Geometric determination of classical actions of heteroclinic and unstable periodic orbits

Jizhou Li and Steven Tomsovic
Department of Physics and Astronomy, Washington State University, Pullman, Washington 99164-2814, USA
(Received 21 March 2017; published 29 June 2017)

Semiclassical sum rules, such as the Gutzwiller trace formula, depend on the properties of periodic, closed, or homoclinic (heteroclinic) orbits. The interferences embedded in such orbit sums are governed by classical action functions and Maslov indices. For chaotic systems, the relative actions of such orbits can be expressed in terms of phase-space areas bounded by segments of stable and unstable manifolds and Moser invariant curves. This also generates direct relations between periodic orbits and homoclinic (heteroclinic) orbit actions. Simpler, explicit approximate expressions following from the exact relations are given with error estimates. They arise from asymptotic scaling of certain bounded phase-space areas. The actions of infinite subsets of periodic orbits are determined by their periods and the locations of the limiting homoclinic points on which they accumulate.

DOI: 10.1103/PhysRevE.95.062224

I. INTRODUCTION

The properties of sets of rare classical orbits can be extremely important in the study of chaotic dynamical systems [1]. For example, classical sum rules over unstable periodic orbits describe various entropies, Lyapunov exponents, escape rates, and the uniformity principle [2]. The information that enters these classical summations are the stability properties and densities. Such sets of orbits are also linked to the properties of the analogous quantized systems through the derivation of semiclassical sum rules. A few cases are given by periodic [3–5] and closed-orbit sum rules [6–8], which determine quantal spectral properties, and homoclinic (heteroclinic) orbit summations [9,10], generating wave-packet propagation approximations. The interferences in such sum rules are governed by the orbits' classical action functions and Maslov indices, and thus this information takes on greater importance in the context of the asymptotic properties of quantum mechanics. Various resummation techniques have been given to work with series, which are often divergent in nature [11–13]. Other studies exploring a fuller understanding of the interferences have also been carried out [14–20]. Our interest in this paper is establishing a framework for understanding the relationships between periodic and homoclinic (heteroclinic) orbit actions and their action correlations.

A periodic orbit in a two-degree-of-freedom system becomes either a single fixed point or an invariant set of points visited periodically in a two-dimensional Poincaré surface of section. It is sufficient to concentrate on symplectic mappings on a plane and study the unstable fixed or periodic points under their application. The fixed points play an advantageous role in this work due to convergence theorems in normal form coordinates. The normal form transformation was first proved by Moser to converge inside a disk-shaped neighborhood of the fixed point denoted by D_0 hereafter [21]. Later, da Silva Ritter *et al*. extended the convergence zone along the stable and unstable manifolds out to infinity [22].

Within the convergence zone are Moser invariant curves, which are images of invariant hyperbolas. As already noted by Birkhoff [21,23,24], the self or mutual intersections between such invariant curves can support periodic orbits with arbitrarily large periods. These periodic orbits accumulate alternatively on one or multiple homoclinic (heteroclinic) points in a homoclinic (heteroclinic) tangle. In the limit of the orbital period going to infinity, the invariant curves become infinitely close to the stable and unstable manifolds of the fixed points. The periodic orbits of this kind are said to be satellite to their respective homoclinic (heteroclinic) points [15,22]. Da Silva Ritter *et al*. developed a method for the numerical computation of satellite orbits supported by such curves in the quadratic map [22]. Therefore, every periodic orbit inside the convergence zone must be satellite to some homoclinic points, with its classical action closely related to that of the homoclinic orbit. Recent work shows that the size of the convergence zone can be quantified in terms of the outermost Moser curves [24], i.e., ones with the largest QP normal form coordinates product, and the convergence zone can be numerically estimated using the outermost Moser curves as boundaries [25].

Assuming a system is fully chaotic, the convergence zone should cover most, if not all, of the accessible phase space. In that case, nearly all of the periodic orbits lie on Moser invariant curves, and each one can be treated as a satellite orbit of some particular set of hyperbolic fixed points. Even if the system is not fully chaotic, the convergence zone can cover nearly all of the available phase space. Figure 4 of Ref. [24] gives an excellent example of the convergence zone covering almost all of the complex region of the homoclinic tangle of the Hénon map [26], avoiding only a small region inside the last KAM curve. Thus, a study of satellite orbits may often encompass nearly all periodic orbits of the system; i.e., satellite orbits are not typically a small subset of the periodic orbits.

In the quantum Baker's map wave-packet autocorrelation functions can equivalently be expressed as a sum over periodic fixed points or homoclinic orbit segments with an exact one-to-one correspondence between terms [27]. Similarly, there is the same, though not exact, correspondence for the stadium billiard as there may be problems with orbits which approach bifurcations points too closely, i.e., some of the orbits that come too close to the joint between the straight edge and curved hard walls [10]. Thus, it is of significant interest to understand how the homoclinic (heteroclinic) and periodic orbits are related.

This work develops a framework for expressing the actions of satellite orbits in terms of the relative actions of homoclinic (heteroclinic) orbits, phase-space areas bounded by stable and

unstable manifolds, and Moser invariant curves. These areas scale down with increasing periods, and the determination of the action of a leading satellite periodic orbit with small period is sufficient to approximate satellite orbits with larger periods; the numerical calculation of individual orbits becomes unnecessary to an excellent approximation. As a final remark, note that Maslov indices can be incorporated into this framework but are not considered in this paper in order to focus on the classical actions. Previous studies of Maslov indices can be found in Refs. [28–30].

This paper is organized as follows. Section II sets the notation and basic definitions of homoclinic (heteroclinic) orbits and their actions. Section III is a generalization of the MacKay-Meiss-Percival action principle [31] for heteroclinic orbits and expresses their actions as phase-space integrals. Section IV concerns relative actions between two hyperbolic fixed points and expresses them as phase-space areas bounded by segments of the stable and unstable manifolds. Section V studies the satellite periodic orbits and expresses their actions using phase-space areas bounded by segments of the Moser invariant curves together with stable and unstable manifolds. An approximation for orbits with large periods is also given, together with a numerical calculation with the Hénon map. Some basic information on homoclinic (heteroclinic) tangles [32–34], the MacKay-Meiss-Percival action principle [31,35], and normal form theory with satellite period orbits [22] can be found in Appendices A and B.

II. HOMOCLINIC (HETEROCLINIC) ORBITS AND RELATIVE ACTIONS

This section lays out the paper's notation and a few basic concepts of homoclinic (heteroclinic) orbits in classical dynamical systems.

A. Homoclinic (heteroclinic) orbits

Let M be an analytic and area-preserving map on the two-dimensional (2D) phase space (q,p), and $x=(q,p)$ be a hyperbolic fixed point under M with stability exponent μ. Denote the unstable and stable manifolds of x by $U(x)$ and $S(x)$, respectively. Typically, its unstable and stable manifolds intersect infinitely many times and form a complicated pattern called a homoclinic tangle [1,33,34] as partially shown in Fig. 1. The intersection points belong to both $U(x)$ and $S(x)$ for all times. The homoclinic orbit, denoted by $\{h_0\}$, is the bi-infinite collection of images:

$$\begin{aligned}\{h_0\} &= \{M^{-\infty}h_0,\cdots,M^{-1}h_0,h_0,Mh_0,\cdots,M^{\infty}h_0\}\\ &= \{h_{-\infty},\cdots,h_{-1},h_0,h_1,\cdots,h_{\infty}\},\end{aligned} \tag{1}$$

where both $h_{-\infty}$ and h_{∞} converge to x. If the unstable and stable segments connecting x with h_0 intersect only at h_0, then $\{h_0\}$ is a primary homoclinic orbit. There must be at least two such orbits [32], such as $\{h_0\}$ and $\{g_0\}$ in Fig. 1. Of particular interest are the unstable segments $U[g_{i-1},g_i]$ and stable segments $S[g_i,g_{i-1}]$, which enclose the so-called "lobe regions" L_i and L'_i, which are extensively studied in transport problems [31,32,34,36]. The region bounded by $U[x,g_0]$ and $S[x,g_0]$ is called the complex region, which is the main region of interest in transport theory. More recent

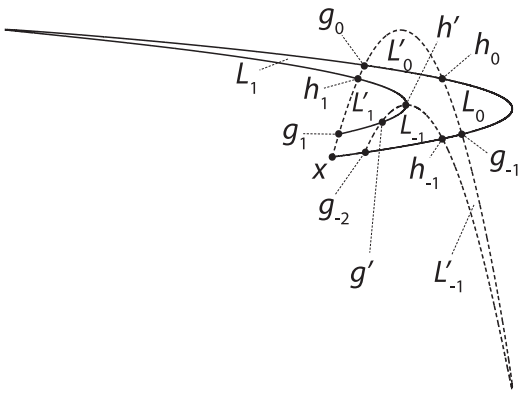

FIG. 1. Example partial homoclinic tangle from the Hénon map [26]. The unstable (stable) manifold is the solid (dashed) curve. There are two primary homoclinic orbits $\{h_0\}$ and $\{g_0\}$. The lobe regions L_0 and L'_0 form a turnstile and govern the transport. In an open system, the lobes L_i (L'_{-i}) may extend out to infinity never to reenter the complex region for $i \geqslant 1$.

works on the topological behavior of lobes resulted in what is termed homotopic lobe dynamics [37–39], which gives rise to fractals in escape time graphs and has been applied to problems such as ionization of hydrogen atoms [40] and escape from a vase-shaped cavity [41,42]. Notice that for open systems such as the Hénon map, any point outside the complex region will escape to inifinity, and thus the lobes L_i and L'_{-i} with $i \geqslant 1$ will extend to infinity and never come back into the complex region. This ensures that there are no homoclinic points on segments $U(g_i,h_{i+1})$ and $S(g_i,h_i)$. The homoclinic points in such systems are distributed only on segments $U[h_i,g_i]$ and $S[h_i,g_{i-1}]$.

A more general scenario is to have two hyperbolic fixed points with their own stable and unstable manifolds intersecting one another, forming a heteroclinic tangle [32]. Consider $x^{(\alpha)}$ and $x^{(\beta)}$, with their unstable [stable] manifolds $U(x^{(\alpha)})$ $[S(x^{(\alpha)})]$ and $U(x^{(\beta)})$ $[S(x^{(\beta)})]$; see Fig. 2. The intersecting stable and unstable manifolds of different fixed points generate heteroclinic orbits. In Fig. 2, $\{h_0^{(\alpha)}\}$ and $\{h_0^{(\beta)}\}$ have the limiting

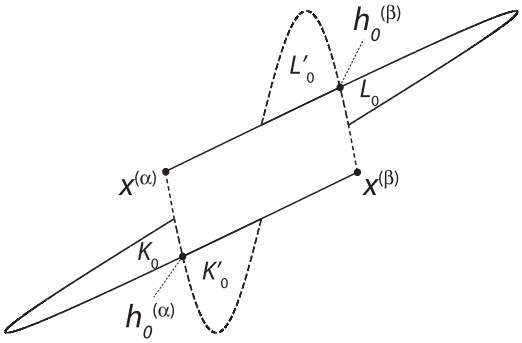

FIG. 2. Schematic partial heteroclinic tangle between $x^{(\alpha)}$ and $x^{(\beta)}$. The unstable (stable) manifolds are plotted in solid (dashed) curves. The "parallelogram" region bounded by segments $U[x^{(\alpha)},h_0^{(\beta)}]$, $S[h_0^{(\beta)},x^{(\beta)}]$, $U[x^{(\beta)},h_0^{(\alpha)}]$, and $S[h_0^{(\alpha)},x^{(\alpha)}]$ is the complex region. The lobes $\{L_0,L'_0\}$ and $\{K_0,K'_0\}$ form the irreducible structures that can be mapped successively to form the entire tangle.

points:

$$\begin{aligned} h_{\infty}^{(\alpha)},\, h_{-\infty}^{(\beta)} &\to x^{(\alpha)}, \\ h_{-\infty}^{(\alpha)},\, h_{\infty}^{(\beta)} &\to x^{(\beta)}. \end{aligned} \tag{2}$$

Unlike homoclinic tangles, there may be only one primary heteroclinic orbit; an example is shown ahead. Homoclinic and heteroclinic orbits play an important role in chaotic dynamics as they provide clues for the entire structure of the chaotic region. As shown in Refs. [22,23], infinite families of satellite periodic orbits accumulate on the homoclinic (heteroclinic) orbits, and the determination of the periodic orbit actions rely on those of the homoclinic (heteroclinic) orbits.

B. Relative actions

The mapping M can be viewed as a canonical transformation that maps a point (q_n, p_n) to (q_{n+1}, p_{n+1}), while preserving the symplectic area, therefore a generating (action) function $F(q_n,q_{n+1})$ can be associated with this process such that [31,35]:

$$\begin{aligned} p_n &= -\partial F/\partial q_n, \\ p_{n+1} &= \partial F/\partial q_{n+1}. \end{aligned} \tag{3}$$

The total action of an orbit $\mathcal{F}$ is the sum of the generating functions,

$$\mathcal{F} = \sum_{n=-\infty}^{\infty} F(q_n,q_{n+1}), \tag{4}$$

and is divergent, in general. However, the MacKay-Meiss-Percival action principle [31,35] can be applied to obtain well-defined action differences for particular pairs of orbits. An important and simple case is the relative action between a fixed point x and any of its homoclinic orbits $\{h_0\}$, which turns out to be equal to an area bounded by unstable and stable manifold segments as

$$\begin{aligned} \Delta\mathcal{F}_{\{h_0\}x} &= \sum_{n=-\infty}^{+\infty} \left[F_{\{h_0\}}(q_n,q_{n+1}) - F_x(q,q)\right] \\ &= \int_{U[x,h_0]} pdq + \int_{S[h_0,x]} pdq = \oint_{US[xh_0]} pdq \\ &= \mathcal{A}^{\circ}_{US[xh_0]}, \end{aligned} \tag{5}$$

where $U[x,h_0]$ is the segment of the unstable manifold from x to h_0, and $S[h_0,x]$ is the segment of the stable manifold from h_0 to x. The $\circ$ superscript from the last line indicates that the area is interior to a path that forms a closed loop, and the subscript indicates the path: $US[xh_0] = U[x,h_0] + S[h_0,x]$. As usual, clockwise enclosure of an area is positive, counterclockwise is negative. $F_{\{h_0\}}(q_n,q_{n+1})$ denotes the generating function along $\{h_0\}$ that maps h_n to h_{n+1}, and $F_x(q,q)$ denotes the generating function of x in one iteration. Likewise, a second important case is for homoclinic orbit pairs, which results in

$$\begin{aligned} \Delta\mathcal{F}_{\{h_0'\}\{h_0\}} &= \sum_{n=-\infty}^{\infty} \left[F_{\{h_0'\}}(q_n,q_{n+1}) - F_{\{h_0\}}(q_n,q_{n+1})\right] \\ &= \int_{U[h_0,h_0']} pdq + \int_{S[h_0',h_0]} pdq = \mathcal{A}^{\circ}_{US[h_0h_0']}, \end{aligned} \tag{6}$$

where $U[h_0,h_0']$ is the segment of the unstable manifold from h_0 to h_0', and $S[h_0',h_0]$ is the segment of the stable manifold from h_0' to h_0. See Appendix A for further details.

It is also desirable to have geometric relations for the differences of any pair of periodic orbits. Since they may not have the same period, comparing each over its primitive period relative to a fixed point suffices. For an l-period orbit, i.e., $M^l(x_0) = x_l = x_0$ and $\{x_0\} = \{x_0,x_1,...x_l\}$,

$$\Delta\mathcal{F}_{\{x_0\}x} = \sum_{n=0}^{l-1} \left[F_{\{x_0\}}(q_n,q_{n+1}) - F_x(q,q)\right]. \tag{7}$$

However, ahead it is shown that the geometric form also requires homoclinic orbits and Moser invariant curves.

III. RELATIVE HETEROCLINIC ORBIT ACTIONS

Consider two hyperbolic fixed points $x^{(\alpha)}$ and $x^{(\beta)}$, and a heteroclinic intersection $h_0^{(\beta)}$; see Fig. 2. Since the infinite past $h_{-\infty}^{(\beta)}$ and the infinite future $h_{\infty}^{(\beta)}$ are asymptotic to different fixed points, it is convenient to consider the heteroclinic orbit in two semi-infinite halves, where $h_0^{(\beta)}$ is the dividing point. The past orbit relative to $h_0^{(\beta)}$ is

$$\left\{h_0^{(\beta)}\right\}^- = \left\{h_{-\infty}^{(\beta)}, \cdots, h_0^{(\beta)}\right\}. \tag{8}$$

The future orbit is similarly

$$\left\{h_0^{(\beta)}\right\}^+ = \left\{h_0^{(\beta)}, \cdots, h_{\infty}^{(\beta)}\right\}. \tag{9}$$

The action of the past [future] orbit is given relative to $x^{(\alpha)}$ [$x^{(\beta)}$]. In particular, the relative action between $\{h_0^{(\beta)}\}^-$ and $\{x^{(\alpha)}\}$ is defined as

$$\Delta\mathcal{F}_{\{h_0^{(\beta)}\}^- x^{(\alpha)}} = \sum_{n=-\infty}^{0} \left[F_{\{h_0^{(\beta)}\}}(q_{n-1},q_n) - F_{x^{(\alpha)}}(q,q)\right], \tag{10}$$

and similarly for the future orbit history,

$$\Delta\mathcal{F}_{\{h_0^{(\beta)}\}^+ x^{(\beta)}} = \sum_{n=0}^{\infty} \left[F_{\{h_0^{(\beta)}\}}(q_n,q_{n+1}) - F_{x^{(\beta)}}(q,q)\right]. \tag{11}$$

The total relative action of $\{h_0^{(\beta)}\}$ is just the sum of the two relative parts. Following Refs. [31,35] again yields, finally,

$$\begin{aligned} &\Delta\mathcal{F}_{\{h_0^{(\beta)}\}^- x^{(\alpha)}} + \Delta\mathcal{F}_{\{h_0^{(\beta)}\}^+ x^{(\beta)}} \\ &= \int_{U[x^{(\alpha)},h_0^{(\beta)}]} pdq + \int_{S[h_0^{(\beta)},x^{(\beta)}]} pdq \\ &= \mathcal{A}_{US[x^{(\alpha)}h_0^{(\beta)}x^{(\beta)}]}, \end{aligned} \tag{12}$$

where the subscript $US[x^{(\alpha)}h_0^{(\beta)}x^{(\beta)}] = U[x^{(\alpha)},h_0^{(\beta)}] + S[h_0^{(\beta)},x^{(\beta)}]$. Since this path is not closed, the final point is

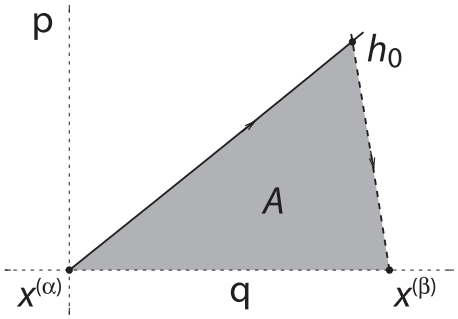

FIG. 3. A primary heteroclinic orbit of the standard map. $x^{(\alpha)} = (0,0)$ and $x^{(\beta)} = (0.5,0)$ are fixed points of the map, and $U[x^{(\alpha)},h_0]$ and $S[h_0,x^{(\beta)}]$ make a primary intersection at h_0. The relative action of $\{h_0\}$ should be equal to the area A.

added to the area notation. The integral gives the algebraic area A in Fig. 3.

This result can be generalized by considering a change in the dividing point $h_0^{(\beta)}$ to some other point $h_k^{(\beta)}$. The form of Eq. (12) must be unchanged. Thus,

$$\Delta\mathcal{F}_{\{h_k^{(\beta)}\}^- x^{(\alpha)}} + \Delta\mathcal{F}_{\{h_k^{(\beta)}\}^+ x^{(\beta)}} = \mathcal{A}_{US[x^{(\alpha)}h_k^{(\beta)}x^{(\beta)}]}, \tag{13}$$

where now the past and future relative actions are defined with respect to $h_k^{(\beta)}$ and the unstable and stable manifold integral paths change accordingly. This simple extension is quite useful ahead.

At this point we would like to make a remark on the difference between the areas defined in Eqs. (5) and (12): upon canonical transformations, the former is a closed area, thus invariant; while the latter is an open algebraic area, therefore not invariant. This is also consistent with the action functions on the left sides of the equations. Despite that the action functions are modified by the canonical transformations, the modifications cancel out between successive steps for the relative homoclinic actions, but not for the heteroclinics, the net change from which should match the change in the algebraic area.

A. Standard map example

Consider the action of a primary heteroclinic orbit of the standard map as an example. The mapping equations are [43]

$$\begin{aligned} p_{n+1} &= p_n - \frac{K}{2\pi}\sin 2\pi q_n, \\ q_{n+1} &= q_n + p_{n+1}, \end{aligned} \tag{14}$$

where our example is for the parameter $K = 8.25$, a value for which the system dynamics are overwhelmingly dominated by chaotic motion. Perhaps the simplest case is that of the two hyperbolic fixed points $x^{(\alpha)} = (0,0)$ and $x^{(\beta)} = (0.5,0)$. The first point is hyperbolic with inversion, which adds a new element to relations coming further ahead. The primary intersection of $x^{(\alpha)}$'s unstable manifold with $x^{(\beta)}$'s stable manifold, and the area A defined in Eq. (12) are drawn in Fig. 3. Calculating numerically the left-hand side of Eq. (12) using the action function for $\{h_0\}$, and the right-hand side using a construction of the manifolds gives $A - \Delta\mathcal{F}_{\{h_0\}^- x^{(\alpha)}} - \Delta\mathcal{F}_{\{h_0\}^+ x^{(\beta)}} = 9.4 \times 10^{-15}$, which is as accurate as one could expect using double precision computation. In this example, the two fixed points both lie on the $p = 0$ axis, and the algebraic area defined by Eq. (12) is relatively simple. Examples of more complicated heteroclinic orbits connecting fixed points with nonzero p values can be found in Fig. 11 of Ref. [44].

The area-relative-action relation has the advantage of giving results without the necessity of calculating the heteroclinic orbit. Only the intersection point h_0 and manifold segments are needed. Otherwise, a long orbit segment of $\{h_0\}$ centered at h_0 must be determined to get high accuracy. As numerical iterations forward and backward of h_0 fail to follow $\{h_0\}$ after a logarithmically short time in the precision divided by the Lyapunov exponent, numerical orbits diverge in this example after just a few iterations. Although techniques can be constructed to evade the divergence problem [44], Eq. (12) makes it unnecessary.

IV. RELATIVE ACTIONS BETWEEN HYPERBOLIC FIXED POINTS

A very interesting relation derives from comparing Eqs. (12) and (13). Subtraction generates a relation between the relative action between two fixed points with an area bounded by unstable and stable manifolds. Defining

$$\Delta\mathcal{F}_{x^{(\beta)}x^{(\alpha)}}(k) = k[F_{x^{(\beta)}}(q,q) - F_{x^{(\alpha)}}(q,q)] \tag{15}$$

gives

$$\Delta\mathcal{F}_{x^{(\beta)}x^{(\alpha)}}(k) = \mathcal{A}_{US[x^{(\alpha)}h_k^{(\beta)}x^{(\beta)}]} - \mathcal{A}_{US[x^{(\alpha)}h_0^{(\beta)}x^{(\beta)}]}. \tag{16}$$

Since

$$\int_{U[x^{(\alpha)},h_k^{(\beta)}]} pdq - \int_{U[x^{(\alpha)},h_0^{(\beta)}]} pdq = \int_{U[h_0^{(\beta)},h_k^{(\beta)}]} pdq \tag{17}$$

and similarly for the stable manifold segments, Eq. (16) simplifies to ($\Delta k = k_2 - k_1$)

$$\begin{aligned} \Delta\mathcal{F}_{x^{(\beta)}x^{(\alpha)}}(k) &= \mathcal{A}^{\circ}_{US[h_0^{(\beta)}h_k^{(\beta)}]}, \\ \Delta\mathcal{F}_{x^{(\beta)}x^{(\alpha)}}(\Delta k) &= \mathcal{A}^{\circ}_{US[h_{k_1}^{(\beta)}h_{k_2}^{(\beta)}]}. \end{aligned} \tag{18}$$

The $k = 1$ case is schematically illustrated in Fig. 4.

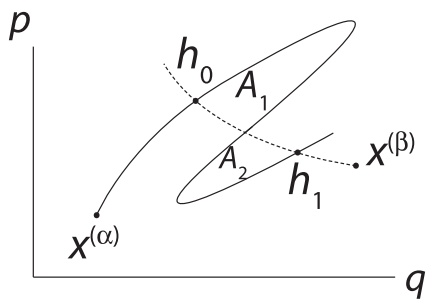

FIG. 4. Schematic partial heteroclinic tangle. The unstable manifold of $x^{(\alpha)}$ and the stable manifold of $x^{(\beta)}$ intersect at h_0, which maps to h_1. Notice that this image requires a second primary heteroclinic orbit, which is not labeled. The algebraic area $A_1 - A_2$ gives the relative action between $x^{(\beta)}$ and $x^{(\alpha)}$.

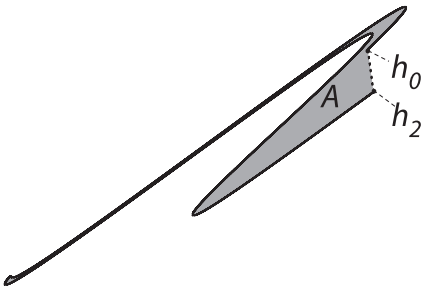

FIG. 5. The fundamental lobe structure of the heteroclinic tangle of $x^{(\alpha)}$ and $x^{(\beta)}$ in the standard map with the torus unfolded. The area of the lobe is enclosed by $U_{x^{(\alpha)}}[h_0,h_2]$ and $S_{x^{(\beta)}}[h_2,h_0]$. The relative action given by the area is $A = 0.835899764985$, which is to be compared with the analytic result using the generating functions of K/π^2. The difference is -6.43×10^{-11} showing that the boundaries of A are well determined numerically. The numerical agreement using double precision is reasonable given the long thin shape of part of the area.

In this case, Eq. (18) reads

$$\Delta\mathcal{F}_{x^{(\beta)}x^{(\alpha)}}(1) = \mathcal{A}^{\circ}_{US[h_0,h_1]} = A_1 - A_2, \tag{19}$$

where the last form is using the areas assigned in Fig. 4. A homoclinic tangle requires $A_1 = A_2$ since $x^{(\alpha)} = x^{(\beta)}$ [31].

A. Standard map example

Applying Eq. (18) to the standard map with the same fixed points as before highlights an intriguing situation due to $x^{(\alpha)}$ being hyperbolic with inversion. It turns out to be convenient to consider the twice-iterated map M^2, under which heteroclinic orbits stay on the same branch of the unstable manifold of $x^{(\alpha)}$. Therefore, to calculate the action difference between the two fixed points, we consider only the Δk-even cases. In addition, there is only one primary heteroclinic orbit, one lobe, and thus one area, not two; see Fig. 5. The fundamental lobe structure for the heteroclinic tangle does not look like a turnstile as it would for a homoclinic tangle [44]. Though not visible in the figure, the unstable manifold wraps counterclockwise around the fixed point $x^{(\alpha)}$ in order for this to be possible.

For k-even all the heteroclinic points map back onto the same branch of the unstable manifold of $x^{(\alpha)}$. Therefore, with $k = 2$, and h_0 and h_2 in Eq. (18),

$$\Delta\mathcal{F}_{x^{(\beta)}x^{(\alpha)}}(2) = \mathcal{A}^{\circ}_{US[h_0h_2]} = -A = \frac{-K}{\pi^2}, \tag{20}$$

where A is defined in Fig. 5, and $\frac{K}{\pi^2}$ comes from the generating function of the standard map. $\Delta\mathcal{F}_{x^{(\beta)}x^{(\alpha)}}(2)$ is the action difference between the two fixed points under M^2, the equality is verified to a high accuracy ($\sim 10^{-11}$).

V. SATELLITE PERIODIC ORBIT ACTIONS

In chaotic dynamical systems there is another generic class of unstable periodic orbits that are of great interest. They are identified as successive points on Moser invariant curves. Certain sequences of these orbits accumulate on particular homoclinic or heteroclinic orbits [15,21–23] and have been referred to as satellite periodic orbits [15]; see Appendix B for more details.

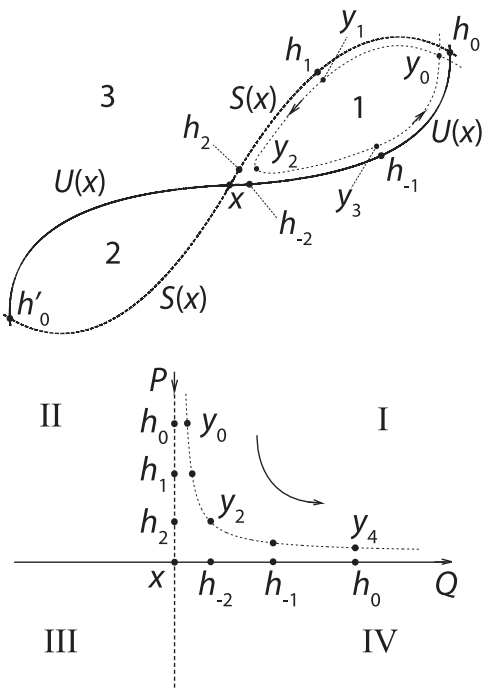

FIG. 6. A satellite orbit $\{y_0\}$ of period-4 associated with the homoclinic orbit segment $\{h_{-2},\cdots,h_2\}$ supported by an invariant Moser curve. Upper panel: $U(x)$ and $S(x)$ intersect at h_0. The Moser curve which supports a periodic-4 satellite orbit $\{y_0\}$, thinner dashed line, is plotted inside the complex region. The orbit segment $\{y_0,y_1,y_2\}$ follows the stable manifold segment $\{h_0,h_1,h_2\}$ for the first 2 iterations, then switches to the unstable manifold segment, such that $\{y_2,y_3,y_0\}$ follows $\{h_{-2},h_{-1},h_0\}$. y_2 is thus a switching point on $\{y_0\}$, where the orbit switches from the future to the past homoclinic segment. Lower panel: Normal form coordinates (Q,P). h_0 is on both axes. y_4 and y_0 correspond to the same point in phase space. y_2 is the switching point, which is associated either with h_2 or h_{-2}.

A. Satellite periodic orbits

Let us consider a periodic orbit associated with a homoclinic orbit segment and supported by an invariant Moser curve; see Fig. 6 for a schematically illustrated example. The Moser curve extends along $U(x)$ and $S(x)$ out to infinity and converges to them. Every homoclinic intersection between the manifolds will produce a self-intersection point on the Moser curve. As argued by Birkhoff [23] and numerically computed by da Silva Ritter *et al.* [22], special choices can be found for each sufficiently large integer N to make $\{y_0\}$ a period-N periodic orbit. As N increases, the corresponding y_0 converges to h_0, and the homoclinic orbit $\{h_0\}$ is itself the limiting case of the period-N periodic orbit $\{y_0\}$ for $N \to \infty$. The set of $y_0(N)$ taken from all integer N periodic orbits gives a sequence converging to h_0. In practice, for any homoclinic orbit, $\{h_0\}$, a truncation into finite segments $\{h_{-l},\cdots,h_0,\cdots,h_k\}$ $(k,l > 0)$ is possible, for which a Newton-Raphson search in its neighborhood can be used to construct the satellite orbit of period $(k+l)$ associated with $\{h_{-l},\cdots,h_0,\cdots,h_k\}$. This

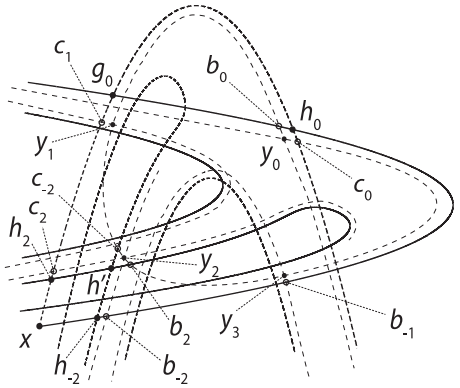

FIG. 7. (Schematic) Expanded view of the homoclinic tangle in Fig. 1. The manifold segments extending out of the drawing are simply connected and left out of this figure. A Moser invariant curve, light dashed line, is shown giving rise to a satellite periodic orbit $\{y_0\}$ of period 4. The invariant curve has been drawn more distant from the actual stable and unstable manifolds for illustration purposes. Every y_i is a self-intersection of the Moser curve. The curve intersects with $U(x)$ $[S(x)]$ at b_i $[c_i]$ near its self-intersection at y_i.

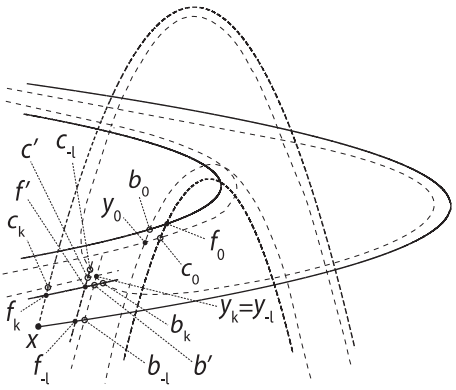

FIG. 8. (Schematic) Satellite orbit $\{y_0\}$ associated with a nonprimary homoclinic orbit segment $\{f_{-l},\cdots,f_0,\cdots,f_k\}$. Since the Moser curve approaches the stable and unstable manifolds, it must intersect with $U(x)$ $[S(x)]$ in the same way that $S(x)$ $[U(x)]$ does. It is under this sense that the homoclinic point f_0 will force two intersections b_0 and c_0 between the Moser curve and the unstable and stable manifolds, respectively.

provides a convenient way to construct the satellite orbits without the need to calculate the normal form series or the Moser curves.

The relative action difference between a given $\{h_0\}$ and its satellite $\{y_0\}$ is determined by a roughly parallelogram shaped region bounded by the manifolds and the Moser curve. To see how this area arises, consider the homoclinic tangle in Fig. 7, which is an expanded view of the tangle in Fig. 1. A Moser invariant curve is drawn which supports a period-4 orbit $\{y_0\}$, satellite to the homoclinic orbit $\{h_0\}$. y_2 is the switching point from the future to the past homoclinic segment. The orbit segment $\{y_0,y_1,y_2\}$ follows $\{h_0,h_1,h_2\}$, then switches at y_2, after which $\{y_2,y_3,y_0\}$ follows $\{h_{-2},h_{-1},h_0\}$.

The relative-action-area-relation derivation makes direct use of Eq. (A3) four times, once for each iteration of the map M:

(1) Starting from the initial point y_0, and map $M(y_0)=y_1$, follow the path $S[x,c_0]+I[c_0,y_0]$; $I[c_0,y_0]$ is the segment of the Moser invariant curve from c_0 to y_0. The path maps to $S[x,c_1]+I[c_1,y_1]$. Substituting the paths into Eq. (A3) yields

$$F_{\{y_0\}}(q_0,q_1)-F_x(q,q)=\mathcal{A}_{\mathrm{ISI}[y_0c_0c_1y_1]}. \tag{21}$$

(2) $M(y_1)=y_2$: Let the paths be $S[x,c_1]+I[c_1,y_1]$ and $S[x,c_2]+I[c_2,y_2]$, giving

$$F_{\{y_0\}}(q_1,q_2)-F_x(q,q)=\mathcal{A}_{\mathrm{ISI}[y_1c_1c_2y_2]}. \tag{22}$$

(3) $M(y_2)=y_3$: Let the paths be $U[x,b_{-2}]+I[b_{-2},y_2]$ and $U[x,b_{-1}]+I[b_{-1},y_3]$, giving

$$F_{\{y_0\}}(q_2,q_3)-F_x(q,q)=\mathcal{A}_{\mathrm{IUI}[y_2b_{-2}b_{-1}y_3]}. \tag{23}$$

(4) $M(y_3)=y_0$: Let the paths be $U[x,b_{-1}]+I[b_{-1},y_3]$ and $U[x,b_0]+I[b_0,y_0]$. This gives

$$F_{\{y_0\}}(q_3,q_0)-F_x(q,q)=\mathcal{A}_{\mathrm{IUI}[y_3b_{-1}b_0y_0]}. \tag{24}$$

The total relative action is, thus,

$$\Delta\mathcal{F}_{\{y_0\}x}=\mathcal{A}^{\circ}_{\mathcal{L}}, \tag{25}$$

where the compound closed path $\mathcal{L}$ is

$$\begin{aligned}\mathcal{L}=&\,I[y_0,c_0]+S[c_0,c_2]+I[c_2,y_2]+I[y_2,b_{-2}]\\&+U[b_{-2},b_0]+I[b_0,y_0].\end{aligned} \tag{26}$$

By adding and subtracting certain path segments, it is possible to deform $\mathcal{L}$ such that it separates into a path for the relative action of the homoclinic orbit and two parallelogram like correction terms. This gives the final desired relation between the relative action of the periodic and homoclinic orbits,

$$\begin{aligned}&\Delta\mathcal{F}_{\{y_0\}x}-\Delta\mathcal{F}_{\{h_0\}x}\\&\quad=\mathcal{A}^{\circ}_{\mathrm{SIIU}[xc_2y_2b_{-2}]}-\mathcal{A}^{\circ}_{\mathrm{SIIU}[h_0c_0y_0b_0]}.\end{aligned} \tag{27}$$

The two areas in the above equation resembles two near-parallelograms bounded by the manifolds and the Moser curves. The satellite orbit action is then

$$\begin{aligned}\mathcal{F}_{\{y_0\}}=&\sum_{n=0}^{3}F_{\{y_0\}}(q_n,q_{n+1})=4F_x(q,q)+\mathcal{A}^{\circ}_{US[xh_0]}\\&+\mathcal{A}^{\circ}_{\mathrm{SIIU}[xc_2y_2b_{-2}]}-\mathcal{A}^{\circ}_{\mathrm{SIIU}[h_0c_0y_0b_0]},\end{aligned} \tag{28}$$

where $\Delta\mathcal{F}_{\{h_0\}x}$ is given by area $\mathcal{A}^{\circ}_{US[xh_0]}$. Although the $\{h_0\}$ segment used here is a primary homoclinic orbit, with a careful definition of the points b_i and c_i near each orbit point y_i, a generalized Eq. (28) applies to satellite orbits associated with any homoclinic orbit segment. Take the example of Fig. 8, where a period-$(k+l)$ satellite orbit $\{y_0\}$ is associated with a nonprimary homoclinic orbit segment $\{f_{-l},\cdots,f_0,\cdots,f_k\}$. Since the Moser curve approaches the stable and unstable manifolds as it extends along them to infinity, it is forced to make a self-intersection at y_0 as the stable and unstable manifolds intersect at f_0. The particular Moser curve is the one for which the $(k+l)$th mapping of y_0 gives back y_0. Therefore, y_0 can be thought as being induced by f_0. Following the same logic, define b_0 to be the intersection between $U(x)$ and the Moser curve that is induced by f_0: as the Moser curve extends along $S(x)$, it intersects with $U(x)$ in the same way that $S(x)$

intersects, so the homoclinic point f_0 induces a b_0 on the Moser curve. Similarly, c_0 is defined as the intersection between $S(x)$ and the Moser curve that is induced by f_0. All b_i and c_i can be located in the same way, using f_i as the inducing point. It follows that all previous derivation steps continue to hold with the resulting more general expression of satellite orbit action:

$$\mathcal{F}_{\{y_0\}} = \sum_{n=0}^{k+l-1} F_{\{y_0\}}(q_n,q_{n+1}) = (k+l)F_x(q,q) + \mathcal{A}^{\circ}_{US[xf_0]} + \mathcal{A}^{\circ}_{\mathrm{SIIU}[xc_k y_k b_{-l}]} - \mathcal{A}^{\circ}_{\mathrm{SIIU}[f_0 c_0 y_0 b_0]}, \tag{29}$$

where f_0 can be any homoclinic point. This formula expresses the satellite action in terms of the fixed point action, the homoclinic relative action, and two four-segmented simple closed curves bounded by stable and unstable manifolds and the Moser curves. The calculation of the two areas requires the construction of the Moser curve, as well as the orbit points y_0 and y_k, which can be difficult to compute. However, a simple approximation scheme is possible.

B. Geometric area approximation

Equation (29) can be approximated with a wedge product form that only requires the location of the homoclinic points f_k and f_{-l}. In this way, it is possible to calculate the full action of a period-N ($N = k + l$) satellite orbit $\{y_0\}$ without its reconstruction or its Moser invariant curve. Assuming the action $F_x(q,q)$ and the area $\mathcal{A}^{\circ}_{US[xf_0]} = \Delta\mathcal{F}_{\{f_0\}x}$ of some homoclinic orbit point f_0 are known, then the first two terms on the right-hand side of Eq. (29) do not depend on knowing $\{y_0\}$, and only the two areas are needed. Notice from Fig. 8 that $\mathcal{A}^{\circ}_{\mathrm{SIIU}[f_0 c_0 y_0 b_0]}$ is mapped to $\mathcal{A}^{\circ}_{\mathrm{SIIU}[f_k c_k y_k b_k]}$ under k iterations, so that the areas of the two are identical. Thus,

$$\begin{aligned}\mathcal{A}^{\circ}_{\mathrm{SIIU}[xc_k y_k b_{-l}]} - \mathcal{A}^{\circ}_{\mathrm{SIIU}[f_0 c_0 y_0 b_0]} &= \mathcal{A}^{\circ}_{\mathrm{SIIU}[xc_k y_k b_{-l}]} - \mathcal{A}^{\circ}_{\mathrm{SIIU}[f_k c_k y_k b_k]} \\ &\approx \mathcal{A}^{\circ}_{\mathrm{SUIU}[xf_k b' b_{-l}]}.\end{aligned} \tag{30}$$

The final approximate closed path has only one side that depends on a Moser invariant curve. Furthermore, as shown in Fig. 8, $I[b',b_{-l}]$ is exceedingly close to $S[f',f_{-l}]$, where f' is a point on a different homoclinic orbit. Consider that

$$\mathcal{A}^{\circ}_{\mathrm{SUIU}[xf_k b' b_{-l}]} = \mathcal{A}^{\circ}_{\mathrm{SUSU}[xf_k f' f_{-l}]} + \mathcal{A}^{\circ}_{\mathrm{UIUS}[f' b' b_{-l} f_{-l}]} \tag{31}$$

and the mean expansion rate of the map is estimated by the positive stability exponent of the fixed point under one iteration of the map, e^{μ}. After $k + l$ iterations, the unstable segment $U[f_{-l},b_{-l}]$ is stretched into $U[f_k,b_k]$ with an expansion factor of roughly $e^{(k+l)\mu}$. This implies that the ratio of areas is

$$\frac{\mathcal{A}^{\circ}_{\mathrm{UIUS}[f' b' b_{-l} f_{-l}]}}{\mathcal{A}^{\circ}_{\mathrm{SUSU}[xf_k f' f_{-l}]}} \sim O(e^{-(k+l)\mu}). \tag{32}$$

For all but the smallest values of $(k + l)$, the small final area term of Eq. (31) can be dropped.

At this point, one can calculate $\mathcal{A}^{\circ}_{\mathrm{SUSU}[xf_k f' f_{-l}]}$ just by following the manifolds, which is very straightforward. However, there is a further approximation one can make. The manifolds are highly constrained in their behaviors in the local neighborhood of x. They must run along nearly parallel, nearly straight lines. This is approximately a parallelogram with area

$$\mathcal{A}^{\circ}_{\mathrm{SUSU}[xf_k f' f_{-l}]} \approx \delta q_{-l}\delta p_k - \delta p_{-l}\delta q_k = \delta f_{-l} \wedge \delta f_k, \tag{33}$$

where $\delta q_k = q_k - q$, $\delta p_k = p_k - p$ and similarly for $(\delta q_{-l},\delta p_{-l})$; i.e., the δ coordinates are just those of f_k and f_{-l} relative to x. With this approximation, to a high degree of accuracy the full satellite orbit action $\mathcal{F}_{\{y_0\}}$ is determined knowing only $F_x(q,q)$, $\mathcal{A}^{\circ}_{US[x,f_0]}$, f_k, and f_{-l}, in general,

$$\mathcal{F}_{\{y_0\}} \approx (k+l)F_x(q,q) + \mathcal{A}^{\circ}_{US[xf_0]} + \delta f_{-l} \wedge \delta f_k, \tag{34}$$

where $\{y_0\}$ is the satellite orbit associated with $\{f_{-l},\cdots,f_0,\cdots,f_k\}$. y_k is the switching point at which the orbit switches from $\{f_0,\cdots,f_k\}$ to $\{f_{-l},\cdots,f_0\}$.

A possible confusion arises from the fact that the same satellite orbit $\{y_0\}$ can also be viewed as associated with any shift in the truncation of the homoclinic orbit: $\{f_{-l+n},\cdots,f_n,\cdots,f_{k+n}\}$, where n is any integer. Furthermore, for $n \ll l,k$, the Newton iteration using $\{f_{-l+n},\cdots,f_n,\cdots,f_{k+n}\}$ as trail orbit will also converge, and one can verify that it leads to the same satellite orbit as using $\{f_{-l},\cdots,f_0,\cdots,f_k\}$. Therefore, the choice of the switching point along the satellite orbit seems not unique. This ambiguity can be resolved by defining the switching point to be the one that minimizes the error from approximation Eq. (34) in the original coordinate system, which is the error from replacing $\mathcal{A}^{\circ}_{\mathrm{SUSU}[xf_k f' f_{-l}]}$ by the wedge product $\delta f_{-l} \wedge \delta f_k$. In practice, the switching point is easy to identify. Since the error is the difference between the curvy "trapezoid" and its linear interpolation, the minimization is achieved by choosing the orbit point that is "closest" to the fixed point. Therefore, in the example of Fig. 7, the switching point can be identified graphically to be y_2, which is the closest point along $\{y_0\}$ relative to x. By ranging over all possible choices of f_0, l, and k, Eq. (34) suffices to calculate the classical actions of all periodic orbits inside the convergence zone.

C. Hénon map example

Consider the action of a satellite orbit in the area-preserving Hénon map [45]:

$$\begin{aligned} p_{n+1} &= q_n, \\ q_{n+1} &= a - q_n^2 - p_n, \end{aligned} \tag{35}$$

with parameter value $a = 10$. We have numerically computed a period-8 orbit $\{y_0\}$ satellite to one of the primary homoclinic orbit segments $\{h_{-4},\cdots,h_4\}$, where $y_0 = (3.1835765543, 3.1835765543)$ and $h_0 = (3.183580560, 3.183580560)$. This gives

$$\mathcal{F}_{\{y_0\}} - 8F_x(q,q) - \mathcal{A}^{\circ}_{US[x,h_0]} = 1.92729 \times 10^{-4}, \tag{36}$$

whereas the wedge product gives

$$\delta h_{-4} \wedge \delta h_4 = 1.92688 \times 10^{-4}. \tag{37}$$

The difference is in the 4th decimal place. On the other hand, without the knowledge of $\{y_0\}$, the full action can be calculated using Eq. (34),

$$\mathcal{F}_{\{y_0\}} = 138.940538512, \tag{38}$$

to be compared with the actual action,

$$\mathcal{F}_{\{y_0\}} = 138.940538553. \tag{39}$$

The relative error equals 3×10^{-10}, demonstrating the high accuracy of the wedge product approximation.

VI. CONCLUSION

The information about classical actions associated with homoclinic, heteroclinic, and periodic orbits that come into various semiclassical sum rules play an important role in the study of quantum chaotic dynamical systems. Although the orbit actions can be calculated from the generating functions, the relations and correlations among their values cannot be discovered without an analysis of the type given in this paper. Furthermore, in the asymptotic limit of semiclassical mechanics, the actions must be known to high precision to understand the interferences that arise in quantum dynamics, and that otherwise requires the accurate determination of long orbit segments. Since any initial deviation due to the machine precision will diverge exponentially, it is *a priori* difficult to compute periodic orbits with long periods. The analysis given here gives an explicit mechanism from which correlations could emerge and avoids the numerical difficulties by making the detailed long orbit calculations unnecessary.

One interesting example is given by the heteroclinic tangle of the standard map, which arises from the two unstable fixed points of the map on the $p = 0$ line. One of the fixed points is hyperbolic with reflection, which generates a single lobe fundamental structure in the tangle under a double iteration of the map. This lobe's area must equal twice the action difference of the fixed points, a nontrivial relation to imagine without generating Eqs. (12) and (13).

For fully chaotic systems, the convergence zone can cover most of the accessible phase space [24], and in that case nearly all of the periodic orbits fall into the category of satellite orbits, to which our analysis applies. Action differences between any pair of the satellite periodic orbits or between them and particular homoclinic (heteroclinic) orbits follow naturally. The simple, rather accurate geometric approximation involving a wedge product generates expressions that do not require the construction of the orbits or Moser invariant curves, only short sections of the stable and unstable manifolds (very simple and stable to calculate) and the endpoints of the homoclinic segments concerned. The error of this approximation scheme decreases exponentially as the length of the orbit increases and the instability exponent of the system increases.

All Moser invariant curves intersect in the untransformed phase space and as shown in Refs. [15,22,23], some satellite orbits lie on more than one Moser curve. In those cases, the actions of the satellite orbits are also related to homoclinic (heteroclinic) orbit actions, and possibly multiple fixed point actions [15]. It is of significant interest to understand the connections of the resulting multiple possible action relations, and this subject is left for future publication.

ACKNOWLEDGMENTS

The authors gratefully acknowledge that Eq. (19) (which relates the action difference of two periodic orbits with the loop structure of their heteroclinic tangle) was derived jointly with Akira Shudo, Hiromitsu Harada, and Kensuke Yoshida during a productive visit to Tokyo Metropolitan University and also gratefully acknowledge support for the travel.

APPENDIX A: MACKAY-MEISS-PERCIVAL ACTION PRINCIPLE

The MacKay-Meiss-Percival action principle discussed in this section was first developed in Ref. [31] for transport theory. A comprehensive review can be found in Ref. [35]. Generalization of the original principle beyond the "twist" and area-preserving conditions is discussed in Ref. [46]. A higher-dimensional generalization using generating 1-forms and phase-space volume forms is discussed in Ref. [47].

Consider an arbitrary point $a_0 = (q_0, p_0)$ and its orbit $\{a_0\}$ in phase space. The twist condition indicates the existence of a generating (action) function $F(q_n, q_{n+1})$, which brings a_n into a_{n+1} under the mapping M, such that

$$\begin{aligned} p_n &= -\partial F/\partial q_n, \\ p_{n+1} &= \partial F/\partial q_{n+1}. \end{aligned} \tag{A1}$$

The total action $\mathcal{F}$ is the sum

$$\mathcal{F} = \sum_{n=-\infty}^{\infty} F(q_n, q_{n+1}). \tag{A2}$$

The central step to obtain the MacKay-Meiss-Percival action principle is demonstrated by Fig. 39 along with Eq. (5.6) in Ref. [35]. Shown here in Fig. 9 are two arbitrary points $a = (q_a, p_a)$, $b = (q_b, p_b)$ and their images $a' = M(a)$, $b' = M(b)$. Let c be an arbitrary curve connecting a and b, which is mapped to a curve $c' = M(c)$, connecting a' and b'. Let A and A' denote the algebraic area under c and c', respectively. Then the difference between these areas is

$$\begin{aligned} A' - A &= \int_{c'} p dq - \int_{c} p dq \\ &= F(q_b, q_{b'}) - F(q_a, q_{a'}), \end{aligned} \tag{A3}$$

i.e., the difference between the two algebraic areas gives the difference between the action functions for one iteration of the map. Starting from this, MacKay *et al.* [31] derived a formula relating the action difference between a pair of homoclinic orbits to the phase space area of a region bounded by stable and unstable manifolds, as demonstrated by Fig. 10. In this

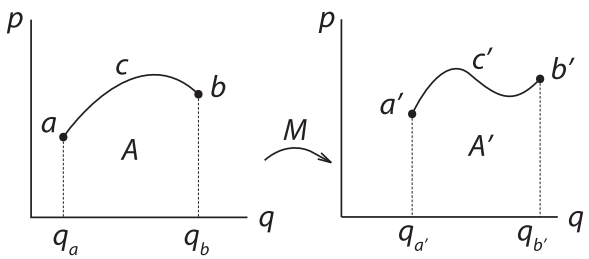

FIG. 9. a and b are arbitrary points and c is a curve connecting them. $a' = M(a)$, $b' = M(b)$, and $c' = M(c)$. Then: $A' - A = F(q_b, q_{b'}) - F(q_a, q_{a'})$.

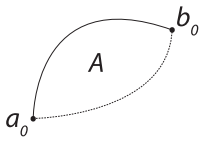

FIG. 10. a_0 and b_0 is a homoclinic pair. They are connected by an unstable segment $U[a_0,b_0]$ (solid) and a stable segment $S[b_0,a_0]$ (dashed). Then the action difference between the homoclinic orbit pair is $\Delta\mathcal{F}_{\{b_0\}\{a_0\}} = A$.

figure, a_0 and b_0 is a pair of homoclinic points:

$$a_{\pm\infty} \to b_{\pm\infty}. \tag{A4}$$

There exist unstable and stable manifolds connecting the two points shown by the solid and dashed curves. Those manifolds could be the manifolds of other fixed points or manifolds associated with a_0 and b_0 themselves. Let $U[a_0,b_0]$ and $S[b_0,a_0]$ be the corresponding segments, then the action difference between $\{a_0\}$ and $\{b_0\}$ is given by

$$\begin{aligned}\Delta\mathcal{F}_{\{b_0\}\{a_0\}} &= \sum_{n=-\infty}^{\infty}\left[F_{\{b_0\}}(q_n,q_{n+1}) - F_{\{a_0\}}(q_n,q_{n+1})\right] \\ &= \int_{U[a_0,b_0]} pdq + \int_{S[b_0,a_0]} pdq = A,\end{aligned} \tag{A5}$$

where A denotes the area shown in Fig. 10.

APPENDIX B: NORMAL FORM COORDINATES, MOSER INVARIANT CURVES, AND SATELLITE PERIODIC ORBITS

There are infinite families of unstable periodic obits accumulating on every homoclinic orbit [15,22]. These orbits are supported by Moser invariant curves, with the orbit points being successive self- or mutual-intersections between the invariant curves. The existence of such curves and orbits is a consequence of the Birkhoff-Moser theorem [21–23]. If the Poincaré map is invertible and analytic, there exists an analytic transformation (normal form transformation) from the normal form coordinates (Q,P) to the neighborhood of stable and unstable manifolds of the hyperbolic fixed point, for which the map takes the simple form

$$\begin{aligned}Q_{n+1} &= \Lambda(Q_nP_n)Q_n, \\ P_{n+1} &= [\Lambda(Q_nP_n)]^{-1}P_n,\end{aligned} \tag{B1}$$

where $\Lambda(Q_nP_n)$ is a polynomial function of the product Q_nP_n [24]:

$$\Lambda(QP) = \lambda + w_2(QP) + w_3(QP)^2 + \ldots, \tag{B2}$$

with $\lambda = e^{\mu}$, where μ is the Lyapunov exponent of the fixed point. The normal form convergence zone was first proved by Moser [21] to be a small disk-shaped region centered at the fixed point and later proved by da Silva Ritter *et al.* [22] to extend along the stable and unstable manifolds into infinity. The extended convergence zone follows hyperbolas to the manifolds ("gets exponentially close" the further out along the manifolds). The stable and unstable manifolds are just images of the P and Q axes, respectively, under the normal form transformation. Every homoclinic intersection point in phase space is mapped to two points, $H_P = (0,P_H)$ and $H_Q = (Q_H,0)$.

All points inside the extended convergence zone near the Q or P axis move along invariant hyperbolas, which are mapped to Moser invariant curves in phase space. Being confined in the extended convergence zone, the Moser invariant curves also get exponentially close to the stable and unstable manifolds while extending along them outward to infinity. In fact, as shown by Ref. [24], the convergence zone can be quantified using the outermost Moser curve with the largest QP product.

Self- and mutual intersections between certain Moser invariant curves give rise to infinite families of periodic orbits. A simple example is shown in Fig. 6. Since the Moser invariant curve (dotted line) extends along $S(x)$ and $U(x)$, intersections between $S(x)$ and $U(x)$ will "force" it to make self-intersections. Its topological behavior is thus determined by the topology of the homoclinic tangle. For example, when $S(x)$ and $U(x)$ make an intersection h_0, it is forced to make a self-intersection at y_0. Thus, one can say that y_0 is induced by h_0. Special choices of Moser curve can be found for each large enough integer N to make $\{y_0\}$ a period-N orbit. The detailed numerical technique is demonstrated in Ref. [22], where the position of y_0 is explicitly calculated using a linearization in the neighborhood of the homoclinic point. The upper panel of Fig. 6 shows a period-4 orbit $\{y_0\}$. The lower panel is the picture in (Q,P). Under four iterations, y_0 is mapped along the hyperbola into y_4. Under the normal form transformation, the P and Q axis become $S(x)$ and $U(x)$, respectively, folding back to intersect each other at h_0. The invariant hyperbolae fold in the same way, with the image of y_4 being identical to one of the images of y_0 due to their being at the self-intersection point. The solution for y_0 is unique for every period N. As N becomes larger, y_0 gets closer to h_0. The homoclinic orbit $\{h_0\}$ is the limiting case of the period-N orbit $\{y_0\}$ when $N \to \infty$:

$$\lim_{N\to\infty}\{y_0\} = \{h_0\}. \tag{B3}$$

The terminology of Refs. [15,22] refers to these $\{y_0\}$ as satellite orbits induced by h_0.

[1] H. Poincaré, *Les méthodes nouvelles de la mécanique céleste*, Vol. 3 (Gauthier-Villars et fils, Paris, 1899).
[2] P. So, Scholarpedia **2**, 1353 (2007).
[3] M. C. Gutzwiller, J. Math. Phys. **12**, 343 (1971), and references therein.
[4] R. Balian and C. Bloch, Ann. Phys. (NY) **63**, 592 (1971).
[5] M. V. Berry and M. Tabor, Proc. R. Soc. Lond. A **349**, 101 (1976).
[6] M. L. Du and J. B. Delos, Phys. Rev. A **38**, 1896 (1988).
[7] M. L. Du and J. B. Delos, Phys. Rev. A **38**, 1913 (1988).
[8] H. Friedrich and D. Wintgen, Phys. Rep. **183**, 37 (1989).
[9] S. Tomsovic and E. J. Heller, Phys. Rev. Lett. **67**, 664 (1991).
[10] S. Tomsovic and E. J. Heller, Phys. Rev. E **47**, 282 (1993).
[11] G. Tanner, P. Scherer, E. B. Bogomonly, B. Eckhardt, and D. Wintgen, Phys. Rev. Lett. **67**, 2410 (1991).
[12] P. Cvitanović and B. Eckhardt, Phys. Rev. Lett. **63**, 823 (1989).

[13] M. V. Berry and J. P. Keating, J. Phys. A: Math. Gen. **23**, 4839 (1990).
[14] N. Argaman, F.-M. Dittes, E. Doron, J. P. Keating, A. Y. Kitaev, M. Sieber, and U. Smilansky, Phys. Rev. Lett. **71**, 4326 (1993).
[15] A. M. Ozorio de Almeida, Nonlinearity **2**, 519 (1989).
[16] E. B. Bogomolny, Chaos **2**, 5 (1992).
[17] M. Sieber and K. Richter, Phys. Scr. **T90**, 128 (2001).
[18] S. Müller, S. Heusler, P. Braun, F. Haake, and A. Altland, Phys. Rev. Lett. **93**, 014103 (2004).
[19] M. Turek, D. Spehner, S. Müller, and K. Richter, Phys. Rev. E **71**, 016210 (2005).
[20] S. Müller, S. Heusler, P. Braun, F. Haake, and A. Altland, Phys. Rev. E **72**, 046207 (2005).
[21] J. Moser, Commun. Pure Appl. Math. **9**, 673 (1956).
[22] G. L. da Silva Ritter, A. M. Ozorio de Almeida, and R. Douady, Physica D **29**, 181 (1987).
[23] G. D. Birkhoff, Acta Math. **50**, 359 (1927).
[24] M. Harsoula, G. Contopoulos, and C. Efthymiopoulos, J. Phys. A: Math. Theor. **48**, 135102 (2015).
[25] G. Contopoulos and M. Harsoula, J. Phys. A: Math. Theor. **48**, 335101 (2015).
[26] M. Hénon, Quart. Appl. Math. **27**, 291 (1969).
[27] P. W. O'Connor, S. Tomsovic, and E. J. Heller, Physica D **55**, 340 (1992).
[28] S. C. Creagh, J. M. Robbins, and R. G. Littlejohn, Phys. Rev. A **42**, 1907 (1990).
[29] I. Esterlis, H. M. Haggard, A. Hedeman, and R. G. Littlejohn, Europhys. Lett. **106**, 50002 (2014).
[30] J.-M. Mao, J. Shaw, and J. B. Delos, J. Stat. Phys. **68**, 51 (1992).
[31] R. S. MacKay, J. D. Meiss, and I. C. Percival, Physica D **13**, 55 (1984).
[32] S. Wiggins, *Chaotic Transport in Dynamical Systems* (Springer, New York, 1992).
[33] R. Easton, Trans. Am. Math. Soc. **294**, 719 (1986).
[34] V. Rom-Kedar, Physica D **43**, 229 (1990).
[35] J. D. Meiss, Rev. Mod. Phys. **64**, 795 (1992).
[36] D. Bensimon and L. P. Kadanoff, Physica D **13**, 82 (1984).
[37] K. A. Mitchell, J. P. Handley, B. Tighe, J. B. Delos, and S. K. Knudson, Chaos **13**, 880 (2003).
[38] K. A. Mitchell, J. P. Handley, J. B. Delos, and S. K. Knudson, Chaos **13**, 892 (2003).
[39] K. A. Mitchell and J. B. Delos, Physica D **221**, 170 (2006).
[40] K. A. Mitchell, J. P. Handley, B. Tighe, A. Flower, and J. B. Delos, Phys. Rev. Lett. **92**, 073001 (2004).
[41] J. Novick, M. L. Keeler, J. Giefer, and J. B. Delos, Phys. Rev. E **85**, 016205 (2012).
[42] J. Novick and J. B. Delos, Phys. Rev. E **85**, 016206 (2012).
[43] B. V. Chirikov, Phys. Rep. **52**, 263 (1979).
[44] J. Li and S. Tomsovic, J. Phys. A: Math. Theor. **50**, 135101 (2017); arXiv:1507.06455
[45] M. Hénon, Comm. Math. Phys. **50**, 69 (1976).
[46] R. Easton, Nonlinearity **4**, 583 (1991).
[47] H. E. Lomelí and J. D. Meiss, Nonlinearity **22**, 1761 (2009).