MATH 590: NONLINEAR DYNAMICS, CHAOS AND APPLICATIONS. LECTURE NOTES

RAFAIL V. ABRAMOV

Contents

1. Week 1: Lorenz "butterfly" and evidence for deterministic chaos

1.1. **The Lorenz "butterfly" system.** Consider the 2D convective equations:

$$\begin{aligned} \frac{\partial}{\partial t}(\Delta\psi) &= -\nabla^{\perp}\psi\cdot\nabla(\Delta\psi)+\nu\Delta^2\psi+g\alpha\frac{\partial\theta}{\partial x},\\ \frac{\partial}{\partial t}\theta &= -\nabla^{\perp}\psi\cdot\nabla\theta+\frac{T}{H}\frac{\partial\psi}{\partial x}+\kappa\Delta\theta, \end{aligned} \tag{1.1}$$

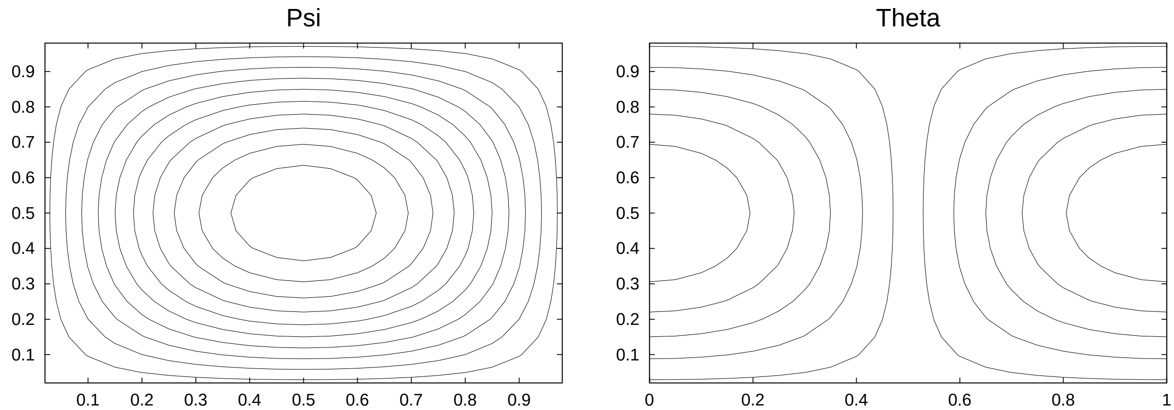

FIGURE 1.1. Rayleigh solutions for (1.1).

where x, y are the horizontal and vertical coordinates, respectively, ψ is the streamfunction, θ is the temperature, and the constants g, α, ν, κ are the gravity acceleration, thermal expansion, viscosity and thermal conductivity. H is the thickness of the fluid layer, and T is the constant temperature difference between the top and bottom fluid surfaces.

Rayleigh has shown that when the quantity

$$\mathcal{R} = \frac{g\alpha H^3 T}{\nu\kappa} \tag{1.2}$$

(now called the Rayleigh number) exceeds critical value

$$\mathcal{R}_0 = \frac{\pi^4}{a^2}(1+a^2)^3, \tag{1.3}$$

where a is a ratio between the horizonal and vertical scales, the state where the streamfunction is constant is no longer stable, and convective motion occurs, yielding a special steady solution of (1.1) in the form

$$\begin{aligned} \psi(x,z) &= \psi_0 \sin\left(\frac{\pi a}{H}x\right)\sin\left(\frac{\pi}{H}z\right), \\ \theta(x,z) &= \theta_0 \cos\left(\frac{\pi a}{H}x\right)\sin\left(\frac{\pi}{H}z\right). \end{aligned} \tag{1.4}$$

This special solution is shown in Figure 1.1. Following Lorenz, here we write a special time-dependent truncation to the convection equation in (1.1), with space-dependence somewhat similar to the Rayleigh's solutions:

$$\begin{aligned} \psi(t,x,z) &= X(t)\frac{\kappa(1+a^2)}{a}\sin\left(\frac{\pi a}{H}x\right)\sin\left(\frac{\pi}{H}z\right), \\ \theta(t,x,z) &= Y(t)\frac{\mathcal{R}_0}{\pi\mathcal{R}}\cos\left(\frac{\pi a}{H}x\right)\sin\left(\frac{\pi}{H}z\right) - Z(t)\frac{\mathcal{R}_0}{\sqrt{2}\pi\mathcal{R}}\sin\left(\frac{2\pi}{H}z\right). \end{aligned} \tag{1.5}$$

Substituting (1.5) into (1.1), one can obtain the following system of ODEs for $X(t), Y(t), Z(t)$ (the Lorenz "butterfly"):

$$\begin{aligned} \dot{X} &= \sigma(Y - X), \\ \dot{Y} &= X(r - Z) - Y, \\ \dot{Z} &= XY - bZ. \end{aligned} \tag{1.6}$$

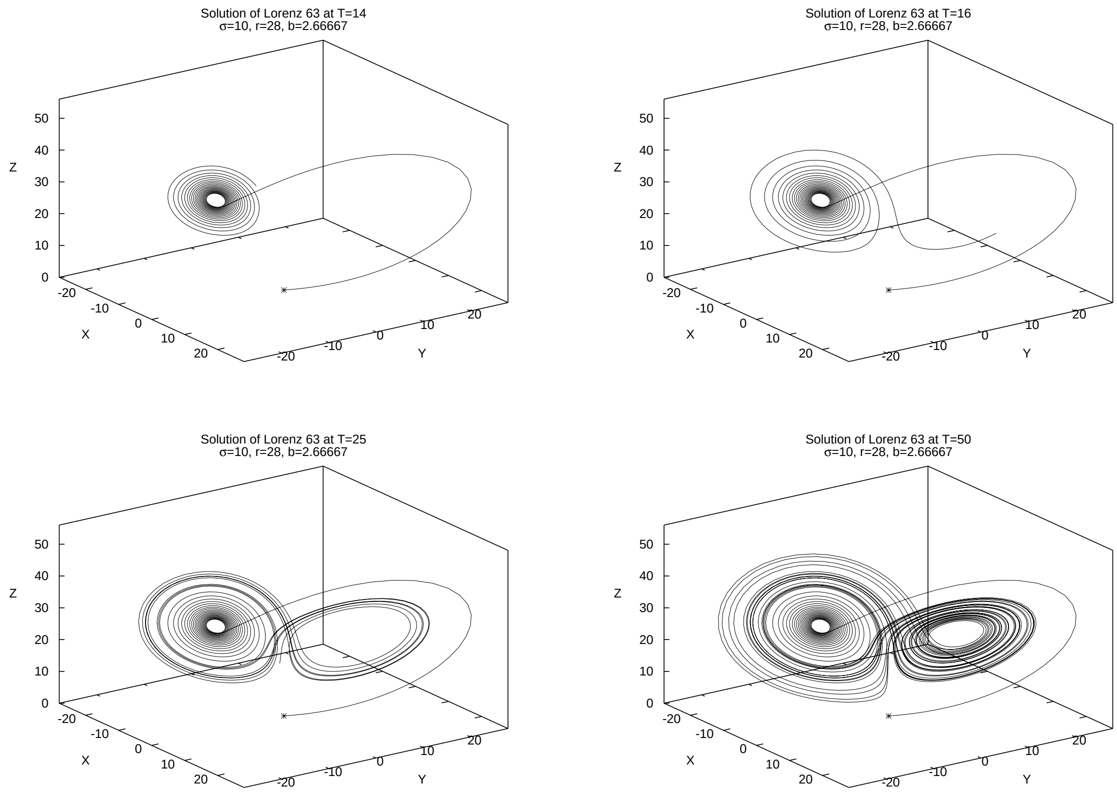

FIGURE 1.2. The solutions of the Lorenz "butterfly" for different times.

Here the time is nondimensionalized, $\tau = \pi^2 H^{-2}(1+a^2)\kappa t$, $\sigma = \kappa^{-2}\nu$ is the Prandtl number, $r = \mathcal{R}\mathcal{R}_0^{-1}$, and $b = 4(1+a^2)^{-1}$.

As we can see, the famous Lorenz "butterfly" in (1.6) is actually a special truncation of the realistic convection equations from (1.1). For the values of parameters $\sigma = 10$, $r = 28$ and $b = 8/3$ one can observe that the solution behaves in a "chaotic" manner, i.e., it does not converge to a fixed point or a limit cycle. The solutions of the Lorenz "butterfly" for different times are shown in Figure 1.2.

So, how do we determine, what is "chaotic"? An intuitive definition is the following: "chaotic" behavior means that the solutions originating from two extremely close (but not identical) initial conditions can diverge significantly from each other. After all, this is where the main difficulty of numerical weather prediction lies; even a small error in initial conditions leads to substantial degradation of the forecast. Here, in Figure 1.3 we start with two extremely close initial conditions and monitor how they behave. Observe that after some time the solutions are far away from each other.

1.2. **Boundedness of solution and linear instability of fixed points.** Hereinafter, we regard a dynamical system as a system of differential equations

$$\frac{\mathrm{d}\vec{x}}{\mathrm{d}t} = \dot{\vec{x}} = \vec{f}(\vec{x}), \tag{1.7}$$

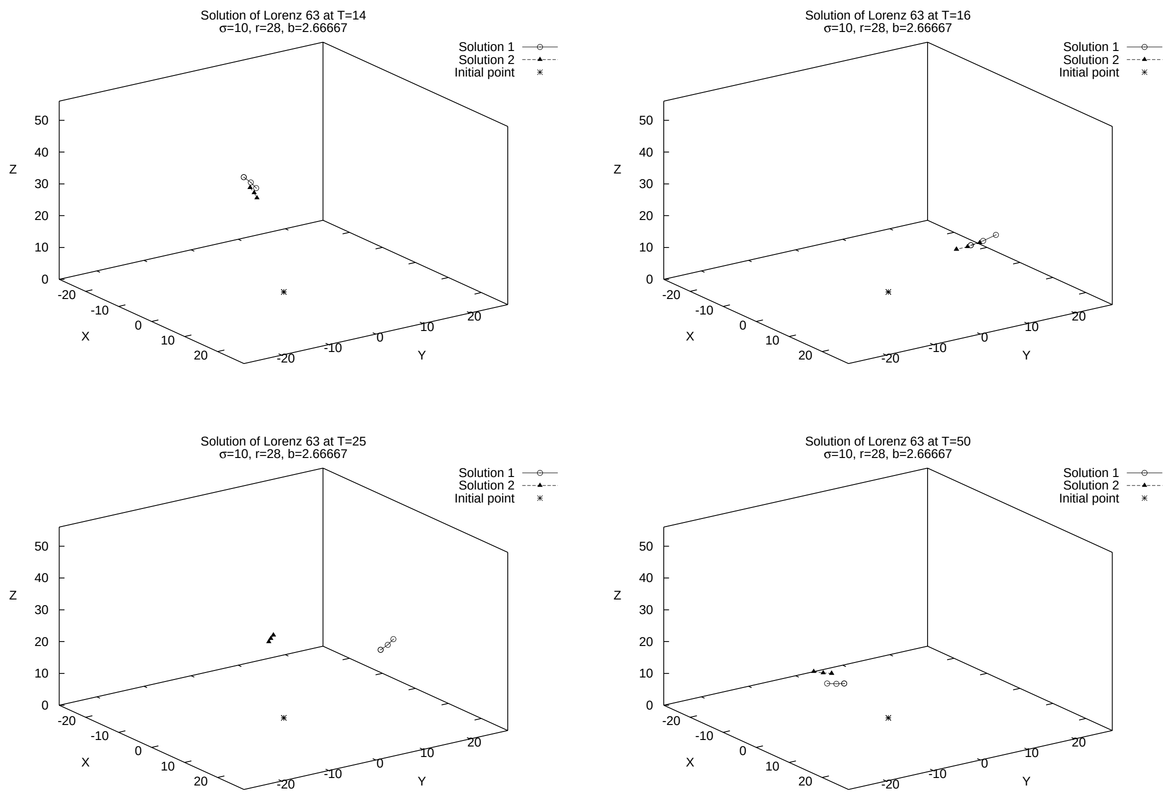

FIGURE 1.3. The solutions of the Lorenz "butterfly" originating from two nearly identical initial conditions. Observe that at $T = 25$ the solutions are far away from each other, although they pass near each other again at $T = 50$.

where $\vec{x} = \vec{x}(t) \in \mathbb{R}^N$ is a vector-valued function of time t, and $\vec{f} : \mathbb{R}^N \to \mathbb{R}^N$ is a smooth nonlinear function. The vector field $\vec{f}$ generates a flow $\vec{\phi}^t : \mathbb{R}^N \to \mathbb{R}^N$, such that the solution $\vec{x}(t)$, originating from an initial point $\vec{x}_0$, is given by

$$\vec{x}(t) = \vec{\phi}^t(\vec{x}_0). \tag{1.8}$$

It is easy to check that

$$\frac{\partial}{\partial t}\vec{\phi}^t(\vec{x}) = \vec{f}(\vec{\phi}^t(\vec{x})). \tag{1.9}$$

The flow $\vec{\phi}^t$ has the group properties:

(1) $\phi^0(\vec{x}) = \vec{x}$,
(2) $\phi^{t_1}(\phi^{t_2}(\vec{x})) = \phi^{t_1+t_2}(\vec{x})$.

Usually, the flow $\vec{\phi}^t(\vec{x})$ is not available explicitly, unlike the vector field $\vec{f}(\vec{x})$.

By looking at the solutions of the Lorenz "butterfly" system, one can wonder "what if the solution eventually becomes infinite, perhaps after very long time?" There are ways to check that, in particular, the Lyapunov function.

Definition 1.1 (Fixed point)**.** *A point* $\bar{\vec{x}} \in \mathbb{R}$ *is said to be* **fixed** *if* $\vec{f}(\bar{\vec{x}}) = 0$*. A fixed point* $\bar{\vec{x}}$ *is said to be* **stable** *if for every neighborhood* U *of* $\bar{\vec{x}}$ *there is a neighborhood* U'*, such that for all* $\vec{x}_0 \in U'$ $\vec{\phi}^t(\vec{x}_0) \in U$ *for all* t*. In addition, if* $\vec{\phi}^t(\vec{x}_0) \to \bar{\vec{x}}$ *as* $t \to \infty$*, then* $\bar{\vec{x}}$ *is* **asymptotically stable***.*

Theorem 1.1 (Hirsch and Smale 1974)**.** *Let* $\bar{\vec{x}}$ *be a fixed point for* (1.6)*, and* $V(\vec{x})$ *be a differentiable function defined on some neighborhood of* $\bar{\vec{x}}$*, such that*

(1) $V(\bar{\vec{x}}) = 0$*,* $V(\vec{x}) > 0$ *for* $\vec{x} \neq \bar{\vec{x}}$*;*

(2) $\dfrac{\mathrm{d}}{\mathrm{d}t}V(\vec{x}) \leq 0$ *for* $\vec{x} \neq \bar{\vec{x}}$*;*

(3) $\dfrac{\mathrm{d}}{\mathrm{d}t}V(\vec{x}) < 0$ *for* $\vec{x} \neq \bar{\vec{x}}$*.*

Then, $V(\vec{x})$ *is called the Lyapunov function. If the first two conditions hold, then* $\vec{x}$ *is stable. If the third condition holds in addition to the first two, then* $\vec{x}$ *is asymptotically stable. Here*

$$\frac{\mathrm{d}}{\mathrm{d}t}V(\vec{x}) = \nabla V(\vec{x}) \cdot \vec{f}(\vec{x}). \tag{1.10}$$

Exercise 1.1. *Exercise 1.0.3 from Guckenheimer.*

With multiple fixed points, one can look for a compact hypersurface $S \subset \mathcal{R}^N$ which encloses a fixed point (or points), such that the vector field $\vec{f}$ is directed everywhere inward on S. If a solution starts inside this surface, it can never leave the interior (apparently), so that the solution will remain bounded forever. In fact, Lorenz [1] has shown that there is a simply connected region $U \in \mathbb{R}^3$, which contains the origin and such that the flow is directed inwards everywhere on ∂U. Now we know that the solution of the Lorenz "butterfly" is bounded for all times. Here is another question: could it, perhaps as time becomes infinite, land on a fixed point somewhere? In order to see whether it is the case, we need to look at the linear stability of fixed points. First, we need to find were the fixed points are.

Exercise 1.2. *Find all fixed points of the Lorenz "butterfly" for positive values of* σ, r, b*. (Answer:* $(0,0,0)$*, and* $(\pm\sqrt{b(r-1)}, \pm\sqrt{b(r-1)}, r-1)$ *if* $r > 1$*.)*

Second, let us linearize the Lorenz "butterfly" near these fixed points and look at the eigenvalues of the linearized system.

Exercise 1.3. *Examine eigenvalues and establish types of the three fixed points of the Lorenz "butterfly" for* $\sigma = 10$*,* $r = 28$ *and* $b = 8/3$*. You do not really need to find the eigenvalues, signs of their real parts will suffice. Answer:*

- $(0,0,0)$ *is a saddle point with one positive eigenvalue*
- $(\pm\sqrt{b(r-1)}, \pm\sqrt{b(r-1)}, r-1)$ *are saddle points with two positive eigenvalues (with rather small real parts)*

Here is what we know at this point: the solution of the Lorenz "butterfly" is bounded, is not (quasi-)periodic, and does not land onto any of the fixed points. On top of that, we observed that trajectories, originating from two very close points, can diverge far away from each other within a short period of time. It is indeed quite peculiar behavior, and we need a tool to quantify it.

1.3. **The tangent map.** As we observed in the case of the Lorenz "butterfly", "chaotic" behavior is characterized by a rapid departure of two initially very close trajectories from each other. In order to quantify this departure, we need a tool which relates initial (infinitesimally) small distance $\delta\vec{x}_0$ between two initial points to distance between trajectories $\delta\vec{x}(t)$. This tool is called the tangent map.

Let us denote by $\vec{x}_0$ and $\vec{x}'_0 = \vec{x}_0 + \delta\vec{x}_0$ two close initial points, then, trajectories $\vec{x}(t)$ and $\vec{x}'(t) = \vec{x}(t) + \delta\vec{x}(t)$, emanating from these two initial points, are given by

$$\vec{x}(t) = \vec{\phi}^t(\vec{x}_0), \qquad \vec{x}'(t) = \vec{\phi}^t(\vec{x}'_0). \tag{1.11}$$

Subtracting $\vec{x}(t)$ from $\vec{x}'(t)$, we obtain

$$\vec{x}'(t) - \vec{x}(t) = \vec{\phi}^t(\vec{x}'_0) - \vec{\phi}^t(\vec{x}_0), \tag{1.12}$$

or

$$\delta\vec{x}(t) = \vec{\phi}^t(\vec{x}_0 + \delta\vec{x}_0) - \vec{\phi}^t(\vec{x}_0) = \left.\frac{\partial\vec{\phi}^t(\vec{x})}{\partial\vec{x}}\right|_{\vec{x}=\vec{x}_0} \delta\vec{x} + O(\|\delta\vec{x}\|^2). \tag{1.13}$$

By the tangent map $(T\phi)^t_{\vec{x}}$ we denote the partial derivative of $\vec{\phi}^t(\vec{x})$:

$$(T\phi)^t_{\vec{x}} = \left.\frac{\partial\vec{\phi}^t(\vec{y})}{\partial\vec{y}}\right|_{\vec{y}=\vec{x}} \tag{1.14}$$

(so that the tangent map is an $N \times N$ matrix) thus obtaining

$$\frac{\delta\vec{x}(t)}{\delta\vec{x}_0} = (T\phi)^t_{\vec{x}_0} + O(\|\delta\vec{x}\|). \tag{1.15}$$

In the limit, as $\delta\vec{x} \to 0$, the term $O(\|\delta\vec{x}\|)$ above vanishes, so one can interpret the tangent map as the derivative of the trajectory with respect to change in its starting point. Further in this notes we will usually omit ϕ from the tangent map notation (that is, denote the tangent map as $T^t_{\vec{x}}$, unless more than one flow is involved).

One can see that, if the tangent map along a trajectory is available, it is possible to get an idea of how rapidly two close trajectories will diverge from each other, which provides an opportunity to quantify "chaos". However, just as the flow $\vec{\phi}^t(\vec{x})$, the tangent map $T^t_{\vec{x}}$ is usually not available to us explicitly. However, one can derive the equation for the tangent map as follows. Differentiating both sides of (1.9) with respect to $\vec{x}$, one obtains

$$\frac{\partial}{\partial\vec{x}}\left(\frac{\partial}{\partial t}\vec{\phi}^t(\vec{x})\right) = \frac{\partial}{\partial\vec{x}}\left(\vec{f}(\vec{\phi}^t(\vec{x}))\right). \tag{1.16}$$

Switching the order of differentiation in the left-hand side, and using the chain rule in the right-hand side, we obtain

$$\frac{\partial}{\partial t}\left(\frac{\partial}{\partial\vec{x}}\vec{\phi}^t(\vec{x})\right) = \frac{\partial\vec{f}(\vec{x})}{\partial\vec{x}}\frac{\partial}{\partial\vec{x}}\vec{\phi}^t(\vec{x}), \tag{1.17}$$

or, by denoting the partial derivative of $\vec{f}$ with respect to $\vec{x}$ by $J(\vec{x})$ (the Jacobian matrix), we obtain the equation for the tangent map:

$$\frac{\mathrm{d}}{\mathrm{d}t}T^t_{\vec{x}_0} = J(\vec{x}(t))T^t_{\vec{x}_0}, \qquad T^0_{\vec{x}_0} = I. \tag{1.18}$$

Note that the equation for the tangent map above in (1.18) must be solved simultaneously with the equation for $\vec{x}(t)$ in (1.7), because the Jacobian $J(\vec{x})$ in (1.18) has to be computed at $\vec{x}(t)$.

2. Week 2. Linear flows and invariant subspaces

3. Week 3. Nonlinear flows and linear and nonlinear maps

3.1. **Nonlinear flows and stable and unstable manifolds near a fixed point.** Given the dynamical system

$$\dot{\vec{x}} = \vec{f}(\vec{x}) \tag{3.1}$$

and a fixed point $\bar{\vec{x}}$, the linearized dynamics for the system is written as

$$\dot{\vec{\xi}} = J(\bar{\vec{x}})\vec{\xi}, \qquad \vec{\xi}(t) = e^{tJ(\bar{\vec{x}})}\vec{\xi}_0, \tag{3.2}$$

where $\vec{\xi} = \vec{x} - \bar{\vec{x}}$. Now the question is: how could the dynamics of the linearized system in (3.2) be related to the dynamics of the original nonlinear system (3.1) in the vicinity of the fixed point $\bar{\vec{x}}$? The answer to this question is given by the following theorem.

Theorem 3.1 (Hartman-Grobman)**.** *If $\bar{\vec{x}}$ is a hyperbolic fixed point, then there exists a neighborhood U of $\bar{\vec{x}}$ and a homeomorphism h, so that $h \circ \vec{f} = J \circ h$.*

In other words, the Hartman-Grobman theorem says that in the vicinity of a hyperbolic fixed point (that is, a fixed point with no eigenvalues with zero real parts) the behavior of the linearized system in (3.2) is qualitatively the same as the behavior of (3.1).

Next, we define nonlinear analogs of invariant stable and unstable subspaces of a linear system.

Definition 3.1 (Local stable and unstable manifolds)**.** *The local stable and unstable manifolds $W^s_{loc}(\bar{\vec{x}})$ and $W^u_{loc}(\bar{\vec{x}})$ of a hyperbolic fixed point $\bar{\vec{x}}$ are defined as follows:*

$$\begin{aligned} W^s_{loc}(\bar{\vec{x}}) &= \{\vec{x} \in \mathbb{R}^N | \phi^t(\vec{x}) \to \bar{\vec{x}} \text{ as } t \to \infty\}, \\ W^u_{loc}(\bar{\vec{x}}) &= \{\vec{x} \in \mathbb{R}^N | \phi^{-t}(\vec{x}) \to \bar{\vec{x}} \text{ as } t \to \infty\}. \end{aligned} \tag{3.3}$$

Theorem 3.2 (Stable/unstable manifold theorem for a fixed point)**.** *For a dynamical system $\dot{\vec{x}} = \vec{f}(\vec{x})$ with a hyperbolic fixed point $\bar{\vec{x}}$ there exist local stable and unstable manifolds $W^s_{loc}(\bar{\vec{x}})$ and $W^u_{loc}(\bar{\vec{x}})$ of the same dimensions N_s and N_u as those of the corresponding eigenspaces E^s and E^u of the linearized system. The $W^s_{loc}(\bar{\vec{x}})$ and $W^u_{loc}(\bar{\vec{x}})$ are tangent to E^s and E^u at $\bar{\vec{x}}$, respectively, and are as smooth as the vectori field $\vec{f}$.*

The local stable and unstable manifolds have their global analogs.

Definition 3.2. *The global stable and unstable manifolds W^s and W^u are defined as follows:*

$$\begin{aligned} W^s(\bar{\vec{x}}) &= \textstyle\bigcup_{t\geq 0} \phi^{-t} W^s_{loc}(\bar{\vec{x}}), \\ W^u(\bar{\vec{x}}) &= \textstyle\bigcup_{t\geq 0} \phi^{t} W^u_{loc}(\bar{\vec{x}}). \end{aligned} \tag{3.4}$$

3.2. **Linear and nonlinear maps.** A linear map is given by the sequence

$$\vec{x}_{n+1} = A\vec{x}_n, \tag{3.5}$$

where $\vec{x} \in \mathbb{R}^N$, and A is an $N \times N$ matrix. Just like for linear flows, there exist stable, unstable and center subspaces for linear maps:

- $E^s = \{N_s$ eigenvectors whose eigenvalues have modulus $< 1\}$,
- $E^u = \{N_u$ eigenvectors whose eigenvalues have modulus $> 1\}$,
- $E^c = \{N_c$ eigenvectors whose eigenvalues have modulus $= 1\}$.

If there are no multiple eigenvalues, then the contraction and expansion of $\vec{x}$ is bounded by the geometric series; i.e., one can find constants $C, \alpha > 0$, such that

$$\begin{aligned} \|\vec{x}^n\| &\le C\alpha^n \|\vec{x}_0\|, \qquad \vec{x}_0 \in E^s, \\ \|\vec{x}^{-n}\| &\le C\alpha^n \|\vec{x}_0\|, \qquad \vec{x}_0 \in E^u. \end{aligned} \tag{3.6}$$

A nonlinear map is given by the sequence

$$\vec{x}_{n+1} = \vec{F}(\vec{x}_n), \tag{3.7}$$

where $\vec{F}$ is a nonlinear vector-valued function. The tangent map $T(\vec{x}_n)$ at $\vec{x}_n$ to $\vec{x}_{n+1}$ is given by the Jacobian (matrix of partial derivatives) of $\vec{F}$:

$$T(\vec{x}_n) = \left.\frac{\partial \vec{F}}{\partial \vec{x}}\right|_{\vec{x}_n}. \tag{3.8}$$

The tangent map at n to $n+m$ is given, according to the chain rule, by the product of corresponding one-step tangent maps:

$$T_n^{n+m} = \prod_{k=n}^{n+m-1} T(\vec{x}_k). \tag{3.9}$$

Just as for flows, there are Hartman-Grobman and stable manifold theorems for nonlinear maps.

Theorem 3.3 (Hartman-Grobman for maps). *Let $\vec{F} : \mathbb{R}^N \to \mathbb{R}^N$ be a diffeomorphism with a hyperbolic fixed point $\bar{\vec{x}}$. Then there exists a homeomorphism h on some neighborhood U of $\bar{\vec{x}}$, so that $h \circ \vec{F} = T(\bar{\vec{x}}) \circ h$ on U.*

Definition 3.3 (Local stable and unstable manifolds for maps). *The local stable and unstable manifolds $W^s_{loc}(\bar{\vec{x}})$ and $W^u_{loc}(\bar{\vec{x}})$ of a hyperbolic fixed point $\bar{\vec{x}}$ are defined as follows:*

$$\begin{aligned} W^s_{loc}(\bar{\vec{x}}) &= \{\vec{x} \in \mathbb{R}^N | \vec{F}^n(\vec{x}) \to \bar{\vec{x}} \text{ as } n \to \infty\}, \\ W^u_{loc}(\bar{\vec{x}}) &= \{\vec{x} \in \mathbb{R}^N | \vec{F}^{-n}(\vec{x}) \to \bar{\vec{x}} \text{ as } n \to \infty\}. \end{aligned} \tag{3.10}$$

Theorem 3.4 (Stable/unstable manifold theorem for a fixed point (maps)). *For a nonlinear map $\vec{x}_{n+1} = \vec{F}(\vec{x}_n)$ with a hyperbolic fixed point $\bar{\vec{x}}$ there exist local stable and unstable manifolds $W^s_{loc}(\bar{\vec{x}})$ and $W^u_{loc}(\bar{\vec{x}})$ of the same dimensions N_s and N_u as those of the corresponding eigenspaces E^s and E^u of $T(\bar{\vec{x}})$. The $W^s_{loc}(\bar{\vec{x}})$ and $W^u_{loc}(\bar{\vec{x}})$ are tangent to E^s and E^u at $\bar{\vec{x}}$, respectively, and are as smooth as $\vec{F}$.*

The local stable and unstable manifolds have their global analogs.

Definition 3.4. *The global stable and unstable manifolds W^s and W^u are defined as follows:*

$$\begin{aligned} W^s(\bar{\vec{x}}) &= \textstyle\bigcup_{n\ge 0} \vec{F}^{-n}(W^s_{loc}(\bar{\vec{x}})), \\ W^u(\bar{\vec{x}}) &= \textstyle\bigcup_{n\ge 0} \vec{F}^{n}(W^u_{loc}(\bar{\vec{x}})). \end{aligned} \tag{3.11}$$

3.3. **Computing the tangent map numerically.** As it is typically impossible to obtain a solution to a system of ODEs like that in (3.1) explicitly, one has to solve this system numerically, discretizing it in time. After the time discretization, we end up with a time-discretized numerical scheme, which in general is a nonlinear map, which has a time-stepping τ and the vector field $\vec{f}$ as parameters:

$$\vec{x}_{n+1} = \vec{G}(\tau, \vec{f}, \vec{x}_n). \tag{3.12}$$

Here we avoid the multistep time discretization schemes, as well as implicit schemes, and assume that the solution at the time step $n+1$ is obtained entirely and explicitly from the solution known at the time step n. Obviously, the solution of the time scheme in (3.12) is not a continuous curve, but a sequence of points $\vec{x}_k$, $k = 0, \dots, K$.

In such a situation, the simplest way to compute the tangent map at the point k to the point $k+m$ is to compute the one-step tangent maps $T(\vec{x})$ of $\vec{G}(\vec{x})$ at the points $k, k+1, \dots, k+m-1$, and then multiply the obtained matrices together:

$$T_k^{k+m} = \prod_{i=k}^{k+m-1} T(\vec{x}_i). \tag{3.13}$$

For the computation of $T(\vec{x})$, let us assume, that the Jacobian $J(\vec{x})$ of $\vec{f}$ can be computed explicitly, and consider two examples that follow.

Example 3.1 (Forward Euler scheme). *In the forward Euler scheme, the nonlinear map $\vec{G}(\tau, \vec{f}, \vec{x})$ is given by*

$$\vec{G}(\tau, \vec{f}, \vec{x}) = \vec{x} + \tau \vec{f}(\vec{x}). \tag{3.14}$$

In this case, the one-step tangent map of $\vec{G}$ is apparently given by

$$T(\tau, \vec{x}) = \frac{\partial \vec{G}}{\partial \vec{x}} = I + \tau J(\vec{x}), \tag{3.15}$$

where $J(\vec{x})$ is the Jacobian of the vector field $\vec{f}$, computed at $\vec{x}$.

Example 3.2 (Second order Runge-Kutta scheme). *The second order Runge-Kutta scheme (RK2) is usually given in textbooks in the following form:*

$$\begin{aligned} \vec{p}_1 &= \vec{f}(\vec{x}_n), \\ \vec{p}_2 &= \vec{f}(\vec{x}_n + \tfrac{\tau}{2}\vec{p}_1), \\ \vec{x}_{n+1} &= \vec{x}_n + \tau \vec{p}_2. \end{aligned} \tag{3.16}$$

In order to obtain the one-step tangent map $T(\tau, \vec{x})$ for the RK2, one has to differentiate all lines in the above scheme in the reverse order, starting with the last:

$$\begin{aligned} T(\tau, \vec{x}_n) &= I + \tau \tfrac{\partial \vec{p}_2}{\partial \vec{x}_n}, \\ \tfrac{\partial \vec{p}_2}{\partial \vec{x}_n} &= J(\vec{x}_n + \tfrac{\tau}{2}\vec{p}_1)(I + \tfrac{\tau}{2}\tfrac{\partial \vec{p}_1}{\partial \vec{x}_n}), \\ \tfrac{\partial \vec{p}_1}{\partial \vec{x}_n} &= J(\vec{x}_n), \end{aligned} \tag{3.17}$$

or, denoting Jacobians of $\vec{p}_1$ and $\vec{p}_2$ as P_1 and P_2, respectively, one can write the following combined scheme, which computes both the next trajectory point and its one-step tangent map simultaneously:

$$\begin{array}{ll} p_1 = \vec{f}(\vec{x}_n) & P_1 = J(\vec{x}_n), \\ p_2 = \vec{f}(\vec{x}_n + \tfrac{\tau}{2}\vec{p}_1) & P_2 = J(\vec{x}_n + \tfrac{\tau}{2}\vec{p}_1)(I + \tfrac{\tau}{2}P_1), \\ \vec{x}_{n+1} = \vec{x}_n + \tau \vec{p}_2 & T(\tau, \vec{x}_n) = I + \tau P_2. \end{array} \tag{3.18}$$

Exercise 3.1. *Derive the time stepping scheme for one-step tangent map for the 4-th order Runge-Kutta method.*

It might be too expensive memory-wise to store the one-step tangent maps $T(\vec{x}_k)$ for each time step. In this case, one can use the chain rule from (3.13) to obtain

$$T_n^{n+k+1} = T(\vec{x}_{n+k}) T_n^{n+k}, \tag{3.19}$$

which allows to compute the tangent map "on the fly", as the solution $\vec{x}_k$ is computed. This approach allows to compute multiple-step tangent maps (that is, tangent maps covering several discrete time steps), which can help save memory.

Often in various applications the vector field $\vec{f}$ (the right-hand side of the original dynamical system in (3.1)) represents a "black box", involving various transformations of coordinates such as the Fourier or wavelet transforms, projections onto spherical harmonics or other bases, which is supplied by the canned routines. In such a case, it quite troublesome to compute the Jacobian of $\vec{f}$ explicitly. However, rather than having the Jacobian in explicit form, in certain cases one can compute its product with an arbitrary vector. In particular, many physical applications have dynamical models of the following form:

$$\vec{f}(\vec{x}) = \vec{B}(\vec{x}, \vec{x}) + \vec{L}(\vec{x}) + \vec{F}, \tag{3.20}$$

where $\vec{B}$ is the bilinear part, which represents nonlinear interactions, $\vec{L}$ is a linear operator representing dissipation or damping, and $\vec{F}$ is $\vec{x}$-independent forcing term. For such form of $\vec{f}$, the product of the Jacobian of $\vec{f}$ with an arbitrary vector $\vec{y} \in \mathbb{R}^N$ can be computed as

$$J(\vec{x})\vec{y} = \vec{B}(\vec{y}, \vec{x}) + \vec{B}(\vec{x}, \vec{y}) + \vec{L}(\vec{y}). \tag{3.21}$$

Note that above the terms $\vec{B}$ and $\vec{L}$ remain "black boxes", i.e., nothing is implied about their structure. The extension to the Jacobian-matrix multiplication (which is, for instance, needed for the RK2 scheme above) is straightforward via Jacobian-vector multiplication used on columns of the matrix.

4. Week 4: Statistical description of chaotic motion: Attractors, invariant measures, characteristic exponents

4.1. Attracting sets and attractors.

Definition 4.1 (Attracting set). *A compact set A in $\mathbb{R}^N$ is an attracting set with a fundamental neighborhood U if*

1. *For every open set $V \supset A$ we have $\phi^t U \subset V$ when t is large enough,*
2. *A is invariant under ϕ, that is, $\phi^t A = A$ for all t.*

The basin of attraction B is defined as $\bigcup_{t>0} \phi^{-t} U$. If the basin of attraction is the whole of $\mathbb{R}^N$, then A is a universal attracting set.

Example 4.1. *For the Lorenz attractor, U is a sufficiently large ball in $\mathbb{R}^3$.*

Some parts of the attracting sets may not be attracting. Consider, for instance, the dynamical system

$$\dot{x} = x - x^3, \qquad \dot{y} = -y. \tag{4.1}$$

The interval $-1 \le x \le 1, y = 0$ is a universal attracting set (by definition), yet, only points $x = \pm 1, y = 0$ are actually attracting. Hence, attractors have to be distinguished from attracting sets.

Definition 4.2 (Attractor: operational definition). *An attractor A is the set, where images $\phi^t x$ of experimental points x accumulate for large t.*

Example 4.2 (Attracting fixed point). *Let $\bar{\vec{x}}$ be a fixed point of the dynamical system, that is $\phi^t \bar{\vec{x}} = \bar{\vec{x}}$ for all t. Then, if all eigenvalues of the Jacobian $J(\bar{\vec{x}})$ have negative real parts, then $\bar{\vec{x}}$ is an attractor and attracting set.*

Example 4.3 (Attracting periodic orbit). *Let there be a point $\vec{a}$ and $\tau > 0$, such that $\phi^\tau \vec{a} = \vec{a}$, but $\phi^t \vec{a} \neq \vec{a}$ when $t \neq \tau$. Then $\vec{a}$ is a periodic point with period τ, and $\Gamma = \{\phi^t \vec{a} : 0 \le t \le \tau\}$ is a corresponding periodic orbit. The tangent map $T^\tau_{\vec{a}}$ has an eigenvalue 1 corresponding to the direction*

$\vec{f}(\vec{a})$. If all other eigenvalues of $T^{\tau}_{\vec{a}}$ are within a circle of radius < 1, then Γ is an attracting periodic orbit. It is also an attracting set and attractor.

Strange attractors. The attractors described above are also attracting sets and nice manifolds. Also note that a small perturbation of the initial point $\delta\vec{x}(0)$ does not grow with time for these attractors, that is $\delta\vec{x}(t) = T^{t}_{\vec{x}(0)}\delta\vec{x}(0)$ does not grow with time.

However, there are attractors, for which a small perturbation of the trajectory leads to its exponential growth in time. These are called strange attractors.

Example 4.4 (Hénon attractor)**.** *The Hénon attractor is given by*

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} 1 + y_n - ax_n^2 \\ bx_n \end{pmatrix}. \tag{4.2}$$

One can find numerically that for $a = 1.4$, $b = 0.3$

$$\delta\vec{x}(t) \approx \delta\vec{x}(0)e^{0.42t}. \tag{4.3}$$

Example 4.5 (Lorenz attractor)**.**

4.2. **Invariant probability measures.** An attractor A provides a global picture of the long-time motion of a dynamical system. However, just knowing what the attractor is does not give any information about the time spent by the solution near different parts of the attractor. A more detailed information is given the probability measure ρ on A, which describes how often a long-term trajectory visits different parts of the attractor. Such a measure is defined as follows: given a function $g(\vec{x})$, its measure on the attractor is defined

$$\rho(g) = \int_{\mathbb{R}^N} g(\vec{x})\rho(\mathrm{d}\vec{x}) = \lim_{s\to\infty}\frac{1}{s}\int_0^s g(\phi^t\vec{x})\mathrm{d}t. \tag{4.4}$$

This measure is invariant under ϕ^t, that is,

$$\rho(g\circ\phi^t) = \rho(g). \tag{4.5}$$

If ρ cannot be written as a superposition of two invariant probability measures ρ_1 and ρ_2 on A, ρ is called ergodic.

Theorem 4.1. *If the compact set A is invariant under the flow ϕ^t, then there is a probability measure ρ on A invariant under ϕ^t with support contained in A. One may choose ρ to be ergodic.*

Theorem 4.2 (Ergodic theorem)**.** *If ρ is ergodic, then the time averages of $\phi^t\vec{x}$ for almost all initial conditions $\vec{x}$ reproduce ρ.*

4.3. **Characteristic exponents.**

Theorem 4.3 (Oseledec)**.** *Let ρ be a probability measure on $\mathbb{R}^N$, and $\vec{F}$ a measure preserving map such that ρ is ergodic. Then, for ρ-almost all $\vec{x}$, the following limit exists:*

$$\lim_{n\to\infty}\left(T^{n\,T}_{\vec{x}}T^{n}_{\vec{x}}\right)^{1/2n} = \Lambda_{\vec{x}}. \tag{4.6}$$

The eigenvalues of $\Lambda_{\vec{x}}$ are ρ-almost everywhere constant.

Let us denote logarithms of the eigenvalues of $\Lambda_{\vec{x}}$ by $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N$. Then, λ_i are called the characteristic, or Lyapunov exponents.

Theorem 4.4. *Let $E^i_{\vec{x}}$ be the subspace of $\mathbb{R}^N$ corresponding to the eigenvalues $\leq e^{\lambda_i}$ of Λ_x. Then, $\mathbb{R}^N = E^1_{\vec{x}} \supset E^2_{\vec{x}} \supset \ldots \supset E^N_{\vec{x}}$ and for ρ-almost all $\vec{x}$*

$$\lim_{n\to\infty} \frac{1}{n} \log \left\| T^n_{\vec{x}} \vec{\xi} \right\| = \lambda_i, \qquad \vec{\xi} \in E^i_{\vec{x}} \backslash E^{i+1}_{\vec{x}}. \tag{4.7}$$

In particular, for almost all $\vec{\xi}$ the limit is the largest characteristic exponent λ_1.

For a continuous-time dynamical system, we have the following corresponding statements:

$$\lim_{t\to\infty} \left(T^{tT}_{\vec{x}} T^t_{\vec{x}} \right)^{1/2t} = \Lambda_{\vec{x}}, \tag{4.8}$$

and

$$\lim_{t\to\infty} \frac{1}{t} \log \left\| T^t_{\vec{x}} \vec{\xi} \right\| = \lambda_i, \qquad \vec{\xi} \in E^i_{\vec{x}} \backslash E^{i+1}_{\vec{x}}. \tag{4.9}$$

Examples:

- Steady state. A steady state is associated with a fixed point $\bar{\vec{x}}$ of the corresponding dynamical system, and is described by the probability measure $\rho = \delta(\bar{\vec{x}})$. Let α_i be the eigenvalues of $T(\bar{\vec{x}})$, then the characteristic exponents are

$$\lambda_i = \log |\alpha_i|. \tag{4.10}$$

  In particular, a stable fixed point has only negative characteristic exponents.
- Periodic orbit. A periodic orbit is $\Gamma = \{\phi^t \vec{a} : 0 \leq t \leq \tau\}$ of the corresponding continuous-time dynamical system. It is described by the probability measure

$$\rho = \delta(\Gamma) = \frac{1}{\tau} \int_0^\tau \delta(\phi^t \vec{a}) \mathrm{d}t. \tag{4.11}$$

  We denote by α_i the eigenvalues of $T^\tau_{\vec{a}}$, then one of these eigenvalues is 1 (corresponding to the direction $\vec{f}(\vec{a})$). The characteristic exponents are then

$$\lambda_i = \frac{1}{\tau} \log |\alpha_i|, \tag{4.12}$$

  and one of them is zero.

Characteristic exponents as indicators of periodic motion.

Theorem 4.5 (continuous-time fixed point)**.** *Assume that all characteristic exponents of a continuous-time dynamical system are nonzero. Then $\rho = \delta(\bar{\vec{x}})$, where $\bar{\vec{x}}$ is a fixed point.*

Theorem 4.6 (discrete-time periodic orbit)**.** *Consider a discrete-time dynamical system and assume that all characteristic exponents are negative. Then*

$$\rho = \frac{1}{M} \sum_{k=1}^{M} \delta(\vec{F}^k \vec{a}), \tag{4.13}$$

where $\{\vec{a}, \vec{F}\vec{a}, \ldots, \vec{F}^{M-1}\vec{a}\}$ is an attracting periodic orbit of period M.

Theorem 4.7 (continuous-time periodic orbit)**.** *Consider a continuous-time dynamical system and assume that all characteristic exponents are negative, except λ_1. Then there are two possibilities:*

(1) $\rho = \delta(\bar{\vec{x}})$, where $\bar{\vec{x}}$ is a fixed point, or
(2) $\rho = \delta(\Gamma)$, where Γ is a periodic orbit, and in this case $\lambda_1 = 0$.

Corollary: chaotic motion is impossible in $\mathbb{R}^2$:

- $\lambda_1 = \lambda_2 = 0$

- $\lambda_1 = 0, \lambda_2 < 0$ – attracting periodic orbit,
- $\lambda_1 > 0, \lambda_2 = 0$ – repelling periodic orbit,
- $\lambda_1 \neq 0$, $\lambda_2 \neq 0$ – fixed point.

4.3.1. *General remarks on characteristic exponents.*
Growth of volume elements. The rate of average exponential growth of a infinitesimal vector $\delta\vec{x}(t)$ is given by the largest characteristic exponent λ_1. The rate of growth of a surface element $\delta\sigma(t) = \delta\vec{x}_1(t) \wedge \delta\vec{x}_2(t)$ is similarly given by the sum of the two largest characteristic exponents $\lambda_1 + \lambda_2$. In general, for a k-volume element $\delta\vec{x}_1(t) \wedge \ldots \wedge \delta\vec{x}_k(t)$ the rate of growth is $\lambda_1 + \ldots + \lambda_k$.

For a dynamical system on $\mathbb{R}^N$ the rate of growth of the N-volume element is the rate of growth of determinant of the tangent map $|T^t_{\vec{x}}|$, and is given by $\lambda_1 + \ldots + \lambda_N$.
Time reflection. Let us assume that the flow ϕ^t is defined for both positive and negative values of t. Then, if ρ is invariant for ϕ^t, it is also invariant for ρ^{-t}. The characteristic exponents for the time-reversed system are those of the original system, but with reversed sign. We also have the corresponding sequence of nested subspaces $\bar{E}^1_{\vec{x}} \subset \bar{E}^2_{\vec{x}} \ldots$ of ϕ^{-t} for almost all $\vec{x}$, such that

$$\lim_{t\to-\infty} \frac{1}{t} \log \|T^t_{\vec{x}}\vec{u}\| = \lambda_i, \quad \vec{u} \in \bar{E}^i_{\vec{x}} \backslash \bar{E}^{i-1}_{\vec{x}}. \tag{4.14}$$

Define $F^i_{\vec{x}} = E^i_{\vec{x}} \bigcap \bar{E}^i_{\vec{x}}$. Then, for ρ-almost all $\vec{x}$, the subspaces $F^i_{\vec{x}}$ span $\mathbb{R}^N$.

5. Week 5: Stable and unstable manifolds, entropy

5.1. **Stable and unstable manifolds.** One can define a nonlinear analog of the linear subspaces $E^i_{\vec{x}}$ which correspond to negative characteristic exponents. Let $\lambda > 0$, $\varepsilon > 0$ and define

$$W^s_{loc}(\vec{x}, \varepsilon, \lambda) = \{\vec{y} : d(\phi^t\vec{x}, \phi^t\vec{y}) \leq \varepsilon e^{-\lambda t}, \quad t > 0\}, \tag{5.1}$$

where $d(\vec{x}, \vec{y})$ is the distance between $\vec{x}$ and $\vec{y}$. If $\lambda_{i-1} > \lambda > \lambda_i$, then W^s_{loc} for ρ-almost all $\vec{x}$ and small ε is a piece of differentiable manifold, called a local stable manifold at $\vec{x}$. It is tangent at $\vec{x}$ to the linear space $E^i_{\vec{x}}$ and has the same dimension.

If the system is defined for negative times, we can define global stable manifolds through

$$W^s(\vec{x}, \lambda, \varepsilon) = \bigcup_{t<0} \phi^t W^s_{loc}(\vec{x}, \lambda, \varepsilon). \tag{5.2}$$

We can also define the stable manifold of $\vec{x}$ as

$$W^s(\vec{x}) = \{y : \lim_{t\to\infty} \frac{1}{t} \log d(\phi^t\vec{x}, \phi^t\vec{y}) < 0\}, \tag{5.3}$$

which is the largest of the stable manifolds, corresponding to the largest negative characteristic exponent. For the systems where negative times are allowed, we obtain unstable manifolds W^u by replacing t with $-t$ in the definitions.

Example 5.1 (Periodic orbit)**.** *Let Γ be a closed orbit for a continuous-time dynamical system. The ergodic invariant measure on Γ is $\delta(\Gamma)$. For $\vec{x} \in \Gamma$ we have*

$$W^s(\vec{x}) = \{\vec{y} : \lim_{t\to\infty} \|\phi^t\vec{y} - \phi^t\vec{x}\| = 0\}. \tag{5.4}$$

This is also called the strong stable manifold of $\vec{x}$, and a stable manifold of Γ is defined as

$$W^s_\Gamma = \bigcup_{x\in\Gamma} W^s(\vec{x}). \tag{5.5}$$

Theorem 5.1. *If A is an attracting set, and $\vec{x} \in A$, then $W^u(\vec{x}) \subset A$.*

Proof. If U is a fundamental neighborhood of A, and $\vec{y} \in W^u(\vec{x})$, then $\phi^{-t}\vec{y} \in U$ for sufficiently large t (because $\phi^{-t}\vec{y}$ approaches $\phi^{-t}\vec{x}$, while $\phi^{-t}\vec{x} \in A$). Therefore, $\vec{y} \in \bigcap_{t>\tau} \phi^t U = A$ □

Corollary 5.1. *The number of positive characteristic exponents for any ergodic invariant measure with support in A is a lower bound to the dimension of A.*

5.2. **Entropy.** If ρ is an ergodic probability measure for a dynamical system, one can introduce a concept of mean rate of creation of information $h(\rho)$, also known as Kolmogorov-Sinai entropy.

Let $\mathcal{A} = (\mathcal{A}_1, \ldots, \mathcal{A}_\alpha)$ be a finite ρ-measurable partition of the support of ρ. Then, we denote $F^{-k}\mathcal{A} = (F^{-k}\mathcal{A}_1, \ldots, F^{-k}\mathcal{A}_\alpha)$, Then, $\mathcal{A}^{(n)}$ is defined as the iterative refinement

$$\mathcal{A}^{(n)} = \mathcal{A} \vee F^{-1}\mathcal{A} \vee \ldots \vee F^{-n+1}\mathcal{A}, \tag{5.6}$$

where the refinement for partitions Q and R is defined as

$$Q \vee R = \{Q_i \cap R_j, \quad 1 \le i, j \le \alpha, \quad \rho(Q_i \cap R_j) > 0\} \tag{5.7}$$

Then, we write

$$H(\mathcal{A}) = -\sum_{i=1}^{\alpha} \rho(\mathcal{A}_i) \log \rho(\mathcal{A}_i). \tag{5.8}$$

Thus, $H(\mathcal{A})$ is the information content of the partition $\mathcal{A}$ with respect to state ρ, while $H(\mathcal{A}^{(n)}$ is the same, over an interval of time of length n. The following limits define $h(\rho, \mathcal{A})$ and $h(\rho)$:

$$\begin{aligned} h(\rho, \mathcal{A}) &= \lim_{n\to\infty} [H(\mathcal{A}^{(n+1)} - H(\mathcal{A}^{(n)})] = \lim_{n\to\infty} \frac{1}{n} H(\mathcal{A}^{(n)}), \\ h(\rho) &= \lim_{diam\mathcal{A}\to 0} h(\rho, \mathcal{A}). \end{aligned} \tag{5.9}$$

Theorem 5.2. *Let $F : \mathbb{R}^N \to \mathbb{R}^N$ be a differentiable map and ρ an ergodic measure with compact support. Then, the sum of the positive characteristic exponents of ρ is an upper bound on the Kolmogorov-Sinai entropy:*

$$h(\rho) \le \sum_{\lambda_i > 0} \lambda_i \tag{5.10}$$

Theorem 5.3. *In addition to the conditions of the previous theorem, if ρ has smooth density w.r.t. Lebesgue measure, then strict equality holds:*

$$h(\rho) = \sum_{\lambda_i > 0} \lambda_i. \tag{5.11}$$

5.3. **Anosov and Axiom-A dynamical systems.**

Definition 5.1. *A point $\vec{a} \in \mathbb{R}^N$ is wandering if there is an open set B containing $\vec{a}$ such that $B \cap F^k B = \emptyset$ for all $k > 0$. The set of points that are not wandering is the nonwandering set Ω. It is closed, $\vec{F}$-invariant subset of $\mathbb{R}^N$.*

Let Λ be a closed $\vec{F}$-invariant subset of $\mathbb{R}^N$, and assume that we have linear subspaces $E^-_{\vec{x}}, E^+_{\vec{x}}$ of $T(\vec{x})$ for each $\vec{x} \in \Lambda$. Also assume that $T_{\vec{x}} E^-_{\vec{x}} = E^-_{\vec{F}(\vec{x})}$, $T_{\vec{x}} E^+_{\vec{x}} = E^+_{\vec{F}(\vec{x})}$ (i.e. there is a continuous invariant splitting of the tangent space). Then, Λ is a hyperbolic set if one can choose constants $C, \theta > 0$ such that

$$\begin{aligned} \|T^n_{\vec{x}} \vec{u}\| &\le C\theta^{-n} \|\vec{u}\| \qquad \vec{u} \in E^-_{\vec{x}}, \\ \|T^{-n}_{\vec{x}} \vec{v}\| &\le C\theta^{-n} \|\vec{v}\| \qquad \vec{v} \in E^+_{\vec{x}}. \end{aligned} \tag{5.12}$$

If the whole $\mathbb{R}^N$ is hyperbolic, then $\vec{F}$ is called an Anosov diffeomorphism. If the nonwandering set Ω is hyperbolic, and if the periodic points are dense in Ω, it is called an Axiom A diffeomorphism.
Continuous systems.

Definition 5.2. *A point $\vec{a} \in \mathbb{R}^N$ is wandering if there is an open set B containing $\vec{a}$ such that $B \cap \phi^t B = \emptyset$ for all $t > 0$. The set of points that are not wandering is the nonwandering set Ω. It is closed, ϕ-invariant subset of $\mathbb{R}^N$.*

Let Λ be a closed F-invariant subset of $\mathbb{R}^N$ containing no fixed point, and assume that we have linear subspaces $E^-_{\vec{x}}, E^0_{\vec{x}}, E^+_{\vec{x}}$ of $T^t_{\vec{x}}$ for each $\vec{x} \in \Lambda$. Also assume that $T^t_{\vec{x}} E^-_{\vec{x}} = E^-_{\phi^t \vec{x}}$, $T^t_{\vec{x}} E^+_{\vec{x}} = E^+_{\phi^t \vec{x}}$ (i.e. there is a continuous invariant splitting of the tangent space). Then, Λ is a hyperbolic set if one can choose constants $C, \theta > 0$ such that

$$\begin{aligned} \|T^t_{\vec{x}} \vec{u}\| &\le C\theta^{-t} \|\vec{u}\| \qquad \vec{u} \in E^-_{\vec{x}}, \\ \|T^{-t}_{\vec{x}} \vec{v}\| &\le C\theta^{-t} \|\vec{v}\| \qquad \vec{u} \in E^+_{\vec{x}}. \end{aligned} \tag{5.13}$$

If the whole $\mathbb{R}^N$ is hyperbolic, then ϕ^t is called an Anosov flow. If the nonwandering set Ω is hyperbolic, and if the periodic points are dense in Ω, it is called an Axiom A flow.

Theorem 5.4. *For an Axiom-A system, Ω is the union of finitely many disjoint closed invariant sets $\Omega_1, \ldots, \Omega_s$, and for each Ω_i there is $\vec{x} \in \Omega_i$ such that the orbit $\phi^t \vec{x}$ is dense in Ω. The decomposition $\Omega = \cup_i \Omega_i$ is unique.*

The sets Ω_i are called basic sets, while those which are attracting sets are attractors. There is always at least on attractor among the basic sets.

6. Week 6. SRB measures, dimensions

6.1. **SRB measures.** As shown before, the attracting sets are the unions of unstable manifolds. On the other hand, in the directions transversal to unstable manifolds, one finds a complicated structure of foliations of these unstable manifolds. This suggests that there are invariant measures on the attracting sets which are smooth along unstable manifolds, and "rough" in the directions transversal to the unstable manifolds. These measures have been shown to exist in special cases, and are called the SRB (Sinai-Ruelle-Bowen) measures.

Given an ergodic invariant measure ρ with compact support, unstable manifolds W^u are defined for almost all $\vec{x}$. Then, the folloing result holds:

Theorem 6.1. *Let $\vec{F}$ be a twice-differentiable diffeomorphism and ρ an ergodic measure with compact support. Then the following conditions are equvalent:*

(1) ρ is an SRB measure, that is, it has absolutely continuous conditional measures on unstable manifolds,

(2)
$$h(\rho) = \sum_{\lambda_i > 0} \lambda_i,$$
where $h(\rho)$ is the Kolmogorov-Sinai entropy of ρ, and λ_i are the characteristic exponents.

Theorem 6.2. *Let $\vec{F}$ be a twice-differentiable diffeomorphism with an Axiom A attractor A and basin of attraction U. Then there is an ergodic $\vec{F}$-invariant probability measure ρ on A such that*

(1) If ρ is SRB, then it is unique,

(2) *There is a set $S \subset U$ such that $U \backslash S$ has zero Lebesgue measure, such that for every continuous observable $g : A \to \mathbb{R}$, we have, for every $x \in S$,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} g(\vec{F}^i(\vec{x})) = \int g(\vec{x})\rho(\mathrm{d}\vec{x}).$$

Theorem 6.3. *Let $\vec{F}$ be a twice-differentiable diffeomorphism on $\mathbb{R}^N$ and ρ an SRB measure such that all characteristic exponents are different from zero. Then there is a set $S \in \mathbb{R}^N$ with nonzero Lebesgue measure such that such that for every continuous observable g, we have, for every $x \in S$,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} g(\vec{F}^i(\vec{x})) = \int g(\vec{x})\rho(\mathrm{d}\vec{x}).$$

What if one of the characteristic exponents is zero? Consider the map

$$x_{n+1} = x_n^3.$$

It has an invariant set $x = 0$ with δ-function as an ergodic measure, with zero characteristic exponent. However, $x = 0$ is weakly repelling, which means that its basin of attraction has zero Lebesgue measure.

6.2. **Dimension.**

Definition 6.1 (Capacity dimension)**.** *Let A be a subset of $\mathbb{R}^N$, and $N(r, A)$ the minimum number of balls of radius r to cover A. Then, the capacity dimension is*

$$\dim A = \lim_{r\to 0} \frac{\log N(r, A)}{\log(1/r)}.$$

Definition 6.2 (Hausdorff measure)**.** *Let A be a subset of $\mathbb{R}^N$, $r > 0$, and σ be a covering of A by a family of sets σ_k with diameter $d_k \leq r$. Given $\alpha > 0$, the Hausdorff measure of A is given by*

$$m^\alpha(A) = \lim_{r\to 0} \inf_\sigma \sum_k d_k^\alpha.$$

Definition 6.3 (Hausdorff dimension)**.** *The Hausdorff dimension of A is defined as*

$$\dim_H(A) = \inf\{\alpha : m^\alpha(A) = 0\}.$$

Definition 6.4 (Information dimension)**.** *Given a probability measure ρ, its information dimension is the Hausdorff dimension of a set with measure* 1.

Theorem 6.4. *Let ρ be a probability measure on $\mathbb{R}^N$. If*

$$\lim_{r\to 0} \frac{\log B_r(x)}{\log r} = \alpha$$

for ρ-almost all x, then $\dim_H \rho = \alpha$.

Relation between characteristic exponents and Hausdorff dimension. Let us denote the following quantity:

$$c_\rho(s) = \sum_{i=1}^{k} \lambda_i + (s-k)\lambda_{k+1}, \qquad k \leq s < k+1. \tag{6.1}$$

For $\mathbb{R}^N$, $c_\rho(s)$ is defined on the interval $[0, N]$. $c_\rho(x)$ has the following properties:

(1) c_ρ is piecewise-linear, with "breaking" points corresponding to integer values of s.
(2) $c_\rho(0) = 0$;

(3) $\max_s c_\rho(s)$ = sum of positive characteristic exponents;
(4) This maximum is achieves at s = number of positive characteristic exponents;
(5) For sufficiently large s, $c_\rho(s)$ becomes negative.

Definition 6.5 (Lyapunov dimension)**.** *The Lyapunov dimension is defined as*

$$\dim_\Lambda \rho = \max\{s : c_\rho(s) \geq 0\},$$

or, equivalently,

$$\dim_\Lambda \rho = k + \frac{c_\rho(k)}{|\lambda_{k+1}|},$$

where k is such that $c_\rho(k) \geq 0$ and $c_\rho(k+1) < 0$.

Theorem 6.5. *Let $\vec{F} : \mathbb{R}^N \to \mathbb{R}^N$ be a twice differentiable map, and let ρ be an ergodic measure with compact support. Then,*

$$\dim_H \rho \leq \dim_\Lambda \rho.$$

Furthermore, if ρ is an SRB measure, then

$$\dim_H \rho \geq \textit{ number of positive characteristic exponents}$$

Conjecture 6.1. *If ρ is an SRB measure, then* $\dim_H \rho = \dim_\Lambda \rho$.

6.3. **Computation of characteristic exponents.**

Conjecture 6.2. *Let $V = [\vec{v}_1, \ldots, \vec{v}_k]$, $Q = [\vec{q}_1, \ldots, \vec{q}_k]$, $\vec{v}_i, \vec{q}_i \in \mathbb{R}^N$, $k \leq N$ be two $N \times K$ matrices, and let R be a $K \times K$ matrix, such that*

$$V = QR.$$

Then

$$vol(\vec{v}_1, \ldots, \vec{v}_k) = \det R \cdot vol(\vec{q}_1, \ldots, \vec{q}_k),$$

where $vol(\vec{v}_1, \ldots, \vec{v}_k)$ is the volume of the k-dimensional parallelogram spanned by $\vec{v}_1, \ldots, \vec{v}_k$.

Sketch of proof: one can consider this situation in $\mathbb{R}^k$, where V and Q are square $K \times K$ matrices, and for apparent reasons $\det V = \det Q \det R$. Then, embed $\mathbb{R}^k$ into $\mathbb{R}^N$.

Now, let $\vec{q}_i, \ldots, \vec{q}_k$ be orthonormal vectors. Then,

$$vol(\vec{v}_1, \ldots, \vec{v}_k) = \det R.$$

At this point, let us assume that R is an upper-triangular matrix with nonnegative diagonal elements. Then,

$$\begin{aligned}
&vol(\vec{v}_1) = R_{11},\\
&vol(\vec{v}_1, \vec{v}_2) = R_{11}R_{22},\\
&vol(\vec{v}_1, \vec{v}_2, \vec{v}_3) = R_{11}R_{22}R_{33},\\
&\cdots\\
&vol(\vec{v}_1, \ldots, \vec{v}_i) = \prod_{j=1}^{i} R_{jj}, \qquad (6.2)\\
&\cdots\\
&vol(\vec{v}_1, \ldots, \vec{v}_k) = \prod_{j=1}^{k} R_{jj} = \det R.
\end{aligned}$$

Now, let us assume that $\vec{v}_1, \ldots, \vec{v}_k$ are obtained from $\vec{q}_1, \ldots, \vec{q}_k$ as follows:

$$\vec{v}_i(n) = T^n \vec{e}_i,$$

where $\vec{e}_i$ is the i-th basis vector in Cartesian coordinates. Then, as $n \to \infty$,

$$\sum_{j=1}^{i} \lambda_j \to \frac{1}{n} \sum_{j=1}^{i} \log R_{jj},$$

and thus

$$\lambda_i \to \frac{1}{n} \log R_{ii}.$$

In practice, the Lyapunov characteristic exponents are computed as follows: recall that $T^n_{\vec{x}} = T(\vec{F}^{n-1}\vec{x}) \dots T(\vec{x})$. Then denote $T(\vec{x}) = Q_1 R_1$, $T_k = T(\vec{F}^{k-1}\vec{x}) Q_{k-1}$ and decompose $T_k = Q_k R_k$. Then, apparently, $T^n_{\vec{x}} = Q_n R_n \dots R_1$, and the Lyapunov exponents can be found as

$$\lambda_i = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \log(R_k)_{ii}. \tag{6.3}$$

Similarly, the nested Lyapunov subspaces $E^i_{\vec{x}}$ are found as follows: consider the last $N - i + 1$ columns of the matrix

$$\mathcal{R}_n = R_1^{-1} \dots R_n^{-1} \Lambda, \tag{6.4}$$

where Λ is the diagonal part of the matrix product $R_n \dots R_1$. Let $E^i_{\vec{x}}(n)$ be the subspace spanned by these columns. Then, $E^i_{\vec{x}} = \lim_{n\to\infty} E^i_{\vec{x}}(n)$.

7. Week 7: Linear response to small external perturbations

7.1. **Linear response for flows.** Consider the following situation: we have a dynamical system of the form

$$\dot{\vec{x}}(t) = \vec{f}(\vec{x}(t)), \qquad \vec{x}(0) = \vec{x}, \qquad \vec{x}(t) = \phi^t \vec{x}, \tag{7.1}$$

with the ergodic invariant probability measure ρ, such that for any observable $A(\vec{x})$

$$\rho(A) = \int A(\vec{x}) \rho(\mathrm{d}x) = \int A(\phi^t \vec{x}) \rho(\mathrm{d}\vec{x}), \quad \forall t. \tag{7.2}$$

Now, assume that a small forcing is introduced into the right-hand side of (7.1), together with a small perturbation in the initial condition:

$$\dot{\vec{x}}(t) = \vec{f}(\vec{x}(t)) + w(\vec{x}(t)) \delta \vec{f}(t), \qquad \vec{x}(0) = \vec{x} + \delta \vec{x}, \qquad \vec{x}(t) = \hat{\phi}^t \vec{x}, \tag{7.3}$$

where $w(\vec{x})$ is a matrix, $\delta \vec{f}(t)$ is a small time-dependent forcing, such that $\delta \vec{f}(t) = 0$ when $t < 0$, and $\hat{\phi}^t$ is the altered flow operator. The question is: how the quantity

$$\delta_t \rho(A) = \int \left[A(\hat{\phi}^t \vec{x}) - A(\phi^t \vec{x}) \right] \rho(\mathrm{d}\vec{x}) \tag{7.4}$$

is going to respond, and is there a suitable linear (with respect to small perturbations) approximation to this response?

Formally expanding (7.4) with respect to $\delta\phi^t \vec{x} = \hat{\phi}^t \vec{x} - \phi^t \vec{x}$ and discarding higher order terms, we obtain

$$\delta_t \rho(A) = \int \nabla A(\phi^t \vec{x}) \delta \phi^t \vec{x} \rho(\mathrm{d}\vec{x}). \tag{7.5}$$

Now one needs to find a suitable linear approximation for $\delta\phi^t \vec{x}$. Subtracting (7.1) from (7.3), expanding with respect to $\delta\phi^t \vec{x}$ and discarding higher order terms, we find

$$\frac{\partial}{\partial t} \delta \phi^t \vec{x} = J(\phi^t \vec{x}) \delta \phi^t \vec{x} + w(\phi^t \vec{x}) \delta \vec{f}(t), \quad \delta \phi^0 \vec{x} = \delta \vec{x}, \tag{7.6}$$

where $J(\vec{x})$ is the Jacobian of $\vec{f}(\vec{x})$, as usual. The formal solution to this equation is given by the Duhamel's principle:

$$\delta\phi^t\vec{x} = e^{\int_0^t J(\phi^s\vec{x})\mathrm{d}s}\delta\vec{x} + \int_0^t e^{\int_\tau^t J(\phi^s\vec{x})\mathrm{d}s} w(\phi^\tau\vec{x})\delta\vec{f}(\tau)\mathrm{d}\tau. \tag{7.7}$$

In order to interpret the semigroup notation above, recall that

$$\frac{\partial}{\partial t} T^{t-\tau}_{\phi^\tau\vec{x}} = J(\phi^t\vec{x}) T^{t-\tau}_{\phi^\tau\vec{x}}, \quad \forall \vec{x}, t, \tau, \tag{7.8}$$

with the formal solution

$$T^{t-\tau}_{\phi^\tau\vec{x}} = e^{\int_\tau^t J(\phi^s\vec{x})\mathrm{d}s}, \tag{7.9}$$

yielding

$$\delta\phi^t\vec{x} = T^t_{\vec{x}_0}\delta\vec{x}_0 + \int_0^t T^{t-\tau}_{\phi^\tau\vec{x}} w(\phi^\tau\vec{x})\delta\vec{f}(\tau)\mathrm{d}\tau. \tag{7.10}$$

Substituting (7.10) into (7.5), one obtains

$$\begin{aligned} \delta_t\rho(A) &= \left[\int \nabla A(\phi^t\vec{x}) T^t_{\vec{x}}\rho(\mathrm{d}\vec{x})\right]\delta\vec{x}_0 + \int_0^t \left[\int \nabla A(\phi^t\vec{x}) T^{t-\tau}_{\phi^\tau\vec{x}} w(\phi^\tau\vec{x})\rho(\mathrm{d}\vec{x})\right]\delta\vec{f}(\tau)\mathrm{d}\tau = \\ &= \left[\int \nabla A(\phi^t\vec{x}) T^t_{\vec{x}}\rho(\mathrm{d}\vec{x})\right]\delta\vec{x}_0 + \int_0^t \left[\int \nabla A(\phi^{t-\tau}\vec{x}) T^{t-\tau}_{\vec{x}} w(\vec{x})\rho(\mathrm{d}\vec{x})\right]\delta\vec{f}(\tau)\mathrm{d}\tau, \end{aligned} \tag{7.11}$$

where in the second line we used the fact that ρ is invariant measure for ϕ^t. Finally, one can write the linear fluctuation-dissipation formula as

$$\begin{aligned} \delta_t\rho(A) &= P(t)\delta\vec{x}_0 + \int_0^t R(t-\tau)\delta\vec{f}(\tau), \\ P(t) &= \int \nabla A(\phi^t\vec{x}) T^t_{\vec{x}}\rho(\mathrm{d}\vec{x}), \\ R(t) &= \int \nabla A(\phi^t\vec{x}) T^t_{\vec{x}} w(\vec{x})\rho(\mathrm{d}\vec{x}). \end{aligned} \tag{7.12}$$

Using ergodicity of ρ and replacing space averages with time averages, one can write the linear response operators $P(t)$ and $R(t)$ as

$$\begin{aligned} P(t) &= \lim_{s\to\infty}\frac{1}{s}\int_0^s \nabla A(\vec{x}(t+\tau)) T^t_{\vec{x}(\tau)}\mathrm{d}\tau, \\ R(t) &= \lim_{s\to\infty}\frac{1}{s}\int_0^s \nabla A(\vec{x}(t+\tau)) T^t_{\vec{x}(\tau)} w(\vec{x}(\tau))\mathrm{d}\tau. \end{aligned} \tag{7.13}$$

7.2. **Linear response for maps.** Consider the following situation: we have a dynamical system of the form

$$\vec{x}_{n+1} = \vec{F}(\vec{x}_n), \tag{7.14}$$

with the ergodic invariant probability measure ρ, such that for any observable $A(\vec{x})$

$$\rho(A) = \int A(\vec{x})\rho(\mathrm{d}x) = \int A(\vec{F}^k(\vec{x}))\rho(\mathrm{d}\vec{x}), \quad \forall k. \tag{7.15}$$

Now, assume that a small forcing is introduced into the right-hand side of (7.14), together with a small perturbation in the initial condition:

$$\vec{x}_{n+1} = \hat{\vec{F}}(\vec{x}_n) = \vec{F}(\vec{x}_n) + w(\vec{x}_n)\delta\vec{F}_n, \tag{7.16}$$

where $w(\vec{x})$ is a matrix, and $\delta\vec{F}_n$ is a small time-dependent forcing. The question is: how the quantity

$$\delta_k\rho(A) = \int A(\hat{\vec{F}}^k(\vec{x}))\rho(\mathrm{d}\vec{x}) - \int A(\vec{F}^k(\vec{x}))\rho(\mathrm{d}\vec{x}) \tag{7.17}$$

is going to respond, and is there a suitable linear (with respect to small perturbations) approximation to this response?

Formally expanding (7.17) with respect to $\delta\vec{x}_k = \hat{\vec{F}}^k(\vec{x}_0) - \vec{F}^k(\vec{x}_0)$ and discarding higher order terms, we obtain

$$\delta_k\rho(A) = \int \nabla A(\vec{F}^k(\vec{x}))\delta\vec{x}_k\rho(\mathrm{d}\vec{x}). \tag{7.18}$$

Now one needs to find a suitable linear approximation for $\delta\vec{x}_k$. Subtracting (7.14) from (7.16), expanding with respect to $\delta\vec{x}$ and discarding higher order terms, we find

$$\delta\vec{x}_{n+1} = T(\vec{x}_n)\delta\vec{x}_n + w(\vec{x}_n)\delta\vec{F}_n, \tag{7.19}$$

where $T(\vec{x}_n)$ is the one-step tangent map of $\vec{F}(\vec{x}_n)$ (Jacobian matrix of $\vec{F}(\vec{x}_n)$, as usual). Due to linearity of (7.19), one finds

$$\delta\vec{x}_n = T^n_{\vec{x}_0}\delta\vec{x}_0 + \sum_{k=0}^{n-1} T^{n-k-1}_{\vec{x}_{k+1}} w(\vec{x}_k)\delta F_k. \tag{7.20}$$

Substituting (7.20) into (7.18), one obtains

(7.21)

$$\delta_n\rho(A) = \left[\int \nabla A(\vec{F}^n(\vec{x}))T^n_{\vec{x}}\rho(\mathrm{d}\vec{x})\right]\delta x_0 + \sum_{k=0}^{n-1}\left[\int \nabla A(\vec{F}^n(\vec{x}))T^{n-k-1}_{\vec{F}^{k+1}(\vec{x})} w(\vec{F}^k(\vec{x}))\rho(\mathrm{d}\vec{x})\right]\delta\vec{F}_k =$$
$$\left[\int \nabla A(\vec{F}^n(\vec{x}))T^n_{\vec{x}}\rho(\mathrm{d}\vec{x})\right]\delta x_0 + \sum_{k=0}^{n-1}\left[\int \nabla A(\vec{F}^{n-k}(\vec{x}))T^{n-k-1}_{\vec{F}(\vec{x})} w(\vec{x})\rho(\mathrm{d}\vec{x})\right]\delta\vec{F}_k,$$

where in the second line we used the fact that ρ is invariant measure for $\vec{F}$. Finally, one can write the linear fluctuation-dissipation formula as

$$\begin{aligned} \delta_n\rho(A) &= P_n\delta\vec{x}_0 + \sum_{k=0}^{n-1} R_{n-k}\delta\vec{F}_k, \\ P_n &= \int \nabla A(\vec{F}^n(\vec{x}))T^n_{\vec{x}}\rho(\mathrm{d}\vec{x}), \\ R_n &= \int \nabla A(\vec{F}^n(\vec{x}))T^{n-1}_{\vec{F}(\vec{x})} w(\vec{x})\rho(\mathrm{d}\vec{x}). \end{aligned} \tag{7.22}$$

Using ergodicity of ρ and replacing space averages with time averages, one can write the linear response operators P_n and R_n as

$$\begin{aligned} P_n &= \lim_{s\to\infty}\frac{1}{s}\sum_{k=1}^{s} \nabla A(\vec{x}_{n+k})T^n_{\vec{x}_k}, \\ R_n &= \lim_{s\to\infty}\frac{1}{s}\sum_{k=1}^{s} \nabla A(\vec{x}_{n+k})T^{n-1}_{\vec{x}_{k+1}} w(\vec{x}_k). \end{aligned} \tag{7.23}$$

For chaotic maps, the computation of P_n and R_n above becomes numerically unstable for large n, due to the fact that some entries of the multistep tangent map $T^n_{\vec{x}}$ grow exponentially fast with increasing n.

References

[1] E. Lorenz. Deterministic nonperiodic flow. *J. Atmos. Sci.*, 20, 1963.