INVARIANT SETS AND MEASURES: A SHORT SURVEY

ASSYR ABDULLE* AND GIACOMO GAREGNANI†

Abstract.

1. Introduction.

2. Notation. Remark: the following notations and definitions summarize [5, 2, 3]. Let us consider the function $f\colon \mathbb{R}^d \to \mathbb{R}^d$ and the autonomous ordinary differential equation (ODE)

$$\begin{aligned} y'(t) &= f(y(t)), \quad t > 0, \\ y(0) &= y_0 \in \mathbb{R}^d. \end{aligned} \tag{1}$$

We consider f to be Lipschitz continuous so that the solution of (1) exists and we can define a family of mappings $\varphi_t\colon \mathbb{R}^d \to \mathbb{R}^d$, for $t \geq 0$, such that

$$y(t) = \varphi_t(y_0). \tag{2}$$

Let us denote by $\mathcal{B}$ the σ-algebra of Borel sets in $\mathbb{R}^d$. For any set $B \in \mathcal{B}$, we denote then by $\varphi_t^{-1}(B)$ the counterimage of B through φ_t, i.e.,

$$\varphi_t^{-1}(B) = \{x \in \mathbb{R}^d \colon \varphi_t(x) \in B\}. \tag{3}$$

Let us furthermore remark that if f is Lipschitz, the mapping φ_t is measurable, i.e., $\varphi_t^{-1}(B) \in \mathcal{B}$ for all Borel sets B. We now give some definitions and introduce notations useful for the following of our analysis.

Definition 2.1. *A set $B \in \mathcal{B}$ is invariant under the family $\{\varphi_t\}_{t\geq 0}$ if*

$$\varphi_t^{-1}(B) = B. \tag{4}$$

We now denote by $\mathcal{M}$ the space of measures on $\mathbb{R}^d$.

Definition 2.2. *A measure $\mu \in \mathcal{M}$ on $(\mathbb{R}^d, \mathcal{B})$ is said invariant for the dynamical system $\{\varphi_t\}_{t\geq 0}$ if for all $B \in \mathcal{B}$ it holds*

$$\mu(\varphi_t^{-1}(B)) = \mu(B). \tag{5}$$

Definition 2.3. *The operator $P\colon \mathcal{M} \to \mathcal{M}$ such that for all $B \in \mathcal{B}$ and $t \geq 0$*

$$(P\mu)(B) = \mu(\varphi_t^{-1}(B)) \tag{6}$$

is called the Frobenius-Perron operator. If $\{\varphi_t\}_{t\geq 0}$ admits an invariant measure μ, then it is trivially a fixed point of P, i.e., $P\mu = \mu$.

Given the definition of invariant sets and measure we can introduce the notion of ergodicity of a dynamical system.

Definition 2.4. *Given the measure space $(\mathbb{R}^d, \mathcal{B}, \mu)$, a dynamical system $\{\varphi_t\}_{t\geq 0}$ is ergodic if for any invariant set $B \in \mathcal{B}$ it holds $\mu(B) = 0$ or $\mu(\mathbb{R}^d \setminus B) = 0$.*

3. Exploring invariant sets and measures. We present two algorithms to explore the invariant set and measure of (1).

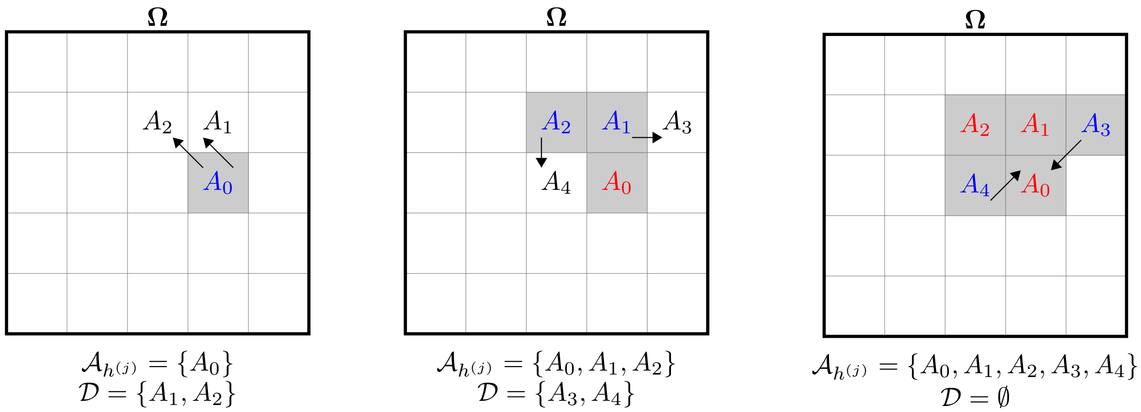


Fig. 1: Graphical representation of the subdivision algorithm in a two-dimensional domain Ω under assumption 3.2. Fleshes indicate the action of the map φ_t. The first square represents initialization, the second a step of the algorithm, while at the third square the algorithm is terminated.

3.1. Subdivision algorithm. A method to approximate the invariant set of a dynamical system of the form (1) has been presented in [2]. We here report the main steps of the procedure. Given a vector with positive entries r in $\mathbb{R}^d$ and a point c in $\mathbb{R}^d$, we consider a rectangular domain $\Omega \in \mathbb{R}^d$, i.e., the set

$$\Omega := \{x \in \mathbb{R}^d \colon |x_i - c_i| \leq r_i,\ i = 1, \dots, d\}, \tag{7}$$

where x_i is the i-th component of an element x of $\mathbb{R}^d$. We denote in the following by $R_r(c)$ a rectangular domain as above. The method is based on the assumption that we have partial knowledge of the geometrical features of (1).

Assumption 3.1. With the notation above, we assume that the domain Ω is chosen such that $B \subset \Omega$, where B is an invariant set of (1).

Given a spatial discretization parameter $h > 0$, the goal is building a collection of non-overlapping rectangles $\mathcal{A}_h$ such that

(i) $\mathcal{A}_h = \{A_i \colon A_i = R_r(c_i),$ where $\|r\|_{l^\infty} \leq h,\ c_i \in \Omega,\ A_i \cap A_j = \emptyset$ for $i \neq j\}_{i=1}^N$,
(ii) $B \subset \bigcup_{A \in \hat{\mathcal{A}}_{h^{(j+1)}}} A$,
(iii) $\mathcal{A}_h \to B$ for $h \to 0$.

In order to build such a set, we consider the following iterative procedure on the discretization parameter. Given $\mathcal{A}_{h^{(0)}} = \{\Omega\}$, and assuming that we have the collection $\mathcal{A}_{h^{(j)}}$ at the j-th step we proceed as follows

(i) Build a refined collection $\hat{\mathcal{A}}_{h^{(j+1)}}$ from $\mathcal{A}_{h^{(j)}}$ such that $h^{(j+1)} \leq h^{(j)}$ and

$$\bigcup_{A \in \hat{\mathcal{A}}_{h^{(j+1)}}} A = \bigcup_{A \in \mathcal{A}_{h^{(j)}}} A.$$

*Mathematics Section, École Polytechnique Fédérale de Lausanne (assyr.abdulle@epfl.ch)
†Mathematics Section, École Polytechnique Fédérale de Lausanne (giacomo.garegnani@epfl.ch)

(ii) Determine $\mathcal{A}_{h^{(j+1)}} \subset \hat{\mathcal{A}}_{h^{(j+1)}}$ selecting the sets that are mapped by φ_t, i.e.,

$$\mathcal{A}_{h^{(j+1)}} = \{A \in \hat{\mathcal{A}}_{h^{(j+1)}} : \varphi_t^{-1}(A) \cap \hat{A} \neq \emptyset \text{ for some } \hat{A} \in \hat{\mathcal{A}}_{h^{(j+1)}}\}.$$

It has been shown that if B is a global attractor for (1) the sequence $\mathcal{A}_{h^{(j)}}$ satisfies

$$\lim_{j\to\infty} d_{\mathrm{H}}(\mathcal{A}_{h^{(j)}}, B) = 0, \tag{8}$$

provided that the sequence $h_j \to 0$ for $j \to \infty$ and where $d_H(\cdot,\cdot)$ is the Hausdorff distance. If we have further knowledge of the geometrical properties of (1), it is possible to build $\mathcal{A}_{h^{(j)}}$ via a more effective scheme.

Assumption 3.2. There exists a point x_0 in the domain Ω known *a priori* such that $x \in B$, where B is an attractor of (1).

Under this assumption, we can build $\mathcal{A}_{h^{(j)}}$ proceeding as follows:

(i) build a collection $\tilde{\mathcal{A}}_{h^{(j)}}$ such that $\bigcup_{A\in\tilde{\mathcal{A}}_{h^{(j)}}} A = \Omega$,

(ii) find the element A of $\tilde{\mathcal{A}}_{h^{(j)}}$ such that $x_0 \in A$ and denote it as A_0, then $\mathcal{A}_{h^{(j)}} = \{A_0\}$,

(iii) choose a set of points $\{x_i\}_{i=1}^M$ inside A_0 and compute the mapped points $\{y_i = \varphi_t(x_i)\}_{i=1}^M$,

(iv) find the smallest subset $\mathcal{D}$ of $\tilde{\mathcal{A}}_{h^{(j)}}$ such that $y_i \in \bigcup_{A\in\mathcal{D}} A$ for all $i = 1,\dots,d$,

(v) update $\mathcal{A}_{h^{(j)}} \leftarrow \mathcal{A}_{h^{(j)}} \cup \mathcal{D}$,

(vi) for each of the elements of $\mathcal{D}$ not considered yet, proceed from (iii) to (v),

(vii) terminate when $\mathcal{D} = \emptyset$.

A graphical representation of this algorithm is depicted in Figure 1.

Once a covering $\mathcal{A}_h$ of the attractor is known, it is possible to find an approximation μ_h of the density of the invariant measure μ of (1) by a Galerkin method [2]. In particular, we search μ_h in the space of densities of probability measure which are constant on each set A in $\mathcal{A}_h$, i.e.,

$$\begin{aligned} \mu_h(x) &= \mu_i \in \mathbb{R}, \quad \forall x \in A_i,\ \forall A_i \in \mathcal{A}_h,\ \ i = 1,\dots,|\mathcal{A}_h|, \\ \int_{\bigcup_{A\in\mathcal{A}_h} A} \mu_h(x)\mathrm{d}x &= 1, \\ \mu_h(x) &= 0, \qquad\quad \forall x \in \Omega \setminus \textstyle\bigcup_{A\in\mathcal{A}_h} A. \end{aligned} \tag{9}$$

In particular, if $|\mathcal{A}_h| = n$, we can uniquely represent μ_h as a vector of $\mathbb{R}^n$. We obtain the approximation by recalling that for $P\mu = \mu$, where P is the Frobenius-Perron operator. We approximate P by a matrix $P_h \in \mathbb{R}^{n\times n}$ whose entries are given by

$$(P_h)_{ij} = \frac{m(\varphi_t^{-1}(A_i) \cap A_j)}{m(A_j)}, \quad \forall A_i, A_j \in \mathcal{A}_h, \tag{10}$$

where m is the Lebesgue measure. Then, we compute μ_h as the eigenvector of P_h associated with the eigenvalue one, i.e.

$$P_h\mu_h = \mu_h. \tag{11}$$

Let us remark that the elements of P_h are approximated in practice by a Monte Carlo average as

$$(P_h)_{ij} \approx M^{-1} \textstyle\sum_{i=1}^M \chi_{A_i}(\varphi_t(x_i)), \tag{12}$$

where the M points x_i are random samples from A_j. Finally, let us remark that due to the local interaction of the flow map, the matrix P_h should be sparse.

3.2. Probabilistic integrator approach. Let us consider (1) and a probabilistic numerical solution Y_N taking values in $\mathbb{R}^d$ obtained either with the additive noise technique [1] or with the random time-stepping scheme. We consider the random variable Y_N to be an approximation of the solution $y(T)$ of (1) at a sufficiently long time $T = N\Delta_t$, where Δ_t is the time step employed to obtain the numerical solution.

We approximate the density μ_h of the invariant measure using samples $Y_N^{(i)}$, $i = 1, \ldots, M$, drawn from the random variable Y_N. In particular, we exploit a kernel density estimation, i.e., for each value $x \in \mathbb{R}^d$ we consider the estimator $\hat{\mu}_h(x)$ given by

$$\hat{\mu}_h(x) = M^{-1} \textstyle\sum_{i=1}^M K(x - Y_N^{(i)}). \tag{13}$$

The function $K\colon \mathbb{R}^d \to \mathbb{R}$ is a Gaussian kernel given by

$$K(x) = (2\pi)^{(-d/2)} \det(H)^{(-1/2)} e^{-x^T H^{-1} x/2}, \tag{14}$$

where the covariance bandwidth H is a diagonal matrix of $\mathbb{R}^{d\times d}$ whose entries are chosen with Silverman's rule of thumb [7], i.e.,

$$H_{i,i}^{1/2} = \Big(\frac{4}{n(d+2)}\Big)^{1/(d+4)} \sigma_i, \tag{15}$$

and where σ_i are the component-wise population standard deviations of the chosen draws.

4. Numerical experiments. We show numerical experiments of the procedures discussed in the above section. The computations are carried out with a self-developed `C++ / Matlab` code.

4.1. Subdivision algorithm. We first consider the FitzHug-Nagumo ODE, which reads

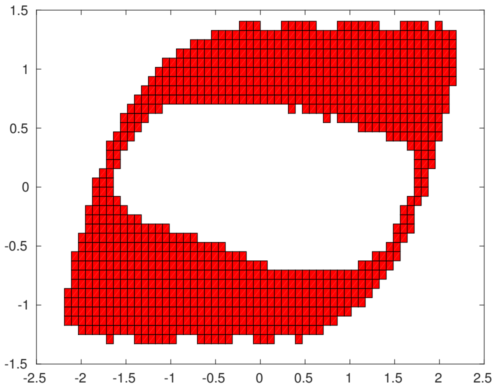
$$\begin{aligned} y_1' &= c\Big(y_1 - \frac{y_1^3}{3} + y_2\Big), \quad && y_1(0) = -1, \\ y_2' &= -\frac{1}{c}(y_1 - a + by_2), && y_2(0) = 1, \end{aligned} \tag{16}$$

where a, b, c are real parameters with values $a = 0.2$, $b = 0.2$, $c = 3$. It is known that with these values for the parameters, the equation admits a limit cycle. We then perform the subdivision algorithm starting from a point laying close to the limit cycle to obtain an approximation $\mathcal{A}_h$ of the attractor A and compute the invariant density with (11). In Figure 2a and 2b, we can see the result obtained with a coarse discretization index h of $\mathcal{A}_h$. In Figure 2c we can see that results are smoother in case of a finer discretization.

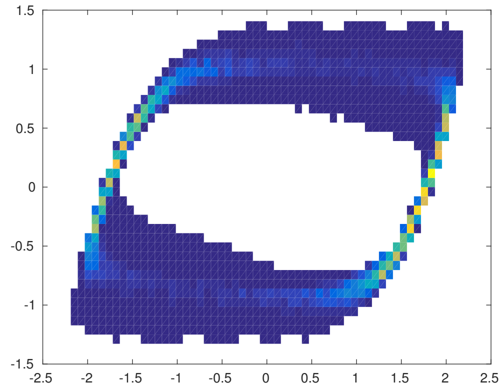
We now consider a three-dimensional model, the Lorenz equations, which is of particular interest. The ODE reads

$$\begin{aligned} y_1' &= \sigma(y_2 - y_1), && y_1(0) = -10, \\ y_2' &= y_1(\rho - y_3) - y_2, \quad && y_2(0) = -1, \\ y_3' &= y_1 y_2 - \beta y_3, && y_3(0) = 40. \end{aligned} \tag{17}$$

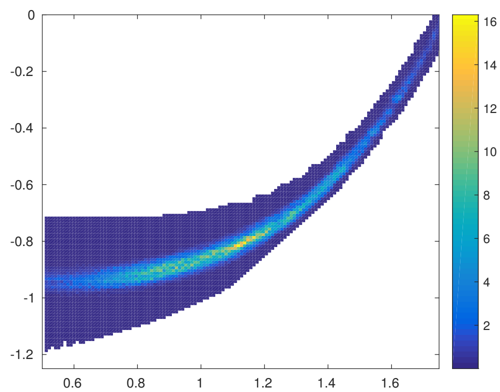
with parameters values. $\sigma = 10$, $\rho = 28$, $\beta = 8/3$. It has been shown [6] that with these values for the parameters, the solution of (17) has a chaotic behavior. Nonetheless, the system admits a strange attractor A and a unique invariant measure μ with support on A [8, 4]. In Figure 3 we show the results for the density estimation on the collection of boxes $\mathcal{A}_h$. The computational time is quite low even in a three-dimensional problem. In this case, we generated 2827 boxes, so the matrix P_h has 7991929 elements. However, storage is not an issue as P_h has only 40180 non-zero entries, thus confirming our hypothesis that P_h should be sparse.



(a) Collection $\mathcal{A}_h$, coarse grid.



(b) Approximation μ_h, coarse grid.



(c) Approximation μ_h, fine grid, zoom.

Fig. 2: Approximation of the attractor A of (16) and of the density of the invariant measure μ.

5. Probabilistic integrator. We consider equation (17) and draw samples from both the additive noise model [1] and with the random time-stepping integrator. We then estimate the density with the kernel estimator, using as evaluation points the draws themselves. Results (Figure 4) show that the random time-stepping and the additive noise methods are consistent, producing similar density estimations. It is possible to remark that the shape of the attractor and the order of magnitude of the density are comparable with respect to the results provided by the subdivision method.

REFERENCES

[1] P. R. Conrad, M. Girolami, S. Särkkä, A. Stuart, and K. Zygalakis, *Statistical analysis of differential equations: introducing probability measures on numerical solutions*, Stat. Comput., (2016).

[2] M. Dellnitz, A. Hohmann, O. Junge, and M. Rumpf, *Exploring invariant sets and invariant measures*, Chaos, 7 (1997), pp. 221–228.

[3] J.-P. Eckmann and D. Ruelle, *Ergodic theory of chaos and strange attractors*, Rev. Modern Phys., 57 (1985),

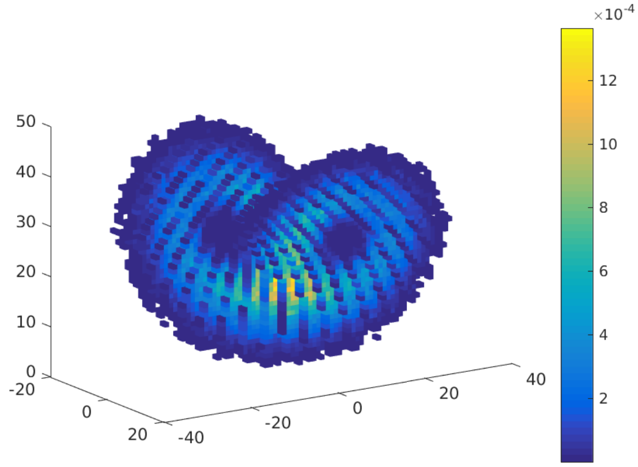
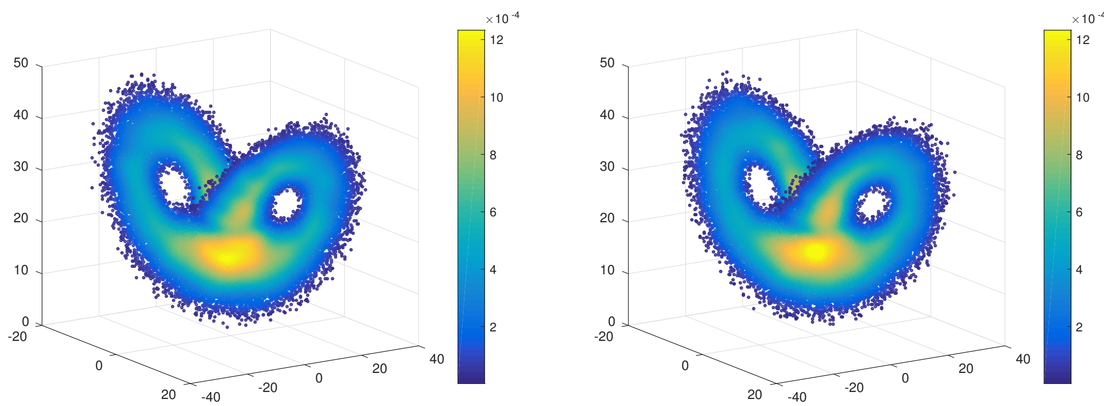


Fig. 3: Approximation of the density of the invariant measure μ for the Lorenz equation with the subdivision method.



(a) Random time-stepping.

(b) Additive noise.

Fig. 4: Approximation of the density of the invariant measure μ for the Lorenz equation with the probabilistic integrator for ODEs.

pp. 617–656.

[4] M. Holland and I. Melbourne, *Central limit theorems and invariance principles for Lorenz attractors*, J. London Math. Soc., 76 (2007), p. 345.

[5] A. Lasota and M. C. Mackey, *Continuous Time Systems: An Introduction*, Springer New York, New York, NY, 1994, pp. 189–249.

[6] E. N. Lorenz, *Deterministic nonperiodic flow*, J. Atmos. Sci., 20 (1963), pp. 130–141.

[7] B. W. Silverman, *Density estimation for statistics and data analysis*, Monographs on Statistics and Applied Probability, Chapman & Hall, London, 1986.

[8] W. Tucker, *The Lorenz attractor exists*, C. R. Acad. Sci. Paris Sér. I Math., 328 (1999), pp. 1197–1202.