\mathcal{A}(x)$ is nonsingular then $U^\top \mathbb{H}(x) U$ is in semi-normal form for the matrix*

$$U = \exp \begin{bmatrix} 0 & \theta \\ -\theta & 0 \end{bmatrix}, \qquad \theta(x) = \tan^{-1}\bigl(\zeta(x)\bigr), \qquad \zeta(x) = \mathcal{Z}(x)\mathcal{A}(x)^{-1}. \tag{2.7}$$

Proof. Define $\zeta = \zeta(x)$ as above. A comparison with (2.6) shows that $\zeta = TS^{-1}$. The conditions in (2.5) on S and T imply that ζ is symmetric, and that $S^\top(\mathrm{I} + \zeta^2)S = \mathrm{I}$. Since $SV = |S|$ for some orthogonal matrix V, we can choose S to be a positive definite symmetric matrix. The choice is then unique: $S = (\mathrm{I} + \zeta^2)^{-1/2}$. Setting $\theta = \tan^{-1}\bigl(\zeta\bigr)$ we obtain $S = \cos(\theta)$ and $T = \zeta S = \sin(\theta)$, which leads to the expression (2.7) for U. **QED**

This offers another way of checking whether H is a shear Hamiltonian. First, we note that a nondegenerate affine-integrable Hamiltonian H is regular: $X(x)$ belongs to the range of $DX(x)$ at every point x where $DX(x)$ has rank n, since $DX(x)X(x) = 0$ by Lemma 2.1.

Lemma 2.7. *Let H be an affine-integrable Hamiltonian. Assume that $\mathcal{A}(x_0)$ is nonsingular at some point x_0. Then H is a shear if and only if ζ is constant near x_0.*

Proof. First, assume that ζ is constant near x_0. So near x_0, the matrix U in Lemma 2.6 is independent of x, and $H \circ U$ is in semi-normal form for a fixed unitary matrix U. By analyticity, this property extends to all $x \in \mathbb{R}^{2n}$. Furthermore, $\mathbb{H}(x_0)$ has rank n, as (2.6) shows, implying that H is regular. So H is a shear Hamiltonian by Lemma 2.2.

Conversely, assume that $(H \circ U)(q,p)$ is independent of q for some linear symplectic matrix U. Then U can in fact be chosen unitary, as was shown the proof of Theorem 1.5. If S and T are the real and imaginary parts of U, as defined by (2.5), then we have $\mathcal{Z}(x)\mathcal{A}^{-1}(x) = TS^{-1}$ at every point x where $\mathcal{A}(x)$ is nonsingular. **QED**

For completeness, let us mention that there is an alternative representation of $\mathbb{H}(x)$ via the shear map $(q,p) \mapsto (q, p+\zeta q)$. Assuming that H is affine-integrable and $\mathcal{A} = \mathcal{A}(x)$ nonsingular,

$$\mathbb{H}(x) = \begin{bmatrix} \mathrm{I} & \zeta \\ 0 & \mathrm{I} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \mathcal{A} \end{bmatrix} \begin{bmatrix} \mathrm{I} & 0 \\ \zeta & \mathrm{I} \end{bmatrix} = \begin{bmatrix} \zeta\mathcal{A}\zeta & \zeta\mathcal{A} \\ \mathcal{A}\zeta & \mathcal{A} \end{bmatrix}. \tag{2.8}$$

This shear is not unitary. But it is symplectic, since ζ is symmetric. Furthermore, these shear maps form a group. Notice also that $\mathcal{B}(x) = \mathcal{Z}(x)\zeta(x)$. So the condition $\mathcal{B}(x) = 0$ in Definition 2.5 is redundant if H is affine-integrable and $\mathcal{A}(x)$ nonsingular.

3. Invariant affine subspaces

In this section we give a proof of Theorem 1.2 and some related results. It is always assumed that H is affine-integrable and real analytic.

Besides the flow Φ for the Hamiltonian H, consider also the flows Ψ_j for the Hamiltonians $\partial_j H$, where $\partial_j H$ denotes the j-th partial derivative of H. By standard ODE results, $\Psi_j^t(x)$ is well defined for all times t in some open neighborhood of zero in $\mathbb{C}$ (which may depend on x). By (2.2) the flows Ψ_j commute with each other and with Φ. So the flow Ψ^w for $\partial_w H$ is given by

$$\Psi^{tw} = \Psi_1^{tw_1} \circ \Psi_2^{tw_2} \circ \cdots \circ \Psi_{2n}^{tw_{2n}}. \tag{3.1}$$

Again, $\Psi^w(x)$ is well defined for all w in some open ball $B(x) \subset \mathbb{C}^{2n}$ centered at the origin. Furthermore, the "group property" $\Psi^u(\Psi^w(x)) = \Psi^{u+w}(x)$ holds whenever $w, w+u \in B(x)$ and $u \in B(\Psi^w(x))$.

Lemma 3.1. *Let $u \in \mathbb{R}^{2n}$. If the derivative of $\partial_u H$ vanishes at some point x, then it vanishes at $\Psi^w(x)$ for every $w \in B(x)$.*

Proof. Define $(G)_t = G \circ \Psi^{tw}$ for any function G on $\mathbb{R}^{2n}$. Let now $G = \partial_u H$. Then

$$\begin{aligned} 0 &= \bigl(\partial_j\{G, \partial_w H\}\bigr)_t = \bigl(\{\partial_j G, \partial_w H\} + \{G, \partial_j \partial_w H\}\bigr)_t \\ &= \{(\partial_j G)_t, \partial_w H\} + \{G, (\partial_j \partial_w H)_t\}. \end{aligned} \tag{3.2}$$

Here, we have used that $\partial_w H$ and G are invariant under the flow Ψ^w, and that the maps Ψ^{tw} are symplectic. Thus, we have

$$\frac{d}{dt}(\partial_j G)_t = \{(\partial_j G)_t, \partial_w H\} = -\{G, (\partial_j \partial_w H)_t\} = -\sum_{\sigma,\tau} (\partial_\sigma G) \mathrm{J}_{\sigma,\tau} \partial_\tau (\partial_j \partial_w H)_t. \tag{3.3}$$

Due to the factors $\partial_\sigma G$ that all vanish at x, the value $(\partial_j G)_t(x)$ is independent of t and thus $(\partial_j G)\bigl(\Psi^{tw}(x)\bigr) = (\partial_j G)_t(x) = (\partial_j G)(x) = 0$, for all t in some open neighborhood of zero. The assertion now follows from the above-mentioned group property of Ψ and the analyticity of G. **QED**

Corollary 3.2. *Let $u \in \mathbb{R}^n$. If $\mathbb{H}u$ vanishes at some point x, then $\mathbb{H}u$ vanishes at $\Psi^w(x)$ for every $w \in B(x)$. In other words, the null space (and thus the range) of $\mathbb{H}$ is invariant under Ψ^w.*

Notice that the same holds for $DX = \mathrm{J}\mathbb{H}$.

Corollary 3.3. *Let $x \in \mathbb{R}^{2n}$ and $R(x) = \mathrm{range}(DX(x))$. Then $\Psi^w(x)$ belongs to the affine space $x + R(x)$ for all $w \in B(x)$. Furthermore, $w \mapsto \Psi^w(x)$ is locally (near zero) invertible as a map from $\mathrm{J}R(x)$ to $x + R(x)$.*

Proof. Consider the curve $u(t) = \Psi^{tw}(x) - x$. Clearly $u(0)$ belongs to $R(x)$. The derivative $u'(t) = DX(\Psi^{tw}(x))w$ belongs to the range of $DX(\Psi^{tw}(x))$, which agrees with $R(x)$ by Corollary 3.2. Thus, $u(t)$ belongs to $R(x)$ whenever $tw \in B(x)$. Since $\mathbb{H}(x)$ is symmetric, $DX(x)$ is invertible as a map from $\mathrm{J}R(x)$ to $R(x)$. Thus, by the implicit function theorem, the same holds locally (near zero) for the map $w \mapsto \Psi^w(x) - x$, whose derivative at $w = 0$ is $DX(x)$. **QED**

Proof of Theorem 1.2. For each y in $x + R(x)$ there exists an open neighborhood B_y of y in $x + R(x)$ that is included in the range of $f_y : w \mapsto \Psi^w(y)$. This follows from Corollary 3.3. The vector field X is constant on each B_y since each component X_j is invariant under the flow Ψ^w. Similarly for H. Furthermore, the open sets B_y cover the affine space $x + R(x)$, and since this space is connected, it follows that X and H are constant on $x + R(x)$.

Assume now that $DX(x)$ has rank n. If $y = x + u + v$, with $u \in R(x)$ and $v \in R(x)^\perp$, then

$$X(y) = X(x) + DX(x + u)v + \mathcal{O}\bigl(|v|^2\bigr) . \tag{3.4}$$

If $y \neq x$ is sufficiently close to x then $|DX(x + u)v|$ is bounded from below by a positive constant times $|v|$, so we have $X(y) = X(x)$ if and only if $y - x = u \in R(x)$. **QED**

4. Proof of Theorem 1.8

We will write the given nilpotency condition on DX in the form

$$\mathbb{H}(x'')\mathrm{J}\mathbb{H}(x)\mathrm{J}\mathbb{H}(x') = 0 , \qquad x'', x, x' \in \mathbb{R}^{2n} . \tag{4.1}$$

It is straightforward to check that this condition is necessary for H to be linear-symplectically conjugate to a Hamiltonian of the form (1.4).

Assume now that H is a nondegenerate (and thus regular) real analytic affine-integrable Hamiltonian that satisfies (4.1). Consider a point $x_0 = (q_0, p_0)$ where $\mathbb{H}(x_0)$ has rank

n. By performing a unitary change of variables, if necessary, we may assume that $\mathbb{H}(x_0)$ is in semi-normal form,

$$\mathbb{H}(x_0) = \begin{bmatrix} 0 & 0 \\ 0 & \mathcal{A}(x_0) \end{bmatrix} . \tag{4.2}$$

In the case where $x' = x'' = x_0$, the property (4.1) implies that $D_q^2 H = 0$, so that

$$\mathbb{H}(x) = \begin{bmatrix} 0 & \mathcal{Z}(x) \\ \mathcal{Z}(x)^\top & \mathcal{A}(x) \end{bmatrix} , \qquad x \in \mathbb{R}^{2n} . \tag{4.3}$$

Thus, our Hamiltonian H has to be of the form

$$H(q,p) = \mathcal{K}(p) + q^\top \mathcal{V}(p) , \tag{4.4}$$

with $\mathcal{K} : \mathbb{R}^n \to \mathbb{R}$ and $\mathcal{V} : \mathbb{R}^n \to \mathbb{R}^n$ real analytic. Another consequence of (4.1) is that the range of

$$\mathrm{J}\mathbb{H}(x)\mathrm{J}\mathbb{H}(x') = \begin{bmatrix} \mathcal{Z}(x)^\top \mathcal{Z}(x')^\top & \mathcal{Z}(x)^\top \mathcal{A}(x') - \mathcal{A}(x)\mathcal{Z}(x') \\ 0 & \mathcal{Z}(x)\mathcal{Z}(x') \end{bmatrix} \tag{4.5}$$

is contained in the null space of $\mathbb{H}(x'')$, for every x''. If we take $x'' = x_0$ then this implies that $\mathcal{Z}(x)\mathcal{Z}(x')$ vanishes. But $\mathcal{Z}(q,p) = D\mathcal{V}(p)$ and thus

$$D\mathcal{V}(p) D\mathcal{V}(p') = 0 , \qquad p, p' \in \mathbb{R}^n . \tag{4.6}$$

Finally, multiplying (4.5) on the left by $\mathbb{H}(x'')$ and using that the result has to be the zero matrix, we find that $\mathcal{Z}(x'')^\top \mathcal{A}(x)\mathcal{Z}(x')$ vanishes. In particular,

$$D\mathcal{V}(p'')^\top \mathbb{K}(p) D\mathcal{V}(p') = 0 , \qquad p'', p, p' \in \mathbb{R}^n , \tag{4.7}$$

where $\mathbb{K} = \nabla D\mathcal{K}$ is the Hessian of $\mathcal{K}$.

Let $\mathcal{R}$ be the linear span of all vectors $\mathcal{V}(p)$, with $p \in \mathbb{R}^n$. Or equivalently, $\mathcal{R}$ is the linear span of all vectors $D\mathcal{V}(p)u$ with $p, u \in \mathbb{R}^n$. Here we have used that $\mathcal{V}(p_0) = 0$. Then (4.6) and (4.7) imply that

$$\mathcal{V}(p+v) = \mathcal{V}(p) , \qquad \mathcal{K}(p+v) = \mathcal{K}(p) + D\mathcal{K}(p)v , \qquad v \in \mathcal{R} . \tag{4.8}$$

Let d be the dimension of $\mathcal{R}$. If $d = 0$ then there is nothing left to prove. Consider now the case where $d > 0$. Since $\mathcal{V}(p) = 0$ for all $p \in \mathcal{R}$, we also have $d < n$. Let $\mathcal{R}^\perp$ be the orthogonal complement of $\mathcal{R}$ in $\mathbb{R}^n$. Notice that $q^\top \mathcal{V}(p)$ vanishes whenever $q \in \mathcal{R}^\perp$.

Next we apply a unitary change of variables. To simplify the description, we rename current quantities by adding a subscript "old". The change of variables is $(q_{\rm old}, p_{\rm old}) = (Sq, Sp)$, with S orthogonal, such that $\mathcal{R} = S^{-1}\mathcal{R}_{\rm old}$ is the span of all vectors $P = (p_1, \ldots, p_d, 0, \ldots, 0)$. Then $\mathcal{R}^\perp$ is the span of all vectors $\bar{p} = (0, \ldots, 0, p_{d+1}, \ldots, p_n)$. And $\mathcal{V} = S^{-1}\mathcal{V}_{\rm old}S$ takes values in $\mathcal{R}$. From (4.8) we see that $\mathcal{V}(p)$ does not depend on P, and that $\mathcal{K}(p)$ is an affine function of P. Setting $Q = (q_1, \ldots, q_d, 0, \ldots, 0)$, the new Hamiltonian H is of the form (1.4). This concludes the proof of Theorem 1.8.

5. Commutators and quadratic functions

The main goal in this section is to give a proof of Theorem 1.9 on quartic Hamiltonians. But some of the observations and computations apply to other Hamiltonians as well.

The Hessian $\mathbb{H}(u)$ at any point $u \in \mathbb{R}^{2n}$ defines a quadratic function $x \mapsto \frac{1}{2}x^\top\mathbb{H}(u)x$. The flow generated by this function is linear in time, since $DX(u)^2 = 0$. It is useful to know how these flows for different vectors u are related, as the hypotheses in Theorem 1.5 and Theorem 1.8 show. What simplifies the situation for homogeneous quartic Hamiltonians is that $x^\top\mathbb{H}(u)x = u^\top\mathbb{H}(x)u = \partial_u^2 H(x)$, and if H is affine-integrable, then $\partial_u^2 H$ commutes with H by (2.2). This fact will be exploited below.

First we note that the Poisson bracket of two homogeneous quadratic functions

$$F(x) = \tfrac{1}{2}x^\top\mathcal{F}x\,, \qquad G(x) = \tfrac{1}{2}x^\top\mathcal{G}x\,, \tag{5.1}$$

is again a homogeneous quadratic function,

$$\{F,G\}(x) = \tfrac{1}{2}x^\top\mathcal{E}x\,, \qquad \mathrm{J}\mathcal{E} = \tfrac{1}{2}(\mathrm{J}\mathcal{F})(\mathrm{J}\mathcal{G}) - \tfrac{1}{2}(\mathrm{J}\mathcal{G})(\mathrm{J}\mathcal{F})\,. \tag{5.2}$$

Here, $\mathcal{F}$, $\mathcal{G}$, and $\mathcal{E}$ are symmetric $2n \times 2n$ matrices. The corresponding matrices $\mathrm{J}\mathcal{F}$, $\mathrm{J}\mathcal{G}$, and $\mathrm{J}\mathcal{E}$ belong to $\mathrm{sp}(2n,\mathbb{R})$.

We also need to compute some double commutators, and not all functions involved are quadratic. To simplify the expressions, we use the operator notation $[\![H]\!]F = \{F,H\}$. A straightforward computation shows that

$$\begin{aligned}\partial_u\partial_v[\![H]\!] &= [\![\partial_v H]\!]\partial_u + [\![\partial_u H]\!]\partial_v + [\![H]\!]\partial_u\partial_v + [\![\partial_u\partial_v H]\!]\\ &= \partial_u[\![\partial_v H]\!] + [\![\partial_u H]\!]\partial_v + [\![H]\!]\partial_u\partial_v\,.\end{aligned} \tag{5.3}$$

Let now F be is a polynomial of degree ≤ 2. Then $\partial_u\partial_v F$ is constant and thus commutes with every function. If in addition F commutes with H, then (5.3) yields

$$\begin{aligned}-[\![\partial_u\partial_v H]\!]F &= [\![\partial_v H]\!]\partial_u F + [\![\partial_u H]\!]\partial_v F\,,\\ 0 &= \partial_u[\![\partial_v H]\!]F + [\![\partial_u H]\!]\partial_v F\,.\end{aligned} \tag{5.4}$$

So far we have not used any properties of H other than differentiability.

Proof of Theorem 1.9. Assume now that H is homogeneous of degree 4. Let F be a polynomial of degree ≤ 2 that commutes with H. Then the second identity in (5.4) implies that

$$[\![\partial_{v'}H]\!]\partial_{u'}[\![\partial_v H]\!]\partial_u F = -[\![\partial_{v'}H]\!][\![\partial_{u'}H]\!]\partial_v\partial_u F = 0\,. \tag{5.5}$$

Assume in addition that H is affine-integrable. Then H satisfies (2.2). By Jacobi's identity for the Poisson bracket, if F and G commute with H, then so does $[\![G]\!]F$. In particular, $[\![\partial_u\partial_v H]\!]F$ commutes with H. Using (5.4) and (5.5), we find that

$$[\![\partial_{u'}\partial_{v'}H]\!][\![\partial_u\partial_v H]\!]F = \bigl([\![\partial_{v'}H]\!]\partial_{u'} + [\![\partial_{u'}H]\!]\partial_{v'}\bigr)\bigl([\![\partial_v H]\!]\partial_u + [\![\partial_u H]\!]\partial_v\bigr)F = 0\,. \tag{5.6}$$

Since H is homogeneous of degree 4, the second derivatives of H are homogeneous quadratic polynomials. They commute with H by (2.2). So as a special case of (5.6) we have

$$[\![\partial_u^2 H]\!][\![\partial_v^2 H]\!]\partial_u^2 H = 0\,. \tag{5.7}$$

Applying (5.2) with $F = \partial_u^2 H$ and $G = \partial_v^2 H$ yields $\mathrm{J}\mathcal{E} = 2\mathrm{J}\mathbb{H}(u)\mathrm{J}\mathbb{H}(v) - 2\mathrm{J}\mathbb{H}(v)\mathrm{J}\mathbb{H}(u)$. According to (5.7) we have $\mathrm{J}\mathcal{E}\mathrm{J}\mathbb{H}(u) - \mathrm{J}\mathbb{H}(u)\mathrm{J}\mathcal{E} = 0$, which implies that

$$\mathbb{H}(u)\mathrm{J}\mathbb{H}(v)\mathrm{J}\mathbb{H}(u) = 0\,, \qquad\qquad u, v \in \mathbb{R}^{2n}\,. \tag{5.8}$$

Here we have used that $\mathbb{H}(u)\mathrm{J}\mathbb{H}(u) = 0$. The identity (5.8) also be written as

$$D^3X(u,u,w)^\top \mathbb{H}(v) D^3X(u,u,z) = 0\,, \qquad\qquad u, v, w, z \in \mathbb{R}^{2n}\,. \tag{5.9}$$

By Lemma 5.1 below, the same holds if each of the arguments u is replaced by a different vector in $\mathbb{R}^{2n}$. In particular, we have (4.1) and thus $DX(x'')DX(x)DX(x') = 0$ for all $x'', x, x' \in \mathbb{R}^{2n}$. The assertion now follows from Theorem 1.8. **QED**

Lemma 5.1. *Let X, Y, Z be vector spaces. Let $\langle \ldots \rangle : X^3 \to Y$ be a symmetric cubic form and $\odot : Y^2 \to Z$ be a symmetric quadratic form. Assume that $\langle p,p,u\rangle \odot \langle p,p,v\rangle = 0$ for all $p, u, v \in X$. Then $\langle u_1,u_2,u_3\rangle \odot \langle v_1,v_2,v_3\rangle = 0$ for any $u_i, v_j \in X$.*

Proof. Under the given assumption we have $\langle p,p,p\rangle \odot \langle p,p,p\rangle = 0$. Now "differentiate" this identity twice: Replace p by $p + u + v$, expand, and then collect all terms that are bilinear in (u, v). The result is

$$12\langle p,u,v\rangle \odot \langle p,p,p\rangle + 18\langle p,p,u\rangle \odot \langle p,p,v\rangle = 0\,. \tag{5.10}$$

By assumption, the second term vanishes, so $\langle p,u,v\rangle \odot \langle p,p,p\rangle = 0$. Differentiating this identity once we get

$$\langle w,u,v\rangle \odot \langle p,p,p\rangle + 3\langle p,u,v\rangle \odot \langle p,p,w\rangle = 0\,. \tag{5.11}$$

Similarly, differentiating $\langle p,p,u\rangle \odot \langle p,p,v\rangle = 0$ once yields

$$2\langle p,w,u\rangle \odot \langle p,p,v\rangle + 2\langle p,p,u\rangle \odot \langle p,w,v\rangle = 0\,. \tag{5.12}$$

Now (5.11) can be used to rewrite (5.12) as

$$-\frac{2}{3}\langle v,w,u\rangle \odot \langle p,p,p\rangle - \frac{2}{3}\langle p,p,p\rangle \odot \langle u,w,v\rangle = 0\,. \tag{5.13}$$

Or simplified, $\langle p,p,p\rangle \odot \langle u,v,w\rangle = 0$. The assertion now follows by polarization. **QED**

6. The example from Theorem 1.6

Consider the Hamiltonian (1.4), with the variables $(Q,\bar{q};P,\bar{p})$ renamed to $x=(q,y;p,z)$,

$$H(x)=K(z)+q^\top V(z)+p^\top W(z)\,, \tag{6.1}$$

where $q,p\in\mathbb{R}^d$ and $y,z\in\mathbb{R}^{n-d}$. The corresponding vector field $X=(\dot{q},\dot{y};\dot{p},\dot{z})$ is given by $\dot{z}=0$ and

$$\dot{q}=W(z)\,,\quad \dot{p}=-V(z)\,,\quad \dot{y}_j=\partial_j K(z)+q^\top\bigl[\partial_j V(z)\bigr]+p^\top\bigl[\partial_j W(z)\bigr]\,. \tag{6.2}$$

Since $\dot{z}=0$, the Hamiltonian H is affine-integrable if and only if

$$\begin{aligned}\ddot{y}_j&=\dot{q}^\top\bigl[\partial_j V(z)\bigr]+\dot{p}^\top\bigl[\partial_j W(z)\bigr]\\&=W(z)^\top\bigl[\partial_j V(z)\bigr]-V(z)^\top\bigl[\partial_j W(z)\bigr]\end{aligned} \tag{6.3}$$

is equal to zero for all j. This is precisely the condition (1.5). According to Theorem 1.5, H is a shear if and only if $DX(x)DX(x')$ vanishes for all x and x'. Or equivalently, if and only if

$$-X(x)^\top \mathrm{J}X(x')=W(z)^\top V(z')-V(z)^\top W(z') \tag{6.4}$$

vanishes for all x and x'.

It should be noted that the case $d=1$ is trivial: If $v(s)=V_1(x+s(x'-x))$ and $w(s)=W_1(x+s(x'-x))$ satisfy $wv'-vw'=0$, then by the quotient rule of differentiation, the functions v and w are constant multiples of each other. So (6.4) follows from (6.3). In this case, $H(q,p)$ can be made either independent of q via a change of variables $(q,p)\mapsto(q,p+cq)$, or independent of p via a change of variables $(q,p)\mapsto(q+cp,p)$. Thus, H is a shear Hamiltonian if $d=1$.

Notice also that, if the right hand side of (6.3) is equal to zero, then it remains zero if ∂_j is replaced by ∂_j^2. Thus, the right hand side of (6.4) is of the order $|z-z'|^3$. This has motivated our choice of V and W below.

The Hamiltonian (1.3) can be written as

$$H(q,y;p,z)=q_1V_1(z)+q_2V_2(z)+p_1W_1(z)+p_2W_2(z)\,, \tag{6.5}$$

where $q,y,p,z\in\mathbb{R}^2$ and

$$V_1(z)=z_1^3\,,\quad W_1(z)=z_2^3\,,\quad V_2(z)=\sqrt{3}z_1^2z_2\,,\quad W_2(z)=-\sqrt{3}z_1z_2^2\,.$$

Let us compute the right hand side of (6.4), with z' replaced by w in order to simplify notation. If $w_2=z_2$ then

$$\begin{aligned}\bigl[W_1(z)V_1(w)+W_2(z)V_2(w)\bigr]&-\bigl[V_1(z)W_1(w)+V_2(z)W_2(w)\bigr]\\&=z_2^3w_1^3-3z_1z_2^2w_1^2z_2-z_1^3z_2^3+3z_1^2z_2w_1z_2^2\\&=\bigl(w_1^3-3z_1w_1^2-z_1^3+3z_1^2w_1\bigr)z_2^3\\&=(w_1-z_1)^3z_2^3\,.\end{aligned} \tag{6.6}$$

This is clearly nonzero at some points, so H cannot be a shear. On the other hand,

$$\big[W_1(z)\partial_j V_1(z) + W_2(z)\partial_j V_2(z)\big] - \big[V_1(z)\partial_j W_1(z) + V_2(z)\partial_j W_2(z)\big] = 0 \tag{6.7}$$

holds for $j = 1$, due to the factor $(w_1 - z_1)^3$ in (6.6). By symmetry, we have an expression analogous to (6.6) if $w_1 = z_1$, with the cubic factor being $(w_2 - z_2)^3$. So (6.7) holds for $j = 2$ as well. Thus, H is a shear Hamiltonian, as claimed in Theorem 1.6.

A straightforward computation shows that the Hamiltonian (6.5) is nondegenerate: The Hessian $\mathbb{H}(x)$ has rank $n = 4$ whenever $z_1 z_2 \neq 0$. Another noteworthy fact is that the matrix $\zeta(x)$ defined in (2.7) depends on x only via the ratio z_1/z_2. But the dependence is nontrivial, so by Lemma 2.7, this shows again that H cannot be a shear Hamiltonian.

7. Elementary factorization

A classical theorem by Jung [1] asserts that the group (under composition) of polynomial maps of the plane $\mathbb{R}^2$ is generated by affine maps and elementary shears $(q,p) \mapsto (q + s(p), p)$. No general result of this type is known in dimensions higher than 2. Theorem 1.10 covers the special case of symplectic maps $F = \mathrm{I} + X$, with X homogeneous of degree 3. Its proof is based on the following observation.

Lemma 7.1. *Let H_0 be a polynomial affine-integrable Hamiltonian of the from (1.4). Write $H_0 = H_1 + H_2$, where $H_1 = Q^\top V(\bar{p})$ and $H_2 = K(\bar{p}) + P^\top W(\bar{p})$. Then H_1 and H_2 Poisson-commute with $H_3 = \frac{1}{2}\{H_1, H_2\}$. Furthermore, H_1 and $H_4 = H_2 - H_3$ are shear Hamiltonians, and the corresponding time-one maps satisfy*

$$\Phi^1_{H_0} = \Phi^1_{H_1} \circ \Phi^1_{H_4}\,. \tag{7.1}$$

Proof. A straightforward calculation yields $H_3 = \frac{1}{2} V(\bar{p})^\top W(\bar{p})$. So each of the Hamiltonians H_j is of the form (1.4). Here, and in what follows, $0 \le j \le 4$. In addition, H_j satisfies the affine-integrability condition (1.5). Thus, the adjoint map $[\![H_j]\!] : G \mapsto \{G, H_j\}$ has the following nilpotency property: If f is any polynomial, then $[\![H_j]\!]^k f = 0$ for sufficiently large k. Furthermore, H_j commutes with H_3, since H_3 only depends on the variable $\bar{p}$, while H_j is independent of the variable $\bar{q}$. Thus, by the Baker-Campbell-Hausdorff formula,

$$e^{t[\![H_0]\!]} f = e^{t[\![H_1]\!]} e^{t[\![H_2]\!]} e^{-t^2[\![H_3]\!]} f = e^{t[\![H_1]\!]} e^{t[\![H_2 - tH_3]\!]} f\,, \tag{7.2}$$

for every polynomial f. To be more precise, (7.2) is an identity for formal power series. But due to the above-mentioned nilpotency property, only finitely many terms of the series are nonzero. So (7.2) holds as an identity between polynomials. Using that $f \circ \Phi^t_{H_j} = e^{t[\![H_j]\!]} f$ for any polynomial f, we obtain (7.1) from (7.2).

Let $j \ge 1$. Then the vector field $X_j = \mathrm{J}\nabla H_j$ satisfies $X_j(x)^\top \mathrm{J} X_j(x') = 0$ for all x and all x', as can be seen from (6.4). This shows that H_j is a shear Hamiltonian. **QED**

Remark 4. The time-one map for H_4 is an elementary shear, $\Phi^1_{H_4}(q,p) = (q + \nabla h_4(p), p)$, where $h_4(p) = H_4(p,p)$. The time-one map for H_1 is unitarily conjugate to an elementary shear $S_1(q,p) = (q + \nabla h_1(p), p)$. A straightforward computation, similar to the one in Example 3, shows that $h_1(p) = H_1(p,p)$.

Proof of Theorem 1.10. Let F be a symplectic map on $\mathbb{R}^{2n}$ such that $X = F - \mathrm{I}$ is a homogeneous polynomial of degree $m \geq 2$. First, we prove that $X = \mathrm{J}\nabla H$ for some affine-integrable Hamiltonian H. The symplecticity condition (1.1) implies that X satisfies the equation (2.1) for $t = 1$. In this equation, the terms in square brackets have to vanish separately, since they have different degrees of homogeneity. The first of the resulting identities implies that the derivative of $\mathrm{J}X$ is a symmetric matrix. Thus, by the Poincaré Lemma, $\mathrm{J}X$ is the gradient of a function $-H$. The second identity implies that $(DX)^2 = 0$. Thus, $(DX)X = 0$, since $X(x) = m^{-1}DX(x)x$ by homogeneity. This shows that the flow for X is linear in time. In conclusion, F is the time-one map of an affine-integrable Hamiltonian H that is homogeneous of degree $m + 1$.

Consider now $m = 3$, and assume that X is nondegenerate. By Theorem 1.9, there exists a linear symplectic map U on $\mathbb{R}^{2n}$, such that $H_0 = H \circ U^{-1}$ is a Hamiltonian of the form (1.4). In fact, U is unitary, as seen in Section 4. Using Lemma 7.1, we have

$$F = \Phi^1_H = U^{-1} \circ \Phi^1_{H_0} \circ U = U^{-1} \circ \Phi^1_{H_1} \circ \Phi^1_{H_4} \circ U = \Phi^1_{H_1 \circ U} \circ \Phi^1_{H_4 \circ U}\,, \tag{7.3}$$

where H_1 and H_4 (and thus $H_1 \circ U$ and $H_4 \circ U$) are shear Hamiltonians. **QED**

References

[1] H.W.E. Jung, *Über ganze birationale Transformationen der Ebene*, J. Reine Angew. Math. **184** 161-174 (1942).

[2] A.J. Dragt and J.M. Finn, *Lie series and invariant functions for analytic symplectic maps*, J. Math. Phys. **17** 2215-2227 (1976).

[3] S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces*, Academic Press, New York (1978).

[4] A.J. Dragt and D.T. Abell, *Jolt Factorization of Symplectic Maps*, Int. J. Mod. Phys. A (Proc. Suppl.) **2B**, 1019-1021 (1993).

[5] J. Moser, *On quadratic symplectic mappings*, Math. Z. **216**, 417-430 (1994).

[6] K. Feng and D.-L. Wang, *Variations on a theme by Euler*, J. Comput. Math. **16**, 97-106 (1998).

[7] M. de Bondt, A. van den Essen, *Singular Hessians*, J. Algebra **282**, 195-204 (2004).

[8] H.E. Lomelí, *Symplectic homogeneous diffeomorphisms, Cremona maps and the jolt representation*, Nonlinearity **18**, 1065-1071 (2005).

[9] E. Forest, *Geometric integration for particle accelerators*, J. Phys. A: Math. Gen. **39**, 5321-5377 (2006)

[10] M. de Bondt, *Quasi-translations and counterexamples to the Homogeneous Dependence Problem*, Proc. Amer. Math. Soc. **134**, 2849-2856 (2006).

[11] R.I. McLachlan, H.Z. Munthe-Kaas, G.R.W. Quispel, A. Zanna, *Explicit Volume-Preserving Splitting Methods for Linear and Quadratic Divergence-Free Vector Fields*, Found. Comput. Math. **8** 335-355 (2008).