1. Lecture I: Introduction

1.1. **Phase Spaces and Phase Flows.** Let X be a space (for us, X will usually be $\mathbb{R}^d$, with all its metric and measurability structures). Let $\phi : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathrm{Hom}(X, X)$ be a map with a "cocycle property"[1]:

$$\begin{aligned} \phi(t,t) &= \mathrm{Id}_X, \\ \phi(t,u) \circ \phi(u,s) &= \phi(t,s) \qquad u,t,s \in \mathbb{R}_{\geq 0}. \end{aligned} \tag{1}$$

Here ϕ takes in a terminal time and an initial time in its first and second arguments respectively, and acts upon an element $\mathbf{x} \in X$ by translating it to another point $\mathbf{y} \in X$. If X is a topological space or a measure space, we will usually want $\phi(t,s)$ respectively to be continuous or measurable. The pair (X, ϕ) is a (continuous-time) dynamical system.

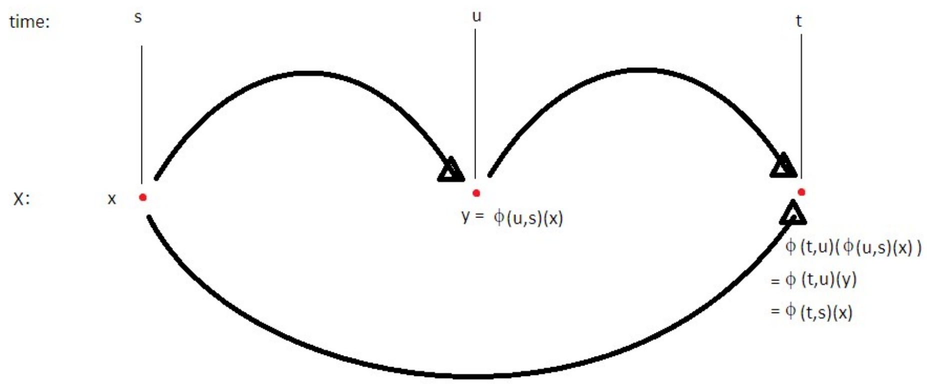

Figure 1. The cocycle property is a consistency criterion.

Example 1.1. Taking $X = \mathbb{R}^d$, we see that the system of first-order differential equations with (locally Lipschitz) $f : [0,T] \times X \to \mathbb{R}^d$,

$$\dot{\mathbf{x}}(t) = f(t, \mathbf{x}(t)), \tag{2}$$

can be construed as a dynamical system (if it has unique continuous solutions) by considering the pair $(\mathbb{R}^d, S)$ where S is the solution map $S : (t, s, \mathbf{x}(s)) \mapsto \mathbf{x}(t)$. (There is good reason not to think of the target space of f as X — it is in fact the tangent space of X, which only happens to be identifiable with X if X is $\mathbb{R}^d$, but this is not true of more general spaces.)

Thinking of ϕ as a map $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times X \to X$, we can also write (2) as

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{(t,s,\mathbf{x}(s))} \phi = f(t, \phi|_{(t,s,\mathbf{x}(s))}). \tag{3}$$

It is sometimes convenient to define ϕ satisfying this equation to be the flow of the vector field f, or of the dynamical system itself. We think of (2) as giving a description of a system for an aggregate of particles from a fixed reference frame, and we think of (3) as a description of the system given by following a specific particle starting at position $\mathbf{x}(s = 0)$. This is known as the

[1] This is not entirely accurate. Let G be a group acting on X. Then a dynamical system in our sense is the group action on the product $X \times_\rho U$, where U is an abelian group, via $g : (\mathbf{x}, u) \mapsto (g\mathbf{x}, u + \rho(g, \mathbf{x}))$. The map $\rho : G \times X \to U$, called the cocyle, satisfies the cocycle equation $\rho(gh, \mathbf{x}) = \rho(g, h\mathbf{x}) + \rho(h, \mathbf{x})$. Where G is the one-parameter group $\{\phi_t\}_{t \in \mathbb{R}}$ generating the flow, we can take $U = \mathbb{R}$ and $\rho(\phi_t, \mathbf{x}) = t$. More generally, we can define a dynamical system as the pair (G, Y), where G is a group that acts on a space Y.

LAGRANGIAN description of a system. In Lectures VIII and IX, we shall be putting this framework into more geometric terms.

We also call $\mathbf{x}(t)$ the INTEGRAL CURVE, or simply an INTEGAL, of f (or of the system), in the sense that the solution of a differential equation is an integral. The space X is called the PHASE SPACE, when it is conctrued to house the curves $\Gamma_+(\mathbf{x}) = \{\phi(t, 0, \mathbf{x}(0)) : t \in \mathbb{R}_{\geq 0}\}$. These curves are the FORWARD ORBITS of the dynamical system, and the plots are known as PHASE PORTRAITS.

The primary object of this module is then twofold:

(i) to study qualitative changes of phase portraits as they depend on f, and
(ii) to study the asymptotic behaviour of the solution/integral curve $\mathbf{x}(t)$.

To these ends we shall be looking at solution curves to differential equations that blow up to infinity asymptotically or in finite time, and spend extra effort exploring the complicated behaviours that can occur when solution curves do neither of the foregoing. Where $d = 2$, it is also practicable to draw phase portraits for an important class of systems, and we shall be applying ourselves to that activity.

We shall first be considering linear systems, i.e., where f is linear in $\mathbf{x}$, and then consider more general systems by looking at linearization of f, i.e., using the approximation

$$f(\mathbf{x}) \approx f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0),$$

in a small neighbourhood of a "critical" point $\mathbf{x}_0$ of particular interest. Some behaviours of nonlinear systems are so complicated that it is impossible to reduce them to their linearizations. That shall also be of interest to us, and we shall look at some techniques to characterize these behaviours.

If we have time we shall also be looking at more general dynamical systems not described by differential equations. Of those that are, we shall restrict our attention to *ordinary* differential equations.

1.2. Structure of this module.

In this module we shall be looking at, in order,

(i) linear systems,
(ii) local theory of nonlinear systems, that is, behaviour of systems around certain "critical" values (via linearization),
(iii) global theory of nonlinear systems (periodicity), and
(iv) bifurcation theory.

1.3. Matrix exponentiation.

The system of equations (2) is quite general in the sense that for a differential equation of any finite order d, it is possible to recast it as a system of d first-order equations by the change-of-variables $x^{i+1} = \mathrm{d}^i x/\mathrm{d}t^i$ for $i = 0, \ldots d-1$, and $\mathbf{x} = (x^1, x^2, \ldots x^d)$ (superscripts are indices, not powers).

As aforementioned, we shall first be studying linear systems. Our attitude towards linear systems in $X = \mathbb{R}^d$ will be to seek to derive solutions explicitly rather than concern ourselves with the general and abstract question of existence and uniqueness.

By linearity we mean

$$f(t, \mathbf{x}(t)) = \mathbf{A}(t)\mathbf{x}(t),$$

where $\mathbf{A}(t)$ is a $d \times d$ matrix, and $\mathbf{x}(t) = (x^1(t), \ldots, x^d(t))^\top$.

First we impose an additional condition of autonomy. We say that a system is AUTONOMOUS if $f(t, \mathbf{x}(t)) = \tilde{f}(\mathbf{x}(t))$, or $\mathbf{A}(t) = \mathbf{A}$ in the linear case, i.e., there is no dependence on t except through the solution $\mathbf{x}(t)$ — there is no direct dependence of the system on t. This means that the rules governing how a particle moves do not themselves change over time. This additional requirement of autonomy (which can be imposed without linearity) reduces the cocycle property (1) on the flow ϕ to a simpler for which $\phi : \mathbb{R}_{\geq 0} \to \mathrm{Hom}(X, X)$, in which the flow only records the *duration* and not the terminal and initial times, and for which

$$\begin{aligned} \phi_0 &= \mathrm{Id}_X, \\ \phi_t \circ \phi_s &= \phi_{t+s} \qquad t, s \in \mathbb{R}_{\geq 0}. \end{aligned} \tag{4}$$

It is conventional to write the temporal argument in the subscript.

Autonomy means further that $\mathbf{A}$ is *time-independent*.

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}(t) = \mathbf{A}\mathbf{x}(t), \tag{5}$$

then, the $d = 1$ case suggests we can write down a solution

$$\mathbf{x}(t) = \exp(\mathbf{A}t)\mathbf{x}(0),$$

if we can properly interpret $\exp(\mathbf{A}t)$. If this were the case, the map $\exp(\mathbf{A}t) : \mathbb{R}^d \to \mathbb{R}^d$ can then be identified with the flow ϕ_t of the autonomous system.

We shall now devote some time to showing what $\exp(\mathbf{A}t)$ must mean and why this intuitive form of the solution is helpful in determining asymptotic behaviour.

Using the Taylor expansion for the exponential, we shall tentatively define

$$\exp(\mathbf{A}t) := \sum_{n=0}^{\infty} \frac{t^n}{n!}\mathbf{A}^n.$$

We shall require a few things:

(i) The series converges in appropriate topologies/norms.
(ii) The series satisfies number of properties expected of exponentials under appropriate conditions.
(iii) The series applied to $\mathbf{x}(0)$ is a solution to the initial value problem $\mathrm{d}\mathbf{x}/\mathrm{d}t = \mathbf{A}\mathbf{x}$.

In which topology do we want the series to converge? Ideally, if $\mathbf{x}_0 \in \mathbb{R}^d$ is a fixed vector, we want the partial sums

$$\mathbf{M}_N = \sum_{n=0}^{N} \frac{t^n}{n!}\mathbf{A}^n$$

to converge in such a way that for some $\mathbf{M}_\infty$

$$\sup_{t\in[0,T]} \|\mathbf{M}_N\mathbf{x} \to \mathbf{M}_\infty\mathbf{x}\|_{\ell^2(\mathbb{R}^d)} \to 0 \tag{6}$$

for any $\mathbf{x} \in \mathbb{R}^d$. We can then set $\exp(\mathbf{A}t) = \mathbf{M}_\infty$ and check that it satisfies the equation (5) on $\mathbb{R}^d$.

If $\mathbf{M}_\infty$ exists, we find

$$\|\mathbf{M}_N\mathbf{x} - \mathbf{M}_\infty\mathbf{x}\|_{\ell^2(\mathbb{R}^d)} \leq \|\mathbf{M}_N - \mathbf{M}_\infty\|_{\ell^2\to\ell^2} \|\mathbf{x}\|_{\ell^2(\mathbb{R}^d)},$$

where

$$\|\mathbf{M}\|_{\ell^2\to\ell^2} := \sup_{|\mathbf{x}|\leq 1} \|\mathbf{M}\mathbf{x}\|_{\ell^2(\mathbb{R}^d)} \left(= \sup_{\mathbf{x}\in\mathbb{R}^d\setminus\{0\}} \frac{\|\mathbf{M}\mathbf{x}\|_{\ell^2(\mathbb{R}^d)}}{\|\mathbf{x}\|_{\ell^2(\mathbb{R}^d)}} \right).$$

is the OPERATOR NORM of linear operators $\ell^2(\mathbb{R}^d) \to \ell^2(\mathbb{R}^d)$. The topology induced by the convergence of this norm is the NORM TOPOLOGY. We see that convergence in the operator norm would imply convergence in the sense (6).

Recall that this sort of convergence can happens to Cauchy sequences in complete metric spaces. Our metric space is $\mathbb{R}^{d\times d}$ with the metric induced by the norm $\|\cdot\|_{\ell^2\to\ell^2}$.

a. the normed space is complete

First, it can be readily verified that this norm is equivalent to the uniform norm:

$$C_d^{-1} \sup_{ij} |\mathbf{M}_{ij}| \leq \|\mathbf{M}\|_{\ell^2\to\ell^2} \leq C_d \sup_{ij} |\mathbf{M}_{ij}|,$$

where C_d is a constant that depends only on the (finite) dimension d.

The space of $d \times d$ real matrices is complete under the uniform norm by the completeness of real numbers, and so it is complete under the $\ell^2 \to \ell^2$ norm (with supremum in t). (Equivalent norms induce equivalent topologies — because of this equivalence we often drop the subscript on the norm.)

b. the sequence is Cauchy

Secondly, by definition,

$$\|\mathbf{MNx}\|_{\ell^2(\mathbb{R}^d)} \leq \|\mathbf{M}\|_{\ell^2\to\ell^2}\|\mathbf{Nx}\|_{\ell^2} \leq \|\mathbf{M}\|_{\ell^2\to\ell^2}\|\mathbf{N}\|_{\ell^2\to\ell^2}\|\mathbf{x}\|_{\ell^2(\mathbb{R}^d)},$$

and therefore, for two $d \times d$ matrices $\mathbf{M}$ and $\mathbf{N}$,

$$\|\mathbf{MN}\|_{\ell^2\to\ell^2} \leq \|\mathbf{M}\|_{\ell^2\to\ell^2}\|\mathbf{N}\|_{\ell^2\to\ell^2}.$$

We say that the linear maps $\ell^2(\mathbb{R}^d) \to \ell^2(\mathbb{R}^d)$ form an ALGEBRA under the operator norm.

We then have

$$\|\mathbf{M}_k\|_{\ell^2\to\ell^2} \leq \sum_{n=0}^{k} \frac{|t^n|}{n!}\|\mathbf{A}\|^n_{\ell^2\to\ell^2},$$

by the triangle inequality and the algebra property.

Next, for any $\varepsilon > 0$, we can find N such that

$$\|\mathbf{M}_M - \mathbf{M}_{M+N}\|_{\ell^2\to\ell^2} = \left\| \sum_{n=M+1}^{M+N} \frac{t^n}{n!}\mathbf{A}^n \right\|_{\ell^2\to\ell^2} \leq \sum_{n=M+1}^{M+N} \frac{t^n}{n!}\|\mathbf{A}\|^n_{\ell^2\to\ell^2}.$$

Therefore $\{M_k\}$ are a Cauchy sequence in the norm and by completeness the sequence, and hence the sum, converges after taking a supremum over $t \in [0, T]$ (in fact, converges absolutely).

This gives us a candidate $\mathbf{M}_\infty$ for the solution $\exp(\mathbf{A}t)$.

Finally, to explain some parts of the next lemma, let us mention that matrix-valued functions $\varphi : \mathbb{R} \to \mathbb{R}^{d\times d}$ are DIFFERENTIABLE at t_0 if there exists a $d \times d$ matrix $\mathbf{N}$ such that

$$\lim_{t\to t_0} \left\|\mathbf{N} - \frac{1}{t - t_0}(\varphi(t) - \varphi(t_0))\ \right\|_{\ell^2\to\ell^2} = 0.$$

We write $\varphi'(t)$ for the matrix $\mathbf{N}$. Note that we could have done this element-wise as we have shown that the $\ell^2 \to \ell^2$ norm and the uniform norm are equivalent norms in finite dimensions.

Now that we have a convergent Taylor series defining $\exp(\mathbf{A}t)$ we can easily show that:

Lemma 1.1. (i) *Let $\varphi(t)$ and $\psi(t)$ be differentiable functions $\mathbb{R} \to \mathbb{R}^{d\times d}$. Then the Leibnitz rule holds that*

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(\varphi(t)\psi(t)\big) = \varphi'(t)\psi(t) + \varphi(t)\psi'(t).$$

(Again, note that ϕ' and ψ', whilst being $d \times d$ real matrices, reside in fact in the tangent space of the space of $d \times d$ matrices.)

(ii) *Let $\mathbf{0}_d$ denote the zero matrix. Then*

$$\exp(\mathbf{0}_d) = \mathbf{I}_d.$$

(iii) *For any $d \times d$ matrix $\mathbf{A}$,*

$$\Big(\exp(\mathbf{A})\Big)^{-1} = \exp(-\mathbf{A}).$$

(iv) *If matrices $\mathbf{A}$ and $\mathbf{B}$ commute, then*

$$\exp(\mathbf{A}t)\exp(\mathbf{B}t) = \exp((\mathbf{A} + \mathbf{B})t).$$

(v) *Let* $\mathbf{P}$ *be a nonsingular* $d \times d$ *matrix, and set* $\mathbf{\Lambda} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. *(That is,* $\mathbf{\Lambda} \sim \mathbf{A}$ *— i.e., the matrices are similar.) Then*

$$\exp(\mathbf{\Lambda}) = \mathbf{P}^{-1}\exp(\mathbf{A})\mathbf{P}.$$

Proof. (i) This holds by direct calculation element-wise.

(ii) The norm $\|\mathbf{0}_d\|_{\ell^2 \to \ell^2}$ can be explicitly found to be 0. The statement follows by the series expansion.

(iii) Take $\psi(t) = \exp(-\mathbf{A}t)\exp(\mathbf{A}t)$ and find that its derivative is 0 via the Leibnitz rule and use (ii) with $t = 0$.

(iv) Take $\psi(t) = \exp(-(\mathbf{A}+\mathbf{B})t)\exp(\mathbf{A}t)\exp(\mathbf{B}t)$ and find that its derivative is 0 via the Leibnitz rule and use (ii). Naturally within the parenthesis $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$. Therefore the result cannot be hoped to be true in general.

(v) Using the series expansion we see that the result is true for a partial sum. Then we take a limit.

□

These properties also allow us to verify that the candidate $\exp(\mathbf{A}t)\mathbf{x}$ is indeed a solution of the equation (5):

Theorem 1.2 (Fundamental theorem for Linear Systems). *For every* $\mathbf{b} \in \mathbb{R}^d$*, the initial value problem*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}(t) = \mathbf{A}\mathbf{x}(t), \qquad \mathbf{x}(0) = \mathbf{b}$$

has a unique solution

$$\mathbf{x}(t) = \exp(\mathbf{A}t)\mathbf{b}.$$

Proof. It can be readily verified that $\mathbf{x}$ as constructed is a solution.

Now suppose $\mathbf{x}$ is a solution. Then set $\mathbf{y}(t) = \exp(-\mathbf{A}t)\mathbf{x}(t)$.

Differentiating $\mathbf{y}(t)$ from first principles we find that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}(t) = -\exp(-\mathbf{A}t)\mathbf{A}\mathbf{x}(t) + \exp(-\mathbf{A}t)\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}(t).$$

We can use the equation to expand the final term to find that the derivative is the zero vector. This implies that

$$\mathbf{y}(t) = \mathbf{y}(0) = \mathbf{x}(0),$$

and by construction $\mathbf{x}(t)$ must take the form required by the theorem statement. □