Quasiperiodic Motions in Superquadratic Time-Periodic Potentials

Mark Levi

Department of Mathematical Sciences, Reusselaer Polytechnic Institute, Troy, NY 12180, USA

Abstract. It is shown that for a large class of potentials on the line with superquadratic growth at infinity and with the additional time-periodic dependence all possible motions under the influence of such potentials are bounded for all time and that most (in a precise sense) motions are in fact quasiperiodic. The class of potentials includes, as very particular examples, the exponential, polynomial and much more. This extends earlier results and gives an answer to a problem posed by Littlewood in the mid 1960's. Along the way machinery is developed for estimating the action-angle transformation directly in terms of the potential and also some apparently new identities involving singular integrals are derived.

1. Introduction and the Results

In the early 1960's Littlewood [LI] asked whether or not the solutions of the Duffing-type equations

$$\ddot{x} + g(x) = p(t), \quad \text{where} \quad p(t + 1) = p(t) \tag{1.0}$$

are bounded for all time, i.e. whether there are resonances that might cause the amplitude of the oscillations to increase without bound. In this paper we study the more general problem of describing the behavior of solutions of the system of the form

$$\ddot{x} + V_x(x, t) = 0, \tag{1.1}$$

which is a Hamiltonian system governing the motion of a particle on the line subject to a time-periodic force. We show that under appropriate growth assumptions on V for large x, the system is near-integrable for large amplitudes in the sense that "most" large amplitude solutions are quasiperiodic and all solutions are bounded for all time with no smallness assumptions on the time-dependence of $V(x, t)$. Intuitively, one might expect that if $V(x, t)$ is superquadratic in x, then the larger amplitude solutions oscillate faster giving rise to a twist in the Poincaré map of the $(x, \dot{x})$ plane and thus a hope of applying Moser's twist theorem to

prove boundedness. This intuition turns out to be too crude: Littlewood [LI] (see also [L2], [LO]) produced an oscillator $\ddot{x} + U'(x) = p(t)$ with U satisfying $\dfrac{U'(x)}{x} \to \infty$ as $x \to \infty$, $U \in C^\infty$, and with a periodic piecewise constant $p(t)$ possessing an unbounded solution. This was accomplished by modifying the quartic potential $U(x) = \frac{1}{4} x^4$ in such a way as to create a resonance for a particular solution while still preserving the superlinear growth of $U'(x)$ at infinity. Further details can be found in [LI, L2, LO].

The problem of boundedness of solutions for (1.1) can be separated into two parts: first, to gain the information on the period map (also called the Poincaré map or the stroboscopic map) $(x, \dot{x})_{t=0} \to (x, \dot{x})_{t=1}$, and second, to use this information to decide whether or not the solutions are bounded. Neither of these two problems is fully tractable. This is due to the implicit character of the relationship between the potential $V(x, t)$ and the Hamiltonian $H(\theta, I, t)$ expressed in the action-angle variables. In other word, Eq. (1.1) cannot be solved explicitly and a qualitative estimate has to be made. Such estimates are provided in this paper.

The second problem, of a more general nature, lies in determining whether or not a given map has escaping orbits. This fundamental classical problem was solved only relatively recently by KAM theory [M1, HE, R1, R2, M2, SZ]: the only known way to assure that all iterates of a map are bounded is to show that there are invariant circles arbitarily close to ∞ in the $(x, \dot{x})$ plane of the map. In fact, the invariant curves are necessary for the boundedness, as was shown by Birkhoff ([B1, B2]); in that sense there is no hope of proving boundedness without demonstrating at the same time the existence of an invariant curve. In fact, earlier attempts to prove boundedness [D1], [D2] without using the invariant tori contain a mistake.

The first boundedness result is due to Morris [MO], who showed that all solutions of $\ddot{x} + x^3 = p(t)$ are bounded, with $p(t) \in C^0$. Subsequently, this result was extended to a wider class of systems with polynomial potentials (of even degree) by Dieckerhoff and Zehnder [DZ1, DZ2]. The restriction on the constancy of the leading coefficient that was required in that latter work was removed by Laederich and the author in [LL], by using a different transformation to a normal form. It might seem surprising at a first glance that all solutions of, say $\ddot{x} + a(t)x^3 = 0$, with a smooth $a(t) = a(t + 1) > 0$ are bounded for all time-after all, the potential $V(x, t) = \frac{1}{4} a(t)x^4$ undergoes a large change in the course of one t-period and thus could (it seems) "pump up" the oscillations of $x(t)$ as in fact happens in the linear case $\ddot{x} + a(t)x = 0$ when there is parametric resonance, for instance, in the Mathieu equation [S]. In the nonlinear case just mentioned, the nonlinearity does not allow the resonance to persist. In contrast to this, Littlewood's counterexample is based on constructing a resonance for a nonlinear case, but this construction violates the monotone twist condition [AD, BE, HA, KA, M], see [L2]. We mention finally an elegant result by Norris [N] for Eq. (1.0) when the x- and t-dependence separate: if $g(x)/x < g'(x)$ (a superquadraticity condition), and if $g(x)$ is analytic in a strip around the x-axis in the complex plane, then the system is near-integrable for large energies; the same result holds for the subquadratic growth $g(x)/x > g'(x)$ as well. (For a geometrical discussion of this condition see the end of this section.)

It should be pointed out that the result of Morris and its generalization by Norris hold with no smoothness assumptions on $p(t)$. This is due to the fact that

the Poincaré map $(x, \dot{x})_{t=0} \rightarrow (x, \dot{x})_{x=1}$ is a small perturbation of the completely integrable Poincaré map of the autonomous equation $\ddot{x} + g(x) = 0$. By contrast to this, in the case of a more general potential $V(x, t)$, with a stronger t-dependence the Poincaré map mentioned above is not close to an integrable one, and higher smoothness in t is necessary to construct the change of variables in which the system is (say) $C^{(4)}$-close to an integrable one. We point out that although the smoothness of $p(t)$ does not affect the *smoothness* of the Poincaré map, it does affect its *proximity* to an integrable map. This remark is illustrated by the counterexample of Littlewood, in which the construction of a non-near-integrable Poincaré map is crucially dependent on the discontinuity of $p(t)$. As soon as we make $p(t)$ smooth, Littlewood's construction (of a "bad" map) fails, as can be shown by more careful estimates; this suggests that **smoother p (for a fixed g) give rise to "nicer" (i.e., closer to integrable) Poincaré maps.**

We conclude this brief historical discussion by mentioning a related problem when $V(x, t)$ in (1.1) is periodic in x as well as in t; the nonlinear pendulum with periodic (zero-average) forcing is a primary example. In this case it has been shown that, as long as $V(x, t)$ is sufficiently smooth, there exist invariant tori $y = f(x, t)$, where $x \bmod 1$, $t \bmod 1$, which are preserved by the flow $\dot{x} = y$, $\dot{t} = 1$, $\dot{y} = -V(x, t)$ in the phase space $\{(x, t, y)\} \equiv T^2 \times \mathbb{R}$, [M3, CZ, L1]. In particular, for any solution there exists a constant $C > 0$ such that for all time $|\dot{x}| < C$.

The main result of this paper is the following

Theorem 1. *Assume that the potential function $V(x, t) = V(x, t+1)$ tends to ∞ as $|x| \rightarrow \infty$ and that it satisfies conditions (1.2), (1.3), and (1.4) below.*

Then Eq. (1.1) is near-integrable for large energies, more precisely for any $0 < \omega < 1$ satisfying the Diophantine conditions

$$\left| \omega - \frac{p}{q} \right| \geq \frac{1}{2} |q|^{-\frac{5}{2}} \quad \text{for all} \quad 0 \neq q, \ p \in Z,$$

the Poincaré's map $P : (x, \dot{x})_{t=0} \rightarrow (x, \dot{x})_{z=1}$ of Eq. (1.1) possesses countably many invariant circles with rotation number ω and these circles cluster at infinity in the $(x, \dot{x})$ plane. These circles are the intersections of the invariant tori in the $(x, \dot{x}, t \bmod 1)$-space with the $\{t = 0\}$-plane, Fig. 1.1; each such torus carries a quasiperiodic solution with basic frequencies 1 and $k + \omega$. All integers $k \geq k_0(\omega)$ are represented by an invariant torus. In particular,
1) *All solutions are bounded for all time:* $\sup_{\mathbb{R}} (|x| + |\dot{x}|) < \infty$.

2) *Most solutions with large amplitude are quasiperiodic, i.e. most initial conditions (in the sense of Lebesque measure) with large $|x(0)| + |\dot{x}(0)|$ give rise to quasiperiodic solutions: $x(t) = f(t, (\omega + k)t)$, where f is a function on a 2-torus.*
3) *Any sufficiently large number is a rotation number for some solution, i.e. there exists ϱ_0 such that for any $\varrho > \varrho_0$ there exists a solution $x_\varrho(t)$ with that rotation number[1]. Furthermore, for any rational $p/q > \varrho_0$ there exists a periodic solution with that rotation number; the corresponding periodic point of the Poincaré map is a Birkhoff periodic point.*

---

[1] The standard definition of the (forward) rotation number is $\varrho = 1/2\pi \lim_{t \to \infty} \arg(x, \dot{x})$. Another interpretation of ϱ is one half the average number of zeros of the solution during the forcing period: $\varrho = \lim_{T \to \infty} 1/2T$ (# of zeros of $x(t)$ on the interval $0 < t < T$)
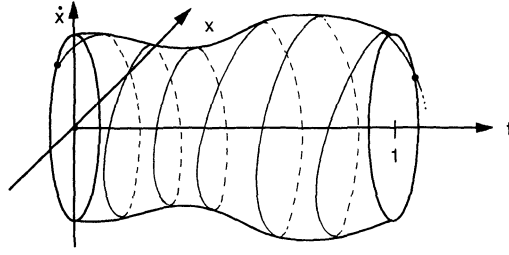
Fig. 1.1. Invariant tori in the extended phase space of Eq. (1.1)

The last statement 3) is a direct consequence of the Aubry-Mather theory; one can say much more by just applying the results of that theory [AD, BE, BH, HA, KA, MA1, MA2] to the Poincaré map of the system, using the information gained here.

Conditions on the potential $V(x, t)$.

For some constants c, c_1, $a > 0$, $0 < \mu < \dfrac{1}{100}\left(\dfrac{1}{2} - a\right)$ and all x and t we have

$$\left|\frac{\partial^{k+\tau}}{\partial^k x \partial^\tau t} V\right| \le c|x|^{-k} V^{1+\mu}, \qquad k + \tau \le 6, \tag{1.2}$$

$$-\frac{1}{2} + c_1 \le W_x \le a < \frac{1}{2}, \quad \text{where} \quad W = V/V_x, \tag{1.3}$$

and

$$|\partial_x^k \partial_t^\tau U|, \quad |\partial_x^k \partial_t^\tau W| \le c|x|^{1-k}, \quad k + \tau \le 5, \quad \text{where} \quad U = V_t/V_x. \tag{1.4}$$

Remark 1. Condition (1.4) can be weakened to read $|\partial^{k,\tau} U|$, $|\partial^{k,\tau} W| < c|x|^{1-k} V^\mu$; furthermore, the above assumptions on V can be weakened to require that these hold for $|x| \ge A$, for any fixed A. These stronger statements do not pose principal new difficulties.

Remark 2. Even a very special case of $\mu = 0$ includes, together with the polynomial, some new examples, e.g. the exponential potential $V(x, t) = (2 + \cos t) \cos hx$. (Condition (1.3) holds after subtracting $2 + \cos t$ from V – this does not affect Eq. (1.1).) For $\mu > 0$ the class widens drastically to allow some oscillation in x. A very particular example is $V(x, t) = p(t)(x + \cos\sqrt{x})^{2n}$, where n is (specifiably) large and $p(t) \in C^{(6)}$ is periodic and positive.

An Outline of the Proof of Theorem 1. There are two parts: first, the formal reduction to normal form and second, the estimates. The formal reduction (Fig. 1.2) consists of the following sequence of transformations; the quadruples list the position, the momentum, the time in that order with the Hamiltonian in the last place:

$$\left(x, \dot{x}, t; H_0 = \frac{\dot{x}^2}{2} + V\right) \xrightarrow{A} (\theta, I, t; H = H_0 + S_t) \xrightarrow{B} (t, H, \theta; I) \xrightarrow{C} (\tau, h, \theta; J),$$

where A is the standard action-angle transformation of $(x, \dot{x})$ into (θ, I) with t as the parameter (hence the extra term S_t in the Hamiltonian), B is the change
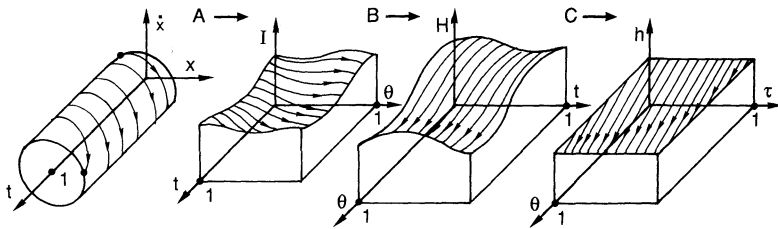
Fig. 1.2. The transformation to normal form and invariant tori

into time and energy as the new position and momentum, with the angle θ now playing the role of new time, and C is another action-angle transformation to make the leading term of the Hamiltonian independent of the position variable.

The estimate of the composition $C \circ B$ of the maps in the above diagram and a subsequent application of Moser's twist theorem will result in Theorem 5.1, which we restate here for convenience:

Theorem 2. *If the Hamiltonian function $H(\theta, I, t) = H_0(I, t) + H_1(\theta, I, t)$ is periodic in θ and t and satisfies for some $\beta > 0$ the growth estimates*

$$|\partial_t^l \partial_I^k H_1(\theta, I, t)| \leq cI^{-k-\beta} H_0(I, t), \quad l + k \leq 5, \tag{5.1}$$

and if the inverse function $I_0(H) = I_0(t, H)$ of $H_0(I, t)$ in H satisfies

$$|\partial_t^j \partial_H^k I_0(t, H)| \leq cH^{-k} I_0(t, H), \quad j + k \leq 5, \tag{5.2}$$

$$cH^{-k} I_0(t, H) \overset{k \leq 2}{\leq} |\partial_H^k I_0(t, H)|, \tag{5.3}$$

then the Hamiltonian system

$$\dot{\theta} = H_I, \quad \dot{I} = H_\theta$$

possesses invariant tori $I = f(\theta, t)$ carrying a quasiperiodic flow, in any region $I > M$, for an arbitrarily large M, Fig. 1.2.

This theorem is actually a combination of Theorem 5.1 and the results of Sect. 4.

The following three theorems estimate the action-angle change A in the diagram; they are restatements of Theorems 6.1–6.3:

Theorem 3. *If the potential $V(x, t)$ satisfies the superquadraticity condition (1.3) together with the bounds (1.2) and (1.4), then there exists $\beta = \beta(\mu, a)$ such that the action-angle Hamiltonian $H_0 + H_1$ satisfies the bound (5.1).*

Theorem 4. *If the potential V satisfies the bounds (1.4), then $H_0(I, t)$ satisfies (5.2), where I_0 is the inverse function of H_0 in I.*

Theorem 5. *If V satisfies (1.2), then $H_0(I, t)$ satisfies the twist conditions (5.3).*

Potentials with Blow-Up and Pulsating Billiards. One can apply the same methods to the cases when the potential blows up at one or two (end)points. For instance, one can show that the potentials

$$V_1(x, t) = \frac{1}{x^2 - a^2(t)} \quad \text{or} \quad V_2(x, t) = \frac{1}{x - b(t)} + a(t)x^4,$$
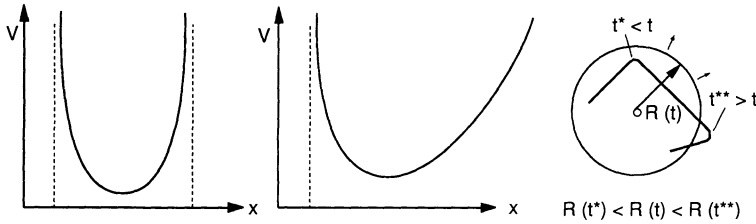
Fig. 1.3. Potentials with finite or semifinite domains and pulsating "billiards"

where $a > 0$, and $a, b \in C^{(5)}$, give rise to near-integrable motion for large energies. The potential V_1 can be viewed as a "softened" potential of the Fermi-Ulam's "ping-pong." The latter problem deals with a particle bouncing between two rigid periodically moving walls. If the motion of the walls is sufficiently smooth then the velocity of any motion is bounded for all time. This was shown by Moser (unpublished lecture notes) and independently by Douady [DO] and Laederich and the author [LL].

This kind of a potential with a blowup at one end arises in the motion of a particle $z \in \mathbb{R}^2$ governed by $\ddot{z} + \mathrm{grad}_z V = 0$, where V has an $SO(1)$ symmetry: $V(z, t) = W(|z|, t)$, with $V(z, t) = V(z, t + 1)$ as usual. The angular momentum $M = z \wedge \dot{z} = Im \, \bar{z}\dot{z}$ is a conserved quantity: $\dot{M} = \dot{z} \wedge \dot{z} + z \wedge \ddot{z} = -z \wedge V_z = 0$, the latter because $V_z \| z$ by symmetry. The polar coordinates $r, \theta : z = re^{i\theta}$ satisfy $r^2\dot{\theta} = M$ and $\ddot{r} - r^2\dot{\theta}^2 + W_r(r, t) = 0$; substituting the expression for $\dot{\theta} = Mr^{-2}$ we obtain the equation for r valid for all motions with the fixed value of M:

$$\ddot{r} + U_r(r, t, M) = 0,$$

where $U = \dfrac{M^2}{2r^2} + W(r, t)$.

One can show that if the effective, or the reduced, potential $U(r, t)$ satisfies the conditions similar to (1.2)–(1.4) then all motions are bounded.

A particular case of this statement gives a stability statement about the "pulsating soft billiards": let $V(z, t) = (z/R(t))^{2N}$, where $R(t)$ is periodic and positive. For N large, the potential is near zero inside the circle away from its boundary and it rises very steeply outside the circle, making this problem the elastic analog of the usual billiard whose potential is zero inside the circle and infinity outside. One can show using similar methods that if $R(t) \in C^{(5)}$, then the energy of any motion will stay bounded for all time, and hence each solution $r(t)$ has an upper and a *positive* lower bound for all time. The same then is true for the angular velocity $\dot{\theta} = \dfrac{M}{r^2}$. An explicit formulation of conditions for near-integrability for the potentials which blow up at finite x will be given elsewhere.

Remark 3. It was stated in [D1] and [D2] that every solution of $\ddot{x} + g(x) = p(t)$ is bounded for all time, provided that $\dfrac{g(x)}{x} \to \infty$ and that p is periodic and continuous. The statement seems to be false without further assumptions. The argument does not use the continuity of $p(t)$ and thus should apply to Littlewood's counterexample, which, however, exhibits unbounded solutions. (The original paper of Littlewood contained some minor errors which were found and corrected

by Long [LO]. A much simplified version of Littlewood's counterexample with an optimal estimate on the growth of $g(x)$ can be found in [L2].)

Other Results: Identities and the Growth Rate for the Period $T(H)$. We give a brief sample of some of the simplest results from the appendix. These auxiliary facts could be of some independent interest. The main results from the appendix are contained in A2 and A4; we do not discuss these here and touch only on a part of Appendix A3. The period of the solution of an autonomous system $\ddot{x} + V'(x) = 0$ with the energy H is given by $\int dt = \int (ds/v)$, or by the singular integral [A1]

$$T(H) = \sqrt{2} \int\limits_{x_-(H)}^{x_+(H)} \frac{dx}{\sqrt{(H - V(x))}},$$

where the limits $x_\pm(H)$ of integration are the points where the denominator vanishes: $V(x_\pm) = H$, and $V(x) < H$ for $x_- < x < x_+$. To be specific, let us also assume that $V(0) = V'(0) = 0$ and that $V'(x) \neq 0$ for all $x \neq 0$. Then one has the following formula for $T'(H)$:

$$T'(H) = \sqrt{2} \frac{1}{2H} \int\limits_{x_-}^{x_+} \left(1 - 2 \frac{V V''}{(V')^2}\right) \frac{dx}{\sqrt{(H - V(x))}}. \tag{1.5}$$

Proof of this is given in Appendix A3.

The relationship between the potential and the period as a function of the energy has been studied starting with Abel [AB]; of the more recent papers we mention [K, M4, SW]. The latter paper contains differentiation formulas for the period as a function of the energy (different from our formulas).

A Remark on the Twist Condition. This is essentially a repetition of an observation made in [L2], where it was pointed out that the period of solutions of $\ddot{x} + V'(x) = 0$ as a function of the energy is a monotone decreasing function if $V''(x) > V'(x)/x$ (the same condition appears in the result of Norris). We note first that this condition follows from the above condition $W_x \equiv (V/V_x)_x < 1/2 - a$, $a > 0$, in the case when $V(0) = V'(0) = 0$ (there is no loss of generality in this assumption). The latter condition on W in turn guarantees the monotonicity $T'(H) < 0$, as follows from Eq. (1.5). We give here a geometrical proof of the fact that the condition $V'' > V'/x$ implies that the period is a decreasing function of the energy. Let $z(t) = (x(t), \dot{x}(t))$ be a solution of $\ddot{x} + V'(x) = 0$ in the phase plane, and let $\zeta = (\xi(t), \eta(t))$ be a solution of the system linearized around $z(t)$. Assume that each solution vector $z(t)$ rotates clockwise in the phase plane. The criterion is based on the following geometrical observation, Fig. 1.4.

Take any solution $z(t)$ and let $\zeta(0)$ be parallel to $z(0)$. If $\zeta(t)$ turns clockwise faster than $z(t)$ at $t = 0$, then the period $T(H)$ of the oscillations is a decreasing function of the energy (or the amplitude, or the area enclosed by the curve). In other words, then the Poincaré map possesses a monotone twist (with radial lines as the reference foliation). Expressing the above idea analytically, we rewrite the condition on the angular velocities of z and ζ in the form

$$\frac{d}{dt}\left(\frac{y}{x}\right) > \frac{d}{dt}\left(\frac{\eta}{\xi}\right), \quad \text{whenever} \quad \frac{y}{x} = \frac{\eta}{\xi}.$$
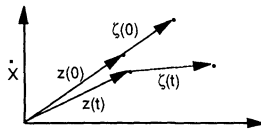
Fig. 1.4. A sufficient condition for monotone twist: any solution vector turns slower than the collinear to it linearized solution vector

Using the governing equations $\dot{x} = y$, $\dot{y} = -V'(x)$, $\dot{\xi} = \eta$, $\dot{\eta} = -V''(x)\xi$, the above criterion reduces to

$$\frac{V'}{x} < V''.$$

Norris [N] gave an alternative proof of this condition using the Sturm comparison argument.

The plan of the paper is as follows.
   Section 2. Reduction to the action-angle variables (step A, Fig. 1.2).
   Section 3. Reduction to the normal form at infinity (steps B and C).
   Section 4. Application of KAM assuming the estimates on normal form.
   Section 5. A transformation lemma: Estimating the normal form in terms of the action-angle Hamiltonian.
   Section 6. Estimating the action-angle Hamiltonian in terms of $V(x, t)$.
   Appendix:
   A1. An inverse function lemma.
   A2. Estimates on the operators $\mathscr{L}$, $\mathscr{M}$, L, M.
   A3. Algebra of the action-angle maps: Differentiation formulas and identities.
   A4. Estimates on the action-angle map.
   By far the main difficulty lies in the estimates on H via V carried out in Sect. 6.

2. Reduction to the Action-Angle Variables

In this section we carry out the standard reduction to the action-angle variables. We refer to [A1] for further details.

2.1. A geometric description. Let us denote the Hamiltonian for Eq. (1.1) by $\mathscr{H}(x, y, t) = \dfrac{y^2}{2} + V(x, t)$. Fixing t as a parameter, we let $I = I(x, y, t)$ denote the area enclosed by that level curve of $\mathscr{H}$ which passes through (x, y), Fig. 2.1. We assume that the level curve in question is (topologically) a circle, at least for I large enough (this is guaranteed if V is a monotone function for large $|x|$, increasing for $x > 0$ and decreasing for $x < 0$). The geometrical meaning of the definition of the action-angle variables is summarized in Fig. 2.1.

   To define the symplectic angle $\theta = \theta(x, y, t)$, we look at the system $\dfrac{d^2x}{ds^2} + V(x, t) = 0$ (with s playing the role of time and $t = $ const), consider that periodic solution which passes through (x, y), Fig. 2.1, and let $\theta = \theta(x, y, t)$ be the ratio of
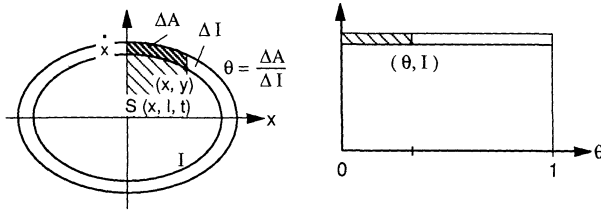
Fig. 2.1. The definition of the action-angle variables. To each point (x, y) we assign the action I given by the area enclosed by the level curve through (x, y) and the angle θ which is given by the proportion of the area of an infinitesimal ring enclosed between the y-axis and (x, y)

the time it takes the solution $\left(x, \dfrac{dx}{ds}\right)$ to travel between the y-axis and the given point (x, y) to the period $T = T(x, y, t)$ of this solution. The transformation $(x, y) \rightarrow (\theta, I)$ is area-preserving, as follows from the definition of I and θ and from the zero-divergence property of the Hamiltonian flow $x'' + V_x(x, t) = 0$. We note that θ is defined mod 1.

2.2. The formal definition. To construct the map $(x, y) \rightarrow (\theta, I)$ more formally, we let $H_0(I, t)$ be the value of $\mathscr{H} = \dfrac{y^2}{2} + V(x, t)$ on that level curve which encloses area I in the (x, y) plane, i.e. we define $H_0(I, t)$ (implicitly) by

$$\oint_{\mathscr{H}(x,y,t)=H_0(I,t)} y\,dx = I, \tag{2.1}$$

with time t as a parameter. We define now the generating function $S(x, I, t)$ as the area (shaded in Fig. 2.1):

$$S(x, I, t) = \int_C y\,dx, \tag{2.2}$$

where C is the part of the level curve $\mathscr{H}(x, y, t) = H_0(I, t)$ connecting the y-axis with the point (x, y), oriented clockwise. This defines S up to an integer multiple of $I = \oint y\,dx$ since C is defined up to an integer number of full trips around the level curve. We define the map $(\theta, I, t) \rightarrow (x, y, t)$ via

$$S_x(x, I, t) = y, \qquad S_I(x, I, t) = \theta; \tag{2.3}$$

it is well-known to be symplectic:

$$dx \wedge dy = dx \wedge (S_{xx}dx + S_{xI}dI) = S_{xI}dx \wedge dI;$$
$$d\theta \wedge dI = (S_{Ix}dx + S_{II}dI) \wedge dI = S_{Ix}dx \wedge dI.$$

The last definition (2.3) is equivalent to the geometric definition above. Equation (1.1) in the new variables (θ, I, t) retains its Hamiltonian character with the new Hamiltonian

$$H(\theta, I, t) = H_0(I, t) + H_1(\theta, I, t), \quad \text{where} \quad H_1 = S_t(x, I, t), \tag{2.4}$$

with $x = x(\theta, I, t)$ defined implicitly by (2.3).

We complete this section with the formulas that will be used later in the estimates.

2.3. Expressions for $H_0(I, t)$ and $H_1(\theta, I, t)$. Rewriting Eq. (2.1) we obtain the *implicit definition of $H_0(I, t)$*:

$$I = 2\sqrt{2} \int\limits_{x_-}^{x_+} \sqrt{H_0(I, t) - V(\xi, t)} d\xi, \qquad (2.5)$$

where $x_-(I, t) < x_+(I, t)$ are given by $V(x_\pm, t) = H_0(I, t)$, Fig. 2.1. Restating this slightly, $H_0(I, t)$ is defined as the inverse function (in I) of

$$I(H, t) = 2\sqrt{2} \int\limits_{x_-}^{x_+} \sqrt{H - V(\xi, t)} d\xi, \qquad (2.6)$$

where $x_-(H, t) < x_+(H, t)$ are given by $V(x_\pm, t) = H$.

To write down H_1 more explicitly, we write the area S as

$$S(x, I, t) = \sqrt{2} \int\limits_{0}^{x} \sqrt{H_0(I, t) - V(\xi, t)} d\xi,$$

so that

$$H_1(\theta, I, t) = \frac{\sqrt{2}}{2} \int\limits_{0}^{x(\theta, I, t)} \frac{\partial_t H_0(I, t) - \partial_t V(\xi, t)}{\sqrt{H_0(I, t) - V(\xi, t)}} d\xi, \qquad (2.7)$$

where $x(\theta, I, t)$ is defined implicitly by $S_I(x, I, t) = \theta$, (cf. 2.3) i.e. by

$$\frac{\sqrt{2}}{2} \partial_I H_0(I, t) \int\limits_{0}^{x} \frac{d\xi}{\sqrt{H_0(I, t) - V(\xi, t)}} = \theta. \qquad (2.8)$$

This formula has a simple dynamical interpretation: $\partial_I H_0$ is the inverse of the period of the solution emclosing area I (t is fixed as a parameter so that the system is autonomous): $\partial_I H_0 = T^{-1}(I, t)$. The rest of the product in the left-hand side of (2.8) is the time $\int \frac{ds}{v}$ it takes to go from $x = 0$ to x; Eq. (2.8) therefore coincides with our first definition of θ.

3. Reduction to the Normal Form at Infinity by Averaging

The Hamiltonian (2.4) is still far from time-independent, even in its leading term, and at this stage it is not clear that it is a small perturbation of an integrable system. In this section we choose new variables that "soak up" the unpleasant time-dependence in the leading term. (The meaning of the word "leading" is formal so far as no estimates have been made as yet.) Following Poincaré, Arnold [A] or Moser [M3], we change the variables by choosing H, t as the new momentum and position, while giving θ the role of the new time. The precise definition of this transformation $(\theta, I, t) \to (q, p, T)$ is as follows: $q = t$, $p = H(\theta, I, t)$ and $T = \theta$. Rather than burdening the reader's memory with new symbols q, p, T, we reserve the old notation (t, H, θ) for the new variables with position, momentum and time occupying the first, second and third place in the triple. The flow in this (t, H, θ)-space is Hamiltonian where the Hamiltonian $I = I(t, H, \theta)$ in the inverse function[2] of $H(\theta, I, t)$: indeed, the canonical 1-forms

---

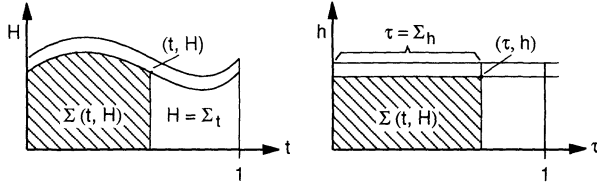[2] with t, θ playing the role of parameters

Fig. 3.1 The Hamiltonian averaging

coincide up to a sign [A]:

$$I d\theta - H(\theta, I, t)dt = -(Hdt - I(t, H, \theta)d\theta).$$

Along with $I(t, H, \theta)$ we define $I_0(t, H)$ as the inverse in H of $H_0(I, t)$ (t is still a parameter); we write

$$I(t, H, \theta) = I_0(t, H) + I_1(t, H, \theta), \tag{3.1}$$

thus defining I_1. Now that t has become a new position variable, it can be eliminated from the leading term in (3.1) by choosing action-angle variables in the t, H-plane as was done by Arnold [A].

We are looking for a function $\Sigma(t, h)$ generating a map $(t, H) \mapsto (\tau, h)$ via

$$\Sigma_t(t, h) = H, \qquad \Sigma_h(t, h) = \tau \tag{3.2}$$

such that the Hamiltonian of the transformed system would be τ-independent in its leading part, that is,

$$I_0(t, \Sigma_t(t, h)) \equiv J_0(h) \tag{3.3}$$

should depend on h alone. As an additional condition, we want the map (3.2) to preserve periodicity: if $(t, H) \mapsto (I, h)$, then $(t + 1, H) \mapsto (\tau + 1, h)$. These two conditions define $J_0(h)$ (up to a constant); specifically, (3.3) gives $\Sigma_t = H_0(H_0(h), t)$ and thus Σ can be taken as

$$\Sigma(t, h) = \int_0^t H_0(J_0(h), \xi)d\xi. \tag{3.4}$$

The periodicity condition translates into

$$\Sigma_h(1, h) - \Sigma_h(0, h) = \frac{\partial}{\partial h} \int_0^1 H_0(J_0(h), \xi)d\xi = \frac{\partial}{\partial J} \int_0^1 H_0(J, \xi)d\xi|_{J=J_0(h)} \cdot J_0'$$

$$\equiv \bar{H}_0'(J)|_{J=J_0(h)} \cdot J_0'(h) = 1,$$

which holds if we choose $J_0(h)$ as the inverse function of the averaged Hamiltonian

$$\bar{H}_0(J) = \int_0^1 H_0(J, \xi)d\xi.$$

Having thus specified the transformation (3.2), we apply it to the flow defined by $I(t, H, \theta)$; the new flow will be Hamiltonian, and the new Hamiltonian

$J(\tau, h, \theta)$ is obtained simply by substituting (3.2) in (3.1), since the change (3.2) is θ-independent:

$$J(\tau, h, \theta) = J_0(h) + J_1(I, h, \theta),\qquad(3.5)$$

where $J_1(t, h, \theta) = I_1(t, H, \theta)$.

This completes the reduction; so far all this is purely formal with no estimates. In the next Sect. 4 we assume some estimates an J and apply KAM theory to prove boundedness of solutions of (1.1). These estimates are proved in Sects. 5 and 6.

4. Application of KAM Theory Assuming the Estimates on Normal Form

We state the key assumptions on the transformed Hamiltonian $J_0 + J_1$ in Sect. 4.1 and we prove the existence of invariant tori using Moser's twist theorem in Sect. 4.2. The main job of proving these assumptions is done in Sects. 5 and 6 below.

4.1. Assume that $J_0(h)$ and $J_1(\tau, h, \theta)$ satisfy the following

Hypothesis on J_0, J_1 – There exist constants $a > 0$, $b > b' > 0$, $\beta > 0$, $C > c > 0$ such that for all h large enough we have for J_0 and J_1 (actually, (4.1) and (4.2) follow from (4.3), but we list them for future use):

$$J_0(h) \geq ch^a,\qquad(4.1)$$

$$J_0'(h) \geq ch^{-b},\qquad(4.2)$$

$$ch^{-k}J_0(h) \overset{k=1,2}{\leq} J_0^{(k)}(h)^{k\leq 5} \leq Ch^{-k}J_0(h),\qquad(4.3)$$

and the main estimate for J_1, J_0:

$$|\partial_\tau^l \partial_h^k J_1(\tau, h, \theta)| \leq ch^{-k}J_0(h)^{1-\beta}\qquad l+k \leq 5.\qquad(4.4)$$

These estimates are proven in Sects. 5 and 6 under the assumptions on $V(x, t)$ stated in Theorem 1. Here we use the above estimates (4.1)–(4.4) to apply Moser's twist theorem.

4.2. Application of KAM. The equations corresponding to the Hamiltonian $J = J_0 + J_1$ are

$$\begin{cases} \dfrac{d\tau}{d\theta} = J_0'(h) + \partial_h J_1(\tau, h, \theta) \\ \dfrac{dh}{d\theta} = -\partial_\tau J_1(\tau, h, \theta), \end{cases}$$

where ∂_h denotes $\dfrac{\partial}{\partial h}$. To bring h down from the neighborhood of ∞ we rescale it by

$$J_0'(h) = \varepsilon r,\qquad 1 \leq r \leq 2.\qquad(4.5)$$

Small ε correspond to large h since $J_0'(h) \to 0$ as $h \to \infty$ by (4.2).

We now have the advantage of the fixed range for the new variable[3] r.

---

[3] The resulting map $(I, h) \mapsto (\tau, r)$ reverses the orientation. We could have chosen $J_0'(h) = \varepsilon r^{-1}$ to obtain an orientation-preserving map

The transformed equations are

$$\begin{cases} \dfrac{d\tau}{d\theta} = \varepsilon r + R_1(\tau, \varepsilon r, \theta) \\[2mm] \dfrac{dr}{d\theta} = R_2(\tau, \varepsilon r, \theta) \end{cases} \tag{4.6}$$

with $R_1 = \partial_h J_1(\tau, h(\varepsilon r), \theta)$ and $R_2 = -\varepsilon \dfrac{J_0''(h(\varepsilon r))}{J_0'(h(\varepsilon r))^2} \partial_\tau J_1(\tau, h(\varepsilon r), \theta)$.

We claim that for some $v > 0$ and $\varepsilon_0 > 0$ the estimates

$$\left| \frac{\partial^{k+l}}{\partial\tau^k \partial r^l} R_i \right| \le \varepsilon^{1+v} \tag{4.7}$$

hold for all $0 \le k + l \le 4$, $i = 1, 2$, $1 \le r \le 2$, $0 < \varepsilon < \varepsilon_0$. Before proving this claim we use it to apply KAM. Integrating Eq. (4.6) from $\theta = 0$ to $\theta = 1$ we obtain the Poincaré map

$$\begin{aligned} \tau_1 &= \tau_0 + \varepsilon r_0 + Q_1(\tau, r, \varepsilon), \\ r_1 &= r_0 + Q_2(\tau, r, \varepsilon) \end{aligned} \tag{4.8}$$

with $Q_{1,2}$ still satisfying the same estimates (4.7) as $R_{1,2}$.

The map (4.8) of the covering strip $\{(\tau, r) : \tau \in \mathbb{R}, 1 \le r \le 2\}$ of the annulus $S^1 \times [1, 2]$ into the covering plane $\mathbb{R} \times \mathbb{R}$ of the cylinder $S^1 \times \mathbb{R}$ satisfies the conditions of Moser's small twist theorem [M1], as extended later by Rüssman to the $C^{(4)}$-case [R1, 2]. We conclude that for small enough ε the map (4.8) has an invariant curve in the annulus. Retracting the sequence of transformations back to the original system, we conlude that there exist invariant curves of the Poincaré map $P : (x, \dot{x})_{t=0} \rightarrow (x, \dot{x})_{t=1}$ of the original system (1.1) arbitrarily far from the origin.

Equations (4.8) show, moreover, that the map satisfies the monotone twist condition, so that we can invoke the Aubry-Mather theory; retracing the transformations we obtain for the map $(x, \dot{x})_{t=0} \rightarrow (x, \dot{x})_{t=1}$ the existence of minimal Mather sets for any rotation number ϱ beyond some ϱ_0, and of Birkhoff periodic orbit for any $p/q > \varrho_0$.

In the rest of this section we prove the estimates (4.7) on the remainders, assuming (4.1)–(4.4).

To estimate $R_1 = \partial_h J_1$ we use (4.4), (4.3), and (4.5) in that order; the abbreviation $h(\varepsilon r) = h$ is used throughout:

$$\begin{aligned} |R_1| &= |\partial_h J_1(\tau, h(\varepsilon r), \theta)| \overset{(4.4)}{\le} |ch^{-1} J_0^{1-\beta}(h)| \\ &= \left| c \frac{h^{-1} J_0}{J_0'} J_0' J_0^{-\beta} \right| \overset{(4.3)}{\le} |c_1 J_0' J_0^{-\beta}| \overset{(4.5)}{=} |e_1 \varepsilon r J_0^{-\beta}|. \end{aligned}$$

We estimate $J_0^{-\beta}$ by using $J_0'(h) \overset{(4.2)}{\ge} ch^{-b}$ to obtain $h \ge c(H_0')^{-\frac{1}{b}}$; together with $J_0 \overset{(4.1)}{\ge} ch^a$ this gives $J_0 \ge c(J_0')^{-\frac{a}{b}} = c(\varepsilon r)^{-\frac{a}{b}}$, and thus the desired estimate $J_0^{-\beta} \le C\varepsilon^{\frac{a}{b}\beta}$. Using this we obtain

$$|R_1| \le c\varepsilon^{1+\frac{a}{b}\beta} \, ;$$

If we take $\nu \geq \dfrac{a}{b}\beta$ and choose $0 < \varepsilon \leq \varepsilon_0$ with ε_0 small enough, we obtain (4.7) for R_1 and for $k = l = 0$. For R_2 we have

$$|R_2| = \left| \frac{J_0''}{(J_1')2} (J_1)_\tau \right| \overset{(4.3)}{\leq} c\varepsilon \frac{h^{-2}J_0}{h^{-2}J_0^2} (J_1)_I \overset{(4.4)}{\leq} ceJ_0^{-1}J_0^{1-\beta}$$

$$= ceJ_0^{-\beta} \leq c\varepsilon^{1+\frac{a}{b}\beta} \leq \varepsilon^{1+\nu},$$

as claimed. The remaining estimates (4.7) on the derivatives of $R_{1,2}$ follow by differentiating these remainders and using the Hypothesis. To be more precise, each differentiation by I does not change the bound $\varepsilon^{1+\nu}$ by (4.4). We show that the application of the r-derivatives to R_1 does not violate the bounds either. The case of R_2 is treated similarly and is omitted. For the sake of brevity we let $F(h) = F(I, h, \theta) = \partial_2 J_1(I, h, \theta)$, so that $R_1 = F(h(\varepsilon r))$ (we suppress the remaining variables).

Differentiating $F(h(\varepsilon r))$ with respect to r we obtain

$$\partial_r F(h(\varepsilon r)) = F'h_r \quad \left(h_r \equiv \frac{\partial}{\partial r} h(\varepsilon r) \right)$$

$$\partial_r^2 F(h(\varepsilon r)) = F''h_r^2 + F'h_{rr} \tag{4.9}$$

$$\partial_r^3 F(h(\varepsilon r)) = F'''h_r^3 + 3F''h_r h_{rr} + F'h_{rrr}$$

$$\partial_r^4 F(h(\varepsilon r)) = F^{(4)}h_r^4 + 6F'''h_r^2 h_{rr} + 3F''h_{rr}^2 + 4F''h_r h_{rrr} + F'h_{rrrr}.$$

Our main estimate (4.4) gives

$$F(h) \leq ch^{-k-1}J_0^{1-\beta}, \qquad k = 1, \ldots, 4, \tag{4.10}$$

and we estimate $\left(\dfrac{\partial}{\partial r} \right)^{(k)} h(\varepsilon r)$. Differentiating the composition $J_0'(h(\varepsilon r))$ in (4.5) we obtain

$$J_0''(h)h_r = \varepsilon,$$

$$J_0''h_{rr} + J_0'''h_r^2 = 0,$$

$$J_0''h_{rrr} + 3J_0'''h_r h_{rr} + J_0^{(4)}h_r^3 = 0,$$

$$J_0''h_{rrrr} + 4J_0'''h_r h_{rr} + 3J_0'''h_{rr}^2 + 6J_0^{(4)}h_r^2 h_{rr} + J_0^{(5)}h_r^4 = 0.$$

Expressing the highest r-derivative from each of these four equations, we obtain

$$\left(\frac{\partial}{\partial r} \right)^k h(\varepsilon r) \leq ch(\varepsilon r), \qquad k = 1, 2, 3, 4; \tag{4.11}$$

indeed,

$$|h_r| = \left| \frac{\varepsilon}{J_0''(h)} \right| = \left| \frac{\varepsilon h}{hJ_0''(h)} \right| \overset{(4.3)}{\leq} c \frac{\varepsilon h}{J_0'(h)} = c \frac{\varepsilon h}{\varepsilon r} \leq ch;$$

using this in

$$h_{rr} = -\frac{J_0'''}{J_0''} h_r^2,$$

together with (4.3): $\left| \dfrac{J_0'''}{J_0''} \right| \leq ch$, we obtain $h_{rr} \leq ch$, etc.

Using now (4.10) and (4.11) in (4.9), we obtain the desired estimate on $\partial_r^{(k)} F$:

$$|\partial_r^k F| \leq ch^{-k-1} J_0^{1-\beta} \cdot h^k = ch^{-1} J_0^{1-\beta} \leq \varepsilon^{1+\nu},$$

the last inequality has been proven above. This completes the estimate on the remainder in (4.6).

5. Estimates on the Reduced Hamiltonian

It remains to prove the estimates (4.1)–(4.4) on $J = J_0 + J_1$. Actually, (4.1) and (4.2) follow from (4.3) at once by integrating the latter differential inequality, and only (4.3) and (4.4) remain to be proven. In this section we reduce these estimates to the ones on $H(I, t, \theta)$. More precisely, we will prove

Theorem 5.1. *If $H_0(I, t)$ and $H_1(\theta, I, t)$ obey, for some $\beta > 0$,*

$$|\partial_t^l \partial_I^k H_1(\theta, I, t)| \leq cI^{-k-\beta} H_0(I, t), \qquad l + k \leq 5, \tag{5.1}$$

and if the inverse function $I_0(H) = I_0(t, H)$ of $H_0(I, t)$ in H satisfies

$$|\partial_t^j \partial_H^k I_0(t, H)| \leq cH^{-k} I_0(t, H), \qquad j + k \leq 5, \tag{5.2}$$

$$cH^{-k} I_0(t, H)^{k \leq 2} \leq |\partial_H^k I_0(t, H)|, \tag{5.3}$$

then the reduced Hamiltonian $J = J_0(h) + J_1(\tau, h, \theta)$ satisfies (4.1)–(4.4).

Proof. We assume that (5.1)–(5.3) hold.

5.1 Proof of (4.4) is the longest and we address it first; we start by proving a similar estimate for I:

$$|\partial_t^l \partial_H^k I_1(t, H, \theta)| \leq cH^{-k} I_0(t, H)^{1-\beta}, \qquad l + k \leq 5, \tag{5.4}$$

and then we show that the action-angle map $(t, H) \rightarrow (\tau, h)$ satisfies

$$|\partial_\tau^i \partial_h^j H(\tau, h)| \leq ch^{1-j}, \tag{5.5}$$

$$|\partial_\tau^i \partial_h^j t(\tau, h)| \leq ch^{-j} \tag{5.6}$$

(these last two estimates depend only on the properties (5.2)–(5.3) of H_0 since the map is defined in terms of H_0 alone). The estimates (5.4)–(5.6) imply (4.4). Indeed, recall that $J_1(\tau, h, \theta) = I_1(\tau, H, \theta)$, where $t = t(\tau, h)$, $H = H(\tau, h)$, and $J_0(h) = \bar{H}_0^{-1}(h)$, the latter denoting the inverse function. Using now (5.4) and (5.5) in $\partial_\tau^l \partial_h^k J_1(\tau, h, \theta) \equiv \partial_t^l \partial_H^k I_1(t, H, \theta)$, where $t = t(\tau, h)$, $H = H(\tau, h)$ together with the induction in $l + k$, we arrive at (4.4). We prove (5.4) and (5.5)–(5.6) in the following two subsections.

An Inverse Function Estimate (5.4). Using the definition of $I(t, H, \theta)$ as the inverse function of $H(\theta, I, t) = H_0(I, t) + H_1(\theta, I, t)$:

$$H_0(I(H)) + H_1(I(H)) = H, \tag{5.7}$$

where we treat H as the independent variable and θ, t as parameters, we obtain

$$I(H) = I_0(H - H_1(I(H))), \tag{5.8}$$

and finally, expanding

$$I_1(H) \equiv I(H) - I_0(H) = I_0(H - H_1) - I_0(H)$$

in Taylor's series:

$$I_1(H) = -I_0'(H)H_1(I(H)) + \int_0^{H_1} sI_0''(H - H_1 + s)ds, \qquad H_1 = H_1(I(H)). \quad (5.9)$$

We will now estimate $H_1(I(H))$ via (5.7), then use it in (5.8) to estimate $I(H)$ and finally use this in (5.9) to estimate $I_1(H)$. To estimate $H_1(I(H))$, note that $I(H) \to \infty$ as $H \to \infty$ [as follows from (5.1) with $l = k = 0$ and (5.3)], and that $|H_1(I)| \leq \frac{1}{2}H_0(I)$ for all I large enough and for all t, θ; consequently,

$$|H_1(I(H))| < \frac{1}{2}H_0(I(H)) \quad (5.10)$$

for all H large enough and for all t, θ. To estimate new $I(H)$ we use (5.10), (5.8) and the monotonicity of I_0 in H [cf. (5.3)], obtaining

$$I_0\left(\frac{1}{2}H\right) < I(H) < I_0\left(\frac{3}{2}H\right), \quad (5.11)$$

which again holds for all sufficiently large H. Wishing to get rid of the factors $\frac{1}{2}$ and $\frac{3}{2}$ above, we use (5.3) to conclude that

$$I_0\left(\frac{1}{2}H\right) > c^{-1}I_0(H) \quad \text{and} \quad I_0\left(\frac{3}{2}H\right) < cI_0(H), \quad (5.12)$$

which leads [using (5.11)] to

$$c^{-1}I_0(H) < I(H) < cI_0(H). \quad (5.13)$$

This in turn is used to bound $H_1(I(H))$ as follows:

$$H_1(I(H) \overset{(5.1)}{\leq} cI(H)^{-\beta}H_0(I(H)) \overset{(5.13)}{\leq} c_1 I_0(H)^{-\beta}H_0(I(H)) \overset{(5.13)}{\leq}$$
$$c_1 I_0(H)^{-\beta}H_0(cI_0(H)).$$

From (5.3) we obtain $H_0(cI_0) \leq c_1 H_0(I_0) \equiv c_1 H$, and thus

$$H_1(I(H)) \leq cHI_0(H)^{-\beta}, \quad (5.14)$$

which we now finally use to estimate $I_1(H)$ from (5.9).

Estimating the first term, we get

$$|I_0'(H)H_1(I(H))| \overset{(5.14)}{\leq} cI_0'(H)HI_0^{-\beta}(H) \overset{(5.3)}{\leq} c_1 I_0^{1-\beta}(H),$$

as desired. The second term in (5.9) is bounded by

$$H^2 \sup_{H-H_1 \leq \tilde{H} \leq H} I_0''(\tilde{H}) \leq H_1^2 \sup_{H-H_1 \leq \tilde{H} \leq H} \tilde{H}^{-2}I_0(\tilde{H}) \overset{(5.1)}{\leq}$$
$$cH_1^2 H^{-2} \sup_{H-H_1 \leq \tilde{H} \leq H} I_0(\tilde{H}) \overset{(5.10,12)}{\leq}$$
$$c_1 H_1^2 H^{-2}I_0(H) \leq c_2 H^2 I_0^{-2\beta}H^{-2}I_0 = c_2 I_0(H)^{1-2\beta}.$$

This completes the proof of $|I_1| \leq cI_0^{1-\beta}$, and it remains to estimate the derivatives of I_1.

Estimating $\partial_H^k I_1(H)$, $1 \leq k \leq 5$. This is done by differentiating (5.9), as follows.

Considering H-derivatives at first, we note that each differentiation of $I_1(H)$ in (5.9) lowers its growth or rather the growth of its upper bound by one power of H, since for $1 \leq m \leq 5$ we have (1) $\partial_H^m I_0'(H) \leq cH^{-m-1} I_0(H)$ by (5.2); (2) $\partial_I^m H_1(I) \leq cI_{-\beta}^{-m} H_0(I)$ [by (5.1)]; and (3) $\partial_H^m I(H) \leq cH^{-m} I(H)$. We prove now this last inequality and thus (5-4) for $l = 0$. Differentiating (5.8):

$$I'(H) = I_0'(H - H_1(I(H))) \cdot (1 - H_1'(I(H))I'(H)),$$

we express

$$I'(H) = I_0'(H - H_1) \frac{1}{1 + I_0'(H - H_1) \cdot H_1'}, \tag{5.15}$$

where $H_1 = H_1(I(H))$. The denominator above is close to one for large H-indeed,

$$I_0'(H - H_1) \cdot H_1' = I_0'(H - H_1) \cdot H_0'(I_0(H - H_1)) \frac{H_1'(I(H))}{H_0'(I_0(H - H_1))}$$

$$= 1 \cdot \frac{H_1'(I(H))}{H_0'(I(H))} \overset{(5.13)}{\to} 0 \quad \text{as} \quad H \to \infty.$$

For H large enough we get

$$I'(H) < 2I_0'(H - H_1) \overset{(5.2)}{<} c(H - H_1)^{-1} I_0(H - H_1)$$

$$\leq c_1 H^{-1} I_0(H - H_1) = c_1 H^{-1} I(H),$$

proving the case $m = 1$. The estimates for higher m follow by differentiating (5.15) and using the estimates obtained in previous differentiations inductively: we note that the differentiation of (5.15) m times gives $I^{(m)}(H)$ in terms of the lower derivatives. We omit the tedious but obvious details. We conclude by pointing out that the differentiation of $I^{(m)}(H)$ with respect to t l times, $1 \leq l \leq 5 - m$ does not increase the order of growth of the upper bound, as follows by application (5.1) and (5.2) after each t-differentiation of (5.15).

5.1.2. Proof of the Transformation Estimate (5.5–6). We first express (τ, h) in terms of (t, H) and then show that the inverse function satisfies the desired estimates.

1. Expressing (I, h) via (t, H): From Eq. (3.2) and (3.4) we have the expressions for τ:

$$\tau = \int_0^t \frac{\partial H_0}{\partial I} (J_0(h), \xi)d\xi \cdot J_0'(h), \tag{5.16}$$

and it remains to express h via (t, H); recall that $I_0(t, H) = J_0(h)$ [Eq. (3.3)], and $J_0(h) = \bar{H}_0^{-1}(h)$, so that

$$h = \bar{H}_0(I_0(t, H)) \overset{\text{def}}{\equiv} G(t, H). \tag{5.17}$$

Substituting the last expression into (5.16) we obtain

$$\tau = \int_0^t \frac{\partial H_0}{\partial I} (I_0(t, H), \xi)d\xi \cdot J_0'(\bar{H}_0(I_0(t, H))) \equiv F(t, H); \tag{5.18}$$

Eqs. (5.18) and (5.17) give an explicit expression for the inverse of the map $(\tau, h) \rightarrow (t, H)$ which we now proceed to estimate.

2. Note that F, G obey the estimate

$$|\partial_t^i \partial_H^j F| \le cH^{-j}F ,$$
$$|\partial_t^i \partial_H^j G| \le cH^{1-j}G , \qquad i+j \le 5, \tag{5.19}$$

as follows from the estimate (5.2) on the unperturbed Hamiltonian. (We recall that the map is determined by H_0 alone.)

3. Denoting $\begin{pmatrix} \tau \\ h \end{pmatrix} = X$, $\begin{pmatrix} t \\ H \end{pmatrix} = Y$, $\begin{pmatrix} F \\ G \end{pmatrix} = \varphi$, we rewrite (5.18) and (5.17) as $X = \varphi(Y)$ and differentiate by $X : \dfrac{\partial Y}{\partial X} = \varphi'(Y)^{-1}$. Since the action-angle map φ is area-preserving, i.e. $\det \varphi'(Y) = 1$, we obtain a simplification for the inverse

$$\frac{\partial Y}{\partial X} = L\varphi'(Y), \tag{5.20}$$

where L is the linear operator acting on 2×2 matrices according to $L\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

4. The last relationship (5.20) used inductively with the estimates (5.19) on φ' and its derivatives gives (5.5) and (5.6).

5.2. Proof of (4.3). We recall that $J_0(h)$ was defined by

$$\bar{H}_0(J_0(h)) = \int\limits_0^1 H_0(J_0(h), t)dt = h ,$$

i.e. $J_0(h) = \bar{H}_0^{-1}(h)$ is the inverse of the *t*-verage of the inverse of I_0, and proving (4.3) amounts to the following general statement on implicit functions:

Lemma 5.1. *If the function $F(x, t)$ satisfies*

$$cF \overset{k \le 2}{\le} x^k |\partial_x^k F(x, t)| \overset{k \le 5}{\le} CF ,$$
$$\textit{with no } |\cdot| \textit{ for } k = 1, \textit{ for } x \textit{ large enough} \tag{5.21}$$

and if

$$|F_t| \le cF \tag{5.22}$$

then the inverse function f of the t-average of the x-inverse of F, defined by

$$\int\limits_0^1 F^{-1}(f(x), t)dt = x , \tag{5.23}$$

where F^{-1} is the inverse function of $F(x, t)$ in x, satisfies

$$cf(x) \overset{k \le 2}{\le} x^k |f^{(k)}(x)| \overset{k \le 5}{\le} Cf(x); \quad \textit{no } |\cdot| \textit{ for } k = 1, \tag{5.24}$$

the same estimate as F.

To apply this lemma we choose $x = h$ and $F(x, t) = I_0(h, t)$; recalling the definition of $J_0(h)$ we have $g(x) \equiv J_0(h)$ and (5.24) translates into the desired estimate (4.3).

Proof of Lemma 5.1 goes by differentiating (5.23).

1. *Case $k = 1$*. First, we obtain $\left(\text{with } ' = \dfrac{\partial}{\partial x} \right)$:

$$f'(x) = \left(\int_0^1 \frac{dt}{F'(F^{-1}(f(x), t), t)} \right)^{-1} ;$$

Using the estimate (5.21) on F' we obtain

$$c \left(\int_0^1 \frac{dt}{F(F^{-1}(f(x), t), t)} \right) \cdot \min_t F^{-1}(f(x), t) \le f'(x)$$

$$\le C \left(\int_0^1 \frac{dt}{F(F^{-1}(f(x), t), t)} \right)^{-1} \cdot \max_t F^{-1}(f(x), t)$$

or

$$cf(x) \min_t F^{-1}(f(x), t) \le f'(x) \le Cf(x) \max_t F^{-1}(f(x), t).$$

It remains to prove that

$$cx \le \min_t F^{-1}(f(x), t) \le \max_t F^{-1}(f(x), t) \le Cx. \tag{5.25}$$

To that end we first show that (5.22) holds for F^{-1} as well:

$$|\partial_t F^{-1}(y, t)| \le cF^{-1}(y, t). \tag{5.26}$$

Differentiation of $F^{-1}(F(x, t), t) = x$ by t gives

$$F_2^{-1}(F(x, t), t) = -F_1^{-1}(F(x, t), t) \cdot F_t(x, t) = -\frac{F_t(x, t)}{F_x(x, t)},$$

where 2 and 1 indicate partial derivatives with respect to the second and the first arguments. Using (5.22) for the numerator and (5.21) for the denominator we obtain

$$|F_2^{-1}(F(x, t), t)| \le cx.$$

For an arbitrary $y > 0$ large enough there exists (correspondingly large) x such that $y = F(x, t)$, and the last estimate gives (5.26), which we use to prove (5.24). Dividing (5.26) by F^{-1} and integrating from $t = t_0$ to t, we obtain

$$F^{-1}(y, t_0)e^{-c(t-t_0)} \le F^{-1}(y, t) \le F^{-1}(y, t_0)e^{c(t-t_0)}.$$

Pick now a specific t_0, namely the one which gives the integrand in (5.23) its average value: $F^{-1}(f(x,), t_0) = x$; setting $y = f(x)$, we get

$$xe^{-c(t-t_0)} \le F^{-1}(f(x), t) \le xe^{c(t-t_0)}.$$

Taking max and min over $0 \le t \le 1$ yields (5.25) thereby proving (5.24) for $k = 1$.

2. *Case $k \geq 2$.* Differentiating the definition (5.23) of the inverse we obtain the expressions for $f^{(k)}$ with the notation $G(f(x), t) \equiv F^{-1}(f(x), t)$:

$$\int_0^1 G' dt \cdot f' = 1,$$

$$\int_0^1 G'' dt \cdot (f')^2 + \int G' dt \cdot f'' = 0,$$

$$\int_0^1 G''' dt \cdot (f')^3 + 3 \int_0^1 G'' dt \cdot f' f'' + \int_0^1 G' dt \cdot f''' = 0, \qquad (5.27)$$

$$\int_0^1 G^{(4)} dt \cdot (f')^4 + 6 \int_0^1 G''' dt \cdot (f')^2 f'' + 3 \int_0^1 G'' dt \cdot (f'')^2$$

$$+ 4 \int_0^1 G'' dt \cdot f' f''' + \int_0^1 G' dt \cdot f^{(4)} = 0,$$

and finally

$$\int_0^1 G^{(5)} dt \cdot (f')^5 + 10 \int_0^1 G^{(4)} dt \cdot (f')^3 f'' + 15 \int_0^1 G''' dt \cdot f' (f'')^2$$

$$+ 10 \int_0^1 G''' dt \cdot (f')^2 f''' + 10 \int_0^1 G'' dt \cdot f'' f'''$$

$$+ 5 \int G'' dt \cdot f' f^{(4)} + \int G' dt \cdot f^{(5)} = 0 ;$$

here $G^{(m)} = \left(\dfrac{\partial}{\partial y} \right)^m G(y, t)|_{y=f(x)}$ and $f \equiv f(x)$.

The above identities (5.27) give us $f^{(k)}(x)$ in terms of its lower derivatives:

$$f^{(k)}(x) = \left(\int_0^1 G' dt \right)^{-1}$$

$$\times \sum_{\substack{m+n_1+\ldots+n_m=k \\ m \geq 2, \, n_i \geq 0}} a_{mn_1 \ldots n_m} \int_0^1 G^{(m)} dt \cdot f^{(n_1+1)} \ldots f^{(n_m+1)}, \qquad (5.28)$$

with integer coefficients $a_{mn_1 \ldots n_m}$, and we estimate $f^{(k)}(x)$ by induction. Assume that (5.24) holds up to $k - 1$ derivatives and prove it for k. We have just shown that the first factor $\left(\int_0^1 G' dt \right)^{-1} = f'$ lies between $cx^{-1}f$ and $Cx^{-1}f$ for large x,

and only the sum in (5.28) needs be estimated. To that end we need an estimate on the inverse function $G(y, t)$ of $F(x, t)$.

Lemma 5.2. *If a real function $\varphi(x)$ satisfies $|\varphi^{(m)}(x)| \leq cx^{-m}\varphi$, $1 \leq m \leq k$ and $\varphi'(x) \geq x^{-1}\varphi(x)$ for all x large enough, then the inverse function $\psi = \varphi^{-1}$ satisfies $|\psi^{(m)}(y)| \leq cy^{-m}\psi(y)$ for $1 \leq m \leq k$. If, moreover, $|\varphi''(x)| \geq cx^{-2}\varphi(x)$, then $|\psi''(y)| \geq cy^{-2}\psi(y)$.*

Proof of this lemma goes by differentiating the identity $\varphi(\psi(y)) = y$ and using the estimates given in the statement of the lemma. We omit the details. □

Applying Lemma 5.2 to $G(y, t) = F^{-1}(y, t)$ we obtain $cy^{-m}G(y, t) \leq |G^{(m)}(y, t)| \leq Cy^{-m}G(y, t)$, where the first inequality holds for $m = 1, 2$ and the second for $1 \leq m \leq 5$, as follows from (5.21). We note that the constants can be chosen independent of t. Substituting $y = f(x)$ in the above inequality we obtain

$$cf^{-m}(x)G(f(x), t) \leq |G^{(m)}(f(x), t)| \leq Cf^{-m}(x)G(f(x), t),$$

which implies, by (5.25), that

$$cxf^{-m}(x) \leq G^m \leq Cxf^{-m}(x),$$

and it remains to estimate the products of the derivatives of f in (5.28).

We note that in (5.28) $n_i + 1 \leq k - m + 1 \leq k - 1$, so that the inductive assumption on the derivatives of $f(x)$ is applicable to (5.28); we obtain (for the upper bound):

$$f^{(k)}(x) \leq Cx^{-1}f(x) \sum_{\substack{m+n_1+\ldots+n_m=k \\ m \geq 2, \, n_i \geq 0}} a_{mn_1\ldots n_m}(C_m x f^{-m}(x))$$

$$\times (C_{n_1} \cdot x^{-n_1-1}f) \cdot \ldots \cdot (C_{n_m}x^{n_m-1}f)$$

$$\leq cx^{-1+1-(n_1+1)-\ldots-(n_m+1)}f^{1-m+1\ldots+1}(x) = cx^{-k}f(x),$$

the lower bound on $f''(x)$ is established in the same way using (5.28) for $k = 2$ [i.e. (5.27b)].

This completes the proof of Lemma 5.1 and thus of the estimate (4.3). □

6. Estimating the Action-Angle Hamiltonian $H(\theta, I, t)$ in Terms of the Potential $V(x, t)$

In the preceding sections we have reduced the problem to the estimates (5.1)–(5.3) on the action-angle Hamiltonian $H = H_0(I, t) + H_1(\theta, I, t)$; (5.1) says that H_1 should not be too large, (5.2) says that H_0 should not behave too wildly and (5.3) is a twist condition, i.e. a convexity condition on the Hamiltonian.

In this section we finally address the main difficulty of the problem: find a set of conditions on $V(x, t)$ which would imply the above estimates (5.1)–(5.3).

Theorem 6.1. *If the potential $V(x, t)$ satisfies the superquadraticity condition*

$$W_x \leq a < \frac{1}{2}, \tag{6.A}$$

together with

$$|\partial_x^p \partial_t^q U|, \quad |\partial_x^p \partial_t^q W| \leq cx^{1-p}, \quad p + q \leq 5, \tag{6.B}$$

where $W = V/V_x$ and $U = V_t/V_x$, and

$$|x^p \partial_x^p \partial_t^q V(x, t)| \le cV^{1+\mu}, \qquad \mu = \text{const} > 0, \qquad p + q \le 5 \qquad (6.\text{C})$$

for all x, t, then there exists $\beta = \beta(\mu, a) > 0$ such that the bound (5.1) holds:

$$|\partial_t^l \partial_I^k H_1(\theta, I, t)| \le cI^{-k-\beta} H_0(I, t), \qquad l + k \le 5. \qquad (5.1)$$

Theorem 6.2. *If the potential V satisfies the above estimates (6.B), then $H_0(I, t)$ satisfies (5.2):*

$$|\partial_t^j \partial_H^k I_0(t, H)| \le cH^{-k} I_0(t, H), \qquad j + k \le 5. \qquad (5.2)$$

Theorem 6.3. *If V satisfies (6.A) then $H_0(I, t)$ satisfies the twist conditions (5.3),*

$$cH^{-k} I_0(t, H)| \overset{k \le 2}{\le} \partial_H^k I_0(t, H). \qquad (5.3)$$

6.1. Proof of Theorem 6.1. 1. We begin by showing that the superquadracity condition (6.A) implies the superlinearity of H_0 in I:

$$H_0(I) \ge cI^{1 + \frac{1-2a}{1+2a}} ; \qquad (6.1.1)$$

we suppress the t-dependence in the notation.

To relate I and $H_0(I)$, we note that the level curve $y^2/2 + V(x) = H_0(I)$ lies inside the rectangle with the sides $2\sqrt{2H_0(I)}$ and $2V^{-1}(H_0(I))$, so that

$$I \le 4\sqrt{2H_0(I)} V^{-1}(H_0(I)).$$

Using now the bound $V \ge cx^{1/a}$ which follows by integrating the superquadraticity condition (6.A), we obtain $cV^{-1} H_0(I)) \le H_0(I)^a$, and thus

$$I \le cH_0(I)^{\frac{1}{2}+a},$$

whic is equivalent to (6.1.1).

2. We show now that if for some $\delta > 0$ the inequality

$$|\partial_t^l \partial_I^k H_1(\theta, I, t)| \le cI^{-k+1} H_0(I, t)^\delta, \qquad l + k \le 5 \qquad (6.1.2)$$

holds, then it together with (6.1.1) implies (5.1) with β given by

$$\beta = \left(1 + \frac{1-2a}{1+2a}\right)(1-\delta) - 1. \qquad (6.1.3)$$

This reduces the proof of (5.1) to showing that (6.1.2) holds with some $0 < \delta < \frac{1}{2} - a$; the last inequality is necessary to have $\beta > 0$.

Indeed, the implication $(6.1.2) \Rightarrow (5.1)$ amounts to showing that (for all I large enough and for all t)

$$I^{-k+1} H_0^\delta \le cI^{-k-\beta} H_0, \qquad (6.1.4)$$

$H_0 = H_0(I, t)$, or equivalently,

$$I^{\frac{1+\beta}{1-\delta}} \le cH_0,$$

which holds because of (6.1.1) if $\dfrac{1+\beta}{1-\delta} = 1 + \dfrac{1-2a}{1+2a}$, which is the same as (6.1.3).

This reduces the estimates (5.1) to (6.1.2) wherein the main difficulty sits. We will prove that (6.1.2) holds with $\delta = 100\mu$.

In order to assure $\beta > 0$ we need $\delta < \frac{1}{2} - a$, or $\mu \leq \frac{1}{100}(\frac{1}{2} - a)$, as in fact is required in the main theorem.

3. *The expression for H_1.* With all the reductions now finally out of the way, we come to the substance, i.e. to the proof that (6.1.2) holds with $\delta = 100\mu$.
Recall that $H_1 = S_t(x, I, t)|_{x=x(\theta, I, t)}$, and

$$S(x, I, t) = \sqrt{2} \int_0^x \sqrt{H(I, t) - V(\xi, t)}d\xi.$$

From now on we write H instead of H_0 without risk of ambiguity. Differentiation of S with respect to the last argument yields

$$H_1(\theta, I, t) = \frac{\sqrt{2}}{2} \int_0^x \frac{H_t(I, t) - V_t(\xi, t)}{\sqrt{H(I, t) - V(\xi, t)}} d\xi. \qquad (6.1.5)$$

It should be emphasized that $x = x(\theta, I, t)$ in the last expression.
Let $K = K(I, t, \xi) = H_t(I, t) - V_t(\xi, t)$; differentiating (6.1.5) with respect to I yields, according to (A3.12) (see Appendix):

$$\partial_I H_1 = \partial_I \left\{ \int_0^x \frac{K(I, t, \xi)}{\sqrt{H - V(\xi, t)}} d\xi \right\} = -K(I, t, x) \int_0^x L(\xi, I, t) \frac{d\xi}{\sqrt{H - V}}$$

$$+ \int_0^x \mathscr{L}(K) \frac{d\xi}{\sqrt{H - V}}; \qquad (6.1.6)$$

L is given in terms of H and V by (see Appendix)

$$-L(x, I, t) = \frac{H_{II}}{H_I} + \frac{H_I}{2H}\left(1 - 2\frac{V V_{xx}}{V_x^2}\right) = \frac{H_{II}}{H_I} + \frac{H_I}{H}\left(-\frac{1}{2} + W_x\right),$$

and $\mathscr{L}$ is the linear differential operator acting on functions of x, I, t, according to

$$\mathscr{L}(f) = \frac{H_I}{H}\left\{\left(f\frac{V}{V_x}\right)_x - \frac{1}{2}f\right\} + f_I.$$

4. *Proof of (6.1.2) with $l = 0$:* it suffices to show that ($\delta = 100\mu$ and $d\xi = d\xi/\sqrt{H - V(\xi)}$ throughout):

$$|\partial_I^k K(x, I, t)| \leq cI^{-k}H^{1+\delta/2}, \qquad k \leq 4, \qquad (6.1.7)$$

$$\left|\partial_I^k \left(\int_0^x L(\xi, I, t)d\xi\right)\right| \leq cI^{-k}H^{-1+\delta/2}, \qquad k \leq 4, \qquad (6.1.8)$$

$$\left|\partial_I^k \left(\int_0^x \mathscr{L}(K)d\xi\right)\right| \leq cI^{-k}H^{\delta}, \qquad k \leq 4, \qquad (6.1.9)$$

where $x = x(\theta, I, t)$ and ∂_I refers to the total derivative: $\partial_I f(x, I) = f_x x_I + f_I$.

4.1. Proof of (6.1.7).

$$|\partial_I^k K(x, I, t)| \le |\partial_I^k \partial_t H(I, t)| + |\partial_I^k V_t(x, t)|.$$

According to Theorem 6.2 and Lemma A1.1, the first term on the right is bounded by $cI^{-k}H(I, t)$, for $k \le 4$; this is better than we need. To estimate the second term $\partial_I^k V_t(x, t)$ we have to differentiate the composition of two functions $(V_t(\cdot, t)$ and $x(\cdot, \theta, t))$; this was already done in (4.9):

$$D_1 \equiv \partial_I V_t(x, t) = V_{xt}x_I,$$
$$D_2 \equiv \partial_I^2 V_t(x, t) = V_{xxt}x_I^2 + V_{xt}x_{II},$$
$$D_3 \equiv \partial_I^3 V_t(x, t) = V_{xxxt}x_I^3 + 3V_{xxt}x_Ix_{II} + V_{xt}x_{III},$$
$$D_4 \equiv \partial_I^4 V_t(x, t) = V_{xxxxt}x_I^4 + 6V_{xxxt}x_I^2x_{II} + 3V_{xxt}x_{II}^2 + 4V_{xxt}x_Ix_{III} + V_{xt}x_{IIII}.$$

Combining (6.C) with the estimates on the I-derivatives of x from Lemma A4.1: $\partial_I^k x(\theta, I, t) \le cI^{-k}xH^\nu$, we obtain, recalling that $V(x, t) \le H(I, t)$:

$$|D_j| \le cI^{-j}H^{\nu+\mu+1}.$$

This proves that (6.1.7) holds with $\delta = 11\mu$.

4.2. Proof of (6.1.8).

Case $k = 1$: We have, in the notation $đx \equiv dx/\sqrt{H - V(x)}$ (Appendix A3):

$$\partial_I \int_0^x L\,đ\xi = L \int_0^x L\,đ\xi + \int_0^x \mathscr{L}(L)đ\xi, \qquad (6.1.10)$$

and since $|L| \le cI^{-1}$ and $\mathscr{L}(L) \le cI^{-2}$ (Lemma A2.1), we obtain

$$\left| \partial_I \int L\,đ\xi \right| \le cI^{-2} \int_0^x đ\xi \le cI^{-2} \int_0^{x_+} đ\xi,$$

where $x_+(I, t) \equiv x_+(\theta = \frac{1}{4}, I, t)$. We note that

$$\frac{1}{\sqrt{2}} \int_0^{x_+} đ\xi = \frac{1}{4} \frac{1}{H_I(I, t)}$$

$(= \frac{1}{4}T(I, t)$, the period for the associated autonomous system, which is the derivative of the area with respect to the energy [A1]).

Since $H_I \ge cI^{-1}H$, we obtain

$$\partial_I \int_0^x L\,đ\xi \le cI^{-1}H^{-1}, \qquad (6.1.11)$$

as desired, completing the case $k = 1$.

Case $k = 2$: Applying ∂_I to (6.1.10), we obtain:

$$\partial_I^2 \int L d\xi = \left\{ (\partial_I L) \int L d\xi, \ L\partial_I \int L d\xi, \ \mathscr{L}(L) \int L d\xi, \ \int \mathscr{L}^2(L) d\xi \right\}, \quad (6.1.12)$$

where $\{f_1, \ldots, f_n\}$ denotes a linear combination of functions $f_1, \ldots, f_n$ with integer coefficients.

First, $|\partial_I L| \le cI^{-2}$ by Lemma A2.1. Furthermore, since

$$\int L d\xi \le cI^{-1} H_I^{-1} \le cH^{-1}, \quad (6.1.13)$$

we obtain the desired bound on the first term in (6.1.12). The same bound is valid for the second term in (6.1.12) using (6.1.11) and the fact that $|L| \le cI^{-1}$. For the remaining two terms we use $|\mathscr{L}(L)| \le cI^{-2}H^\nu$ and $\mathscr{L}^2(L) \le cI^{-3}H^\nu$ (Lemma A2.1). This completes the case $k = 2$.

Case $k = 3$: Applying ∂_I to (6.1.12) we obtain

$$\partial_I^3 \int L d\xi = \left\{ (\partial_I^2 L) \int L, \ (\partial_I L) \left(\partial_I \int L \right), \ L \left(\partial_I^2 \int L \right), \ (\partial_I \mathscr{L}(L)) \int L, \right.$$
$$\left. \mathscr{L}(L)\partial_I \int L, \ \mathscr{L}^2(L) \int L, \ \int \mathscr{L}^3(L) \right\}, \quad (6.1.14)$$

where $\int L \equiv \int L d\xi$, etc.; we prove that all these terms are bounded by $cI^{-3}H$. Considering the first one we have by Lemma A2.1:

$$|\partial_I^2 L| = \left| \partial_I^2 \left(\frac{H_{II}}{H_I} + \frac{H_I}{2H} \right) + \partial_I^2 \left(\frac{H_I}{H} W_x \right) \right| \le cI^{-3} + \frac{H_I}{H} |\partial_I^2 W_x|, \quad (6.1.15)$$

where

$$|\partial_I^2 W_x| = |\partial_I (W_{xx} x_I)| = |W_{xxx} x_I^2 + W_{xx} x_{II}| \le cI^{-2}H^{2\nu},$$

using the assumptions on W and the estimates on x_I, x_{II} from Lemma A4.1. This gives $|\partial_I^2 L| \le cI^{-3}H^{2\nu}$. Using $|\int L| \le cH^{-1}$ we obtain the desired bound on $\partial_I^2 L \int L$. That the same bound $cI^{-3}H^{-1+2\nu}$ holds for the second, third and fifth terms in braces in (6.1.14) follows from the estimates already mentioned in previous steps. The remaining terms in (6.1.14) satisfy the same bound according to Lemma A2.1.

Case $k = 4$: From (6.1.14) we get

$$\partial_I^{k+1} \int L = \left\{ \partial_I^n \mathscr{L}^m(L) \partial_I^{k-m-n} \int L, \ \int \mathscr{L}^{k+1}(L) \ 0 \le m + n \le k \right\},$$

or more explicitly,

$$\partial_I^4 \int L d\xi = \left\{ (\partial_I^n L)\partial_I^{3-n} \int L \ (n = 0, 1, 2, 3), \ \partial_I^2 \mathscr{L}(L) \int L, \ (\partial_I \mathscr{L}(L))\partial_I \int L, \right.$$
$$\mathscr{L}(L)\partial_I^2 \int L, \ (\partial_I \mathscr{L}^2(L)) \int L, \ \mathscr{L}^2(L)\partial_I \int L,$$
$$\left. \mathscr{L}^3(L) \int L, \ \int \mathscr{L}^4(L) \right\}. \quad (6.1.16)$$

We have $|(\partial_I^n L)(\partial_n^{3-n} \int L)| \leq cI^{-4}H^{-1+3\nu}$ – for $n = 0, 1, 2$ this follows from our previous steps, and for $n = 3$ we have, using (6.1.15) and Theorem 6.2:

$$|\partial_I^3 L| \leq cI^{-4} + \left|\frac{H_I}{H}\partial_I^2 W_x\right|,$$

and $|\partial_I^3 W_x| = |W_{xxxx}x_I^3 + 3W_{xxx}x_I x_{II} + W_{xx}x_{III} \leq cI^{-3}H^{3\nu}$ by the asumptions on W and Lemma A4.1, so that

$$|\partial_I^3 L| \leq cI^{-4}H^{3\nu}.$$

The remaining terms in (6.1.16) are estimated in the same way as above using Lemma A2.1.

This completes the proof of (6.1.8)

4.3. *Proof of (6.1.9).* Abbreviating $\mathscr{L}(K) = \mathscr{L}$, $\mathscr{L}^2(K) = \mathscr{L}(\mathscr{L}(K)) = \mathscr{L}^2$, etc., we use the differentiation formula to obtain

$$\partial_I^p \int \mathscr{L}(K)d\xi = \left\{(\partial_I^i \mathscr{L}^m(K))\left(\partial_I^j \int Ld\xi\right)_{i+j+m=p}, \int \mathscr{L}^{p+1}d\xi\right\},$$

where $m \geq 1$, $i, j \geq 0$ and $i + j \leq p = 1, 2, 3, 4$. Applying Lemma A2.2 to this formula, we obtain

$$\left|(\partial_I^i \mathscr{L}^m)\left(\partial_I^i \int L\right)\right| \leq cI^{-i-m}H^{1+\nu}I^{-j}H^{-1+\nu} = cI^{-p}H^{2\nu}.$$

Using Lemma A2.2 again we estimate ($p \leq 4$):

$$\int \mathscr{L}^{p+1}d\xi \leq cI^{-p-1}H^{1+\nu}\int d\xi \leq cI^{-p-1}H^{1+\nu}\frac{1}{H_I} \leq cI^{-p}H^{\nu}.$$

This completes the proof of (6.1.9) and thus of (6.1.2).

4. *Proof of (6.1.2) with $k = 0$.* Differentiating H_1 we obtain

$$\partial_t H_1 = \partial_t \int K d\xi = K \int M d\xi + \int \mathscr{M}(K)d\xi, \qquad (6.1.17)$$

where M and $\mathscr{M}$ are given by (A3.5) and (A3.6). The desired first estimate $|\partial_t H_1| \leq cIH_0^\delta$ follows from $|H_t| \leq cH$ (Theorem 6.2), $|V_t| \leq cVH^\mu$ (assumption), $|M| \leq cH^\mu$, $\mathscr{M}(K) \leq cH^{1+\nu}$ (Lemma A2.2) – these estimates give

$$|\partial_t H_1| \leq cH^{1+2\nu}\int d\xi \equiv cH^{1+2\nu}\frac{1}{\dfrac{\partial H}{\partial I}} \leq cIH^{2\nu} < cIH^\mu,$$

the inequality before last holding by Theorem 6.2.

Differentiating now (6.1.17) in order to estimate higher derivatives, it suffices to prove that (here $x = x(\theta, I, t)$, as usual):

$$|\partial_t^l K(x, I, t)| \leq cH^{1+\delta/2}, \qquad (6.1.18)$$

$$\left|\partial_t^l \int_0^x M(\xi, I, t)d\xi\right| \leq cH^{-1+\delta/2}, \qquad (6.1.19)$$

$$\left|\partial_t^l \int \mathscr{M}(K)d\xi\right| \leq cH^\delta, \qquad (6.1.20)$$

with $\delta \leq 100\mu$.

5.1. Proof of (6.1.18). $\partial_t^l K = \partial_t^{l+1} H(I, t) - \partial_t^l V_t(x, t)$; the first term is bounded by cH according to Theorem 5.2 and Lemma A1.1. The second term is given by

$$d_t^l V_t(x, t) = \sum_{\substack{r_1 + \ldots + r_q + s = l \\ r_j \geq 1, q \geq 0, s \geq 0}} C_{qrs} \frac{\partial^{q+s+1} V}{\partial x^q \partial t^{s+1}} \partial_t^{r_1} x \cdot \ldots \cdot \partial_t^{r_q} x, \tag{6.1.21}$$

where C_{qrs} are integers and d_t denotes the total derivative. The value $q = 0$ corresponds to the term $\dfrac{\partial^{l+1} V}{\partial t^{l+1}}$ in (6.1.21) (with no factors involving x).

By Lemma A4, $|\partial_t^{r_1} x \cdot \ldots \cdot \partial_t^{r_q} x| \leq x^q H^{qv}$. Using this inequality together with (6.C) in (6.1.21) we obtain (6.1.18) with $\delta = \mu + 4v = 81\mu$.

5.2. Proof of (6.1.19). The proof is almost identical to that of (6.1.8).

5.3. Proof of (6.1.20). The proof is analogous to that of (6.1.9).

6. Proof of (6.1.2) for all k, l. We represent $\partial_t^l \partial_I^k H_1$ by the points in the integer lattice

It remains to estimate the terms not lying on the lines $l = 0$ and $k = 0$; we proceed in the direction indicated by the arrows, considering the four columns in the diagram.

For the column $k = 1$ we need to estimate

$$\partial_t^l \partial_I H_1 = \partial_t^l \partial_I \int K = \partial_t^l \left(K \int L + \int \mathscr{L}(K) \right), \quad 1 \leq l \leq 4,$$

and it suffices to prove [cf. (6.1.7)–(6.1.9) that for $l = 1, 2, 3, 4$,

$$|\partial_t^l K| \leq c H^{1+\delta/2}, \tag{6.1.22}$$

$$\left| \partial_t^l \int L \right| \leq c H^{-1+\delta/2}, \tag{6.1.23}$$

$$\left| \partial_t^l \int \mathscr{L}(K) \right| \leq c H^\delta. \tag{6.1.24}$$

(6.1.22) has already been proven.

For (6.1.23) we use $\partial_t^l \int L = \partial_t^{l-1}(L \int M + \int \mathscr{M}(L))$, and estimate the t-derivatives of each term in the last expression. By Lemma A2.1 we have

$$|\partial_t^{l-1} L| \leq cI^{-1} H^\nu .\qquad (6.1.25)$$

The t-derivatives of $\int M$ are given by (6.1.19), and we estimate the remaining term

$$\left|\partial_t^{l-1} \int \mathscr{M}(L)\right| \leq cH^{-1+\nu}, \qquad l = 1, 2, 3, 4,$$

using Lemma A2.1.

This completes the proof of (6.1.23).

Proof of (6.1.24). We use the identity

$$\partial_t^l \int \mathscr{L}(K) = \left\{ [\partial_t^i \mathscr{M}^{m-1} \mathscr{L}(K)] \left[\partial_t^n \int M\right]_{i+m+n=l}, \int \mathscr{M}^l \mathscr{L}(K)\right\} ;$$

applying Lemma A2.2 we obtain (6.1.24).

The remaining columns $k = 2, 3, 4$ are estimated in the same way, using the results of Lemmas A2.1, A2.2, and A4.1. The cases $k = 2, 3, 4$ are done analogously. This completes the proof of Theorem 6.1. $\square$

6.2. Proof of Theorem 6.2. We recall that $I_0(t, H) = 4 \int\limits_0^{x_+(H,t)} \sqrt{2(H - V(\xi, t))}d\xi.$

Let us denote the operators involved in the differentiation formulas (A3.2) and (A3.3) by $\mathscr{H}$ and $\mathscr{T}$:

$$\mathscr{H}(K) = \frac{1}{2H} K + K_H + \frac{1}{H}(KW)_x, \qquad W = \frac{V}{V_x},\qquad (6.2.1)$$

$$\mathscr{T}(K) = K_t - (KU)_x, \qquad U = \frac{V_t}{V_x},\qquad (6.2.2)$$

so that

$$\frac{\partial}{\partial H} I_0(t, H) = 4\sqrt{2} \int\limits_0^{x(H,t)} \mathscr{H}(1)\sqrt{H - V} \, d\xi$$

and

$$\frac{\partial}{\partial t} I_0 = 4\sqrt{2} \int \mathscr{T}(1)\sqrt{H - V} \, d\xi ,$$

and the proof reduces to showing that any composition of $\mathscr{H}$ and $\mathscr{T}$ of length ≤ 5 applied to $K = 1$ gives a function of (t, H, x), $0 \leq x \leq x_+(H, t)$ bounded by cH^{-k}, k being the number of occurrences of $\mathscr{H}$ in the composition. Note first that since $\frac{\partial}{\partial H}$ and $\frac{\partial}{\partial t}$ commute, so do $\mathscr{T}$ and $\mathscr{H}$, and it suffices to show that

$$\left| \underbrace{\mathscr{T} \circ \ldots \circ \mathscr{T}}_{j} \circ \underbrace{\mathscr{H} \circ \ldots \circ \mathscr{H}}_{k}(1)\right| \leq cH^{-k}, \qquad i + j \leq 5.\qquad (6.2.3)$$

Lemma 6.2.1. *The kth composition of $\mathscr{H}$ applied to $K = 1$ is given by*

$$\mathscr{H}^k(1) = \frac{1}{H^k} P_k(W, \partial_x W, \ldots, \partial_x^k W) ,\qquad (6.2.4)$$

where P_k is a polynomial of degree k in its $k+1$ variables, and where each monomial $W^{(i_1)} \cdot \ldots \cdot W^{(i_n)}$ ($n \leq k$) is such that the number of differentiations equals the number of factors: $\sum\limits_{p=1}^{n} i_p = n$.

Proof of the lemma. For $k = 1$ the statement is obvious. Equation (6.2.1) gives the recurrence relation

$$P_k = \left(\frac{1}{2} + W_x\right) P_{k-1} + W \frac{d}{dx} P_{k-1} - \frac{1}{k-1}, \qquad (6.2.5)$$

which shows that if the statement of the lemma holds for $k - 1$ then it does for k. $\square$

Lemma 6.2.2. *The compositions of $\mathscr{T}$ and $\mathscr{H}$ applied to $K = 1$ is given by[4]*

$$\mathscr{T}^j \circ \mathscr{H}^k(1) = \frac{1}{H^k} P_{jk}(U^{(p,q)}, W^{(r,s)}) : \begin{cases} p + q \leq j, & r + s \leq j + k \\ p \leq j - 1, & r \leq j, \end{cases} \qquad (6.2.6)$$

where P_{jk} is a polynomial of degree $j + k$ and each monomial

$$U^{(p_1,q_1)} \cdot \ldots \cdot U^{(p_m,q_m)} W^{(r_1,s_1)} \cdot \ldots \cdot W^{(r_n,s_n)} (m + n \leq j + k)$$

has the property: the number of factors equals the number of x-differentiations:

$$\sum_{1}^{m} q + \sum_{1}^{n} s = m + n \qquad (6.2.7)$$

Proof. The statement holds for $j = 0$ by the previous lemma (there is no U-dependence at all). We show now that if it holds for some j then it does for $j + 1 (\leq 5)$. We have by (6.2.2):

$$H^k \mathscr{T}^{j+1} \mathscr{H}^k(1) = -U_x P_{jk} + \partial_t P_{jk} - U \partial_x P_{jk}.$$

We see that as the result of one application of $\mathscr{T}$ to $\frac{1}{H^k} P_{jk}$, each monomial of P_{jk} gets acted upon by the operators $-U_x Id$, ∂_t and $-U \partial_x$. Each of these operations preserves the property (6.2.7) and the resulting expression $P_{j+1,k}$ is a polynomial in $U^{p,q}$, $W^{r,s}$ and $p, q, r, 2$ are in the range

$$p + q \leq j + 1, \quad p \leq j, \quad r + s \leq j + 1 + k, \quad r \leq j + 1. \qquad \square$$

Lemma 6.2.2 and the estimates on U, W from the statement of Theorem 6.2 complete the proof of the latter. $\square$

6.3. Proof of Theorem 6.3. Differentiating $I_0 = 4\sqrt{2} \int\limits_0^{x_+(H,t)} \sqrt{H - V(\xi, t)} d\xi$ we obtain, using (A3.2):

$$\frac{\partial I_0}{\partial H} = 4\sqrt{2} \int\limits_0^{x_+(H,t)} \mathscr{H}(1) \sqrt{H - V(\xi, t)} d\xi. \qquad (6.3.1)$$

---

[4] $U^{(p,q)}$ denotes $\partial_t^p \partial_x^q U$

where $\mathscr{H}(1) = \dfrac{1}{H}\left(\dfrac{1}{2} + W_x\right)$, and since $W_x > -\dfrac{1}{2} + C_1$, we obtain $\dfrac{\partial I_0}{\partial H} \geq$

$C_1 H^{-1} I_0$. To estimate the second derivative we could use A3.2 again: $\dfrac{\partial^2 I_0}{\partial H^2} =$

$4\sqrt{2}\int \mathscr{H}^2(1)\sqrt{H - V}\,d\xi$, where

$$\mathscr{H}^2(1) = \dfrac{1}{H^2}\left[\left(-\dfrac{1}{2} + W_x\right)\left(\dfrac{1}{2} + W_x\right) + W W_{xx}\right],$$

and to estimate the second derivative it is sufficient to have the lower bound on $\mathscr{H}^2(1)$; however it is not necessary: the bound can be obtained (as the theorem states) with only one derivative of W by using the identity (A3.14):

$$\int_0^{x_+(H)} \mathscr{H}^2(1)\sqrt{H - V}\,d\xi = \dfrac{1}{H}\int_0^{x(H)}\left(-\dfrac{1}{2} + W_x\right)d\xi, \qquad (6.3.2)$$

and

$$\dfrac{\partial I_0}{\partial H} = 2\sqrt{2}\int_0^{x_+(H)} \dbar\xi, \qquad (6.3.3)$$

where $\dbar\xi \equiv \dfrac{d\xi}{\sqrt{H - V}}$.

Since $W_x < \dfrac{1}{2} - a$ we have, using (6.3.2) and (6.3.3):

$$\dfrac{\partial^2 I_0}{\partial H^2} = \dfrac{4\sqrt{2}}{H}\int\left(-\dfrac{1}{2} + W_x\right)d\xi < \dfrac{4\sqrt{2}a}{H}\int \dbar\xi = 2a\,\dfrac{\partial I_0}{\partial H}. \qquad \Box$$

Appendix

A1. An Inverse Function Lemma. As elsewhere in this paper, $c > 0$ denotes a constant and different constants are denoted by the same letter c unless their value is of more than transient interest.

Lemma A1.1. *If a real function f of two real variables x, t (t viewed as a parameter) satisfies for some $c > 0$ and $N \in \mathbb{N}$:*

$$|\partial_x^k \partial_t^i f(x, t)| \leq cx^{-k}f(x, t) \qquad (A1.1)$$

for all $x > 0$ large enough and for all k, $i : k + i \leq N$ and if, moreover,

$$\partial_x f(x, t) \geq cx^{-1}f(x, t) > 0 \qquad (A1.2)$$

for all $x > 0$ large enough, then the inverse function $g(y, t)$ of f in x satisfies

$$|\partial_y^k \partial_t^i g(y, t)| \leq cy^{-k}g(y, t), \qquad (A1.3)$$

for all $k + i \leq N$ and for all y large enough.

Proof. We proceed by induction in N. For $i + k = 1$ (A1.3) is easy to check: first, for $i = 0$, $k = 1$ we have $g_y(y, t) = \dfrac{1}{f_x(x, t)}$ and (A1.2) gives

$$|g_y(y, t)| \le \frac{cx}{f(x, t)} = cy^{-1} g(y, t),\tag{A1.4}$$

as desired. Second, for $i = 1$, $k = 0$ the definition $g(f(x, t), t) = x$ gives $g_t(y, t) = -g_y(y, t)f_t(x, t)$, where $x = g(y, t)$, and the desired estimate follows from (A1.4) and (A1.1):

$$|g_t(y, t)| \le cy^{-1} g(y, t) \cdot cf(x, t) = cg(y, t).$$

To carry out the induction step we show that if the implication (A1.1) & (A1.2) $\Rightarrow$(A1.3) holds for $i + k \le N - 1 (N \ge 2)$ then it holds for $i + k = N$ as well. To that end we derive a formula for $\partial_t^k \partial_t^i g(y, t)$.

Differentiating $k > 0$ times the definition $f(f(x, t), t) = x$ by x and expressing the highest derivative, we obtain with $y \equiv f(x, t)$:

$$\partial_y^k g(y, t) = \frac{1}{(\partial_x f(x, t))^k}$$
$$\times \sum_{1 \le r < k} \sum_{\substack{m_j \ge 1 \\ m_1 + \ldots + m_r = k}} a_{r, m}^k (\partial_y^r g(y, t)) \partial_x^{m_1} f(x, t) \cdot \ldots \cdot \partial_x^{m_r} f(x, t),\tag{A1.5}$$

[cf. (4.9)] where $a_{r, m}^k \equiv a_{r, m_1, \ldots, m_r}^k$ are integers. Differentiation of (A1.5) by t is ambiguous unless we specify which variable is independent; let it be y; then $x \equiv x(y, t) = g(y, t)$ and the application of ∂_t^i to (A1.5) gives

$$\partial_t^i \partial_y^k g(y, t) = \sum a_{r, m}^k \sum_{p + q + s = i} \binom{pqs}{i} \left(\partial_t^p \frac{1}{(\partial_x f(g(y, t), t))^k} \right)$$
$$\times (\partial_t^q \partial_y^r g) \left(\partial_t^s \prod_{1 \le j \le r} \partial_x^{m_j} f(g(y, t), t) \right),\tag{A1.6}$$

Here we used the Leibnitz differentiation formula for the ith derivative of the product of three functions. For the case $k = 0$ we have a separate formula, by differentiating $g_t(y, t) = -g_y(y, t)f_t(x, t)$:

$$\partial_t^i g(y, t) = -\sum_{r=0}^{i-1} \binom{r}{i-1} (\partial_t^r \partial_y g(y, t)) \left(\left(\frac{d}{dt} \right)^{i-1-r} f_2(g(y, t), t) \right),\tag{A1.7}$$

where $\dfrac{d}{dt}$ denotes the total derivative and the subscript 2 indicates the differentiation with respect to the second argument.

We activate now the inductive assumption: (A1.1) & (A1.2)$\Rightarrow$(A1.3) for $i + k \le N - 1$. Assume that (A1.1) holds for $i + k = N$ and prove (A1.3) for $i + k = N$. First let $k > 0$, so that (A1.6) is applicable.

First, we have $p = i - q - s = N - k - q - s < N$, and

$$\left| \partial_t^p \frac{1}{f_x(g(y, t), t)^k} \right| < cy^{-k} g(y, t)^k \tag{A1.8}$$

by induction on p. Next, $q + r = i - p - s + r < i - p - s + k \le N$, and thus

$$|\partial_t^q \partial_y^r g(y, t)| \le cy^{-r} g(y, t) \tag{A1.9}$$

by the inductive assumption on g.

Finally, $s + m_j \le i + k = N$, and

$$\left| \partial_t^s \prod_{1 \le j \le r} \partial_x^{m_j} f(g(y, t), t) \right| \le cg(y, t)^{(-\Sigma m_j)} \prod_{1 \le j \le r} f(g(y, t), t)$$

$$= cg(y, t)^{-k} y^r, \tag{A1.10}$$

using (A1.1) and the inductive assumption on g.

Using now (A1.8)–(A1.10) in (A1.6) we obtain (A1.3) with $i + k = N$, as desired. It remains only to consider the case $k = 0$; this is done in a similar way using (A1.7). $\square$

A2. Estimates on $\mathscr{L}, \mathscr{M}, L, M$. The following expressions $L, M, \mathscr{L}$ and $\mathscr{M}$ arise in the process of differentiation of the action-angle transformation and are estimated in this section. These expressions are given by the formulas (A3.6–9) below.

Lemma A2.1. *If the potential function $V(x, t)$ satisfies $|x^q \partial_x^q \partial_t^\tau V| \le cV^{1+\mu}$,*

$$|\partial_t^\tau \partial_x^j W| \le c|x|^{1-j}, \quad \tau + j \le 5, \tag{A2.1}$$

and

$$|\partial_t^\tau \partial_x^j U| \le c|x|^{1-j}, \quad \tau + j \le 5, \tag{A2.2}$$

where $W = V/V_x$, $U = V_t/V_x$, then there exists a constant $\nu < 20\mu$ such that for all $\tau + i + l \le 3$ we have

$$|\partial_t^\tau \partial_I^i \mathscr{L}^l(L)| \le cI^{-i-l-1} H^\nu, \tag{A2.3}$$

$$|\partial_t^\tau \partial_I^i \mathscr{M}^m(M)| \le cI^{-i} H^\nu, \tag{A2.4}$$

$$|\partial_t^\tau \mathscr{M}^m \mathscr{L}^l(L)| \le cI^{-l-1} H^\nu, \tag{A2.5}$$

and

$$\mathscr{L}^4(L) \le cI^{-5}, \tag{A2.6}$$

$$\mathscr{M}^4(M) \le c, \tag{A2.7}$$

where $L, M, \mathscr{L}$, and $\mathscr{M}$ are given by the formulas (A3.6–9) below.

Lemma A2.2. *Under the asumptions of Lemma A2.1, the kernel $K(x, I, t) = H_t(I, t) - V_t(x, t)$ satisfies for some $\nu < 20\mu$:*

$$|\partial_t^\tau \partial_I^i \mathscr{M}^m \mathscr{L}^l(K)| \le cI^{-i-l} H^{1+\nu}, \quad \tau + i + m + l \le 4,$$

and

$$|\mathscr{M}^m \mathscr{L}^l(K)| \le cI^{-l} H^{1+\nu}, \quad m + l = 5.$$

Proof of Lemma A2.1. We start with the proof of A2.6; the proof of A2.7 is virtually identical and is omitted.

The main ingredient in the proof is a tractable expression for the composition $\mathscr{L}^4 = \mathscr{L} \circ \ldots \circ \mathscr{L}$, which we obtain by first writing $\mathscr{L}$ as the sum $\mathscr{L} = \mathscr{L}_1 + \mathscr{L}_2 + \mathscr{L}_3$ of three linear operators acting on $f = f(x, t)$ via

$$\mathscr{L}_1(f) = h(Wf)_x, \quad \mathscr{L}_2(f) = -\frac{h}{2}f, \quad \mathscr{L}_3 f = f_I,$$

where $h = H_I/H$.

We have

$$\mathscr{L}^4 = (\mathscr{L}_1 + \mathscr{L}_2 + \mathscr{L}_3)^4 = \sum_{i_j = 1, 2, 3} \mathscr{L}_{i_1} \circ \mathscr{L}_{i_2} \circ \mathscr{L}_{i_3} \circ \mathscr{L}_{i_4},$$

and proceed now to obtain an expression for the compositions of the elementary operators $\mathscr{L}_{i_j}$. Using the commutation relations $[\mathscr{L}_1, \mathscr{L}_2] = 0$, $[\mathscr{L}_1, \mathscr{L}_3] = -(h_I/h)\mathscr{L}_1$, and $[\mathscr{L}_2, \mathscr{L}_3] = -(h_I/2h)\mathscr{L}_2$, we obtain

$$\mathscr{L}_{i_1} \circ \mathscr{L}_{i_2} \circ \mathscr{L}_{i_3} \circ \mathscr{L}_{i_4} = \mathscr{L}_1^p \mathscr{L}_2^q \mathscr{L}_3^r + \sum_{\substack{p'+q'+r'<p+q+r, \\ s \leq 3}} a_{p'q'r'} h^{(s)} \mathscr{L}_1^{p'} \mathscr{L}_2^{q'} \mathscr{L}_3^{r'},$$

where p is the number of times $\mathscr{L}_1$ occurs in the composition, similarly for q and r, where $h^{(s)} = \partial_I^s h$ and where $a_{p'q'r'}$ are integers.

Using induction, one shows that

$$\mathscr{L}_i^p f = h^p \sum_{i_1 + \ldots + i_p + j = p} A_{pij} \partial_x^{i_1} W \cdot \ldots \cdot \partial_x^{i_p} W \partial_x^j f,$$

where A_{pij} are integers.

The proof of (A2.6) now follows at once from the estimate (A2.1) on W and Theorem (6.2) according to which $\partial_I^s h \leq cI^{-s-1}$, $s = 1, 2, 3, 4$.

The remaining inequalities in the Lemma are proven in a similar way by using in addition, when necessary, Lemma A4.1.

Proof of Lemma A2.2 is completely analogous and we omit it. $\square$

A3. Differentiation Formulas and Identities
Derivatives of the Action $I(t, H)$. For simplicity we assume further that $V(x, t) = V(-x, t)$; furthermore, there is no essential loss of generality in assuming that $V(x, t)$ satisfies $V(0, t) = V_x(0, t) = 0$, $V_{xx}(x, t) > 0$. For any $H > 0$ we define $x_+(H, t) > 0$ by $V(x, t) = H$. Let $K = K(x, t, H)$ be a real function continuously differentiable in x, t, and H. Define

$$\mathscr{I}(t, H) = \int_0^{x_+(H,t)} K(\xi, t, H)\sqrt{H - V(\xi, t)}\,d\xi. \tag{A3.1}$$

Then

$$\frac{\partial}{\partial H}\mathscr{I}(t, H) = \frac{1}{H}\int_0^{x_+(H,t)}\left[HK_H + \frac{1}{2}K + \left(K\frac{V}{V_x}\right)_x\right]\sqrt{H - V}\,d\xi$$

$$\equiv \int \mathscr{H}(K)\sqrt{H - V}\,d\xi \tag{A3.2}$$

and

$$
\frac{\partial}{\partial t} \mathscr{I}(t, H) = \int_0^{x_+(H,t)} \left[K_t - \left(K \frac{V_t}{V_x} \right)_x \right] \sqrt{H - V} \, d\xi
$$

$$
\equiv \int \mathscr{T}(K) \sqrt{H - V} \, d\xi . \tag{A3.3}
$$

When $K = 1$, we obtain the action integral: $(1/4\sqrt{2}) \mathscr{I}(t, H) = I_0(t, H)$.

Derivatives of the Period $T(H, t) = \dfrac{4}{\sqrt{2}} \displaystyle\int_0^{x_+(H,t)} \dfrac{d\xi}{\sqrt{H - V(\xi, t)}}$ *are given by*

$$
\frac{\sqrt{2}}{4} \frac{\partial}{\partial H} T(H, t) = \frac{1}{2H} \int_0^{x_+(H,t)} \left(1 - 2 \frac{V V_{\xi\xi}}{V_\xi^2} \right) \frac{d\xi}{\sqrt{H - V(\xi, t)}} \tag{A3.4}
$$

and

$$
\frac{\sqrt{2}}{4} \frac{\partial}{\partial T} T(H, t) = \int_0^{x_+(H,t)} \left(\frac{V V_{\xi\xi}}{V_\xi^2} \frac{V_t}{V} - \frac{V_{\xi t}}{V_\xi} \right) \frac{d\xi}{\sqrt{H - V(\xi, t)}} . \tag{A3.5}
$$

Derivatives of the Action-Angle Map $(x, y) \mapsto (\theta, I)$: *(A3.10) and (A3.11).* Let $(x(\theta, I, t), y(\theta, I, t))$ be the point in the xy-lane corresponding to the action-angle variables I, θ as described in Sect. 2.

The following expressions L, M and the operators $\mathscr{L}$, $\mathscr{M}$ will play a key role:

$$
L = L(x, t, I) = -\frac{H_I I}{H_I} + \frac{H_I}{H} \left(W_x - \frac{1}{2} \right)
$$

$$
= -\frac{H_{II}}{H_I} - \frac{H_I}{2H} \left(1 - 2 \frac{V V_{xx}}{V_x^2} \right), \tag{A3.6}
$$

$$
M = M(x, t, I) = -\frac{H_{It}}{H_I} - \frac{H_t}{2H} \left(1 - 2 \frac{V V_{xx}}{V_x^2} \right) + \frac{V_{xt}}{V_x} . \tag{A3.7}
$$

The linear differential operators $\mathscr{L}$ and $\mathscr{M}$ act on functions $f(x, t, I)$ according to

$$
\mathscr{L}(f) = \frac{H_I}{2H} \left(1 - 2 \frac{V V_{xx}}{V_x^2} \right) f + \frac{H_I}{H} \frac{V}{V_x} f_x + f_I
$$

$$
\equiv \frac{H_I}{H} \left((f W)_x - \frac{1}{2} f \right) + f_I , \tag{A3.8}
$$

and

$$
\mathscr{M}(f) = \frac{H_t}{2H} \left(1 - 2 \frac{V V_{xx}}{V_x^2} \right) f - \frac{V_{xt}}{V_x} f + \frac{H_t}{H} \frac{V}{V_x} f_x + f_t
$$

$$
\equiv \frac{H_t}{H} \left((f W)_x - \frac{1}{2} f \right) + f_t - \frac{V_{xt}}{V_x} f . \tag{A3.9}
$$

We point out the close similarity between $\mathscr{L}$ and $\mathscr{M}$ – in fact, one obtains $\mathscr{L}$ by replacing the t-differentiations in $\mathscr{M}$ by the I-differentiations.

Derivatives of the action-angle map are given by $\left(\text{as usual, } W = \dfrac{V}{V_x} \text{ and } U = \dfrac{V_t}{V_x}\right)$:

$$\frac{\partial}{\partial I} x(\theta, I, t) = \sqrt{H - V} \int L \frac{d\xi}{\sqrt{H - V}} + \frac{H_I}{H} W, \tag{A3.10}$$

and

$$\frac{\partial}{\partial t} x(\theta, I, t) = \sqrt{H - V} \int M \frac{d\xi}{\sqrt{H - V}} + \frac{H_t}{H} W - U. \tag{A3.11}$$

One could easily obtain similar expressions for $\partial_t y$, $\partial_I y$.

Derivatives of Singular Integrals. Let again $K = K(x, t, I)$ be a smooth function of three variables.

We have the key identities:

$$\frac{\partial}{\partial I} \int_0^{x(\theta, I, t)} K(\xi, t, I) \frac{d\xi}{\sqrt{H(I, t) - V(\xi, t)}} = L \int K + \int \mathscr{L}(K), \tag{A3.12}$$

where $L = L(x, I, t)$, is given by (A3.6), $\int K = \int_0^x K(\xi, I, t)\,đ\xi$, with $đ = d\xi / \sqrt{H - V}$, $\mathscr{L}$ given by (A3.8), and $\int \mathscr{L}(K) = \int_0^x \mathscr{L}(K)\,(\xi, I, t)\,đ\xi$. Similarly for the t-derivative:

$$\frac{\partial}{\partial t} \int K = K \int M + \int \mathscr{M}(K), \tag{A3.13}$$

with the abbreviations as above.

Some Interesting Identities. Under the above assumptions on V, we have (suppressing the t-dependence which plays no role here):

$$\int_0^{x_+(H)} \left(\left(W_x + \frac{1}{2}\right)\left(W_x - \frac{1}{2}\right) + W W_x \right) \sqrt{H - V}\,d\xi$$

$$= H \int_0^{x_+(H)} \left(W_x - \frac{1}{2}\right) \frac{d\xi}{\sqrt{H - V}}, \tag{A3.14}$$

where $W = \dfrac{V}{V_x}$,

$$H \int_0^{x_+(H)} \frac{d\xi}{\sqrt{H - V(\xi)}} = \sqrt{2} \int_0^{x_+(H)} \left(\frac{3}{2} - \frac{V V_{xx}}{V_x^2} \right) \sqrt{H - V}\,d\xi \tag{A3.15}$$

and

$$H \int_0^x \frac{d\xi}{(H - V(\xi, t))^{3/2}} = - \int_0^x \left(-\frac{1}{2} + W_\xi \right) \frac{d\xi}{\sqrt{H - V}}$$

$$+ 2W \frac{1}{\sqrt{H - V(x)}}. \tag{A3.16}$$

It should be noted that $x \in (0, x_+(H))$ in the last identity is arbitrary, while in the previous two identities the upper limit of integration is $x = x_+$.

An infinite hierarchy of further identities of this kind can be obtained using the method of proof outlined below.

The last identity can be used to study the blowing up of the integral on the left-hand side of (A3.16) as $x \uparrow x_+(H)$.

Proofs

Proof of (A3.2). Let us choose $\dfrac{V(\xi, t)}{H} = \sigma \in [0, 1]$ as the new variable of integration; (A3.1) becomes

$$I(t, H) = H\sqrt{H} \int\limits_0^1 \frac{K(\xi(\delta, t, H), t, H)}{V_x(\xi(\sigma, t, H), t)} \sqrt{1 - \sigma}\, d\sigma \; ; \qquad (A3.17)$$

differentiating by H, observing that $\dfrac{\partial}{\partial H} \xi(\sigma, t, H) = \dfrac{\sigma}{V_x}$ and simplifying we obtain (A3.2).   □

Proof of (A3.3). Applying $\dfrac{\partial}{\partial t}$ to (A3.17), we use $\dfrac{\partial}{\partial t} \xi(\sigma, t, H) = -\dfrac{V_t}{V_x}$.   □

Proof of (A3.10)–(A3.12).

$$\partial_t \int\limits_0^{x(\theta, I, t)} K(I, t, \xi) \frac{d\xi}{\sqrt{H - V}} = \frac{K(I, t, x)}{\sqrt{H - V}} \partial_t x + \int\limits_0^x \partial_t \frac{K(I, t, \xi)d\xi}{\sqrt{H(I, t) - V(\xi, t)}} . \quad (A3.18)$$

To find the expression for $\partial_t x(\theta, I, t)$ we differentiate the identity $H_I \int\limits_0^x$

$$\times \frac{d\xi}{\sqrt{2(H - V)}} = \theta, \text{ where } T = \frac{1}{\partial_I H(I, t)} \text{ is the period of the frozen system.}$$

We obtain

$$H_{It} \int\limits_0^x \frac{d\xi}{\sqrt{H - V}} + \frac{H_I}{\sqrt{H - V}} x_t + H_I \int\limits_0^x \partial_t \frac{d\xi}{\sqrt{H(I, t) - V(\xi, t)}} = 0,$$

and

$$\frac{x_t}{\sqrt{H - V}} = -\frac{H_{It}}{H_I} \int\limits_0^x d\xi - \int\limits_0^x \partial_t d\xi . \qquad (A3.19)$$

Using the differentiation formula (A3.20) below for singular integrals, we obtain for the last term in (A3.18),

$$\int\limits_0^x \partial_t(K d\xi) = \frac{H_t}{2H} \int\limits_0^x K d\xi + \frac{H_t}{H} \int\limits_0^x \left(\frac{K}{V_x}\right)_x V d\xi$$

$$+ \int\limits_0^x \left(\frac{K}{V_x}\right)_t V_x d\xi + H \left(\frac{V}{H}\right)_t \frac{K}{V_x} \frac{1}{\sqrt{H - V}} \Bigg|_{\xi = x} ;$$

we used the assumption $V(0, t) = V_x(0, t) = 0$. Setting $K \equiv 1$ we obtain the expression for the last term in (A3.19) as well:

$$\int_0^x \partial_t d\xi = \frac{H_t}{2H} \int d\xi + \frac{H_t}{H} \int_0^x \left(\frac{1}{V_x} \right)_x V d\xi$$

$$+ \int_0^x \left(\frac{1}{V_x} \right)_t V_x d\xi + H \left(\frac{V}{H} \right)_t \frac{1}{V_x} \frac{1}{\sqrt{H - V}} \bigg|_{\xi = x}.$$

Substitution of the last two expressions into (A3.18) leads, mercifully, to the cancellation of the boundary terms, resulting in (A3.12). The remaining identities are proven similarly. □

An Auxiliary Formula. Let $A(\lambda, x)$, $B(\lambda)$, $W(\lambda, x)$ be real functions of independent real variables x and λ such that the expressions below are well defined. Then

$$\frac{\partial}{\partial \lambda} \int_0^x \frac{A(\lambda, \xi) d\xi}{\sqrt{B(\lambda) - W(\lambda, \xi)}} = \int_0^x \left[\frac{B_\lambda}{2B} \left(A + 2 \left(\frac{A}{W_\xi} \right)_\xi W \right) + \left(\frac{A}{W_\xi} \right)_\lambda W_\xi \right]$$

$$\times \frac{d\xi}{\sqrt{B - W}} + B \left(\frac{W}{B} \right)_\lambda \frac{A}{W_x} \frac{1}{\sqrt{B - V}} \bigg|_{\xi = 0}. \quad \text{(A3.20)}$$

Proof goes by first choosing the new variable $\sigma = \dfrac{W(\xi)}{B}$ in the integral and then differentiating the integral. □

A4. Estimates on the Derivatives of the Action-Angle Map: $\partial_t^\tau \partial_I^i x(\theta, I, t)$. In this section, we use the above differentiation formulas together with the assumptions on the potential function V to obtain the desired estimates on the derivatives in terms of the action-angle variables.

Lemma A4.1. *Assume that the functions* $W = \dfrac{V}{V_x}$, $U = \dfrac{V_t}{V_x}$ *satisfy the bounds*

$$|\partial_x^p \partial_t^\tau W|, \quad |\partial_x^p \partial_t^\tau U| \le c x^{1-p}, \quad p + \tau \le 5,$$

$$|x^q \partial_x^q \partial_t^\tau V| \le c V^{\mu+1}, \quad p + \tau \le 5;$$

then

$$|\partial_t^\tau \partial_I^i x(\theta, I, t)| \le c I^{-i} x(\theta, I, t) H^\nu, \quad i + \tau \le 4,$$

where $\nu = \nu(\mu) < 20\mu$ *is a constant[5].*

Proof. 1a. The formula for x_I: We recall (A3.10):

$$x_I = \sqrt{H - V} \int_0^x L \frac{d\xi}{\sqrt{H - V}} + h(I, t) W(x, t), \quad \text{(A4.1)}$$

where $h = H_I/H$, $W = V/V_x$.

---

[5] Whose value can be estimated much more precisely. For $\tau + i \le 2$, for instance, the estimate holds with $\nu = 0$. We sacrifice this precision to avoid further complications in the exposition

1b. The estimate of x_I: For $0 \leq \xi \leq x$ we have $V(\xi, t) \leq V(x, t)$, and thus $\dfrac{\sqrt{H - V(x)}}{\sqrt{H - V(\xi)}} \leq 1$. Now, $|L| \leq cI^{-1}$ using the assumption on W which for $p = 1$ gives $\dfrac{V V_{xx}}{V_x^2} \leq C$. Using this in (A4.1) we obtain

$$x_I \leq cI^{-1}x.$$

2a. The formula for x_{II}: We first differentiate $\sqrt{H - V}$ in (A4.1) and obtain using (A4.1):

$$\frac{\partial}{\partial I} \sqrt{H - V} = h\sqrt{H - V} - \frac{1}{2} V_x \int_0^x L\, d\xi, \qquad (A4.2)$$

where $d\xi = \dfrac{d\xi}{\sqrt{H - V(\xi)}}$. Using this and the differentiation formula (A3.12) we obtain

$$x_{II} = \left(h\sqrt{H - V} - \frac{1}{2} V_x \int_0^x L\, d\xi \right) \int_0^x L\, d\xi + \sqrt{H - V}\, L \int_0^x L\, d\xi$$

$$+ \sqrt{H - V} \int_0^x \mathscr{L}(L)\, d\xi + h_I W + h W_x x_I. \qquad (A4.3)$$

2b. Estimating x_{II}: It suffices to prove that

$$\left| V_x \int_0^x L\, d\xi \right| \leq cI^{-1}\sqrt{H - V} \qquad (A4.4)$$

[for $0 \leq x \leq x_+(I, t)$] – indeed, all the remaining terms in (A4.3) are estimated by $cI^{-2}x$ in the same way as in 1b – one only needs to use the bound $|\mathscr{L}(L)| \leq cI^{-2}$ (Lemma A2.1) for the fourth term in (A4.3) and $W_x \leq c$ in the last term in (A4.3), together with the last estimate on x_I. Now, (A4.4) is equivalent to

$$-B(x) \equiv -cI^{-1} \frac{1}{V_x} \sqrt{H - V} \leq A(x)$$

$$\equiv \int_0^x L\, d\xi \leq cI^{-1} \frac{1}{V_x} \sqrt{H - V} \equiv B(x), \qquad (A4.5)$$

which holds for $x = x_+$: $A(x_+) = B(x_+) = 0$, and it remains to prove that $-B'(x) \geq A'(x) \geq B'(x)$ for $0 \leq x \leq x_+$.

After multiplying by $\sqrt{H - V}$ both sides in $-B' \geq A' \geq B'$ these inequalities reduce to

$$cI^{-1} \left[\frac{V_{xx}}{V_x^2} (H - V) + \frac{1}{2} \right] \geq L(I, x, t) \geq cI^{-1} \left[-\frac{V_{xx}}{V_x^2} (H - V) - \frac{1}{2} \right],$$

here $V = V(x, t)$. Since $\dfrac{V_{xx}}{V_x^2} (H - V) \geq 0$, this holds, if we have

$$\frac{1}{2} cI^{-1} \geq L \geq -\frac{1}{2} cI^{-1},$$

which we do if c is fixed at a large enough value.

3a. The formula: differentiating (A4.3) we obtain

$$x_{III} = \partial_I \left[\left(h\sqrt{H-V} - V_x \int_0^x L d\xi \right) \int_0^x L d\xi \right] \tag{A4.6a}$$

$$+ \partial_I \left[\sqrt{H-V} L \int_0^x {_I} d\xi \right] + \partial_I \left(\sqrt{H-V} \int_0^x \mathscr{L}(L) d\xi \right) \tag{A4.6b}$$

$$+ \partial_I (h_I W) + \partial_I (h W_x x_I). \tag{A4.6c}$$

3b. The estimate: We show that (A4.6a) is bounded by $cI^{-3} x H^\nu$; indeed, $|\partial_I h| \leq cI^{-2}$, $|\partial_I \sqrt{H-V}|^{(A4.2\,\&\,4)} \leq cI^{-1} \sqrt{H-V}$, $\left| \partial_I \left(\int_0^x L d\xi \right) \right| \leq cI^{-2} \left| \int_0^x d\xi \right|$ (checking the latter, $\left| \partial_I \int_0^x L d\xi \right| = \left| L \int_0^x L d\xi + \int_0^x \mathscr{L}(L) d\xi \right| \leq cI^{-2} \int_0^x d\xi$), and it remains to show that $\left| V_{xx} x_I \int_0^x L d\xi \right| \leq cI^{-2} \sqrt{H-V} H^\nu$. To that end, we write

$$V_{xx} x_I \int_0^x L d\xi \leq c V_{xx} x I^{-1} \int_0^x L d\xi = c \frac{x V_{xx}}{V_x} \left(V_x \int_0^x L d\xi \right) I^{-1}$$

$$\leq c_1 I^{-1} H^\mu V_x \int_0^x L d\xi \leq c_1 I^{-2} H^\mu \sqrt{H-V}.$$

Proving the desired bound on (A4.6b) and (A4.6c) is straightforward; in addition to the estimates mentioned above we use $|\partial_I L| \leq cI^{-2}$, $\left| \partial_I \int \mathscr{L}(L) d\xi \right| \leq cI^{-3} \int d\xi$ [Indeed, $\left| \partial_I \int_0^x \mathscr{L}(L) d\xi \right| = \left| \mathscr{L}(L) \int_0^x L d\xi + \int_0^x \mathscr{L}^2(L) d\xi \right| \leq cI^{-3} \int_0^x d\xi$ by the differentiation formula and the estimates on $\mathscr{L}(L)$ and $\mathscr{L}^2(L)$ from Lemma A2.1], $|h_{II}| < cI^{-2} H$ (Theorem 6.2), and $W_{xx} \leq cx^{-1}$. Further estimates proceed in the same way and pose no new difficulties. $\square$

References

[AB]   Abel, N.H.: Resolution d'un probleme mechanique. J. Reine Angew. Math. **1**, 13–18 (1826)

[A]   Arnol'd, V.I.: On the behavior of an adiabatic invariant under a slow periodic change of the Hamiltonian. DAN **142**, (4) 758–761 (1962), (Transl. Sov. Math. Dokl. **3**, 136–139)

[A1]   Arnol'd, V.I.: Mathematical methods of classical mechanics. Berlin, Heidelberg, New York: Springer 1978

[AD]   Aubry, S., LeDaeron, P.Y.: The discrete Frenkel-Kontorova model and its extensions. I. Exact results for the ground states. Physica **8** D, 381–422 (1983)

[BE]   Bernstein, D.: Birkhoff periodic orbits for twist maps with the graph intersection property.
         Erg. Th. Dyn. Sys. **5**, 531–537 (1985)
[B1]   Birkhoff, G.D.: Surface transformations and their dynamical properties. Acta Math. **43**,
         1–119 (1920)
[B2]   Birkhoff, G.D.: Sur quelques courbes fermees remarquables. Bull. Soc. Math. France **60**,
         1–26 (1932)
[BH]   Boyland, P.L., Hall, G.R.: Invariant circles and the order structure of periodic orbits in
         monotone twist maps. Topology **26**, 21–35 (1987)
[CZ]   Chiercia, L., Zehnder, E.: On asymptotic expansion of quasiperiodic solutions. Annali della
         Sc. Norm. Sup. di Pisa Serie IV Vol. XVI, 245–258 (1989)
[DZ1]   Dieckerhoff, R., Zehnder, E.: An "a priori" estimate for oscillatory equation. Dyn. systems
         and bifurcations. Groningen, 1984. Lecture Notes in Mathematics, Vol. 1125, pp. 9–14.
         Berlin, Heidelberg, New York: Springer 1985
[DZ2]   Dieckerhoff, R., Zehnder, E.: Boundedness of solutions via the Twist theorem. Ann. Scuola
         Norm. Sup. Pisa **14**(1), 79–85 (1987)
[DO]   Douady, R.: Docteur D'etat Thesis, Ecole Polytechnique
[D1]   Tongren Ding: Boundedness of solutions of Duffing's equation. JDE **61**, 178–207 (1986)
[D2]   Tongren Ding: An answer to Littlewood's problem on boundedness for super-linear Duff-
         ing's equations. JDE **73**, 269–287 (1988)
[HA]   Hall, G.R.: A topological version of a theorem of Mather on twist maps. Erg. Th. Dyn.
         Sys. **4**, 585–603 (1984)
[HE]   Herman, M.R.: Sur les Courbes invariantes par les diffeomorphismes de l'anneau, Vol. 1
         Asterisque, 103–104, 1983, and Vol. 2, Asterisque **144**, 1986
[KA]   Katok, A.: Some remarks on Birkhoff and Mather twist map theorems. Erg. Th. Dyn. Sys.
         2, 185–194 (1982)
[K]   Keller, J.B.: Inverse problems. The American Mathematical Monthly **83**, 107–118 (1976)
[LL]   Laederich, S., Levi, M.: Invariant curves and time-dependent potentials. Erg. Th. Dyn. Sys.
         (to appear)
[L1]   Levi, M.: KAM Theory for particles in Periodic Potentials. To appear in Erg. Th. Dyn. Sys.
         (1989)
[L2]   Levi, M.: On the Littlewood's counterexample of unbounded motions in superquadratic
         potentials. ETH preprint, October 1988; to appear in Dynamics Reported
[LI]   Littlewood, J.E.: Unbounded solutions of an equation $\ddot{y} + g(y) = p(t)$, with $p(t)$ periodic
         and bounded and $g(y)/y \to \infty$ as $y \to \pm\infty$. J. Lond. Math. Soc. **41**, 497–507 (1966)
[LB]   Liu Bin: On Moser's boundedness problem for solutions of Duffing's equation. Preprint,
         Math. Dept., Peking University, 1987
[LO]   Yiming Long: Unbounded solutions of a superlinear Duffing's equation, Preprint, Nankai
         Institute of Mathematics, Tianjun, China, 1989
[MA1]   Mather, J.N.: Existence of quasi-periodic orbits for twist homeomorphisms of the annulus.
         Topology **21**, 457–476 (1982)
[MA2]   Mather, J.: More Denjoy minimal sets for area-preserving diffeomorphisms. Comment.
         Math. Helv. **60**, 508–557 (1985)
[MO]   Morris, G.R.: A case of boundedness in Littlewood's problem on oscillatory differential
         equations. Bull. Austr. Math. Soc. **14**, 71–93 (1976)
[M1]   Moser, J.K.: On invariant curves of area preserving mappings of an annulus. Nachr. Akad.
         Wiss. Göttingen Math. Phys. Kl. II, 1–20 (1962)
[M2]   Moser, J.K.: A stability theorem for minimal foliations of the torus. Erg. Th. Dyn. Sys. **8**,
         151–188 (1988)
[M3]   Moser, J.K.: Quasi-periodic solutions of nonlinear elliptic partial differential equations. Bol.
         Soc. Bras. Mat., **20**(1), 29–45
[M4]   Moser, J.K.: The scattering problem for some particle systems on the line. Lecture Notes
         Math., vol. 597, pp. 441–463. Berlin, Heidelberg, New York: Springer 1977
[N]   Norris, J.: Boundedness in periodically forced second order conservative system. Preprint,
         Aberystwyth, 1990
[R1]   Rüssman, H.: Über invariante Kurven differenzierbarer Abbildungen eines Kreisringes.
         Nachr. Akad. Wiss. Göttingen, Math. Phys. Kl. II, 67–105 (1970)

[R2]   Rüssman, H.: On the existence of invariant curves of twist mapping of an annulus. In Geometric dynamics. Palis, J. (ed.). Lecture Notes in Mathematics, vol. 1007, pp. 677–718. Berlin, Heidelberg, New York: Springer 1981

[S]   Salamon, D., Zehnder, E.: KAM theory in configuration space. Commun. Math. Helv. **64**, 84–132 (1989)

[S1]   Stoker, J.J.: Nonlinear vibrations in mechanical and electrical systems. New York: Interscience 1950

[SW]   Smoller, J., Wasserman, A.: Global bifurcations of steady state solutions. J. Diff. Eq. **39**, 269–290 (1981), or in the book Smoller, J.: Shock waves and reaction-diffusion equations. New York, Heidelberg, Berlin: Springer 1983

Communicated by M. Herman