

FIRST EXAMPLES OF DISCRETE DYNAMICAL SYSTEMS

We shall follow Zehnder's 'Lectures on dynamical systems' at first but include additional sources as the course progresses. There will be lecture notes accompanying the course.

1. INTRODUCTION

The origins of the modern theory of dynamical systems can be traced back in part to the studies of solutions of systems of ordinary differential equations at the end of the nineteenth century. Assume that

$$F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

is a map in $C^1(\mathbb{R}^n; \mathbb{R}^n)$. Given $x_0 \in \mathbb{R}^n$, we seek to describe the solution of the system of first order differential equations

$$(1.1) \quad \dot{x}(t) = F(x(t)), \quad x(0) = x_0,$$

where the function

$$x(\cdot) : I \longrightarrow \mathbb{R}^n$$

is defined at least on some open interval $I \ni 0$. That the system (1.1) has at least a local solution $x \in C^1(I; \mathbb{R}^n)$, which is unique on its interval of existence is well-known from the basic existence theory of ODE.

In this course, the interest is directed toward the *long-time qualitative behavior* of the solutions. More specifically, we are interested in developing abstract tools, which may be useful to make progress towards this issue.

Assume that we know that solutions of (1.1) are known to exist a priori for all $t \in \mathbb{R}$ (i. e. they are *global solutions*). We can then set

$$x(1) = T(x_0),$$

where again by basic local existence theory of ordinary differential equations the map $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is continuous and in fact a homeomorphism. To understand the long time behavior of $x(t)$, it then makes sense to understand the *iterates*

$$x(n) = (T \circ T \circ \dots \circ T)(x_0) =: T^n(x_0).$$

In the first part of the course, we will do this in a much *more abstract setting*. Thus we shall consider maps $T : X \longrightarrow X$, where X will be a set, which can be equipped with additional structures, namely either topological, or measure theoretical, or both, and only later do we consider more refined structures, such as differentiability, to infer deeper conclusions.

Our first task will be to understand some very basic *generic phenomena* by means of carefully chosen examples.

2. ABSTRACT DISCRETE DYNAMICAL SYSTEMS

Let X be a (non-empty) set, and $T : X \longrightarrow X$ a map. Given $x \in X$, we can define the images

$$x_0 := T^0(x) := x, \quad x_j = T(x_{j-1}) = T^j(x), \quad j \geq 1.$$

Then we make the

Definition 2.1. *We call the sequence $\{T^j(x)\}_{j \geq 0}$ the parametrized orbit, and the set*

$$\mathcal{O}^+(x) := \cup_{j \geq 0} T^j(x)$$

the un-parametrized orbit of x .

If T is bijective, so that $T^{-1}(y)$ is defined for all $y \in X$, we call $\{T^{-j}(x)\}_{j \geq 0}$ the parametrized negative orbit, and

$$\mathcal{O}^-(x) := \cup_{j \geq 0} T^{-j}(x)$$

the un-parametrized negative orbit. We then also call $\{T^j(x)\}_{j \in \mathbb{Z}}$ the full parametrized orbit, and analogously for $\mathcal{O}(x)$.

A very particular situation is associated with

Definition 2.2. A point x is called periodic of period $N \geq 1$ for T , provided we have

$$T^N(x) = x, T^j(x) \neq x, j \in \{1, 2, \dots, N-1\}.$$

If x is a periodic point of period N for T , then clearly we have

$$\mathcal{O}^+(x) = \{x, T(x), \dots, T^{N-1}(x)\}.$$

A very basic and not particularly interesting (from the point of view of dynamical systems) example is the following:

Example 1: Let (X, d) be a complete metric space, and let $T : X \rightarrow X$ a contraction. Thus $\exists a \in [0, 1)$ such that

$$d(T(x), T(y)) \leq a \cdot d(x, y) \forall x, y \in X.$$

Then by the Banach fixed point theorem $\exists! x_* \in X$ such that $T(x_*) = x_*$; in fact, $x_* = \lim_{r \rightarrow +\infty} T^r(x)$ for all $x \in X$. This is the unique periodic point (of period $N = 1$) for T since if $x = T^k(x)$ for some $k \geq 1$, then

$$x_* = \lim_{l \rightarrow +\infty} T^{kl}(x) = x.$$

A more interesting example in our present context, and which will accompany us through a portion of this course, concerns *rigid rotations of the circle*.

Before we introduce this, we recall some basic facts from abstract topology, which will be useful in the sequel. First

Definition 2.3. Let X, Y be topological spaces, and let $\pi : X \rightarrow Y$ a continuous map. Then (X, π) is called a covering of Y , provided for every $y \in Y$, there exists an open neighborhood $U_y \subset Y$ of y , such that

$$\pi^{-1}(U_y) = \cup_{a \in \mathcal{A}} V_a$$

where the $\{V_a\}_{a \in \mathcal{A}}$ form a disjoint collection of open sets in X , and $\pi|_{V_a} : V_a \rightarrow U_y$ is a homeomorphism.

Remark 2.4. For us Y will always be connected, in which case it can be shown that π is automatically surjective and the cardinality of $\mathcal{A}$ does not depend on $y \in Y$.

Example: Consider the map $\pi : \mathbb{R} \rightarrow \mathbb{T} := S^1 \hookrightarrow \mathbb{C}$ given by $x \mapsto e^{2\pi i x}$. This is a covering map. More generally, the map $\pi : \mathbb{R}^n \rightarrow \mathbb{T}^n$ given by $(x_1, \dots, x_n) \mapsto (e^{2\pi i x_1}, \dots, e^{2\pi i x_n})$ is a covering map.

In the sequel, we shall sometimes use the *lifting property* for such coverings. An abstract version is given by the following

Theorem 2.5. Assume that X, Y, W are pathwise connected and locally pathwise connected, Further assume W is simply connected. Then if

$$\pi : X \rightarrow Y$$

is a covering map, $w_0 \in W, y_0 \in Y, x_0 \in X$ are given with $\pi(x_0) = y_0$, and

$$f : W \rightarrow Y, f(w_0) = y_0,$$

is a given continuous map, there is a unique 'lifting'

$$g : W \rightarrow X$$

such that

$$\pi \circ g = f, g(w_0) = x_0.$$

Example: Let $\pi : \mathbb{R} \rightarrow S^1$ be given by $\pi(x) = e^{2\pi i x}$. Then if $f : S^1 \rightarrow S^1$ is a continuous map, and $x_0, y_0 \in \mathbb{R}$ with $f(e^{2\pi i x_0}) = e^{2\pi i y_0}$, there is a unique continuous map

$$g : \mathbb{R} \rightarrow \mathbb{R}$$

with the properties that

$$\pi \circ g = f \circ \pi, \quad g(x_0) = y_0.$$

We now turn to the second abstract discrete dynamical system:

Example 2: Let $S^1 \hookrightarrow \mathbb{C}$ the unit circle, and let $\alpha \in \mathbb{R}$. Then we define the *rotation map by angle α* as follows:

$$\phi(z) := e^{2\pi i \alpha} \cdot z, \quad z \in S^1.$$

Of course if $z = e^{2\pi i x}$, then $\phi(z) = e^{2\pi i(x+\alpha)}$. It follows that the map

$$\Phi : \mathbb{R} \rightarrow \mathbb{R}, \quad \Phi(x) = x + \alpha$$

is a lifting of the map ϕ , in the sense that $\phi \circ \pi = \pi \circ \Phi$.

The orbits in this second example are of fundamentally different character than in the first example. In fact, depending on the character of α , one distinguishes between two types of scenarios:

Proposition 2.6. (i) Assume that $\alpha = \frac{p}{q} \in \mathbb{Q}$, with $p \in \mathbb{Z}, q \in \mathbb{N}$ relatively prime. Then every orbit of ϕ is periodic with period q .

(ii) If $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ is irrational, then every orbit of ϕ is dense:

$$\overline{\mathcal{O}^+(z)} = S^1 \quad \forall z \in S^1.$$

Proof. (i) Note that $(e^{2\pi i \alpha})^l = e^{2\pi i l \cdot \frac{p}{q}} \neq 1$ for $l = 1, 2, \dots, q-1$, while $(e^{2\pi i \alpha})^q = 1$. This proves part (i).

(ii) Let $x \in \mathbb{R}$, and consider the 'lifted' orbit $\{\Phi^n(x)\}_{n \geq 0} = \{x + n\alpha\}_{n \geq 0}$. We consider the 'reduction' of these mod 1, i. e. we consider their projection in $\mathbb{R} \setminus \mathbb{Z} \simeq [0, 1)$. Given $\varepsilon = \frac{1}{N} > 0$, $N \gg 1$, divide $[0, 1)$ into equal intervals of length ε . Then amongst the points $\{\Phi^n(x)\}_{N \geq n \geq 0}$, at least two, say $\Phi^j(x)$, $\Phi^k(x)$, $N \geq k > j \geq 0$, fall into the same interval I . Hence

$$x + j\alpha = x + k\alpha + c\varepsilon \pmod{1}$$

for some c with $|c| \in (0, 1]$, whence $(k-j)\alpha = c\varepsilon \pmod{1}$. This means that the point

$$\phi^{k-j}(e^{2\pi i x}) = e^{2\pi i[(k-j)\alpha + x]}$$

on the circle S^1 is at a distance of at most $2\pi\varepsilon$ from the $\phi^0(x)$. Furthermore, the points

$$\phi^{l(k-j)}(e^{2\pi i x}), \quad l = 1, 2, \dots,$$

advance by steps of length $2\pi c\varepsilon$ for a fixed $c \in (0, 1]$, either clockwise or counter clockwise around the circle. Thus given $e^{2\pi i x_*} \in S^1$, there is a $l_* \in \mathbb{N}$ such that

$$|\phi^{l_*(k-j)}(e^{2\pi i x}) - e^{2\pi i x_*}| \leq 2\pi\varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, the density of the orbit follows. $\square$

3. MORE DETAILED CONSIDERATIONS ON THE PRECEDING EXAMPLES; STRUCTURAL STABILITY

Aside from the fact that the structure of the orbits of the preceding examples 1 and 2 is quite distinct, they are also different in terms of their *structural stability*. By this we loosely mean that if we add a small perturbation to the map T , replacing it by $\tilde{T}$, then the orbit structure of $\tilde{T}$ is the same.

To see that this is the case for example 1, we can for example state the following immediate

Lemma 3.1. Let (X, d) be a complete metric space, and $T : X \rightarrow X$ a contraction. Then if $\tilde{T} : X \rightarrow X$ is close to T in the precise sense that

$$|d(T(x), T(y)) - d(\tilde{T}(x), \tilde{T}(y))| \leq c \cdot d(x, y)$$

for some c sufficiently small (depending on T), then $\tilde{T}$ is also a contraction.

Proof. It suffices to pick $c < 1 - \lambda$, where $\lambda \in [0, 1)$ is the contraction parameter for T . $\square$

By contrast, circle rotations are *structurally unstable*, by the simple fact that any rational number is arbitrarily well approximated by irrational numbers, and vice versa.

4. MORE DETAILED CONSIDERATIONS ON THE PRECEDING EXAMPLES; EQUIDISTRIBUTION FOR CIRCLE ROTATIONS

Let $\phi = \phi_\alpha$ be the circle rotation by angle $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. We know that all its orbits are dense in S^1 , and it is natural to ask more refined questions. For example, are there regions on S^1 where 'more points of an orbit accumulate asymptotically' than other regions, depending on α ?

A beautiful classical theorem of Weyl affirms that this is *not the case*, and that the points in each orbit are *equidistributed asymptotically*, in a precise quantitative way. This result is important in particular since it also foreshadows the deep *Birkhoff ergodic theorem*, covering vastly more general dynamical systems.

We shall rely on some basic facts from the theory of Fourier series on S^1 . In particular, we rely on *Fejer's theorem*, whose proof is contained in a supplementary document. We recall

Definition 4.1. Let $f \in L^1(S^1)$. Then we define the Fourier coefficients $\hat{f}(n)$, $n \in \mathbb{Z}$, by the formula

$$\hat{f}(n) = \int_0^1 e^{-2\pi i n x} f(x) dx,$$

where we identify f with its periodic lifting on $\mathbb{R}$. We denote the formal Fourier series associated to f by

$$Sf(x) := \sum_{n \in \mathbb{Z}} e^{2\pi i n x} \cdot \hat{f}(n),$$

and the truncated Fourier series by

$$S_N f(x) := \sum_{|n| \leq N} e^{2\pi i n x} \cdot \hat{f}(n), \quad N \geq 0.$$

As far as convergence of Fourier series to a function are concerned, we have the following *averaged version*:

Theorem 4.2. (Fejer) Assume that $f \in C^0(S^1)$. Then we have

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=0}^{N-1} S_j f(x) = f(x)$$

uniformly in $x \in S^1$.

Remark 4.3. In general it is not true that

$$\lim_{N \rightarrow \infty} S_N f(x) = f(x)$$

for every $x \in S^1$ if we only require that $f \in C^0(S^1)$.

With these preparations, we now state the following

Theorem 4.4. (Weyl) Let f be Riemann integrable on S^1 , and assume $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. Then the relation

$$\lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{j=0}^{N-1} f(\phi^j(z)) = \frac{1}{2\pi} \int_{S^1} f(w) d\sigma \quad \forall z \in S^1$$

holds.

Proof. It suffices to work with the 'lift' of f to $\mathbb{R}$, meaning we think of f as a periodic function of $f : \mathbb{R} \rightarrow \mathbb{R}$ which is periodic, of period 1. Moreover, the map ϕ is then given by $\phi(x) = x + \alpha$, and the integral average becomes

$$\int_0^1 f(x) dx.$$

The proof now proceeds in two steps:

Step 1: assume that f is a trigonometric polynomial. More specifically assume $f(x) = e^{2\pi i n x}$, $n \in \mathbb{Z}$.

(1): $n = 0$. Here we obviously have

$$\lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{j=0}^{N-1} f(\phi^j(z)) = 1 = \int_0^1 f(x) dx.$$

(2): $n \in \mathbb{Z} \setminus \{0\}$. Here we evaluate the geometric series (for $z \in \mathbb{R}$)

$$\sum_{j=0}^{N-1} f(\phi^j(z)) = \sum_{j=0}^{N-1} e^{2\pi i n(j\alpha + z)} = e^{2\pi i n z} \cdot \frac{e^{2\pi i N n \alpha} - 1}{e^{2\pi i n \alpha} - 1}.$$

Since by assumption $e^{2\pi i n \alpha} \neq 1$, and $|e^{2\pi i N n \alpha} - 1| \leq 2$, we infer that

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=0}^{N-1} f(\phi^j(z)) = 0 = \int_0^1 f(x) dx.$$

By linear dependence of both the average as well as the integral on the function, it follows that the desired identity also holds for trigonometric polynomials.

Step 2: assume that f is a general Riemann integrable function. In particular, we assume that f is bounded. Given $\varepsilon > 0$, we can find a partition $\mathcal{P} = \{I_j\}_{j=1}^N$ of $[0, 1]$ such that we have

$$0 \leq \mathcal{U}(f, \mathcal{P}) - \mathcal{L}(f, \mathcal{P}) < \varepsilon,$$

where

$$\mathcal{U}(f, \mathcal{P}) = \sum_{j=1}^N |I_j| \cdot \sup_{t \in I_j} f(t), \quad \mathcal{L}(f, \mathcal{P}) = \sum_{j=1}^N |I_j| \cdot \inf_{t \in I_j} f(t)$$

Now we approximate the step function $f^+ := \delta + \sum_{j=1}^N \chi_{I_j} \cdot \sup_{t \in I_j} f(t)$, where $\delta > 0$ is a small parameter, from above by a continuous function g^+ and the step function $f^- := -\delta + \sum_{j=1}^N \chi_{I_j} \cdot \inf_{t \in I_j} f(t)$ from below by the continuous function g^- such that we have

$$0 < \int_0^1 g^+(x) dx - \int_0^1 g^-(x) dx < 2\varepsilon,$$

and furthermore by construction we have

$$g^+ - \delta \geq f \geq g^- + \delta.$$

This can be arranged by choosing $\delta > 0$ small enough. From Fejer's theorem we know that there exist trigonometric polynomials $T_{1,2}(x)$ such that

$$|g^+(x) - T_1(x)| < \delta, \quad |g^-(x) - T_2(x)| < \delta \quad \forall x \in [0, 1],$$

and furthermore

$$0 \leq \int_0^1 T_1(x) dx - \int_0^1 T_2(x) dx < 3\varepsilon.$$

Moreover, since $T_2(x) \leq f(x) \leq T_1(x)$ for all $x \in [0, 1]$, we have

$$\int_0^1 T_2(x) dx \leq \int_0^1 f(x) dx \leq \int_0^1 T_1(x) dx.$$

By step 1, we can pick $M \gg 1$ such that

$$\left| \frac{1}{M} \sum_{j=0}^{M-1} T_k(\phi^j(x)) - \int_0^1 T_k(x) dx \right| < \varepsilon, \quad k = 1, 2.$$

Then we finally infer

$$\begin{aligned}
-\varepsilon &\leq \frac{1}{M} \sum_{j=0}^{M-1} T_2(\phi^j(x)) - \int_0^1 T_2(x) dx \\
&\leq \frac{1}{M} \sum_{j=0}^{M-1} f(\phi^j(x)) - \int_0^1 f(x) dx + 3\varepsilon \\
&\leq \frac{1}{M} \sum_{j=0}^{M-1} T_1(\phi^j(x)) - \int_0^1 T_1(x) dx + 6\varepsilon \\
&\leq 7\varepsilon.
\end{aligned}$$

Thus we get

$$\left| \frac{1}{M} \sum_{j=0}^{M-1} f(\phi^j(x)) - \int_0^1 f(x) dx \right| \leq 4\varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, the claim follows. $\square$

Corollary 4.5. *Let $I \subset S^1$ an open interval, and $x \in S^1$. Then we have*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \cdot \left| \{n \in \{1, 2, \dots, N-1\} \mid x + n\alpha - \lfloor x + n\alpha \rfloor \in I\} \right| = |I|.$$

Proof. Apply the preceding corollary to the characteristic function $f = \chi_I$. $\square$

Remark 4.6. The preceding theorem is false in general if f is only Lebesgue integrable. However, a far reaching generalisation, namely the Birkhoff ergodic theorem, shows that the conclusion still holds for most orbits in a suitable sense.

The essentially identical argument yields the following higher dimensional generalisation. Denote by $\mathbb{T}^n = S^1 \times \dots \times S^1$ the n -dimensional torus. The natural projection

$$\pi : \mathbb{R}^n \longrightarrow \mathbb{T}^n$$

is a covering map. For $(\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$, we let

$$\phi : \mathbb{T}^n \longrightarrow \mathbb{T}^n$$

the 'rotation map'

$$\mathbf{z} := (z_1, \dots, z_n) \in \mathbb{T}^n \longrightarrow \phi(\mathbf{z}) := (e^{2\pi i \alpha_1} z_1, \dots, e^{2\pi i \alpha_n} z_n).$$

Then the following holds:

Theorem 4.7. *Let $\alpha := (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ such that*

$$\sum_{k=1}^n \alpha_k \cdot j_k \notin \mathbb{Z} \forall \mathbf{j} = (j_1, \dots, j_n) \in \mathbb{Z}^n \setminus \{\mathbf{0}\}.$$

Then if $f : \mathbb{T}^n \rightarrow \mathbb{R}$ is Riemann integrable (in the sense that its representation on $[0, 1]^n$ is Riemann integrable), then the limiting relation (where $\mathbf{z} \in \mathbb{T}^n$ is arbitrary)

$$\lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{k=0}^{N-1} f(\phi^k(\mathbf{z})) = \frac{1}{(2\pi)^N} \cdot \int_{\mathbb{T}^n} f(w) d\sigma.$$

Corollary 4.8. *Let α be as in the theorem. Then for any $\mathbf{z} \in \mathbb{T}^n$, the orbit $\{\phi^j(\mathbf{z})\}_{j \geq 0}$ is dense in $\mathbb{T}^n$. More precisely, if $\mathbf{I} \subset \mathbb{T}^n$ is an open interval, i. e.*

$$\mathbf{I} = \pi(I_1 \times I_2 \times \dots \times I_n)$$

with $I_j \subset [0, 1)$ open intervals, then

$$\lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{k=0}^{N-1} \chi_{\mathbf{I}}(\phi^k(\mathbf{z})) = \frac{|\mathbf{I}|}{|\mathbb{T}^n|}$$