Appendix C

Ordinary differential equations

Although the subject of these notes is not dynamical systems, in order to say some elementary but useful things about the behaviour of Lagrangian systems, it is essential to have on hand a collection of standard tools for handling differential equations. We assume that the reader knows what a differential equation is, and is aware of the existence and uniqueness theory for such.

C.1 Linear ordinary differential equations

Although linear equations themselves are not necessarily interesting for us—few Lagrangian systems are actually linear—when linearising differential equations, one naturally obtains linear equations (of course). Thus we record some basic facts about linear differential equations.

Let $V$ be an $n$-dimensional $\mathbb{R}$-vector space. A linear ordinary differential equation with constant coefficients on $V$ is a differential equation of the form

$$\dot{x} = Ax, \quad x(0) = x_0$$

(C.1)

for a curve $t \mapsto x(t) \in V$ and where $A : V \to V$ is a linear transformation. Given a linear transformation $A$ we define a linear transformation $e^A$ by

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}.$$

This series may be shown to converge using the fact that the series for the usual exponential of a real variable converges. We next claim that the solution to (C.1) is $x(t) = e^{At}x_0$. To see this we simply substitute this proposed solution into the differential equation:

$$\dot{x}(t) = \frac{d}{dt} \sum_{k=1}^{\infty} \frac{A^k t^k}{k!} = \sum_{k=1}^{\infty} \frac{A^k t^{k-1}}{k!} = A \sum_{k=1}^{\infty} \frac{A^{k-1} t^{k-1}}{(k-1)!} = A \sum_{\ell=0}^{\infty} \frac{A^\ell t^\ell}{\ell!} = Ax(t).$$

Thus $x(t)$ satisfies the differential equation. It is also evident that $x(t)$ satisfies the initial conditions, and so by uniqueness of solutions, we are justified in saying that the solution of the initial value problem (C.1) is indeed $x(t) = e^{At}x_0$.

Let us turn to the matter of computing $e^{At}$. First note that if $P$ is an invertible linear transformation on $V$ then one readily shows that $e^{P(A)t} = Pe^{At}P^{-1}$. Thus $e^{At}$ is independent of similarity transformations, and so we may simplify things by choosing a basis $\{e_1, \ldots, e_n\}$ for $V$ so that the initial value problem (C.1) becomes the ordinary differential equation

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

(C.2)

in $\mathbb{R}^n$. To compute $e^{At}$ one proceeds as follows. First one computes the eigenvalues for $A$. There will be $n$ of these in total, counting algebraic multiplicities and complex conjugate pairs. One treats each eigenvalue separately. For a real eigenvalue $\lambda_0$ with algebraic multiplicity $k = m_0(\lambda_0)$, one must compute $k$ linearly independent solutions. For a complex eigenvalue $\lambda_0$ with algebraic multiplicity $\ell = m_0(\lambda_0)$, one must compute $2\ell$ linearly independent solutions, since $\lambda_0$ is also necessarily an eigenvalue with algebraic multiplicity $\ell$.

We first look at how to deal with real eigenvalues. Let $\lambda_0$ be one such object with algebraic multiplicity $k$. It is a fact that the matrix $(A - \lambda_0 I_n)^k$ will have rank $n - k$, and so will have a kernel of dimension $k$ by the Rank-Nullity Theorem. Let $u_1, \ldots, u_k$ be a basis for ker$((A - \lambda_0 I_n)^k)$. We call each of these vectors a generalised eigenvector.

If the geometric multiplicity of $\lambda_0$ is also $k$, then the generalised eigenvectors will simply be the usual eigenvectors. If $m_0(\lambda_0) < m_0(\lambda_0)$ then a generalised eigenvector may or may not be an eigenvector. Corresponding to each generalised eigenvector $u_i$, $i = 1, \ldots, k$, we will define a solution to (C.2) by

$$x_i(t) = e^{\lambda_0 t} \exp((A - \lambda_0 I_n)t)u_i.$$ 

Note that because $u_i$ is a generalised eigenvector, the infinite series $\exp((A - \lambda_0 I_n)t)u_i$ will have only a finite number of terms—at most $k$ in fact. Indeed we have

$$\exp((A - \lambda_0 I_n)t)u_i = \left( I_n + t(A - \lambda_0 I_n) + \frac{t^2}{2!}(A - \lambda_0 I_n)^2 + \cdots + \frac{t^{k-1}}{(k-1)!}(A - \lambda_0 I_n)^{k-1} \right) u_i,$$

since the remaining terms in the series will be zero. In any case, it turns out that the $k$ vector functions $x_1(t), \ldots, x_k(t)$ so constructed will be linearly independent solutions of (C.2). This tells us how to manage the real case.

Now let us look at the complex case. Thus let $\lambda_0$ be a complex eigenvalue (with nonzero imaginary part) of algebraic multiplicity $\ell$. This means that $\lambda_0$ will also be an eigenvalue of algebraic multiplicity $\ell$ since $A$, and hence $P(A)$, is real. Thus we need to find $2\ell$ linearly independent solutions. We do this by following the exact same idea as in the real case, except that we think of $A$ as being a complex matrix for the moment. In this case it is still true that the matrix $(A - \lambda_0 I_n)^\ell$ will have an $\ell$-dimensional kernel, and we can take vectors $u_1, \ldots, u_\ell$ as a basis for this kernel. Note, however, that since $(A - \lambda_0 I_n)^\ell$ is complex, these vectors will also be complex. But the procedure is otherwise identical to the real case. One then constructs $\ell$ complex vector functions

$$z_j(t) = e^{\lambda_0 t} \exp((A - \lambda_0 I_n)t)u_j.$$ 

Each such complex vector function will be a sum of its real and imaginary parts: $z_j(t) = x_j(t) + iy_j(t)$. It turns out that the $2\ell$ real vector functions $x_1(t), \ldots, x_\ell(t), y_1(t), \ldots, y_\ell(t)$ are linearly independent solutions to (C.2).
We still haven’t gotten to the matrix exponential yet, but all the hard work is done. Using the above methodology we may in principle compute for any $n \times n$ matrix $A$, $n$ linearly independent solutions $x_1, \ldots, x_n(t)$. If we assemble the resulting solutions into the columns of a matrix $X(t)$:

$$X(t) = \begin{bmatrix} x_1(t) & \cdots & x_n(t) \end{bmatrix},$$

the resulting matrix is an example of a fundamental matrix. Generally, a fundamental matrix is any $n \times n$ matrix function of $t$ whose columns form $n$ linearly independent solutions to (C.2). What we have done above is give a recipe for computing a fundamental matrix (there are an infinite number of these). The following result connects the construction of a fundamental matrix with the matrix exponential.

**C.1.1 Theorem** Given any fundamental matrix $X(t)$ we have $e^{At} = X(t)X^{-1}(0)$.

Thus, once we have a fundamental matrix, the computation of the matrix exponential is just algebra, although computing inverses of matrices of any size is a task best left to the computer.

One of the essential observations from the above discussion is that the behaviour of the solutions to the differential equation (C.1) are largely governed by the eigenvalues of $A$.

### C.2 Fixed points for ordinary differential equations

In this section we consider the differential equation

$$\dot{x}(t) = f(x(t))$$  \hfill (C.3)

for $x(t) \in \mathbb{R}^n$ and with $f: \mathbb{R}^n \to \mathbb{R}^n$ a smooth map. A fixed point for (C.3) is a point $x_0$ for which $f(x_0) = 0$. Thus if $x_0$ is a fixed point, the trivial curve $t \mapsto x_0$ is a solution to the differential equation. A fixed point $x_0$ is **stable** if for each $\epsilon > 0$ there exists $\delta > 0$ so that if $\|x(0) - x_0\| < \delta$, then $\|x(t) - x_0\| < \epsilon$ for all $t > 0$. The fixed point $x_0$ is **asymptotically stable** if there exists $\delta > 0$ so that if $\|x(0) - x_0\| < \delta$ then $\lim_{t \to \infty} x(t) = x_0$. These notions of stability are often said to be “in the sense of Liapunov,” to distinguish them from other definitions of stability. In Figure C.1 we give some intuition concerning our definitions.

As a first pass at trying to determine when a fixed point is stable or asymptotically stable, one linearises the differential equation (C.3) about $x_0$. Thus one has a solution $\bar{x}(t)$ of the differential equation, and uses the Taylor expansion to obtain an approximate expression for the solution:

$$\frac{d}{dt}(x(t) - x_0) = f(x_0) + Df(x_0) \cdot (x(t) - x_0) + \cdots$$

$$\Rightarrow \quad \xi(t) = Df(x_0) \cdot \xi(t) + \cdots$$

where $\xi(t) = x(t) - x_0$. Thus linearisation leads us to think of the differential equation

$$\dot{\xi}(t) = Df(x_0) \cdot \xi(t)$$  \hfill (C.4)

as somehow approximating the actual differential equation near the fixed point $x_0$. Let us define notions of stability of $x_0$ which are related only to the linearisation. We say that $x_0$ is **spectrally stable if** $Df(x_0)$ has no eigenvalues in the positive complex plane, and that $x_0$ is **linearly stable** (resp. **linearly asymptotically stable**) if the linear system (C.4) is stable (resp. asymptotically stable). Let us introduce the notation

$$C_+ = \{ z \in \mathbb{C} \mid \text{Re}(z) > 0 \}$$

$$C_0 = \{ z \in \mathbb{C} \mid \text{Re}(z) \geq 0 \}$$

$$C_- = \{ z \in \mathbb{C} \mid \text{Re}(z) < 0 \}$$

$$C_{-\infty} = \{ z \in \mathbb{C} \mid \text{Re}(z) \leq 0 \}.$$

From our discussion of linear ordinary differential equations in Section C.1 we have the following result.

**C.2.1 Proposition** A fixed point $x_0$ is linearly stable if and only if the following two conditions hold:

(i) $Df(x_0)$ has no eigenvalues in $C_+$, and

(ii) all eigenvalues of $Df(x_0)$ with zero real part have equal geometric and algebraic multiplicities.

The point $x_0$ is linearly asymptotically stable if and only if all eigenvalues of $Df(x_0)$ lie in $C_-\infty$.

A question one can ask is how much stability of the linearisation has to do with stability of the actual fixed point. More generally, one can speculate on how the solutions of the linearisation are related to the actual solutions. The following important theorem due to Hartman and Grobman tells us when we can expect the solutions to (C.3) near $x_0$ to “look like” those of the linear system. The statement of the result uses the notion of a flow which we define in Section 2.2.

**C.2.2 Theorem** (Hartman-Grobman Theorem) Let $x_0$ be a fixed point for the differential equation (C.3) and suppose that the $n \times n$ matrix $Df(x_0)$ has no eigenvalues on the imaginary axis. Let $F_t$ denote the flow associated with the differential equation (C.3). Then there exists a neighbourhood $V$ of $0 \in \mathbb{R}^n$ and a neighbourhood $U$ of $x_0 \in \mathbb{R}^n$, and a homeomorphism (i.e., a continuous bijection) $\phi: V \to U$ with the property that $\phi(e^{Dt}(x_0)) = F_t(\phi(x_0))$. 

![Figure C.1 A stable fixed point (left) and an asymptotically stable fixed point (right)](image)
The idea is that when the eigenvalue all have nonzero real part, then one can say that the flow for the nonlinear system (C.3) looks like the flow of the linear system (C.4) in a neighbourhood of \( x_0 \). When \( Df(x_0) \) has eigenvalues on the imaginary axis, one cannot make any statements about relating the flow of the nonlinear system with its linear counterpart. In such cases, one really has to look at the nonlinear dynamics, and this becomes difficult.

A proof of the Hartman-Grobman theorem can be found in [Palis, Jr. and de Melo 1982].

Let us be more specific about the relationships which can be made about the behaviour of the nonlinear system (C.3) and its linear counterpart (C.4). Let \( E^s(x_0) \) be the subspace of \( \mathbb{R}^n \) containing all generalised eigenvectors for eigenvalues of \( Df(x_0) \) in \( C_- \), and let \( E^u(x_0) \) be the subspace of \( \mathbb{R}^n \) containing all generalised eigenvectors for eigenvalues of \( Df(x_0) \) in \( C_+ \). \( E^s(x_0) \) is called the stable subspace at \( x_0 \), and \( E^u(x_0) \) is called the linear unstable subspace at \( x_0 \). For the linearised system (C.4), the subspaces \( E^s(x_0) \) and \( E^u(x_0) \) will be invariant sets (i.e., if one starts with an initial condition in one of these subspaces, the solution of the differential equation will remain on that same subspace). Indeed, initial conditions in \( E^s(x_0) \) will tend to 0 as \( t \to \infty \), and initial conditions in \( E^u(x_0) \) will explode as \( t \to \infty \). The following result says that analogues of \( E^s(x_0) \) and \( E^u(x_0) \) exist.

**C.2.3 Theorem (Stable and Unstable Manifold Theorem)** Let \( x_0 \) be a fixed point for the differential equation (C.3).

(i) There exists a subset \( W^s(x_0) \) of \( \mathbb{R}^n \) with the following properties:

- \( W^s(x_0) \) is invariant under the flow \( F_t \);
- \( E^s(x_0) \) forms the tangent space to \( W^s(x_0) \) at \( x_0 \);
- \( \lim_{t \to -\infty} F_t(t, x) = x_0 \) for all \( x \in W^s(x_0) \).

(ii) There exists a subset \( W^u(x_0) \) of \( \mathbb{R}^n \) with the following properties:

- \( W^u(x_0) \) is invariant under the flow \( F_t \);
- \( E^u(x_0) \) forms the tangent space to \( W^u(x_0) \) at \( x_0 \);
- \( \lim_{t \to -\infty} F_t(t, x) = x_0 \) for all \( x \in W^u(x_0) \).

\( W^s(x_0) \) is called the **stable manifold** for the fixed point \( x_0 \), and \( W^u(x_0) \) is called the **unstable manifold** for the fixed point \( x_0 \).

Again, we refer to [Palis, Jr. and de Melo 1982] for a proof. Some simple examples of stable and unstable manifolds can be found in Section 3.4. A picture of what is stated in this result is provided in Figure C.2. The idea is that the invariant sets \( E^s(x_0) \) and \( E^u(x_0) \) for the linear system do have counterparts in the nonlinear case. Near \( x_0 \) they follow the linear subspaces, but when we go away from \( x_0 \), we cannot expect things to look at all like the linear case, and this is exhibited even in the simple examples of Section 3.4.