

A method for visualization of invariant sets of dynamical systems based on the ergodic partition

Igor Mezić^{a)}

Department of Mechanical and Environmental Engineering and Department of Mathematics,
University of California, Santa Barbara, California 91306-5070

Stephen Wiggins

Control and Dynamical Systems 107-81, Caltech, Pasadena, California 91125

(Received 13 April 1998; accepted for publication 29 October 1998)

We provide an algorithm for visualization of invariant sets of dynamical systems with a smooth invariant measure. The algorithm is based on a constructive proof of the ergodic partition theorem for automorphisms of compact metric spaces. The ergodic partition of a compact metric space A , under the dynamics of a continuous automorphism T , is shown to be the product of measurable partitions of the space induced by the time averages of a set of functions on A . The numerical algorithm consists of computing the time averages of a chosen set of functions and partitioning the phase space into their level sets. The method is applied to the three-dimensional ABC map for which the dynamics was visualized by other methods in Feingold *et al.* [J. Stat. Phys. **50**, 529 (1988)].
© 1999 American Institute of Physics. [S1054-1500(99)00801-0]

In this paper we address the question of the visualization of phase space of measure-preserving dynamical systems. For the purpose of visualizing the phase space of dimension bigger or equal to three, it would be useful to provide a method that color-codes invariant sets of the system. Then, two-dimensional slices through the phase space can be taken. A question arises: there are many invariant sets in measure-preserving dynamical systems—which are the ones that we want to identify? We visualize the invariant sets on which the dynamical system is ergodic. Our method is based on the notion of ergodic partition. The computation involves choosing a set of functions and calculating their time-averages along the orbits. The invariant ergodic sets are then identified as those on which the time-averages of these functions are constant.

I. INTRODUCTION

In this paper we consider the issue of the visualization of invariant sets of dynamical systems possessing a smooth invariant measure. Probably the most widespread method for visualizing phase space structure is simply plotting the trajectories associated with as many initial conditions, and for as long a time, as possible. Another approach is that of constructing *exit times plots* [see, e.g., Easton *et al.* (1993)].

Both of these methods have certain shortcomings. The method of plotting trajectories for visualization runs into difficulties when the phase space has a dimension equal to or larger than three. There have been several strategies to overcome this difficulty. For example, Feingold *et al.* (1988) have used the method of recording the location of the orbit in thin slices of the three-dimensional phase space of the ABC

map. A common method for the visualization of phase space structure in four-dimensional dynamical systems involves the color-coding of invariant surfaces (in particular, of stable and unstable manifolds of certain phase-space objects) according to the magnitude of the fourth coordinate [see, e.g., Haller and Wiggins (1995)]. The method of plotting exit times can be applied for dynamical systems of arbitrary dimension, as the phase-space is color coded with a particular color assigned depending on the time it takes to get out of the prescribed, bounded set in the phase space (see Easton *et al.*, 1993). The shortcoming of exit times plots is that it is not generally true that the points that are assigned the same color (i.e., for which it takes an equal amount of time to exit the bounded set) are, in fact, in the same invariant set for the dynamics.

Here we suggest a different approach that includes the features of both of these methods and is based on some ergodic theory concepts. We consider partitions of the phase space A , on which a dynamical system T that preserves a smooth measure μ on A acts, into invariant sets induced by functions that are constant on the orbits of the dynamical system. To identify the invariants of motion, we consider a set of functions f_i , $i \in N$ such that their linear combinations are dense in L^1 . If T is a map, we define the associated set of functions f_i^* by

$$f_i^* = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} f_i \circ T^j,$$

so that f_i^* is the time average of f under T , T^j being the j -th iterate of T . Now the sets $C_v = \{x \in A | f_i^*(x) = v\}$ are invariant under the dynamics of T . Thus, color-coding the phase space according to the value of v provides a partition of the phase space into invariant sets. However, there are many invariant sets for a measure-preserving dynamical system and the following question arises: which invariant sets is the

^{a)}Electronic mail: mezić@engineering.ucsb.edu

method identifying? The invariant sets in our partition of the phase space are not just any invariant sets. Rather, they are invariant sets on which the dynamics is ergodic. In an ergodic invariant set all points are “accessible,” in the sense that the initial conditions in that set thoroughly sample the set. Hence, one could view the ergodic partition as a partition as a partition of the phase space into accessible regions. This point of view is very important, and very natural, for many applications. For example, in the dynamical systems approach to transport in fluid flows the collection of ergodic components would provide a “road map” for the transport. Since the different invariant sets that are identified are ergodic, a statistical theory of mixing on each component would then follow immediately. In molecular dynamics one is interested in the exchange of energy or angular momentum in phase space between different modes of rotation and vibration. In this setting the ergodic components show the regions where such exchanges can, and cannot, take place. Moreover, molecular dynamics is an area of application where higher-dimensional visualization techniques are a necessity. The method developed in this paper provides direct way of addressing important issues arising in both of these areas of applications (as well as others), that are not immediately answered by other visualization methodologies. Of course, computational issues are very important. The numerical results that we present here show that even using a small number of functions f_i and short computational time, a good visual description of the phase space is obtained.

In this sense, the method retains good properties both of the method of plotting directly the trajectories of the system, as every set in the partition consists of the orbits of the system, and of the exit time plot: because of color-coding, the constructed partition can be sliced in various ways (in particular, by two-dimensional surfaces) to provide information about the invariant sets. The most “detailed” (with respect to the dynamics) of such partitions into invariant sets is the *ergodic partition* which is a partition of the phase space into sets on which the map is ergodic, and the ergodic measures on these sets are consistent with the smooth measure that the system preserves. We provide a numerical algorithm for calculating the ergodic partition based on the theorem that the ergodic partition is equivalent to the product of invariant partitions induced by the time averages of a countable set of functions f_i that we prove in Sec. III. That this theorem could be used in visualization was suggested in Mezić (1994). Von Milczewski *et al.* (1996) have also plotted average quantities to determine phase-space structure. Dellnitz *et al.* (1997) have developed visualization methods for dissipative systems based on a discretization of the Frobenius–Perron operator. In Sec. IV we discuss the numerical algorithm and its application to the ABC map [Feingold *et al.* (1988)]. The ABC map is a map on a three-dimensional torus that was studied to provide an example of the application of our algorithm to maps with a smooth invariant measure in dimensions bigger than two, where plotting the orbits of a map on a screen often results in a scatter of points that is not convenient for analysis. We chose the ABC map because we could compare our results with the visualization of Feingold *et al.* (1988). The numerical results

are somewhat preliminary and much of research is needed to improve the efficiency of computation and investigate its properties in detail. Still, the computations that we present show definite promise.

II. SET UP AND DEFINITIONS

Let A be a compact metric space, μ a complete measure defined on the Borel σ -algebra \mathcal{A} , and T a measure-preserving continuous automorphism on A . We denote the space of all integrable functions on A by $L^1(A)$, the set of all real continuous functions on A by $C(A)$, and a dense countable subset of $C(A)$ by S . We introduce the notions of partition and measurable partition in the following definition.

Definition 2.1 (partition, measurable partition): A family ζ of disjoint sets C_α (α an element of some, arbitrary, indexing set) whose union is all of A is called a partition of A . We call C_α an element of ζ . Unions of elements of ζ are called ζ -sets. A partition is called measurable if there exists a countable family Δ of measurable sets $\{D_i\}$ such that every D_i is a union of elements of ζ , and for every pair C_1, C_2 of elements of ζ there exists $D_j \in \Delta$ such that $C_1 \subset D_j$, $C_2 \subset D_j^c$. We call Δ a basis for the partition ζ .

We will omit the index from the notation for the element of the partition, where that does not cause confusion. A partition ζ_1 is said to be finer than a partition ζ_2 if every component of ζ_2 is a ζ_1 -set. We now define a *product* operation on the set of measurable partitions of A .

Definition 2.2 (product): Let ζ_1 and ζ_2 be two measurable partitions. Denote the element of ζ_1 by C_1 and the element of ζ_2 by C_2 . Let ζ be the family of all sets of the form $C = C_1 \cap C_2$. Then we call the measurable partition ζ a product of ζ_1 and ζ_2 , and write $\zeta = \zeta_1 \vee \zeta_2$. For a finite or countable product, we may write $\zeta = \bigvee_{i=1}^n \zeta_i$, n finite or ∞ .

Now we turn to the construction of ergodic partitions.

III. ERGODIC PARTITIONS CONSTRUCTIVELY

We start with the following definitions.

Definition 3.1 (stationary partition, ergodic partition): A measurable partition ζ is called stationary if every element of ζ is invariant under T . The stationary partition ζ is called ergodic if, for almost every (with respect to μ) element C of ζ , there is an invariant measure μ_C on C such that the restriction of T to C , denoted T_C is an ergodic automorphism on C , with respect to some probability measure μ_C on C , and, for every $f \in L^1(A)$,

$$\int_A f d\mu = \int_A \left[\int_C f|_C d\mu_C \right] d\mu, \quad (1)$$

where $f|_C$ denotes the restriction of f to the ergodic component C .

The notion of the ergodic partition of automorphisms of Lebesgue spaces has its origin in the works of von Neumann (1932); Halmos (1941, 1949); and Rokhlin (1949). Note that the term *ergodic decomposition* is used usually when the decomposed (or partitioned) object is a measure. In this work we are mainly concerned with the partition of the space A into disjoint ergodic components [or ergodic fibers, as in Denker *et al.* (1976)]. As this is more of a topological sub-

ject matter, the term *partition* would be more appropriate. In the above definition, integration with respect to μ is performed so that to every $x \in A$ we associate $C(x)$, the element of the ergodic partition to which x belongs. We shall call the function f^* the time average of a function f under T if

$$f^*(x) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x)),$$

almost everywhere (a.e.) with respect to μ . Note that by Birkhoff's pointwise ergodic theorem, f^* exists for every function $f \in L^1(A)$. Denote by Σ the set of all $x \in A$ such that f^* exists for every $f \in C(A)$, and by $\Sigma(f)$ the set of all $x \in A$ such that f^* exists for a particular $f \in C(A)$. The following lemma is standard [see Mañé (1987)].

Lemma 3.1:

$$\Sigma = \bigcap_{f \in S} \Sigma(f),$$

where S is some dense, countable subset of $C(A)$.

Now we have a set Σ such that the time averages of all continuous functions on A are well-defined on Σ . Σ^c is of measure zero, as by the Birkhoff's Ergodic Theorem each $\Sigma(f)^c$ is of measure zero, and thus Σ^c is a countable union of measure zero sets, which is again of measure zero. The lemma shows that the time averages of measurable, bounded functions induce measurable partitions on Σ .

Lemma 3.2: Let f be a measurable, bounded function on Σ . The family of sets C_α , $\alpha \in \mathbb{R}$ such that $C_\alpha = (f^*)^{-1}(\alpha)$ is a measurable partition of Σ . We denote this partition by ζ_f and call it the partition induced by f .

Proof: As $F_n = (1/n) \sum_{i=0}^{n-1} f(T^i(x))$ is a measurable function on A , and $\lim_{n \rightarrow \infty} F_n = f^*$, f^* is a measurable function, too. The fact that ζ_f is a measurable partition follows by taking Δ_f to be the collection of preimages under f^* of open intervals with rational endpoints in \mathbb{R} . As f^* is measurable, each $(f^*)^{-1}[a, b]$ is measurable. Sets of this type, where a and b are rational, clearly separate sets of the form $(f^*)^{-1}\{c\}$, $c \in \mathbb{R}$. \square

Now we can prove our main theorem.

Theorem 3.1: Let

$$\zeta_e = \vee_{f \in S} \zeta_f,$$

be a measurable partition on Σ . Then ζ_e together with Σ^c is the ergodic partition of A with respect to T .

Proof: Let C be an element of ζ_e . Define a linear functional L_C on $C(A)$ by

$$L_C(f) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x)), \quad \text{where } x \in C.$$

L_C is well-defined, as the right hand side is the same for any $x \in C$. As $L_C(1) = 1$, by Riesz's representation theorem, there exists a measure μ_C on A , such that

$$L_C(f) = \int_A f d\mu_C.$$

Note that μ_C is invariant under T . To prove this, we have, for every continuous f ,

$$\int_A f \circ T d\mu_C = L_C(f \circ T) = L_C(f) = \int_A f d\mu_C. \tag{2}$$

For the above operation we needed continuity of T . As continuous functions are dense in L^1 , μ_C is invariant under T .

Now we show that μ_C is a probabilistic measure on C . There is a sequence of compact sets C_n^c , subsets of C^c , such that

$$C_1^c \subset C_2^c \subset \dots,$$

and

$$\mu_C(C^c \cup \bigcup_{n \geq 1} C_n^c) = 0. \tag{3}$$

Further, we can show that for every C_n^c , $\mu_C(C_n^c) = 0$. To do this, note that, by Urysohn's Lemma, for every C_n^c there is a continuous, positive function f_n equal to one on C_n^c , and equal to zero outside of C_{n+1}^c . Clearly, as C_i^c 's are subsets of C^c , $f_n = 0$ on C . Therefore,

$$\mu_C(C_n^c) \leq \int_A f_n d\mu_C = L_C(f_n) = 0 \Rightarrow \mu_C(C_n^c) = 0.$$

As the measure of a union of the countable number of sets of measure zero is zero,

$$\mu_C(\bigcup_{n \geq 1} C_n^c) = 0. \tag{4}$$

Therefore, by (3) and (4), $\mu_C(C^c) = 0$, and we are done with the proof of the fact that μ_C is a probability measure on C .

We have shown that μ_C is a probability measure when restricted to C . Let us show that it is an ergodic measure on C . Observe that the set of all restrictions of functions in $C(A)$ to C , denoted $C(A)|_C$ forms a dense set in the set of all integrable (with respect to μ_C) functions on C . To show this, note that $C(A)$ is dense in $L^1(A)$. Let f be an element of $L^1(C)$. Consider its extension to all of A , \bar{f} , such that $\bar{f} = f$ on C and $\bar{f} = 0$ elsewhere. As

$$\int_A \bar{f} d\mu_C = \int_C f d\mu_C,$$

exists, $\bar{f} \in L^1_{\mu_C}(A)$. Therefore, there is a sequence of functions in $C(A)$, $\{f_n\}$ converging to \bar{f} . But then the corresponding sequence of restrictions, $\{(f_n)_C\}$ converges to f .

Now, for every $f \in C(A)|_C$, we have

$$f^*(x) = \int_C f d\mu_C. \tag{5}$$

Recalling (Mañé (1987), Chap. II, Proposition 2.2) that if (5) holds for a dense subset of $L^1(C)$, then T_C is ergodic, this completes the proof that μ_C is an ergodic measure, for every C .

What is left is to observe that the equality (1) is proven in Mañé (1987), Chap. II, Theorem 6.4. So, we are done with the proof of the theorem. \square

As we already mentioned, our interest in the problem of ergodic partitioning of the phase space stems from the fact that the concept of an ergodic partition can be used as a visualization tool. But, it is not always necessary to try to construct numerically the ergodic partition, which would

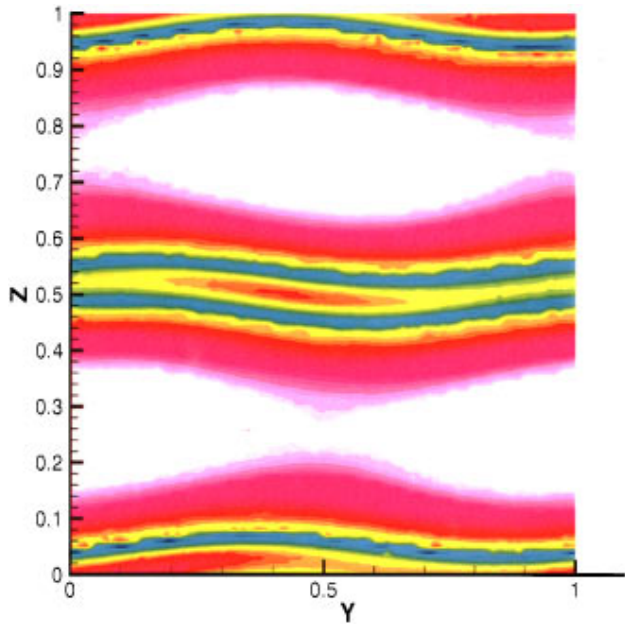


FIG. 1. Contour plots for g .

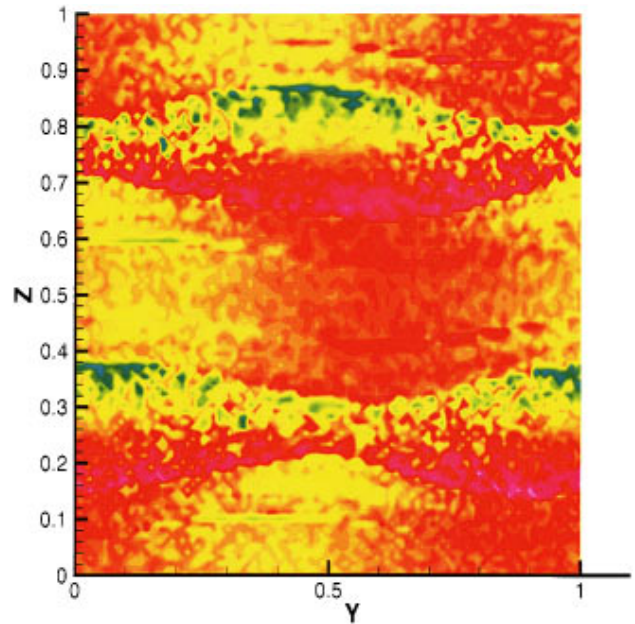


FIG. 3. Contour plot for the time average of $\sin y$.

consist of computing the time averages of a countably infinite number of functions. Invariant sets can be detected by computing the time averages of just one function, and approximating the partition ζ_f .

IV. ALGORITHM AND NUMERICAL ASPECTS

Based on the construction of the ergodic partition described above, we can approximate the ergodic partition numerically.

The basic idea of the algorithm is as follows.

- Start with a regular grid of initial conditions over the phase space, A ; one could choose less initial points in

the chaotic region, as we know that the time averages are constant on orbits. Thus, one initial condition can provide a lot of data in the calculation. We did not implement that possibility here.

- Find the finite time averages of the (finite) set of functions, $\{e_j\}_{j=1}^M$, whose partitions are used, at each point of the grid.
- Once these averages have been obtained, we must graphically represent the partition. We examine two-dimensional slices of these (potentially) multidimensional data sets. We form a norm of the finite time averages,

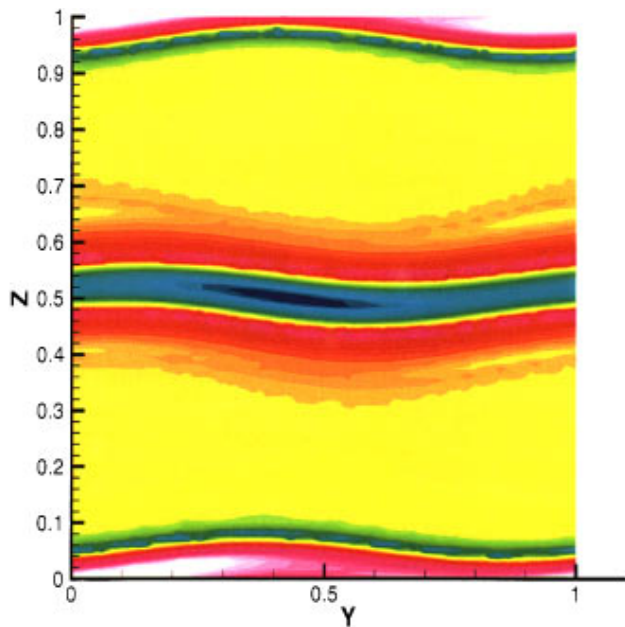


FIG. 2. Contour plot for the time average of $\sin x$.

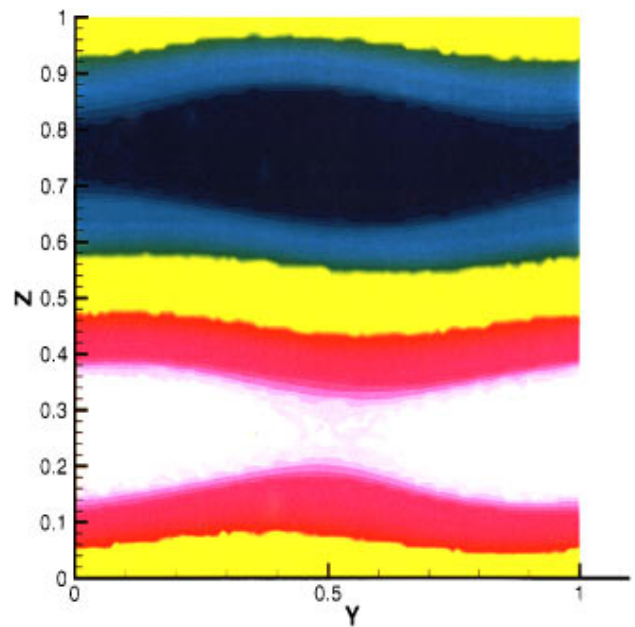


FIG. 4. Contour plot for the time average of $\sin z$.

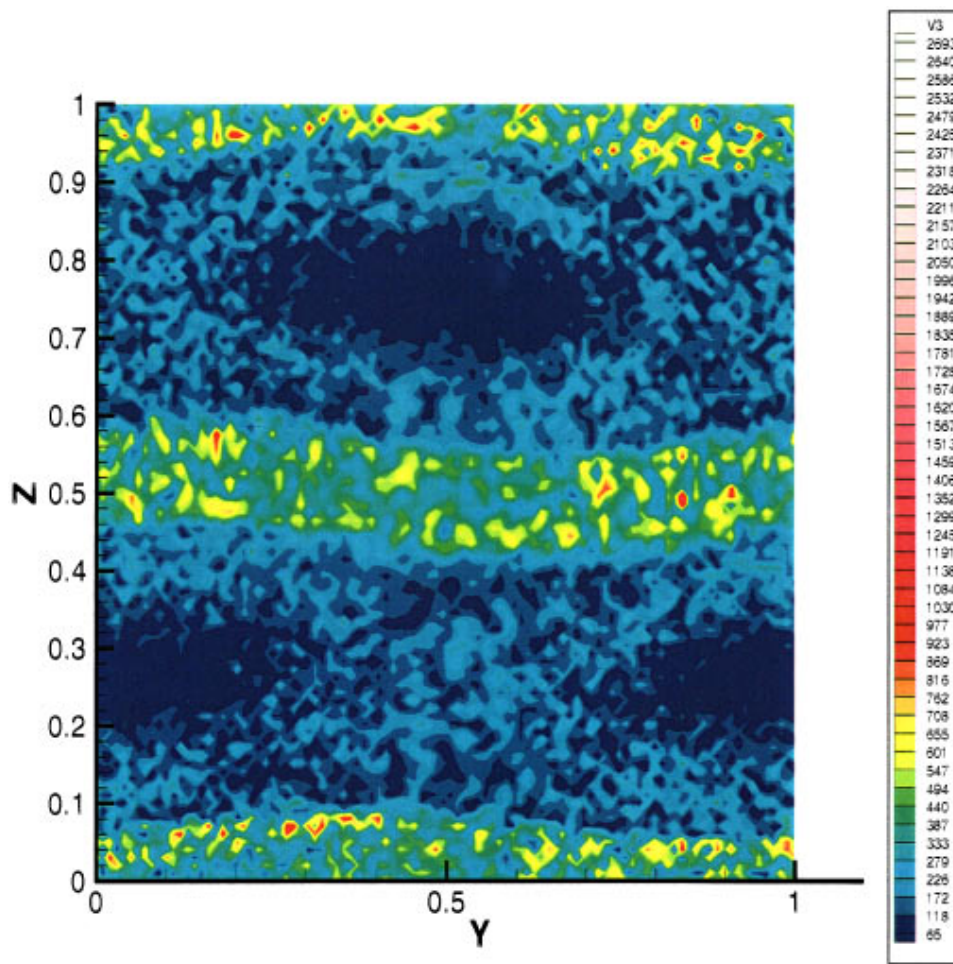


FIG. 5. The convergence of time averages.

$$g = \sqrt{\sum_{i=1}^N (e_i^{*fin})^2},$$

and form a density plot of this function.

It is well-known that there is no uniformity in the convergence of the time-averages [Petersen (1983)]. Questions regarding the number of iterations needed to numerically obtain ergodicity on a component (not relating to the function which is being averaged) are discussed by Meiss (1994). In his study, he looks at the behavior of one orbit in a positive area ergodic component of the Standard Map and finds that only after times as long as 10^{10} iterations does the orbit closely exhibit ergodic behavior. Given the number of initial conditions that we use, that is much too long a time to employ in the method described here. On the other hand, the success of the method in outlining the rough features of invariant ergodic regions within 1000 iterates for most of the points in our example described below (see the figures in comparison with the results existing in the literature; in particular, Fig. 5 has a legend that indicates the number of iterates until the calculation has converged) indicates that although the convergence might not be very good, different components will exhibit sufficiently different finite time averages.

A. The ABC map

The ABC map was introduced by Feingold, Kadanoff and Piro (1988) in order to investigate the dynamical properties of three-dimensional, volume-preserving maps. It is given by

$$\begin{aligned} x' &= x + A \sin z + C \cos y, \\ y' &= y + B \sin x' + A \cos z, \\ z' &= z + C \sin y' + B \cos x'. \end{aligned} \tag{6}$$

We have performed calculations of finite time averages along trajectories for functions $\sin x$, $\sin y$ and $\sin z$ starting from 10000 initial conditions (y, z) in the plane $x=0.01$ on the uniform grid, for parameter values $A=1.5$, $B=0.08$, $C=0.16$. The computation starting from a particular initial condition was stopped when the relative difference between the value of the average at the previous time step and the current time average was less than 10^{-6} . The function g was defined, as above, to be

$$g(y, z) = \sqrt{[\sin^{*fin}(x)]^2 + [\sin^{*fin}(y)]^2 + [\sin^{*fin}(z)]^2}.$$

In the above definition,

$$\sin^{*fin}(x) = \frac{1}{n} \sum_{i=0}^{N(y,z)-1} \sin(x_i),$$

and $N(y,z)$ is the number of iterates it takes for the finite time average of $\sin x$ to converge starting from the point (y,z) in the $x=0.01$ plane. Similar expressions are used for $\sin^{*fin}(y)$ and $\sin^{*fin}(z)$. In Fig. 1 we show the density plot of the function g in the plane $x=0.01$. This figure should be compared with an equivalent, Fig. 2 in Feingold *et al.* Figures 2, 3, and 4 show the density plots of the time averages of $\sin x$, $\sin y$ and $\sin z$, respectively. It is seen that different functions contribute different features to the density plot of g . Figure 5 shows the time to convergence of g to within 10^{-6} . Less than 10^3 iterates is sufficient for most of the initial conditions to converge. The computation took a couple of minutes of computing time on a workstation.

V. CONCLUSIONS

In this paper we have suggested a method for the visualization of dynamical systems that preserve a smooth invariant measure which should be useful for the detection of ergodic invariant sets—especially for systems with high dimensionality. Previously used methods for visualization were the following:

- Plotting of orbits, which can be used effectively only in two-dimensional systems—already in three dimensions we are forced to project to two-dimensional planes of the computer screen and paper.
- Exit time plots. There is no guarantee that the region of the constant exit time is an invariant region.

Our algorithm consists of making a density or contour plot of the time averages of functions in a two-dimensional slice of the phase space. The algorithm is backed up by rigorous results that we proved in the first part of the paper. We have shown that ergodic partition can be constructed out of sets on which the time-averages of a dense, countable set of functions in the space of all continuous functions are constant.

In the numerical implementation of the algorithm we can take only finite-time averages. The questions of convergence thus come into question. In all the cases we considered, coarse properties of invariant sets are detected in a reasonable number of iterations. In some sense, our algorithm can be considered as a finite-time measurement of the system in which we are interested in the average properties. We have tested the method on the three-dimensional, volume-

preserving ABC map introduced by Feingold *et al.* and shown that our method of visualization produces good contrast in reasonable iteration time.

There are many issues that we only touched upon here, like the effect of dimensionality and the variation of functions on the convergence of the time averages, the choice of functions, the numerical implementation of the product operation, etc. These deserve further investigation and need to be clarified if the method is to become useful.

ACKNOWLEDGMENTS

We would like to thank Jean-Christophe Nave of the University of California, Santa Barbara for coding the program to compute the partition and producing the figures. This research was partially supported by ONR Grant No. N00014-98-1-0056, AFOSR Grant No. F49620-98-1-0146, and National Science Foundation Grant No. DMS-9803555 to IM and ONR Grant No. N00014-97-1-0071 to SW.

- Dellnitz, M., Hohmann, A., Junge, O., and Rumpf, M. (1997). "Exploring invariant sets and invariant measures," *Chaos* **7**, 221–228.
- Denker, M., Grillenberger, C., and Sigmund, K. (1976). *Ergodic Theory on Compact Spaces*, Springer Lecture Notes in Mathematics, Vol. 527 (Springer-Verlag, New York).
- Easton, R. W., Meiss, J. D., and Carver, S. (1993). "Exit times and transport for symplectic twist maps," *Chaos* **3**, 153–165.
- Feingold, M., Kadanoff, L. P., and Piro, O. (1988). "Passive Scalars, three-dimensional volume-preserving maps and chaos," *J. Stat. Phys.* **50**, 529–565.
- Haller, G., and Wiggins, S. (1995). "Multi-pulse jumping orbits and homoclinic trees in truncation of the damped-forced nonlinear Schrödinger equation," *Physica D* **85**.
- Halmos, P. R. (1941). "The decomposition of measures," *Duke Math. J.* **8**, 386–392.
- Halmos, P. R. (1949). "On the theorem of Dieudonné," *Proc. Natl. Acad. Sci. USA* **35**, 38–42.
- Mañé, R. (1987). *Ergodic Theory and Differentiable Dynamics* (Springer, New York).
- Meiss, J. D. (1994). "Transient measures in the standard map," *Physica D* **74**, 254–267.
- Mezić, I. (1994). "On the geometrical and statistical properties of dynamical systems: Theory and applications," Caltech Ph.D. thesis.
- Von Milczewski, J., Dierksen, G. H. F., and Uzer, T. (1996). "Computation of the Arnol'd Web for the hydrogen atom in crossed electric and magnetic fields," *Phys. Rev. Lett.* **76**, 2890–2893.
- von Neumann, J. (1932). "Zur operatorenmethode in der klassischen mechanik," *Ann. Math.* **33**, 587–642.
- Petersen, K. (1983). *Ergodic Theory* (Cambridge University Press, Cambridge).
- Rokhlin, V. A. (1949). "On the fundamental ideas of measure theory," *Mat. Sb.* **25**, 107–150 [(1962). *Am. Math. Soc. Trans. Ser. 1* **10**, 1–54]; (1949). "Selected topics in the metric theory of dynamical systems," *Usp. Mat. Nauk* **4**, 57–128 [(1966). *Am. Math. Soc. Trans. Ser. 2* **49**, 171–240].
- Rudin, W. (1987). *Real and Complex Analysis* (McGraw-Hill, New York).