Lectures on Dynamical Systems

Fraydoun Rezakhanlou
Departmet of Mathematics, UC Berkeley

December 3, 2010

PART I

1 Systems of differential equations

Consider the differential equation

$$(1.2) \qquad \frac{dx}{dt} = f(x,t), \;\; x \in U, \;\; t \in \mathbb{R},$$

where $U \subseteq \mathbb{R}^d$ is an open set, $x : [t_1, t_2] \to U$ is a differentiable function and $f : U \times \mathbb{R} \to \mathbb{R}^d$ is a continuous function. The main question is this: Given an initial condition $x(t_0) = a$, does (2.1) possess a unique solution that is defined for all times? Before answering this, let us examine some examples.

Example 1.1.

(i) Assume $d = 1$ and consider $\frac{dx}{dt} = x^2$ subject to the initial condition $x(0) = a$ with $a \neq 0$. Then $x(t) = (a^{-1} - t)^{-1}$ is a solution. This solution blows up at time $\frac{1}{a}$. Hence the ODE does not have a globally defined solution.

(ii) Consider $\frac{dx}{dt} = \sqrt{|x|}$ again in dimension 1. Consider the initial condition $x(0) = 0$. Pick $\alpha_1 < 0 < \alpha_2$ and define $x(t) = \begin{cases} \frac{1}{4}(t-\alpha_2)^2 & \text{for } t \geq \alpha_2 \\ -\frac{1}{4}(t-\alpha_1)^2 & \text{for } t \leq \alpha_1 \end{cases}$, and $x(t) = 0$ for $t \in (\alpha_1, \alpha_2)$. This $x(\cdot)$ is a solution. Here for the initial condition $x(0) = 0$, there are infinitely many solutions. □

From the above examples we learn that some conditions are needed on the function f in order to guarantee the existence of a unique globally defined solution for a given initial condition. What is responsible for the blow-up in Example 1(i) is the fact that the velocity or the growth rate is quadratic. Less is needed to have a blow-up as the following Exercise indicates.

Exercise 1.2. Take any continuous function $f : \mathbb{R} \to (0, \infty)$. Consider the equation $\frac{dx}{dt} = f(x)$, subject to the initial condition $x(0) = a$. Show that we have a blow-up if and only if $\int_a^\infty \frac{dx}{f(x)} < \infty$.

The non-uniqueness in Example 1.1(ii) stems from the fact that the function $f = \sqrt{|x|}$ has a cusp at 0, i.e., $f'(0\pm) = \pm\infty$. To avoid both cusp and super linear growth, it suffices to assume that the function f is Lipschitz. More precisely, we say f is (uniformly) *Lipschitz* if there exists a constant L such that for every $x, y \in U$ and $t \in \mathbb{R}$,

$$(1.2) \qquad |f(x,t) - f(y,t)| \leq L|x - y|.$$

Theorem 1.3. *Suppose f is Lipschitz. Then* (1.1) *has a unique solution for every initial condition.*

Let us first address the question of uniqueness.

Lemma 1.4. *Let x and y be two solutions with f satisfying* (1.2). *Then*

$$|x(t) - y(t)| \leq e^{L|t-t_0|}|x(t_0) - y(t_0)|. \tag{1.3}$$

Note that (1.3) implies the uniqueness part of Theorem 1.3 by assuming $x(t_0) = y(t_0) = a$.

Proof of Lemma 1.4. Set $\varphi(t) = |x(t) - y(t)|^2$. We have

$$\begin{aligned}
\varphi'(t) &= 2(x(t) - y(t)) \cdot \left(\tfrac{dx}{dt}(t) - \tfrac{dy}{dt}(t)\right) \\
&= 2(x(t) - y(t)) \cdot (f(x(t), t) - f(y(t), t) \\
&\leq 2L|x(t) - y(t)|^2 = 2L\varphi(t)
\end{aligned}$$

by Schwartz Inequality and (1.2). As a result,

$$\frac{d}{dt}(\varphi(t)e^{-2tL}) \leq 0.$$

This means that $\varphi(t)e^{-2tL}$ is a non-increasing function of t. Hence for $t > t_0$,

$$\varphi(t)e^{-2tL} \leq \varphi(t_0)e^{-2t_0L}.$$

This implies (1.3) for $t > t_0$. For $t < t_0$, first observe that we also have $\varphi' \geq -2L\varphi$. So, the function $\varphi(t)e^{2Lt}$ is now non-decreasing. Hence,

$$\varphi(t)e^{2tL} \leq \varphi(t_0)e^{2t_0L},$$

whenever $t < t_0$. This completes the proof. □

Before turning to the question of existence, let us mention that in practice we may only know an approximation of what appears on the right-hand side of (1.1). The following lemma asserts that by solving (1.1) with an error both on the right-hand side and the initial data, we are making a small error on the solution.

Lemma 1.5. *Assume*

$$\frac{dx}{dt} = f(x, t) + E_1 \quad \frac{dy}{dt} = f(y, t) + E_2$$

with $|E_1| + |E_2| < \epsilon$. *Then*

$$|x(t) - y(t)| \le |x(t_0) - y(t_0)|e^{L|t-t_0|} + \frac{\epsilon}{L}(e^{L|t-t_0|} - 1). \tag{1.4}$$

Proof. We may try $\varphi(t) = |x(t) - y(t)|^2$ as in Lemma 1.4 to write

$$\begin{aligned} \varphi'(t) &= 2(x(t) - y(t)) \cdot (f(x(t), t) - f(y(t), t) + E_1 - E_2) \\ &\le 2L|x(t) - y(t)|^2 + 2\epsilon|x(t) - y(t)| \\ &= 2L\varphi(t) + 2\epsilon\sqrt{\varphi(t)}. \end{aligned}$$

This does not work for us as before because $\sqrt{\varphi(t)}$ could be much larger than $\varphi(t)$ if $\varphi(t)$ is small. Instead we set $\psi(t) = |x(t) - y(t)|$ and write

$$\begin{aligned} \psi(t) &= \left| x(t_0) - y(t_0) + \textstyle\int_{t_0}^t [f(x(s), s) - f(y(s), s) + E_1 - E_2]ds \right| \\ &\le \psi(t_0) + L \textstyle\int_{t_0}^t \psi(s)ds + \epsilon(t - t_0). \end{aligned} \tag{1.5}$$

The good news is that $\sqrt{\psi}$ does not show up as in the previous attempt. The bad news is that ψ is no longer a differentiable function because of the absolute value. However if we set

$$D^+\psi(t_0) = \limsup_{t \downarrow t_0} \frac{\psi(t) - \psi(t_0)}{t - t_0}$$

then (1.5) implies that for every t_0,

$$D^+\psi(t_0) \le L\psi(t_0) + \epsilon. \tag{1.6}$$

By Grownall's inequality, (1.6) implies that

$$\psi(t) \le e^{L|t-t_0|}\psi(t_0) + \frac{\epsilon}{L}(e^{L|t-t_0|} - 1) \tag{1.7}$$

which is exactly (1.4). To establish (1.7) first observe

$$D^+(e^{-Lt}\psi(t)) \le \epsilon e^{-Lt}$$

and this in turn implies

$$D^+\left(e^{-Lt}\psi(t) + \frac{\epsilon}{L}e^{-Lt}\right) \le 0.$$

This and Exercise 1.6 below implies that if $t > t_0$, then

$$e^{-Lt}\psi(t) + \frac{\epsilon}{L}e^{-Lt} \le e^{-Lt_0}\psi(t_0) + \frac{\epsilon}{L}e^{-Lt_0}$$

and this is exactly (1.4). $\square$

Exercise 1.6. Let ψ be a continuous function with $D^+\psi \leq 0$. Show that ψ is non-increasing. (*Hint:* First define $\psi^\delta(t) = \psi(t) - \delta t$ with $\delta > 0$. Show that ψ^δ is decreasing. Then send δ to 0.)

We now turn to the question of existence. For simplicity let us assume that $f(x,t) = f(x)$ is independent of t. Also assume that $x(0) = a$. To find a solution, let us design an approximation scheme. For $n > 0$, define $x^n(\cdot)$ by the requirement that $x_n(0) = a$, and

$$\frac{dx^n}{dt}(t) = \begin{cases} f\left(x_n\left(\frac{j}{n}\right)\right) & \text{if } \frac{j}{n} < t < \frac{j+1}{n},\ j \geq 0, \\ f\left(x_n\left(\frac{j+1}{n}\right)\right) & \text{if } \frac{j}{n} < t < \frac{j+1}{n},\ j < 0. \end{cases}$$

Clearly such x_n is piecewise linear and can be constructed. The existence of a solution is an immediate consequence of Lemma 1.7.

Lemma 1.7. *The sequence $\{x_n\}$ is Cauchy and if $x_n \to x$, then x solves* (1.1)*.*

Proof. First we establish the equicontinuity of the sequence $\{x_n\}$. For this it suffices to show that the sequence $\left\{x_n^j = x_n\left(\frac{j}{n}\right)\right\}$ is uniformly bounded for $\left|\frac{j}{n}\right| \leq T$. To see this, observe that if $j > 0$, then

$$\begin{aligned} |x_n^j - a| &= |x_n^j - x_n^{j-1}| + |x_n^{j-1} - a| = \left|\tfrac{1}{n} f(x_n^{j-1})\right| + |x_n^{j-1} - a| \\ &\leq \tfrac{1}{n}|f(a)| + \tfrac{1}{n}|f(x_n^{j-1}) - f(a)| + |x_n^{j-1} - a| \\ &\leq \tfrac{1}{n}|f(a)| + \left(\tfrac{L}{n} + 1\right)|x_n^{j-1} - a| \leq \dots \\ &\leq \tfrac{1}{n}|f(a)|\left(1 + \left(1 + \tfrac{L}{n}\right) + \cdots + \left(1 + \tfrac{L}{n}\right)^{j-1}\right) \\ &= \tfrac{|f(a)|}{L}\left(\left(1 + \tfrac{L}{n}\right)^j - 1\right) \leq \tfrac{|f(a)|}{L}(e^{Lj/n} - 1) \\ &\leq \tfrac{|f(a)|}{L}(e^{LT} - 1). \end{aligned}$$

The case $j < 0$ can be treated likewise. By Ascoli's theorem, we can find a convergent subsequence of $\{x_n\}$. We continue to write $\{x_n\}$ for such a subsequence. Note that

$$\begin{aligned} \left|\tfrac{dx_n}{dt} - f(x_n(t))\right| &= |f(x_n(j/n)) - f(x_n(t))| \leq L|x_n(j/n) - x_n(t)| \\ &\leq \tfrac{L}{n}|f(x_n(j/n))|, \end{aligned}$$

provided that $t \in \left(\frac{j}{n}, \frac{j+1}{n}\right)$, $j \geq 0$. Hence

$$x_n(t) = a + \int_0^t f(x_n(s))ds + O\left(\frac{1}{n}\right).$$

If $x_n \to x$, then

$$x(t) = a + \int_0^t f(x(s))ds,$$

for $t > 0$. The case $t < 0$ can be treated likewise. $\square$

Exercise 1.8.

- (i) Carry out the proof for the time-dependent case.
- (ii) Given a square matrix A, define its norm $\|A\| = \max_v |Av|/|v|$. Show that $\|A + B\| \leq \|A\| + \|B\|$ and $\|AB\| \leq \|A\| \, \|B\|$. Moreover, if A is a symmetric matrix, then $\|A\|$ is the absolute value of the largest eigenvalue of A.

Assuming that f is continuous and x-Lipschitz, we have showed the existence of a unique solution. Let us write $\phi_{t_0}^t(a)$ for such a solution, to display its dependence on the initial data a. Since both $x(t) = \phi_{t_0}^t(a)$ and $y(t) = \phi_{t_1}^t(\phi_{t_0}^{t_1}(a))$ solve (1.1) and satisfy $x(t_1) = y(t_1)$, we deduce

$$\phi_{t_0}^t(a) = \phi_{t_1}^{t_1} \circ \phi_{t_0}^{t_1}(a), \tag{1.8}$$

whenever $t_0 < t_1 < t$. This is the *group property* of the family $\{\phi_{t_0}^{t_1}\}$.

Remark 1.9. When $U = \mathbb{R}^d$ and f is Lipschitz, our existence proof implies that the solutions exist for all time. This may not be true if $U \neq \mathbb{R}^d$. If (t_1, t_2) is the largest existence interval with, say, $t_2 < \infty$, then $\lim_{t \to t_2} x(t)$ exists and belongs to ∂U. $\square$

Note that $\phi_{t_0}^t = \phi_0^{t-t_0}$ when f is independent of t. In this case, we simply write ϕ^t for ϕ_0^t Now (1.8) becomes

$$\phi^t \circ \phi^s = \phi^{t+s}. \tag{1.9}$$

Also note that by Lemma 1.4,

$$|\phi^t(x) - \phi^t(y)| \leq e^{Lt}|x - y|. \tag{1.10}$$

This certainly implies the Lipschitzness of ϕ in the x-variable. We can say more if f is a differentiable function.

Theorem 1.10. *If f is C^k (k-times continuously differentiable), then ϕ is C^k.*

Proof. If we already know that ϕ is differentiable in x, then the ODE

$$\frac{d}{dt}\phi^t(x) = f(\phi^t(x))$$

can be differentiated to yield

$$\frac{d}{dt}D\phi^t(x) = Df(\phi^t(x))D\phi^t(x). \tag{1.11}$$

If $B(t) = Df(\phi^t(x))$ and $A(t) = D\phi^t(x)$ for a given x, then (1.11) means that A is a (matrix-valued) solution to the linear ODE

$$\begin{cases} \frac{dA}{dt} = B(t)A \\ A(0) = I \end{cases} \tag{1.12}$$

Note that the function $(A, t) \mapsto B(t)A$ satisfies the Lipschitz property (1.2) so long as t is restricted to a bounded interval. Hence (1.12) must have a unique solution. Hence we already have a candidate for $D\phi^t(x)$, namely the solution of (1.12). From this we expect to have the differentiability of ϕ^t with respect to x.

To turn the above heuristic reasoning into a rigorous proof, we go back to Lemma 1.7 and use $x_n(\cdot)$. In fact the approximation scheme used in Lemma 1.7 produces a flow that is denoted by $\phi_n^t(x)$. One can readily check that $D\phi_n^t(x)$ exists. Indeed

$$\frac{d}{dt}D\phi_n^t(x) = \begin{cases} Df\left(\phi_n^{\frac{j}{n}}(x)\right) D\phi_n^{\frac{j}{n}}(x) & t \in \left(\frac{j}{n}, \frac{j+1}{n}\right), \ j \geq 0, \\ Df\left(\phi_n^{\frac{j+1}{n}}(x)\right) D\phi_n^{\frac{j+1}{n}}(x) & t \in \left(\frac{j}{n}, \frac{j+1}{n}\right), \ j < 0. \end{cases} \tag{1.13}$$

Note that by Lemma 1.7, if $\left|\frac{j}{n}\right| \leq T$ and $|x| \leq T$, then

$$\left|\phi_n^{\frac{j}{n}}(x)\right| \leq \left(\sup_{|x| \leq T} |f(x)|\right)(e^{LT} - 1) + T =: B_T.$$

We also set

$$C_T = \max\{\|Df(x)\| : |x| \leq B_T\}.$$

This implies that

$$\begin{aligned} \left|D\phi_n^{\frac{j}{n}}(x)\right| &\leq \left|D\phi_n^{\frac{j}{n}}(x) - D\phi_n^{\frac{j-1}{n}}(x)\right| + \left|D\phi_n^{\frac{j-1}{n}}(x)\right| \\ &\leq \tfrac{1}{n}C_T\left|D\phi_n^{\frac{j-1}{n}}(x)\right| + \left|D\phi_n^{\frac{j-1}{n}}(x)\right| \\ &= \left(1 + \tfrac{1}{n}C_T\right)\left|D\phi_n^{\frac{j-1}{n}}(x)\right| \\ &\leq \cdots \leq \left(1 + \tfrac{1}{n}C_T\right)^j \leq e^{\frac{j}{n}C_T} \leq e^{TC_T}. \end{aligned}$$

From this and (1.13) we deduce the equicontinuity of $D_x\phi^n(x,t)$ in the t-variable. This allows us to pass to the limit as $n\to\infty$. Let $\psi(x,t)$ be a limit along a subsequence for a fixed x. Then

$$\text{(1.14)}\qquad \begin{cases} \frac{d\psi}{dt}(x,t) = Df(\phi^t(x))\psi(x,t) \\ \psi(x,0) = I. \end{cases}$$

But (1.14) has a unique solution. Hence the full sequence $\{D_x\phi^{(n)}\}$ converges to ψ uniformly in t for every x. On the other hand $\phi^{(n)}$ converges uniformly to ϕ. Since the derivatives converge to ψ, we deduce that $D_x\phi$ must exist and that $D_x\phi = \psi$. As a result, ϕ is differentiable in x.

As our next step we would like to show that indeed $\psi = D_x\phi$ is a continuous function. Recall that we are assuming $f\in C^1$. We have

$$\|\psi(x,t)-\psi(y,t)\| \le \left\| \int_{t_0}^t (Df(\phi^s(x))\psi(x,s) - Df(\phi^s(y))\psi(y,s))ds \right\| + \|\psi(x,t_0)-\psi(y,t_0)\|.$$

Set $\tau(t) = \|\psi(x,t)-\psi(y,t)\|$. Set

$$C_1 = \sup_{|s|,|x|\le T} \|Df(\phi^s(x))\|,\ C_2 = \sup_{|s|,|x|\le T} \|\psi(x,s)\|.$$

Then

$$\tau(t) \le \tau(t_0) + C_1\int_{t_0}^t \tau(s)ds + C_2\int_{t_0}^t \|Df(\phi^s(x)) - Df(\phi^s(y))\|.$$

We know that f is Lipschitz with a Lipschitz constant L. Hence

$$|\phi^s(x)-\phi^s(y)| \le e^{L|s|}|x-y|.$$

Since Df is continuous, we learn that for every $\epsilon>0$, there exists $\delta>0$ such that if $|x-y|<\delta$, then

$$\|Df(\phi^s(x)) - Df(\phi^s(y))\| \le \epsilon$$

for s in a bounded interval. Hence for $|x-y|<\delta$,

$$\tau(t) \le \tau(t_0) + C_1\int_{t_0}^t \tau(s)ds + C_2\epsilon(t-t_0)$$

for $t>t_0$ and both t and t_0 in a bounded interval. Using Grownwall's inequality,

$$\tau(t) \le e^{C_1|t-t_0|}\tau(t_0) + \frac{C_2\epsilon}{C_1}(e^{C_1|t-t_0|}-1).$$

Note $\tau(0) = 0$. Thus

$$\|\psi(x,t) - \psi(y,t)\| \le \frac{C_2\epsilon}{C_1}(e^{C_1|t|} - 1)$$

for $|x - y| < \delta$. This proves the continuity of ψ.

So far we have established the theorem for the case $k = 1$. For higher k, consider the system

$$\frac{dx}{dt} = f(x), \quad \frac{d\xi}{dt} = Df(x)\xi. \tag{1.15}$$

If f is C^2, then (f, Df) is C^1. Hence (1.15) has a C^1-flow. If its flow is denoted by $\hat{\phi}^t$, then $\hat{\phi}^t(a, I) = (\phi^t(a), D\phi^t(a))$. Since $\xi = D_x\phi^t$, we learn that ϕ^t is C^2 in x-variable. Using (1.15) again we can show that ϕ^t is C^2 in both variables. This proves the theorem for $k = 2$. The larger k can be treated by induction. □

The equation (1.14) plays an important role in studying the stability of the equation $dx/dt = f(x)$. More precisely, if x, y are two solutions to

$$\frac{dx}{dt} = f(x) \tag{1.16}$$

with $y = x + \delta v$ then $\frac{d}{dt}(\delta v) \approx Df(x)(\delta v)$. Hence the relevant problem to study is now

$$\frac{dv}{dt} = Df(x)v \tag{1.17}$$

where $x = x(t)$ is a solution to (1.16). This is closely related to (1.15) or (1.14) because $v(t) = D\phi^t(a)v(0)$ where $a = x(0)$. If $x(\cdot)$ happens to be a fixed point a, then (1.17) becomes

$$\frac{dv}{dt} = Av \tag{1.18}$$

with $A = Df(a)$. The equation (1.18) will be studied in Section 2. The study of the equation (1.16) when $x(\cdot)$ is a periodic orbit, is the subject of Section 3 and is known as Floquet Theory.

In the case of a discrete dynamical system, we start with a function $f : \mathbb{R}^d \to \mathbb{R}^d$ and set $x_{n+1} = f(x_n)$, so that $x_n = f^n(a)$ for an initial choice a. Now if $\xi_n = Df^n(a)$, then

$$\xi_{n+1} = Df(x_n)\xi_n. \tag{1.19}$$

Imagine now that $\{y_n\}$ is another orbit and that y_n is close to x_n. If $y_n = x_n + \delta v_n$, then $y_{n+1} = f(x_n + \delta v_n) \approx Df(x_n)\delta v_n + f(x_n)$. Since $y_{n+1} = x_{n+1} + \delta v_{n+1}$, we deduce that

$\delta v_{n+1} = Df(x_n)\delta v_n$. Motivated by this, let us study the non-autonomous linear dynamical system

$$v_{n+1} = Df(x_n)v_n. \tag{1.20}$$

provided that $\{x_n\}$ is an orbit of the original dynamical system $x_{n+1} = f(x_n)$. If $\{x_n\}$ is a fixed point, i.e., $f(a) = a$ and $x_n = a$ for all n, then (1.20) becomes $v_{n+1} = Av_n$ for $A = Df(a)$. Hence $v_n = A^n v$ for an initial choice v. The behavior of v_n as $n \to \infty$ depends on the eigenvalues of A and will be treated in Section 2. Again, if $\{x_n\}$ is a periodic orbit, then the sequence $\{v_n\}$ is analyzed by Floquet Theory and will be discussed in Section 3.

2 Linear Systems

In this section we study linear dynamical systems. In the discrete case we have a $d \times d$ matrix and we are interested in the behavior of the sequence $(A^n x : n \in \mathbb{Z})$ for a nonzero vector $x \in \mathbb{R}^d$. In the continuous case we study the flow of the ODE

$$\frac{dx}{dt} = Ax. \tag{2.1}$$

We start with the discrete case. Using a Jordan normal form we can find a basis of $\mathbb{R}^d$ such that the transformation $x \mapsto Ax$ has the following matrix representation:

$$A = \begin{bmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_n \end{bmatrix} \tag{2.2}$$

where each block is either of the form

$$A_j = \begin{bmatrix} \lambda & 0 & \dots & 0 & 0 \\ 1 & \lambda & \dots & 0 & 0 \\ \vdots & & & \vdots & \vdots \\ 0 & 0 & \dots & 1 & \lambda \end{bmatrix} \tag{2.3}$$

with λ a real eigenvalue of A, or else of the form

$$\begin{bmatrix} \alpha & -\beta & 0 & 0 & & & & & 0 \\ \beta & \alpha & 0 & 0 & & & & & \\ 1 & 0 & \alpha & -\beta & & & & & \\ 0 & 1 & \beta & \alpha & & & & & \\ & & & & \ddots & & & & \\ & & & & & 1 & 0 & \alpha & -\beta \\ 0 & & & & & 0 & 1 & \beta & \alpha \end{bmatrix}$$

with $\lambda = \alpha + i\beta$, $\bar{\lambda} = \alpha - i\beta$ a pair of complex eigenvalues of A. For each eigenvalue λ we write E_λ for the corresponding invariant subspace. In the case of real λ,

$$E_\lambda = \{v \in \mathbb{R}^d : (A - \lambda)^k v = 0 \text{ for some } k \in \mathbb{N}\}.$$

Indeed if $v \in E_\lambda$, then $(A-\lambda)^k Av = A(A-\lambda)^k v = 0$. We now set

$$E^- = \bigoplus_{|\lambda|<1} E_\lambda, \quad E^0 = \bigoplus_{|\lambda|=1} E_\lambda, \quad E^+ = \bigoplus_{|\lambda|>1} E_\lambda. \tag{2.4}$$

Since $T(E_\lambda) \subseteq E_\lambda$, we have that $T(E^\pm) \subseteq E^\pm$, and $T(E^0) \subseteq E^0$.

Theorem 2.1. *If $x \in E^-$, then $\lim_{n\to\infty} T^n(x) = 0$ exponentially fast. If $x \in E^+$, then $\lim_{n\to\infty} |T^n(x)| = \infty$ exponentially fast.*

Proof. By invariance, it suffices to verify the theorem when A is of the form (2.3) or (2.2). First assume that A is of the form (2.2) and that $x \in E_\lambda$ with $|\lambda| < 1$. For such a transformation,

$$T \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ x_1 + \lambda x_2 \\ \vdots \\ x_{d-1} + \lambda x_d \end{bmatrix}.$$

To obtain a contraction, we would like to replace the off-diagonal entries with some small number. To achieve this, let us try a diagonal change of coordinates of the form

$$\varphi \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix} = \begin{bmatrix} \mu_1 x_1 \\ \vdots \\ \mu_d x_d \end{bmatrix}.$$

We have

$$\varphi^{-1} T \varphi \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \frac{\mu_1}{\mu_2} x_1 + \lambda x_2 \\ \vdots \\ \frac{\mu_{d-1}}{\mu_d} x_{d-1} + \lambda x_d \end{bmatrix}.$$

This suggests choosing μ_i so that $\frac{\mu_1}{\mu_2}, \dots, \frac{\mu_{d-1}}{\mu_d}$ are small. This can be done if $\mu_j = \delta^{-j}$ for a small δ. Set $\hat{T} = \varphi^{-1} T \varphi$. It suffices to verify the theorem for

$$\hat{T} \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \delta x_1 + \lambda x_2 \\ \vdots \\ \delta x_{d-1} + \lambda x_d \end{bmatrix}.$$

Here λ is real and $|\lambda| < 1$. Pick $\gamma \in (0, |\lambda|)$. Then we can choose δ sufficiently small so that

$$|\hat{T}(x)| \le \gamma |x|,$$

simply because $\sum_i (\delta x_j + \lambda x_{j+1})^2 \le \delta^2 |x|^2 + \lambda^2 |x|^2 + 2\lambda\delta |x|^2$. Hence $|\hat{T}^n(x)| \le \gamma^n |x|$ for some $0 < \gamma < 1$.

The case $|\lambda| > 1$ can be treated likewise. Also for A of the form (2.4), a similar argument applies. □

Remark 2.1. Our proof indicates that the following is true: Pick γ such that $\gamma < |\lambda|$ for every eigenvalue λ with $|\lambda| < 1$. Then there exists a basis $\{u_1, \dots, u_r\}$ for E^+ such that if $|\sum_i x_i u_i| = (\sum_i x_i^2)^{1/2}$ denotes the standard norm with respect to this basis, then $|A^n x| \le \gamma^n |x|$ for every $n \in \mathbb{N}$ and $x \in E^-$. A similar comment applies to the space E^+. □

Exercise 2.2. Suppose A is of the form (2.3) with $\lambda = \pm 1$. Show that $|A^n x| = O(n^{k-1})$ where k is the size of the matrix A. If A is of the form (2.4) with $|\lambda| = 1$, then show that $|A^n x| = O(n^{k-1})$ where now A is of the size $2k \times 2k$.

We say a linear transformation T is *hyperbolic* if $E^0 = \{0\}$. From Theorem 2.2 we learn that we have a simple picture for the behavior of $T^n(x)$. If $x = x^+ + x^-$ with $x^+ \in E^+$, $x^- \in E^-$, then $T^n(x) = T^n(x^+) + T^n(x^-)$, with $|T^n(x^+)| \to \infty$ and $|T^n(x^-)| \to 0$ as $n \to +\infty$.

The situation is drastically different if T is not hyperbolic. For example if $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_1 \\ -x_2 \end{bmatrix}$, then the only eigenvalue is -1 and every orbit $(T^n(x) : n \in \mathbb{Z})$ with $x \neq 0$ is periodic of period 2. To have another example, assume that T corresponds to a matrix of the form

$$\begin{bmatrix} R_1 & & 0 \\ & \ddots & \\ 0 & & R_n \end{bmatrix}$$

where each $R_i = \begin{bmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{bmatrix}$. In this case all eigenvalues are of norm 1 and if $x = [x_1, \ldots, x_n]^t$, with $x_j \in \mathbb{R}^2$, then $T(x) = [R_1 x_1, \ldots, R_n x_n]^t$. Since each R_i is a rotation, the torus

$$\mathbb{T}(\rho_1, \ldots, \rho_n) = \{(x_1, \ldots, x_n) : |x_1| = \rho_1, \ldots, |x_n| = \rho_n\},$$

is invariant. If we assume that all ρ_j's are nonzero, then $\mathbb{T}$ is an n-dimensional torus. The effect of T on $\mathbb{T}(\rho_1, \ldots, \rho_n)$ is simply a translation in the following sense: If $x_j = \rho_j e^{ia_j}$ then

$$T(x) = \begin{bmatrix} \rho_1 e^{i(a_1+\theta_1)} \\ \vdots \\ \rho_n e^{i(a_n+\theta_n)}. \end{bmatrix}$$

So in terms of angles, we simply have

$$(a_1, a_2, \ldots, a_n) \mapsto (a_1 + \theta_1, \ldots, a_n + \theta_n).$$

As we will see later, the orbits of this transformation are dense if $\theta_1, \ldots, \theta_n, 1$ are rationally independent. That is, if $c_1\theta_1 + \cdots + c_n\theta_n + c_{n+1} = 0$ for integers $c_1, \ldots, c_{n+1}$, then $c_1 = \cdots = c_{n+1} = 0$.

We now turn to (2.1). We again use the form (2.2) with A_j as in (2.3) and (2.4). But this time

$$F^+ = \bigoplus_{\operatorname{Re}\lambda > 0} E_\lambda \quad F^0 = \bigoplus_{\operatorname{Re}\lambda = 0} E_\lambda \quad F^- = \bigoplus_{\operatorname{Re}\lambda < 0} E_\lambda. \tag{2.5}$$

We now say the system (2.1) is *hyperbolic* if $F^0 = \{0\}$.

Theorem 2.3. *If $x \in F^+$, then $\lim_{t\to+\infty} |\phi^t(x)| = +\infty$ and $\lim_{t\to-\infty} |\phi^t(x)| = -\infty$.*

Proof. Let us assume that T is given by a matrix A of the form

$$\begin{bmatrix} R & 0 & \ldots & 0 & 0 \\ \delta I & R & \ldots & 0 & 0 \\ & \ddots & & & \\ 0 & 0 & \ldots & \delta I & R \end{bmatrix}$$

where $R = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}$ with $\alpha > 0$ and $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Set

$$\tilde{A} = \begin{bmatrix} R & 0 & \ldots & 0 & 0 \\ 0 & R & \ldots & 0 & 0 \\ & & & \ddots & \\ 0 & 0 & \ldots & 0 & R \end{bmatrix}.$$

Let ϕ^t and $\tilde{\phi}^t$ be the flow of A and $\tilde{A}$ respectively. To compare ϕ^t with $\tilde{\phi}^t$, we calculate

$$\frac{d}{dt}|\phi^t\tilde{\phi}^{-t}a|^2 = 2(\phi^t\tilde{\phi}^{-t}a)\cdot(A-\tilde{A})(\phi^t\tilde{\phi}^{-t}a).$$

Since

$$(A-\tilde{A})\begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix} = -\delta \begin{bmatrix} 0 \\ x_1 \\ \vdots \\ x_{d-1} \end{bmatrix},$$

we have

$$|(A-\tilde{A})(x)| \le \delta|x|.$$

As a result,

$$\frac{d}{dt}|\phi^t\tilde{\phi}^{-t}a|^2 \le 2\delta|\phi^t\tilde{\phi}^{-t}a|^2.$$

Hence

$$|\phi^t\tilde{\phi}^{-t}a| \le e^{\delta t}|a|, \ t>0,$$

for every a. This can be rephrased as

$$|\phi^t a| \le e^{\delta t}|\tilde{\phi}^t a|, \ t>0.$$

In the same fashion

$$|\tilde{\phi}^t a| \le e^{\delta t}|\phi^t a|, \ t>0.$$

From this we can readily deduce

$$|\phi^t a| \le e^{(Re\lambda+\delta)t}|a|$$

for $t>0$. Now if $\operatorname{Re}\lambda<0$,then we have that $\lim_{t\to+\infty}|\phi^t a| = 0$. Since

$$|\phi^t a| \ge e^{(Re\lambda-\delta)t}|a|,$$

we deduce that

$$\lim_{t\to+\infty}|\phi^t a| = +\infty,$$

whenever $\operatorname{Re}\lambda > 0$. □

The linear systems of the form we have studied so far can be used to study nonlinear systems near a fixed point. Theorem 2.1 can be used to study the orbits of the system $x_{n+1} = f(x_n)$ near a fixed point. A celebrated theorem of *Hartman–Grobman* asserts that the nonlinear system and its linearization near a hyperbolic fixed point are equivalent:

Theorem 2.4. *Assume f is smooth with $f(a) = a$ and $A = Df(a)$ hyperbolic. Then there exists a homeomorphism h from a neighborhood U of a onto a neighborhood of the origin such that $f = h^{-1} \circ T \circ h$ in U where $T(x) = Ax$.*

Note that if $v_n = A^n v$ belongs to $h(U)$ then $h^{-1}(A^n v) = x_n$ belongs to U and $\{x_n\}$ is an orbit for the f-system. The reason is simply $f^n = h^{-1} \circ T^n \circ h$.

Given the equation $dx/dt = f(x)$ with $f(a) = 0$ and set $A = Df(a)$. Let us write ϕ^t and ψ^t for the flow associated with $dx/dt = f(x)$ and $dx/dt = Ax$ respectively. In the continuous case, the analogue of Theorem 2.4 is this:

Theorem 2.5. *Assume f is smooth and $f(a) = 0$. Assume that $A = Df(a)$ is hyperbolic in the sense that A has no purely imaginary eigenvalue. Then there exists a homeomorphism h from a neighborhood of a onto a neighborhood of the origin such that*

$$\phi^t = h^{-1} \circ \psi^t \circ h. \tag{2.10}$$

Exercise 2.6. Given a matrix A, use Jordan Normal Form Theorem to find a collection of numbers $l_1 < l_2 < \cdots < l_k$, positive integers $n_1, n_2, \ldots, n_k$ and linear subspaces $G^1, G^2, \ldots, G^k$ such that $\dim G^j = n_j$ and that if $v \in (G^1 \oplus G^2 \oplus \cdots \oplus G^j) - (G^1 \oplus \cdots \oplus G^{j-1})$ then

$$\lim_{n \to \infty} \frac{1}{n} \log |A^n v| = l_j.$$

(The numbers l_j are known as *Lyapunov exponents.*) □

Observe that if

$$A = \begin{bmatrix} R_1 & & 0 \\ & \ddots & \\ 0 & & R_n \end{bmatrix}$$

with

$$R_j = \begin{bmatrix} 0 & -\beta_j \\ \beta_j & 0 \end{bmatrix},$$

then

$$\phi_t \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \psi_t^1 x_1 \\ \vdots \\ \psi_t^n x_n \end{bmatrix},$$

where $x_1, \dots, x_n \in \mathbb{R}^2$ and $\psi_t^j z = \begin{bmatrix} \cos \beta_j t & -\sin \beta_j t \\ \sin \beta_j t & \cos \beta_j t \end{bmatrix} z$. Again the torus (4.5) is invariant and the flow on this torus is simply

$$(a_1, a_2, \dots, a_n) \mapsto (a_1 + \beta_1 t, \dots, a_n + \beta_n t).$$

In other words, we have a free motion on this torus. It turns out that the flow restricted to $\mathbb{T}^d$ is dense if and only if $(\beta_1, \dots, \beta_n)$ are linearly independent over rationals. Consider $T : \mathbb{T}^d \to \mathbb{T}^d$ defined by $T(a_1, \dots, a_n) = (a_1+\theta_1, \dots, a_n+\theta_n)$. We have the following theorem.

Theorem 2.7. *Suppose* $\theta_1, \theta_2, \dots, \theta_n, 1$ *are* rationally independent, *i.e.,* $\sum_{j=1}^d \lambda_j \theta_j$ *is not an integer for any* $\lambda = (\lambda_1, \dots, \lambda_d) \in \mathbb{Z}^d$ *with* $\lambda \neq 0$. *Then the orbit* $(T^n x : n \in \mathbb{Z}^+)$ *is dense in* $\mathbb{T}^d$ *for every* $x \in \mathbb{T}^d$. *Conversely, if an orbit is dense, then* $(\theta_1, \dots, \theta_n, 1)$ *are rationally independent.*

Let us use this theorem as an excuse to make a definition. We say a transformation $T : X \to X$ is *topologically transitive* if $(T^n(x) : n \in \mathbb{Z}^+)$ is dense for some $x \in X$.

Theorem 2.8. *Assume that* X *is locally compact, second countable, with no isolated point. Suppose* $T : X \to X$ *is continuous. Then the following statements are equivalent:*

(i) T *is topologically transitive.*

(ii) *For any pair of nonempty open sets* U *and* V, *there exists an integer* $N \geq 0$ *such that* $T^{-N}(U) \cap V \neq \emptyset$.

(iii) *If* U *is an open and* $T^{-1}(U) \subseteq U$, *then either* $U = \phi$ *or* U *is dense.*

Proof. (i) $\Rightarrow$ (ii). Take $x \in X$ with $\{T^n(x) : n \in \mathbb{Z}^+\}$ is dense. Then there are infinitely many indices $n_1, n_2 \in \mathbb{Z}^+$ such that $T^{n_1}(x) \in U$ and $T^{n_2}(x) \in V$. We may assume $n_1 \geq n_2$ so that $T^{n_2}(x) \in V \cap T^{-N}(U)$ for $N = n_1 - n_2$.

(ii) $\Rightarrow$ (i). Let $\{U_1, U_2, \dots\}$ be a countable open base for X. Find n_1 such that $T^{-n_1} U_2 \cap U_1 \neq \emptyset$. Find an open set V_2 such that $\bar{V}_2 \subseteq T^{-n_1} U_2 \cap U_1$ and $V_2 \neq \emptyset$. We then pick $n_2 \in \mathbb{Z}^+$

such that $T^{-n_2}U_3 \cap V_2 \neq \emptyset$ and a nonempty open set V_3 such that $\bar{V}_3 \subseteq T^{-n_2}U_3 \cap V_2$ etc. By induction, we find a sequence of open sets $U_1 = V_1 \supseteq \bar{V}_2 \supseteq V_2 \supseteq \bar{V}_3 \supseteq \dots$ with $V_j \neq \emptyset$ and a sequence of positive integers such that $\bar{V}_j \subseteq V_{j-1} \cap T^{-n_{j-1}}U_j$. Without loss of generality, we may assume $\bar{V}_1$ is compact. Let $A = \bigcap_{j=1}^{\infty} \bar{V}_j$. Evidently $A \neq \emptyset$ and if $x \in A$ then $T^{n_j}(x) \in U_{j+1}$ for all j, and $x \in U_1$. Hence $(T^n(x) : n \in \mathbb{Z}^+)$ is dense.

(ii) $\Rightarrow$ (iii). Suppose $U \neq \emptyset$ is open and invariant. We have $T^{-1}(U) \subseteq U$ which implies $T^{-N}(U) \subseteq U$ for $N \geq 0$. Take a nonempty open set V. For some $N \geq 0$ we have $\emptyset \neq T^{-N}(U) \cap V \subseteq U \cap V$. Since $U \cap V \neq \emptyset$ for every nonempty open set V, the set U is dense.

(iii) $\Rightarrow$ (ii). Let U and V be two nonempty open sets and set $\hat{U} = \bigcup_{n\geq 0} T^{-n}(U)$. Clearly $\hat{U}$ is open and invariant. Hence $\hat{U}$ is dense and $\hat{U} \cap V \neq \emptyset$. This implies $T^{-n}(U) \cap V \neq \emptyset$ for some $n \geq 0$. □

If $T : X \to X$ is a homeomorphism, then we can talk about the full orbit $O(x) = (T^n(x) : n \in \mathbb{Z})$. The proof of Theorem 2.8 implies this.

Corollary 2.9. *Let X be as in Theorem* 2.8 *and assume $T : X \to X$ is a homeomorphism. Then the following statements are equivalent:*

(i) *For some x, $O(x)$ is dense.*

(ii) *For every nonempty open sets U and V, there exists $N \in \mathbb{Z}$ such that $T^N(U) \cap V \neq \emptyset$.*

(iii) *If U is open and $T(U) = U$, then either $U = \emptyset$ or U is dense.*

Remark 2.2. Note that if T is invertible and instead of (i) we have

(i′) there exists $x \in X$ such that $O^-(x) = \{T^{-n}(x) : n \in \mathbb{N}\}$ is dense,

then as in Theorem 2.8 we can show that (i′) $\Rightarrow$ (ii). (We simply assume $n_2 \geq n_1$.) Since (ii) $\Rightarrow$ (i), we deduce that (i′) $\Rightarrow$ (i).

We now show that the denseness of the full orbit is equivalent to the topological transivity.

Lemma 2.10. *Let X be as in Theorem* 2.8. *Suppose $T : X \to X$ is a homeomorphism. If $(T^n(x) : n \in \mathbb{Z})$ is dense for some x, then T is topologically transitive.*

Proof. Suppose that $(T^n(x) : n \in \mathbb{Z})$ is dense for some x. Let $\omega(x)$ and $\alpha(x)$ be the set of limit points of $O^+(x)$ and $O^-(x)$ respectively. Note that both $\omega(x)$ and $\alpha(x)$ are closed sets. By assumption $\alpha(x) \cup \omega(x) = X$. Hence either $x \in \omega(x)$ or $x \in \alpha(x)$. In the former case, $O(x) \subseteq \omega(x)$ and this implies that $\omega(x) = X$. In the latter case $X = O(x) \subseteq \alpha(x)$. By the previous remark, we deduce that T is topologically transitive. □

We are now ready to prove Theorem 2.7.

Proof of Theorem 2.7. Suppose $\theta_1, \dots, \theta_n, 1$ are not rationally dependent. Since a translation of a dense set is dense, it suffices to show that the corresponding transformation T is topologically transitive. Using Corollary 2.9 and Lemma 2.10, it suffices to show that if U is open with $T(U) = U$, then either $U = \emptyset$ or U is dense. Set $f = 1\!\!1_U$. We may regard $f : \mathbb{R}^d \to \mathbb{R}$ as a periodic function (here we lifted f to a transformation on $\mathbb{R}^d$.) Since $T(U) = U$, we have $f \circ T = f$. Write the Fourier expansion of f:

$$f(x) = \sum_{n_1,\dots,n_d \in \mathbb{Z}^d} a(n_1, \dots, n_d) \exp(2\pi i(n_1 x_1 + \dots + n_d x_d)).$$

From $f = f \circ T$ we learn

$$\sum_n a(n) \exp(2\pi i x \cdot n) = \sum_n a(n) \exp(2\pi i x \cdot n) \exp(2\pi i \theta \cdot n).$$

By uniqueness of the Fourier coefficients,

$$a(n) = a(n) e^{2\pi i \theta \cdot n}$$

for every $n \in \mathbb{Z}^d$. Since $\theta \cdot n$ is never an integer for $n \neq 0$, $e^{2\pi i \theta \cdot n}$ is never 1 for $n \neq 0$. As a result $a(n) = 0$ whenever $n \neq 0$. Thus f is a constant almost everywhere. Since U is open, we deduce that either $U = \emptyset$ or U is dense.

Conversely, assume $\bar{n} \cdot \theta$ is an integer for some $\bar{n} \neq 0$. Define $f(x) = \sin(2\pi \bar{n} \cdot x)$. This induces a transformation on $\mathbb{T}^d$ because $\bar{n} \in \mathbb{Z}^d$. Moreover, $f \circ T = f$ because $\bar{n} \cdot \theta = 0$. Since $\bar{n} \neq 0$, f is not a constant function. As a result, the sets

$$U = \{x \in \mathbb{T}^d : f(x) > 0\}, \; V = \{x \in \mathbb{T}^d : f(x) < 0\}$$

are two nonempty invariant open sets. From Corollary 2.11 $(T^n(x) : n \in \mathbb{Z})$ is not dense for every $x \in \mathbb{T}^d$. □

Example 2.11. (Free motion on a torus). The ODE $\frac{dx}{dt} = v$, $v \in \mathbb{R}^d$, has a simple flow: $\phi_t(x) = x + tv$. This induces a flow on the torus $\mathbb{T}^d$ by $\hat{\phi}_t(x) = x + tv (\mathrm{mod}\, 1)$.

Exercise 2.12. Show that if $v = (v_1, \dots, v_d)$ and $v_1, \dots, v_d$ are not rationally dependent, then $(\hat{\phi}_t(x) : t \in \mathbb{R}^+)$ is dense for every $x \in \mathbb{T}^d$. Conversely, if $v_1, \dots, v_d$ are rationally dependent, then $(\hat{\phi}_t(x) : t \in \mathbb{R})$ is never dense when $d \geq 2$.

Exercise 2.13. Define $T : \mathbb{T}^2 \to \mathbb{T}^2$ by $T(x, y) = (x + \alpha, x + y)(\mathrm{mod}\, 1)$. Show that T is topologically transitive iff α is irrational.

Exercise 2.14. Show that the decimal expansion of 2^n may start with any finite combination of digits. (*Hint*: use $T : \mathbb{T}^1 \to \mathbb{T}^1$ defined by $T(x) = x + \alpha (\mathrm{mod}\, 1)$ with $\alpha = \log_{10} 2$.)

Example 2.15. A function $f : [0,1] \to \mathbb{R}$ with $f(0) = 0$, $f'(x) > 1$ for all $x \in [0,1]$, induces an expanding map $T : \mathbb{T}^1 \to \mathbb{T}^1$ by $T(x) = f(x)(\text{mod}\, 1)$. Such a function T is an example of a chaotic dynamical system and its orbit structure is significantly more complex than translations. As an example, take $f(x) = mx$ with $m > 1$ an integer. If we identify $\mathbb{T}^1$ with the set of complex numbers z such that $|z| = 1$, then T corresponds to the transformation $z \mapsto z^m$.

Theorem 2.16. *T is topologically transitive and periodic points of T are dense in $\mathbb{T}^1$.*

Proof. Since $T^n(z) = z^{m^n}$, the condition $T^n(z) = z$ means $z^{m^n-1} = 1$. As a result, there are exactly $m^n - 1$ points of period at most n. There are roots of unity and are uniformly spread over $\mathbb{T}^1$. Hence we have a dense set of periodic orbits.

If we represent $x = a_1 m^{-1} + a_2 m^{-2} + \dots + a_k m^{-k} + \dots$, $a_1, a_2, \dots \in \{0, 1, \dots, m-1\}$, then $T(x) = a_1 + a_2 m^{-1} + \dots (\text{mod}\, 1) = a_2 m^{-1} + \dots$. Now let $A = \left[\frac{i}{m^k}, \frac{i+1}{m^k}\right)$, $B = \left[\frac{j}{m^k}, \frac{j+1}{m^k}\right)$ for some $i, j \in \{0, 1, \dots, m^k - 1\}$. In base m representation

$$A = \{x : x = \cdot a_1 a_2 \dots a_k * * * \dots\},\ B = \{x : x = \cdot b_1 b_2 \dots b_k * * * \dots\}$$

for some $a_1 a_2 \dots a_k b_1 \dots b_k \in \{0, 1, \dots, m-1\}$. Since

$$T^{-n}(A) = \{x : x = \cdot \overbrace{* * \cdots *}^{n} a_1 a_2 \dots a_k * * \dots\}.$$

Now it is clear that if $n \geq k$ then

$$T^{-n}(A) \cap B = \{x : x = \cdot \underbrace{b_1 \dots b_k * * \cdots *}_{n} a_1 \dots a_k * * \dots\} \neq \emptyset.$$

From this and Theorem 4.8, we can readily deduce that T is topologically transitive. $\square$

Theorem 2.16 implies that some orbits are periodic and there exists a dense orbit. Do we have any other type of an orbit? We now claim that, for example, when $m = 3$, then there exists a point x for which $\omega(x)$ is the Cantor set. To see this, set

$$K = \{a_1 3^{-1} + a_2 3^{-1} + \dots : a_j = 0 \text{ or } 2 \text{ for all } j\},$$

and define $h : K \to [0,1]$ by $h(a_1 3^{-1} + a_2 3^{-1} + \dots) = \frac{a_1}{2} 2^{-1} + \frac{a_2}{2} 2^{-2} + \dots$.

In fact h is continuous and strictly increasing except for points of finite expansion. Let us write T^m for $z \mapsto z^m$. We can now see that in fact $h \circ T_3 = T_2 \circ h$ in K. Since T_2 is topologically transitive, there exists x with $\{T_2^n(x) : n \in \mathbb{N}\}$ dense. Each $T_2^n(x)$ cannot be a dyadic rational because dyadic rationals have finite orbits. Set $y = h^{-1}(x)$. Then $\{T_3^n y : n \in \mathbb{N}\}$ is dense in K.

In fact what we have shown for T_m is stronger than topological transitivity. Namely, T_m is *topologically mixing* in the following sense:

If U and V are two nonempty open sets, then there exists $N = N(U,V)$ such that $T^{-n}(U) \cap V \neq \emptyset$ for $n > N$.

3 Floquet Theory

In this section we study the orbits of a linear system with periodic coefficients. This is the subject of the Floquet Theory.

In the discrete case, we are interested in the behavior of the sequence $(x_n : n \in \mathbb{N})$ with $x_{n+1} = A_n x_n$ for a given collection of invertible matrices A_n satisfying

(3.1), $$A_{n+N} = A_n$$

for every n. In other words we have a non-autonomous system with periodic coefficients. If the period $N = 1$ then A_n is independent of n and we studied the corresponding problem in Section 2. In the continuous setting, we are interested in the flow of the dynamical system

(3.2) $$\frac{dx}{dt} = A(t)x$$

with A satisfying

(3.3) $$A(t+T) = A(t)$$

for all t.

Recall that if $x_n = f^n(a)$, then its variation satisfies $v_{n+1} = Df(x_n)v_n$. If x_n is a periodic orbit of period N then we have (3.1) for $A_n = Df(x_n)$. Similarly if we start with a nonlinear problem of the form $\frac{dx}{dt} = f(x)$ and look at its variation $\frac{dv}{dt} = Df(x(t))v$ then $A(t) = Df(x(t))$ satisfies (3.3) whenever $x(\cdot)$ is periodic of period T.

We start with the discrete problem

(3.4) $$v_{n+1} = A_n v_n, \;\; A_{n+T} = A_n.$$

Set

(3.5) $$R_n = A_n A_{n-1} \dots A_1.$$

We certainly have

$$v_n = R_n v_0,\ R_{n+N} = R_n R_N. \tag{3.6}$$

We start with a simple fact.

Proposition 3.1. *Assume that all the eigenvalues of R_N belong to the set $\{z : |z| < 1\}$. Then $\lim_{n\to\infty} v_n = 0$ exponentially fast. In fact $\lim_{n\to\infty} R_n = 0$ exponentially fast.*

Proof. If $n = Nk + r$ with $r \in \{0, 1, \dots, N-1\}$, then

$$R_n = A_r \dots A_2 A_1 R_N^k$$

if $r > 0$ and $R_n = R_N^k$ otherwise. Hence

$$\|R_n\| \le \|A_r \dots A_1\|\ \|R_N^k\|.$$

Set $c_0 = \max\{1, \|A_1\|, \dots, \|A_{N-1} \dots A_1\|\}$. If all the eigenvalues of R_N belong to $\{z : |z| < 1\}$, then $R_N^k \to 0$ as $k \to +\infty$ exponentially fast. This completes the proof. □

In this context, we would like to have a result similar to Theorem 2.1. It turns out that Exercise 2.6 has a generalization.

Theorem 3.2. *There exists a collection of numbers $l_1 < l_2 < \dots < l_k$, positive integers $n_1, n_2 \dots n_k$ and linear subspaces $G^1, \dots, G^k$, such that $n_1 + \dots + n_k = d$, and if $v \in (G^1 \oplus \dots \oplus G^j) - (G^1 \oplus \dots \oplus G^{j-1})$, then*

$$\lim_{n\to\infty} \frac{1}{n} \log |R_n v| = l_j.$$

Proof. If $n = kN + r$ with $r \in \{0, 1, \dots, N-1\}$, then $R_n = R_r R_N^k$. Hence

$$|R_n v| \le c_0 |R_N^k v|$$

where $c_0 = \max\{1, \|R_1\|, \dots, \|R_{N-1}\|\}$. Similarly

$$|R_N^{k+1} v| = |A_N \dots A_{r+1} R_n v| \le \|A_N \dots A_{r+1}\| |R_n v| \le c_1 |R_n v|,$$

where $c_1 = \max\{1, \|A_N\|, \|A_N A_{N-1}\|, \dots, \|A_N A_{N-1} \dots A_1\|\}$. As a result,

$$\lim_{n\to\infty} \frac{1}{n} \log |R_n v| = \frac{1}{N} \lim_{k\to\infty} \frac{1}{k} \log |R_N^k v|.$$

We now apply Exercise 2.6 to R_N. □

Remark 3.3. In fact $l_1 < \cdots < l_k$ are the numbers $\left\{\frac{1}{N}\log|\lambda_1|, \ldots, \frac{1}{N}\log|\lambda_n|\right\}$ written in increasing order, where $\lambda_1 \ldots \lambda_k$ are the eigenvalues of R_N.

More can be said about the orbit of (3.4). For this let us discuss a useful linear algebra lemma.

Lemma 3.4. *Let R be an invertible matrix. Then there exists a matrix C such that* $\exp C = R$.

Proof. In some sense $C =$ "$\log R$" and in fact we can define $f(R)$ for any analytic f. More precisely let us take an analytic function $f : \Omega \to \mathbb{C}$, where Ω is a domain in $\mathbb{C}$. Assume $\mathrm{Spect}(R) \subseteq \Omega$ where

$$\mathrm{Spect}(R) = \{z : zI - R \text{ is not invertible}\}.$$

We then use Cauchy's formula to define $f(R)$. For this let γ be any closed curve γ in Ω that winds once around $\mathrm{Spect}(R)$. Define

$$f(R) = \frac{1}{2\pi i}\int_\gamma (zI - R)^{-1} f(z)\,dz. \tag{3.7}$$

It is straightforward to find a simple expression for $f(R)$. Note that if $\hat{R} = P^{-1}RP$ then $f(\hat{R}) = P^{-1}f(R)P$ because $(zI - P^{-1}RP)^{-1} = P^{-1}(zI - R)^{-1}P$. Hence for (3.7) we may assume that R is in a Jordan Normal Form. As a result, it suffices to calculate $f(R)$ when R is of the form

$$\begin{bmatrix} \lambda & & & 0 \\ 1 & \ddots & & \\ & \ddots & & \\ 0 & & 1 & \lambda \end{bmatrix} =: B + \lambda I.$$

Here λ is an eigenvalue (possibly complex) and the matrix B satisfies $B^d = 0$ where R is a $d \times d$ matrix. In this case

$$\begin{aligned}(zI - R)^{-1} = ((z-\lambda)I - B)^{-1} &= \frac{1}{(z-\lambda)}\left(I - \frac{1}{z-\lambda}B\right)^{-1} \\ &= \frac{1}{z-\lambda}I + \frac{1}{(2-\lambda)^2}B + \cdots + \frac{1}{(z-\lambda)^d}B^{d-1}.\end{aligned}$$

Therefore

$$\begin{aligned} f(R) &= \frac{1}{2\pi i}\int_\gamma \left(\frac{f(z)}{z-\lambda}I + \cdots + \frac{f(z)}{(z-\lambda)^d}B^{d-1}\right)dz \\ &= f(\lambda)I + f'(\lambda)B + \cdots + \frac{1}{(d-1)!}f^{(d-1)}(\lambda)B^{d-1}. \end{aligned} \tag{3.8}$$

This formula yields a candidate for $f(R)$ for every analytic function f. If f is simply a polynomial, then we already know how to calculate $f(R)$ and we now would like to show that this calculation is consistent with (3.8). We only need to verify this when $f(z) = z^n$. The verification in this case is left to the reader. Also if f is given by $\sum_0^\infty a_n z^n$ over a region containing γ, then (3.8) is verified by approximating f by polynomials.

We are now ready to define $\log R$ for an invertible matrix R. Since R is invertible, $0 \notin \mathrm{Spect}(R)$. Pick a half-line L such that $L \cap \mathrm{Spect}(R) = \emptyset$. Set $\Omega = \mathbb{C} - L$. Take a branch of $\log z$ defined in Ω. Use this branch for f in (3.7) to define $C = \log R$. More precisely,

$$\log R = (\log \lambda)I + \frac{1}{\lambda}B - \frac{1}{\lambda^2}B^2 + \cdots + \frac{(-1)^{d-1}}{\lambda^{d-1}}B^{d-1}. \tag{3.9}$$

This is simply obtained by using the expansion of log and using the fact $B^d = 0$:

$$\log(\lambda I + B) = \log \lambda + \log\left(I + \frac{1}{\lambda}B\right) = \log \lambda + \sum_{j=1}^{\infty} \left(\frac{1}{\lambda}B\right)^j (-1)^{j-1}.$$

By direct calculation one can show that indeed $e^C = R$. □

Even when R is a real matrix in Lemma 3.4, there might not exist a real C such that $e^C = R$. For a real $\log R$ we need additional conditions.

Lemma 3.5. *Let R be a real invertible matrix. There exists a real matrix C with $e^C = R$ if R has no negative real eigenvalue. Moreover we can always find a real Z such that $e^Z = R^2$.*

Proof. We would like to use (3.7):

$$C = \frac{1}{2\pi i}\int_\gamma (zI - R)^{-1} \log z dz.$$

Since R has no negative eigenvalue, we may choose the standard branch of log. That is $\Omega = \mathbb{C} - \{x : x \le 0\}$ and $\log(\rho e^{i\theta}) = \log \rho + i\theta$ for $\theta \in (-\pi, \pi)$. Note that $\overline{\log z} = \log \rho - i\theta =$

$\log \bar{z}$. As a result

$$\bar{C} = \frac{-1}{2\pi i} \int_{\gamma} (\bar{z}I - R)^{-1} \log \bar{z} d\bar{z}.$$

Recall γ winds around Spect(R) once. Since R is real, $\overline{\text{Spect}(R)} = \text{Spect}(R)$. Hence the curve $\bar{\gamma}$ winds once clockwise around Spect(R). As a result

$$\bar{C} = \frac{1}{2\pi i} \int_{-\bar{\gamma}} (zI - R)^{-1} \log z dz = C.$$

For the existence of Z, without loss of generality we assume that $R = \begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_k \end{bmatrix}$ is its Jordan Normal Form. We find $Z = \begin{bmatrix} Z_1 & & 0 \\ & \ddots & \\ 0 & & Z_k \end{bmatrix}$ such that $e^{Z_j} = A_j^2$. If A_j corresponds to a pair of complex conjugate eigenvalues $\alpha \pm i\beta$ with $\beta \neq 0$, then we can find C_j such that $e^{C_j} = A_j$ with C_j real. We then set $Z_j = 2C_j$ in this case. If $A_j = \begin{bmatrix} \lambda & & & 0 \\ 1 & \ddots & & \\ & \ddots & & \\ 0 & & 1 & \lambda \end{bmatrix}$ corresponds to a real eigenvalue, then A_j^2 has no negative eigenvalue, so we can find Z_j such that $e^{Z_j} = A_j^2$. □

We are now ready for the Floquet representation.

Theorem 3.6. *Let $R_n = A_n \dots A_1$ with $A_{n+N} = A_n$. Assume that R_N is invertible. Then there exist (possibly complex) matrices P_n and Z such that $P_{n+N} = P_n$ and*

$$R_n = P_n Z^n. \tag{3.10}$$

Also there exist real matrices $\hat{P}_n$ and $\hat{Z}$ such that $\hat{P}_{n+2N} = \hat{P}_n$ and

$$R_n = \hat{P}_n \hat{Z}^n. \tag{3.11}$$

Proof. For (3.10) we simply choose $Z = \frac{1}{N} \log R_N$ and set

$$P_n = R_n Z^{-n}.$$

We then have

$$P_{n+N} = R_{n+N}Z^{-n-N} = R_n R_N Z^{-N} Z^{-n} = R_n Z^{-n} = P_n.$$

For (3.11), observe that the matrix R_N may not have real log but R_N^2 always has a real log. So choose $\hat{Z} = \frac{1}{2N}\log R_N^2 = \frac{1}{2N}\log R_{2N}$. We certainly have $\hat{Z}^{2N} = R_{2N}$. We then set $\hat{P}_n = R_n\hat{Z}^{-n}$ so that (3.11) holds and

$$\hat{P}_{n+2N} = R_{n+2N}\hat{Z}^{-2N}\hat{Z}^{-n} = R_n R_{2N}\hat{Z}^{-2N}\hat{Z}^{-n} = \hat{P}_n.$$

□

The statement (3.10) is often phrased as the existence of a periodic change of coordinates $x = P_n y$ that transforms the system $v_n = R_n v_0$ to the system $w_n = Z^n v_0$. Note that the latter is linear with constant coefficients.

We now turn to the continuous problem (3.2)–(3.3). First observe that if we solve the matrix equation

$$\begin{cases} \frac{d}{dt}R(t) = A(t)R(t) \\ R(0) = I \end{cases} \tag{3.12}$$

with $R(t)$ a $d \times d$ matrix for each t, then $v(t) = R(t)v_0$ solves

$$\frac{dv(t)}{dt} = A(t)v(t),\ v(0) = v_0.$$

We now argue that if (3.3) holds, then

$$R(t+T) = R(t)R(T). \tag{3.13}$$

This follows from the uniqueness; both $R_1(t) = R(t+T)$ and $R_2(t) = R(t)R(T)$ solve

$$\begin{cases} \frac{dX}{dt} = AX \\ X(0) = R(T). \end{cases}$$

Throughout we assume that $A(\cdot)$ is continuous so that $R(\cdot)$ is also continuous.

Proposition 3.7. *If all the eigenvalues of $R(T)$ belongs to $\{z : |z| < 1\}$, then $\lim_{t\to+\infty} R(t) = 0$ exponentially fast.*

The proof of Proposition 3.7 is very similar to the proof of Proposition 3.1 and is omitted.

Before stating and proving the analogue of Theorem 3.2, let us observe that $R(t)$ is always invertible. In fact we have a candidate for $B = R^{-1}$:

$$\begin{aligned} \frac{d}{dt}B(t) &= -R(t)^{-1}\frac{d}{dt}R(t)R(t)^{-1} \\ &= -R(t)^{-1}A(t)R(t)R(t)^{-1} \\ &= -B(t)A(t). \end{aligned}$$

Hence if B solves

$$\begin{cases} \frac{d}{dt}B(t) = -B(t)A(t) \\ B(0) = I, \end{cases}$$

then

$$\frac{d}{dt}BR = -BAR + BAR = 0.$$

Hence $BR = I$, i.e., R is invertible. Obviously $R(t) = \phi_0^t$ and $R(t)^{-1} = \phi_t^0$.

Theorem 3.8. *There exists a collection of numbers $l_1 < \cdots < l_k$, positive integers $n_1, \ldots, n_k$ and linear subspaces $H^1, \ldots, H^k$, such that $n_1 + \cdots + n_k = d$, and if $v \in H^1 \oplus \cdots \oplus H^j - H^1 \oplus \cdots \oplus H^{j-1}$, then*

$$\lim_{t\to+\infty} \frac{1}{t}\log|R(t)v| = l_j.$$

Proof. If $t = kT + r$ with $r \in [0, T)$, then

$$|R(t)v| = |R(r)R(T)^k v| \le \|R(r)\||R(T)^k v| \le c_0|R(T)^k v| \tag{3.14}$$

where $\delta = \max_{r\in[0,T]} \|R(r)\|$. On the other hand, since

$$\begin{aligned} R(t) &= R(r)R(T)^k, \\ R(T)^k &= R(r)^{-1}R(t). \end{aligned}$$

Hence

$$|R(T)^k r| \le c_1|R(t)r|$$

where $c_1 = \max_{r\in[0,T]} \|R(r)^{-1}\|$. From this and (3.14) we deduce that

$$\lim_{t\to\infty} \frac{1}{t}\log|R(t)r| = \frac{1}{T}\lim_{k\to\infty}\frac{1}{R}|R(T)^k r|.$$

We now apply Exercise 2.6 to the matrix $R(T)$. $\square$

Remark 3.9. If $\lambda_1, \dots, \lambda_n$ are the eigenvalues of $R(T)$, then the set $\{l_1, \dots, l_k\}$ coincides with the set $\left\{\frac{1}{T}\log|\lambda_1|, \dots, \frac{1}{T}\log|\lambda_n|\right\}$. □

Theorem 3.10. *There exists matrices $P(t)$ and C such that $R(t) = P(t)e^{tC}$ and $P(t+T) = P(t)$. Moreover there exist real matrices $\hat{P}(t)$ and $\hat{C}$ such that $\hat{P}(t+2T) = \hat{P}(t)$ and $R(t) = \hat{P}(t)e^{t\hat{C}}$.*

Proof. Since $R(T)$ is invertible, there exists a matrix C such that $R(T) = e^{TC}$. Set $P(t) = R(t)e^{-TC}$. We have

$$\begin{aligned} P(t+T) &= R(t+T)e^{-TC}e^{-tC} = R(t)R(T)e^{-TC}e^{-tC} \\ &= P(t). \end{aligned}$$

Since $R(2T) = R(T)^2$, we can find a real matrix $\hat{C}$ such that $R(2T) = \exp(2T\hat{C})$. Set $\hat{P}(t) = R(t)e^{-t\hat{C}}$. Then

$$\begin{aligned} \hat{P}(t+2T) &= R(t+2T)e^{-2T\hat{C}}e^{-t\hat{C}} \\ &= R(t)R(T)^2e^{-2T\hat{C}}e^{-t\hat{C}} \\ &= \hat{P}(t). \end{aligned}$$

□

The eigenvalues of $R(T)$ are the *Floquet multipliers*. In practice it is hard to calculate them. The following lemma is useful in some cases.

Lemma 3.11. *We have a solution $x(t) = p(t)\lambda^t$ with $p(t+T) = p(t)$ if and only if λ^T is an eigenvalue of $R(T)$.*

Proof. Suppose $x(t) = p(t)\lambda^t$ is a solution with p a T-periodic function. Recall $x(t) = P(t)e^{tC}x_0$ with $P(\cdot)$ periodic. We have

$$\begin{aligned} P(t+T)e^{(t+T)C}x_0 &= \lambda^T P(t)e^{tC}x_0, \\ P(t)e^{tC}(e^{TC} - \lambda^T I)x_0 &= 0. \end{aligned}$$

This implies that $e^{TC} - \lambda^T I$ is not invertible. (Recall that $R(t) = P(t)e^{tC}$ is invertible.) Hence λ^T is an eigenvalue of $R(T) = e^{TC}$.

Conversely if λ^T is an eigenvalue of e^{TC}, then we may choose μ such that μ is an eigenvalue of C and $Cx_0 = \mu x_0$, $\lambda^T = e^{T\mu}$. We then set $p(t) = P(t)x_0$ so that $p(t+T) = p(t)$ and if $x(t) = p(t)e^{t\mu}$ then

$$P(t)e^{tC}x_0 = P(t)e^{t\mu}x_0 = p(t)e^{t\mu} = x(t).$$

So $x(t)$ is a solution. □

Lemma 3.12. *If* $z(t) = \det R(t)$, *then* $z(t) = e^{\int_0^t \operatorname{tr}(A(s))ds}$.

Proof. It is well known that as $\delta \to 0$,

$$\det(I + \delta A) = 1 + \delta \operatorname{tr}(A) + O(\delta^2). \tag{3.15}$$

Because of this,

$$\frac{d}{dt}\det(R(t)) = \operatorname{tr} A(t) \det R(t),$$

or $\dot{z}(t) = \operatorname{tr}(A(t))z(t)$ with $z(0) = 1$. This implies the lemma. □

Exercise 3.11. Verify (3.15).

As a consequence of Lemma 3.10, if $\lambda_1, \dots, \lambda_n$ are the eigenvalue of $R(T)$ then

$$\lambda_1 \dots \lambda_n = \exp\left(\int_0^T \operatorname{tr}(A(s))ds\right). \tag{3.16}$$

Exercise 3.12. Consider (3.2) with

$$A(t) = \begin{bmatrix} -1 + \frac{3}{2}\cos^2 t & 1 - \frac{3}{2}\cos t \sin t \\ -1 - \frac{3}{2}\sin t \cos t & -1 + \frac{3}{2}\sin^2 t \end{bmatrix}.$$

Show that $v(t) = (-\cos t, \sin t)e^{t/2}$ is a solution to (3.2). Find the Floquet multipliers. (*Hint:* Use Lemma 3.9 and (3.16).)

Exercise 3.13. Consider the ODE $\dot{x} = f(x)$ with

$$f(x_1, x_2) = (x_1 - x_2 - x_1(x_1^2 + x_2^2), x_1 + x_2 - x_2(x_1^2 + x_2^2)).$$

Show that $\bar{x}(t) = (\sin t, -\cos t)$ is a solution. Find the Floquet multipliers for the variation equation

$$\frac{dv}{dt} = Df(\bar{x}(t))v.$$

4 Planar Dynamical Systems

In Section 2 we learned that if $f : \mathbb{R}^d \to \mathbb{R}^d$ is a Lipschitz cotinuous function, then the ODE

$$\frac{dx}{dt} = f(x) \tag{6.1}$$

produces a Lipschitz flow ϕ_t satisfying $\phi_{t+s} = \phi_t \circ \phi_s$.

As it turns out, in the very low dimensional cases, the orbit structure of (6.1) cannot be too complex.

Exercise 4.1. Let $f : \mathbb{R} \to \mathbb{R}$, $f(a) = f(b) = 0$, $a < b$ and $f > 0$ on (a, b). Show that for every $x \in (a, b)$,

$$\omega(x) = \{b\},\ \alpha(x) = \{a\}.$$

□

In this section we study the orbits of (4.1) when $d = 2$. A lot can be said in this case because of *Jordan Curve Theorem*, a closed simple curve in the plane divides the plane into two connected components.

Theorem 4.2 (Poincaré–Bendixon). *Assume $\omega(x) \neq \emptyset$ and is bounded with no fixed point. Then $\omega(x)$ is a closed orbit.*

We can generalize this further.

Theorem 4.3. *Suppose that $\{\phi_t(x) : t \geq 0\}$ is a subset of a closed bounded set K with K having only finitely many fixed points. Then either $\omega(x)$ is a fixed point, or $\omega(x)$ is a closed orbit, or $\omega(x)$ contains a finite number of fixed points and a set of orbits $\gamma_1, \ldots, \gamma_l$ with $\omega(\gamma_j)$ and $\alpha(\gamma_j)$ a fixed point.*

Set $O^+(x) = \{\phi_t(x) : t \geq 0\}$, $O^-(x) = \{\phi_t(x) : t \leq 0\}$.

Lemma 4.4. *Suppose $O^+(x) \cap \omega(x) \neq \emptyset$. Then either x is a fixed point or a closed orbit.*

To prepare for the proof of Lemma 4.4, let us make some observations. Suppose we have an ODE in $\mathbb{R}^d$ as in (4.1) and let $\gamma : U \to \mathbb{R}^d$ be a parametrization of a surface of codimension one in $\mathbb{R}^d$. We assume that $\gamma(0) = x^0$ and that $f(x^0) =$ is not tangent to $\Gamma = \gamma(U)$. Define

$$F : U \times \mathbb{R} \to \mathbb{R}^d,\ U \subseteq \mathbb{R}^{d-1}$$

by $F(y, t) = \phi_t(\gamma(y))$

We have

$$DF(y,t) = [D\phi_t(\gamma(y))D\gamma(y),\ f(\phi_t(\gamma(y)))],$$
$$DF(0,0) = [D\gamma(0), f(x^0)]$$

because $\phi_0(x) = x$ and $D\phi_0(x) =$ Identity. Since $f(x^0)$ is not tangent to Γ, we have that $DF(0,0)$ is invertible. Hence in a neighborhood of $(0,0)$, say $U_0 \times (-\delta, \delta)$ we have that F is an invertible differentiable map with a differentiable inverse.

We now consider a planar dynamical system with Γ a curve *transverse* to f, i.e., $f(x)$ is not tangent to Γ at every $x \in \Gamma$. Let γ be a parametrization of Γ with I an interval. We also assume that we have a flow box (chart) about Γ.

Lemma 4.5. *Assume that $\gamma : I \to \mathbb{R}^2$ with $\gamma(I) = \Gamma$ transverse to f as above. Let $\{y_n\}$ be a distinct collection of points in Γ with $\phi_{t_n}(y_0) = y_n$, and $0 < t_1 < t_2 < \dots$. Then $\{y_n\}$ is a monotone sequence on Γ, i.e., $y_n = \gamma(s_n)$ with $\{s_n\}$ a monotone sequence in I.*

Proof. It suffices to show that y_1 is between y_0 and y_2. Let β be a simple curve made up of $\{\phi_t(y_0) : 0 \le t \le t_1\}$ and the segment L of Γ between y_0 and y_1. By Jordan Curve Theorem, β divides $\mathbb{R}^2$ into a bounded region Ω_1 and unbounded region Ω_2.

Now either $\phi_t(y_1)$ enters Ω_1, i.e., $\phi_t(y_1) \in \Omega_1$ for $t > 0$ and small or $\phi_t(y_1)$ enters Ω_2. Let us assume the former. Set $E_1 = \{y \in L : \phi_t(y) \text{ enters } \Omega_1, \text{ for } t > 0\}$ and $E_2 = L - E_1$. Since we have a flow chart about Γ, it is not hard to see that both E_1 and E_2 are open in L. Since L is connected and by assumption $E_1 \ne \emptyset$, we deduce that $\phi_t(y)$ enters Ω_1 for every $y \in L$. From this we deduce that in fact $\{\phi_t(y_1) : t > 0\} \subseteq \Omega_1$ because $\phi_t(y_1)$ cannot exit Ω_1 through L, and cannot exit through $\beta - L$ by the uniqueness of our ODE.

The complement of L in Γ consists of two connected arcs Γ_0 and Γ_1 with y_0 an endpoint of Γ_0 and y_1 an endpoint of Γ_1. If we can show that in fact $\Gamma_0 \subseteq \Omega_2$, $\Gamma_1 \subseteq \Omega_1$, we are done because $y_2 = \phi_t(y_0) = \phi_{t-t_1}(y_1) \in \Gamma_1$ which means that y_1 is between y_0 and y_2. It is not hard to see $\Gamma_1 \subseteq \Omega_1$ because near Γ we have a flow box and in the box Γ_1 and $\{\phi_t(y_1) : t > 0, t \text{ small}\}$ belong to the same connected component.

□

Proof of Lemma 4.4. Assume that $O^+(x) \cap \omega(x) \neq \emptyset$. Let $a \in O^+(x) \cap \omega(x)$ and if a is a fixed point, then we are done. If a is not a fixed point, erect a transverse L through a. Since $a \in \omega(x) = \omega(a)$, there exist $t_n \to +\infty$ with $\phi_{t_n}(a) = a_n \in L$ and $a_n \to a$. If $a_n = a_m$ for some $n \neq m$, then a is a periodic point which implies that x is a periodic point and we are done. If a_n's are distinct, then by Lemma 6.5 $\{a_n\}$ is a monotone sequence. But this is impossible because $\lim_{n\to\infty} a_n = a_0$. □

Recall that an invariant set A is called *minimal* if it has no proper invariant subset. As a corollary to Lemma 4.4 we have this:

Corollary 4.6. *If A is minimal and compact, then A is either a fixed point or a periodic orbit.*

Proof. Let A be a minimal set and let $a \in A$. By invariance $O^+(a) \subseteq A$. By compactness $\omega(a) \subseteq A$. Since $\omega(a)$ is invariant, $\omega(a) = A$. Since $a \in \omega(a) \cap O^+(a)$, we use Lemma 4.4 to deduce that a is a fixed point or a periodic point. □

Note that if A is compact and invariant, then we can use Zorn's lemma to deduce that A has a nonempty compact subset that is minimal. (Here we are using the fact that the intersection of a finite number of invariant sets is again invariant.)

Let us state an exercise regarding the connectedness of $\omega(x)$:

Exercise 4.7.

(i) Show that $\omega(x) = \bigcap_{\tau>0} \overline{\{\phi_t(x) : t \geq \tau\}}$.

(ii) Use (i) to show that if $O^+(x)$ is bounded then $\omega(x)$ is connected.

(iii) Give an example of a disconnected ω-set in $\mathbb{R}^2$. □

Proof of Theorem 4.2. Assume that $\omega(x)$ is bounded with no fixed point. Let γ_0 be a minimal subset of $\omega(x)$. Then γ_0 must be a periodic orbit by Corollary 4.6. Let $a \in \gamma_0$

and erect a transverse L through a. Since $a \in \omega(x)$, we can find $t_n \to +\infty$ such that $\phi_{t_n}(x) = x_n \in L$ and $x_n \to a$. By Lemma 6.5, the sequence x_n is monotone. (Note that if $x_n = x_m$ for some $t_n \neq t_m$, then x is a periodic point and $a = x_n$ for all n. In this case we are done already.) If γ_1 is another minimal set and if γ_1 intersects L as well at a point b, then $x_n \to b$ and a monotone x_n cannot converge to two distinct points $a \neq b$. Hence there exists a neighborhood B_a of a so that $B_a \cap \omega(x) = B_a \cap \gamma_0$. This is true for every $a \in \gamma_0$. By compactness of γ_0, we can find a neighborhood B of γ such that $B \cap \omega(x) = B \cap \gamma_0$. Since $\omega(x)$ is connected, we must have $\omega(x) = \gamma_0$. □

Proof of Theorem 4.3. Assume that $\omega(x)$ is neither a fixed point nor a periodic orbit. In fact the proof of Theorem 4.2 reveals that if $\omega(x)$ contains a periodic orbit, then it must be equal to it. Hence $\omega(x)$ does not contain any periodic orbit. Since $\omega(x)$ is connected, it cannot consist of fixed points only. Let $y \in \omega(x)$ is not a fixed point. Evidently $\omega(y) \subseteq \omega(x)$, $\alpha(y) \subseteq \omega(x)$. To complete the proof, it suffices to show that $\omega(y)$ consists of a single fixed point and the same is true for $\alpha(y)$. We only verify the former. Indeed if $z \in \omega(y)$ and z is not a fixed point, then we can erect a transverse L at z. We can repeat the proof of Theorem 4.2 to deduce that $\omega(x) \cap L = \omega(y) \cap L = \{z\}$ because a transverse can only have one point of $\omega(x)$. Also $O^+(y)$ must intersect L at some point, say at y_0. But $O^+(y) \subseteq \omega(x)$. So $y_0 = z$. As a result, $O^+(y) \cap \omega(y) \ni z$ and by Lemma 4.4, y must be a periodic point. This contradicts our assumption that $\omega(x)$ is not a periodic orbit. Hence $\omega(y)$ is a fixed point, completing the proof. □

Example 4.8. Consider the system

$$\begin{cases} \dot{x}_1 = -x_2 + x_1(1 - r^2), \\ \dot{x}_2 = x_1 + x_2(1 - r^2). \end{cases}$$

In polar coordinates (r, θ), we have

$$\dot{\theta} = 1, \ \dot{r} = r(1 - r^2).$$

Now if $a \neq 0$ does not lie on the circle $r = 1$ then $\omega(a)$ is a single periodic orbit, namely the circle $r = 1$.

Exercise 4.9. Consider a system that in polar coordinates is given by

$$\begin{cases} \dot{r} = r(1 - r) \\ \dot{\theta} = \sin^2\theta + |1 - r|^\alpha. \end{cases}$$

Assume that $0 < |a| < 1$. Find $\omega(a)$. □

Consider the linear equation

$$\frac{dx}{dt} = Ax \tag{4.2}$$

with A a 2×2 matrix. If $A = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}$, then

$$x(t) = e^{\alpha t} \begin{bmatrix} \cos \beta t & -\sin \beta t \\ \sin \beta t & \cos \beta t \end{bmatrix} x(0).$$

If we write $x(t) = \rho(t) \begin{bmatrix} \cos \theta(t) \\ \sin \theta(t) \end{bmatrix}$, then

$$\lim_{t \to \infty} \frac{1}{t} \log \rho(t) = \alpha, \ \lim_{t \to \infty} \frac{\theta(t)}{t} = \beta$$

where θ is the lifted angle. Note that $\alpha \pm i\beta$ are the eigenvalues of A and α measures the exponential rate of increase in ρ and β measures the linear growth rate of θ.

Exercise 4.10. Show that if $A = \begin{bmatrix} \lambda & 0 \\ 1 & \lambda \end{bmatrix}$ in (4.1) and $x(t) = \rho(t) \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}$, then $\frac{1}{t} \log \rho(t) \to \lambda$ and $\frac{1}{t}\theta(t) \to 0$ as $t \to \infty$. □

What we learn is that in (4.2) we always have

$$\lim_{t \to \infty} \frac{1}{t} \theta(t) = \operatorname{Im} \lambda$$

where λ is the eigenvalue of A.

We now turn to

$$\frac{dx}{dt} = A(t)x \tag{4.3}$$

with A a 2×2 and T-periodic matrix-valued continuous function. Again we write

$$x = \rho \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}$$

with $t \mapsto \theta(t)$ lifted, i.e., $\theta(t) \in \mathbb{R}$ and $t \mapsto \theta(t)$ continuous. We can readily come up with equations for ρ and θ; if $u = \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}$, $u^\perp = \begin{bmatrix} -\sin \theta \\ \cos \theta \end{bmatrix}$, then (4.3) means

$$\dot{\rho} u + \rho \dot{u} = \rho A(t) u, \tag{4.4}$$

multiplying both sides by $u^\perp$ yields

$$u^\perp \cdot \dot{u} = A(t)u \cdot u^\perp.$$

Since $\dot{u} = \dot{\theta} u^{\perp}$, we obtain

$$\dot{\theta} = A(t)u \cdot u^{\perp}. \tag{4.5}$$

Note that this equation is a first order nonlinear equation in θ and does not depend on ρ. Multiplying both sides of (4.4) by u yields

$$\dot{\rho} = \rho(A(t)u \cdot u) \tag{4.6}$$

or

$$\rho(t) = e^{\int_0^t A(t')u(\theta(t'))\cdot u(\theta(t'))dt'} \rho(0). \tag{4.7}$$

We are interested in

$$\lim_{t\to\infty} \frac{1}{t} \log \rho(t) = \lim_{t\to\infty} \frac{1}{t} \int_0^t A(s)u(s) \cdot u(s) ds =: \bar{\rho} \tag{4.8}$$

$$\lim_{t\to\infty} \frac{1}{t} \theta(t) =: \bar{\theta} \tag{4.9}$$

where for simplicity we wrote $u(s)$ for $u(\theta(s)) = \begin{bmatrix} \cos\theta(s) \\ \sin\theta(s) \end{bmatrix}$. In fact we already know what $\bar{\rho}$ is. Recall that by Floquet Theory,

$$x(t) = P(t)e^{tC}x(0) \tag{4.10}$$

with $P(t)$ periodic in t, and $e^{2TC} = R(2T)$ where $R(t)$ is the fundamental solution. If λ_1 and λ_2 are the Floquet multipliers, then $\bar{\rho}$ exists and belongs to the set $\left\{\frac{1}{T}\log|\lambda_1|,\ \frac{1}{T}\log|\lambda_2|\right\}$. Let μ_1 and μ_2 denote the eigenvalues of the real matrix C. Then $\frac{1}{T}\log|\lambda_j| = \operatorname{Re}\mu_j$. When $A(t) \equiv A$ is independent of t, then μ_1, μ_2 are simply the eigenvalues of A.

Theorem 4.11. *If A is T-periodic, then the* rotation number $\bar{\theta}$ *exists and equals* $\operatorname{Im}\mu_j$ *or* $-\operatorname{Im}\mu_j$.

Proof. We first verify the Theorem when $A(t) \equiv A$ is independent of t. Then $C = A$. If C is in Jordan Normal Form

$$C = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} \text{ or } C = \begin{bmatrix} \lambda & 0 \\ 1 & \lambda \end{bmatrix},$$

then the result follows from Exercise 4.10 and the preceding discussion. In this case $\bar{\theta} = \operatorname{Im}\mu$ where $\mu = \alpha + i\beta$ or λ. If C is not in Jordan Normal Form, we can find a matrix Q such that $C = Q\hat{C}Q^{-1}$ with $\hat{C}$ in Jordan Form. So,

$$e^{tC} = Qe^{t\hat{C}}Q^{-1}.$$

We already know that if $e^{t\hat{C}}v = \rho(t)\begin{bmatrix}\cos\theta(t)\\ \sin\theta(t)\end{bmatrix}$, then $\frac{1}{t}\theta(t) \to \bar{\theta}$ as $t \to +\infty$. Hence we only need to make sure that if

$$Q\begin{bmatrix}\cos\theta(t)\\ \sin\theta(t)\end{bmatrix} = \hat{\rho}(t)\begin{bmatrix}\cos\hat{\theta}(t)\\ \sin\hat{\theta}(t)\end{bmatrix},$$

then we still have $\frac{1}{t}\hat{\theta}(t) \to \pm\bar{\theta}$. For this define $f : \mathbb{T}^1 \to \mathbb{T}^1$ by $f(z) = \frac{Qz}{|Qz|}$ where z is a vector of length 1. Clearly f is a continuous function. Also f is a homeomorphism because Q is invertible. Indeed $f^{-1}(z) = \frac{Q^{-1}z}{|Q^{-1}z|}$. As a result, f has a continuous lift $F : \mathbb{R} \to \mathbb{R}$. Then

$$\hat{\theta}(t) = F(\theta(t)).$$

Here we regard $\mathbb{T}^1$ as the interval $[0, 2\pi]$ with $0 = 2\pi$ an F enjoys the property $F(\theta + 2\pi) = F(\theta) \pm 2\pi$. (Note that we either have $\deg F = 1$ or $\deg F = -1$ by Lemma 5.3.) It is not hard to see that

$$\lim_{\theta\to\pm\infty} \frac{F(\theta)}{\theta} = \pm\deg F = \pm 1.$$

Since we know that $\frac{\theta(t)}{t} \to \bar{\theta}$ as $t \to +\infty$, we learn that $\frac{\hat{\theta}(t)}{t} \to \pm\bar{\theta}$ as $t \to +\infty$, we learn that $\frac{\hat{\theta}(t)}{t} \to \pm\bar{\theta}$. Of course we get $\bar{\theta}$ if the matrix Q does not reverse the orientation.

We now turn to the general periodic case. We know (4.10) and by the previous argument, if $y(t) = e^{tC}x(0)$ and $y(t) = \rho(t)\begin{bmatrix}\cos\theta(t)\\ \sin\theta(t)\end{bmatrix}$, then $\lim_{t\to+\infty}\frac{\theta(t)}{t} = \bar{\theta}$ exists. We only need to make sure that the matrices $P(t)$ do not change the angles by much, i.e., if

$$P(t)\begin{bmatrix}\cos\theta(t)\\ \sin\theta(t)\end{bmatrix} = r(t)\begin{bmatrix}\cos\hat{\theta}(t)\\ \sin\hat{\theta}(t)\end{bmatrix},$$

then we still have $\lim_{t\to\infty}\frac{\hat{\theta}(t)}{t} = \pm\bar{\theta}$. Indeed if $f(t,\cdot) : \mathbb{T} \to \mathbb{T}$ is defined by $f(t,x) = \frac{P(t)x}{|P(t)x|}$, then $f(t,\cdot)$ has a lift $F(t,\cdot) : \mathbb{R} \to \mathbb{R}$ and

$$\hat{\theta}(t) = F(t, \theta(t)).$$

It is not hard to show that F can be chosen to be continuous in t, and T-periodic. By continuity, $\deg F(t,\cdot) \equiv 1$ for all t, or $\deg F(t,\cdot) \equiv -1$ for all t. In the former case,

$$\sup_{t,\theta} |F(t,\theta) - \theta| < \infty$$

because $F(t,\theta) - \theta$ is T-periodic in t and 2π-periodic in θ. This immediately implies that

$$\lim_{t\to\infty}\frac{\hat{\theta}(t)}{t} = \lim_{t\to\infty}\frac{\theta(t)}{t}.$$

Similarly if $\deg F \equiv -1$, then

$$\lim_{t\to\infty} \frac{\hat{\theta}(t)}{t} = -\lim_{t\to\infty} \frac{\theta(t)}{t}.$$

□