Contents

List of Figures

Chapter 1

Introduction to Dynamics

1.1 Introduction

1.1.1 Phase space and phase curves

Dynamics is the study of motion through phase space. The phase space of a given dynamical system is described as an N-dimensional manifold, $\mathcal{M}$. A (differentiable) manifold $\mathcal{M}$ is a topological space that is locally diffeomorphic to $\mathbb{R}^N$.[1] Typically in this course $\mathcal{M}$ will $\mathbb{R}^N$ itself, but other common examples include the circle $\mathbb{S}^1$, the torus $\mathbb{T}^2$, the sphere $\mathbb{S}^2$, *etc.*

Let $g_\tau : \mathcal{M} \to \mathcal{M}$ be a one-parameter family of transformations from $\mathcal{M}$ to itself, with $g_{\tau=0} = 1$, the identity, and $\tau \in \mathbb{R}$. We call g_τ the τ-*advance mapping*. It satisfies the composition rule

$$g_\tau \, g_\sigma = g_{\tau+\sigma} \quad . \tag{1.1}$$

Let us choose an arbitrary point $\varphi_0 = \varphi(t_0) \in \mathcal{M}$ to be the phase space coordinate at an initial time t_0. We then write $\varphi(t_0 + \tau) \equiv g_\tau \, \varphi(t_0)$, which also is in $\mathcal{M}$. The set

$$\{ g_\tau \, \varphi_0 \, | \, \tau \in \mathbb{R} \, , \, \varphi_0 \in \mathcal{M} \} \tag{1.2}$$

is called a *phase curve*. A graph of the motion $\varphi(t)$ in the product space $\mathcal{M} \times \mathbb{R}$ is called an *integral curve*. In a more general setting, we could define the two-parameter dynamical map $G_{t_0}^{t_0+\tau}$, which evolves the phase space coordinate from $\varphi(t_0)$ to $\varphi(t_0 + \tau)$ to depend on t_0 as well. Thus

$$G_{t_0}^{t_0+\tau} \varphi(t_0) = \varphi(t_0 + \tau) \quad . \tag{1.3}$$

But then $G_t^t = 1$ must be the identity for all t, and $G_t^{t'}$ must satisfy a composition rule,

$$G_{t_0}^{t_2} = G_{t_1}^{t_2} \, G_{t_0}^{t_1} \tag{1.4}$$

---

[1] A *diffeomorphism* $F \colon \mathcal{M} \to \mathcal{N}$ is a differentiable map with a differentiable inverse. This is a special type of *homeomorphism*, which is a continuous map with a continuous inverse.
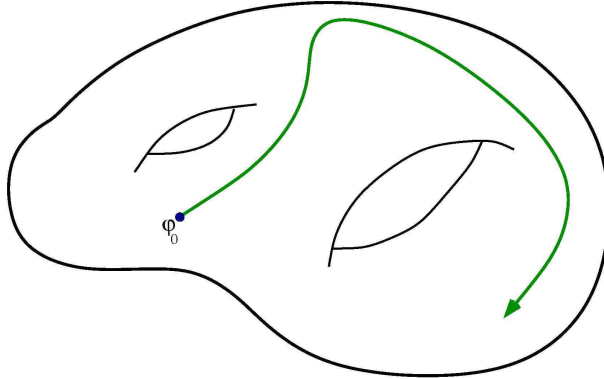
Figure 1.1: An example of a phase curve.

for any time t_1. This says that to evolve the phase space coordinate φ between times t_0 and t_2, we can first evolve from t_0 to t_1 and then from t_1 to t_2, for any t_1. By extending the phase space $\mathcal{M}$ to $\mathcal{M}' \equiv \mathcal{M} \times \mathbb{R}$, where $\dim(\mathcal{M}') = N + 1$, then defining the $(N + 1)$-dimensional phase space vector $\psi^{\mathsf{T}} = (\varphi_1, \ldots, \varphi_N, t)$, we define the one-parameter map on $\mathcal{M}'$, $\tilde{g}_\tau$, by

$$
\tilde{g}_\tau \begin{pmatrix} \varphi \\ t \end{pmatrix} = \begin{pmatrix} G_t^{t+\tau} \varphi \\ t + \tau \end{pmatrix} \quad . \tag{1.5}
$$

Unless otherwise stated, and without loss of generality, we will always consider our dynamical systems to be governed by a one-parameter τ-advance map g_τ.

1.1.2 Vector fields

The *velocity* vector $\boldsymbol{V}(\varphi)$ is given by the derivative

$$
\boldsymbol{V}(\varphi) = \frac{d}{dt}\bigg|_{t=0} g_t\, \varphi \quad . \tag{1.6}
$$

The velocity $\boldsymbol{V}(\varphi)$ is an element of the *tangent space* to $\mathcal{M}$ at φ, abbreviated $\mathsf{T}\mathcal{M}_\varphi$. If $\mathcal{M}$ is N-dimensional, then so is each $\mathsf{T}\mathcal{M}_\varphi$ (for all φ). However, $\mathcal{M}$ and $\mathsf{T}\mathcal{M}_\varphi$ may differ topologically. For example, if $\mathcal{M} = \mathbb{S}^1$, the circle, the tangent space at any point is isomorphic to $\mathbb{R}$.

For our purposes, we will take $\varphi = (\varphi_1, \ldots, \varphi_N)$ to be an N-tuple, *i.e.* a point in $\mathbb{R}^N$. The equation of motion is then

$$
\frac{d}{dt}\, \varphi(t) = \boldsymbol{V}\big(\varphi(t)\big) \quad . \tag{1.7}
$$

Note that any N^{th} order ODE, of the general form

$$
\frac{d^N x}{dt^N} = F\left(x, \frac{dx}{dt}, \ldots, \frac{d^{N-1}x}{dt^{N-1}}\right) \quad , \tag{1.8}
$$

may be represented by the first order system $\dot{\boldsymbol{\varphi}} = \boldsymbol{V}(\boldsymbol{\varphi})$. To see this, define $\varphi_k = d^{k-1}x/dt^{k-1}$, with $k = 1, \ldots, N$. Thus, for $j < N$ we have $\dot{\varphi}_j = \varphi_{j+1}$, and $\dot{\varphi}_N = f$. In other words,

$$\frac{d}{dt} \overbrace{\begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_{N-1} \\ \varphi_N \end{pmatrix}}^{\dot{\boldsymbol{\varphi}}} = \overbrace{\begin{pmatrix} \varphi_2 \\ \vdots \\ \varphi_N \\ F(\varphi_1, \ldots, \varphi_N) \end{pmatrix}}^{\boldsymbol{V}(\boldsymbol{\varphi})} \quad . \tag{1.9}$$

1.1.3 Existence / uniqueness / extension theorems

Theorem : Given $\dot{\boldsymbol{\varphi}} = \boldsymbol{V}(\boldsymbol{\varphi})$ and $\boldsymbol{\varphi}(0)$, if each $\boldsymbol{V}(\boldsymbol{\varphi})$ is a smooth vector field over some open set $\mathcal{D} \in \mathcal{M}$, then for $\boldsymbol{\varphi}(0) \in \mathcal{D}$ the initial value problem has a solution on some finite time interval $(-\tau, +\tau)$ and the solution is unique. Furthermore, the solution has a unique extension forward or backward in time, either indefinitely or until $\boldsymbol{\varphi}(t)$ reaches the boundary of $\mathcal{D}$.

Corollary : *Different trajectories never intersect!*

More generally, we might ask the following: *under what conditions does the dynamical system $\dot{\boldsymbol{\varphi}} = \boldsymbol{\mathcal{V}}(\boldsymbol{\varphi}, t)$ with initial conditions $\boldsymbol{\varphi}(0) = \boldsymbol{\varphi}_0$ have a unique solution?*. This is specified by the *Picard-Lindelöf theorem*[2], which says that if $D \subseteq \mathbb{R}^N \times \mathbb{R}$, is a closed rectangle and $\boldsymbol{\mathcal{V}} : D \to \mathbb{R}^N$ is a function which is continuous in t and *Lipschitz continuous* in $\boldsymbol{\varphi}$, then there exists some $\varepsilon > 0$ such that the initial value problem $\dot{\boldsymbol{\varphi}}(t) = \boldsymbol{\mathcal{V}}(\boldsymbol{\varphi}(t), t)$ with $(\boldsymbol{\varphi}_0, t_0) \in D$, has a unique solution $\boldsymbol{\varphi}(t)$ on the interval $t \in [t_0 - \varepsilon, t_0 + \varepsilon]$. In this case the solution is

$$\boldsymbol{\varphi}(t) = \boldsymbol{\varphi}(t_0) + \int_{t_0}^{t} ds \, \boldsymbol{\mathcal{V}}(\boldsymbol{\varphi}(s), s) \quad . \tag{1.10}$$

In general, Lipschitz continuity[3] is a condition on a function $f : \mathcal{X} \to \mathcal{Y}$ between two metric spaces $(\mathcal{X}, d_{\mathcal{X}})$ and $(\mathcal{Y}, d_{\mathcal{Y}})$, where $d_{\mathcal{X}}(X_1, X_2)$ and $d_{\mathcal{Y}}(Y_1, Y_2)$ are distance functions on $\mathcal{X}$ and $\mathcal{Y}$, respectively, with $X_{1,2} \in \mathcal{X}$ and $Y_{1,2} \in \mathcal{Y}$. The function f is then Lipschitz continuous if there exists a real constant $K \geq 0$ such that

$$d_{\mathcal{Y}}(f(X_1), f(X_2)) \leq K \, d_{\mathcal{X}}(X_1, X_2) \quad . \tag{1.11}$$

For our application, this means $\exists \, K \geq 0$ such that

$$\left| \boldsymbol{\mathcal{V}}(\boldsymbol{\varphi}_1, t) - \boldsymbol{\mathcal{V}}(\boldsymbol{\varphi}_2, t) \right| \leq K \left| \boldsymbol{\varphi}_1 - \boldsymbol{\varphi}_2 \right| \tag{1.12}$$

for all $t \in [t_0 - \varepsilon, t_0 + \varepsilon]$, where

$$|\boldsymbol{\varphi}' - \boldsymbol{\varphi}| = \sqrt{\sum_{n=1}^{N} (\varphi'_n - \varphi_n)^2} \tag{1.13}$$

is the Euclidean distance in phase space. An example where this fails is given in §1.2.3 below.

---

[2]See Wikipedia: https://en.wikipedia.org/wiki/Picard?Lindelf_theorem.
[3]See https://en.wikipedia.org/wiki/Lipschitz_continuity.

1.1.4 Linear differential equations

A homogeneous linear N^{th} order ODE,

$$\frac{d^N x}{dt^N} + c_{N-1} \frac{d^{N-1} x}{dt^{N-1}} + \ldots + c_1 \frac{dx}{dt} + c_0\, x = 0 \tag{1.14}$$

may be written in matrix form, as

$$\frac{d}{dt} \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_N \end{pmatrix} = \overbrace{\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -c_0 & -c_1 & -c_2 & \cdots & -c_{N-1} \end{pmatrix}}^{M} \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_N \end{pmatrix} \quad . \tag{1.15}$$

Thus,

$$\dot{\boldsymbol{\varphi}} = M \boldsymbol{\varphi} \quad , \tag{1.16}$$

and if the coefficients c_k are time-independent, *i.e.* the ODE is *autonomous*, the solution is obtained by exponentiating the constant matrix Q:

$$\boldsymbol{\varphi}(t) = \exp(Mt)\, \boldsymbol{\varphi}(0) ; \tag{1.17}$$

the exponential of a matrix may be given meaning by its Taylor series expansion. If the ODE is not autonomous, then $M = M(t)$ is time-dependent, and the solution is given by the path-ordered exponential,

$$\boldsymbol{\varphi}(t) = \mathcal{P} \exp \left\{ \int_0^t dt'\, M(t') \right\} \boldsymbol{\varphi}(0) \quad , \tag{1.18}$$

As defined, the equation $\dot{\boldsymbol{\varphi}} = \boldsymbol{V}(\boldsymbol{\varphi})$ is autonomous, since g_t depends only on t and on no other time variable. However, by extending the phase space from $\mathcal{M}$ to $\mathbb{R} \times \mathcal{M}$, which is of dimension $(N+1)$, one can describe arbitrary time-dependent ODEs.

1.1.5 Lyapunov functions

For a general dynamical system $\dot{\boldsymbol{\varphi}} = \boldsymbol{V}(\boldsymbol{\varphi})$, a *Lyapunov function* $L(\boldsymbol{\varphi})$ is a function which satisfies

$$\boldsymbol{\nabla} L(\boldsymbol{\varphi}) \cdot \boldsymbol{V}(\boldsymbol{\varphi}) \leq 0 \quad . \tag{1.19}$$

There is no simple way to determine whether a Lyapunov function exists for a given dynamical system, or, if it does exist, what the Lyapunov function is. However, if a Lyapunov function can be found, then this severely limits the possible behavior of the system. This is because $L\big(\boldsymbol{\varphi}(t)\big)$ must be a monotonic function of time:

$$\frac{d}{dt}\, L\big(\boldsymbol{\varphi}(t)\big) = \boldsymbol{\nabla} L \cdot \frac{d\boldsymbol{\varphi}}{dt} = \boldsymbol{\nabla} L(\boldsymbol{\varphi}) \cdot \boldsymbol{V}(\boldsymbol{\varphi}) \leq 0 \quad . \tag{1.20}$$

Thus, the system evolves toward a local minimum of the Lyapunov function. In general this means that oscillations are impossible in systems for which a Lyapunov function exists. For example, the relaxational dynamics of the magnetization M of a system are sometimes modeled by the equation

$$\frac{dM}{dt} = -\Gamma \frac{\partial F}{\partial M} \quad , \tag{1.21}$$

where $F(M, T)$ is the *free energy* of the system. In this model, assuming constant temperature T, $\dot{F} = F'(M) \dot{M} = -\Gamma \left[F'(M) \right]^2 \leq 0$. So the free energy $F(M)$ itself is a Lyapunov function, and it monotonically decreases during the evolution of the system. We shall meet up with this example again in the next chapter when we discuss imperfect bifurcations.

1.2 $N = 1$ **Systems**

We now study phase flows in a one-dimensional phase space, governed by the equation

$$\frac{du}{dt} = f(u) \quad . \tag{1.22}$$

Again, the equation $\dot{u} = h(u, t)$ is first order, but not autonomous, and it corresponds to the $N = 2$ system,

$$\frac{d}{dt} \begin{pmatrix} u \\ t \end{pmatrix} = \begin{pmatrix} h(u, t) \\ 1 \end{pmatrix} \quad . \tag{1.23}$$

The equation 1.22 is easily integrated:

$$\frac{du}{f(u)} = dt \qquad \Longrightarrow \qquad t - t_0 = \int_{u_0}^{u} \frac{du'}{f(u')} \quad . \tag{1.24}$$

This gives $t(u)$; we must then invert this relationship to obtain $u(t)$.

Example : Suppose $f(u) = a - bu$, with a and b constant. Then

$$dt = \frac{du}{a - bu} = -b^{-1} \, d\ln(a - bu) \tag{1.25}$$

whence

$$t = \frac{1}{b} \ln \left(\frac{a - b\,u(0)}{a - b\,u(t)} \right) \qquad \Longrightarrow \qquad u(t) = \frac{a}{b} + \left(u(0) - \frac{a}{b} \right) \exp(-bt) \quad . \tag{1.26}$$

Even if one cannot analytically obtain $u(t)$, the behavior is very simple, and easily obtained by graphical analysis. Sketch the function $f(u)$. Then note that

$$\dot{u} = f(u) \qquad \Longrightarrow \qquad \begin{cases} f(u) > 0 \quad \dot{u} > 0 \quad \Rightarrow \quad \text{move to right} \\ f(u) < 0 \quad \dot{u} < 0 \quad \Rightarrow \quad \text{move to left} \\ f(u) = 0 \quad \dot{u} = 0 \quad \Rightarrow \quad \text{fixed point} \end{cases} \tag{1.27}$$
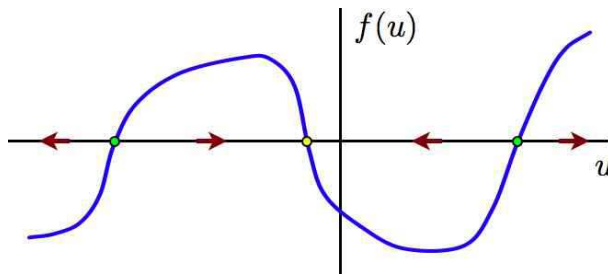
Figure 1.2: Phase flow for an $N = 1$ system.

The behavior of $N = 1$ systems is particularly simple: $u(t)$ flows to the first stable fixed point encountered, where it then (after a logarithmically infinite time) stops. The motion is monotonic – the velocity $\dot{u}$ never changes sign. Thus, *oscillations never occur for $N = 1$ phase flows.*[4]

1.2.1 Classification of fixed points ($N = 1$)

A *fixed point* u^* satisfies $f(u^*) = 0$. Generically, $f'(u^*) \neq 0$ at a fixed point.[5] Suppose $f'(u^*) < 0$. Then to the left of the fixed point, the function $f(u < u^*)$ is positive, and the flow is to the right, *i.e.* toward u^*. To the right of the fixed point, the function $f(u > u^*)$ is negative, and the flow is to the left, *i.e.* again toward u^*. Thus, when $f'(u^*) < 0$ the fixed point is said to be *stable*, since the flow in the vicinity of u^* is to u^*. Conversely, when $f'(u^*) > 0$, the flow is always away from u^*, and the fixed point is then said to be *unstable*. Indeed, if we linearize about the fixed point, and let $\epsilon \equiv u - u^*$, then

$$\dot{\epsilon} = f'(u^*)\,\epsilon + \tfrac{1}{2}\,f''(u^*)\,\epsilon^2 + \mathcal{O}(\epsilon^3) \quad , \tag{1.28}$$

and dropping all terms past the first on the RHS gives

$$\epsilon(t) = \exp\left[f'(u^*)\,t\right]\epsilon(0) \quad . \tag{1.29}$$

The deviation decreases exponentially for $f'(u^*) < 0$ and increases exponentially for $f(u^*) > 0$. Note that

$$t(\epsilon) = \frac{1}{f'(u^*)}\,\ln\left(\frac{\epsilon}{\epsilon(0)}\right) \quad , \tag{1.30}$$

so the approach to a stable fixed point takes a logarithmically infinite time. For the unstable case, the deviation grows exponentially, until eventually the linearization itself fails.

1.2.2 Logistic equation

This model for population growth was first proposed by Verhulst in 1838. Let N denote the population in question. The dynamics are modeled by the first order ODE,

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right) \quad , \tag{1.31}$$

---

[4]When I say 'never' I mean 'sometimes' – see the section 1.3.
[5]The system $f(u^*) = 0$ and $f'(u^*) = 0$ is overdetermined, with two equations for the single variable u^*.
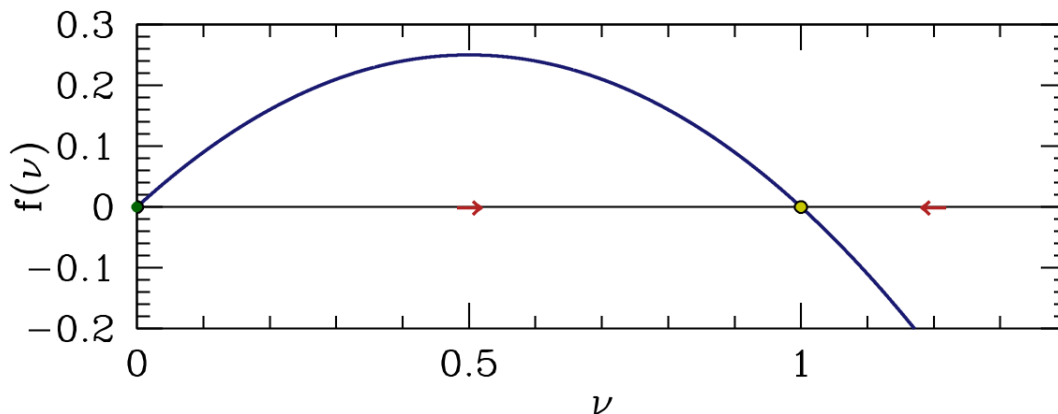
Figure 1.3: Flow diagram for the logistic equation.

where N, r, and K are all positive. For $N \ll K$ the growth rate is r, but as N increases a quadratic nonlinearity kicks in and the rate vanishes for $N = K$ and is negative for $N > K$. The nonlinearity models the effects of competition between the organisms for food, shelter, or other resources. Or maybe they crap all over each other and get sick. Whatever.

There are two fixed points, one at $N^* = 0$, which is unstable ($f'(0) = r > 0$). The other, at $N^* = K$, is stable ($f'(K) = -r$). The equation is adimensionalized by defining $\nu = N/K$ and $s = rt$, whence

$$\dot{\nu} = \nu(1 - \nu) \quad . \tag{1.32}$$

Integrating,

$$\frac{d\nu}{\nu(1 - \nu)} = d\ln\left(\frac{\nu}{1 - \nu}\right) = ds \quad \Longrightarrow \quad \nu(s) = \frac{\nu_0}{\nu_0 + (1 - \nu_0)\exp(-s)} \quad . \tag{1.33}$$

As $s \to \infty$, $\nu(s) = 1 - \left(\nu_0^{-1} - 1\right) e^{-s} + \mathcal{O}(e^{-2s})$, and the relaxation to equilibrium ($\nu^* = 1$) is exponential, as usual.

Another application of this model is to a simple autocatalytic reaction, such as

$$A + X \rightleftharpoons 2X \quad , \tag{1.34}$$

i.e. X catalyses the reaction $A \longrightarrow X$. Assuming a fixed concentration of A, we have

$$\dot{x} = \kappa_+ \, a \, x - \kappa_- \, x^2 \quad , \tag{1.35}$$

where x is the concentration of X, and $\kappa_\pm$ are the forward and backward reaction rates.

1.2.3 Singular $f(u)$

Suppose that in the vicinity of a fixed point we have $f(u) = A\left|u - u^*\right|^\alpha$, with $A, \alpha > 0$. We now analyze both sides of the fixed point.
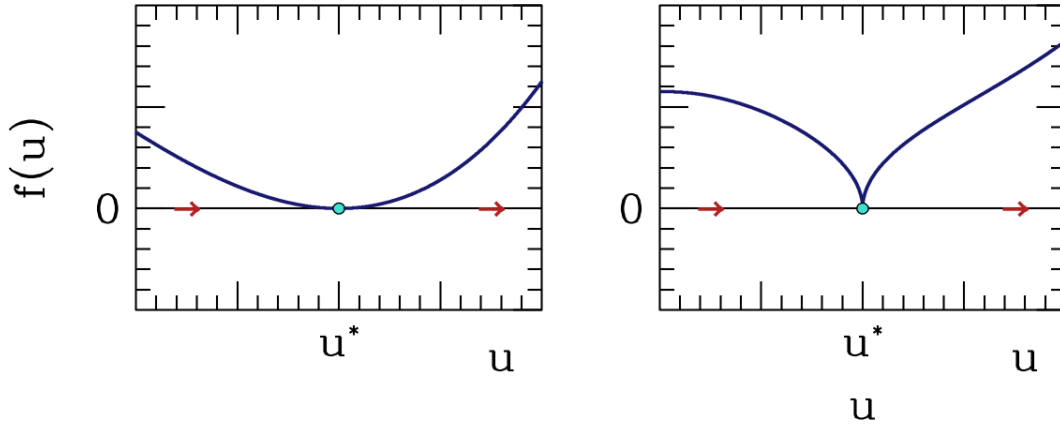
Figure 1.4: $f(u) = A\left|u - u^*\right|^\alpha$ for $\alpha > 1$ (left) and $0 < \alpha < 1$ (right).

$u < u^*$: Let $\epsilon = u^* - u$. Then

$$\dot\epsilon = -A\,\epsilon^\alpha \quad \Longrightarrow \quad \frac{\epsilon^{1-\alpha}}{1-\alpha} = \frac{\epsilon_0^{1-\alpha}}{1-\alpha} - At \quad , \tag{1.36}$$

hence

$$\epsilon(t) = \left[\epsilon_0^{1-\alpha} + (\alpha-1)A\,t\right]^{1/(1-\alpha)} \quad . \tag{1.37}$$

This, for $\alpha < 1$ the fixed point $\epsilon = 0$ is reached in a finite time: $\epsilon(t_{\rm c}) = 0$, with

$$t_{\rm c} = \frac{\epsilon_0^{1-\alpha}}{(1-\alpha)\,A} \quad . \tag{1.38}$$

For $\alpha > 1$, we have $\lim_{t\to\infty}\epsilon(t) = 0$, but $\epsilon(t) > 0$ for all finite t.

The fixed point $u = u^*$ is now *half-stable* – the flow from the left is toward u^* but from the right is away from u^*. Let's now analyze the flow on either side of u^*.

$u > u^*$: Let $\epsilon = u - u^*$. Then $\dot\epsilon = A\,\epsilon^\alpha$, and

$$\epsilon(t) = \left[\epsilon_0^{1-\alpha} + (1-\alpha)A\,t\right]^{1/(1-\alpha)} \quad . \tag{1.39}$$

For $\alpha < 1$, $\epsilon(t)$ escapes to $\epsilon = \infty$ only after an infinite time. For $\alpha > 1$, the escape to infinity takes a finite time: $\epsilon(t_{\rm c}) = \infty$, with

$$t_{\rm c} = \frac{\epsilon_0^{1-\alpha}}{(\alpha-1)\,A} \quad . \tag{1.40}$$

In both cases, higher order terms in the (nonanalytic) expansion of $f(u)$ about $u = u^*$ will eventually come into play.

The case $\alpha < 1$ provides a nice illustration of the Picard-Lindelöf theorem in §1.1.3, for consider the equation $dx/ds = -A|x|^\alpha$ with $\alpha \in (0,1)$ and initial condition $x(s=0) = x_0 > 0$. The solution is

$$x(s) = \left[x_0^{1-\alpha} - (1-\alpha)As\right]^{1/(1-\alpha)} \Theta(t_{\rm c} - s) \quad , \tag{1.41}$$
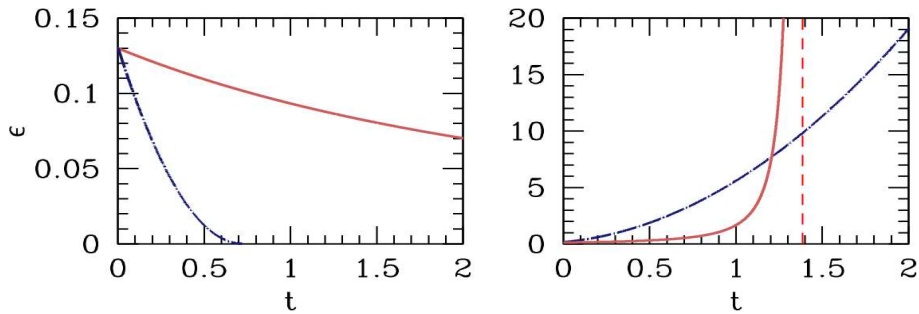
Figure 1.5: Solutions to $\dot{\epsilon} = \mp A \, \epsilon^\alpha$. Left panel: $\epsilon = u^* - u$, with $\alpha = 1.5$ (solid red) and $\alpha = 0.5$ (dot-dashed blue); $A = 1$ in both cases. Right panel: $\epsilon = u - u^*$, $\alpha = 1.5$ (solid red) and $\alpha = 0.5$ (dot-dashed blue); $A = 4$ in both cases.

where t_c is given above. Now let's run time backwards and define $t \equiv -s + t_\mathrm{c} + \tau$ with $\tau > 0$. The dynamical system is then $dx/dt = A \, |x|^\alpha$ and the solution is

$$x(t) = \left[x_0^{1-\alpha} - (1 - \alpha)A \, (t_\mathrm{c} + \tau - t) \right]^{1/(1-\alpha)} \Theta(t - \tau) \quad, \tag{1.42}$$

where $x(t_\mathrm{c} + \tau) = x_0$. This is a valid solution for *any* $\tau \geq 0$. Thus for any positive τ, we have a solution to the equation $\dot{x} = A \, |x|^\alpha$ with initial condition $x(t = 0) = 0$, and therefore *the solution to the initial value problem is not unique!* The difficulty can be traced to the fact that the vector field $V(x) = A|x|^\alpha$ is not Lipschitz continuous at $x = 0$, since for any $K > 0$ we may find an x such that

$$\left| V(x) - V(0) \right| = A \, |x|^\alpha > K|x| \quad, \tag{1.43}$$

which is the case for all $|x| < (A/K)^{1/(1-\alpha)}$.

1.2.4 Recommended exercises

It is constructive to sketch the phase flows for the following examples:

$$\dot{v} = -g \qquad\qquad \dot{u} = A \sin(u)$$
$$m\dot{v} = -mg - \gamma v \qquad\qquad \dot{u} = A \, (u - a)(u - b)(u - c)$$
$$m\dot{v} = -mg - cv^2 \, \mathrm{sgn}(v) \qquad\qquad \dot{u} = au^2 - bu^3 \quad.$$

In each case, identify all the fixed points and assess their stability. Assume all constants A, a, b, c, γ, etc. are positive.

1.2.5 Non-autonomous ODEs

Non-autonomous ODEs of the form $\dot{u} = h(u, t)$ are in general impossible to solve by quadratures. One can always go to the computer, but it is worth noting that in the *separable* case, $h(u, t) = f(u) \, g(t)$, one

can obtain the solution

$$\frac{du}{f(u)} = g(t)\,dt \quad \Longrightarrow \quad \int_{u_0}^{u} \frac{du'}{f(u')} = \int_{0}^{t} dt'\,g(t') \quad , \tag{1.44}$$

which implicitly gives $u(t)$. Note that $\dot{u}$ may now change sign, and $u(t)$ may even oscillate. For an explicit example, consider the equation

$$\dot{u} = A\,(u+1)\,\sin(\beta t) \quad , \tag{1.45}$$

the solution of which is

$$u(t) = -1 + (u_0 + 1)\,\exp\left\{\frac{A}{\beta}\left[1 - \cos(\beta t)\right]\right\} \quad . \tag{1.46}$$

In general, the non-autonomous case defies analytic solution. Many have been studied, such as the Riccati equation,

$$\frac{du}{dt} = P(t)u^2 + Q(t)u + R(t) \quad . \tag{1.47}$$

Riccati equations have the special and remarkable property that one can generate *all* solutions (*i.e.* with arbitrary boundary condition $u(0) = u_0$) from *any* given solution (*i.e.* with any boundary condition).

1.3   Flows on the Circle

We had remarked that oscillations are impossible for the equation $\dot{u} = f(u)$ because the flow is to the first stable fixed point encountered. If there are no stable fixed points, the flow is unbounded. However, suppose phase space itself is bounded, *e.g.* a circle $\mathbb{S}^1$ rather than the real line $\mathbb{R}$. Thus,

$$\dot{\theta} = f(\theta) \quad , \tag{1.48}$$

with $f(\theta + 2\pi) = f(\theta)$. Now if there are no fixed points, $\theta(t)$ endlessly winds around the circle, and in this sense we can have oscillations.

1.3.1   Nonuniform oscillator

A particularly common example is that of the nonuniform oscillator,

$$\dot{\theta} = \omega - \sin\theta \quad , \tag{1.49}$$

which has applications to electronics, biology, classical mechanics, and condensed matter physics. Note that the general equation $\dot{\theta} = \omega - A\sin\theta$ may be rescaled to the above form. A simple application is to the dynamics of a driven, overdamped pendulum. The equation of motion is

$$I\ddot{\theta} + b\,\dot{\theta} + I\omega_0^2\,\sin\theta = N \quad , \tag{1.50}$$
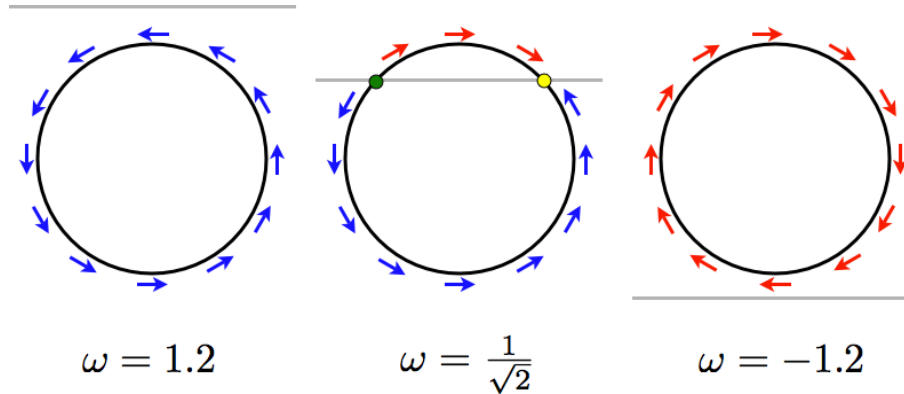
$$\omega = 1.2 \qquad\qquad \omega = \frac{1}{\sqrt{2}} \qquad\qquad \omega = -1.2$$

Figure 1.6: Flow for the nonuniform oscillator $\dot\theta = \omega - \sin\theta$ for three characteristic values of ω.

where I is the moment of inertia, b is the damping parameter, N is the external torque (presumed constant), and ω_0 is the frequency of small oscillations when $b = N = 0$. When b is large, the inertial term $I\ddot\theta$ may be neglected, and after rescaling we arrive at eqn. 1.49.

The book by Strogatz provides a biological example of the nonuniform oscillator: fireflies. An individual firefly will on its own flash at some frequency f. This can be modeled by the equation $\dot\phi = \beta$, where $\beta = 2\pi f$ is the angular frequency. A flash occurs when $\phi = 2\pi n$ for $n \in \mathbb{Z}$. When subjected to a periodic stimulus, fireflies will attempt to synchronize their flash to the flash of the stimulus. Suppose the stimulus is periodic with angular frequency Ω. The firefly synchronization is then modeled by the equation

$$\dot\phi = \beta - A\sin(\phi - \Omega t) \quad . \tag{1.51}$$

Here, A is a measure of the firefly's ability to modify its natural frequency in response to the stimulus. Note that when $0 < \phi - \Omega t < \pi$, *i.e.* when the firefly is leading the stimulus, the dynamics tell the firefly to slow down. Conversely, when $-\pi < \phi - \Omega t < 0$, the firefly is lagging the stimulus, the the dynamics tell it to speed up. Now focus on the difference $\theta \equiv \phi - \Omega t$. We have

$$\dot\theta = \beta - \Omega - A\sin\theta \quad , \tag{1.52}$$

which is the nonuniform oscillator. We can adimensionalize by defining

$$s \equiv At \quad , \quad \omega \equiv \frac{\beta - \Omega}{A} \quad , \tag{1.53}$$

yielding $\frac{d\theta}{ds} = f(\theta) = \omega - \sin\theta$.

Fixed points θ^* occur only for $|\omega| < 1$, at $\sin\theta^* = \omega$, in which case $f'(\theta) = -\cos\theta^*$. As we have seen above, stability requires $f'(\theta^*) < 0$, which means $\theta^* \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, *i.e.* θ^* must lie on the right half of the circle. For $|\omega| > 1$, the angular velocity never vanishes anywhere along the circle, and there are no fixed points. In this case the motion is eternally clockwise ($\omega < -1$) or counterclockwise ($\omega > +1$). The situation is depicted in Fig. 1.6.

To integrate, set $z = \exp(i\theta)$, in which case

$$\frac{dz}{ds} = -\tfrac{1}{2}(z^2 - 2i\omega z - 1) = -\tfrac{1}{2}(z - \xi_-)(z - \xi_+) \quad , \tag{1.54}$$

where $\xi_{\pm} = i\omega \pm \sqrt{1 - \omega^2} \equiv i\omega \pm \nu$ with $\nu \equiv \sqrt{1 - \omega^2}$. This yields

$$d \log\left(\frac{z - \xi_+}{z - \xi_-}\right) = \tfrac{1}{2}\left(\xi_- - \xi_+\right) ds = -\nu\, ds \tag{1.55}$$

which integrates to

$$z(s) = \frac{(\xi_+ - e^{-\nu s}\,\xi_-)\, z(0) + (e^{-\nu s} - 1)\,\xi_+\xi_-}{(1 - e^{-\nu s})\, z(0) + (\xi_+\, e^{-\nu s} - \xi_-)}\quad . \tag{1.56}$$

When $\omega^2 > 1$, ν is pure imaginary and $\exp(-\nu s)$ continually winds about the unit circle. When $\omega^2 < 1$, ν is real and positive. We then have that $z(s \to \infty) = \xi_+$ while $z(s \to -\infty) = \xi_-$. Note that $\xi_{\pm}$ lie on the appropriate halves of the circle as depicted in fig. 1.6.

For $|\omega| > 1$, the motion is periodic, with period

$$T = \int\limits_{0}^{2\pi} \frac{d\theta}{|\omega| - \sin\theta} = \frac{2\pi}{\sqrt{\omega^2 - 1}}\quad . \tag{1.57}$$

1.4   Appendix I : Evolution of Phase Space Volumes

Recall the general form of a dynamical system, $\dot{\boldsymbol{\varphi}} = \boldsymbol{V}(\boldsymbol{\varphi})$. Usually we are interested in finding integral curves $\boldsymbol{\varphi}(t)$. However, consider for the moment a collection of points in phase space comprising a region $\mathcal{R}$. As the dynamical system evolves, this region will also evolve, so that $\mathcal{R} = \mathcal{R}(t)$. We now ask: how does the volume of $\mathcal{R}(t)$,

$$\text{vol}\big[\mathcal{R}(t)\big] = \int\limits_{\mathcal{R}(t)} d\mu \quad , \tag{1.58}$$

where $d\mu = d\varphi_1 d\varphi_2 \cdots d\varphi_N$ is the phase space measure, change with time. We have, explicitly,

$$\text{vol}\big[\mathcal{R}(t + dt)\big] = \int\limits_{\mathcal{R}(t+dt)} d\mu = \int\limits_{\mathcal{R}(t)} d\mu \left\|\frac{\partial\varphi_i(t + dt)}{\partial\varphi_j(t)}\right\|$$
$$= \int\limits_{\mathcal{R}(t)} d\mu \left\{1 + \boldsymbol{\nabla}\cdot\boldsymbol{V}\, dt + \mathcal{O}\big((dt)^2\big)\right\}\quad , \tag{1.59}$$

since

$$\frac{\partial\varphi_i(t + dt)}{\partial\varphi_j(t)} = \delta_{ij} + \frac{\partial V_i}{\partial\varphi_j}\bigg|_{\boldsymbol{\varphi}(t)} dt + \mathcal{O}\big((dt)^2\big)\quad , \tag{1.60}$$

and, using $\ln \det M = \text{Tr}\ln M$,

$$\det(1 + \epsilon A) = 1 + \epsilon\,\text{Tr}A + \mathcal{O}(\epsilon^2)\quad . \tag{1.61}$$

Thus,

$$\frac{d}{dt}\,\text{vol}\big[\mathcal{R}(t)\big] = \int\limits_{\mathcal{R}(t)} d\mu\,\boldsymbol{\nabla}\cdot\boldsymbol{V} = \int\limits_{\partial\mathcal{R}(t)} d\Sigma\,\hat{\boldsymbol{n}}\cdot\boldsymbol{V}\quad , \tag{1.62}$$

where in the last line we have used Stokes' theorem to convert the volume integral over $\mathcal{R}$ to a surface integral over its boundary $\partial\mathcal{R}$.

1.5 Appendix II : Lyapunov Characteristic Exponents

Suppose $\varphi(t)$ is an integral curve – *i.e.* a solution of $\dot{\varphi} = \boldsymbol{V}(\varphi)$. We now ask: how do nearby trajectories behave? Do they always remain close to $\varphi(t)$ for all t? To answer this, we write $\widetilde{\varphi}(t) \equiv \varphi(t) + \boldsymbol{\eta}(t)$, in which case

$$\frac{d}{dt}\,\eta_i(t) = M_{ij}(t)\,\eta_j(t) + \mathcal{O}(\eta^2) \quad , \tag{1.63}$$

where

$$M_{ij}(t) = \frac{\partial V_i}{\partial \varphi_j}\bigg|_{\varphi(t)} \quad . \tag{1.64}$$

The solution, valid to first order in $\delta\varphi$, is

$$\eta_i(t) = Q_{ij}(t, t_0)\,\eta_j(t_0) \quad , \tag{1.65}$$

where the matrix $Q(t, t_0)$ is given by the *path ordered exponential*,

$$Q(t, t_0) = \mathcal{P}\exp\left\{\int\limits_{t_0}^{t} dt'\, M(t')\right\}$$

$$\equiv \lim_{N\to\infty}\left(1 + \frac{\Delta t}{N}\,M(t_{N-1})\right)\cdots\left(1 + \frac{\Delta t}{N}\,M(t_1)\right)\left(1 + \frac{\Delta t}{N}\,M(t_0)\right) \quad , \tag{1.66}$$

with $\Delta t = t - t_0$ and $t_j = t_0 + (j/N)\Delta t$. $\mathcal{P}$ is the *path ordering operator*, which places earlier times to the right:

$$\mathcal{P}\,A(t)\,B(t') = \begin{cases} A(t)\,B(t') & \text{if } t > t' \\ B(t')\,A(t) & \text{if } t < t' \end{cases} \quad . \tag{1.67}$$

The distinction is important if $\big[A(t), B(t')\big] \neq 0$. Note that Q satisfies the composition property,

$$Q(t, t_0) = Q(t, t_1)\,Q(t_1, t_0) \tag{1.68}$$

for any $t_1 \in [t_0, t]$. When M is time-independent, as in the case of a *fixed point* where $\boldsymbol{V}(\varphi^*) = 0$, the path ordered exponential reduces to the ordinary exponential, and $Q(t, t_0) = \exp\big(M(t - t_0)\big)$.

Generally it is impossible to analytically compute path-ordered exponentials. However, the following example may be instructive. Suppose

$$M(t) = \begin{cases} M_1 & \text{if } t/T \in \big[2j, 2j+1\big] \\ \\ M_2 & \text{if } t/T \in \big[2j+1, 2j+2\big] \end{cases} \quad , \tag{1.69}$$

for all integer j. $M(t)$ is a 'matrix-valued square wave', with period $2T$. Then, integrating over one period, from $t = 0$ to $t = 2T$, we have

$$A \equiv \exp\left\{\int\limits_{0}^{2T} dt\, M(t)\right\} = e^{(M_1 + M_2)T}$$

$$A_{\mathcal{P}} \equiv \mathcal{P}\exp\left\{\int\limits_{0}^{2T} dt\, M(t)\right\} = e^{M_2 T}\,e^{M_1 T} \quad . \tag{1.70}$$

In general, $A \neq A_\mathcal{P}$, so the path ordering has a nontrivial effect[6].

The Lyapunov exponents are defined in the following manner. Let $\hat{e}$ be an N-dimensional unit vector. Define

$$\Lambda(\boldsymbol{\varphi}_0, \hat{e}) \equiv \lim_{t \to \infty} \lim_{b \to 0} \frac{1}{t - t_0} \ln \left(\frac{\|\boldsymbol{\eta}(t)\|}{\|\boldsymbol{\eta}(t_0)\|} \right)_{\boldsymbol{\eta}(t_0) = b \hat{e}} \quad , \tag{1.71}$$

where $\| \cdot \|$ denotes the Euclidean norm of a vector, and where $\boldsymbol{\varphi}_0 = \boldsymbol{\varphi}(t_0)$. A theorem due to Oseledec guarantees that there are N such values $\Lambda_i(\boldsymbol{\varphi}_0)$, depending on the choice of $\hat{e}$, for a given $\boldsymbol{\varphi}_0$. Specifically, the theorem guarantees that the matrix

$$W \equiv \left(Q^\mathsf{T} Q \right)^{1/(t - t_0)} \tag{1.72}$$

converges in the limit $t \to \infty$ for almost all $\boldsymbol{\varphi}_0$. The eigenvalues Λ_i correspond to the different eigenspaces of W. Oseledec's theorem (also called the 'multiplicative ergodic theorem') guarantees that the eigenspaces of W either grow ($\Lambda_i > 1$) or shrink ($\Lambda_i < 1$) *exponentially* fast. That is, the norm any vector lying in the ith eigenspace of W will behave as $\Lambda_i^t = \exp(t \ln \Lambda_i)$ as $t \to \infty$.

Note that while $W = W^\mathsf{T}$ is symmetric by construction, Q is simply a general real-valued $N \times N$ matrix. The left and right eigenvectors of a matrix $M \in \mathsf{GL}(N, \mathbb{R})$ will in general be different. The set of eigenvalues λ_α is, however, common to both sets of eigenvectors. Let $\{\psi_\alpha\}$ be the right eigenvectors and $\{\chi_\alpha^*\}$ the left eigenvectors, such that

$$\begin{aligned} M_{ij} \, \psi_{\alpha,j} &= \lambda_\alpha \, \psi_{\alpha,i} \\ \chi_{\alpha,i}^* \, M_{ij} &= \lambda_\alpha \, \chi_{\alpha,j}^* \quad . \end{aligned} \tag{1.73}$$

We can always choose the left and right eigenvectors to be orthonormal, *viz.*

$$\left\langle \chi_\alpha \,\middle|\, \psi_\beta \right\rangle = \chi_{\alpha,i}^* \, \psi_{\beta,j} = \delta_{\alpha\beta} \quad . \tag{1.74}$$

Indeed, we can define the matrix $S_{i\alpha} = \psi_{\alpha,i}$, in which case $S_{\alpha j}^{-1} = \chi_{\alpha,j}^*$, and

$$S^{-1} \, M \, S = \mathrm{diag}\left(\lambda_1, \ldots, \lambda_N \right) \quad . \tag{1.75}$$

The matrix M can always be decomposed into its eigenvectors, as

$$M_{ij} = \sum_\alpha \lambda_\alpha \, \psi_{\alpha,i} \, \chi_{\alpha,j}^* \quad . \tag{1.76}$$

If we expand $\boldsymbol{u}$ in terms of the right eigenvectors,

$$\boldsymbol{\eta}(t) = \sum_\beta C_\beta(t) \, \boldsymbol{\psi}_\beta(t) \quad , \tag{1.77}$$

then upon taking the inner product with χ_α, we find that C_α obeys

$$\dot{C}_\alpha + \left\langle \chi_\alpha \,\middle|\, \dot{\psi}_\beta \right\rangle C_\beta = \lambda_\alpha \, C_\alpha \quad . \tag{1.78}$$

---

[6]If $\left[M_1, M_2 \right] = 0$ then $A = A_\mathcal{P}$.

If $\dot{\psi}_\beta = 0$, *e.g.* if M is time-independent, then $C_\alpha(t) = C_\alpha(0)\, e^{\lambda_\alpha t}$, and

$$
\eta_i(t) = \sum_\alpha \sum_j \overbrace{\eta_j(0)\, \chi^*_{\alpha,j}}^{C_\alpha(0)}\, e^{\lambda_\alpha t}\, \psi_{\alpha,i} \quad . \tag{1.79}
$$

Thus, the component of $\boldsymbol{\eta}(t)$ along $\boldsymbol{\psi}_\alpha$ increases exponentially with time if $\mathrm{Re}(\lambda_\alpha) > 0$, and decreases exponentially if $\mathrm{Re}(\lambda_\alpha) < 0$.

Nota bene: If $M \in \mathsf{GL}(N, \mathbb{R})$ is not symmetric – more generally if it does not commute with its transpose M^{T} – then it may be that the right eigenvectors of M do not span $\mathbb{R}^N$. In this case, the canonical decomposition of M contains one or more *Jordan blocks*. See §1.6 below.

1.6   Appendix III : Normal Matrices, Non-Normal Matrices, and Jordan Blocks

Normal matrices and eigenspectra : Quantum mechanical Hamiltonians can be represented as Hermitian matrices. In elementary school linear algebra class, we all learned that any Hermitian matrix H is diagonalizable by a unitary transformation, its eigenvalues are real, and eigenvectors corresponding to different eigenvalues are necessarily orthogonal. In the case of degenerate eigenvalues, their associated eigenvectors may be chosen to be mutually orthogonal via the Gram-Schmidt process. In the following discussion, we will assume our matrices are in general complex, but we can of course restrict to the real case, as is appropriate for real linear dynamical systems.

Any complex square matrix A which satisfies $A^\dagger A = A A^\dagger$ is called *normal*. Hermitian matrices are normal, but so are antihermitian and unitary matrices[7]. Real symmetric, antisymmetric, and orthogonal matrices satisfy $A^{\mathsf{T}}A = A A^{\mathsf{T}}$. The *Schur decomposition theorem* guarantees that any $n \times n$ matrix A may be decomposed as $A = V T V^\dagger$, where $V \in \mathsf{U}(n)$ and T is upper triangular. Now if A is normal, $[A, A^\dagger] = V\,[T, T^\dagger]\, V^\dagger = 0$, hence T is normal. However, it is easy to show that any normal upper triangular matrix must be diagonal[8], so $A = V D V^\dagger$, which means $D = V^\dagger A V$ is the diagonal matrix of eigenvalues of A. Conversely, if $A = V D V^\dagger$ is unitarily equivalent to a diagonal matrix, it is trivial to show that A is normal. Thus *any $n \times n$ matrix A is diagonalizable by a unitary transformation if and only if A is normal.*

There is a real version of Schur decomposition whereby a real matrix B satisfying $B^{\mathsf{T}}B = B B^{\mathsf{T}}$ may be decomposed as $B = R S R^{\mathsf{T}}$, where R is a real orthogonal matrix, and S is *block upper triangular*. The diagonal blocks of S are either 1×1, corresponding to real eigenvalues, or 2×2, corresponding to complex eigenvalues. One eventually concludes that real symmetric matrices have real eigenvalues, real antisymmetric matrices have pure imaginary (or zero) eigenvalues, and real orthogonal matrices have unimodular complex eigenvalues.

---

[7]There are many examples of normal matrices which are neither Hermitian, antihermitian, nor unitary. For example, any diagonal matrix with arbitrary complex diagonal entries is normal.

[8]$T^\dagger T = T T^\dagger$ says that $\sum_j |T_{ij}|^2 = \sum_j |T_{ji}|^2$, *i.e.* the sum of the square moduli of the elements in the ith row is the same as that for the ith column. Starting with $i = 1$, the only possible nonzero entry in the first column is $T_{1,1}$, hence all the remaining entries in the first row must vanish. Filling in all these zeros, proceed to $i = 2$. Since we just showed $T_{1,2} = 0$, we conclude that the only possible nonzero entry in the second column is $T_{2,2}$, hence all remaining entries in the second row must vanish. Continuing in this manner, we conclude that T is diagonal if it is both normal and upper triangular.

Now let's set $A = VDV^\dagger$ and consider different classes of matrix A. If A is Hermitian, $A = A^\dagger$ immediately yields $D = D^\dagger$, which says that all the eigenvalues of A must be real. If $A^\dagger = -A$, then $D^\dagger = -D$ and all the eigenvalues are purely imaginary. And if $A^\dagger = A^{-1}$, then $D^\dagger = D^{-1}$ and we conclude that all the eigenvalues are unimodular, *i.e.* of the form $e^{i\omega_j}$. This analysis also tells us that any unitary matrix U can be written in the form $U = \exp(iH)$ for some Hermitian matrix H.

Jordan blocks: What happens when an $n \times n$ matrix A is not normal? In this case A is not diagonalizable by a unitary transformation, and while the sum of the dimensions of its eigenspaces is generically equal to the matrix dimension $\dim(A) = n$, this is not guaranteed; it may be less than n. For example, consider the matrix

$$A = \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix} \quad . \tag{1.80}$$

The eigenvalues are solutions to $\det(\lambda I - A) = 0$, hence $\lambda = a$, but there is only one eigenvector, $\psi = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. What is always true for any complex matrix A is that it can be brought to *Jordan canonical form* by a similarity transformation $J = P^{-1}AP$, where P is invertible, and

$$J = \begin{pmatrix} J^{(1)} & & \\ & \ddots & \\ & & J^{(b)} \end{pmatrix} \quad , \tag{1.81}$$

where b is the number of *Jordan blocks* and where each block $J^{(r)}$ is an $n_r \times n_r$ matrix of the form

$$J^{(r)} = \begin{pmatrix} \lambda_r & 1 & & \\ & \lambda_r & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_r \end{pmatrix} \quad . \tag{1.82}$$

Thus each $J^{(r)}$ is tridiagonal, with diagonal elements all given by λ_r and each element directly above the diagonal equal to one. Clearly $J^{(r)}$ has only one eigenvalue, λ_r, and writing the corresponding *right* eigenvector as $\vec{R}^{(r)}$, the condition $J^{(r)} \vec{R}^{(r)} = \lambda^{(r)} \vec{R}^{(r)}$ yields the equations

$$\lambda_r R_1^{(r)} + R_2^{(r)} = \lambda_r R_1^{(r)} \quad , \quad \lambda_r R_2^{(r)} + \psi_3 = \lambda_r R_2^{(r)} \quad \cdots \quad \lambda_r R_{n_r - 1}^{(r)} + R_{n_r}^{(r)} = \lambda_r R_{n_r - 1}^{(r)} \quad , \tag{1.83}$$

where $n_r = \dim(J^{(r)})$. These equations entail $R_2^{(r)} = R_3^{(r)} = \cdots = R_{n_r}^{(r)} = 0$, which says that there is only one such eigenvector, whose components are $R_j^{(r)} = \delta_{j,1}$. Note that the corresponding *left* eigenvector $\vec{L}^{(r)}$ then has components $L_j^{(r)} = \delta_{j,n_r}$. If $n_r > 1$ we then have $\langle L^{(r)} | R^{(r)} \rangle \equiv \vec{L}^{(r)} \cdot \vec{R}^{(r)} = 0$, which means that the left and right eigenvectors of A which correspond to the Jordan blocks with $n_r > 1$ are orthogonal. *Nota bene*: It may be the case that there are degeneracies among the eigenvalues $\{\lambda_r\}$.

To summarize[9], for every general complex $n \times n$ matrix A,

---

[9]See https://en.wikipedia.org/wiki/Jordan_normal_form.

- A may be brought to Jordan canonical form by a similarity transformation $J = P^{-1}AP$, where $J = \mathsf{bdiag}\big(J^{(1)}, \ldots, J^{(b)}\big)$ is block diagonal, with each $(J^{(r)})_{ij} = \lambda_r\, \delta_{i,j} + \delta_{i,j-1}$ with $\dim\big(J^{(r)}\big) = n_r$, for $r \in \{1, \ldots, b\}$.

- There are $b \leq n$ eigenvalues $\{\lambda_1, \ldots, \lambda_b\}$ (again, not necessarily all distinct) and b corresponding eigenvectors $\big\{\vec{R}^{(1)}, \ldots, \vec{R}^{(b)}\big\}$. If $b = n$ then the matrix is diagonalizable.

- The dimension n of the matrix A satisfies $n = n_1 + \ldots + n_b$, *i.e.* it is the sum of the dimensions of all its Jordan blocks.

- Let $\lambda \in \{\lambda_1, \ldots, \lambda_b\}$ be an eigenvalue, and define

$$\mathsf{t}_k(\lambda) = \dim \ker(A - \lambda)^k \quad , \tag{1.84}$$

which is the *dimension of the null space* of the matrix $A - \lambda I$. Then

  ◇ $\mathsf{t}_k(\lambda)$ is the number of Jordan blocks corresponding to the eigenvalue λ.
  ◇ The number of Jordan blocks of size greater than k is $\mathsf{t}_{k+1}(\lambda) - \mathsf{t}_k(\lambda)$. Thus the number of Jordan blocks of size k for the eigenvalue λ is

$$\mathsf{N}_k(\lambda) = 2\,\mathsf{t}_k(\lambda) - \mathsf{t}_{k+1}(\lambda) - \mathsf{t}_{k-1}(\lambda) \quad . \tag{1.85}$$

Singular value decomposition : Note the difference between the decomposition into Jordan canonical form and singular value decomposition (SVD), in which we write an $m \times n$ matrix A as $A = U S V^\dagger$, where U is $m \times k$, V is $n \times k$ (hence $V^\dagger$ is $k \times n$), $U^\dagger U = V^\dagger V = \mathbb{I}_{k \times k}$, and $S = \mathsf{diag}(s_1, \ldots, s_k)$ is a $k \times k$ real matrix with $k \leq \min(m, n)$ and each $s_j > 0$. The elements s_j are the singular values and the rows of U and V are the singular vectors. Note that $A^\dagger A = V S^2 V^\dagger$ is $n \times n$ and $A A^\dagger = U S^2 U^\dagger$ is $m \times m$. If we define

$$F(\lambda) = \prod_{j=1}^{k} \big(\lambda - s_j^2\big) \quad , \tag{1.86}$$

Then

$$P(\lambda) \equiv \det(\lambda - A^\dagger A) = \lambda^{n-k} F(\lambda) \qquad , \qquad Q(\lambda) \equiv \det(\lambda - A A^\dagger) = \lambda^{m-k} F(\lambda) \quad . \tag{1.87}$$

Some comments:

- When $A \in \mathbb{R}$ is real, then both U and V may be chosen to be real, and we may write $A = U S V^\mathsf{T}$.

- We may also adopt a convention where U is $m \times m$, V is $n \times n$, and S to be $m \times n$, where only the first k diagonal elements S_{ii} are the (nonzero and real) singular values. In this case, $U^\dagger U = \mathbb{I}_{m \times m}$ and $V^\dagger V = \mathbb{I}_{n \times n}$.

- For any square $n \times n$ complex matrix A we therefore have two decompositions, via JCF and SVD, *viz.*

$$A = P J P^{-1} = U S V^\dagger \quad , \tag{1.88}$$

where J is the Jordan canonical form of A. When A is normal, $k = n$ and $U = V = P$, which says that the two decompositions are equivalent.

Example : As an example highlighting the difference between eigenvalues and singular values, consider

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} = RDL^{\mathsf{T}} = USV^{\mathsf{T}} \quad , \tag{1.89}$$

where

$$L^{\mathsf{T}} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad , \quad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad , \quad R = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.90}$$

and

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \\ 0 & 0 \end{pmatrix} \quad , \quad S = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{pmatrix} \quad , \quad V^{\mathsf{T}} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 & 1 & 2 \\ -\sqrt{3} & \sqrt{3} & 0 \end{pmatrix} \quad . \tag{1.91}$$

Note that U and V^{T} are both chosen to be real, which is a consequence of the fact that A itself is real. One can check that R is the matrix or right column eigenvectors, L^{T} is the matrix of left row eigenvectors, and Λ is the matrix of eigenvalues. Thus, the three eigenvalues are $\{\lambda_1, \lambda_2, \lambda_3\} = \{1, 1, 0\}$. One also has $L^{\mathsf{T}} R = \mathbb{I}$, *i.e.* $L^{\mathsf{T}} = R^{-1}$, which says that $L_i^{(a)} R_i^{(b)} = \delta^{ab}$ – the row and column eigenvectors satisfy orthonormality. Thus $R^{-1}AR = D$ and A is diagonalizable by R, which is a consequence of there being no Jordan blocks. Note that there are only two singular values, $\{s_1, s_2\} = \{\sqrt{3}, 1\}$, and that U has dimensions 3×2 while V^{T} has dimensions 2×3. One can further check that $U^{\mathsf{T}}U = V^{\mathsf{T}}V = \mathbb{I}_{2\times 2}$.

Had we adopted the convention where both U and V are square, we would have

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} \quad , \quad S = \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad , \quad V^{\mathsf{T}} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 & 1 & 2 \\ -\sqrt{3} & \sqrt{3} & 0 \\ -\sqrt{2} & -\sqrt{2} & \sqrt{2} \end{pmatrix} \quad , \tag{1.92}$$

for which $U^{\mathsf{T}}U = V^{\mathsf{T}}V = \mathbb{I}_{3\times 3}$. The extra zeroes in the matrix S are padding, and there are only two singular values, $\sqrt{3}$ and 1.

For this example, both the set of eigenvalues and the set of singular values are distinct. Furthermore,

$$A^{\dagger}A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix} = VS^2 V^{\mathsf{T}} \quad , \quad AA^{\dagger} = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} = US^2 U^{\dagger} \quad . \tag{1.93}$$

The singular values of A are thus the positive square roots of the eigenvalues of the nonnegative definite Hermitian matrix $A^{\dagger}A$ (or, equivalently, of $AA^{\dagger}$). In general, the eigenvalues λ_j of a non-normal matrix A may not be real, even if $A \in \mathsf{GL}(n, \mathbb{R})$ is itself real. (In this case the eigenvalues are either real or come in complex conjugate pairs.) The singular values, however, are always real and positive.

As a second example, consider the matrix

$$B = \begin{pmatrix} 1 & 1 \\ -3 & 3 \end{pmatrix} = RDL^{\mathsf{T}} = USV^{\mathsf{T}} \quad , \tag{1.94}$$

where

$$L^{\mathsf{T}} = \frac{i}{2\sqrt{2}} \begin{pmatrix} +\sqrt{3}\,e^{-i\phi} & -1 \\ -\sqrt{3}\,e^{+i\phi} & 1 \end{pmatrix} \quad , \quad D = \begin{pmatrix} 2+i\sqrt{2} & 0 \\ 0 & 2-i\sqrt{2} \end{pmatrix} \quad , \quad R = \begin{pmatrix} 1 & 1 \\ \sqrt{3}\,e^{+i\phi} & \sqrt{3}\,e^{-i\phi} \end{pmatrix} \quad (1.95)$$

with $e^{i\phi} = \frac{1}{\sqrt{3}}\left(1+i\sqrt{2}\right)$ and

$$U = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad , \quad S = \begin{pmatrix} 3\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix} \quad , \quad V^{\mathsf{T}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \quad . \quad (1.96)$$

Note that the two right eigenvectors form a complex conjugate pair, as do the two left eigenvectors. This situation pertains for every complex eigenvalue, since if $\lambda_r \in \mathbb{C}$ is an eigenvalue then so is λ_r^*. Again, since $B \in \mathsf{GL}(2,\mathbb{R})$, the U and V matrices may be chosen real. There are two singular values $\{s_1, s_2\} = \{3\sqrt{2}, \sqrt{2}\}$. But unlike the matrix A in the previous example, B has complex eigenvalues $\lambda_\pm = 2 \pm i\sqrt{2}$, and the matrices L^{T} and R of the left (row) and right (column) eigenvectors are complex. As in the previous case, $L^{\mathsf{T}} = R^{-1}$, hence $R^{-1}BR = D$, *i.e.* B is diagonalized by the matrix R, which is possible because there are no nontrivial Jordan blocks when it is brought to canonical form.