Dynamical systems: some basic facts

Harald Hanche-Olsen
hanche@math.ntnu.no

An autonomous[1] *dynamical system* is specified by an ordinary differential equation

$$\dot{x} = f(x)$$

where $f\colon \Omega \to \mathbb{R}^n$ is a smooth function, $\Omega \subset \mathbb{R}^n$ is open, and the unknown function x is supposed to be a function of a single real variable t with values in Ω.

More precisely, the minimum requirement on f is that it be a *locally Lipschitz* function, which means that whenever $K \subset \Omega$ is a compact set there exists a constant L so that $|f(x) - f(y)| \le L|x-y|$ whenever $x, y \in K$.[2] Usually, however, we simply assume the condition that $f \in C^1$ (i.e., its partial derivatives of first order exist and are continuous).

Clearly, f in the above definition is a *vector field*, so in a sense, a dynamical system is nothing but a smooth vector field. However, we usually think of dynamical systems in somewhat different terms than vector fields.

The basic existence and uniqueness theorem on ordinary differential equations states that the equation $\dot{x} = f(x)$ with an initial condition on the form $x(0) = x_0$ has a unique solution in a neighbourhood of 0. Moreover, this solution can be extended (still in a unique way) to a maximal open interval containing 0, so that the solution $x(t)$ is defined for $a < t < b$ where $-\infty \le a < 0 < b \le \infty$. If $b < \infty$, it must be because either $x(t)$ is unbounded, or $x(t)$ gets arbitrarily close to the boundary of Ω as $t \to b$ (roughly speaking, $x(t)$ leaves Ω or escapes to infinity). A similar statement goes in case $a > -\infty$.

The *flow* of the dynamical system is the function Φ defined by[3] $\Phi_t(x_0) = x(t)$, where x is the solution of $\dot{x} = f(x)$ satisfying $x(0) = x_0$. For any given x_0, $\Phi_t(x_0)$ is defined for t in the maximal interval of existence for this initial value problem.

Put differently, the flow is defined by

$$\Phi_0(x) = x, \quad \frac{\partial}{\partial t}\Phi_t(x) = f\big(\Phi_t(x)\big).$$

[1]A non-autonomous dynamical system is given by $\dot{x} = f(x,t)$. It can be rewritten as an autonomous system by adding the equation $\dot{\tau} = 1$ and replacing x by (x,τ).

[2]f is called (globally) Lipschitz if the same "Lipschitz constant" L can be chosen for all of Ω.

[3]A notation like $\Phi(x_0, t)$ may seem more natural, but the present notation is the more conventional, and in fact more convenient.

Moreover, it is a basic fact of life that the flow satisfies the following group property:

$$\Phi_t \circ \Phi_s = \Phi_{s+t}$$

wherever both sides are defined: For $\Phi_{s+t}(x_0) = \Phi_t\big(\Phi_s(x_0)\big)$ when $t = 0$, and both sides satisfy the differential equation $\dot{x} = f(x)$.

The flow of a C^1 vector field is itself a C^1 function of t and x. In particular, each Φ_t is a C^1 function, with a C^1 inverse Phi_{-t}.

Given any point $x \in \Omega$, the set of all points $\Phi_t(x)$ (where that is defined) is called an *orbit* of the dynamical system. It is either a single point (if x is an equilibrium point, see the next section) or a curve. Any two orbits having a single common point are in fact identical.

A diagram showing a representative set of orbits (with direction) is called a *phase diagram* of the dynamical system. The study of the phase diagram is an indispensible tool in understanding the qualitative behaviour of solutions to the system. Maple can draw phase diagrams, and so can many other systems.

Equilibrium points

An *equilibrium point* of a dynamical system is a point x_0 where $f(x_0) = 0$. Equivalently, $\Phi_t(x_0) = x_0$ for all t.

To understand equilibrium points, let us digress into Taylor's formula for a moment. If g is a C^2 function of one variable, we can write

$$\begin{aligned} g(1) &= g(0) + \int_0^1 g'(t)\,dt = g(0) + \Big[(t-1)g'(t)\Big]_0^1 - \int_0^1 (t-1)g''(t)\,dt \\ &= g(0) + g'(0) + \int_0^1 (1-t)g''(t)\,dt \end{aligned} \tag{1}$$

by a not-so-obvious integration by parts. Once we're past that point, the obvious induction leads to

$$g(1) = \sum_{k=0}^{n} \frac{g^{(k)}(0)}{k!} + \frac{1}{n!}\int_0^1 (1-t)^k g^{(n+1)}(t)\,dt.$$

If f is any function, we can apply this to the function $g(t) = f(x+th)$ and obtain the usual Taylor's formula with an exact formula for the remainder term, but that is a digression we shall skip for now.

Instead, we do the same thing for a vector field f, assumed to belong to C^2, and apply (1). The result can be written

$$\begin{aligned} f(x+h) &= f(x) + \sum_j \frac{\partial f(x)}{\partial x_j} h_j + \int_0^1 (1-t) \sum_{j,k} \frac{\partial^2 f(x+th)}{\partial x_j \partial x_k} h_j h_k \, dt \\ &= f(x) + f'(x)h + O(|h|^2), \end{aligned}$$

where h is supposed to be a column vector and $f'(x)$ is the matrix with (i,j) component $\partial f_i / \partial x_j$.

Now assume that x_0 is an equilibrium point, so $f(x_0) = 0$. Write $A = f'(x_0)$. Then the above formula becomes $f(x_0 + h) = Ah + O(|h|^2)$.

Consider now a solution to $\dot{x} = f(x)$ in a neighbourhood of x_0. That is, we can write $x = x_0 + \varepsilon X$ where $0 < \varepsilon \ll 1$ and $X = O(1)$. The scaled equation becomes $\dot{X} = AX + O(\varepsilon)$. Throwing away the final term, then, we arrive at the *linearization* of the original equation $\dot{x} = f(x)$ at the equlibrium point x_0: Namely

$$\dot{X} = AX, \qquad A = f'(x_0).$$

The **Hartman–Grobman** theorem states that, if $f \in C^2$ and x_0 is an equilibrium point, and if the matrix $A = f'(x_0)$ has *no eigenvalues with real part zero*, then the flow of $\dot{x} = f(x)$ in a neighbourhood of x_0 is similar to the flow of the linearized system $\dot{X} = AX$ in the following sense: There is a C^1 diffeomorphism[4] from a neighbourhood of x_0 to a neighbourhood of 0 mapping orbits of one system to orbits of the other, with the same orientation of the orbits.[5]

Because of this, classifying linear systems is quite interesting, since a classification of many equilibrium points follows for free.

We can state three definitions for an equilibrium x_0:

– It is called *stable* if, for each $\varepsilon > 0$ there is a $\delta > 0$ so that whenever $|x - x_0| < \delta$ then $|\Phi_t(x) - x_0| < \varepsilon$ for all $t > 0$;

– it is called *unstable* if it is not stable;

– and it is called *asymptotically stable* if it is stable and there is some $\delta > 0$ so that $\Phi_t(x) \to x_0$ when $t \to \infty$ for all x with $|x - x_0| < \delta$.

From the Hartman–Grobman theorem we immediately conclude that x_0 is asymptotically stable if all eigenvalues of A have negative real part.

[4] A C^1 map with a C^1 inverse.

[5] That is perhaps a bit more precise than we need here, but there you have it.

Also, x_0 is unstable if any eigenvalue has a positive real part. (This does not follow from Hartman–Grobman in the case where some eigenvalue has real part zero, but it is true never the less.)

Classification of equilibrium points in the plane. For two-dimensional systems, the classification of most eqlibria boils down to the study of 2×2-matrices. The characteristic equation of such a matrix can be written

$$\det(A - \lambda I) = \lambda^2 - \tau\lambda + \delta = (\lambda - \lambda_1)(\lambda - \lambda_2)$$

where λ_1 and λ_2 are the eigenvalues, $\delta = \lambda_1 \lambda 2 = \det A$, and $\tau = \lambda_1 + \lambda 2 = \operatorname{tr} A$.

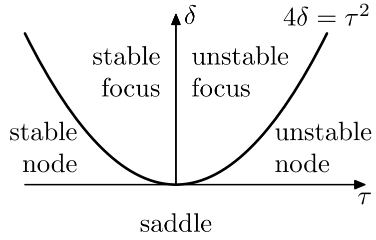

Figure 1: Classification of equilibrium points

The two eigenvalues can also be written

$$\lambda_{\pm} = \tfrac{1}{2}\Bigl(\tau \pm \sqrt{\tau^2 - 4\delta}\Bigr),$$

so if $4\delta > \tau^2$ then the eigenvalues are complex with real part τ, and the equilibrium is asymptotically stable if $\tau < 0$ and stable if $\tau > 0$. We call such an equilibrium a *focus*.

When $4\delta \le \tau^2$ then both eigenvalues are real. If $\delta < 0$ they have opposite signs, so the equilibrium is unstable. Such an equilibrium is called a *saddle point*.

When $4\delta \le \tau^2$ and $\delta > 0$ then we have real eigenvalues of the same sign, so the system is asymptotically stable if $\tau < 0$ and unstable if $\tau > 0$. These equilibria are called *nodes*.

To sum up, saddles are alway unstable, while a focus or a node can be stable[6] or unstable.

[6] If a focus or node is stable, it is asymptotically stable, so we don't bother to mention the word "asymptotically" in this case.

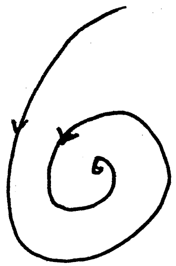

Figure 2: A stable focus

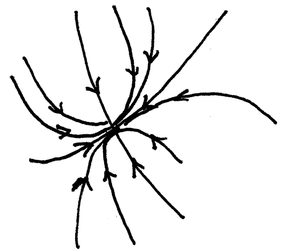

Figure 3: A stable node

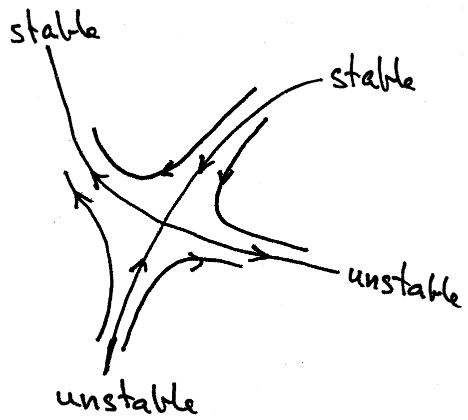

Figure 4: A saddle point

Saddle points are especially interesting, since there are four special orbits, or separatrices associated with a saddle point: These are defined by $\Phi_t(x) \to x_0$ as $t \to \pm\infty$. The points x for which $\lim_{t\to\infty} = x_0$ belong to the *stable manifold* or a *stable separatrix* of x_0, whil the points x for which $\lim_{t\to-\infty} = x_0$ belong to the *unstable manifold* or an *unstable separatrix* of x_0.

If x and y are close together, but on different sides of a stable separatrix, then as t grows, $\Phi_t(x)$ and $\Phi_t(y)$ will first travel together towards x_0, but then they will shoot off in roughly opposite directions: They must separate, and hence the name separatrix.

Cycles

A *cycle* of a dynamical system is simply a periodic orbit: There is a period $T > 0$ so that $\Phi_T(x) = x$ for any x in such an orbit. The existence (or not) of periodic orbits is often of great interest. One often encounters periodic orbits in predator-prey systems: The prey species multiplies strongly while predators are few; then with plenty of prey, the predators multiply too, then the number of predators grows too great, they almost eradicate the prey species, after which the predators starve. Now the prey species is free to multiply again, and the cycle repeats.

The **Poincaré–Bendixsson** theorem provides a sufficient condition for the existence of a cycle for a dynamical system in the plane.

We call a subset $K \subset \Omega$ *forward invariant* if a point once in K never leaves it: If $x \in K$ then $\Phi_t(x) \in K$ for every $t > 0$. The Poincaré–Bendixsson theorem states that, if K is a compact forward invariant subset of Ω, and if there are no equilibrium points in K, then K contains a cycle.

(The start of the proof is simple enough: Take any point in K and follow it forward in time. By compactness, $\Phi_t(x)$ must have a limit point in K as $t \to \infty$. It turns out that this limit point must belong to a cycle. It is essential that we are working in the plane; in three or more dimensions, the proof – and the result – breaks down in quite spectacular ways. Keywords: chaos, strange attractors, Lorenz system.)

As an example, consider the **van der Pol** equation:

$$\ddot{x} + \varepsilon(3x^2 - 1)\dot{x} + x = 0, \qquad \varepsilon > 0.$$

It can be written as a system of two first order equations in the obvious way ($\dot{x} = y$, $\dot{y} = \cdots$), but another way turns out to be more useful: Rewrite the equation as

$$\frac{d}{dt}\big(\dot{x}+\varepsilon(x^3-x)\big)+x=0$$

Then write $u=\dot{x}+\varepsilon(x^3-x)$ and rewrite as

$$\begin{pmatrix}\dot{x}\\ \dot{u}\end{pmatrix}=\begin{pmatrix}\varepsilon(x-x^3)+u\\ -x\end{pmatrix} \tag{2}$$

Note that $\dot{u}<0$ in the right half plane and $\dot{u}>0$ in the left half plane, while $\dot{x}>0$ above the curve $u=\varepsilon(x^3-x)$ and $\dot{x}<0$ below it. The net effect is that the flow moves in a generally clockwise direction.

It seems reasonable to expect that sometimes, the x^3 term will dominate in the first equation, so that x behaves like a solution to $\dot{x}=-\varepsilon x^3$. That is a separable equation, with the solution $x=\pm 1/\sqrt{2\varepsilon(t-t_0)}$, where t_0 is an integration constant. We note that this solution goes to zero as $t\to\infty$, which violates the assumption that the x^3 term dominates – so it is not a good approximation forward in time. Backward in time, however, the solution goes to $\pm\infty$ in finite time (it blows up at $t=t_0$), so this looks more plausible. If we solve $\dot{u}=-x$ given this supposed approximation to x, we find $u=c\mp\sqrt{2(t-t_0)/\varepsilon}$, which stays bounded as $t\to t_0$. All of this makes it very plausible that we can find a solution that, for example, stays in the second quadrant but where x escapes to $-\infty$ when we move back in time, like the upper curve in figure 5.

As we move *forward* along that same curve, we must hit the u axis, since x is negative and increasing, and $\dot{u}=-x$ shows that u itself cannot go to infinity because x is bounded. The curve crosses the u axis horizontally ($\dot{u}=0$ there) and then starts moving down ($\dot{u}=-x<0$). After crossing the curve $u=\varepsilon(x^3-x)$ vertically ($\dot{x}=0$), it must hit the x axis at some point. Let us stop following it at that point.

Next, draw the mirror image (reflected through the origin, (x,u) replaced by $(-x,-u)$) of the curve. It is a solution too, since the dynamical system (2) is invariant under this transformation.

Finally, draw two vertical lines as in figure 5. The two solution curves and the vertical lines bound a region from which the solution can never escape as we move forward in time, since the flow points into the region along the vertical lines, and one solution curve cannot cross another.

But also, the origin is an unstable focus (or unstable node ε is large enough), so we can draw a small circle around the origin, and the region outside this circle will also be forward invariant. This is quite easy to prove directly, since

a simple calculation shows

$$\frac{d}{dt}\big(\tfrac{1}{2}(x^2+u^2)\big)=x\dot{x}+u\dot{u}=\varepsilon x^2(1-x^2)$$

which is positive when $|x|<1$, so the exterior of any circle with radius less than 1 is forward invariant.

Since there are no equilibrium points in the intersection of the two regions (i.e., inside the outer, complicated curve and outside the inner circle), the Poincaré–Bendixsson theorem guarantees the existence of a cycle in the region.

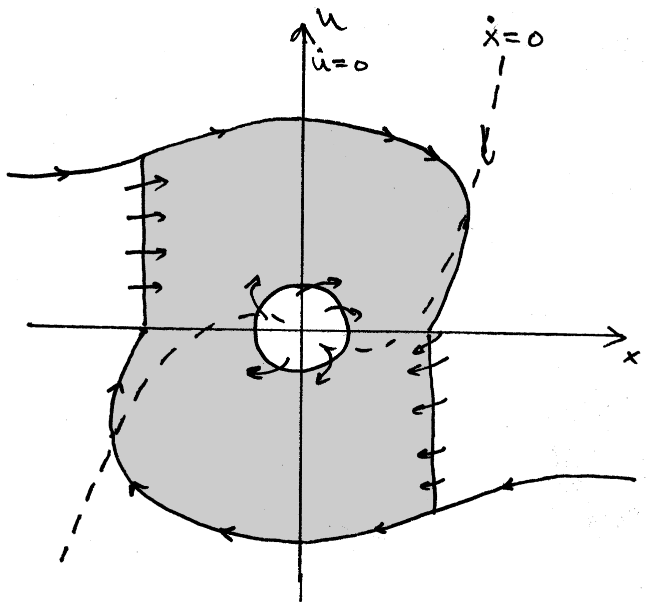

Figure 5: Existence of a cycle for the van der Pol equation