

Deformational rigidity of integrable metrics on the torus

JOSCHA HENHEIK 

*Institute of Science and Technology Austria, Klosterneuburg 3400, Austria
(e-mail: joscha.henheik@ist.ac.at)*

(Received 13 February 2024 and accepted in revised form 29 May 2024)

Abstract. It is conjectured that the only integrable metrics on the two-dimensional torus are Liouville metrics. In this paper, we study a deformative version of this conjecture: we consider integrable deformations of a non-flat Liouville metric in a conformal class and show that for a fairly large class of such deformations, the deformed metric is again Liouville. The principal idea of the argument is that the preservation of rational invariant tori in the foliation of the phase space forces a linear combination on the Fourier coefficients of the deformation to vanish. Showing that the resulting linear system is non-degenerate will then yield the claim. Since our method of proof immediately carries over to higher dimensional tori, we obtain analogous statements in this more general case. To put our results in perspective, we review existing results about integrable metrics on the torus.

Key words: Liouville metrics, deformational rigidity, geodesic flow, weak KAM theory
2020 Mathematics Subject Classification: 37J35, 37J39 (Primary); 53D25, 37C83, 37J40, 70H06 (Secondary)

1. Introduction

Let $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ be the two-dimensional torus being equipped with a C^2 -smooth global Liouville metric g , that is, having line element

$$ds^2 = (f_1(x^1) + f_2(x^2))((dx^1)^2 + (dx^2)^2), \quad (1.1)$$

where $(x^1, x^2) \in \mathbb{T}^2$ are the standard periodic coordinates and $f_1, f_2 \in C^2(\mathbb{T})$ are positive Morse functions or positive constants and thus ‘non-degenerate’. (Recall that Morse functions on a manifold M are characterized by having no degenerate critical points. They form a dense and open set in $C^2(M)$ and are thus ‘generic’.) The corresponding geodesic flow (see §1.1) is well known to be integrable and a longstanding folklore conjecture says that *Liouville metrics are the only integrable metrics on \mathbb{T}^2* . We emphasize that, in this context, integrability always allows for singularities in the foliation of the phase space of the naturally associated Hamiltonian system, which is made precise in Definition 1.2 below.



Although the validity of the folklore conjecture appeared conceivable for a long time, there is strong indication for it being false in its very general form, as shown in [35]: here, the authors constructed a Hamiltonian counterexample which is *locally* integrable in a p -cone in the cotangent bundle. This means that, on a fixed energy level, there exists an analytic change of variables, transforming the Hamiltonian with non-Liouville potential to the standard form $(p_1^2 + p_2^2)/2$ but only for p_i in a certain cone in \mathbb{R}^2 (see also Theorem 3.8 below for a more precise statement). However, despite this delicate example, certain suitably weakened conjectures are still believed to be true, which is supported by a variety of partial results obtained in this direction, starting from classical ones by Dini [43], Darboux [37], and Birkhoff [22] and further developed in [11, 62, 64]. In particular, several works by Bialy, Mironov [13, 17–19], Denisova, Kozlov, Treshev [39–42, 70], Mironov [79], and others [2, 11, 64, 87] strongly indicate the validity of the following (yet unproven) conjecture: *Every polynomially integrable metric g on \mathbb{T}^2 is of Liouville type.* We refer to §3 for details. (See also [24, 31] for recent surveys on open problems and questions concerning geodesics and integrability of finite-dimensional systems.)

In this paper, we are concerned with a *perturbative* version of the folklore conjecture: Let $(g_\varepsilon)_{|\varepsilon| \leq \varepsilon_0}$ for some small $\varepsilon_0 > 0$ be a family of perturbations of $g \equiv g_0$ in the same conformal class (note that on the torus, there exist global isothermal coordinates [26, Ch. 11]) having line-element

$$ds_\varepsilon^2 = (f_1(x^1) + f_2(x^2) + \varepsilon\lambda(x^1, x^2))((dx^1)^2 + (dx^2)^2), \quad (1.2)$$

where $\lambda \in C^2(\mathbb{T}^2)$ is assumed to be a Morse function (or constant) and have an absolutely convergent Fourier series. We will assume that the perturbed family g_ε *remains integrable*, meaning that within the foliation of the phase space for the *unperturbed* Liouville metric in equation (1.1), the deformation in equation (1.2) preserves sufficiently many *rational invariant tori* (see Assumption (P) below for a precise formulation of our requirement on the preservation of these tori). Then we obtain that λ is necessarily *separable* in a sum of two single-valued functions, that is,

$$\lambda(x^1, x^2) = \lambda_1(x^1) + \lambda_2(x^2)$$

for some $\lambda_1, \lambda_2 \in C^2(\mathbb{T})$. Therefore, our main results formulated (somewhat informally) below assert the following.

The class of Liouville metrics is deformationally rigid under a fairly wide class of integrable conformal perturbations.

To the best of our knowledge, this is the first instance of a rigidity result for (not necessarily analytically) integrable dynamical systems allowing for singularities in the invariant foliation of the unperturbed system. The precise statements of our main results are given in Theorems 2.2, 2.3, and 2.4 in §2.

Main Results. Let g be a non-degenerate Liouville metric on \mathbb{T}^2 as in equation (1.1) and assume that the family $(g_\varepsilon)_{|\varepsilon| \leq \varepsilon_0}$ of perturbations defined in equation (1.2) remains integrable. Then we have the following.

- (i) In the case where $f_1, f_2 \equiv \text{const.}$, then λ is separable.

- (ii) In the case where $f_1 \equiv \text{const.}$, λ is a trigonometric polynomial in x^2 , and the relative difference μ_2 between f_2 and its mean $\int_{\mathbb{T}} f_2$, that is, $\mu_2 := \|f_2 - \int_{\mathbb{T}} f_2\|_{C^0} / \int_{\mathbb{T}} f_2$ is small, then λ is separable.

If, additionally, f_2 is analytic, we have that λ is separable, irrespective of the size μ_2 of the fluctuations of f_2 (but only for μ_2 outside of an exceptional (Lebesgue) null-set).

- (iii) In general, if λ is a trigonometric polynomial and the relative differences μ_i , $i = 1, 2$, between the f_i and their means $\int_{\mathbb{T}} f_i$, that is, the $\mu_i := \|f_i - \int_{\mathbb{T}} f_i\|_{C^0} / \int_{\mathbb{T}} f_i$ are small, then λ is separable.

If, additionally, f_i is analytic (for one or both $i = 1, 2$), we have that λ is separable, irrespective of the size μ_i of the fluctuations of f_i (outside of an exceptional null-set).

It is straightforward to generalize our results to higher dimensional tori $\mathbb{T}^d = \mathbb{R}^d / \mathbb{Z}^d$. To ease notation and make the presentation clearer, we only mention it here and postpone a more detailed discussion to Appendix A.

Remark 1.1. (Generalization to higher dimensions) Analogously to equation (1.1), let \mathbb{T}^d be equipped with a C^2 -smooth global Liouville metric g having line element

$$ds^2 = (f_1(x^1) + \dots + f_d(x^d))((dx^1)^2 + \dots + (dx^d)^2), \tag{1.3}$$

where $x = (x^1, \dots, x^d) \in \mathbb{T}^d$ are standard periodic coordinates and $f_i \in C^2(\mathbb{T})$ for $1 \leq i \leq d$ are positive Morse functions or constants. Again, it is easy to see that the geodesic flow is integrable. Just as in equation (1.2), we now perturb equation (1.3) in the same conformal class by some $\lambda \in C^2(\mathbb{T}^d)$ having an absolutely convergent Fourier series.

Under the assumption that the family of perturbed metrics $(g_\varepsilon)_{|\varepsilon| \leq \varepsilon_0}$ remains integrable, we have the following (somewhat informal) rigidity result.

Let $f_i \equiv \text{const.}$ for the first $0 \leq d_{\text{flat}} \leq d$ indices, and f_j be analytic for the last $0 \leq d_{\text{analyt}} \leq d - d_{\text{flat}}$ indices. Then, if λ is a trigonometric polynomial in x^k for $k \in \{d_{\text{flat}} + 1, \dots, d\}$, and the relative differences between $f_{d_{\text{flat}}+1}, \dots, f_{d-d_{\text{analyt}}}$ and their mean values are small, we have that λ is separable, irrespective of size μ_j of the fluctuations of f_j (outside of a null-set).

This result unifies and generalizes the three separate statements given above. A precise formulation is given in Theorem A.1 in Appendix A.

The present paper is not the first study on *rigidity* of important integrable systems. In [10, 60], Avila, de Simoi, Kaloshin and Sorrentino recently solved both, a deformative and a perturbative version of the famous *Birkhoff conjecture* concerning integrable billiards in two dimensions. In a nutshell, their result says that *a strictly convex domain with integrable billiard dynamics sufficiently close to an ellipse is necessarily an ellipse*. This can be viewed as an analogue of the perturbative version of the folklore conjecture formulated above [61]. More precisely, our main results concerning general $f_i \in C^2(\mathbb{T})$ are similar—in spirit—to the deformational rigidity for ellipses of *small* eccentricity (cf. f_1, f_2 in equation (1.1) having small fluctuations), which has been shown

first in [10], later extended by Huang, Kaloshin, and Sorrentino [59] to a local notion of integrability, and finally significantly improved in [66]. The overall strategy pursued in [10, 59, 60] also inspired the arguments employed in the present paper.

In a more recent work, Arnaud, Massetti, and Sorrentino [4] (replacing the earlier preprint [74]) studied the rigidity of integrable symplectic twist maps on the $2d$ -dimensional annulus $\mathbb{T}^d \times \mathbb{R}^d$. More precisely, they consider one-parameter families $(f_\varepsilon)_{\varepsilon \in \mathbb{R}}$ of symplectic twist maps $f_\varepsilon(x, p) = f_0(x, p + \varepsilon \nabla G(x))$ and prove two main rigidity results. First, in the analytic category for f_0 and the perturbation G , if a *single* rational invariant Lagrangian graph of f_ε exists for infinitely many values of ε (e.g., an interval around zero), then G must necessarily be constant. Second, if f_0 is analytic and *completely integrable* (that is, not plagued with singularities in the invariant foliation of the phase space, see [16, 86]), G is of class C^2 , and sufficiently (infinitely) many rational invariant Lagrangian graphs of f_ε persist for small $\varepsilon \neq 0$, then G must necessarily be constant. Note that in this second result, the entire phase space is foliated by invariant tori, and the perturbation solely depends on the angle variables of the dynamical system. In this sense, Theorem 2.2 can—morally—be viewed as a special case of the second result in [4] (see also [74, Theorem 2]), but Theorems 2.3 and 2.4 generalize this statement to more general functional dependencies of the perturbation. Apart from this, our general results (that is, those not concerning analytic functions f_i) do not require any regularity beyond the standard C^2 .

As mentioned above, by assuming that the family of metrics $(g_\varepsilon)_{|\varepsilon| \leq \varepsilon_0}$ remains integrable, we mean that, in particular, sufficiently many rational invariant tori in an isoenergy manifold of the Hamiltonians associated to the metric by the *Maupertuis principle* (see §1.2) are preserved. This will be made precise in Assumption (P) below. As we will show, the preservation of an (n, m) -rational invariant torus ‘annihilates’ the Fourier coefficients λ_{k_1, k_2} with indices $(k_1, k_2) \in \{(n, m)\}^\perp$ of

$$\lambda(x, y) = \sum_{(k_1, k_2) \in \mathbb{Z}^2} \lambda_{k_1, k_2} e^{2\pi i(k_1 x + k_2 y)},$$

or of the corresponding perturbing mechanical potential, denoted by U later on. We already noted that, contrary to items (ii) and (iii), the unperturbed metric in our first result is guaranteed to be *completely integrable*. Moreover, the perturbation λ depends solely on the *angular* but not the *action* coordinates of the unperturbed problem. Although the analog of this result for symplectic twist maps in this peculiar setting has already been shown in [4, 74] by methods similar to ours, we reprove it by pursuing an only slightly different but original strategy, which is suitable for certain inevitable modifications for the proofs of the more general statements under items (ii) and (iii). These two cases (corresponding to surfaces of revolution and general Liouville metrics, see §3) build on perturbative estimates for (possibly infinitely many) systems of linear equations for the Fourier coefficients. These are obtained from the first-order term of an expansion in ε , somewhat similar to the (*subharmonic*) *Melnikov potential* in the Poincaré–Melnikov method [8, 55, 91]. Establishing this expansion as well as proving that the resulting systems of linear equations are of full rank requires perturbative estimates on action-angle coordinates and certain basic objects from weak KAM theory [85]. Finally, the extension of

our results for analytic functions f_i beyond the perturbative regime are proven by exploiting the analytic dependence of the linear system on the size μ_i of the fluctuations of f_i (see Appendix C).

In the remainder of this introduction, we recall basic notions in geometry and dynamical systems, which are frequently used in this paper, and introduce the problem of classifying integrable metrics on Riemannian manifolds, in particular, the torus \mathbb{T}^2 , as formulated in Questions (Q1) and (Q2) below. In §2, we formulate our main results in Theorems 2.2, 2.3, and 2.4. In §3, we present related existing results and known partial answers on the classification problem for integrable metrics on the torus \mathbb{T}^2 (a few of which have already been mentioned above) to put our results into context. In §4, we give the proofs of our main results, and, finally, comment on possible generalizations, different approaches, and a list of open problems in §5. As already mentioned above, the precise formulation of our result for higher dimensions is given in Theorem A.1 in Appendix A. A fundamental perturbation theoretic lemma on action-angle coordinates, a concise study on important analyticity properties of these, and a brief overview of the relevant aspects of weak KAM theory are presented in three further appendices.

An extended version of this paper containing more details and background can be found at [arXiv: 2210.02961](https://arxiv.org/abs/2210.02961).

1.1. *Geodesic flow and integrability.* Let (M, g) be a (compact) C^2 -smooth n -dimensional connected Riemannian manifold without boundary equipped with the Riemannian metric $g = (g_{ij}(x))_{ij}$. Geodesics of the given metric g are defined as smooth parameterized curves $\gamma(t) = (x^1(t), \dots, x^n(t))$ that are solutions to the system of differential equations

$$\nabla_{\dot{\gamma}} \dot{\gamma} = 0, \quad (1.4)$$

where $\dot{\gamma} = (d\gamma/dt)$ denotes the velocity vector of the curve γ , and ∇ is the covariant derivative operator related to the Levi-Civita connection associated with the metric g .

It is well known that the geodesic equation in equation (1.4) can also be viewed as a Hamiltonian system on the cotangent bundle T^*M , and the geodesics γ themselves can be regarded as projections of trajectories of the Hamiltonian system onto M . Therefore, let x and p be natural coordinates on the cotangent bundle T^*M and $\omega = dx \wedge dp$ denote the standard symplectic structure. Then, the Hamiltonian function $H \in C^2(T^*M)$ is defined as

$$H(x, p) = \frac{1}{2} \sum_{ij} g^{ij}(x) p_i p_j = \frac{1}{2} |p|_g^2. \quad (1.5)$$

A trajectory $(x(t), p(t))$ is an integral curve for the Hamiltonian vector field if and only if the following Hamiltonian system of differential equations (written in local coordinates) is satisfied:

$$\begin{cases} \dot{p}_i = -\frac{\partial H}{\partial x^i}, \\ \dot{x}^i = \frac{\partial H}{\partial p_i}. \end{cases} \quad (1.6)$$

In view of this connection to Hamiltonian dynamics, it is natural to ask for a classification of Riemannian manifolds (M, g) , for which the geodesic equations in equation (1.4) can be solved explicitly. More precisely, we can formulate the following questions.

- (Q1) On which manifolds do there exist Riemannian metrics whose (co-)geodesic flow is integrable?
- (Q2) Given such a manifold, how does one characterize the class of metrics with integrable geodesic flow?

Clearly, the answers and their complexity hinge on the notion of integrability for the Hamiltonian system (see §3). In this paper, we will be concerned with the standard notion, that is, *Liouville integrability*, which we recall for the readers convenience.

Definition 1.2. The geodesic flow on (M, g) is called *Liouville integrable* if there exist n functions $F_1, \dots, F_n \in C^2(T^*M)$ (called *first integrals*) that are:

- (i) functionally independent on T^*M , that is, the vector fields $X_{F_1}(x, p), \dots, X_{F_n}(x, p)$ are linear independent in $T_{(x,p)}(T^*M)$ for all $(x, p) \in \mathcal{M} \subset T^*M$, where \mathcal{M} is some open and everywhere dense set of full measure (cf. the restriction to Morse functions);
- (ii) pairwise in involution, that is,

$$\{F_k, F_l\} := \omega(X_{F_k}, X_{F_l}) = \sum_i \left(\frac{\partial F_k}{\partial x^i} \frac{\partial F_l}{\partial p_i} - \frac{\partial F_k}{\partial p_i} \frac{\partial F_l}{\partial x^i} \right) = 0.$$

Whenever the geodesic flow on (M, g) is Liouville integrable, we call g an *integrable metric* on M . Moreover, we call the Hamiltonian system in equation (1.6) (or the corresponding Hamiltonian in equation (1.5) itself) *integrable*, whenever the associated metric g is integrable on M .

Remark 1.3. Whenever the first integrals F_1, \dots, F_n can be chosen to be functions that are polynomially in the momentum variables, the metric is often called *polynomially integrable* or *algebraically integrable*. If we aim at indicating the order of the polynomial, we speak of *linearly/quadratically/... integrable* metrics.

Remark 1.4. Note that since one can always choose $H = F_1$ as a first integral for the geodesic flow, the question of integrability for one-dimensional manifolds is completely answered. Therefore, the simplest manifolds, for which the answers to Questions (Q1) and (Q2) are non-trivial, are two-dimensional.

In this work, we are mainly concerned with a characterization of integrable metrics in the sense of Question (Q2) for the two-dimensional torus \mathbb{T}^2 . In this case, the largest known class of such metrics g are so-called *Liouville metrics*, where the line element takes the form in equation (1.1) in appropriate global coordinates (x^1, x^2) , and where f_1 and f_2 are sufficiently regular positive periodic functions. See §3.2 for more details.

1.2. *Maupertuis principle.* To approach Questions (Q1) and (Q2), we will use the *Maupertuis principle* (see, e.g., [28]): for a compact Riemannian manifold, (M, g) , let

$$H(x, p) = \frac{1}{2} \sum_{ij} g^{ij}(x) p_i p_j - V(x) \tag{1.7}$$

be a natural mechanical Hamiltonian function on T^*M , where $V \in C^2(M)$ denotes some potential function. Moreover, let $T_h = \{H(x, p) = h\}$ be an isoenergy submanifold for some $h > -\min_x V(x)$ and note that T_h is also an isoenergy submanifold for another system with Hamiltonian function

$$\tilde{H}(x, p) = \frac{1}{2} \sum_{ij} \frac{g^{ij}(x)}{h + V(x)} p_i p_j,$$

that is, $T_h = \{\tilde{H}(x, p) = 1\}$. Now, the *Maupertuis principle* states that the integral curves for the Hamiltonian vector fields X_H and $X_{\tilde{H}}$ on the fixed isoenergy submanifold T_h coincide. Moreover, if there exists an additional first integral F for H on T_h , then there also exists a first integral \tilde{F} for \tilde{H} on the *whole* of T^*M (except, potentially, at the zero section). Finally, note that the vector field $X_{\tilde{H}}$ gives rise to the geodesic flow of the Riemannian metric \tilde{g} with

$$\tilde{g}_{ij}(x) = (h + V(x))g_{ij}(x), \tag{1.8}$$

which is the correspondence between Hamiltonian systems and geodesic flows we will use.

2. Main results

The main results of this paper are rigidity results in the sense of Question (Q2) for classes of integrable metrics on the two-torus $\mathbb{T}^2 = \mathbb{R}^2 / \Gamma$, initially equipped with the flat metric, and hence obtained by a Hamiltonian defined on $T^*\mathbb{T}^2$ by means of the Maupertuis principle. In general, $\Gamma \subset \mathbb{R}^2$ is an arbitrary lattice, but we focus on the case $\Gamma = \mathbb{Z}^2$ here. We define the Hamiltonian function

$$H_0(x, p) = \frac{p_1^2}{2} + \frac{p_2^2}{2} - \mu_1 V_1(x^1) - \mu_2 V_2(x^2) \tag{2.1}$$

on $T^*\mathbb{T}^2$, where $\mu_i \in [0, \infty)$ are parameters, and $V_i \in C^2(\mathbb{T})$ with $V_i \geq 0$ and $\|V_i\|_{C^0} \leq C_i$ are Morse functions (or constant). We may assume without loss of generality that $\min_{x^i} V_i(x^i) = 0$. This includes, e.g., the situation of two pendulums, that is, $V_i(x^i) = 1 - \cos(2\pi x^i)$. The torus coordinates are denoted by $x = (x^1, x^2) \in \mathbb{T}^2$ and the conjugate coordinate pairs are (x^1, p_1) and (x^2, p_2) . By the Maupertuis principle, for fixed $e > 0$, the Hamiltonian flow on the isoenergy manifold $T_e = \{H_0 = e\}$ coincides with the geodesic flow on \mathbb{T}^2 with the Liouville metric g_e (see equation (1.1) and §3.2 for more details) having line element

$$ds_e^2 = (e + \mu_1 V_1(x^1) + \mu_2 V_2(x^2))(dx^1)^2 + (dx^2)^2.$$

The system with the Hamiltonian function in equation (2.1) is clearly integrable in the sense of Definition 1.2, since an additional conserved quantity can easily be found as

$$F_1(x, p) = \frac{p_1^2}{2} - \mu_1 V_1(x^1). \tag{2.2}$$

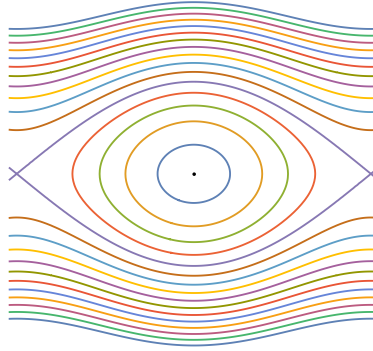


FIGURE 1. Schematic picture of the Liouville foliation of the phase space $T^*\mathbb{T} \cong \mathbb{T} \times \mathbb{R}$ for the classical one-dimensional pendulum system described by the Hamiltonian function $H(x, p) = p^2/2 - (1 - \cos(2\pi x))$. The horizontal direction covers slightly more than one period of length one.

The Liouville foliation of T_e has the following qualitative structure that is similar to the phase portrait of the pendulum. The common level surface

$$T_{(e,f)} = \{H_0 = e, F_1 = f\}$$

differs in shape, depending on the values of e and f . Recall that $e > 0$ and $V_i \geq 0$. If (i) $f \in (-\mu_1 \max_{x^1} V_1(x^1), 0)$ and $e - f > 0$, $T_{(e,f)}$ is an *annulus*; if (ii) $f > 0$ and $e - f > 0$, $T_{(e,f)}$ is a *torus*; if (iii) $f > 0$ and $e - f \in (-\mu_2 \max_{x^2} V_2(x^2), 0)$, $T_{(e,f)}$ is an *annulus*. Therefore, if V_1 and V_2 are both non-constant, the foliation qualitatively exhibits a pendulum-like phase portrait (see Figure 1).

2.1. *Definitions and assumptions.* Our main results concern perturbations of the Hamiltonian function in equation (2.1) in the class of mechanical systems as

$$H_\varepsilon(x, p) = H_0(x, p) + \varepsilon U(x), \tag{2.3}$$

where $\varepsilon \in \mathbb{R}$ and $U \in C^2(\mathbb{T}^2)$ denotes a perturbing potential, which is assumed to be a Morse function (or a constant) and have an absolutely convergent Fourier series:

$$U(x) = \sum_{k_1 \in \mathbb{Z}} U_{k_1}(x^2) e^{i2\pi k_1 x^1} = \sum_{(k_1, k_2) \in \mathbb{Z}^2} U_{k_1, k_2} e^{i2\pi(k_1 x^1 + k_2 x^2)}.$$

(Note that in two dimensions, C^2 -regularity is not sufficient for ensuring an absolutely convergent Fourier series, although in one dimension it is). In the following, we introduce several subsets of \mathbb{Z}^2 in such a way that their definitions immediately carry over in arbitrary dimension $d \in \mathbb{N}$ (see Remark 1.1). First, we define the *spectrum* of U , that is, the set of non-vanishing Fourier coefficients, as

$$S_U := \{\mathbf{k} = (k_1, k_2) \in \mathbb{Z}^2 : U_{\mathbf{k}} \neq 0\}, \tag{2.4}$$

while the *non-singular* spectrum is denoted by

$$S_{U,0} := \{\mathbf{k} \in S_U : \text{there exists } i \neq j \text{ such that } k_i \cdot k_j \neq 0\}. \tag{2.5}$$

Moreover, we define the *coprime set of the orthogonal complement of S_U* as well as its *non-singular subset* via

$$\mathcal{B}(S_U^\perp) := \{\mathbf{b} \in S_U^\perp : \mathbf{b} \text{ coprime}\} \quad \text{and} \quad \mathcal{B}_0(S_U^\perp) := \left\{ \mathbf{b} \in \mathcal{B}(S_U^\perp) : \prod_i b_i \neq 0 \right\}, \quad (2.6)$$

respectively. Note that the orthogonal complement is taken within \mathbb{Z}^2 . For the proofs in §4 and the generalization in Appendix A, it is important to observe that for every $\mathbf{k} \in S_{U,0}$ exists some $\mathbf{b} \in \mathcal{B}_0(S_U^\perp)$ such that $\mathbf{b} \cdot \mathbf{k} = 0$.

Our main results will be formulated under the following assumptions.

Assumptions on the perturbed Hamiltonian function in equation (2.3). Let $H_0 \in C^2(T^*\mathbb{T}^2)$ denote the Hamiltonian function from equation (2.1) with $\min V_i = 0$, $\|V_i\|_{C^0} \leq C_i$ and $\mu_i \in [0, \tilde{\mu}_i]$ for some $\tilde{\mu}_i \in [0, \infty)$, $i \in \{1, 2\}$, and U be a perturbing potential as in equation (2.3), which satisfies one of the following assumptions.

- (A1) If $\tilde{\mu}_1 = \tilde{\mu}_2 = 0$, we have $U \in C^2(\mathbb{T}^2)$.
- (A2) If, without loss of generality, $\tilde{\mu}_1 = 0$ and $\tilde{\mu}_2 > 0$, we have $U \in C^2(\mathbb{T}^2)$ and there exists $d^{(2)} \geq 0$ such that

$$S_U \subset \mathbb{Z} \times [-d^{(2)}, d^{(2)}], \quad (2.7)$$

that is, U is a trigonometric polynomial in the second variable x^2 .

- (A3) If $\tilde{\mu}_1, \tilde{\mu}_2 > 0$, we have $U \in C^2(\mathbb{T}^2)$ and there exist $d^{(1)}, d^{(2)} \geq 0$ such that

$$S_U \subset [-d^{(1)}, d^{(1)}] \times [-d^{(2)}, d^{(2)}], \quad (2.8)$$

that is, U is a trigonometric polynomial.

We denote the minimum over all $d^{(i)}$ such that equation (2.7) (respectively equation (2.8)) holds as $\text{deg}_U^{(i)}$ and call it the *i -degree of U* . Whenever we refer to one of the Assumptions (A1), (A2), or (A3), we implicitly assume that $H_0 \in C^2(T^*\mathbb{T}^2)$ is of the form in equation (2.1).

Note that the assumption on the spectrum in equation (2.4) of U is more restrictive when we include more general potentials $\mu_1 V_1$ and $\mu_2 V_2$ in the unperturbed Hamiltonian H_0 in equation (2.1).

The following basic proposition is fundamental for the precise formulation of our assumptions concerning preservation of integrability. It rephrases certain aspects of the standard Liouville–Arnold theorem [7] in our concrete setting using standard notions from weak KAM theory (see Appendix D and its extension in the arXiv: 2210.02961 version of this article).

PROPOSITION 2.1. (Liouville–Arnold theorem and weak KAM theory [85]) *Let $H_0 \in C^2(T^*\mathbb{T}^2)$ be the Hamiltonian function from equation (2.1).*

- (a) *In the region of phase space, where $f > 0$ as well as $e - f > 0$, each of the two connected components of a Liouville torus $T_{(e,f)}$ (again denoted by $T_{(e,f)}$) is a Lipschitz (we will see in Appendix D that $u_c \in C^3(\mathbb{T}^2)$, so the regularity of $T_{(e,f)}$)*

is in fact C^2) Lagrangian graph, that is,

$$T_{(e,f)} = \{(x, \mathbf{c} + \nabla_x u_{\mathbf{c}}) : x \in \mathbb{T}^2\}$$

for a unique cohomology class $\mathbf{c} \in H^1(\mathbb{T}^2, \mathbb{R}) \cong \mathbb{R}^2$ with $|c_i| > \sqrt{\mu_i} \mathfrak{c}(V_i)$ and $u_{\mathbf{c}} \in C^{1,1}(\mathbb{T}^2)$ so we may equivalently write $T_{(e,f)} \equiv T_{\mathbf{c}}$. Here, $\mathfrak{c}(V_i) := \int_0^1 \sqrt{2 V_i(x^i)} dx^i$ (see Appendix D) and $C^{1,1}$ denotes the functions in C^1 with Lipschitz derivative. The function $u_{\mathbf{c}} \in C^{1,1}(\mathbb{T}^2)$ is a classical solution of the Hamilton–Jacobi equation

$$\alpha(\mathbf{c}) = H_0(x, \mathbf{c} + \nabla_x u_{\mathbf{c}}(x)),$$

where the left-hand side is Mather’s α -function (see Appendix D).

- (b) The Hamiltonian flow on $T_{\mathbf{c}}$ is conjugated to a rotation on \mathbb{T}^2 , that is, there exists a diffeomorphism $\varphi : \mathbb{T}^2 \rightarrow T_{\mathbf{c}}$ such that $\varphi^{-1} \circ \Phi_t^{X_H} \circ \varphi = R_t^{\omega}$ for all $t \in \mathbb{R}$, where $R_t^{\omega} : \mathbb{T}^2 \rightarrow \mathbb{T}^2, x \mapsto (x + \omega t \bmod \mathbb{Z}^2)$ for some rotation vector $\omega \in \mathbb{R}^2$.

An invariant Liouville torus $T_{\mathbf{c}}$ is called *irrational* or *non-resonant* if $\mathbf{k} \cdot \omega \neq 0$ for all $\mathbf{k} \in \mathbb{Z}^2 \setminus \{0\}$. If this is not the case, the invariant torus is *rational* or *resonant*. For two-dimensional manifolds (and if $\omega_2 \neq 0$), this can be phrased as a distinction between $\omega_1/\omega_2 \notin \mathbb{Q}$ and $\omega_1/\omega_2 \in \mathbb{Q}$.

Assumptions on the preserved integrability of equation (2.3). Let $H_0 \in C^2(T^*\mathbb{T}^2)$ denote the Hamiltonian function from equation (2.1) satisfying one of the Assumptions (A1)–(A3), and U a perturbing potential as in equation (2.3) such that the following statement concerning the perturbed Hamilton–Jacobi equation (HJE):

$$\alpha_{\varepsilon}(\mathbf{c}) = H_{\varepsilon}(x, \mathbf{c} + \nabla_x u_{\varepsilon, \mathbf{c}}(x)) \tag{2.9}$$

as well as the preserved integrability of H_{ε} is satisfied.

- (P) There exists an energy $e > 0$ such that for every $(n, m) \in \mathcal{B}_0(\mathcal{S}_T^{\perp})$ (recall equation (2.6)) and $\mu_i \in [0, \tilde{\mu}_i], i \in \{1, 2\}$, there exists a sequence $(\varepsilon_k)_{k \in \mathbb{N}}$ with $\varepsilon_k \neq 0$ but $\varepsilon_k \rightarrow 0$ such that we have the following.

- (i) The resonant torus from Proposition 2.1, characterized by $\mathbf{c} \in H^1(\mathbb{T}^2, \mathbb{R})$ with

$$|c_i| > \sqrt{\mu_i} \mathfrak{c}(V_i) \tag{2.10}$$

in the isoenergy submanifold T_e having rotation vector proportional to (n, m) , is preserved under the sequence of deformations $(H_{\varepsilon_k})_{k \in \mathbb{N}}$.

- (ii) For $\mathbf{c} \in H^1(\mathbb{T}^2, \mathbb{R})$ satisfying equation (2.10), Mather’s α -function and a solution $u_{\varepsilon, \mathbf{c}} : \mathbb{T}^2 \rightarrow \mathbb{R}$ of the HJE in equation (2.9) can be expanded to first order in ε , that is,

$$u_{\varepsilon, \mathbf{c}} = u_{\mathbf{c}}^{(0)} + \varepsilon u_{\mathbf{c}}^{(1)} + \mathcal{O}_{\mathbf{c}}(\varepsilon^2) \quad \text{and} \quad \alpha_{\varepsilon} = \alpha^{(0)} + \varepsilon \alpha^{(1)} + \mathcal{O}(\varepsilon^2), \tag{2.11}$$

where $u_{\mathbf{c}}^{(0)}, u_{\mathbf{c}}^{(1)} \in C^{1,1}(\mathbb{T}^2)$ and $\mathcal{O}_{\mathbf{c}}(\varepsilon^2)$ is understood in $C^{1,1}$ -sense. (Having C^1 -regularity here would be sufficient for our proofs in §4. However, we chose $C^{1,1}$ -regularity for the formulation of Assumption (P) to be in agreement with the statement from Proposition 2.1(b). More precisely, $C^{1,1}$ -regularity

is kind of a compromise between the true C^3 -regularity of u_c and the required C^1 -regularity of $u_{\varepsilon,c}$. In addition, $C^{1,1}$ is the optimal regularity for subsolutions of equation (2.9), which exist, even if the Hamiltonian H_ε is *not* integrable (see [12, 46]).

We comment on the validity of assuming Assumption (P) in Remark D.1 in Appendix D. Moreover, we shall also discuss an alternative to equation (2.11) in Remark D.3. Finally, one can easily see from the proofs given in §4 that the condition on a *fixed* isoenergy manifold $\{H_\varepsilon = e\}$ can be relaxed to having preservation of invariant tori in isoenergy manifolds characterized by energies $e \geq e_0$ for some fixed $e_0 > 0$.

Note that the rational invariant tori are the most ‘fragile’ objects of an integrable system as the KAM theorem [5, 63, 80] predicts that general (non-integrable) perturbations preserve only ‘sufficiently irrational’ (*Diophantine*) invariant tori.

2.2. *Results.* As mentioned above, our main results in Theorems 2.2, 2.3, and 2.4 concern rigidity of certain deformations of integrable metrics (in the sense of Question (Q2)), which, by means of the Maupertuis principle, correspond to perturbations of the form in equation (2.3). More precisely, under the assumptions formulated above, our results show that the perturbed Hamiltonian function in equation (2.3) has to be of the same general form as the unperturbed Hamiltonian function in equation (2.1). This means that the potential U is *separable*, that is, there exist $U_1, U_2 \in C^2(\mathbb{T}^2)$ such that

$$U(x) = U_1(x^1) + U_2(x^2).$$

THEOREM 2.2. *Let H_ε from equation (2.3) satisfy Assumption (A1) and Assumption (P) for some energy $e > 0$. Then U is separable in a sum of two single-valued functions.*

Put briefly, in view of the Maupertuis principle, this means that integrable deformations in the same conformal class of a flat metric are Liouville metrics. Now, Theorem 2.3 generalizes Theorem 2.2 to Hamiltonian functions which depend on one toral position variable via a mechanical potential.

THEOREM 2.3. *Let H_ε from equation (2.3) satisfy Assumption (A2) and Assumption (P) for some energy $e > 0$. Then the following hold.*

- (a) *If $\tilde{\mu}_2 = \tilde{\mu}_2(C_2, \deg_U^{(2)}, e) > 0$ is small enough (see Lemma 4.2), we have that U is separable in a sum of two single-valued functions.*
- (b) *If, additionally, V_2 is analytic, then U is separable, irrespective of $\tilde{\mu}_2 > 0$, but only for $\mu_2 \in [0, \tilde{\mu}_2]$ outside of an exceptional null-set.*

Therefore, by means of the Maupertuis principle, we infer that integrable deformations in the same conformal class of metrics realizing surfaces of revolution (see §3.2) are Liouville metrics. Finally, Theorem 2.4 generalizes the above results to Hamiltonian functions, which correspond to arbitrary Liouville metrics by means of the Maupertuis principle.

THEOREM 2.4. *Let H_ε from equation (2.3) satisfy Assumption (A3) and Assumption (P) for some energy $e > 0$. Then the following hold.*

- (a) If $\tilde{\mu}_1 = \tilde{\mu}_1(C_1, \deg_U^{(1)}, \deg_U^{(2)}, e) > 0$ and $\tilde{\mu}_2 = \tilde{\mu}_2(C_2, \deg_U^{(1)}, \deg_U^{(2)}, e) > 0$ are small enough (see Lemma 4.3), we have that U is separable in a sum of two single-valued functions.
- (b) If, additionally, V_2 is analytic and $\tilde{\mu}_1 = \tilde{\mu}_1(C_2, \deg_U^{(1)}, \deg_U^{(2)}, e) > 0$ is small enough, then U is separable, irrespective of $\tilde{\mu}_2 > 0$, but only for $\mu_2 \in [0, \tilde{\mu}_2]$ outside of an exceptional one-dimensional null-set (depending on $\mu_1 \in [0, \tilde{\mu}_1]$).
- (c) If both V_i for $i = 1, 2$ are analytic, then U is separable, irrespective of $\tilde{\mu}_1, \tilde{\mu}_2 > 0$, but only for $(\mu_1, \mu_2) \in [0, \tilde{\mu}_1] \times [0, \tilde{\mu}_2]$ outside of an exceptional two-dimensional null-set.

Our results formulated in Theorems 2.2, 2.3, and 2.4 can each be viewed as a verification of a special case of the following conjecture, saying that ‘(nice) integrable deformations of Liouville metrics are Liouville metrics’.

Conjecture. (Deformational rigidity of Liouville metrics) Let g be a Liouville metric on \mathbb{T}^2 and let $(g_t)_{t \in [0,1]}$ with $g_0 = g$ be a deformation that preserves all rational invariant tori (except finitely many). Then g_t is a Liouville metric for all $t \in [0, 1]$.

This conjecture is in strong analogy to the perturbative Birkhoff conjecture for integrable billiards, which is discussed in §3.4 below.

3. Literature review: integrable metrics on the torus

As pointed out in §1.1, integrability of metrics on one-dimensional manifolds is not questionable and the first non-trivial examples occur whenever M has dimension two. Recall from Definition 1.2 that integrability of metrics on two-dimensional manifolds requires the existence of only one additional first integral (beside the Hamiltonian).

3.1. *Topological obstructions.* The following theorem due to Kozlov [68, 69] (see [15] for a strengthened version of this result) categorizes two-dimensional compact manifolds regarding the possibility to endow them with an integrable metric (see Question (Q1)).

THEOREM 3.1. (Kozlov [68, 69]) *Let M be a two-dimensional compact and real-analytic manifold that is endowed with a real-analytic Riemannian metric g . If the Euler characteristic χ_M of M is negative, then there exists no other non-trivial real-analytic first integral.*

A result similar to Theorem 3.1 holds for polynomially integrable geodesic flows.

THEOREM 3.2. (Kolokoltsov [64]) *There exist no polynomially integrable geodesic flow on a closed two-dimensional Riemannian manifold M with negative Euler characteristic χ_M .*

Recall that any two-dimensional compact manifold M can be represented either as the sphere with handles or the sphere with Möbius strips, in the orientable and non-orientable case, respectively. The Euler characteristic χ_M can be computed as

$$\chi_M = 2 - 2g \quad \text{respectively} \quad \chi_M = 2 - m,$$

where g is the number of handles (the genus) and m is the number of Möbius strips. To have integrability, the above theorem imposes the condition $\chi_M \geq 0$ on M and we thus know that the number of handles is at most 1 and the number of Möbius strips is not greater than 2. Therefore, any real-analytic two-dimensional compact Riemannian manifold (M, g) with real-analytic (or polynomial) additional integral is either the sphere \mathbb{S}^2 or the torus \mathbb{T}^2 (in the orientable case), or the projective plane \mathbb{RP}^2 or the Klein bottle \mathbb{K}^2 (in the non-orientable case). In [29], Bolsinov and Taimanov give a striking example of a real-analytic Riemannian manifold of dimension three, whose geodesic flow has the peculiar property, and that it is smoothly (but not analytically) integrable although it has positive topological entropy [1]. The problem of proving (non-)existence of smoothly (but not analytically) integrable geodesic flows on compact surfaces of genus $g > 1$ is widely open (see [31]).

In this work, we focus on integrable metrics on the torus \mathbb{T}^2 and refer to works by Bolsinov, Fomenko, Matveev, Kolokoltsov, and others [27, 48, 64, 78, 81] for studies on integrable metrics on the sphere, the projective plane, and the Klein bottle. See [24, 31] for recent surveys on open problems, and questions concerning geodesics and integrability of finite-dimensional systems in general.

3.2. *Linearly and quadratically integrable metrics.* The first non-trivial class of integrable metrics on the torus \mathbb{T}^2 is *surfaces of revolution*. Consider a two-dimensional surface $M \subset \mathbb{R}^3$ given by the equation $r = r(z)$ in standard cylindrical coordinates $(r, \varphi, z) \in (0, \infty) \times [0, 2\pi) \times \mathbb{R}$. As local coordinates on M , we take z and φ . In the case where $r(z)$ is L -periodic and we identify 0 and L , then M is diffeomorphic to the torus \mathbb{T}^2 and the Riemannian metric induced on M by the Euclidean metric on \mathbb{R}^3 has line element

$$ds^2 = (1 + r'(z)^2)dz^2 + r(z)^2d\varphi^2. \tag{3.1}$$

Since the corresponding Hamiltonian function in equation (1.5) is independent of φ , its associated momentum variable p_φ is an additional first integral and thus the metric in equation (3.1) is integrable. Note that the additional first integral is linear in the momentum variables.

As discussed earlier, a Riemannian metric g on \mathbb{T}^2 is called a *Liouville metric*, whenever its line element can be written in the form in equation (1.1) in appropriate global coordinates (x^1, x^2) , and where f_1 and f_2 are smooth positive periodic functions. The corresponding Hamiltonian function in equation (1.5) is given by

$$H(x^1, x^2, p_1, p_2) = \frac{p_1^2 + p_2^2}{2(f_1(x^1) + f_2(x^2))}$$

and an additional first integral can easily be obtained as

$$F(x^1, x^2, p_1, p_2) = p_1^2 - f_1(x^1)H(x^1, x^2, p_1, p_2).$$

Therefore, clearly, also Liouville metrics are integrable. Note that the additional first integral F is quadratic in the momentum variables. It is not hard to see that a surface of revolution is just a particular case of a Liouville metric, where one can choose, e.g., $f_2 \equiv 0$, by employing a simple change of variables.

The following proposition also provides the converse to the observation that surfaces of revolution and Liouville metrics admit additional first integrals which are linear and quadratic in the momenta, respectively. It collects several statements that have been proven in early works by Dini [43], Darboux [37], and Birkhoff [22], and were further developed by Babenko and Nekhoroshev [11], Kiyohara [62], Kolokoltsov [64], and others.

PROPOSITION 3.3. (Linear and quadratic first integrals [11, 22, 37, 43, 62, 64])

- (a) *Let the metric g on \mathbb{T}^2 possess an additional first integral F that is linear in the momenta. Then there exist global periodic coordinates (x^1, x^2) on the torus such that the line element of g takes the form*

$$ds^2 = f(x^1)(a(dx^1)^2 + c dx^1 dx^2 + b(dx^2)^2),$$

where f is some positive periodic function and $a, b, c \in \mathbb{R}$ such that the quadratic form $a(dx^1)^2 + c dx^1 dx^2 + b(dx^2)^2$ is positive definite.

Conversely, any such metric on the torus \mathbb{T}^2 admits an additional first integral that is linear in the momentum variables.

In case a linear in momenta F exists locally near a point $q \in \mathbb{T}^2$, then there exists local coordinates (x^1, x^2) near q such that the line element of g reads

$$ds^2 = f(x^1)((dx^1)^2 + (dx^2)^2).$$

- (b) *A metric g on \mathbb{T}^2 possess an additional first integral F that is quadratic in the momenta if and only if there exists a finite-sheeted covering $\pi : \tilde{\mathbb{T}}^2 \rightarrow \mathbb{T}^2$ by another torus, such that the lifted metric $\tilde{g} = \pi^*g$ is globally Liouville, that is, there exist global periodic coordinates (x^1, x^2) on $\tilde{\mathbb{T}}^2$ and smooth positive periodic functions f_1 and f_2 such that the line element of \tilde{g} takes the form in equation (1.1).*

There exist Riemannian metrics g on \mathbb{T}^2 which are not globally Liouville but have an additional first integral that is quadratic in the momentum variables.

In case a quadratic in momenta F exists locally near a point $q \in \mathbb{T}^2$, then there exist local coordinates (x^1, x^2) near q such that the line element of g takes the form in equation (1.1).

This classical result completely characterizes the integrable metrics g on \mathbb{T}^2 that admit an additional first integral that is linear or quadratic in the momentum variables. Similar results hold for Riemannian metrics on general two-dimensional manifolds [27, 48, 64, 81].

3.3. *Polynomially integrable metrics of higher degree.* In the case of a sphere \mathbb{S}^2 , one can easily construct examples of metrics which admit an additional first integral that is cubic respectively quartic in the momentum variables. Using the Maupertuis principle, these can be obtained from the metrics constructed from Goryachev and Chaplygin [33, 54], and Kovalevskaya [67] in the situation of the dynamics of a rigid body. Therefore, let $h > 1$ be large enough (cf. equation (1.8)) and define the metrics g_3 and g_4 on \mathbb{R}^3 via their respective line elements

$$ds_3^2 = \frac{h - x^1}{4} \frac{(dx^1)^2 + (dx^2)^2 + 4(dx^3)^2}{(x^1)^2 + (x^2)^2 + (x^3)^2/4}, \quad ds_4^2 = \frac{h - x^1}{2} \frac{(dx^1)^2 + (dx^2)^2 + 2(dx^3)^2}{(x^1)^2 + (x^2)^2 + (x^3)^2/2}.$$

By restriction of g_3 and g_4 to the unit sphere $\mathbb{S}^2 \subset \mathbb{R}^3$, the resulting metrics admit an additional first integral that is cubic respectively quartic in the momentum variables. It was shown by Bolsinov, Fomenko, and Kozlov [25, 28] that these cannot be reduced to first integrals that are polynomially in the momentum variables of a lower degree, that is, they are not linearly or quadratically integrable. Since all attempts to construct such examples for the case of the torus have failed so far, the following folklore conjecture emerged.

Folklore Conjecture. Liouville metrics are the only integrable metrics on \mathbb{T}^2 .

In this general form, there is strong indication for the conjecture being false, as to be shown below (see Theorem 3.8). We will, however, provide existing results, which indicate that a certain weaker version of this conjecture, also formulated below, is indeed true.

It was proven by Korn and Lichtenstein [65, 71] that on every point on a two-dimensional Riemannian manifold (M, g) , there exist *locally* isothermal coordinates, that is, locally, the line element takes the form

$$ds^2 = \lambda(x^1, x^2)((dx^1)^2 + (dx^2)^2), \quad (3.2)$$

where λ is some smooth positive function. In the case of a torus, it can be shown (by virtue of the uniformization theorem) that there exist *global* isothermal coordinates (not necessarily periodic), so the metric g is conformal equivalent to the Euclidean metric g_{eucl} . In particular, assuming that (x^1, x^2) are just the angular coordinates on the torus \mathbb{T}^2 and in the special case of λ being a trigonometric polynomial (this means that the spectrum \mathcal{S}_λ defined in equation (2.4) is bounded), we have the following result due to Denisova and Kozlov.

THEOREM 3.4. (Denisova and Kozlov [39]) *Let λ from equation (3.2) be a trigonometric polynomial and assume that the geodesic flow on \mathbb{T}^2 is polynomially integrable. Then there exists an additional polynomial first integral of degree at most two.*

Note that by Weierstrass's theorem, any conformal factor λ can be approximated as closely as required by a trigonometric polynomial. However, in the case of a general conformal factor λ , there is the following theorem, again due to Denisova and Kozlov [40].

THEOREM 3.5. (Denisova and Kozlov [40]) *Assume that the geodesic flow on (\mathbb{T}^2, g) is polynomially integrable with first integral F of degree n such that:*

- (a) *if n is even, then F is an even function of p_1 and p_2 ;*
 - (b) *if n is odd, then F is an even function of p_1 (or p_2) and an odd function of p_2 (or p_1).*
- Then there exists an polynomial first integral of degree at most two.*

In the following theorem, we collect several results from Bialy [13], Denisova, Kozlov [41] and Treshchev [42], Agapov and Aleksandrov [2], and Mironov [79].

THEOREM 3.6. *Let H be a natural mechanical Hamiltonian (see equation (1.7)) on the torus \mathbb{T}^2 equipped with the flat metric g_{eucl} . Assume that H is polynomially integrable of degree n . If $n = 3, 4$, there exists another polynomial first integral of degree at most two. Whenever H is a real-analytic Hamiltonian, this is also true for $n = 5$.*

Kozlov and Treshchev [70] considered the problem from yet another point of view. They investigated the case of a mechanical Hamiltonian

$$H = \frac{1}{2} \sum_{ij} a_{ij} p_i p_j + V(x^1, \dots, x^n),$$

where $A = (a_{ij})_{ij}$ is a positive definite matrix and V is a trigonometric polynomial of $(x^1, \dots, x^n) \in \mathbb{T}^n$. On the one hand, they show that there exist n polynomial first integrals if and only if the spectrum \mathcal{S}_V of V is contained in $m \leq n$ mutually orthogonal lines meeting at the origin. On the other hand, they showed that whenever there exist n polynomial integrals with independent forms of highest degree, then there exist n independent involutive polynomial first integrals of degree at most two. In the case where $a_{ij} = \delta_{i,j}$ (which can be achieved by diagonalization and scaling), Combot [34] improved the first result from the assumption of *polynomial* integrability to *rational* integrability, that is, the additional first integrals being rational functions of p_i and $e^{i2\pi x^i}$. More recently [57, 83, 84], the problem was rephrased in the language of Killing tensor fields on \mathbb{T}^2 , where the *order* of an additional (polynomial) first integral is replaced by the *rank* of a Killing tensor field.

The results of Theorems 3.4, 3.5, and 3.6 support the validity of the following weaker version of the folklore conjecture formulated by Denisova and Kozlov [39].

Conjecture. [39] If g is a metric on \mathbb{T}^2 that is polynomially integrable, then there exists an additional polynomial first integral of degree at most two.

By Proposition 3.3, this means that polynomially integrable metrics on \mathbb{T}^2 are Liouville metrics. However, beside the partial results given above, a proof of this conjecture is still open. The numerous attempts on proving it used methods of complex analysis [11, 22] and the theory of partial differential equations (PDEs) [17, 19]. More precisely, it is shown by Kolokoltsov [64] that there exists an additional first integral quadratic in the momenta if and only if there exists a holomorphic function $R(z) = R_1(z) + iR_2(z)$, with real valued R_1 and R_2 , and $z = x^1 + ix^2$, which solves

$$R_2(\partial_{x^2}^2 \lambda - \partial_{x^1}^2 \lambda) + R_1(\partial_{x^1} \partial_{x^2} \lambda) - 3(\partial_{x^1} R_2)(\partial_{x^1} \lambda) + 3(\partial_{x^2} R_2)(\partial_{x^2} \lambda) + 2(\partial_{x^2}^2 R_2)\lambda = 0, \quad (3.3)$$

where λ denotes the conformal factor from equation (3.2). Note that the second term in equation (3.3) disappears whenever λ is the conformal factor of a Liouville metric. In this situation, the linear PDE in equation (3.3) always has a holomorphic solution $R = R_1 + iR_2$. The existence of first integrals of higher degree turns out to be equivalent to delicate questions about nonlinear PDEs of hydrodynamic type [17–19]. The PDE approach has also successfully been applied to generate new examples of integrable *magnetic* geodesic flows as analytic deformations of Liouville metrics on \mathbb{T}^2 without magnetic field (see [3]). In fact, the examples from [3] disprove the folklore conjecture when understood in the larger class of magnetic geodesic flows.

However, even for the *original* folklore conjecture stated above, there is a result due to Corsi and Kaloshin [35], which indicates it being false in the following (considerably weaker) sense.

THEOREM 3.8. (Corsi and Kaloshin [35]) *There exists a real-analytic mechanical Hamiltonian*

$$H_\varepsilon(x^1, x^2, p_1, p_2) = \frac{p_1^2 + p_2^2}{2} + U(x^1, x^2; \varepsilon)$$

with a non-separable potential U and an analytic change of variables Φ such that $H_\varepsilon \circ \Phi = (p_1^2 + p_2^2)/2$ on the energy surface $\{H_\varepsilon = 1/2\}$ and $p \in \mathcal{P}$, where \mathcal{P} denotes a certain cone in the action space. (The function U is called non-separable whenever it cannot be written as a sum of two single-valued functions.)

If one assumes that the whole phase space $T^*\mathbb{T}^2$ is foliated by two-dimensional invariant Liouville tori (which is often called C^0 -integrability or *complete integrability*), then it follows from Hopf conjecture [30, 58] that the associated metric must be flat. This notion of integrability is thus too strong for a meaningful characterization of integrable metrics on \mathbb{T}^2 . (Similar results have been shown for geodesic flows of more general Finsler metrics on \mathbb{T}^2 preserving a sufficiently regular foliation of the phase space [52, 53].)

3.4. Analogy to integrable billiards. The fundamental Question (Q2) of characterizing integrable metrics on the torus \mathbb{T}^2 can be thought of as an analogue of identifying the class of integrable billiards [61]. For billiards, integrability is understood in a similar way as for the geodesic flow (see Definition 1.2). More precisely, integrability is characterized either through the existence of an integral of motion (near the boundary of the billiard table) for the so-called billiard ball map, or the existence of a foliation of the phase space (globally or near the boundary), consisting of invariant curves. The *classical Birkhoff conjecture* [23, 82] states that the boundary of a strictly convex integrable billiard table is necessarily an ellipse. This corresponds to the folklore conjecture formulated above. Remarkably, while the Birkhoff conjecture is believed to be true, and there is strong evidence that this indeed the case [10, 20, 49, 60] (on the opposite side, Treshev constructed a non-elliptic billiard table which is *formally* integrable close to a two-periodic orbit [88–90]. This formal power series has recently been shown to be of Gervé class of order $\sigma > 9/4$ [93]), the folklore conjecture in its general form was shown to be false by Theorem 3.8.

However, recall that if one assumes C^0 -integrability of a metric on \mathbb{T}^2 , the metric is actually flat [30, 58]. This corresponds to the following result from Bialy in the case of billiards.

THEOREM 3.9. (Bialy [14]) *If the phase space of the billiard ball map is completely foliated by continuous invariant curves which are all not null-homotopic, then the boundary of the billiard table is a circle.*

Following a similar strategy leading to Theorem 3.9, Bialy and Mironov [21] proved the Birkhoff conjecture for centrally symmetric billiards, assuming only *local* C^0 -integrability, that is, the foliation of a suitable open *proper* subset of the phase space. In addition to this, the weakened version of the folklore conjecture (polynomial integrals can be reduced to integrals of degree at most two) corresponds to the so-called *algebraic Birkhoff conjecture*, which has recently been proven [20, 49].

The main results of this paper in Theorems 2.2, 2.3, and 2.4 prove special cases of our conjecture that integrable deformations of Liouville metrics which preserve all (but finitely many) rational invariant tori are again Liouville metrics. This is related to the following conjecture in the case of billiards.

Perturbative Birkhoff Conjecture. [61] A smooth strictly convex domain that is sufficiently close to an ellipse and whose corresponding billiard ball map is integrable is necessarily an ellipse.

A first result in this direction was obtained by Delshams and Ramírez-Ros [38]. More recently, Avila, De Simoi, and Kaloshin [10] proved the conjecture for domains which are sufficiently close to a circle. The complete proof for domains sufficiently close to an ellipse of any eccentricity is given by Kaloshin and Sorrentino in [60]. Both works require the preservation of rational caustics (a curve Γ is a caustic for the billiard in the domain Ω if every time a trajectory is tangent to it, then it remains tangent after every reflection according to the billiard ball map), which can be thought of as an analogue for the preservation of rational invariant tori as a fundamental assumption of our main results from §2. The result in [10] was later extended by Huang, Kaloshin, and Sorrentino [59] to the case of *local integrability* close to the boundary and finally significantly improved by Koval [66].

Finally, as shown by Vedyushkina and Fomenko [92], linearly and quadratically integrable geodesic flows on orientable two-dimensional Riemannian manifolds are Liouville equivalent to topological billiards, glued from planar billiards bounded by concentric circles and arcs of confocal quadrics, respectively.

4. Proofs

In this section, we prove our main result as formulated in Theorems 2.2, 2.3, and 2.4. All proofs will, in general, follow the same three-step strategy.

- (i) Transform the unperturbed system H_0 in action-angle coordinates.
- (ii) Derive a first-order harmonic equation (that is, concerning the Fourier coefficients) for the perturbation by Assumption (P).
- (iii) Annihilate sufficiently many Fourier coefficients of the perturbing potential by proving a certain *full-rank* condition for a naturally associated linear system for each of the three theorems separately (cf. Lemmas 4.1, 4.2, and 4.3). Finally, for analytic potentials V_i , the extensions of our results beyond the perturbative regime are proven by exploiting the analytic dependence of the linear system on μ_i (see Appendix C).

4.1. *Proof of Theorem 2.2.* The argument is divided into three steps.

Step (i). Fix an energy $e > 0$. Since the Hamiltonian is already in action-angle coordinates, we simply change notation and write $(x^i, p_i) = (\theta^i, I_i)$ for $i = 1, 2$ as well as $\theta = (\theta^1, \theta^2)$ and $I = (I_1, I_2)$, such that the perturbed Hamiltonian function H_ε takes the form

$$H_\varepsilon(\theta, I) = \frac{I_1^2}{2} + \frac{I_2^2}{2} + \varepsilon U(\theta).$$

Step (ii). By Assumption (P), for any $(n, m) \in \mathcal{B}_0(S_U^\perp)$ (recall equation (2.6)), we can find (in the isoenergy manifold T_{e_ε} with energy $e = e_\varepsilon$ and $\varepsilon = \varepsilon_k$ for some $k \in \mathbb{N}$) a rational invariant Liouville torus with rotation vector $\omega = (\omega_1, \omega_2)$ which satisfies

$$\frac{\omega_1}{\omega_2} = \frac{n}{m} \in \mathbb{Q}. \tag{4.1}$$

Moreover, we fix $\mathbf{c} \in H^1(\mathbb{T}^2, \mathbb{R}) \cong \mathbb{R}^2$ to be given by $\mathbf{c} = (\omega_1, \omega_2)$. We make this choice to cancel the average over a trajectory in equation (4.3) of the first term on the right-hand side of equation (4.2) (cf. also equations (4.6)–(4.8) below).

Using Assumption (P) again, we can expand the Hamilton–Jacobi equation in equation (2.9) as

$$\begin{aligned} \alpha_\varepsilon(\mathbf{c}) &= H_\varepsilon(\theta, \mathbf{c} + \nabla u_{\varepsilon, \mathbf{c}}(\theta)) \\ &= \frac{|\partial_{\theta^1} u_{\varepsilon, \mathbf{c}}(\theta) + c_1|^2}{2} + \frac{|\partial_{\theta^2} u_{\varepsilon, \mathbf{c}}(\theta) + c_2|^2}{2} + \varepsilon U(\theta) \\ &= \frac{c_1^2}{2} + \frac{c_2^2}{2} + \langle \mathbf{c}, \nabla u_{\varepsilon, \mathbf{c}}(\theta) \rangle + \varepsilon U(\theta) + \frac{(\partial_{\theta^1} u_{\varepsilon, \mathbf{c}}(\theta))^2}{2} + \frac{(\partial_{\theta^2} u_{\varepsilon, \mathbf{c}}(\theta))^2}{2}, \end{aligned}$$

and it holds that

$$u_{\varepsilon, \mathbf{c}} = u_{\mathbf{c}}^{(0)} + \varepsilon u_{\mathbf{c}}^{(1)} + \mathcal{O}_{\mathbf{c}}(\varepsilon^2)$$

with $u_{\mathbf{c}}^{(0)} = u_{0, \mathbf{c}}$. Since $H_0(\theta, I)$ is integrable (and written in action-angle coordinates), one can choose $u_{0, \mathbf{c}} \equiv 0$. By equation (D.6) in Proposition D.2 (see also [50]), we have $\alpha^{(1)}(\mathbf{c}) = [U]_0$, where

$$[U]_0 = \int_{\mathbb{T}^2} U(x^1, x^2) dx^1 \wedge dx^2.$$

Since the sequence $(\varepsilon_k)_{k \in \mathbb{N}}$ from Assumption (P) converges to zero, we compare coefficients and establish the first-order equation

$$[U]_0 = \alpha^{(1)}(\mathbf{c}) = \langle \mathbf{c}, \nabla u_{\mathbf{c}}^{(1)}(\theta) \rangle + U(\theta). \tag{4.2}$$

Averaging equation (4.2) over the trajectory $\theta(t) = \theta_0 + \omega t \in \mathbb{T}^2$, with initial position $\theta_0 \in \mathbb{T}^2$ and where $\omega = \mathbf{c}$ is chosen according to equation (4.1), such that the period T_ω satisfies $T_\omega \cdot \omega = (n, m)$, we get

$$[U]_0 = \frac{1}{T_\omega} \int_0^{T_\omega} \frac{d}{dt} u_{\varepsilon, \mathbf{c}}^{(1)}(\theta(t)) dt + \frac{1}{T_\omega} \int_0^{T_\omega} U(\theta(t)) dt. \tag{4.3}$$

The first integral vanishes since $\theta(0) = \theta(T_\omega)$ such that we are left with

$$\int_0^1 (U(\theta_0^1 + nt, \theta_0^2 + mt) - [U]_0) dt = 0 \tag{4.4}$$

for all $\theta_0 = (\theta_0^1, \theta_0^2) \in \mathbb{T}^2$, which easily follows from equation (4.3) after a change of variables.

Before continuing with the third and final step, we have two important observations. First, by replacing $U \rightarrow U - [U]_0$, we can assume without loss of generality that

$[U]_0 = 0$. Second, we define the *separable part*, U_{sep} , of U as

$$U_{\text{sep}}(x^1, x^2) := \sum_{(k_1, k_2) \in \mathcal{S}_U \setminus \mathcal{S}_{U,0}} U_{k_1, k_2} e^{i2\pi k_1 x^1} e^{i2\pi k_2 x^2} \tag{4.5}$$

(recall the definition of the spectrum and the non-singular spectrum in equations (2.4) and (2.5)). Then, after a simple computation, we find that

$$\int_0^1 U_{\text{sep}}(\theta_0^1 + nt, \theta_0^2 + mt) dt = [U_{\text{sep}}]_0 \quad \text{for all } (\theta_0^1, \theta_0^2) \in \mathbb{T}^2$$

holds *generally* (that is, independent of the first-order relation in equation (4.2)) by means of equation (D.6) in Proposition D.2 (see also Remark D.1). We can thus split off the separable part and assume that $\mathcal{S}_U = \mathcal{S}_{U,0}$ in the following. Hence, the third step consists of showing that $\mathcal{S}_U = \mathcal{S}_{U,0} = \emptyset$.

Step (iii). The goal of this final step is to establish the following lemma.

LEMMA 4.1. *Let $(n, m) \in \mathcal{B}_0(\mathcal{S}_U^\perp)$ as in equation (4.1) from Step (ii). Then $U_{jm, -jn} = 0$ for all $j \in \mathbb{Z} \setminus \{0\}$.*

Since $(n, m) \in \mathcal{B}_0(\mathcal{S}_U^\perp)$ were arbitrary, this proves that

$$\mathcal{S}_U \subset (\{0\} \times \mathbb{Z}) \cup (\mathbb{Z} \times \{0\}),$$

or equivalently $\mathcal{S}_{U,0} = \emptyset$ and we have shown Theorem 2.2. It remains to prove Lemma 4.1.

Proof of Lemma 4.1. Starting from equation (4.4), we perform a Fourier decomposition to infer

$$\sum_{k_1, k_2 \neq 0} \left[U_{k_1, k_2} \int_0^1 e^{i2\pi k_1 nt} e^{i2\pi k_2 mt} dt \right] e^{i2\pi k_1 \theta_0^1} e^{i2\pi k_2 \theta_0^2} = 0 \quad \text{for all } (\theta_0^1, \theta_0^2) \in \mathbb{T}^2,$$

which implies that

$$U_{k_1, k_2} \cdot \delta_{k_1 n + k_2 m, 0} = 0. \tag{□}$$

Applying Lemma 4.1 for every $(n, m) \in \mathcal{B}_0(\mathcal{S}_U^\perp)$, we find that $\mathcal{S}_{U,0} = \emptyset$, which finishes the proof of Theorem 2.2.

4.2. *Proof of Theorem 2.3.* For notational simplicity, we write $\mu \equiv \mu_2 > 0$ and $V \equiv V_2 \in C^2(\mathbb{T})$.

Step (i). We fix an energy $e > 0$ and consider the region of the phase space, where the subsystem in the second pair of coordinates is rotating, that is,

$$\frac{p_2^2}{2} - \mu V(x^2) = e^{(2)} > 0$$

and for $p_1^2/2 = e^{(1)} > 0$, we have $e = e^{(1)} + e^{(2)}$. In a neighborhood of each of the two Liouville tori characterized by $H_0 = e$ and $p_1^2/2 = e^{(1)}$, we can find a change of variables $(x^2, p_2) = \Phi_\mu^{(2)}(\theta^2, I_2)$ (and we denote $(x^1, p_1) = (\theta^1, I_1)$) such that the Hamiltonian

function H_0 gets transformed in action-angle coordinates, that is,

$$H_0(\theta^1, I_1, \Phi_\mu^{(2)}(\theta^2, I_2)) = \frac{I_1^2}{2} + h_\mu^{(2)}(I_2)$$

for some smooth function $h_\mu^{(2)}$ agreeing with Mather’s α -function for the one-dimensional subsystem described by the Hamiltonian $p_2^2/2 - \mu V(x^2)$ (see Appendix D). The change in the order of the four arguments of H_0 should not lead to confusion. Now, the perturbed Hamiltonian takes the form

$$H_\varepsilon(\theta^1, I_1, \Phi_\mu^{(2)}(\theta^2, I_2)) = \frac{I_1^2}{2} + h_\mu^{(2)}(I_2) + \varepsilon U(\theta^1, x^2(\theta^2, I_2, \mu)),$$

where we write $x_\mu^2(\theta^2, I_2)$ for the first component of $\Phi_\mu^{(2)}(\theta^2, I_2)$.

Step (ii). Assume without loss of generality that the 2-degree $\text{deg}_U^{(2)}$ of U is at least 1 (recall equation (2.7)), as otherwise, we had $U(x) = U_1(x^1)$ and Theorem 2.3 was proven. Then, for any $(n, m) \in \mathcal{B}_0(\mathcal{S}_U^\perp)$, in particular with $|n| \leq \text{deg}_U^{(2)}$, we can find (in the isoenergy manifold T_{e_ε} with energy $e = e_\varepsilon$ and $\varepsilon = \varepsilon_k$ for some $k \in \mathbb{N}$) a rational invariant Liouville torus with rotation vector $\omega = (\omega_1, \omega_2)$, which satisfies

$$\frac{\omega_1}{\omega_2} = \frac{n}{m} \in \mathbb{Q} \quad \text{and} \quad \omega = (c_1, \nabla h_\mu^{(2)}(c_2)) \tag{4.6}$$

for some $c \in H^1(\mathbb{T}^2, \mathbb{R}) \cong \mathbb{R}^2$ with $c_1 = \omega_1$ (as around equation (4.1)) and $|c_2| > \gamma + \sqrt{\mu}c(V)$ for some $\gamma = \gamma(e, \text{deg}_U^{(2)}) > 0$, which we fix now. This new parameter γ quantifies a safe distance (depending on the total energy $e > 0$ and the degree of the trigonometric polynomial) to the region, *opposite* to where (i) the change of variables $\Phi_\mu^{(2)}$ has bounded derivative (cf. equation (4.7)) and (ii) the function $h_\mu^{(2)}$ is bounded from below (cf. equation (4.12)). In §4.3, we will have two such parameters, γ_1, γ_2 , for both coordinates directions which get transformed by some Φ .

By Assumption (P), we have

$$u_{\varepsilon,c} = u_c^{(0)} + \varepsilon u_c^{(1)} + \mathcal{O}_c(\varepsilon^2)$$

with $u_c^{(0)} = u_{0,c}$ and since $H_0(\theta, I)$ is integrable (and written in action-angle coordinates), one can choose $u_{0,c} \equiv 0$. Therefore, by Assumption (P) again, we expand the Hamilton Jacobi equation in equation (2.9) as

$$\begin{aligned} \alpha_\varepsilon(c) &= H_\varepsilon(\theta, c + \nabla u_{\varepsilon,c}(\theta)) \\ &= \frac{|\partial_{\theta^1} u_{\varepsilon,c}(\theta) + c_1|^2}{2} + h_\mu^{(2)}(\partial_{\theta^2} u_{\varepsilon,c}(\theta) + c_2) + \varepsilon U(\theta^1, x_\mu^2(\theta^2, \partial_{\theta^2} u_{\varepsilon,c}(\theta) + c_2)) \\ &= \frac{c_1^2}{2} + h_\mu^{(2)}(c_2) + \varepsilon \langle (c_1, \nabla h_\mu^{(2)}(c_2)), \nabla u_{\varepsilon,c}^{(1)}(\theta) \rangle + \varepsilon U(\theta^1, x_\mu^2(\theta^2, c_2)) \\ &\quad + \mathcal{O}(\|(\nabla^2 h_\mu^{(2)})|_{\{|c_2| > \gamma + \sqrt{\mu}c(V)\}}\|_{C^0} \varepsilon^2) + \mathcal{O}(\|(\partial_{I_2} \Phi_\mu^{(2)})|_{\{|c_2| > \gamma + \sqrt{\mu}c(V)\}}\|_{C^0} \varepsilon^2). \end{aligned} \tag{4.7}$$

Since $|c_2| > \gamma + \sqrt{\mu}c(V)$, both error terms are of the order $\mathcal{O}_\gamma(\varepsilon^2)$.

Analogously to the proof of Theorem 2.2, we thus obtain the first-order equation

$$[U]_0 = \langle (c_1, \nabla h_\mu^{(2)}(c_2)), \nabla u_{\varepsilon,c}^{(1)}(\theta) \rangle + U(\theta^1, x_\mu^2(\theta^2, c_2)), \tag{4.8}$$

where the constant $\alpha^{(1)} \equiv [U]_0$ is given in equation (D.6) in Proposition D.2 (see also [50]). Just as in the proof of Theorem 2.2, after averaging equation (4.8) over the trajectory $\theta(t) = \theta_0 + \omega t \in \mathbb{T}^2$, with initial position $\theta_0 \in \mathbb{T}^2$ and where ω is chosen according to equation (4.6), such that the period T_ω satisfies $T_\omega \cdot \omega = (n, m)$, we find

$$\int_0^1 (U(\theta_0^1 + nt, x_\mu^2(\theta_0^2 + mt, c_2)) - [U]_0) dt = 0 \tag{4.9}$$

for all $\theta_0 = (\theta_0^1, \theta_0^2) \in \mathbb{T}^2$.

Finally, analogously to §4.1, we may assume without loss of generality $[U]_0 = 0$ and observe that

$$\int_0^1 U_{\text{sep}}(\theta_0^1 + nt, x_\mu^2(\theta_0^2 + mt, c_2)) dt = [U_{\text{sep}}]_0 \quad \text{for all } (\theta_0^1, \theta_0^2) \in \mathbb{T}^2$$

holds *generally* (that is, independent of the first-order relation in equation (4.8)) by a simple calculation based on equation (D.6) in Proposition D.2 (see also Remark D.1). We can thus split off the separable part U_{sep} of U defined in equation (4.5) and assume that $S_U = S_{U,0}$ in the following. Hence, the third step consists of showing that $S_U = S_{U,0} = \emptyset$.

Step (iii). We begin this final step with performing a Fourier decomposition in equation (4.9) such that we obtain

$$\sum_{k_1 \neq 0} \left[\sum_{0 \neq |k_2| \leq \text{deg}_U^{(2)}} U_{k_1, k_2} \int_0^1 e^{i2\pi k_1 nt} e^{i2\pi k_2 x_\mu^2(\theta_0^2 + mt, c_2)} dt \right] e^{i2\pi k_1 \theta_0^1} = 0$$

for all $(\theta_0^1, \theta_0^2) \in \mathbb{T}^2$,

which implies that $[\dots] = 0$ for every $k_1 \in \mathbb{Z} \setminus \{0\}$ and $\theta_0^2 \in \mathbb{T}$.

After having eliminated $\theta_0^1 \in \mathbb{T}$, we now fix some $k_1 \in \mathbb{Z} \setminus \{0\}$ and consider the family of functions $(f_{k_2}^{(k_1, \mu)})_{0 \neq |k_2| \leq \text{deg}_U^{(2)}}$ in the Hilbert space $L^2(\mathbb{T})$, where

$$f_{k_2}^{(k_1, \mu)} : \mathbb{T} \rightarrow \mathbb{C}, \quad \theta_0^2 \mapsto \sum_{\substack{(n,m) \in \mathcal{B}_0(S_U^\perp) \\ \exists 0 \neq |\tilde{k}_2| \leq \text{deg}_U^{(2)} : k_1 n + \tilde{k}_2 m = 0}} \int_0^1 e^{i2\pi k_1 nt} e^{i2\pi k_2 x_\mu^2(\theta_0^2 + mt, c_2)} dt. \tag{4.10}$$

Note that the sum in equation (4.10) is finite by Assumption (A2) (more precisely, it ranges over at most $2 \cdot \text{deg}_U^{(2)}$ elements from $\mathcal{B}_0(S_U^\perp)$) and we suppressed the dependence of $|c_2| > \gamma + \sqrt{\mu}c(V)$ on $(n, m) \in \mathcal{B}_0(S_U^\perp)$ from the notation (recall equation (4.6)).

In this way, the problem of proving Theorem 2.3, that is, justifying $S_{U,0} = \emptyset$, reduces to a question about linear independence for the family of functions in equation (4.10) in the Hilbert space $L^2(\mathbb{T})$. Recall that the family $(f_{k_2}^{(k_1, \mu)})_{0 \neq |k_2| \leq \text{deg}_U^{(2)}}$ being linearly independent is equivalent to the *Gram matrix*

$$G^{(k_1, \mu)} = (G_{k_2, k_2'}^{(k_1, \mu)})_{0 \neq |k_2|, |k_2'| \leq \text{deg}_U^{(2)}} \quad \text{with } G_{k_2, k_2'}^{(k_1, \mu)} := \langle f_{k_2}^{(k_1, \mu)}, f_{k_2'}^{(k_1, \mu)} \rangle_{L^2(\mathbb{T})} \tag{4.11}$$

being of full rank, where $\langle g, h \rangle_{L^2(\mathbb{T})}$ denotes the standard inner product of $g, h \in L^2(\mathbb{T})$.

LEMMA 4.2. *There exists $\tilde{\mu} = \tilde{\mu}(C_2, \text{deg}_U^{(2)}, e) > 0$ such that for all $\mu \in [0, \tilde{\mu}]$, the Gram matrix $G^{(k_1, \mu)}$ from equation (4.11) is of full rank.*

Proof. Using the version of Lemma B.1 for the inverse function, we find that

$$\|e^{i2\pi k_2 x_\mu^2(\cdot, c_2)} - e^{i2\pi k_2 \cdot}\|_{C^0} = \mathcal{O}\left(\text{deg}_U^{(2)} \frac{\mu \|V\|_{C^0}}{h_\mu(\gamma + \sqrt{\mu}c(V))}\right) =: \mathcal{O}(\mu) \tag{4.12}$$

uniformly in $|k_2| \leq \text{deg}_U^{(2)}$ and $(n, m) \in \mathcal{B}_0(S_U^\perp)$.

With a slight abuse of notation for the error term, the elements $G_{k_2, k_2'}^{(k_1, \mu)}$ of the Gram matrix can thus be computed as

$$\int_0^1 d\theta_0^2 \left(\left[\sum_{(n, m)} \int_0^1 dt e^{-i2\pi k_1 n t} (e^{-i2\pi k_2 m t} + \mathcal{O}(\mu)) \right] e^{-i2\pi k_2 \theta_0^2} \right. \\ \left. \times e^{i2\pi k_2' \theta_0^2} \left[\sum_{(n', m')} \int_0^1 dt' e^{i2\pi k_1 n' t'} (e^{i2\pi k_2' m' t'} + \mathcal{O}(\mu)) \right] \right),$$

where the summations over (n, m) and (n', m') are understood as in equation (4.10). Using that for every $(k_1, k_2) \in \mathcal{S}_{U,0}$, there exist exactly two elements from $\mathcal{B}_0(S_U^\perp)$ (differing by a sign), we can evaluate both brackets $[\dots]$ being equal to $2 + \mathcal{O}(\text{deg}_U^{(2)} \mu)$.

From this, we conclude that

$$G_{k_2, k_2'}^{(k_1, \mu)} = \int_0^1 d\theta_0^2 [2 + \mathcal{O}(\text{deg}_U^{(2)} \mu)] e^{i2\pi(k_2' - k_2)\theta_0^2} [2 + \mathcal{O}(\text{deg}_U^{(2)} \mu)] = 4\delta_{k_2, k_2'} + \mathcal{O}(\text{deg}_U^{(2)} \mu).$$

Therefore, going back to equation (4.12), we infer the existence of $\tilde{\mu} = \tilde{\mu}(C_2, \text{deg}_U^{(2)}, e) > 0$ such that for all $\mu \in [0, \tilde{\mu}]$, the Gram matrix $G^{(k_1, \mu)}$ from equation (4.11) is of full rank. □

Since $k_1 \in \mathbb{Z} \setminus \{0\}$ was arbitrary and Lemma 4.2 is independent of k_1 , this concludes the proof of Theorem 2.3(a).

For part (b), we note that $e^{i2\pi k_2 x_\mu^2(\theta_0^2 + mt, c_2)}$ from equation (4.10) depends analytically on μ (see Appendix C). Therefore, the function $\mu \mapsto G^{(k_1, \mu)}$ mapping to the Gram matrix in equation (4.11), for every fixed $k_1 \in \mathbb{Z} \setminus \{0\}$, is also analytic. (Using joint continuity of $(u, \mu) \mapsto e^{i2\pi k_2 x_\mu^2(u, c_2)}$, it is an elementary exercise to show that the integrals over t and θ_0^2 do not disturb the analyticity in μ .) This in turn implies that $\det(G^{(k_1, \mu)})$ is analytic in μ and thus, since $\det(G^{(k_1, \mu)}) \neq 0$ for $\mu \in (0, \tilde{\mu})$ (see Lemma 4.2), we find that the zero set

$$\mathcal{E}_0^{(k_1)} := \{\mu \in (0, \infty) \mid \det(G^{(k_1, \mu)}) = 0\} \subset (\tilde{\mu}, \infty)$$

of $\mu \mapsto \det(G^{(k_1, \mu)})$ is at most countable (finite in every compact subset), that is, in particular, a set of zero measure. Finally, setting

$$\mathcal{E}_0 := \bigcup_{k_1 \in \mathbb{Z} \setminus \{0\}} \mathcal{E}_0^{(k_1)},$$

we constructed the exceptional null set, for which the conclusion $\mathcal{S}_{U,0} = \emptyset$ is not valid.

This finishes the proof of Theorem 2.3(b).

4.3. *Proof of Theorem 2.4.* As above, the argument is divided into three steps.

Step (i). We fix an energy $e > 0$ and consider the region of the phase space, where both one-dimensional subsystems are rotating, that is,

$$\frac{p_1^2}{2} - \mu_1 V_1(x^1) = e^{(1)} > 0 \quad \text{and} \quad \frac{p_2^2}{2} - \mu_2 V_2(x^2) = e^{(2)} > 0,$$

such that we have $e = e^{(1)} + e^{(2)}$. In a neighborhood of each of the two Liouville tori characterized by $H_0 = e$ and $(p_1^2/2) - \mu_1 V_1(x^1) = e^{(1)}$, we can find two changes of variables $(x^1, p_1) = \Phi_{\mu_1}^{(1)}(\theta^1, I_1)$ and $(x^2, p_2) = \Phi_{\mu_2}^{(2)}(\theta^2, I_2)$ such that the Hamiltonian function H_0 gets transformed in action-angle coordinates, that is,

$$H_0(\Phi_{\mu_1}^{(1)}(\theta^1, I_1), \Phi_{\mu_2}^{(2)}(\theta^2, I_2)) = h_{\mu_1}^{(1)}(I_1) + h_{\mu_2}^{(2)}(I_2)$$

for some smooth functions $h_{\mu_1}^{(1)}$ and $h_{\mu_2}^{(2)}$, which agree with Mather’s α -functions for the one-dimensional subsystem described by the Hamiltonians $(p_1^2/2) - \mu_1 V(x^1)$, respectively $(p_2^2/2) - \mu_2 V(x^2)$ (see Appendix D). As in the proof of Theorem 2.3, the change in the order of the four arguments of H_0 should not lead to confusion.

Now, the perturbed Hamiltonian takes the form

$$H_\varepsilon(\Phi_{\mu_1}^{(1)}(\theta^1, I_1), \Phi_{\mu_2}^{(2)}(\theta^2, I_2)) = h_{\mu_1}^{(1)}(I_1) + h_{\mu_2}^{(2)}(I_2) + \varepsilon U(x_{\mu_1}^1(\theta^1, I_1), x_{\mu_2}^2(\theta^2, I_2)),$$

where we write $x_{\mu_i}^i(\theta^i, I_i)$ for the first component of $\Phi_{\mu_i}^{(i)}(\theta^i, I_i)$, $i \in \{1, 2\}$.

Step (ii). Analogously to the proof of Theorem 2.3, we assume without loss of generality that the 1- and 2-degree $\text{deg}_U^{(1)}$ and $\text{deg}_U^{(2)}$ of U are at least 1 (recall equation (2.8)), as otherwise, we had $U(x) = U_2(x^2)$ or $U(x) = U_1(x^1)$ and Theorem 2.4 was proven. Then, for any $(n, m) \in \mathcal{B}_0(\mathcal{S}_U^1)$, in particular with $|m| \leq \text{deg}_U^{(1)}$ and $|n| \leq \text{deg}_U^{(2)}$, we can find (in the isoenergy manifold T_{e_ε} with energy $e = e_\varepsilon$ and $\varepsilon = \varepsilon_k$ for some $k \in \mathbb{N}$) a rational invariant Liouville torus with rotation vector $\omega = (\omega_1, \omega_2)$ which satisfies

$$\frac{\omega_1}{\omega_2} = \frac{n}{m} \in \mathbb{Q} \quad \text{and} \quad \omega = (\nabla h_{\mu_1}^{(1)}(c_1), \nabla h_{\mu_2}^{(2)}(c_2)) \tag{4.13}$$

for some $c \in H^1(\mathbb{T}^2, \mathbb{R}) \cong \mathbb{R}^2$ with $|c_1| > \gamma_1 + \sqrt{\mu_1}c(V_1)$ and $|c_2| > \gamma_2 + \sqrt{\mu_2}c(V_2)$ for some $\gamma_1 = \gamma_1(e, \text{deg}_U^{(1)}) > 0$, respectively $\gamma_2 = \gamma_2(e, \text{deg}_U^{(2)}) > 0$, which we fix now (see the paragraph below equation (4.6) for a discussion of the γ parameters).

By Assumption (P), we have

$$u_{\varepsilon, c} = u_c^{(0)} + \varepsilon u_c^{(1)} + O_c(\varepsilon^2)$$

with $u_c^{(0)} = u_{0, c}$ and since $H_0(\theta, I)$ is integrable (and written in action-angle coordinates), one can choose $u_{0, c} \equiv 0$. Therefore, by Assumption (P) again, we expand the Hamilton Jacobi in equation (2.9) as

$$\begin{aligned} \alpha_\varepsilon(c) &= H_\varepsilon(\theta, c + \nabla u_{\varepsilon, c}(\theta)) \\ &= h_{\mu_1}^{(1)}(\partial_{\theta^1} u_{\varepsilon, c}(\theta) + c_1) + h_{\mu_2}^{(2)}(\partial_{\theta^2} u_{\varepsilon, c}(\theta) + c_2) \\ &\quad + \varepsilon U(x_{\mu_1}^1(\theta^1, \partial_{\theta^1} u_{\varepsilon, c}(\theta) + c_1), x_{\mu_2}^2(\theta^2, \partial_{\theta^2} u_{\varepsilon, c}(\theta) + c_2)) \end{aligned}$$

$$\begin{aligned}
 &= \sum_{i=1}^2 h_{\mu_i}^{(i)}(c_i) + \varepsilon \langle (\nabla h_{\mu_1}^{(1)}(c_1), \nabla h_{\mu_2}^{(2)}(c_2)), \nabla u_{\varepsilon, c}^{(1)}(\theta) \rangle + \varepsilon U(x_{\mu_1}^1(\theta^1, c_1), x_{\mu_2}^2(\theta^2, c_2)) \\
 &\quad + \mathcal{O}\left(\sum_{i=1}^2 (\|\nabla^2 h_{\mu_i}^{(i)}\|_{\{|c_i| > \gamma_i + \sqrt{\mu_i} c(V_i)\}} \|_{C^0} + \|\partial_{t_i} \Phi_{\mu_i}^{(i)}\|_{\{|c_i| > \gamma_i + \sqrt{\mu_i} c(V_i)\}} \|_{C^0}) \varepsilon^2\right).
 \end{aligned}$$

Since $|c_i| > \gamma_i + \sqrt{\mu_i} c(V_i)$, the error term is of order $\mathcal{O}_{\gamma_i}(\varepsilon^2)$.

Analogously to the proofs of Theorems 2.2 and 2.3, we thus obtain the first-order equation

$$[U]_0 = \langle (\nabla h_{\mu_1}^{(1)}(c_1), \nabla h_{\mu_2}^{(2)}(c_2)), \nabla u_{\varepsilon, c}^{(1)}(\theta) \rangle + U(x_{\mu_1}^1(\theta^1, c_1), x_{\mu_2}^2(\theta^2, c_2)), \tag{4.14}$$

where the constant $\alpha^{(1)} \equiv [U]_0$ is again given by equation (D.6) in Proposition D.2 (see also [50]). Just as in the proof of Theorems 2.2 and 2.3, after averaging equation (4.14) over the trajectory $\theta(t) = \theta_0 + \omega t \in \mathbb{T}^2$, with initial position $\theta_0 \in \mathbb{T}^2$ and where ω is chosen according to equation (4.13) such that the period T_ω satisfies $T_\omega \cdot \omega = (n, m)$,

$$\int_0^1 (U(x_{\mu_1}^1(\theta_0^1 + nt, c_1), x_{\mu_2}^2(\theta_0^2 + mt, c_2)) - [U]_0) dt = 0 \tag{4.15}$$

for all $\theta_0 = (\theta_0^1, \theta_0^2) \in \mathbb{T}^2$.

Finally, analogously to §§4.1 and 4.2, we may assume without loss of generality $[U]_0 = 0$ and observe that

$$\int_0^1 U_{\text{sep}}(x_{\mu_1}^1(\theta_0^1 + nt, c_1), x_{\mu_2}^2(\theta_0^2 + mt, c_2)) dt = [U_{\text{sep}}]_0 \quad \text{for all } (\theta_0^1, \theta_0^2) \in \mathbb{T}^2$$

holds *generally* (that is, independent of the first-order relation in equation (4.14)) by a simple calculation based on equation (D.6) in Proposition D.2 (see also Remark D.1). We can thus split off the separable part U_{sep} of U defined in equation (4.5) and assume that $\mathcal{S}_U = \mathcal{S}_{U,0}$ in the following. Hence, the third step consists of showing that $\mathcal{S}_U = \mathcal{S}_{U,0} = \emptyset$.

Step (iii). We begin this final step with performing a Fourier decomposition in equation (4.15), such that we obtain

$$\sum_{\substack{0 \neq |k_1| \leq \text{deg}_U^{(1)} \\ 0 \neq |k_2| \leq \text{deg}_U^{(2)}}} U_{k_1, k_2} \int_0^1 e^{i2\pi k_1 x_{\mu_1}^1(\theta_0^1 + nt, c_1)} e^{i2\pi k_2 x_{\mu_2}^2(\theta_0^2 + mt, c_2)} dt = 0 \quad \text{for all } (\theta_0^1, \theta_0^2) \in \mathbb{T}^2.$$

Analogously to the proof of Theorem 2.3, we now consider the family of functions

$$(f_{k_1, k_2}^{(\mu_1, \mu_2)})_{0 \neq |k_1| \leq \text{deg}_U^{(1)}, 0 \neq |k_2| \leq \text{deg}_U^{(2)}}$$

in the Hilbert space $L^2(\mathbb{T}^2)$, where

$$f_{k_1, k_2}^{(\mu_1, \mu_2)} : \mathbb{T}^2 \rightarrow \mathbb{C}, \quad (\theta_0^1, \theta_0^2) \mapsto \sum_{(n, m) \in \mathcal{B}_0(\mathcal{S}_U^1)} \int_0^1 e^{i2\pi k_1 x_{\mu_1}^1(\theta_0^1 + nt, c_1)} e^{i2\pi k_2 x_{\mu_2}^2(\theta_0^2 + mt, c_2)} dt. \tag{4.16}$$

Note that the sum in equation (4.16) is finite by Assumption (A3) (more precisely, it ranges over the at most $(2 \deg_U^{(1)}) \cdot (2 \deg_U^{(2)})$ elements from $\mathcal{B}_0(\mathcal{S}_U^\perp)$ and we suppressed the dependence of $|c_i| > \gamma_i + \sqrt{\mu_i} \mathfrak{c}(V_i)$ on $(n, m) \in \mathcal{B}_0(\mathcal{S}_U^\perp)$ from the notation (recall equation (4.13)).

In this way, the problem of proving Theorem 2.4, that is, justifying $\mathcal{S}_{U,0} = \emptyset$, reduces to a question about linear independence for the family of functions in equation (4.16) in the Hilbert space $L^2(\mathbb{T}^2)$. Recall that the family $(f_{k_1, k_2}^{(\mu)})_{(k_1, k_2)}$ being linearly independent is equivalent to the Gram matrix $G^{(\mu)}$ with entries

$$G_{(k_1, k_2), (k'_1, k'_2)}^{(\mu_1, \mu_2)} := \langle f_{k_1, k_2}^{(\mu_1, \mu_2)}, f_{k'_1, k'_2}^{(\mu_1, \mu_2)} \rangle_{L^2(\mathbb{T}^2)} \quad \text{for } 0 \neq |k_i|, |k'_i| \leq \deg_U^{(i)}, i \in \{1, 2\}, \tag{4.17}$$

being of full rank, where $\langle g, h \rangle_{L^2(\mathbb{T}^2)}$ denotes the standard inner product of $g, h \in L^2(\mathbb{T}^2)$.

LEMMA 4.3. *There exist $\tilde{\mu}_i = \tilde{\mu}(C_i, \deg_U^{(1)}, \deg_U^{(2)}, e) > 0$ such that for all $\mu_i \in [0, \tilde{\mu}_i]$, $i \in \{1, 2\}$, the Gram matrix $G^{(\mu_1, \mu_2)}$ from equation (4.17) is of full rank.*

Proof. Using the version of Lemma B.1 for the inverse function, we find that

$$\|e^{i2\pi k_i x_i^{i, c_i}} - e^{i2\pi k_i \cdot}\|_{C^0} = \mathcal{O}\left(\deg_U^{(i)} \frac{\mu_i \|V_i\|_{C^0}}{h_{\mu_i}^{(i)}(\gamma_i + \sqrt{\mu_i} \mathfrak{c}(V_i))}\right) =: \mathcal{O}(\mu_i) \tag{4.18}$$

uniformly in $|k_i| \leq \deg_U^{(i)}$ and $(n, m) \in \mathcal{B}_0(\mathcal{S}_U^\perp)$.

Similarly to Lemma 4.2, with a slight abuse of notation for the error term, the elements $G_{(k_1, k_2), (k'_1, k'_2)}^{(\mu_1, \mu_2)}$ of the Gram matrix can thus be computed as

$$\int_0^1 d\theta_0^1 \int_0^1 d\theta_0^2 \left(\left[\sum_{(n, m)} \int_0^1 dt (e^{-i2\pi k_1 n t} + \mathcal{O}(\mu_1))(e^{-i2\pi k_2 m t} + \mathcal{O}(\mu_2)) \right] e^{-i2\pi k_1 \theta_0^1} e^{-i2\pi k_2 \theta_0^2} \right. \\ \left. \times e^{i2\pi k'_1 \theta_0^1} e^{i2\pi k'_2 \theta_0^2} \left[\sum_{(n', m')} \int_0^1 dt' (e^{i2\pi k'_1 n' t'} + \mathcal{O}(\mu_1))(e^{i2\pi k'_2 m' t'} + \mathcal{O}(\mu_2)) \right] \right),$$

where the summations over (n, m) and (n', m') are understood as in equation (4.16). Using that for every $(k_1, k_2) \in \mathcal{S}_{U,0}$, there exist exactly two elements from $\mathcal{B}_0(\mathcal{S}_U^\perp)$ (differing by a sign), we can evaluate both brackets $[\cdot \cdot \cdot]$ being given by

$$2 + \mathcal{O}(\deg_U^{(1)} \deg_U^{(2)} \mu_1) + \mathcal{O}(\deg_U^{(1)} \deg_U^{(2)} \mu_2) =: 2 + \mathcal{O}(\deg_U^{(1)} \deg_U^{(2)} (\mu_1 + \mu_2)),$$

after absorption of the second-order error in the first-order ones.

From this, we conclude that

$$G_{(k_1, k_2), (k'_1, k'_2)}^{(\mu_1, \mu_2)} = \int_0^1 d\theta_0^1 \int_0^1 d\theta_0^2 ([2 + \mathcal{O}(\deg_U^{(1)} \deg_U^{(2)} (\mu_1 + \mu_2))] e^{i2\pi(k'_1 - k_1)\theta_0^1} \\ \times e^{i2\pi(k'_2 - k_2)\theta_0^2} [2 + \mathcal{O}(\deg_U^{(1)} \deg_U^{(2)} (\mu_1 + \mu_2))]) \\ = 4\delta_{k_1, k'_1} \delta_{k_2, k'_2} + \mathcal{O}(\deg_U^{(1)} \deg_U^{(2)} (\mu_1 + \mu_2)).$$

Therefore, going back to equation (4.18), we infer the existence of $\tilde{\mu}_i = \tilde{\mu}(\mathcal{C}_i, \deg_U^{(1)}, \deg_U^{(2)}, e) > 0, i \in \{1, 2\}$, such that for all $\mu_i \in [0, \tilde{\mu}_i]$, the Gram matrix $G^{(\mu_1, \mu_2)}$ from equation (4.17) is of full rank. \square

This finishes the proof of Theorem 2.4(a). For part (b), similarly to the proof of Theorem 2.3(b), we observe that for every fixed $\mu_1 \in [0, \tilde{\mu}_1]$, the function $\mu_2 \mapsto \det(G^{(\mu_1, \mu_2)})$ is analytic. Since $\det(G^{(\mu_1, \mu_2)}) \neq 0$ for $\mu_2 \in (0, \tilde{\mu}_2)$ (see Lemma 4.3), we find that the zero set

$$\mathcal{E}_0^{(\mu_1)} := \{\mu_2 \in (0, \infty) \mid \det(G^{(\mu_1, \mu_2)}) = 0\} \subset (\tilde{\mu}_2, \infty)$$

of $\mu_2 \mapsto \det(G^{(\mu_1, \mu_2)})$ is at most countable (finite in every compact subset), that is, in particular, a (one-dimensional) set of zero measure.

Finally, for part (c), we note that, similarly to the proof of Theorem 2.3(b) and by means of Hartogs’s theorem on separate analyticity [56] (a separately analytic function is jointly analytic), the function $(\mu_1, \mu_2) \mapsto \det(G^{(\mu_1, \mu_2)})$ is (jointly) analytic. Since $\det(G^{(\mu_1, \mu_2)}) \neq 0$ for $(\mu_1, \mu_2) \in (0, \tilde{\mu}_1) \times (0, \tilde{\mu}_2)$ (see Lemma 4.3), we find that the zero set

$$\mathcal{E}_0 := \{(\mu_1, \mu_2) \in (0, \infty) \times (0, \infty) \mid \det(G^{(\mu_1, \mu_2)}) = 0\} \subset (\tilde{\mu}_1, \infty) \times (\tilde{\mu}_2, \infty)$$

of $(\mu_1, \mu_2) \mapsto \det(G^{(\mu_1, \mu_2)})$ is a (two-dimensional) set of zero measure.

This concludes the proof of Theorem 2.4(c).

5. Concluding remarks and outlook

We have shown that integrable deformations of Liouville metrics on \mathbb{T}^2 are Liouville metrics—at least when more restrictive conditions on the unperturbed metric are balanced with more general conditions on the perturbation. Removing this balancing, that is, showing that *arbitrary* integrable deformations of *arbitrary* Liouville metrics remain of Liouville type, is an interesting problem for future investigations resolving the conjecture proposed at the end of §2. This would require stronger versions of Lemmas 4.2 and 4.3 in two senses.

- (a) Allow for possibly *infinitely many non-zero Fourier coefficients* and refrain from restricting to trigonometric polynomials. A resolution of this issue has been found in the context of the *perturbative Birkhoff conjecture* [10, 60] concerning integrable billiards. Here, the authors studied the matrix of correlations between the standard basis $(e^{i2\pi kx})_{k \in \mathbb{Z}}$ of $L^2(\mathbb{T})$ and certain deformed dynamical modes (given as some kind of Jacobi elliptic function, see Appendix C of the [arXiv: 2210.02961](https://arxiv.org/abs/2210.02961) version of this article), corresponding to $e^{i2\pi k_i x_{\mu_i}^i(\cdot, c_i)}$ in Lemmas 4.2 and 4.3. Exponential estimates for the entries of this matrix (obtained from considering the maximal width of a strip of analyticity around the real axis for the dynamical modes) allowed to prove a suitable full-rank lemma, also for infinitely many coefficients.
- (b) Allow *arbitrary* $\tilde{\mu}_i > 0$ and refrain from restricting to small ones. Also for this issue, a potential resolution might be found by analytically extending action-angle coordinates to the complex plane and exploiting their singularities away from the

real axis. However, this requires the potentials V_i in the unperturbed Hamiltonian to be restrictions of holomorphic functions and, as such, way more special than generic $V_i \in C^2(\mathbb{T})$.

Moreover, we note that in [60], the authors also outlined a potential strategy for proving the classical (non-perturbative) Birkhoff conjecture, which might possibly be adapted for proving a suitably weakened version of the folklore conjecture given in §3.

We end this section with a brief list of open problems being related to the main results of the present paper.

- (i) As described above, it is a natural follow-up problem to extend our results to the situation where *arbitrary* integrable deformations of *arbitrary* Liouville metrics remain of Liouville type, that is, remove the restricting assumptions from Assumptions (A1)–(A3) and prove the conjecture formulated at the end of §2.
- (ii) In particular, starting with (the time-independent version of) Arnold's example [6] for diffusion,

$$H_0(x, p) = \frac{p_1^2}{2} + \frac{p_2^2}{2} - \mu(1 - \cos(2\pi x^2)),$$

is it possible to deduce rigidity, similarly to Theorem 2.3, but without restricting to the perturbation being a trigonometric polynomial in x^2 and any smallness condition on $\mu \in [0, 1]$? In this case, the *full rank lemma* might be obtained by proving non-degeneracy of certain infinite-dimensional matrices, which have Fourier coefficients of powers of Jacobi elliptic functions (see Appendix C of the [arXiv: 2210.02961](https://arxiv.org/abs/2210.02961) version of this article) as their entries.

- (iii) In view of the non-trivial examples of magnetic geodesic flows found in [3] and the potential counterexample constructed in [35], it is a major task to completely settle the *folklore conjecture* mentioned in §§1 and 3, that is, clarify which part is only 'folklore' and which part is 'real'.
- (iv) In particular, the main result of [35], which we stated in Theorem 3.8, should be extended to show that the system is really integrable on an open set in the phase space and not only on an isoenergy manifold. Furthermore, it remains open, whether the PDEs underlying the examples in [3] can be solved with zero magnetic fields or not and thus potentially disproves the folklore conjecture.
- (v) For our main results, we assumed the preservation of rational invariant tori 'outside the eye of the pendulum' (cf. Figure 1). Can one obtain the same result, if only tori 'inside the eye' are preserved?
- (vi) An alternative approach to the one chosen here could be to study perturbations of the additional first integral in equation (2.2), that is, write $F_\varepsilon = F_0 + \varepsilon F_1 + \mathcal{O}(\varepsilon^2)$ and use the vanishing of the Poisson bracket $\{H_\varepsilon, F_\varepsilon\} = 0$ with $H_\varepsilon = H_0 + \varepsilon U$ to obtain the first-order equation

$$\{H_0, F_1\} + \{U, F_0\} = 0$$

for the perturbing potential U .

- (vii) Does there exist a Riemannian metric g on \mathbb{T}^2 such that its geodesic flow admits hyperbolic periodic orbits of *at least three* different homotopy types? If yes, does

there exist a Liouville metric with this property? (These questions were suggested by Vadim Kaloshin.)

Acknowledgments. I am very grateful to Vadim Kaloshin for suggesting the topic, his guidance during this project, and many helpful comments on an earlier version of the manuscript. Moreover, I would like to thank Comlan Edmond Koudjina and Volodymyr Riabov for interesting discussions. Partial financial support by the ERC Advanced Grant ‘RMTBeyond’ No. 101020331 is gratefully acknowledged. This project received funding from the European Research Council (ERC) ERC Grant No. 885707.

A. Appendix. Generalization to higher dimensions

Our results from §2 immediately generalize to higher dimensions $d \geq 3$. In this setting, we define the Hamiltonian function

$$H_0(x, p) = \sum_{i=1}^d \left(\frac{p_i^2}{2} - \mu_i V_i(x^i) \right) \tag{A.1}$$

on $T^*\mathbb{T}^d$, where $\mu_i \in [0, \infty)$ are parameters, and $V_i \in C^2(\mathbb{T})$ with $\|V_i\| \leq C_i$, $V_i \geq 0$ are Morse functions (or constant). We may assume without loss of generality that $\min_{x^i} V_i(x^i) = 0$. The system in equation (A.1) is clearly integrable, since additional first integrals can easily be found as

$$F_i(x, p) = \frac{p_i^2}{2} - \mu_i V_i(x^i), \quad i \in \{1, \dots, d-1\}.$$

Completely analogous to §2, we perturb the integrable system in equation (A.1) as $H_\varepsilon = H_0 + \varepsilon U$ with $\varepsilon \in \mathbb{R}$ by an additive potential $U \in C^2(\mathbb{T}^d)$, which we assume to have an absolutely convergent Fourier series.

Now, the analogs of the assumptions in §2 read as follows.

(1) *Assumptions on the perturbed Hamiltonian function H_ε .* Let $H_0 \in C^2(T^*\mathbb{T}^d)$ denote the Hamiltonian function from equation (A.1) with $\|V_i\| \leq C_i$ and $\mu_i \in [0, \tilde{\mu}_i]$ for some $\tilde{\mu}_i \in [0, \infty)$, $i \in \{1, \dots, d\}$, and $U \in C^2(\mathbb{T}^d)$ be a perturbing potential, which satisfies the following assumption.

(A4) If $\tilde{\mu}_i = 0$ for the first $0 \leq d_{\text{flat}} \leq d$ indices, there exist $d^{(k)} \geq 0$ for $k \in \{d_{\text{flat}} + 1, \dots, d\}$ such that

$$S_U \subset \mathbb{Z}^{d_{\text{flat}}} \times ([-d^{(d_{\text{flat}}+1)}, d^{(d_{\text{flat}}+1)}] \times \dots \times [-d^{(d)}, d^{(d)}]) \tag{A.2}$$

that is $U \in C^2(\mathbb{T}^d)$ is a trigonometric polynomial in the last $(d - d_{\text{flat}})$ variables.

As in §2, we denote the minimum over all $d^{(i)}$ such that equation (A.2) holds as $\text{deg}_U^{(i)}$ and call it the i -degree of U .

Note that Proposition 2.1 immediately generalizes to higher dimensions, such that we can formulate the analog of Assumption (P) as follows.

(2) *Assumptions on the preserved integrability of H_ε .* Let $H_0 \in C^2(T^*\mathbb{T}^d)$ denote the Hamiltonian function from equation (A.1) satisfying Assumption (A4), and U a perturbing potential, such that the following statement concerning the perturbed Hamilton–Jacobi

equation (HJE)

$$\alpha_\varepsilon(\mathbf{c}) = H_\varepsilon(x, \mathbf{c} + \nabla_x u_{\varepsilon, \mathbf{c}}(x)) \tag{A.3}$$

as well as the preserved integrability of H_ε is satisfied.

(P') There exists an energy $e > 0$, such that for every $\mathbf{b} \in \mathcal{B}_0(\mathcal{S}_U^\perp)$ (recall equation (2.6)) there exists a sequence $(\varepsilon_k)_{k \in \mathbb{N}}$ with $\varepsilon_k \neq 0$ but $\varepsilon_k \rightarrow 0$ such that for any $\mu_i \in [0, \tilde{\mu}_i]$ we have the following:

- (i) The \mathbf{b} -torus from (the analog of) Proposition 2.1 characterized by $\mathbf{c} \in H^1(\mathbb{T}^d, \mathbb{R}) \cong \mathbb{R}^d$ with

$$|c_i| > \sqrt{\mu_i} c(V_i) \tag{A.4}$$

in the isoenergy submanifold T_e is preserved under the sequence of deformations $(H_{\varepsilon_k})_{k \in \mathbb{N}}$, where $c(V_i)$ is defined in equation (D.3).

- (ii) For $\mathbf{c} \in H^1(\mathbb{T}^d, \mathbb{R})$ satisfying equation (A.4), Mather's α -function and a solution $u_{\varepsilon, \mathbf{c}} : \mathbb{T}^d \rightarrow \mathbb{R}$ of the HJE in equation (A.3) can be expanded to first order in ε , that is,

$$u_{\varepsilon, \mathbf{c}} = u_{\mathbf{c}}^{(0)} + \varepsilon u_{\mathbf{c}}^{(1)} + \mathcal{O}_{\mathbf{c}}(\varepsilon^2) \quad \text{and} \quad \alpha_\varepsilon = \alpha^{(0)} + \varepsilon \alpha^{(1)} + \mathcal{O}(\varepsilon^2),$$

where $u_{\mathbf{c}}^{(0)}, u_{\mathbf{c}}^{(1)} \in C^{1,1}(\mathbb{T}^d)$ and $\mathcal{O}_{\mathbf{c}}(\varepsilon^2)$ is understood in $C^{1,1}$ -sense.

We can now formulate our generalized main result.

THEOREM A.1. *Let H_ε satisfy Assumption (A4) and Assumption (P') for some energy $e > 0$. If V_j is analytic for $j \in \{d - d_{\text{anlyt}} + 1, \dots, d\}$, where $0 \leq d_{\text{anlyt}} \leq d - d_{\text{flat}}$, and $\tilde{\mu}_k = \tilde{\mu}_k(\mathcal{C}_k, \text{deg}_U^{(d_{\text{flat}}+1)}, \dots, \text{deg}_U^{(d)}, e) > 0$ for $k \in \{d_{\text{flat}} + 1, \dots, d - d_{\text{anlyt}}\}$ are small enough, then U is separable, that is, there exist $U_1, \dots, U_d \in C^2(\mathbb{T})$ such that*

$$U(x^1, \dots, x^d) = U_1(x^1) + \dots + U_d(x^d) \quad \text{for all } (x^1, \dots, x^d) \in \mathbb{T}^d.$$

This is irrespective of $\tilde{\mu}_j > 0$ for $j \in \{d - d_{\text{anlyt}} + 1, \dots, d\}$, but only for

$$(\mu_{d-d_{\text{anlyt}}+1}, \dots, \mu_d) \in [0, \tilde{\mu}_{d-d_{\text{anlyt}}+1}] \times \dots \times [0, \tilde{\mu}_d]$$

outside of an exceptional d_{anlyt} -dimensional null-set (depending on $(\mu_{d_{\text{flat}}+1}, \dots, \mu_{d-d_{\text{anlyt}}})$).

B. Appendix. Basic perturbation lemma

In this appendix, we state a basic perturbation lemma, which is instrumental in the continuity arguments required for the proofs of Lemmas 4.2 and 4.3. Its proof is given Appendix B of the [arXiv: 2210.02961](https://arxiv.org/abs/2210.02961) version of this article.

LEMMA B.1. *Let $V \in C^1(\mathbb{T})$ be a non-negative function with $\min V = 0$, $\mu \in [0, 1]$, and define the Hamiltonian function*

$$H_\mu(p, x) = \frac{p^2}{2} - \mu V(x) \tag{B.1}$$

on the cotangent bundle $T^*\mathbb{T}$. In the neighborhood of a fixed energy $E > 0$, we can find action-angle coordinates (I, θ) of equation (B.1) as

$$I = \pm \int_0^1 \sqrt{2(E + \mu V(x))} \, dx, \quad \theta = \pm \frac{\int_0^x (dx'/\sqrt{1 + \mu V(x')/E})}{\int_0^1 (dx'/\sqrt{1 + \mu V(x')/E})}. \tag{B.2}$$

Regarding $\theta = \theta(x)$ as a function on \mathbb{T} , we have $\theta \in C^1(\mathbb{T})$ and

$$\|\theta \mp x\|_{C^1} = \mathcal{O}\left(\frac{\mu \|V\|_{C^0}}{E}\right) \quad \text{as } \mu \rightarrow 0. \tag{B.3}$$

The same holds true if we regard $x = x(\theta)$ as a function on \mathbb{T} .

C. Appendix. Action-angle coordinates and analyticity

This appendix is concerned with analyticity properties of action-angle coordinates for one-dimensional Hamiltonian system

$$H_\mu(p, x) = \frac{p^2}{2} - \mu V(x) \tag{C.1}$$

being defined on the cotangent bundle $T^*\mathbb{T}$, where μ is a positive parameter and $V \geq 0$ an analytic function. Just as in Appendix B, in the neighborhood of a fixed energy $E > 0$, we can find action-angle coordinates (I, θ) of equation (C.1) as given in equation (B.2). From now on, we shall restrict to the first sign choice in equation (B.2).

In our proofs of the analyticity cases in Theorems 2.3 and 2.4, we shall exploit the fact that the function

$$\theta : (x, \mu) \mapsto \frac{\int_0^x (dx'/\sqrt{1 + \mu V(x')/E})}{\int_0^1 (dx'/\sqrt{1 + \mu V(x')/E})} \tag{C.2}$$

is analytic in both variables. (Note that the further implicit dependence on μ via $E = E(I)$ is also analytic.) Now, for every fixed $\mu > 0$, the function $x \mapsto \theta(x, \mu)$ is analytic and invertible, and we denote its analytic inverse by $\theta \mapsto x_\mu(\theta)$ (cf. Step (i) in the proofs of Theorems 2.3 and 2.4). Moreover, most importantly, also the function

$$(\theta, \mu) \mapsto x_\mu(\theta)$$

is analytic in μ , as shown in the following simple lemma applied to $f(z, w) \equiv \theta(x, \mu)$ in equation (C.2). Its elementary proof, based on Hartogs’s theorem, is given in Appendix C of the [arXiv: 2210.02961](https://arxiv.org/abs/2210.02961) version of this article.

LEMMA C.1. *Let $D_z, D_w \subset \mathbb{R}$ be open sets and*

$$f : D_z \times D_w \rightarrow \mathbb{R}, \quad (z, w) \mapsto f(z, w) \tag{C.3}$$

an analytic function. Moreover, assume that the one-variable restriction $f(\cdot, w) : D_z \rightarrow \mathbb{R}$ is invertible and satisfies $f(D_z, w) = D$ for every fixed $w \in D_w$ and some open $D \subset \mathbb{R}$, such that we can write its analytic inverse function as

$$f^{-1}(\cdot, w) : D \rightarrow D_z, \quad \zeta \mapsto f^{-1}(\zeta, w).$$

Then it holds that, with a slight abuse of notation, also

$$f^{-1} : D \times D_w \rightarrow D_z, \quad (\zeta, w) \mapsto f^{-1}(\zeta, w)$$

is an analytic function.

We note that although θ from equation (C.2) is always analytic in μ , the lower regularity in x for a general $V \in C^2(\mathbb{T})$ prevents the analyticity in μ to carry over to the inverse function.

D. Appendix. Weak KAM theory

In this appendix, we provide a brief overview on basic results of weak KAM theory and Aubry–Mather theory, which are relevant in the proofs of our main results. More details and background information can be found in extended version of this appendix in the [arXiv: 2210.02961](https://arxiv.org/abs/2210.02961) version of this article or lecture notes from Sorrentino [85], which build on seminal works from Mather [75–77], Aubry [9], Mañé [73], Fathi [44, 45], Siconolfi [46, 47], Bernard [12], and others [32, 36, 72].

D.1. *Aubry–Mather theory in one dimension.* In the following, we briefly discuss Aubry–Mather theory for the one-dimensional example of a mechanical Hamiltonian on $M = \mathbb{T}$. Note that the unperturbed Hamiltonian in equation (2.1) in the formulation of our main results is a sum of two such one-dimensional systems. Let $V \in C^2(\mathbb{T})$ be a non-negative Morse function with $\min_{x \in \mathbb{T}} V(x) = 0$, $\mu \in (0, 1]$, and consider the Hamiltonian

$$H : T^*\mathbb{T} \rightarrow \mathbb{R}, \quad (x, p) \mapsto \frac{p^2}{2} - \mu V(x), \tag{D.1}$$

whose corresponding Lagrangian can easily be obtained as $L(x, v) = v^2/2 + \mu V(x)$.

We first note that the (co)tangent bundle and the (co)homology group of \mathbb{T} are given by

$$T\mathbb{T} \cong T^*\mathbb{T} \cong \mathbb{T} \times \mathbb{R} \quad \text{and} \quad H_1(\mathbb{T}, \mathbb{R}) \cong H^1(\mathbb{T}, \mathbb{R}) \cong \mathbb{R},$$

respectively. Next, we find the *Mather set* $\widetilde{\mathcal{M}}_c$ and *Mather’s α -function* $\alpha(c)$ (the energy level of a Mather set) at cohomology $c \in \mathbb{R}$ to be given by

$$\widetilde{\mathcal{M}}_c = \begin{cases} \{V = 0\} \times \{0\} & \text{if } |c| \leq \sqrt{\mu}c(V), \\ \mathcal{P}_{E(|c|)}^{\text{sgn}(c)} & \text{if } |c| > \sqrt{\mu}c(V), \end{cases} \quad \alpha(c) = \begin{cases} 0 & \text{if } |c| \leq \sqrt{\mu}c(V), \\ E(|c|) & \text{if } |c| > \sqrt{\mu}c(V), \end{cases} \tag{D.2}$$

respectively, where $\text{sgn}(c)$ denotes the sign of c . We now explain the various notation used in equation (D.2). For energy $E > 0$, we denoted the two homotopically non-trivial periodic orbits contained in the energy level $\{H(x, p) = E\}$ by

$$\mathcal{P}_E^\pm := \{(x, p) : p = \pm\sqrt{2(E + \mu V(x))}, x \in \mathbb{T}\}.$$

The cohomology class of the closed 1-form $\eta_E^+ := \sqrt{2(E + \mu V(x))} dx$ corresponding to the orbit in \mathcal{P}_E^+ is given by $c^+(E) = [\eta_E^+] := \int_0^1 \sqrt{2(E + \mu V(x))} dx$. This function is

continuous, strictly increasing for $E > 0$ and we have

$$c^+(E) \longrightarrow \sqrt{\mu} \int_0^1 \sqrt{2V(x)} \, dx =: \sqrt{\mu} \, c(V) \quad \text{as } E \rightarrow 0. \tag{D.3}$$

Therefore, this defines an invertible function $c^+ : (0, \infty) \rightarrow (\sqrt{\mu} \, c(V), \infty)$, whose inverse we denote by $E(c)$.

Remark D.1. Using equation (D.2) for the two independent dimensions of equation (2.1), we obtain Proposition 2.1(a). More precisely, this follows after realizing that $\alpha(\mathbf{c}) = \alpha_1(c_1) + \alpha_2(c_2)$, where α_i is the α -function of the one-dimensional system with coordinates labeled by i , and taking $u_{\mathbf{c}} \in C^3(\mathbb{T}^2)$ with $|c_i| > \sqrt{\mu_i} c(V_i)$ according to

$$\nabla_x u_{\mathbf{c}}(x) = -\mathbf{c} \pm \left(\begin{array}{c} \sqrt{2(\alpha_1(c_1) + \mu_1 V_1(x^1))} \\ \sqrt{2(\alpha_2(c_2) + \mu_2 V_2(x^2))} \end{array} \right),$$

(recall $V_i \in C^2(\mathbb{T})$ is a non-negative Morse function and $\alpha_i(c_i) > 0$) such that the Hamilton–Jacobi equation

$$\alpha(\mathbf{c}) = H_0(x, \mathbf{c} + \nabla_x u_{\mathbf{c}}(x))$$

is satisfied. Moreover, in the case where U as in equation (2.3) is actually separable, one can employ the explicit forms for $c^+(E)$ as the inverse of the α -function and $\nabla u_{\mathbf{c}}$ to prove the validity of Assumption (P), simply by using the same expansions leading to the proof of Lemma B.1. This means that separable systems satisfy Assumption (P), which shows consistency with our main results.

D.2. Fathi’s weak KAM theory and perturbations. For concreteness, we specialize to $M = \mathbb{T}^2$, in which case $H^1(\mathbb{T}^2, \mathbb{R}) \cong T_x^* \mathbb{T}^2 \cong \mathbb{R}^2$ for every $x \in \mathbb{T}^2$, such that we can identify $\mathbf{c} \in H^1(\mathbb{T}^2, \mathbb{R})$ with a closed 1-form of cohomology class \mathbf{c} . The central object of investigation in Fathi’s weak KAM theory is the HJE

$$H(x, \mathbf{c} + \nabla_x u) = k, \quad k \in \mathbb{R}, \tag{D.4}$$

where H is a Tonelli Hamiltonian on $T^* \mathbb{T}^2$ with associated Tonelli Lagrangian L .

For classical solutions, that is, C^1 -functions $u : \mathbb{T}^2 \rightarrow \mathbb{R}$ solving equation (D.4), it is immediate to check that there is at most one value $k \in \mathbb{R}$, for which such a C^1 -solution may exist. In fact, this value agrees with Mather’s α -function mentioned above. The following proposition contains perturbative properties of weak KAM solutions u_ε and Mather’s α -function α_ε for systems of the form

$$H_\varepsilon(x, p) = H_0(x, p) + \varepsilon H_1(x, p).$$

PROPOSITION D.2. (Gomes [50]) *Let $H_0 : T^* \mathbb{T}^2 \rightarrow \mathbb{R}$ be an integrable Tonelli Hamiltonian and $u^{(0)}$ a (classical) C^1 -solution of the HJE $H_0(x, \mathbf{c} + \nabla_x u^{(0)}) = \alpha^{(0)}(\mathbf{c})$. Moreover, let $\nu^{(0)}$ denote the projection of a Mather measure with cohomology class \mathbf{c} . Suppose there*

exists a function $u^{(1)} \in C^1(\mathbb{T}^2)$ and a number $\alpha^{(1)}(\mathbf{c})$ such that

$$\alpha^{(1)}(\mathbf{c}) = \langle (\nabla_p H_0)(x, \mathbf{c} + \nabla_x u^{(0)}, \nabla_x u^{(1)}) + H_1(x, \mathbf{c} + \nabla_x u^{(0)}) \rangle \quad \text{for all } x \in \mathbb{T}^2. \quad (\text{D.5})$$

Then

$$\alpha^{(1)}(\mathbf{c}) = \int_{\mathbb{T}^2} H_1(x, \mathbf{c} + \nabla_x u^{(0)}) dv^{(0)} \quad \text{and} \quad \alpha_\varepsilon(\mathbf{c}) = \alpha^{(0)}(\mathbf{c}) + \varepsilon \alpha^{(1)}(\mathbf{c}) + \mathcal{O}_\varepsilon(\varepsilon^2). \quad (\text{D.6})$$

Remark D.3. By invoking Remark D.1, the above proposition provides a converse to equation (2.11) in Assumption (P). In fact, the transport-type equation in equation (D.5) for the unknown $u^{(1)}$ (with so far unspecified constant $\alpha^{(1)}(\mathbf{c})$) is exactly the first-order expansion obtained in equations (4.2), (4.8), and (4.14) in §4 and also fixes $\alpha^{(1)}(\mathbf{c})$ to be given by equation (D.6). Moreover, equation (D.5) coincides with the relation, which the correction term $u^{(1)}$ of an approximate solution $\tilde{u}_\varepsilon = u^{(0)} + \varepsilon u^{(1)}$ to the HJE

$$H_\varepsilon(x, \mathbf{c} + \nabla_x u_\varepsilon) = k$$

of order one has to satisfy (see [50]). The approximate solution $\tilde{u}_\varepsilon = u^{(0)} + \varepsilon u^{(1)}$ also coincides with the first-order truncation of the so-called *Lindstedt series* [8, 51], a not necessarily convergent perturbative expansion similar to those in KAM theory [5, 63, 80] or the Poincaré–Melnikov method [8, 55, 91]. Finally, it is interesting to note that if $H_1(x, p) = W(x)$ is independent of the p -variables, then $\alpha_\varepsilon(\mathbf{c})$ is a convex function of ε and thus almost everywhere twice differentiable—yielding the expansion in equation (D.6) at every such point.

REFERENCES

- [1] R. L. Adler, A. G. Konheim and M. H. McAndrew. Topological entropy. *Trans. Amer. Math. Soc.* **114** (1965), 309–319.
- [2] S. V. Agapov and D. N. Aleksandrov. Fourth-degree polynomial integrals of a natural mechanical system on a two-dimensional torus. *Mat. Z.* **93** (2013), 790–793; Engl. transl. *Math. Notes* **93** (2013), 780–783.
- [3] S. V. Agapov, M. Bialy and A. E. Mironov. Integrable magnetic geodesic flows on 2-torus: New examples via quasi-linear system of PDEs. *Comm. Math. Phys.* **351** (2017), 993–1007.
- [4] M.-C. Arnaud, J. E. Massetti and A. Sorrentino. On the persistence of periodic tori for symplectic twist maps and the rigidity of integrable twist maps. *Preprint*, 2023, [arXiv:2202.00313](https://arxiv.org/abs/2202.00313).
- [5] V. I. Arnold. Proof of a Theorem of A. N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian. *Uspekhi Mat. Nauk* **18** (1963), 13–40; Engl. transl. *Russ. Math. Surv.* **18** (1963), 9–36.
- [6] V. I. Arnold. Instability of dynamical systems with several degrees of freedom. *Sov. Math.* **5** (1964), 9–12.
- [7] V. I. Arnold. *Mathematical Methods of Classical Mechanics*, 2nd edn. Springer, New York, 1989.
- [8] V. I. Arnold, V. V. Kozlov and A. I. Neishtadt. *Mathematical Aspects of Classical and Celestial Mechanics*. Springer, Berlin, 2007.
- [9] S. Aubry and P. Y. Le Daeron. The discrete Frenkel–Kontorova model and its extensions: I. Exact results for the ground-states. *Phys. D* **8** (1983), 381–422.
- [10] A. Avila, J. De Simoi and V. Kaloshin. An integrable deformation of an ellipse of small eccentricity is an ellipse. *Ann. of Math. (2)* **184** (2016), 527–558.
- [11] I. K. Babenko and N. N. Nekhoroshev. On complex structures on two-dimensional tori admitting metrics with nontrivial quadratic integral. *Math. Notes* **58** (1995), 1129–1135.
- [12] P. Bernard. Existence of $C^{1,1}$ critical sub-solutions of the Hamilton–Jacobi equation on compact manifolds. *Ann. Sci. Éc. Norm. Supér. (4)* **40** (2007), 445–452.

- [13] M. Bialy. First integrals that are polynomial in the momenta for a mechanical system on the two-dimensional torus. *Funktsional. Anal. i Prilozhen.* **21** (1987), 64–65; Engl. transl. *Funct. Anal. Appl.* **21** (1987), 310–312.
- [14] M. Bialy. Convex billiards and a Theorem by E. Hopf. *Math. Z.* **124** (1993), 147–154.
- [15] M. Bialy. Integrable geodesic flows on surfaces. *Geom. Funct. Anal.* **20** (2010), 357–367.
- [16] M. Bialy and R. MacKay. Symplectic twist maps without conjugate points. *Israel J. Math.* **141** (2004), 235–247.
- [17] M. Bialy and A. E. Mironov. Rich quasi-linear system for integrable geodesic flows on 2-torus. *Discrete Contin. Dyn. Syst. Ser. A* **29** (2011), 81–90.
- [18] M. Bialy and A. E. Mironov. Cubic and quartic Integrals for geodesic flow on 2-torus via a system of the hydrodynamic type. *Nonlinearity* **24** (2011), 3541–3554.
- [19] M. Bialy and A. E. Mironov. Integrable geodesic flows on 2-torus: Formal solutions and variational principle. *J. Geom. Phys.* **87** (2015), 39–47.
- [20] M. Bialy and A. E. Mironov. Angular billiard and algebraic Birkhoff conjecture. *Adv. Math.* **313** (2017), 102–126.
- [21] M. Bialy and A. E. Mironov. The Birkhoff–Poritsky conjecture for centrally-symmetric billiard tables. *Ann. of Math. (2)* **196** (2022), 389–413.
- [22] G. D. Birkhoff. *Dynamical Systems (Colloquium Publications, 9)*. American Mathematical Society, Providence, RI, 1927.
- [23] G. D. Birkhoff. On the periodic motions of dynamical systems. *Acta Math.* **50** (1927), 359–379.
- [24] A. Bolsinov, V. S. Matveev, E. Miranda and S. Tabachnikov. Open problems, questions and challenges in finite-dimensional integrable systems. *Philos. Trans. Roy. Soc. A* **376** (2018), 20170430.
- [25] A. V. Bolsinov and A. T. Fomenko. Integrable geodesic flows on the sphere, generated by Goryachev–Chaplygin and Kowalewski systems in the dynamics of a rigid body. *Math. Notes* **56** (1994), 859–861.
- [26] A. V. Bolsinov and A. T. Fomenko. *Integrable Hamiltonian Systems. Geometry, Topology, Classification*. CRC Press, Boca Raton, FL, 2004.
- [27] A. V. Bolsinov, A. T. Fomenko and V. S. Matveev. Two-dimensional Riemannian metrics with integrable geodesic flows. Local and global geometry. *Sb. Math.* **189** (1998), 1441–1466.
- [28] A. V. Bolsinov, V. V. Kozlov and A. T. Fomenko. The Maupertuis principle and geodesic flows on the sphere arising from integrable cases in the dynamics of a rigid body. *Russ. Math. Surv.* **50** (1995), 473–501.
- [29] A. V. Bolsinov and I. A. Taimanov. Integrable geodesic flows with positive topological entropy. *Invent. Math.* **140** (2000), 639–650.
- [30] D. Burago and S. Ivanov. Riemannian tori without conjugate points are flat. *Geom. Funct. Anal.* **4** (1994), 259–269.
- [31] K. Burns and V. S. Matveev. Open problems and questions about geodesics. *Ergod. Th. & Dynam. Sys.* **41** (2021), 641–684.
- [32] M. J. D. Carneiro. On minimizing measures of the action of autonomous Lagrangians. *Nonlinearity* **8** (1995), 1077–1085.
- [33] S. A. Chaplygin. New special solution of the problem of rigid body motion in a fluid. *Selected Papers I*. OGIZ, Moscow–Leningrad, 1948 (in Russian).
- [34] T. Combet. Rational integrability of trigonometric polynomial potentials on the flat torus. *Regul. Chaotic Dyn.* **22**(4) (2017), 386–407.
- [35] L. Corsi and V. Kaloshin. Locally integrable non-Liouville analytic geodesic flows. *Preprint*, 2018, arXiv:1803.01222.
- [36] M. G. Crandall and P.-L. Lions. Viscosity solutions of Hamilton–Jacobi equations. *Trans. Amer. Math. Soc.* **277** (1983), 1–42.
- [37] G. Darboux. *Leçons sur la Théorie Générale des Surfaces et les Applications Géométriques du Calcul Infinitésimal*. Gauthier-Villars, Paris, 1891 (in French).
- [38] A. Delshams and R. Ramírez-Ros. Poincaré–Melnikov–Arnold method for analytic planar maps. *Nonlinearity* **9** (1996), 1–26.
- [39] N. V. Denisova and V. V. Kozlov. Polynomial integrals of geodesic flows on a two-dimensional torus. *Mat. Sb.* **185** (1994), 49–64; Engl. transl. *Russian Acad. Sci. Sb. Math.* **83** (1995), 469–481.
- [40] N. V. Denisova and V. V. Kozlov. Symmetries and the topology of dynamical systems with two degrees of freedom. *Sb. Math.* **80** (1995), 105–124.
- [41] N. V. Denisova and V. V. Kozlov. Polynomial integrals of reversible mechanical systems with a two-dimensional torus as the configuration space. *Sb. Math.* **191** (2000), 189–208.
- [42] N. V. Denisova, V. V. Kozlov and D. V. Treshev. Remarks on polynomial integrals of higher degree for reversible systems with a toral configuration space. *Izv. Ross. Akad. Nauk Ser. Mat.* **76** (2012), 907–921; Engl. transl. *Izv. Math.* **76** (2012), 57–72.

- [43] U. Dini. Sopra un problema che si presenta nella theoria generale delle rappresentazioni geografiche di una superficie su di un'altra. *Annali di Matematica* 3 (1869), 269–293 (in Italian).
- [44] A. Fathi. Théoreme KAM faible et théorie de Mather sur les systemes lagrangiens. *C. R. Acad. Sci. Ser. I – Math.* **324** (1997), 1043–1046 (in French).
- [45] A. Fathi. *Weak KAM theorem in Lagrangian dynamics*. 10th preliminary version, 2008; available from https://www.math.u-bordeaux.fr/~pthieull/Recherche/KamFaible/Publications/Fathi2008_01.pdf.
- [46] A. Fathi and A. Siconolfi. Existence of C^1 critical subsolutions of the Hamilton–Jacobi equation. *Invent. Math.* **155** (2004), 363–388.
- [47] A. Fathi and A. Siconolfi. PDE aspects of Aubry–Mather theory for quasiconvex Hamiltonians. *Calc. Var. Partial Differential Equations* **22** (2005), 185–228.
- [48] A. T. Fomenko, S. V. Matveev. Morse-type theory for integrable Hamiltonian systems with tame Integrals. *Mat. Zametki* **43** (1988), 382–386; Engl. transl. *Math. Notes* **43** (1988), 663–671.
- [49] A. Glutsyuk. On polynomially integrable Birkhoff billiards on surfaces of constant curvature. *J. Eur. Math. Soc. (JEMS)* **23** (2020), 995–1049.
- [50] D. A. Gomes. Regularity theory for Hamilton–Jacobi equations. *J. Differential Equations* **187** (2003), 359–374.
- [51] D. A. Gomes. Perturbation theory for viscosity solutions of Hamilton–Jacobi equations and stability of Aubry–Mather sets. *SIAM J. Math. Anal.* **35** (2003), 135–147.
- [52] J. B. Gomes, M. J. Dias Carneiro and R. O. Ruggiero. Hopf conjecture holds for analytic, k -basic Finsler tori without conjugate points. *Bull. Braz. Math. Soc. (N.S.)* **46** (2015), 621–644.
- [53] J. B. Gomes and R. O. Ruggiero. Weak integrability of Hamiltonians on the two torus and rigidity. *Nonlinearity* **26**(7) (2013), 2109–2120.
- [54] D. N. Goryachev. New cases of integrability of Euler’s dynamical equations. *Izv. Warsaw Univ.* **3** (1916), 1–13 (in Russian).
- [55] J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Springer, New York, 2013.
- [56] F. Hartogs. Zur Theorie der analytischen Funktionen mehrerer unabhängiger Veränderlichen, insbesondere über die Darstellung derselben durch Reihen, welche nach Potenzen einer Veränderlichen fortschreiten. *Math. Ann.* **62** (1906), 1–88.
- [57] K. Heil, A. Moroianu and U. Semmelmann. Killing tensors on tori. *J. Geom. Phys.* **117** (2017), 1–6.
- [58] E. Hopf. Closed surfaces without conjugate points. *Proc. Natl. Acad. Sci. USA* **34** (1948), 47–51.
- [59] G. Huang, V. Kaloshin and A. Sorrentino. Nearly circular domains which are integrable close to the boundary are ellipses. *Geom. Funct. Anal.* **28** (2018), 334–392.
- [60] V. Kaloshin and A. Sorrentino. On the local Birkhoff conjecture for convex billiards. *Ann. of Math. (2)* **188** (2018), 315–380.
- [61] V. Kaloshin and A. Sorrentino. On the integrability of Birkhoff billiards. *Philos. Trans. Roy. Soc. A* **376** (2018), 20170419.
- [62] K. Kiyohara. Compact Liouville surfaces. *J. Math. Soc. Japan* **43** (1991), 555–591.
- [63] A. N. Kolmogorov. On the conservation of conditionally periodic motions under small perturbation of the Hamiltonian. *Dokl. Akad. Nauk SSSR* **98** (1954), 527–530.
- [64] V. N. Kolokoltsov. Geodesic flows on two-dimensional manifolds with an additional first integral polynomial in velocities. *Izv. Akad. Nauk SSSR* **46** (1982), 291–306.
- [65] A. Korn. *Zwei Anwendungen der Methode der sukzessiven Annäherungen*. Schwarz Abhandlungen, Springer, Berlin, 1916 (in German).
- [66] I. Koval. Local strong Birkhoff conjecture and local spectral rigidity of almost every ellipse. *Preprint*, 2021, [arXiv:2111.12171](https://arxiv.org/abs/2111.12171).
- [67] S. Kovalevskaya. Sur le Problème de la Rotation d’un Corps Solide Autour d’un Point Fixe. *Acta Math.* **2** (1889), 177–232 (in French).
- [68] V. V. Kozlov. Topological obstacles to integrability of natural mechanical systems. *Dokl. Akad. Nauk SSSR* **249** (1979), 1299–1302.
- [69] V. V. Kozlov. Integrable and non-integrable Hamiltonian systems. *Sov. Sci. Rev. C Math. Phys.* **8** (1988), 1–81.
- [70] V. V. Kozlov and D. V. Treshev. On the integrability of Hamiltonian systems with toral position space. *Math. USSR-Sb.* **63** (1989), 121–139.
- [71] L. Lichtenstein. Zur Theorie der konformen Abbildung. *Bull. Int. Acad. Sci. Cracovie Cl. Sci. Math. Nat. Sér.A* (1916), 192–217 (in German).
- [72] P.-L. Lions, G. Papanicolaou and S. R. S. Varadhan. Homogenization of Hamilton–Jacobi equations, unpublished manuscript, 1986.
- [73] R. Mañé. Lagrangian flows: the dynamics of globally minimizing orbits. *Bull. Braz. Math. Soc. (N.S.)* **28** (1997), 141–153.

- [74] J. E. Massetti and A. Sorrentino. On the rigidity of integrable deformations of twist maps and Hamiltonian flows. *Preprint*, 2022. [arXiv:2011.10997](https://arxiv.org/abs/2011.10997) (later replaced by [4]).
- [75] J. N. Mather. Existence of quasi-periodic orbits for twist homeomorphisms of the annulus. *Topology* **21** (1982), 457–467.
- [76] J. N. Mather. Action minimizing invariant measures for positive definite Lagrangian systems. *Math. Z.* **207** (1991), 169–207.
- [77] J. N. Mather. Variational construction of connecting orbits. *Ann. Inst. Fourier (Grenoble)* **43** (1993), 1349–1386.
- [78] V. S. Matveev. Quadratically integrable geodesic flows on the torus and on the Klein bottle. *Regul. Chaotic Dyn.* **2** (1997), 96–102.
- [79] A. E. Mironov. Polynomial integrals of a mechanical system on a two-dimensional torus. *Izv. Ross. Akad. Nauk Ser. Mat.* **74** (2010), 805–817; Engl. transl. *Izv. Math.* **74** (2010), 145–156.
- [80] J. Moser. On invariant curves of area-preserving mappings of an annulus. *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II* (1962), 1–20.
- [81] T. Z. Nguyen, L. S. Polyakova and E. N. Selivanova. Topological classification of integrable geodesic flows on orientable two-dimensional Riemannian manifolds with additional integral depending linearly or quadratically on the momenta. *Funktional. Anal. i Prilozhen.* **27** (1993), 186–196; Engl. transl. *Funct. Anal. Appl.* **27** (1993), 42–56.
- [82] H. Poritsky. The billiard ball problem on a table with a convex boundary—An illustrative dynamical problem. *Ann. of Math. (2)* **51** (1950), 446–470.
- [83] V. A. Sharafutdinov. Killing tensor fields on the 2-torus. *Sib. Math. J.* **57** (2016), 155–173.
- [84] V. A. Sharafutdinov. Killing tensor fields of third rank on a two-dimensional Riemannian torus. *Preprint*, 2020, [arXiv:2011.09603](https://arxiv.org/abs/2011.09603).
- [85] A. Sorrentino. *Action-minimizing Methods in Hamiltonian Dynamics (Mathematical Notes, 50)*. Princeton University Press, Princeton, NJ, 2015.
- [86] Y. B. Suris. Integrable mappings of the standard type. *Funct. Anal. Appl.* **23** (1989), 74–76.
- [87] I. S. Taimanov. On first integrals of geodesic flows on a two-torus. *Proc. Steklov Inst. Math.* **295** (2016), 225–242.
- [88] D. Treshev. Billiard map and rigid rotation. *Phys. D* **255** (2013), 31–34.
- [89] D. Treshev. On a conjugacy problem in billiard dynamics. *Proc. Steklov Inst. Math.* **289** (2015), 291–299.
- [90] D. Treshev. A locally integrable multi-dimensional billiard system. *Discrete Contin. Dyn. Syst.* **37** (2017), 5271–5284.
- [91] D. Treshev and O. Zubelevich. *Introduction to the Perturbation Theory of Hamiltonian Systems*. Springer, Berlin, 2009.
- [92] V. V. Vedyushkina (Fokicheva) and A. T. Fomenko. Integrable geodesic flows on orientable two-dimensional surfaces and topological billiards. *Izv. Math.* **83** (2019), 1137–1173.
- [93] Q. Wang and K. Zhang. Gevrey regularity for the formally linearizable billiard of Treshev. *Preprint*, 2022, [arXiv:2211.03182](https://arxiv.org/abs/2211.03182).