

Symbolic encoding in symplectic maps

Freddy Christiansen† and Antonio Politi†‡

† Istituto Nazionale di Ottica, I-50125 Firenze, Italy

‡ INFN Sezione di Firenze, Italy

Received 17 October 1995, in final form 8 July 1996

Recommended by P Cvitanović

Abstract. A general procedure to construct a generating partition in 2D symplectic maps is introduced. The implementation of the method, specifically discussed with reference to the standard map, can be easily extended to any model where chaos originates from a horseshoe-type mechanism. Symmetries arising from the symplectic structure of the dynamics are exploited to eliminate the remaining ambiguities of the encoding procedure, so that the resulting symbolic dynamics possesses the same symmetry as that of the original model. Moreover, the dividing line of the partition turns out to pass through the stability islands, in such a way as to yield a proper representation of the quasiperiodic dynamics as well as of the chaotic component. As a final confirmation of the correctness of our approach, we construct the associated pruning front and show that it is monotonous.

PACS numbers: 0545

1. Introduction

Symbolic dynamics is a powerful tool for investigating statistical properties of dynamical systems. In fact, the representation of any trajectory as an infinite sequence of symbols allows a fruitful application of thermodynamic formalism [1] and an effective computation of statistical averages. Among the various approaches introduced to obtain a meaningful symbolic representation [2], the direct construction of a generating partition from homoclinic tangencies [3] seems to be the most general and powerful one. One of its advantages is the possibility to construct the corresponding pruning front [4] and to study the grammar of the underlying language [5] from it, as well as the possibility to obtain a faster convergence of thermodynamic variables [6].

The method has been successfully applied to 2D maps [3] and 3D flows [7] with a relatively strong contraction of volumes in phase space. More recently, an extension to symplectic maps has been proposed in [8] which is able to account for the evolution of the main ergodic component, while leaving open the problem of describing the evolution in the stability islands. In this paper, we solve this last problem by properly taking into account the symmetry properties of the evolution. Moreover, our approach allows removing the ambiguities detected in [8], obtaining a symbolic sequence which is invariant under time-reversal as is the evolution of the map.

We have chosen to work with the most general prototype of Hamiltonian chaos, namely the standard map

$$\begin{aligned}x_{n+1} &= y_n \\ y_{n+1} &= -x_n + 2y_n - \alpha \cos(y_n) \bmod 2\pi,\end{aligned}\tag{1}$$

which, in a shorthand notation, will be denoted as $F(\cdot)$. In place of the usual representation of the dynamics in terms of the variables θ and ρ , we refer to $x = \theta$ and $y = \rho + \theta$, since the map, in this representation, transforms horizontal into vertical lines (as, for instance, in the Hénon map), thus providing a more natural description of the horseshoe-type transformation.

There are two relevant symmetries characterizing the action of map (1). Firstly, the map is invariant under the composition of time reversal with the involution $T(x, y) = (y, x)$ exchanging x and y variables, i.e., $T \circ F \circ T = F^{-1}$. Secondly, the map is invariant under the transformation $S(x, y) = (\pi - x, \pi - y) \bmod 2\pi$, namely, $F = S \circ F \circ S$.

We believe that the arguments developed in this paper are general enough to be applicable to any 2D symplectic map (and to Hamiltonians with 2 degrees of freedom), provided that the corresponding symmetries are suitably taken into account. In fact, the problems that we face while discussing the geometric structure of the sequence of primary homoclinic tangencies are by no means peculiar to the standard map.

In section 2, we briefly describe the action of the map, introducing a first approximation of the partition. In the following section, we discuss how homoclinic tangencies (HTs) can be computed and used to construct a symbolic representation. Section 4 is devoted to the illustration of the main difficulty arising in the identification of the so-called primary tangencies. In section 5, we make use of the symmetry properties to get rid of the ambiguities discussed in the previous section, arriving at a prototype of generating partition which takes into account the properties of the main chaotic component. In section 6, we discuss stability islands, showing that also such trajectories can be properly included into the previous framework. In section 7, we use the generating partition to construct the corresponding symbol-plane and the pruning front. In section 8, we change the nonlinearity to show that we are still able to cope with different structures of the stability islands, reinforcing the conjectured generality of the method herein proposed.

2. General remarks

A partition is said to be generating if for every bi-infinite sequence of symbols there may at most exist one physical trajectory. A general guideline to construct a generating partition consists in looking for the regions where the map folds the phase-space, so that the support of the measure remains confined notwithstanding the expansion along the unstable direction. This is extremely clear in the logistic map, where the position of the maximum, i.e. the folding point, is the only ingredient required to construct the generating partition which turns out to consist of the two subintervals lying to the left and to the right of the maximum, respectively [9].

In two dimensions, the construction of a partition is not equally simple but, as long as the chaotic evolution originates from a horseshoe-type mechanism, the same ideas proved to be very powerful in a series of dissipative models [3, 7]. Accordingly, it is natural to start the analysis of the standard map from this point of view. We have chosen to work with a high value of the nonlinearity, namely $\alpha = 6$, since stability islands, which certainly require a different treatment, cover a tiny portion of the phase space, as can be seen from figure 1, where the largest islands (one period-2 and two period-5) have been reported. In a first approximation, one can hope to construct a partition just disregarding such stability islands.

The folding regions of the map are approximately situated at the vertical lines defined by $x = \sin^{-1}(-2/\alpha)$, as can be seen by looking at the image of a generic vertical line. Let us arbitrarily choose the locus of primary tangencies to be the two lines corresponding to the two roots of the above equation, namely $L_1 : (x = 3.481 \dots)$ and $L_2 : (x = 5.943 \dots)$ (see figure 2). At variance with simple horseshoes (like the Hénon map), the folding lines

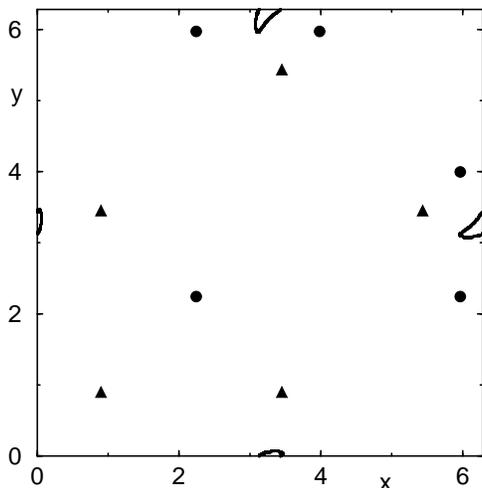


Figure 1. The largest periodic stability islands of the standard map for $\alpha = 6$: the continuous line represents the border of a single period-2 island; crosses and triangles indicate two, reciprocally symmetric, period-5 islands, which are too small to be resolved on the actual scale.

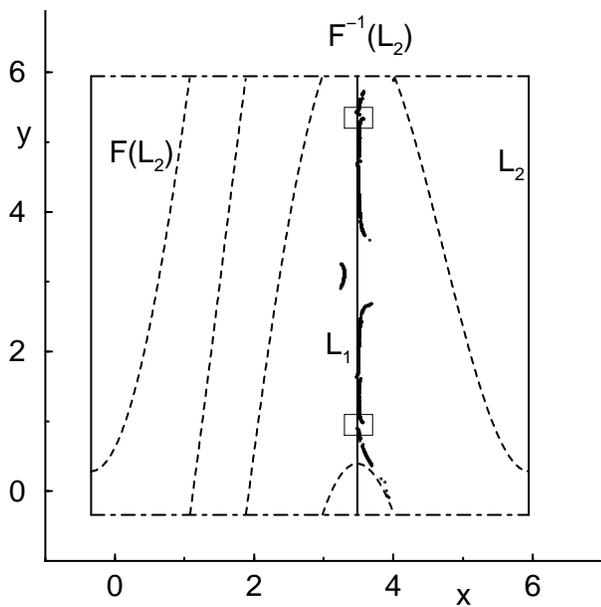


Figure 2. The approximate generating partition obtained starting from the vertical lines L_1 and L_2 (solid lines), the pre-image of L_2 , $F^{-1}(L_2)$ (dashed-dotted lines), and its image $F(L_2)$ (dashed line). Dots denote homoclinic tangencies classified as primary. The small rectangular boxes delimit the regions reported in figure 3.

are not sufficient to split the phase space into disjoint sets because of the topology (the phase space is a torus with no natural boundaries along both x and y directions). One must, therefore, break the continuity by introducing two sets of transversal lines, a distance 2π

apart horizontally and vertically, respectively. We find it convenient to use the folding lines themselves. For instance, this can be done by using the horizontal pre-image $F^{-1}(L_2)$ of L_2 . As a result, the plane is partitioned into infinitely many equivalent squares S . Any other pair of transversal lines is, in principle, equivalent; the idea of using L_1 and L_2 is inspired by the attempt to make the partition as simple as possible. As a result, the square is split into two subsets. The resulting partition is not sufficiently fine-grained to account for the multiplicity of trajectories generated by map (1). In fact, the (pre-)images of the two subsets intersect different copies of S . One is therefore led to further divide each of the two subsets into as many elements as the number of copies of S which are visited. This is automatically obtained by using $F(L_2)$ as a further dividing line (dashed line in figure 2). The resulting partition turns out to be approximately generating, in the sense that a small fraction of periodic orbits is characterized by the same symbolic sequence [8].

Thus we can conclude this preliminary discussion by stating that the two folding lines account for the basic mechanisms involved in the underlying dynamics and we can consider the partition of figure 2 as a tentative solution to be improved on the basis of more rigorous arguments.

3. Homoclinic tangencies

The most general method to construct seemingly exact generating partitions in dissipative dynamical systems is based on the refinement of the concept of folding points, i.e. on the identification of the ‘primary’ homoclinic tangencies and on the subsequent connection of all such points with one or more continuous curves representing the border of the elements of the partition. Such a strategy has been successfully applied to moderately and strongly dissipative systems, including both maps and flows [10, 7].

The main justification for this method comes from the following argument. Because of the folding process, if a fibre of the unstable manifold [11] (w_u) intersects a fibre of the stable manifold (w_s), it does so twice (see figure 3). When this is the case, let us call two such trajectories *companion*, as e.g., the orbits stemming from P_1 and P_2 in figure 3. Two companion orbits approach each other both in the past and in the future, as they belong to the same fibres of unstable and stable manifold. For a partition to be generating, it is necessary that a border separates $F^n(P_1)$ from $F^n(P_2)$ for some n (either negative or positive), since the same reasoning applies to any pair of intersections, no matter how close they are. Thus, the only way to distinguish the corresponding symbolic sequences is to set the border of the partition exactly on the tangency point P_t , or on some backward (forward) image of it. As long as one limits the analysis to just one fibre, all choices are equivalent. However, the partition of phase space into distinct elements requires taking all fibres simultaneously into account. As a consequence, one is faced with the problem of implementing a consistent procedure.

Any attempt of putting the above considerations on a rigorous ground faces two problems: (i) there is no argument to justify, *a priori*, the existence of a subset of HTs (the primary ones) aligning along more or less smooth curves such that the phase space is split in a meaningful way (i.e., so that the elements are compact sets); (ii) there is no clear-cut definition of ‘primary’ tangencies: they are usually identified *a posteriori*, after implementing a careful trial and error approach.

Before discussing in detail how the above problems arise in the context of the standard map, let us briefly define the numerical methods adopted for identifying HTs. First of all, one should notice that the line L_2 can be mapped onto L_1 by exploiting the symmetry of map (1). We need, therefore, to study only one folding line, namely L_1 . Moreover, since

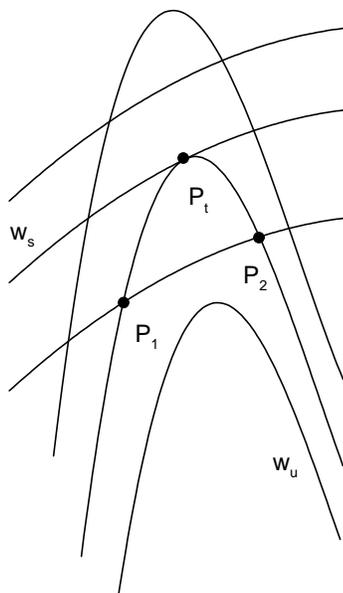


Figure 3. Sketch of the intersections between stable (w_s) and unstable (w_u) manifolds. The trajectories originating from P_1 , P_2 approach each other both in the past and in the future. The point P_t indicate a homoclinic tangency.

any piece of unstable manifold eventually fills the main ergodic component densely, we can restrict ourselves to the unstable manifold w_u of the hyperbolic fixed point $O = (\pi/2, \pi/2)$.

The manifold can be formally parametrized as $w_u(s) = (x_u(s), y_u(s))$ and the functions $x_u(s)$, $y_u(s)$ expanded in a power series of s . The coefficients of the series can then be determined by demanding that the curve is left invariant by the map, i.e.

$$F(x_u(s), y_u(s)) = (x_u(\lambda s), y_u(\lambda s)) \tag{2}$$

where λ is the unstable eigenvalue of the stability matrix for the fixed point $O = (x_u(0), y_u(0))$. HTs are identified by iterating a piece of unstable manifold and looking for those points where the local curvature appears to diverge[7]. The curvature $C(s)$ in the point $w_u(s)$ can be determined from the standard formula

$$C(s) \equiv \frac{\|\mathbf{u} \times \mathbf{v}\|}{\|\mathbf{u}\|^3}, \tag{3}$$

where \mathbf{u} and \mathbf{v} are the first two coefficients of the power expansion of w_u around s [12],

$$w_u(s + \delta s) = w_u(s) + \mathbf{u}\delta s + \mathbf{v}\frac{\delta s^2}{2} + \mathcal{O}(\delta s^3). \tag{4}$$

The above procedure leads to a tentative set of primary HTs which are defined as such if they lie close to L_1 . One can see that actually most of the tangencies are really close to the first approximation of the dividing line, confirming the suitability of the former approach and also suggesting we are on the right track towards the identification of the truly ‘primary’ tangencies. However, problems are expected to arise whenever a supposedly primary tangency returns to the folding region, because then it is not clear which of the two points should be used to identify the border of the partition. Moreover, practically speaking, if one tries to connect all the tangencies with a single curve, one cannot avoid discontinuities

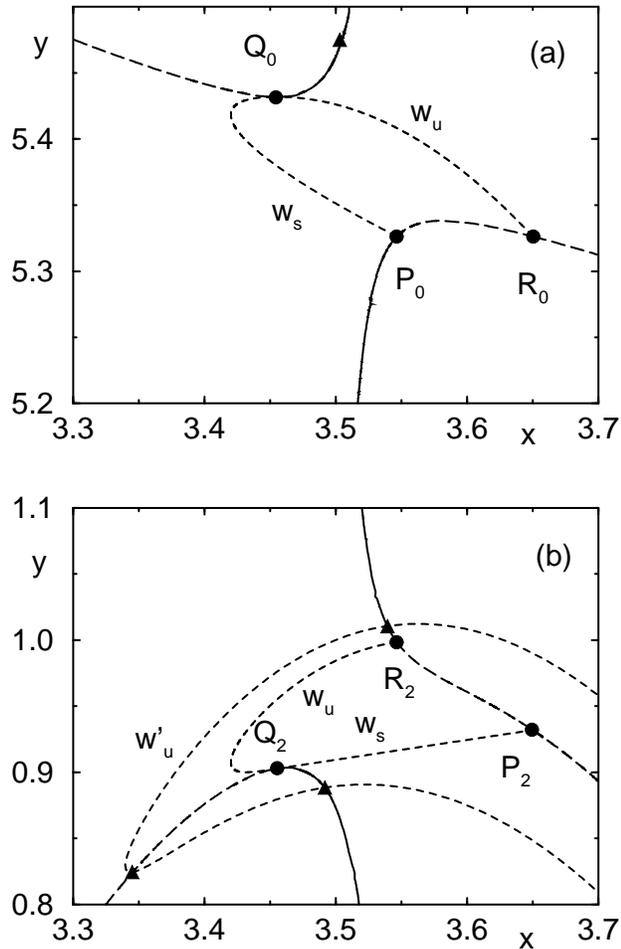


Figure 4. Enlarged picture of the avoided crossings contained in the small boxes of figure 2. The region in (b) is the second forward image of (a). Solid lines denote those HTs which are unambiguously classified as primary. Out of the other tangencies (dotted lines), the points along P_0R_0 may or may not be classified as primary; when not, their second image, lying on R_2P_2 , must be taken as primary. Q_0 and its second image Q_2 are identified as the points where two sequences of HTs meet and collapse. The stable and unstable manifolds (dashed lines) departing from Q_0 (Q_2) intersect the strand of HTs in P_0 (P_2) and R_0 (R_2), respectively. A branch of the unstable manifold containing three tangencies (triangles) is also shown in (b). The second pre-image of the intermediate tangency is also reported in (a) (full triangle).

and relatively big jumps (again, see figure 2). In dissipative systems, this is not considered to be a serious problem. Since the attractor does not fill the whole phase space, one has a large degree of freedom in connecting HTs that are far apart, as long as such lines do not intersect other pieces of the attractor. This is no longer true in a conservative map, where the entire phase space is typically filled by a single ergodic component (with the exception of stability islands which will be considered separately).

4. Discontinuities

In order to better clarify what happens around each discontinuity, let us look closer at one example, namely within the two boxes of figure 2, enlarged in figure 4 (the region in figure 4(b) is the second iterate of that depicted in figure 4(a)). We realize that the jump is the consequence of an ‘avoided crossing’ between two lines of HTs. The discontinuity is in fact caused by the intersection of the border of our generating partition with some forward or backward image of itself. Such a phenomenon is clearly seen in figure 4, where the second forward and second backward images (long dashed lines) of the ‘primary’ tangencies (solid lines) have been added. One is clearly faced with the question of where to stop considering a line of HTs as primary.

In figure 4(b) it is seen that three distinct tangencies can be identified on those fibres of the unstable manifold which are not too close to the jump (see, e.g., the triangles on w'_u). The first and the last of such points are unambiguously classified as primary points, whereas the middle one corresponds to the 2nd iterate of a tangency classified as primary in the upper part (see the triangle in figure 4(a)). Upon shifting the fibre of reference w'_u towards the region with the avoided crossing, the two lower HTs meet to disappear afterwards, thus preventing a continuation of the dividing line. This happens in the special point Q_2 , where stable and unstable manifolds are not only tangent, but have also the same curvature.

The construction of a meaningful partition now requires bridging the gap between the lower part of the dividing line, which just stops in Q_2 , with the above line, which lies a finite distance apart. Two reasonable candidates are represented by the two branches of invariant manifold which, departing from Q_2 , cross the upper line of HTs (see dashed lines) in R_2 and Q_2 , respectively. Two such fibres, together with the piece of HTs connecting P_2 with R_2 (which will be denoted by P_2R_2) delimit a closed region U . Let us now focus our attention on a trajectory visiting U and a nearby trajectory on the opposite side of P_2R_2 . It is clear from figure 4, that an orbit visiting U can be discriminated from its companion trajectory either when the orbit visits $F^{-2}(U)$ (the region encircled by the curves passing through Q_0 , P_0 and R_0 in figure 4(a)), or when the orbit lies in U itself. As an example, if we choose to use the fibre of the unstable manifold, w_u , to connect Q_2 to R_2 , the two trajectories will not be discriminated in figure 4(b). However they will be on opposite sides of the partition in figure 4(a). In [8], it was conjectured that any curve C lying in U and connecting Q_2 with a point S on P_2R_2 is appropriate, provided that $F^{-2}(C)$ also is used in $F^{-2}(U)$ in a self-consistent manner. The border itself of U , i.e. the manifolds w_u and w_s , can be used to construct the partition.

In any case, we see that an intrinsic and unavoidable ambiguity is associated with the existence of avoided crossings of the dividing line. This is the same phenomenon observed for some parameter values in dissipative maps, where it was found that some periodic orbits can be encoded in different ways [13, 14]. In a conservative system, like the standard map under investigation, the same problem occurs for any parameter value, since moving with continuity across the fibres of w_u is like changing a parameter of the dynamics.

The above described scenario is nothing but a single example of a phenomenon which occurs any time an image of the dividing line returns to the folding region, i.e. infinitely often. However, from the series of supposedly HTs reported in figure 2 one can see that the size of the jumps appear to diminish with the respective number of iterates needed to return to the folding region. Accordingly, it is tempting to conjecture that the whole procedure can in principle be implemented until an infinitesimal resolution is reached without yielding too wild a separation line.

5. Symmetry lines

The partition constructed from homoclinic tangencies plus pieces of either stable or unstable manifolds suffer from three problems: (i) the pieces of manifold do not reflect the symmetries of the dynamical system; (ii) stable islands are not taken into account and orbits inside a stable island are therefore not distinguished by the partition; (iii) in a section near the stable period-2 island the partition is constructed by using the same tangencies twice, i.e. both primary tangencies and one of their forward images which, at least superficially, seems redundant.

When we searched for HTs we looked solely at the unstable manifold of one of the unstable fixed points. We might equally well have done the same for the stable manifold and would then have gotten another set of tangencies. Due to the time reversal symmetry of the dynamics this set would have been just the original set of HTs transformed by T , i.e. a reflection in the line $x = y$. Since the tangencies found by looking at the stable manifold are by their very nature also points on the unstable manifold, the two sets must be related to each other by the dynamics. In fact, almost all HTs obtained by monitoring the curvature of the stable manifold are pre-images of the tangencies determined from the unstable manifold. This is a straightforward consequence of the invariance of the map under time reversal.

The identification of primary tangencies is more vague in regions where the folding process is not strong, for example in the area near the stable period-2 island. Here it is much more likely than elsewhere to find tangencies where both the stable and the unstable manifolds have a noticeable curvature and where this picture persists for several iterations of the map, introducing some ambiguity in the notion of primary tangency.

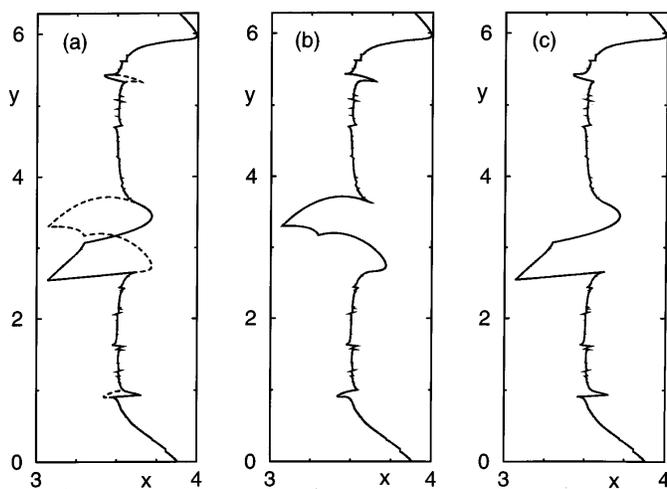


Figure 5. (a) The generating partitions as constructed from primary HTs and suitable pieces of stable (full line) and unstable (dashed line) manifolds. Near the avoided crossings, specifically the large gap near (π, π) , the set of HTs used here is slightly different from that of figure 2 in order for the HTs to match up with the pieces of stable and unstable manifold originating from the main line of HTs [8]. (b) only the partition constructed with pieces of unstable manifold. (c) The partition in (b) transformed first by the involution T (interchange of x and y) and then by the map F . Note that (c) is exactly equal to the full line in (a).

In figure 5(a) we show the two partitions obtained from primary HTs of the unstable manifold plus pieces of the stable and unstable manifolds respectively. In figure 5(b) only

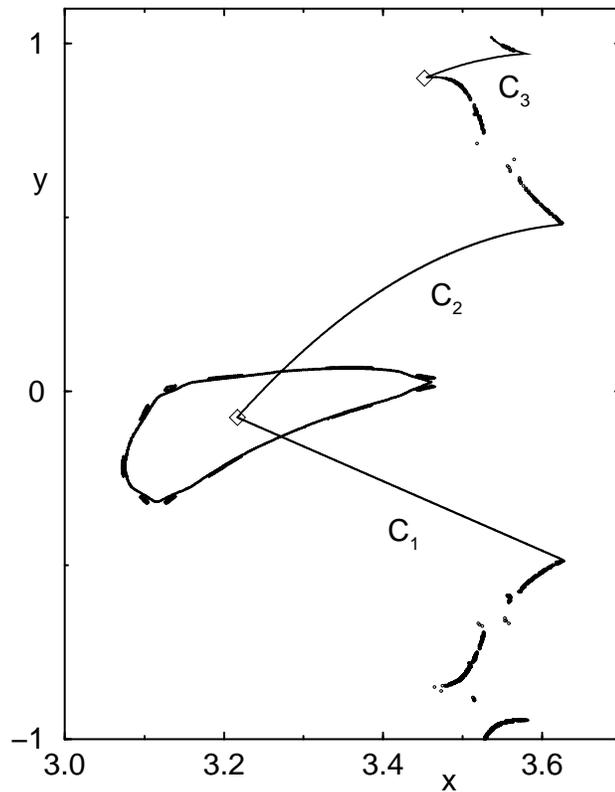


Figure 6. The region near the stable period-2 island. The curve C_1 is given by $y = \pi - x$. C_2 is found as $S \circ F(C_1)$. The two curves together define a winding number partition of the stable island, at the same time as they eliminate the need for redundant use of homoclinic tangencies in the partition. C_3 bridges the gap of figure 3.

the partition using pieces of unstable manifold is shown. This should be compared to figure 5(c) which is the partition of figure 5(b) first transformed under the involution T , and then mapped once forward by the map F . This curve is exactly identical to the partition obtained by combining tangencies with pieces of stable manifold.

The choice of using pieces of stable or unstable manifolds or other curves that may bridge the gap is rather arbitrary and gives us some freedom to obtain the following: that the whole partition line, not just the unambiguously primary tangencies, will be left invariant under the transformation $F \circ T$.

Recall that a gap in the curve of HTs is related to a crossing of the partition with some image, F^n , of itself. The gap is therefore related to a similar gap somewhere on the partition and we already saw that the curves that bridge the two gaps must also be related as F^n -(pre-)images of each other. The two gaps are furthermore related by the $F \circ T$ symmetry and if we want the corresponding curves to have the same symmetry, we end up with the relation

$$F^n(x, y) = F \circ T(x, y) . \tag{5}$$

Therefore, the curves are parts of the set of fixed points for the mapping

$$\phi_n(x, y) = T \circ F^{n-1}(x, y) , \tag{6}$$

which turns out to be an involution since

$$\phi_n \circ \phi_n = T \circ F^{n-1} \circ T \circ F^{n-1} = T \circ T \circ F^{1-n} \circ F^{n-1} = \text{id} \quad \det \frac{d\phi_n(x, y)}{d(x, y)} = -1. \quad (7)$$

We will call this set of fixed points the symmetry lines of ϕ_n [15]. Note that where two such symmetry lines cross we find a periodic orbit of F since $\phi_n \circ \phi_m(x, y) = F^{m-n}(x, y)$. Let us take a look at the simplest of these curves, namely for $n = 1$ and $n = 2$. The curve for $n = 1$ relates to the large gap near (π, π) . For $n = 1$ the symmetry line of equation (7) is trivially found to be the time reversal symmetry line of F , $x = y$, which will give us the upper line in the central gap. The lower curve is then found as the first forward image of $x = y$ or more explicitly $y = x - \alpha \cos(x) \bmod 2\pi$. For ϕ_2 the symmetry line is determined by $-x + 2y - \alpha \cos(y) = x \bmod 2\pi$ or $x = y - \frac{\alpha}{2} \cos(y) \bmod \pi$. This curve will bridge the gap shown in figure 3(a). The second forward image of this will then be the curve that bridges the gap in figure 3(b). The symmetry line for ϕ_2 is also related to two smaller gaps and can be related to the stable period-2 island, but for this some ambiguity still exists.

There is another possibility for the creation of a gap in the line of HTs, namely that the partition line crosses a symmetric image of itself, i.e. the two corresponding gaps are related by the transformation $S \circ F^n$, where S gives the spatial symmetry of the map. They are still related by the symmetry $F \circ T$ so the corresponding symmetry curves are fixed point of

$$\phi_n^{(S)}(x, y) = S \circ T \circ F^{n-1}(x, y). \quad (8)$$

For $n = 1$ the symmetry line is given by $y = (\pi - x) \bmod 2\pi$ and for $n = 2$ we find the curve as $y = \pi/2 \bmod \pi$. The first of these curves is related to the period-2 island, whereas the second bridges some smaller gaps in the partition.

Fixed curves for other values of n can in principle be found directly from the corresponding involutions ϕ_n or $\phi_n^{(S)}$ but it is in fact much easier than that in [15]. All the curves for n odd are found as (pre-)images of the curve for $n = 1$ and all the curves for n even are found as (pre-)images of $n = 2$. In this way it is easy to bridge all significant gaps in the partition line and we notice that the size of the gaps decrease fast with rising n .

6. Stable islands

As we noted in the previous section, there is still some ambiguity associated with, e.g., the stable period-2 island: many of the homoclinic tangencies were used twice in the original partition. Furthermore, we are interested in having the partition pass through the island in order to distinguish orbits inside the island, something which clearly cannot be done with a line of points on the unstable manifold of an orbit outside the island, as this manifold will never penetrate the island. We are going to deal with this by including two of the symmetry lines determined in the previous section. But also here some free choice exists because many of the curves pass through the period-2 orbit and the island associated with it. In fact, since the period-2 orbit is invariant under not only the transformation $F \circ T$, as all period-2 orbits must be, but also under the symmetry operation $F \circ S$, all symmetry lines of ϕ_n for n even and all symmetry lines of $\phi_n^{(S)}$ for n odd will pass through the island. For the moment we will just use the simplest choice, namely that of $\phi_1^{(S)}$. As given in the previous section, this involution has the fixed curve $y = \pi - x$. In figure 6, we see how this curve, C_1 , will connect the HTs with the period-2 orbit. The period-2 orbit is situated where the partition line crosses the pre-image of the symmetric partition. To reconnect the partition line to the HTs we must therefore use the curve obtained from $F \circ S(C_1)$ or, more precisely, the curve $y = (3x - \alpha \cos(x) - \pi) \bmod 2\pi$. This curve is C_2 in figure 6.

Including C_1 and C_2 in the partition has had the effect of excluding exactly a part of the original partition which contains one copy of the HTs that was used twice and we have thus rid the partition of this redundancy. At the same time we have made a partition of the stable period-2 island as shown in figure 6. This partition works very much like a winding number description inside the island: we have chosen a partition line from the central point passing through the border of the island. In addition, we have another line which is the forward image of the first line (albeit related through a symmetry). Any point in the island inside the small section to the right of the partition is therefore mapped out of this section by $F \circ S$. Any orbit outside the small section must however map into it to go around the central point. To distinguish orbits inside the island, one may therefore start by counting the average number of times they hit the small section, i.e. we can construct a winding number. This winding number varies monotonically along the partition and we are therefore able to distinguish quasiperiodic orbits inside the island, but it is possible that we may not be able to take into account motion in sub-islands inside the main island. We will return to this later.

In the top of figure 6 we see the gap of figure 3. This is bridged by the curve, C_3 . This is the second forward image of the symmetry line of ϕ_2 or, equivalently, the symmetry line of ϕ_{-2} . It is interesting to notice that the symmetry lines turn out to touch the sequence of primary HTs exactly in the higher-order points where two primary tangencies collide (like Q_2 in figure 3(b)). This result represents a strong indication of the internal consistency of our construction of a generating partition. In fact, it is not *a priori* obvious that two seemingly unrelated ingredients such as homoclinic points and symmetries match perfectly together.

Very near this gap there is a stable period-5 island. We would like to have the partition pass through this island in the same way as for the period-2 island. However, C_3 stops short just before reaching the island, but a further refinement of the partition will solve this problem. We will return to this later.

We are now ready to produce the partition that we have sought. This is shown in figure 7. We have used a similar notation to figure 2. We use the the dividing line D_2 plus its pre-image to make up the border of the primary region of phase space. This is partitioned by D_1 . We must also consider that the points in the primary region originate from different regions, i.e. we must consider the effect of the ‘mod 2π ’ in the map (1). The forward image of D_2 separates points that are mapped in one iteration into the primary region. This gives in total seven elements in the primary region and the dynamics can be described topologically by a seven letter alphabet.

7. A closer look at the stability islands

In order for the partition to be truly generating, the partition line must pass not only through every primary stable island, but also through all subislands inside these. In figure 8 we show that such a partition is possible. In figure 8(a) we show the previous choice of partition (solid lines) with an alternative choice (dashed lines). The solid curve is the symmetry line of $\phi_1^{(S)}(C_1)$ of figure 6 and the symmetric part of its forward image $C_2 = F \circ S(C_1)$ as mentioned in the previous section. The dashed line is another possible choice constructed from the symmetry line of ϕ_{-2} plus its symmetric forward image. This other choice means that the symmetry lines intersect the set of HTs in different places. However, the extra HTs used in the lower part of the partition in figure 8(a) is exactly the symmetric pre-image of the HTs not needed in the upper part of the partition. We are therefore just shifting our definition of which tangencies are primary in a region where this notion is, as already noted,

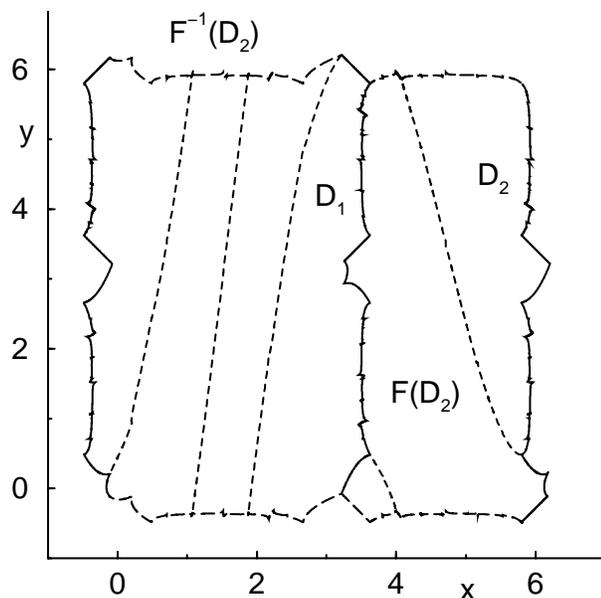


Figure 7. The generating partition. In analogy to figure 2, we have used the dividing line D_2 (solid) and its pre-image (dotted) to define the primary region of phase space. This is then partitioned by D_1 (solid) and by the image of D_2 (dashed).

ambiguous. Figure 8(b) is a magnification of the stable island that shows why the second choice of partition might be more useful. Inside the island we have plotted a number of periodic orbits, both stable (diamonds) and unstable (plusses). As can be seen, the symmetry line of $\phi_1^{(S)}$ goes through both stable and unstable orbits. The symmetry line of ϕ_{-2} , on the other hand, passes through only the stable orbits. In fact, it passes through a point of every stable periodic orbit that we have found inside the period-2 island. This indicates that we can use the same line that defines a winding number for the major island to also define a winding number on the subislands and that this definition can therefore distinguish orbits even in the subislands. One can then hope that this construct will continue through the hierarchy of sub-subislands. In order for this partition to also work successfully in the chaotic regions inside the island we would need to refine it by finding homoclinic tangencies for every chaotic layer. This is numerically a difficult task due to the very slow stretching and folding that takes place in the chaotic layers but could in principle be done. We will, however, note that by symmetry considerations the primary tangencies must lie approximately midway between points in the main unstable periodic orbit associated with the chaotic layer, i.e. where the symmetry line passes through the stable subislands.

8. The pruning front

A useful representation of the dynamics is obtained by constructing a piecewise-linear map with the same topological properties of the initial transformation. This can be done by first identifying the qualitative structure of the standard map and then by using the symbolic encoding to construct two suitable variables δ and γ the evolution of which keeps the relevant 'ordering' properties of the dynamics.

As a first step, we schematically represent the action of map (1) on the square S as

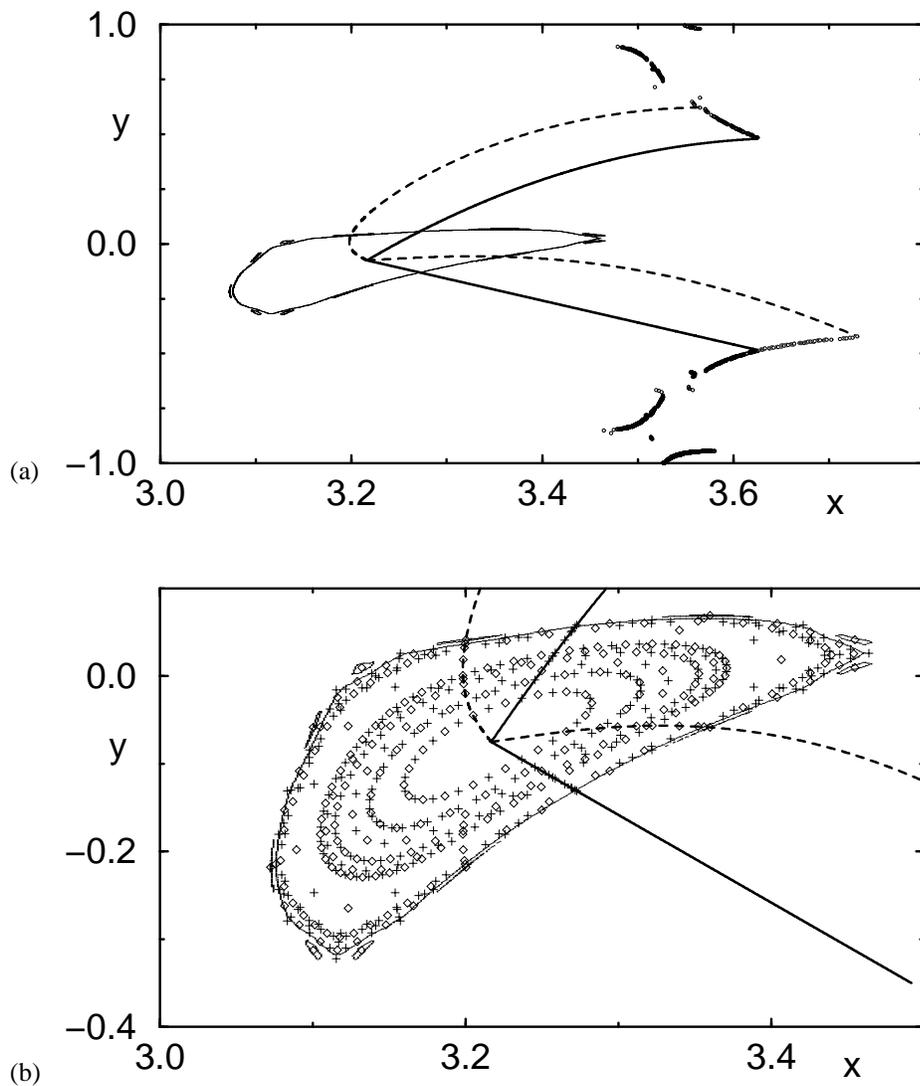


Figure 8. The stable period-2 island revisited. In (a) we show two possible choices for a partition: our original choice, C_1 - C_2 of figure 5 (solid lines) and curves found from (7) with $n =$ (dashed line). (b) is a magnification of the stable island including both stable (diamonds) and unstable (pluses) periodic orbits. It is seen that the dashed partition has the advantage that it crosses a stable island in every chain of subislands. This holds true for all of the approximately 50 subchains we have been able to locate.

described in figure 9. The reader can recognize in the 7 horizontal stripes A–G the pre-images of the 7 different regions present in figure 7 (apart from an irrelevant exchange of x and y directions). They are mapped into 5 different copies of the square in the order reported in the central part of the figure, with their local orientation given by the orientation of the corresponding letters. Finally, the equivalence among the copies of the squares allows interpretation of the action of the mapping as in the rightmost part of the figure.

According to the above picture, all symbols imply an expansion by a factor 7 (i.e. to

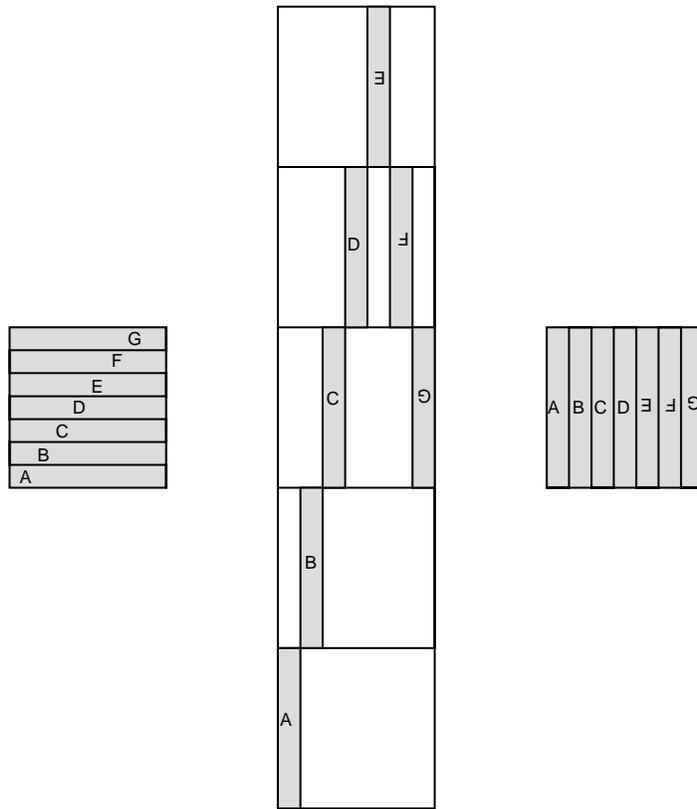


Figure 9. Schematic representation of the map F as a generalized baker's transformation. The unit square is divided into seven sections corresponding to the elements of the generating partition. The square is stretched and folded as shown in the intermediate stage and finally put back into the unit square by the modulus operation.

a shift of the corresponding variable, if represented in a 7-nary alphabet), but while $A-D$ (i.e., symbols $s = 0, 1, 2, 3$) preserve the vertical ordering, $E-G$ ($s = 4, 5, 6$) reverse it. Before defining γ_n , we need the auxiliary variable

$$a_n = \sum_{i \leq n} b_i \text{ mod } 2 \tag{9}$$

where $b_i = 0$ if $s_i \leq 3$, and equal to 1 otherwise. In words, b_n says whether the ordering is inverted or not at time n , while a_n conveys the absolute information about the local ordering. In this sense there is no difference with the standard horseshoe [4], where there are two symbols only, one of the two inverting the order. Finally, we obtain γ_n as,

$$\gamma_n = \sum_{j \geq n} c_j 7^{n-j} \tag{10}$$

where the integer c_j is equal to s_j , if a_j is zero (meaning that the direction has been changed an even number of times), while $s_j = 6 - s_j$, otherwise.

A very similar definition applies to δ_n , once the time direction has been changed and the sums run from -1 towards $-\infty$. Accordingly, any trajectory of the map can be first encoded in a sequence of symbols s_n and then transformed into a sequence of variables

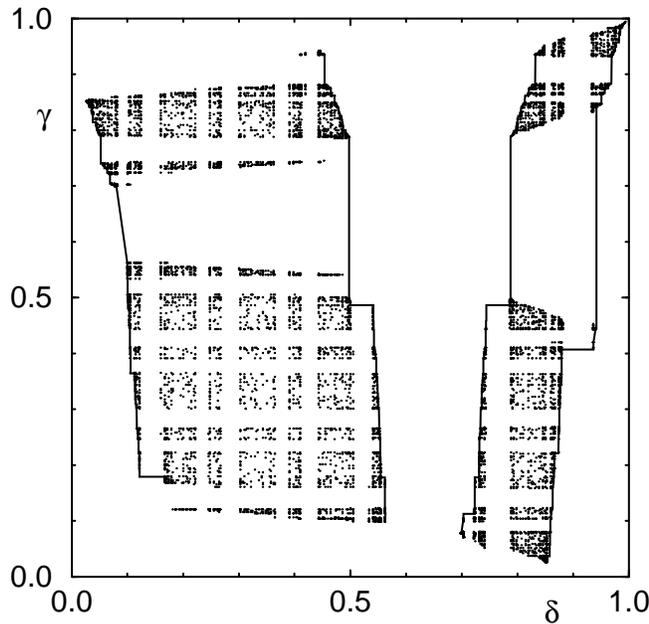


Figure 10. The symbolic plane with pruning fronts. The dots correspond to points in a long trajectory. In order to also include points from the stable islands which the chaotic trajectory cannot reach we have also used periodic orbit from inside the stable islands. The full lines are the pruning fronts: the symbolic representation of the generating partition. The generating partition is constructed from two dividing lines. For each of these one has the choice of assigning the symbol of the elements either to the left or to the right of the dividing lines, leading to in all four pruning fronts.

(γ_n, δ_n) bounded between 0 and 1, which provide a symbolic representation of the original dynamics. The result is reported in figure 10. One can recognize the typical fractal structure which arises because of the presence of forbidden sequences.

After reproducing the main chaotic component in the symbol plane, it is very important to obtain the coordinates γ , δ of the dividing lines, as they represent the so called pruning front, a continuous line which allows determining the forbidden sequences [4] and, in turn, to study, the complexity of the dynamical system [18]. The solid lines in figure 10 represent our front, which indeed is seen to separate the primary forbidden regions. In fact, we should remind the reader that once a region in the symbol plane is recognized as forbidden, its forward and backward images are forbidden as well (this is the reason for observing a fractal structure). Hence, it is obviously sufficient to locate one such region for each forbidden sequence in order to be able to provide a complete characterization of the dynamical system.

To check that the pruning front does indeed determine the forbidden region we have found all periodic orbits up to length 9 (approximately 30 000 orbits) and their corresponding symbolic sequences. They were all inside the two regions bounded by the four pruning fronts in figure 8. The same check was made with a long trajectory (100 000 points).

An important property of the pruning front which is invoked as a guarantee for determining the forbidden sequences is its monotonicity [4]. We have checked that this is indeed true also in our case (see figure 11, for an enlargement of one such piece). This further evidence of consistency with previous results about the dynamics in the symbol plane can be considered as both a confirmation of the correctness of our approach and of

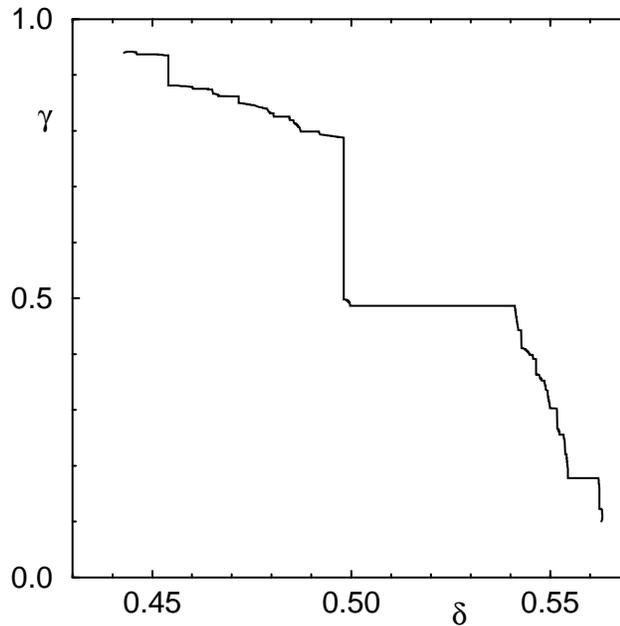


Figure 11. A magnification of one of the pruning fronts of figure 8. One can see both the monotonicity of the pruning front and the symmetry which is related to the symmetry of the dividing line under the transformation $F \circ T$.

the monotonicity of the pruning front which has been rigorously proved in very few cases (see [16] for the proof in the Lozi map).

Among the novel features of the pruning front reported in figure 11, with respect to examples previously discussed in the literature, notice that it includes pieces of symmetry lines and accounts for the quasiperiodic behaviour in the stability islands. More precisely, each island corresponds to a specific region in the symbol plane where the lack of chaos, presumably induces a nearly vanishing fractal dimension.

9. Concluding remarks

In view of a statement about the general validity of our approach, we still need to address the question of what happens upon changing the control parameter. Here, we will take a brief look at the effect of a period doubling bifurcation of the stable period-2 island. When the parameter α is increased to 2π , the stable period-2 orbit turns unstable and a bifurcation spawns a new stable period-4 orbit. At the moment of creation, the partition constructed from pieces of the symmetry lines of ϕ_{-2} touches the period 4 islands, but for higher values of α they do not enter the islands. A kind of bifurcation of the partition therefore also takes place. The simplest connection we have been able to make in the gap of HTs going through one of the islands of the period 4 orbit is shown in figure 12. It consists of two pieces of the symmetry line of ϕ_{-3} and two corresponding pieces of the symmetry line of ϕ_5 that are 4th forward iterates of the former.

A similar thing happens in the major central gap where, also at $\alpha = 2\pi$, a tangent bifurcation gives rise to a new pair of fixed points; one stable, the other unstable. The two symmetry lines bridging the gap meet at the point of the tangent bifurcation, whereas the

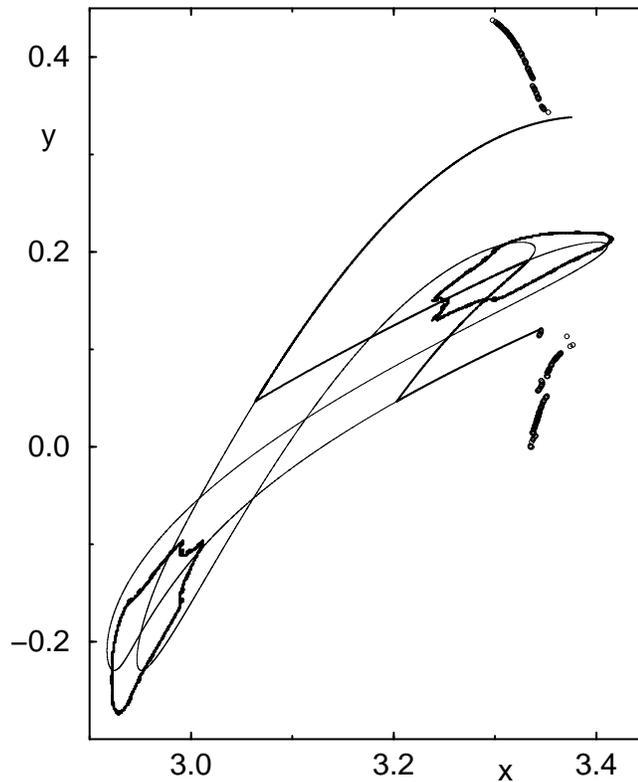


Figure 12. A section of the partition for $\alpha = 6.4$. At $\alpha = 2\pi$ the stable period-2 island goes through a period doubling bifurcation creating a stable period-4 orbit of which two of the islands are shown here. The pieces of fixed curves that connect different portions of the HTs likewise undergo something similar to period doubling. Instead of using a piece of the curve found from (7) with $n = 4$ we use 2 pieces of the curve with $n = 4$ plus their 8th forward image.

small set of HTs inbetween disappear. After the bifurcation, the two symmetry lines meet at the central fixed point of the newly created stable island.

By recalling the success of methods based on the identification of HTs in several dissipative systems, and noticing that we have not exploited any specific feature of the standard map [19] it appears plausible to conjecture that, via this method, a generating partition can be constructed in generic Hamiltonian systems as well. The procedure is not straightforward in the sense that several trials with errors are typically required, before converging to the asymptotic result. However, there is convincing evidence that all possible problems arising during the construction can be eventually solved. Obviously, it would be much nicer, if all the questions could be solved *a priori*, with a general algorithm. The eventual complete success of this strategy crucially depends on the possibility to give a more formal definition of primary homoclinic tangencies, and on a better understanding of the dynamics in the vicinity of stable islands. In particular, a problem that even from a heuristic point is not completely solved, concerns the observation that primary tangencies appear to end up precisely in correspondence of the symmetry lines. We have no explanation for this nice phenomenon which is the key ingredient that allows reconciling under the same approach both the chaotic and ordered component.

Acknowledgments

FC acknowledges support from the European Community Human Capital and Mobility program, grant No ERBCHICT920122 and The Max Planck Institute for Complex Systems, Dresden. Both authors wish to thank ISI-Torino for the hospitality which allowed completion of this work.

References

- [1] Sinai Ya G 1972 *Russ. Math. Surv.* **27** 21
Ruelle D 1978 *Thermodynamic Formalism* (Reading, MA : Addison-Wesley)
- [2] Politi A 1994 *From Statistical Physics to Statistical Inference and Back* ed J-P Nadal and P Grassberger (Dordrecht: Kluwer) p 293
- [3] Grassberger P and Kantz H 1985 *Phys. Lett.* **113A** 235
- [4] Cvitanović P, Gunaratne G H and Procaccia I 1988 *Phys. Rev. A* **38** 1503
- [5] Hopcroft H E and Ullman J D 1979 *Introduction to Automata Theory, Languages and Computation* (Reading, MA: Addison-Wesley)
- [6] D'Alessandro G, Grassberger P, Isola S and Politi A 1990 *J. Phys. A: Math. Gen.* **23** 5285
- [7] Giovannini F and Politi A 1991 *J. Phys. A: Math. Gen.* **24** 1837
- [8] Christiansen F and Politi A 1995 *Phys. Rev. E* **51** R3811
- [9] Within a stability window, the partition turns out to be generating only for the set of all unstable periodic orbits, i.e. for the dynamics on the strange repeller; see also Collet P and Eckmann J-P 1980 *Iterated Maps on the Interval as Dynamical Systems* (Boston: Birkhäuser)
- [10] Grassberger P, Kantz H and Moenig U 1989 *J. Phys. A: Math. Gen.* **43** 5217
- [11] By fibre, here and in the following we mean a small piece of manifold.
- [12] For a detailed description of the procedure to determine the coefficients, see [7].
- [13] Giovannini F and Politi A 1992 *Phys. Lett. A* **161** 332
- [14] Hansen K 1992 *Phys. Lett. A* **165** 100
- [15] Richter P H, Scholz H-J and Wittek A 1990 *Nonlinearity* **3** 45
- [16] D'Alessandro G, Isola S and Politi A 1991 *Prog. Theor. Phys.* **86** 1149
- [17] Ruelle D and Eckmann J-P 1985 *Rev. Mod. Phys.* **57** 617
- [18] Cvetkovic D, Doob M and Sachs H 1980 *Spectra of Graphs* (New York: Academic)
- [19] Apart from its symmetry properties of which only the time reversal symmetry is of fundamental importance.
In fact, any spatial symmetry merely adds on more features to the symmetry line treatment.