Chapter 2

Linear maps and ODEs

The aim of this chapter is to study *linear* dynamical systems and in particular

- to obtain insight in the qualitative features of their phase portrait and how these relate to eigenvalues of the defining matrices;
- to derive estimates that settle the stability issue for the zero steady state;
- to discuss the more subtle issue of topological equivalence (conjugacy) of these systems.

2.1 Linear maps in $\mathbb{R}^n$

Consider a linear map

$$x \mapsto Ax, \quad x \in \mathbb{R}^n, \tag{2.1}$$

where $A = (a_{ij})_{i,j=1,\ldots,n}$ is an $n \times n$ real matrix and Ax is the product of the matrix A and the vector $x = (x_1, x_2, \ldots, x_n)^T$:

$$(Ax)_i = \sum_{j=1}^{n} a_{ij} x_j, \quad i = 1, \ldots, n.$$

The map (2.1) will be denoted by A as well. Note however, that if another basis in $\mathbb{R}^n$ is selected, the same map will be represented by a different matrix. We extend (2.1) to complex vectors by the rule

$$A(u + iv) = Au + iAv, \quad u + iv \in \mathbb{C}^n, \quad u, v \in \mathbb{R}^n.$$

Now consider a discrete-time dynamical system $\{\mathbb{N}, \mathbb{R}^n, A^k\}$ or $\{\mathbb{Z}, \mathbb{R}^n, A^k\}$ defined by iteration of the map A (and – if defined – its inverse).

2.1.1 Dynamics in eigenspaces

Recall that an *eigenvalue* of A is a number $\lambda \in \mathbb{C}$ satisfying

$$Av = \lambda v,$$

for a nonzero vector $v \in \mathbb{C}^n$, which is called a corresponding *eigenvector*. The eigenvalues of a real matrix A can be complex but then occur in complex-conjugate pairs. The set of all eigenvalues of A is denoted by $\sigma(A)$ and is called the *spectrum* of A. It should be known to the reader that $\lambda \in \mathbb{C}$ is an eigenvalue of A if and only if it is a root of the *characteristic polynomial*

$$h(\lambda) = \det(\lambda I - A), \tag{2.2}$$

where I is the unit $n \times n$ matrix. The polynomial h and hence the *characteristic equation*

$$h(\lambda) = 0 \tag{2.3}$$

are independent of the choice of the basis, so one can speak about the eigenvalues of a linear map A. Note that $h(A) = 0$ (*Hamilton-Cayley Theorem*).

First consider algebraically simple eigenvalues of A, i.e. simple roots of the characteristic polynomial.

(*Simple real eigenvalue*) Suppose that λ is a simple real eigenvalue of A with the eigenvector $v \in \mathbb{R}^n$, i.e. $Av = \lambda v$. The line

$$X_\lambda = \{x \in \mathbb{R}^n : x = sv,\ s \in \mathbb{R}\}$$

is invariant under A, since $x \in X_\lambda$ implies $Ax \in X_\lambda$. If $x = sv$, then

$$Ax = A\ sv = sAv = \lambda sv.$$

This means that on the line X_λ the dynamics is simply the multiplication by the number λ in every time step: if $x = s_0 v$ then $A^k x = s_k v$ where $s_k = \lambda s_{k-1}$, so that

$$s_k = \lambda^k s_0.$$

If $|\lambda| < 1$, $|s_k| \to 0$ as $k \to +\infty$, while $|s_k| \to +\infty$ as $k \to -\infty$ if $s_0 \neq 0$. In contrast, if $|\lambda| > 1$ and $s_0 \neq 0$ then $|s_k| \to +\infty$ as $k \to +\infty$, while $|s_k| \to 0$ as $k \to -\infty$. Notice that if $\lambda > 0$, then the time-series is monotone, while if $\lambda < 0$, the time-series exhibits oscillation. Also notice that for $\lambda = 0$ the decay of s_k to zero is instantaneous. Finally, note that we did not use the assumption that λ is simple. The assumption is neverthelass made, because for simple eigenvalues the formulated results provide the whole story, while for multiple eigenvalues they capture only part of it, see below.

(*Simple complex-conjugate pair of eigenvalues*) Suppose $\lambda, \bar{\lambda}$ is a pair of simple complex eigenvalues with eigenvectors $v, \bar{v} \in \mathbb{C}^n$, i.e.

$$Av = \lambda v, \quad A\bar{v} = \bar{\lambda}\bar{v}.$$

(Again, the observations below are correct even if λ and $\bar{\lambda}$ are not simple.) Since $\lambda \neq \bar{\lambda}$ the vectors v and $\bar{v}$ are complex-linearly-independent (see Exercise 2.5.2) and the plane in $\mathbb{R}^n$

$$X_{\lambda,\bar{\lambda}} = \{x \in \mathbb{R}^n : x = zv + \bar{z}\bar{v}, \quad z \in \mathbb{C}\}$$

is invariant under A. Indeed, if $x = zv + \bar{z}\bar{v}$ then

$$Ax = A(zv + \bar{z}\bar{v}) = zAv + \bar{z}A\bar{v} = \lambda zv + \bar{\lambda}\bar{z}\bar{v}.$$

This means that A acts as

$$z \mapsto z' = \lambda z,$$

i.e. as a multiplication by $\lambda \in \mathbb{C}$. If $\lambda = \rho e^{i\psi}$ and $z = re^{i\varphi}$, then

$$z' = (\rho r)e^{i(\psi+\varphi)}.$$

The motion in $X_{\lambda,\bar{\lambda}}$ is therefore a superposition of the rotation through the angle $\psi = \arg\lambda$ and the radial expansion ($\rho > 1$) or contraction ($\rho < 1$) with factor $\rho = |\lambda|$.

Alternatively (but equivalently) we can work with real coordinates (a, b) and write

$$X_{\lambda,\bar{\lambda}} = \{x \in \mathbb{R}^n : x = a \text{ Re } v - b \text{ Im } v,\ a, b \in \mathbb{R}\}.$$

The relationship between the coordinates is given by

$$z = \frac{1}{2}(a + ib).$$

Writing $Av = \lambda v$ as

$$A(\text{Re } v + i \text{ Im } v) = (\text{Re } \lambda + i \text{ Im } \lambda)(\text{Re } v + i \text{ Im } v)$$

we obtain

$$\begin{aligned} A(\text{Re } v) &= \text{Re } \lambda \text{ Re } v - \text{Im } \lambda \text{ Im } v, \\ A(\text{Im } v) &= \text{Im } \lambda \text{ Re } v + \text{Re } \lambda \text{ Im } v. \end{aligned}$$

Therefore, in the coordinates (a, b) the map A restricted to $X_{\lambda,\bar{\lambda}}$ is represented by the matrix

$$\begin{pmatrix} \text{Re } \lambda & -\text{Im } \lambda \\ \text{Im } \lambda & \text{Re } \lambda \end{pmatrix} = \rho \begin{pmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{pmatrix}.$$

Whenever the characteristic polynomial (2.2) has n distinct (and therefore necessarily simple) roots λ_i, $i = 1, \ldots, n$, we can combine our understanding of the dynamics in each of the invariant subspaces X_{λ_i} (for real λ_i) or $X_{\lambda_i,\bar{\lambda}_i}$ (for complex λ_i) to obtain an overall description of the dynamics. The key point is that $\mathbb{R}^n$ can be decomposed into these subspaces.

Definition 2.1 *Let $X_1, \ldots, X_r$ be linear subspaces of a linear space X. We say that X is the* ***direct sum*** *of $X_1, \ldots, X_r$, and write*

$$X = X_1 \oplus X_2 \oplus \cdots \oplus X_r,$$

whenever every $x \in X$ can be written in a unique way as

$$x = x_1 + x_2 + \cdots + x_r$$

with $x_i \in X_i$ for $i = 1, 2, \ldots, r$.

As our building blocks X_{λ_i} and $X_{\lambda_i,\bar{\lambda}_i}$ are linearly independent and together have dimension n, they form a direct sum decomposition of $\mathbb{R}^n$. So an arbitrary $x \in \mathbb{R}^n$ can be uniquely written as a linear combination of elements of X_{λ_i} and $X_{\lambda_i,\bar{\lambda}_i}$. The action of A on x can then be simply reconstructed from that on the invariant subspaces by superposition. Since a generic matrix A has only simple eigenvalues, the above consideration completely describes the dynamics generated by such a map. If there are no eigenvalues with $|\lambda| = 1$, we get exponential contraction/expansion on the linear eigenspaces, combined with rotation in the case of nonreal eigenvalues.

What happens when the characteristic equation (2.3) has multiple roots?

(*Real multiple eigenvalue*) First assume that the multiple (i.e., not simple) eigenvalue λ is *real*. The preceding analysis still applies when there are n linearly-independent eigenvectors. But this is an exceptional case. Generically, there are $m \geq 1$ *Jordan chains* of linearly-independent vectors $v^{(j,k)} \in \mathbb{R}^n$ such that

$$\begin{cases} Av^{(j,1)} &= \lambda v^{(j,1)}, \\ Av^{(j,2)} &= \lambda v^{(j,2)} + v^{(j,1)}, \\ \quad \cdots & \\ Av^{(j,n_j)} &= \lambda v^{(j,n_j)} + v^{(j,n_j-1)}, \end{cases} \tag{2.4}$$

for $j = 1, 2, \ldots, m$. The integer n_j is called the *length* of the Jordan chain. The number of Jordan chains with length l can be computed by the formula

$$N(l, \lambda) = R_{l+1} - 2R_l + R_{l-1},$$

where $R_0 = n$ and $R_k = \text{rank } (A - \lambda I)^k$ for $k \geq 1$. The vectors $v^{(j,1)}$ are the eigenvectors of A corresponding to the eigenvalue λ. The vectors $v^{(j,2)}, v^{(j,3)}, \ldots, v^{(j,n_j)}$ are called *generalized eigenvectors* corresponding to the eigenvalue λ. It follows from (2.4) that these vectors are null-vectors of the matrix $(A - \lambda I)^{n_j}$.

We now have an n_j-dimensional subspace of $\mathbb{R}^n$ defined by

$$X_\lambda^{(j)} = \{x \in \mathbb{R}^n : x = \sum_{k=1}^{n_j} c_k^{(j)} v^{(j,k)}, \quad c_k^{(j)} \in \mathbb{R}\} \tag{2.5}$$

that is invariant under (2.1). A coordinate vector $c = (c_1, c_2, \ldots, c_{n_j})^T \in \mathbb{R}^{n_j}$ is mapped to

$$c' = Jc,$$

where the $n_j \times n_j$ matrix J is given by the *Jordan block*

$$J = \begin{pmatrix} \lambda & 1 & & & 0 \\ & \lambda & 1 & & \\ & & & & \\ & & & \lambda & 1 \\ 0 & & & & \lambda \end{pmatrix}. \tag{2.6}$$

Therefore, to understand how orbits of (2.1) behave in $X_\lambda^{(j)}$, we must consider J^k.

The low-dimensional examples (see Exercise 2.5.8 for the general case)

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}^k = \begin{pmatrix} \lambda^k & k\lambda^{k-1} \\ 0 & \lambda^k \end{pmatrix}$$

and

$$\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}^k = \begin{pmatrix} \lambda^k & k\lambda^{k-1} & \frac{k(k-1)}{2}\lambda^{k-2} \\ 0 & \lambda^k & k\lambda^{k-1} \\ 0 & 0 & \lambda^k \end{pmatrix}$$

clearly indicate that all elements of J^k tend to zero as $k \to +\infty$ when $|\lambda| < 1$, whereas some elements of J^k diverge as $k \to +\infty$ when $|\lambda| > 1$. Moreover, in the critical case $|\lambda| = 1$ only higher multiplicity can and, for suitable initial vectors, will lead to divergent orbits.

The direct sum of all linear subspaces $X_\lambda^{(j)}$,

$$X_\lambda = X_\lambda^{(1)} \oplus X_\lambda^{(2)} \oplus \cdots \oplus X_\lambda^{(m)},$$

is an invariant linear subspace of $\mathbb{R}^n$ called the *generalized eigenspace* of A corresponding to the eigenvalue λ. Its dimension is $(n_1 + \cdots + n_m)$. Note that X_λ is the null-space of $(A - \lambda I)^l$ for $l = \max_{1 \le j \le m} n_j$.

(*Complex multiple eigenvalue*) A similar construction applies when λ is a multiple complex eigenvalue. In this case, $v^{(j,k)}$, $j = 1, 2, \ldots, m$, $k = 1, 2, \ldots, n_j$, are complex vectors and the Jordan chains corresponding to the eigenvalue $\bar{\lambda}$ can be composed of complex-conjugate vectors $\bar{v}^{(j,k)}$. Then

$$X_{\lambda,\bar{\lambda}}^{(j)} = \{x \in \mathbb{R}^n : x = \sum_{k=1}^{n_j} \left(c_k^{(j)} v^{(j,k)} + \bar{c}_k^{(j)} \bar{v}^{(j,k)} \right), \quad c_k^{(j)} \in \mathbb{C}\} \tag{2.7}$$

is a real $2n_j$-dimensional invariant subspace of $\mathbb{R}^n$ parameterized by $c \in \mathbb{C}^{n_j}$. Also in this case, $\|J^k c\| \to 0$ as $k \to +\infty$ when $|\lambda| < 1$, and $\|J^k c\| \to \infty$ as $k \to +\infty$ when $|\lambda| > 1$ and $c \neq 0$. The direct sum of all linear subspaces $X_{\lambda,\bar{\lambda}}^{(j)}$,

$$X_{\lambda,\bar{\lambda}} = X_{\lambda,\bar{\lambda}}^{(1)} \oplus X_{\lambda,\bar{\lambda}}^{(2)} \oplus \cdots \oplus X_{\lambda,\bar{\lambda}}^{(m)},$$

is an invariant linear subspace of $\mathbb{R}^n$ called the *generalized eigenspace* of A corresponding to the eigenvalues $\lambda, \bar{\lambda}$. Its dimension is $2(n_1 + \cdots + n_m)$.

It is also possible to define

$$X_\lambda^{(j)} = \{x \in \mathbb{C}^n : x = \sum_{k=1}^{n_j} c_k^{(j)} v^{(j,k)}, \quad c_k^{(j)} \in \mathbb{C}\}$$

and consider the generalized eigenspace corresponding to λ

$$X_\lambda = X_\lambda^{(1)} \oplus X_\lambda^{(2)} \oplus \cdots \oplus X_\lambda^{(m)}$$

as a complex-linear subspace of $\mathbb{C}^n$. Then $X_{\lambda,\bar{\lambda}} = (X_\lambda \oplus X_{\bar{\lambda}}) \cap \mathbb{R}^n$.

Since $\mathbb{R}^n$ can be decomposed into the direct sum of all generalized eigenspaces X_λ for real eigenvalues λ and $X_{\lambda,\bar{\lambda}}$ for complex-conjugated eigenvalue pairs $\lambda, \bar{\lambda}$, we may summarize the main results of this section as follows.

Theorem 2.2 *If all eigenvalues of A are located strictly inside the unit circle in the complex plane, then for any $x \in \mathbb{R}^n$ we have that*

$$\|A^k x\| \to 0,$$

as $k \to +\infty$.

If there is an eigenvalue outside the unit circle, or at least one generalized eigenvector corresponding to some eigenvalue on the unit circle, then there exists $x \in \mathbb{R}^n$ such that

$$\|A^k x\| \to \infty,$$

as $k \to +\infty$. $\square$

In the next section we shall prove a somewhat stronger result using more general techniques. Moreover, each statement of the above theorem has a suitable converse, see Exercise 2.5.10.

Example 2.3 (Dynamics generated by planar linear maps)

When $n = 2$, equation (2.1) takes the form

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

and defines a linear dynamical system $\{\mathbb{N}, \mathbb{R}^2, A^k\}$, where

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{2.8}$$

Let $\tau = \mathrm{Tr}(A) = a_{11} + a_{22}$, $\Delta = \det(A) = a_{11}a_{22} - a_{21}a_{12}$. Then the characteristic

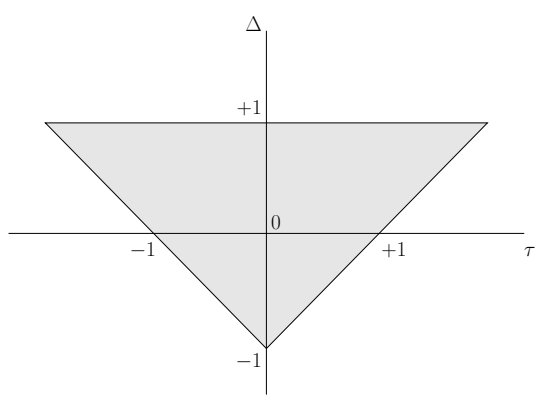

Figure 2.1: Stability triangle for planar linear maps.

equation (2.3) can be written as

$$\lambda^2 - \tau\lambda + \Delta = 0$$

and has two (possibly complex) solutions

$$\lambda_{1,2} = \frac{\tau}{2} \pm \sqrt{\frac{\tau^2}{4} - \Delta}.$$

It can easily be shown that both eigenvalues λ_j satisfy $|\lambda| < 1$ if and only if the point (τ, Δ) falls inside the triangle

$$\{(\tau, \Delta) : \Delta < 1,\ \Delta > -\tau - 1,\ \Delta > \tau - 1\}$$

depicted in Figure 2.1. In this case, in accordance with Theorem 2.2, $\|A^k x\| \to 0$ for all $x \in \mathbb{R}^2$.

The reader is also invited to verify that A has

– an eigenvalue $\lambda_1 = 1$ along the boundary where $\Delta = \tau - 1$;
– an eigenvalue $\lambda_1 = -1$ along the boundary where $\Delta = -\tau - 1$;
– eigenvalues $\lambda_{1,2} = e^{\pm i\theta}$ with $\cos\theta = \frac{1}{2}\tau$ along the boundary where $\Delta = 1$. $\diamond$

2.1.2 Growth estimates and the spectral radius

For any norm in $\mathbb{R}^n$ we define the associated *operator norm* for linear maps:

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\|=1} \|Ax\| \ . \tag{2.9}$$

The number $\|A\|$ describes a rate of expansion/contraction of the map A that depends on the choice of the norm in $\mathbb{R}^n$. A norm independent measure is given by the spectral radius.

Definition 2.4 *The **spectral radius** of a linear map A is defined by*

$$r(A) = \sup_{\lambda \in \sigma(A)} |\lambda| \ .$$

The relation between these quantities is specified by the following lemma, which once again shows that the eigenvalues of A yield information about the growth or decay of the time-series obtained by iterating A.

Lemma 2.5 (Gelfand's formula)

$$r(A) = \lim_{k \to \infty} \|A^k\|^{1/k} = \inf_{k \geq 1} \|A^k\|^{1/k}. \tag{2.10}$$

Proof: Define

$$r = \inf_{k \geq 1} \|A^k\|^{1/k}.$$

Then for all $\varepsilon > 0$ there exists $m = m(\varepsilon)$ such that $\|A^m\|^{1/m} \leq r + \varepsilon$. For arbitrary k, put $k = pm + q$ with $0 \leq q \leq m - 1$. Then

$$\|A^k\|^{1/k} \leq \|A^m\|^{p/k} \|A\|^{q/k} \leq (r + \varepsilon)^{pm/k} \|A\|^{q/k}.$$

As $k \to \infty$, necessarily $q/k \to 0$ and $pm/k \to 1$, hence

$$\limsup_{k\to\infty} \|A^k\|^{1/k} \le r + \varepsilon.$$

By letting $\varepsilon \downarrow 0$ we deduce that

$$\limsup_{k\to\infty} \|A^k\|^{1/k} \le r = \inf_{k\ge 1} \|A^k\|^{1/k},$$

from which we conclude that the limit of $\|A^k\|^{1/k}$ for $k \to \infty$ exists and that this limit equals the infimum.

For $|\lambda| > r$ the series

$$\lambda^{-1} \sum_{k=0}^{\infty} \lambda^{-k} A^k \tag{2.11}$$

converges absolutely. If we multiply (2.11), either from the left or from the right, by $(\lambda I - A)$, we obtain

$$\sum_{k=0}^{\infty} \lambda^{-k} A^k - \sum_{k=0}^{\infty} \lambda^{-(k+1)} A^{k+1} = \lambda^0 A^0 = I.$$

Thus the series (2.11) defines the inverse $(\lambda I - A)^{-1}$ and, accordingly, $\lambda \notin \sigma(A)$. It follows that $r(A) \le r$.

It remains to exclude the possibility that $r(A) < r$. The matrix-valued function

$$f(z) = (I - zA)^{-1}$$

is analytic in $\{z \in \mathbb{C} : |z| < 1/r(A)\}$. Since $f^{(k)}(0) = k!A^k$, the *Cauchy integral formula* yields the identity

$$A^k = \frac{1}{2\pi i} \int_\gamma \frac{f(z)}{z^{k+1}}\, dz\ ,$$

where γ is a circle centered at the origin with radius $(r(A)+\varepsilon)^{-1}$ for arbitrary $\varepsilon > 0$. It follows that

$$\|A^k\| \le (r(A) + \varepsilon)^k \sup_{z\in\gamma} \|f(z)\|$$

and hence

$$r = \lim_{k\to\infty} \|A^k\|^{1/k} \le r(A) + \varepsilon\ .$$

By letting $\varepsilon \downarrow 0$, we obtain that $r \le r(A)$. $\square$

Remark: The matrix-valued function $\lambda \mapsto (\lambda I - A)^{-1}$ is called the *resolvent* of A. The series representation above is the Taylor series at $\lambda = \infty$. By working with $z = \lambda^{-1}$ we avoid having to deal with the point at infinity. $\diamond$

Theorem 2.6 *For every $\rho > r(A)$ there exists $M \ge 1$ such that*

$$\|A^k\| \le M\rho^k$$

for all $k \ge 1$.

Proof: If $\rho > r(A)$ then Lemma 2.5 implies $\|A^k\| \le \rho^k$ for *large* k and consequently

$$M = \sup_{k\ge 0} \rho^{-k}\|A^k\| < \infty \ .$$

Thus *for all* $k \ge 0$ the inequality

$$\|A^k\| \le M\rho^k$$

holds. For $k = 0$ this gives $M \ge 1$, since $A^0 = I$ and $\|I\| = 1$. $\square$

Using (2.9), we find that

$$\|A^k x\| \le M\rho^k\|x\|, \qquad x \in \mathbb{R}^n, k \in \mathbb{N}. \tag{2.12}$$

This estimate establishes the global asymptotic stability of the origin when $r(A) < 1$. Indeed, in this case we can choose ρ such that $r(A) < \rho < 1$. Then for any point $x \in \mathbb{R}^n$ with $\|x\| \le \varepsilon/M$ the estimate $\|A^k x\| \le \rho^k\varepsilon$ holds, which implies the (global) asymptotic stability of $x = 0$. In Exercise 2.5.12 the reader is asked to explore the geometric manifestation of the fact that, possibly, $M\rho > 1$.

We now show that one can introduce an equivalent norm[1], tailor-made for the linear map A, such that M reduces to one (the advantage being that the estimate for $k = 1$ carries all information, in the sense that the estimate for arbitrary $k > 1$ is obtained by iteration of the one for $k = 1$).

Theorem 2.7 *Let $\rho > r(A)$. There exists an equivalent norm $\|\cdot\|_1$ on $\mathbb{R}^n$ such that $\|A\|_1 \le \rho$.*

Proof: Define $\|\cdot\|_1$ for $x \in \mathbb{R}^n$ by the formula:

$$\|x\|_1 = \sum_{k=0}^{\infty} \rho^{-k}\|A^k x\|.$$

Formula (2.10) implies that this series converges. Indeed, for k sufficiently large and some $q < 1$,

$$\rho^{-1}\|A^k\|^{1/k} \le q$$

and hence

$$\rho^{-k}\|A^k x\| \le \rho^{-k}\|A^k\|\,\|x\| \le \|x\|q^k \ .$$

Clearly, $\|x\|_1 \ge 0$ for all $x \in \mathbb{R}^n$ and $\|x\|_1 = 0$ if and only if $x = 0$. Likewise the properties $\|\alpha x\|_1 = |\alpha|\|x\|_1$ and $\|x + y\|_1 \le \|x\|_1 + \|y\|_1$ for $x, y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$ follow straightaway from the corresponding properties of $\|\cdot\|$. So $\|\cdot\|_1$ is a norm on $\mathbb{R}^n$. Moreover, we have

$$\|x\| \le \|x\|_1 \le \left(\sum_{k=0}^{\infty} \left(\rho^{-1}\|A^k\|^{1/k}\right)^k\right)\|x\|,$$

[1]Recall that two norms on a linear space E are called *equivalent*, if there are two constants, $C_2 \ge C_1 > 0$, such that for all $x \in E$: $C_1\|x\| \le \|x\|_1 \le C_2\|x\|$.

i.e. $\|\cdot\|_1$ is *equivalent* to $\|\cdot\|$ on $\mathbb{R}^n$. Now, for $x \in \mathbb{R}^n$,

$$\|Ax\|_1 = \sum_{k=0}^{\infty} \rho^{-k} \|A^{k+1}x\| = \rho(\|x\|_1 - \|x\|)$$

so that

$$\|Ax\|_1 \leq \rho \|x\|_1, \quad x \in \mathbb{R}^n. \quad \square$$

Note that by induction the estimate $\|A\|_1 \leq \rho$ implies that

$$\|A^k x\|_1 \leq \rho^k \|x\|_1, \quad x \in \mathbb{R}^n,$$

for all $k \geq 0$.

Definition 2.8 *The map A is called a* ***(strict) linear contraction*** *if for some ρ with $0 < \rho < 1$ we have*

$$\|Ax\|_1 \leq \rho \|x\|_1, \quad x \in \mathbb{R}^n.$$

Since a linear contraction obviously has the origin as its asymptotically stable fixed point, we have proved the following criterion, which we have already mentioned before.

Theorem 2.9 *If all eigenvalues of a matrix A are located strictly inside the unit circle, then A defines a linear contraction with respect to an appropriate norm and the origin is globally asymptotically stable.* $\square$

As with Theorem 2.2 a certain converse of the statements above holds (see Exercise 2.5.10).

Since $\sigma(A^{-1}) = \{\lambda^{-1} : \lambda \in \sigma(A)\}$ whenever $0 \notin \sigma(A)$, we can obtain similar bounds for A^{-1} in terms of $\inf_{\lambda \in \sigma(A)} |\lambda|$.

Theorem 2.10 *Assume that $|\lambda| > \tilde{\rho} > 0$ for all $\lambda \in \sigma(A)$. Then there exists $M \geq 1$ such that*

$$\|A^{-k}\| \leq M \tilde{\rho}^{-k}, \quad k \geq 0.$$

Moreover, there exists an equivalent norm $\|\cdot\|_1$ on $\mathbb{R}^n$ such that

$$\|A^{-1}\|_1 \leq \tilde{\rho}^{-1}$$

and hence $\|A^{-k}\|_1 \leq \tilde{\rho}^{-k}$ for all $k \geq 0$.

2.1.3 Hyperbolic linear maps and spectral projectors

In this section we focus on the situation when

$$\sigma(A) = \sigma_s(A) \cup \sigma_u(A),$$

where $\sigma_s(A)$ is located strictly inside the unit circle but does not contain zero, while $\sigma_u(A)$ is located strictly outside the unit circle.

Definition 2.11 *A linear map $x \mapsto Ax$ in $\mathbb{R}^n$ is called* ***hyperbolic*** *if A is invertible and has no eigenvalue λ with $|\lambda| = 1$.*

The direct sum of all generalized eigenspaces of A corresponding to $\sigma_s(A)$ form an invariant subspace T^s of A, called the *stable eigenspace*, while the direct sum of all generalized eigenspaces of A corresponding to σ_u form the *unstable eigenspace* T^u of A. Moreover, we have $T^s \oplus T^u = \mathbb{R}^n$, meaning that any $x \in \mathbb{R}^n$ can be uniquely decomposed as

$$x = x_s + x_u$$

with $x_s \in T^s$ and $x_u \in T^u$. Clearly, $\sigma(A|_{T^s}) = \sigma_s(A)$ and $\sigma(A|_{T^u}) = \sigma_u(A)$.

There are explicit methods to find x_s (and hence x_u) for a given x, using the so called *spectral projectors*. We begin by assuming that $\sigma_s(A)$ consists of just one simple eigenvalue $\lambda \in \mathbb{R}$ with $|\lambda| < 1$. Let $v \in \mathbb{R}^n$ be the corresponding eigenvector, so that

$$Av = \lambda v.$$

Let A^T denote as usual the transposed matrix corresponding to A. It also has a simple eigenvalue λ with some eigenvector $w \in \mathbb{R}^n$,

$$A^T w = \lambda w.$$

Since the eigenvalue λ is simple, the *Fredholm decomposition* for linear maps (see Lemma 6.14 in Chapter 6) tells us that $\langle w, v\rangle \neq 0$ and therefore w can be selected to satisfy the normalization condition

$$\langle w, v\rangle = 1.$$

Now the map

$$x \mapsto Px = \langle w, x\rangle v$$

assigns to each $x \in \mathbb{R}^n$ a vector $x_s = Px \in T^s \subset \mathbb{R}^n$. The map P is linear and has obviously the property $P^2 = P$. We say that P *projects* $\mathbb{R}^n$ onto T^s. Note that

$$T^s = \{x \in \mathbb{R}^n : Px = x\} = \{sv : s \in \mathbb{R}\}.$$

The linear map $Q = I - P$ or, in more detail,

$$x \mapsto Qx = x - \langle w, x\rangle v$$

projects $\mathbb{R}^n$ onto T^u and

$$T^u = \{x \in \mathbb{R}^n : Px = 0\} = \{x \in \mathbb{R}^n : \langle w, x\rangle = 0\} = w^\perp.$$

If σ_s is composed of n_s eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_{n_s}$, all of which are simple and satisfy $|\lambda_i| < 1$, the corresponding projectors can be defined by

$$P = \sum_{i=1}^{n_s} P_i, \quad Q = I - P,$$

where

$$P_i x = \langle w^{(i)}, x \rangle v^{(i)}, \tag{2.13}$$

is the projection of $\mathbb{R}^n$ onto the one-dimensional invariant eigenspace X_{λ_i} corresponding to the eigenvalue λ_i. Here $Av^{(i)} = \lambda_i v^{(i)}$ and $A^T w^{(i)} = \lambda_i w^{(i)}$ for $i = 1, 2, \ldots, n_s$, and

$$\langle w^{(i)}, v^{(j)} \rangle = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

(Verify!) The projectors defined by (2.13) have the following properties:

$$P_i^2 = P_i, \quad P_i P_j = 0, \quad \text{for} \quad j \neq i. \tag{2.14}$$

Notice that the eigenvalues (and the corresponding eigenvectors) can be complex, provided that $\langle \cdot, \cdot \rangle$ is interpreted as the standard scalar product in $\mathbb{C}^n$.

There is an alternative method to construct a projector P_i onto the generalized eigenspace corresponding to an eigenvalue $\lambda_i \in \mathbb{C}$, which is applicable for both simple and multiple eigenvalues and does not require computing any eigenvector or generalized eigenvector. In this method, the projection matrix P_i is constructed as a *polynomial function* of the matrix A. We describe this method briefly below. Actually, we explain how to compute P_i corresponding to any eigenvalue λ_i of A.

Suppose that the characteristic polynomial $h(\lambda) = \det(\lambda I - A)$ has d distinct roots: $\lambda_1, \lambda_2, \ldots, \lambda_d \in \mathbb{C}$. Then, it can be written as

$$h(\lambda) = \prod_{i=1}^{d} (\lambda - \lambda_i)^{m_i},$$

where m_i are the algebraic multiplicities of the eigenvalues and hence $m_1 + \cdots + m_d = n$. This implies that the following *partial fraction decomposition* holds:

$$\frac{1}{h(\lambda)} = \sum_{i=1}^{d} \frac{g_i(\lambda)}{(\lambda - \lambda_i)^{m_i}}, \tag{2.15}$$

where $g_i(\lambda)$ are polynomials of degree $m_i - 1$ which can be found, for example, by the method of unknown coefficients.

Multiplying both sides of (2.15) by $h(\lambda)$ leads to the equation

$$1 = \sum_{i=1}^{d} p_i(\lambda), \tag{2.16}$$

where

$$p_i(\lambda) = \frac{h(\lambda) g_i(\lambda)}{(\lambda - \lambda_i)^{m_i}} = g_i(\lambda) \prod_{j \neq i} (\lambda - \lambda_j)^{m_j}. \tag{2.17}$$

Each p_i is a polynomial of degree less than n in λ. Moreover, all the polynomials $p_i^2 - p_i$ and $p_i p_j$ with $j \neq i$ are *divisible* by the characteristic polynomial h. Indeed,

$$p_i(\lambda) p_j(\lambda) = \frac{g_i(\lambda) g_j(\lambda) h^2(\lambda)}{(\lambda - \lambda_i)^{m_i} (\lambda - \lambda_j)^{m_j}} = h(\lambda) \left(g_i(\lambda) g_j(\lambda) \prod_{k \neq i, k \neq j} (\lambda - \lambda_k)^{m_k} \right).$$

Moreover,

$$p_i^2(\lambda) - p_i(\lambda) = p_i(\lambda)(p_i(\lambda) - 1) = -p_i(\lambda) \sum_{j \neq i} p_j(\lambda) = -\sum_{j \neq i} p_i(\lambda) p_j(\lambda),$$

with each term in the sum divisible by h due to the previous formula.

Define now

$$P_i = p_i(A), \quad i = 1, 2, \ldots, d. \tag{2.18}$$

Since $h(A) = 0$ by the Hamilton-Cayley Theorem, we see that the linear maps P_i thus constructed satisfy (2.14), i.e., are projectors. Because each P_i is a polynomial function of A, it is linear and commutes with A. This implies that each range $P_i(\mathbb{C}^n)$ is an invariant subspace of A. The formula (2.17) implies that $P_i(A)x = 0$ if $Ax = \lambda_j$ with $j \neq i$. Therefore, since $\sum_{i=1}^d P_i = I$, P_i must project $\mathbb{C}^n$ onto the generalized eigenspace of A corresponding to the eigenvalue λ_i:

$$X_{\lambda_i} = P_i(\mathbb{C}^n).$$

If there are d_s distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_{d_s}$ satisfying $|\lambda_i| < 1$, then the projectors onto T^s and T^u are given by

$$P = \sum_{i=1}^{d_s} P_i \quad \text{and} \quad Q = I - P,$$

repectively.

Example 2.12 (Practical computation of spectral projectors)

Consider

$$A = \begin{pmatrix} -1 & -3 \\ \frac{3}{2} & \frac{7}{2} \end{pmatrix}.$$

Then

$$\lambda I - A = \begin{pmatrix} \lambda + 1 & 3 \\ -\frac{3}{2} & \lambda - \frac{7}{2} \end{pmatrix}$$

so that the characteristic polynomial is

$$h(\lambda) = \det(\lambda I - A) = \lambda^2 - \tfrac{5}{2}\lambda + 1 = (\lambda - \tfrac{1}{2})(\lambda - 2)$$

and the eigenvalues of A are $\lambda_1 = \frac{1}{2}$ and $\lambda_2 = 2$. Hence A is hyperbolic.

The vector

$$v = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

is an eigenvector corresponding to the eigenvalue $\lambda_1 = \frac{1}{2}$, while the vector

$$u = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

is an eigenvector corresponding to the eigenvalue $\lambda_2 = 2$. The transposed matrix

$$A^T = \begin{pmatrix} -1 & \frac{3}{2} \\ -3 & \frac{7}{2} \end{pmatrix}$$

has

$$w = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

as an eigenvector corresponding to its eigenvalue $\lambda_1 = \frac{1}{2}$. Notice that

$$\langle w, v \rangle = 1, \quad \langle w, u \rangle = 0.$$

For any vector

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2,$$

we have $\langle w, x \rangle = x_1 + x_2$ and, therefore,

$$Px = \langle w, x \rangle v = \begin{pmatrix} 2x_1 + 2x_2 \\ -x_1 - x_2 \end{pmatrix}.$$

This gives

$$P = \begin{pmatrix} 2 & 2 \\ -1 & -1 \end{pmatrix} \tag{2.19}$$

which is indeed a projection matrix:

$$P^2 = \begin{pmatrix} 2 & 2 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} 2 & 2 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ -1 & -1 \end{pmatrix} = P.$$

The map $x \mapsto Px$ projects the space $\mathbb{R}^2$ onto the eigenline T^s spanned by the eigenvector v. The matrix

$$Q = I - P = \begin{pmatrix} -1 & -2 \\ 1 & 2 \end{pmatrix}$$

defines a projection $x \mapsto Qx$ onto the eigenline T^u spanned by the eigenvector u. Clearly, $PQ = QP = 0$.

To find the spectral projectors P and Q using the partial fraction decomposition (2.15), write

$$\frac{1}{h(\lambda)} = \frac{1}{(\lambda - \frac{1}{2})(\lambda - 2)} = \frac{-\frac{2}{3}}{\lambda - \frac{1}{2}} + \frac{\frac{2}{3}}{\lambda - 2}.$$

This gives $1 = p_1(\lambda) + p_2(\lambda)$, where

$$p_1(\lambda) = -\tfrac{2}{3}(\lambda - 2), \quad p_2(\lambda) = \tfrac{2}{3}(\lambda - \tfrac{1}{2}).$$

Therefore

$$P = p_1(A) = -\tfrac{2}{3}(A - 2I) = -\tfrac{2}{3} \begin{pmatrix} -3 & -3 \\ \frac{3}{2} & \frac{3}{2} \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ -1 & -1 \end{pmatrix},$$

which (re-assuringly) is the same projector (2.19) as before. For the projector Q, we have

$$Q = p_2(A) = \tfrac{2}{3}\,(A - \tfrac{1}{2}\,I) = \tfrac{2}{3}\begin{pmatrix} -\frac{3}{2} & -3 \\ \frac{3}{2} & 3 \end{pmatrix} = \begin{pmatrix} -1 & -2 \\ 1 & 2 \end{pmatrix}.$$

Note that $P + Q = I$, so that the computation of Q as $p_2(A)$ is redundant: It can be found by merely evaluating $I - P$. $\diamond$

Remark:

There is another alternative formula for the projector P_i, namely:

$$P_i = \frac{1}{2\pi i}\int_{\gamma_i} (\lambda I - A)^{-1}\, d\lambda. \tag{2.20}$$

Here the integral is interpreted (for each matrix element) as the contour integral in the complex plane along the small circle γ_i around λ_i. It can be computed by the method of *residues*. Notice that the formula (2.20) is valid in case of multiple eigenvalues as well. Moreover, by a suitable choice of contour γ it works for any spectral set, i.e. any subset of $\sigma(A)$. For example, $\gamma = S^1$, where S^1 is the unit circle, gives the projector onto T^s of a hyperbolic map A. The formula (2.20) can also be generalized to an operator A in a Banach space. $\diamond$

Theorem 2.13 *For any hyperbolic map A, there is an equivalent norm $\|\cdot\|_1$ in which $A|_{T^s}$ and $A^{-1}|_{T^u}$ become linear contractions, i.e. for some ρ with $0 < \rho < 1$ we have*

$$\|Ax\|_1 \le \rho\|x\|_1, \quad x \in T^s, \tag{2.21}$$

and

$$\|A^{-1}x\|_1 \le \rho\|x\|_1, \quad x \in T^u. \tag{2.22}$$

Proof:

Take some ρ such that $0 \le r(A|_{T^s}) < \rho < 1$. According to Theorem 2.7, there exists an equivalent norm $\|\cdot\|_1$ on T^s such that (2.21) holds.

Next note that $A^{-1}|_{T^u}$ is defined and that $\sigma(A^{-1}|_{T^u})$ is contained strictly inside the unit circle. Therefore, it follows from Theorem 2.10 that there exists an equivalent norm $\|\cdot\|_1$ on T^u such that

$$\|A^{-1}x\|_1 \le \tilde{\rho}\|x\|_1, \quad x \in T^u,$$

for some $\tilde{\rho} < 1$. To obtain the estimates (2.21) and (2.22) with the same ρ, we can simply take $\rho := \max\{\rho, \tilde{\rho}\}$.

Finally, extend the norm to all of $\mathbb{R}^n$ by setting

$$\|x\|_1 = \|x_s\|_1 + \|x_u\|_1 \quad \text{for} \quad x = x_s + x_u \in T^s \oplus T^u. \quad \square$$

Remarks:

(1) Note that, by induction,

$$\|A^k x\|_1 \leq \rho^k \|x\|_1, \quad x \in T^s,$$

and

$$\|A^{-k} x\|_1 \leq \rho^k \|x\|_1, \quad x \in T^u.$$

(2) Note that we did not use in the proof the invertibility of A, i.e. the absence of eigenvalue 0. All that matters for the estimates is that there are no eigenvalues on the unit circle. Indeed, we only used $A^{-1}|_{T^u}$ and we can replace this by $(A|_{T^u})^{-1}$ since $0 \notin \sigma(A|_{T^u})$. $\diamond$

Due to Theorem 2.13 we can give an alternative definition of a hyperbolic linear map, which does not use eigenvalues.

Definition 2.14 *An invertible linear continuous map $A : X \to X$ on a Banach space is called* ***hyperbolic*** *if $X = T^s \oplus T^u$, where both T^s and T^u are invariant under A, and with respect to an equivalent norm $\|\cdot\|_1$ we have*

$$\|A|_{T^s}\|_1 \leq \rho, \quad \|A^{-1}|_{T^u}\|_1 \leq \rho,$$

for some $0 < \rho < 1$.

Remark:

Not all authors include "invertibility" in the definition of "hyperbolicity". We do this to facilitate the formulation of a topological equivalence result ahead (see Theorem 2.30). For linear maps, the sign of an eigenvalue has no impact on growth or decay of the image length, but it does have an impact on orientation. In particular, the border between orientation preserving and orientation reversing linear contractions corresponds to zero eigenvalue. $\diamond$

2.2 Linear autonomous systems of ODEs

Consider an autonomous system of linear ODEs

$$\dot{x} = Ax, \quad x \in \mathbb{R}^n. \tag{2.23}$$

The associated flow is given by $\varphi^t(x) = e^{tA}x$ where

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

(Note in particular the group property $e^{(t+s)A} = e^{tA}e^{sA}$.) Now consider the linear continuous-time dynamical system $\{\mathbb{R}, \mathbb{R}^n, \varphi^t\}$.

For the spectra we have

$$\sigma(e^{tA}) = e^{t\sigma(A)} = \{\mu \in \mathbb{C} : \mu = e^{t\lambda},\ \lambda \in \sigma(A)\},$$

for any $t \in \mathbb{R}$, so in particular zero is never an eigenvalue of e^{tA} (indeed, e^{tA} has inverse e^{-tA}).

Note that $z \mapsto e^{tz}$ maps the imaginary axis onto the unit circle and the left half plane inside the unit circle. Accordingly, the condition $|\lambda| < 1$ translates in the condition $\mathrm{Re}\ \lambda < 0$.

2.2.1 Dynamics in the eigenspaces

First consider the dynamics in one- and two-dimensional invariant eigenspaces associated to, respectively, a simple real eigenvalue and a simple complex pair of eigenvalues.

(*Simple real eigenvalue*) Let λ be a simple real eigenvalue of A with eigenvector $v \in \mathbb{R}^n$, *i.e.* $Av = \lambda v$. The line

$$X_\lambda = \{x \in \mathbb{R}^n : x = sv,\ s \in \mathbb{R}\}$$

is invariant with respect to the flow, since $x \in X_\lambda$ implies $\dot{x} \in X_\lambda$. The function $x(t) = s(t)v \in X_\lambda$ satisfies the equation (2.23) if and only if

$$\dot{s} = \lambda s, \tag{2.24}$$

since

$$\dot{s}v = \dot{x} = Ax = Asv = sAv = \lambda sv.$$

Thus we find that

$$s(t) = s_0 e^{\lambda t}$$

and that $s(t) \to 0$ as $t \to +\infty$ if $\lambda < 0$, while $s(t) \to 0$ as $t \to -\infty$ if $\lambda > 0$.

(*Simple complex-conjugate pair of eigenvalues*) Let $\lambda, \bar{\lambda}$ be simple nonreal eigenvalues with the eigenvectors $v, \bar{v} \in \mathbb{C}^n$, *i.e.*

$$Av = \lambda v, \quad A\bar{v} = \bar{\lambda}\bar{v}.$$

The plane

$$X_{\lambda,\bar{\lambda}} = \{x \in \mathbb{R}^n : x = zv + \bar{z}\bar{v}, \quad z \in \mathbb{C}\} \subset \mathbb{R}^n$$

is invariant with respect to the flow generated by (2.23). Moreover, $x(t) = z(t)v + \bar{z}(t)\bar{v}$ is a solution if and only if

$$\dot{z} = \lambda z.$$

This implies

$$z(t) = z_0 e^{\lambda t}.$$

If $\lambda = \alpha + i\omega$ and $z_0 = r_0 e^{i\varphi_0}$, then

$$z(t) = (r_0 e^{\alpha t}) e^{i(\varphi_0 + \omega t)}.$$

If $\alpha \neq 0$, the motion in $X_{\lambda,\bar{\lambda}}$ is a superposition of rotation (with angular speed ω) and radial divergence (if $\alpha > 0$) or convergence (if $\alpha < 0$).

If $\alpha = \mathrm{Re}\ \lambda = 0$ then the invariant subspace $X_{\lambda,\bar{\lambda}}$ is filled with periodic orbits parametrized by $r_0 > 0$. The corresponding periodic solutions are given by

$$x(t) = r_0 \left(e^{i(\varphi_0 + \omega t)} v + e^{-i(\varphi_0 + \omega t)} \bar{v} \right),$$

where the phase φ_0 determines the position of $x(0)$ on the closed orbit. Of course, the period of the orbit is $p = \frac{2\pi}{\omega}$.

Since a generic $n \times n$ matrix A has n different simple eigenvalues and therefore n linearly independent eigenvectors, the above consideration completely describes the dynamics of a generic linear system (2.23). If there are no eigenvalues with $\mathrm{Re}\ \lambda = 0$, we get exponential contraction/expansion on the linear eigenspaces, combined with rotations in the case of nonreal eigenvalues.

There is a possibility, however, that the matrix A has multiple eigenvalues.

(*Real multiple eigenvalue*) Assume first that the eigenvalue $\lambda \in \mathbb{R}$. If there is a basis of eigenvectors in the corresponding invariant subspace, then the solutions behave as in the simple real case. Otherwise, there are Jordan chains (2.4) of vectors $v^{(j,k)} \in \mathbb{R}^n,\ j = 1, 2, \ldots, m,\ k = 1, 2, \ldots, n_j$. The n_j-dimensional subspace $X_\lambda^{(j)}$ defined by (2.5) is invariant and $c = (c_1, c_2, \ldots, c_{n_j})^T \in \mathbb{R}^{n_j}$ satisfies

$$\dot{c} = Jc,$$

where J is again given by (2.6). Therefore, any solution component is the sum of scalar multiples of $t^k e^{\lambda t}$ with $k < n_j$ (see Exercise 2.5.8 for an explicit expression for e^{tJ}).

(*Complex multiple eigenvalue*) A similar construction applies when $\lambda \in \mathbb{C}$ is a multiple eigenvalue. In this case, $v^{(j,k)}$ are complex vectors and the Jordan chains corresponding to the eigenvalue $\bar{\lambda}$ can be composed of complex-conjugate vectors $\bar{v}^{(j,k)}$. Then $X_{\lambda,\bar{\lambda}}$ defined by (2.7) is a real $2n_j$-dimensional invariant subspace parameterized by $c \in \mathbb{C}^{n_j}$. All components of $\mathrm{Re}\ c(t)$ and $\mathrm{Im}\ c(t)$ are the sums of scalar multiples of $t^k e^{\alpha t} \sin(\omega t), t^k e^{\alpha t} \cos(\omega t)$, where $\lambda = \alpha + i\omega,\ k < n_j$.

Thus, the following theorem holds.

Theorem 2.15 *If all eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of the matrix A satisfy* $\mathrm{Re}\ \lambda_i < 0$, *then for any $x \in \mathbb{R}^n$*

$$\|e^{tA}x\| \to 0,$$

as $t \to \infty$.

If at least one eigenvalue of A satisfies $\mathrm{Re}\ \lambda > 0$ *or there is at least one generalized eigenvector with corresponding eigenvalue* $\mathrm{Re}\ \lambda = 0$, *then there exists $x \in \mathbb{R}^n$ such that*

$$\|e^{tA}x\| \to \infty,$$

as $t \to \infty$. $\square$

Remark:

Notice that when $\|x(t)\| \to 0$ as $t \to \infty$, the decay need not to be monotone, see Exercise 2.5.13. Moreover, as in the discrete time case the statements in this theorem can be reversed, see Exercise 2.5.11. $\diamond$

Example 2.16 (Phase portraits of planar linear autonomous ODEs)

When $n = 2$, equation (2.23) takes the form

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

or

$$\begin{cases} \dot{x}_1 &= a_{11}x_1 + a_{12}x_2, \\ \dot{x}_2 &= a_{21}x_1 + a_{22}x_2. \end{cases} \tag{2.25}$$

Let

$$\tau = \mathrm{Tr}(A) = a_{11} + a_{22}, \ \Delta = \det(A) = a_{11}a_{22} - a_{21}a_{12}.$$

Then the characteristic equation (2.3) can be written as

$$\lambda^2 - \tau\lambda + \Delta = 0$$

and has two (possibly complex) solutions

$$\lambda_{1,2} = \frac{\tau}{2} \pm \sqrt{\frac{\tau^2}{4} - \Delta}.$$

If follows right away that both eigenvalues λ_j satisfy $\mathrm{Re}\,\lambda < 0$ if and only if

$$\tau < 0 \quad \text{and} \quad \Delta > 0.$$

Under this condition, Theorem 2.15 implies that $\|e^{tA}x\| \to 0$ as $t \to \infty$ for any $x \in \mathbb{R}^2$.

Traditionally, three generic cases are distinguished, characterized by inequalities in terms of τ and Δ:

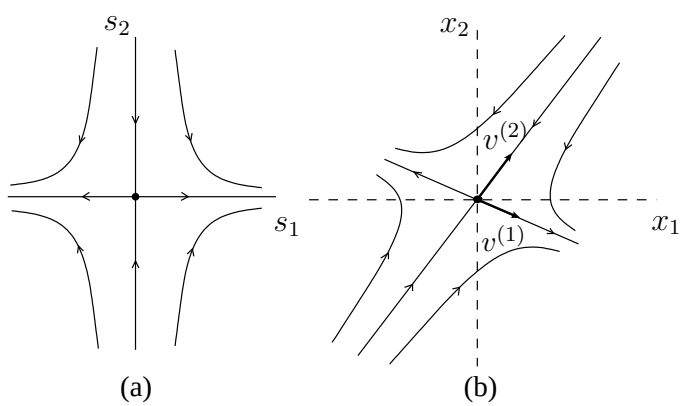

Figure 2.2: Canonical (a) and original (b) saddles.

(*Saddle*: $\Delta < 0$) The matrix A has one positive and one negative eigenvalue: $\lambda_2 < 0 < \lambda_1$. Let $v^{(1)}$ and $v^{(2)}$ be the corresponding eigenvectors:

$$Av^{(j)} = \lambda_j v^{(j)}, \quad j = 1, 2.$$

Let us use these eigenvectors as new basis vectors in $\mathbb{R}^2$, i.e. write any $x \in \mathbb{R}^2$ in the form

$$x = s_1 v^{(1)} + s_2 v^{(2)},$$

with some $s_{1,2} \in \mathbb{R}$. Notice that $s_{1,2}$ can be considered as new coordinates in $\mathbb{R}^2$. In these new coordinates, the system (2.25) takes a particularly simple form, namely

$$\begin{cases} \dot{s}_1 &= \lambda_1 s_1, \\ \dot{s}_2 &= \lambda_2 s_2. \end{cases} \tag{2.26}$$

So each of the equations can be solved independently. This gives

$$s_1(t) = s_1(0)e^{\lambda_1 t}, \quad s_2(t) = s_2(0)e^{\lambda_2 t},$$

yielding a phase portrait in the (s_1, s_2)-plane presented in Figure 2.2(a). The phase portrait in the (x_1, x_2)-plane (see Figure 2.2(b)) appears then by applying a linear transformation to the phase portrait in the (s_1, s_2)-plane. The equilibrium at the origin is a *saddle*.

(*Node*: $\tau^2 > 4\Delta > 0$) The matrix A has either two positive eigenvalues $\lambda_1 > \lambda_2 > 0$ (when $\tau > 0$) or two negative eigenvalues $\lambda_2 < \lambda_1 < 0$ (when $\tau < 0$). The analysis here is very similar to that in the saddle case. Using the eigenvectors $v^{(j)}$ as

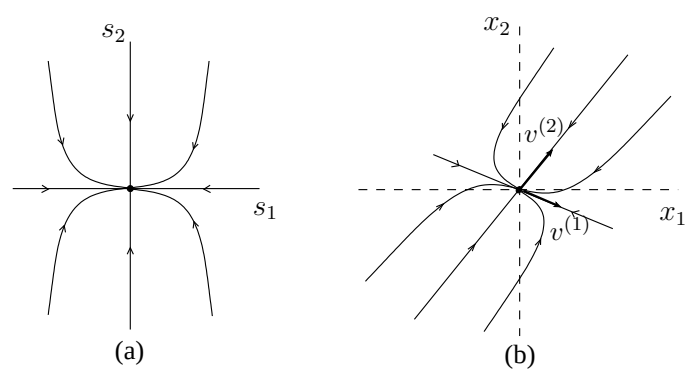

Figure 2.3: Canonical (a) and original (b) stable nodes.

new basis vectors in $\mathbb{R}^2$, we obtain once again system (2.26) and solve it as above. When $\tau < 0$, this leads to a phase portrait shown in Figure 2.3(a), since both $s_1(t)$ and $s_2(t)$ tend to zero exponentially fast as $t \to \infty$. Notice, however, that due to the difference in the convergence rates, a generic orbit of (2.26) tends to the origin horizontally. The phase portrait in the (x_1, x_2)-plane (see Figure 2.3(b)) is obtained by a linear transformation. The equilibrium in the origin is a *stable node*. The case $\tau > 0$ gives an *unstable node*. The corresponding phase portrait can be obtained from that of the stable node by reversing time.

(*Focus*: $0 < \tau^2 < 4\Delta$) The matrix A has now two complex-conjugate eigenvalues $\lambda_{1,2} = \alpha \pm i\omega$ with

$$\alpha = \frac{\tau}{2}, \quad \omega = \frac{1}{2}\sqrt{4\Delta - \tau^2}.$$

The corresponding eigenvectors $v^{(1)}$ and $v^{(2)}$ are also complex but can be selected to be complex-conjugate:

$$v^{(1)} = \bar{v}^{(2)} = w \in \mathbb{C}^2.$$

The vector w can be used to introduce a new coordinate $z \in \mathbb{C}$ which specifies a point $x \in \mathbb{R}^2$ according to the formula

$$x = zw + \bar{z}\bar{w}.$$

As we have already seen, this coordinate satisfies

$$\dot{z} = \lambda z$$

with $\lambda = \alpha + i\omega$, so that

$$z(t) = r_0 e^{\alpha t} e^{i(\varphi_0 + \omega t)},$$

where r_0 and φ_0 are specified by the initial point at $t = 0$. For $s_1 = \text{Re } z$ and $s_2 = \text{Im } z$, we get oscillatory behaviour superimposed with growth ($\alpha > 0$) or decay ($\alpha < 0$):

$$\begin{cases} s_1(t) &= r_0 e^{\alpha t} \cos(\varphi_0 + \omega t), \\ s_2(t) &= r_0 e^{\alpha t} \sin(\varphi_0 + \omega t). \end{cases} \tag{2.27}$$

When $\tau < 0$, we have $\alpha < 0$ and the phase portrait in the (s_1, s_2)-plane is as shown

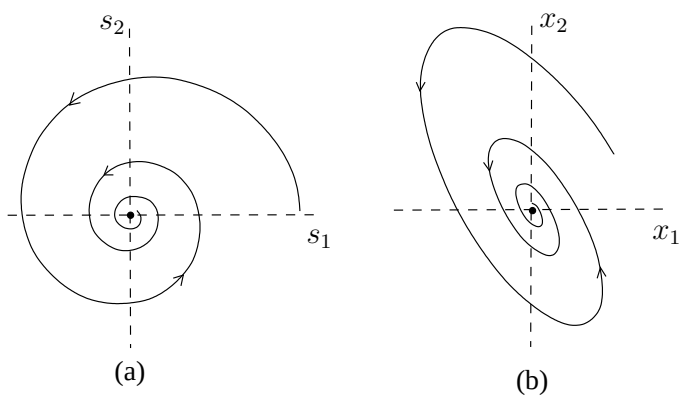

Figure 2.4: Canonical (a) and original (b) stable foci.

in Figure 2.4(a). All orbits spiral to the origin. A linear transformation produces the phase portrait in the (x_1, x_2)-plane (see Figure 2.4(b)). The equilibrium at the origin is a *stable focus*. When $\tau > 0$, one obtains an *unstable focus*. As we have already noted (see Example 1.17 and Exercises 1.5.13 and 1.5.14 in Chapter 1), a node and a focus of the same stability type are topologically equivalent. We return to this issue in Section 2.3.1 ahead.

We leave it to the reader to carry out the analysis of all critical cases (defined by at least one equality relation involving τ and/or Δ). Here we only mention the so-called *center*, which occurs when $\tau = 0$ but $\Delta > 0$. In this case both eigenvalues of A are purely imaginary: $\lambda_{1,2} = \pm i\omega$ with $\omega = \sqrt{\Delta} > 0$. From (2.27) it follows that all solutions of (2.25) are p_0-periodic, where

$$p_0 = \frac{2\pi}{\omega}$$

and, therefore, all nonequilibrium orbits of (2.25) are closed (see Figure 2.5). $\diamond$

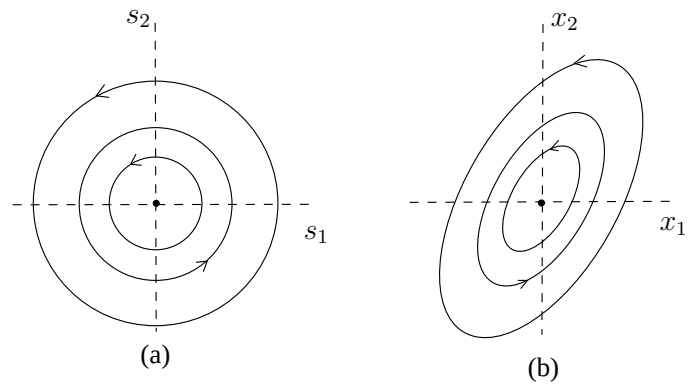

Figure 2.5: Canonical (a) and original (b) centers.

2.2.2 Growth estimates and the spectral bound

Motivated by the observations so far, we can provide the following definition.

Definition 2.17 *The **spectral bound** of a linear map A is defined by*

$$s(A) = \sup_{\lambda \in \sigma(A)} \mathrm{Re}(\lambda).$$

Theorem 2.18 *For every $\omega > s(A)$ there exists $M \geq 1$ such that*

$$\|e^{tA}\| \leq M e^{\omega t}$$

for $t \geq 0$.

Proof: Define $B = -\omega I + A$. Since $\sigma(e^B) = e^{-\omega + \sigma(A)}$ and $|e^\lambda| = e^{\mathrm{Re}(\lambda)}$, we know that all eigenvalues of e^B are inside the unit circle. Let ρ be such that $r(e^B) < \rho < 1$. By Theorem 2.7 we know that

$$\|x\|_1 = \sum_{k=0}^{\infty} \rho^{-k} \|e^{kB} x\|$$

defines an equivalent norm such that $\|e^B\|_1 \leq \rho$.

We claim that, for some $\tilde{M} \geq 1$, the inequality

$$\|e^{tB}\|_1 \leq \rho^t \tilde{M}$$

holds for all $t \geq 0$. Indeed, write $t = m - \nu$ with $m \in \mathbb{N}$ and $\nu \in (0, 1]$ and define

$$\tilde{M} = \sup_{0 < s \leq 1} \|e^{-sB}\|_1.$$

Then $1 \leq \tilde{M} < \infty$ and, moreover,

$$\|e^{tB}\|_1 = \|e^{(m-\nu)B}\|_1 \leq \|e^{mB}\|_1 \|e^{-\nu B}\|_1 \leq \rho^m \tilde{M} \leq \rho^{m-\nu} \tilde{M} = \rho^t \tilde{M}.$$

It follows that

$$\|e^{tA}\|_1 \leq e^{(\omega+\ln\rho)t}\tilde{M} \leq e^{\omega t}\tilde{M}$$

(here we use that $\ln\rho < 0$). Finally, the equivalence of $\|\cdot\|_1$ and $\|\cdot\|$ yields the estimate for $\|e^{tA}\|$ with an adjusted constant M. $\square$

Like in the discrete time case, we can eliminate the constant M by modifying the norm on $\mathbb{R}^n$ such that it is tailor-made for A.

Theorem 2.19 *Let $\omega > s(A)$. There exists an equivalent norm $\|\cdot\|_2$ on $\mathbb{R}^n$ such that*

$$\|e^{tA}\|_2 \leq e^{\omega t}$$

for all $t \geq 0$.

Proof: As in the proof of Theorem 2.18, let

$$\|x\|_1 = \sum_{k=0}^{\infty} \rho^{-k}\|e^{kB}x\|$$

where $B = -\omega I + A$. Now define

$$\|x\|_2 = \int_0^\infty \|e^{\tau B}x\|_1 d\tau.$$

The integral converges and

$$\|x\|_2 = \int_0^\infty \|e^{\tau B}x\|_1 d\tau \leq \int_0^\infty \|e^{\tau B}\|_1\|x\|_1 d\tau \leq M\left(\int_0^\infty \rho^\tau d\tau\right)\|x\|_1 = -\frac{M}{\ln(\rho)}\|x\|_1\,,$$

i.e.

$$\|x\|_2 \leq -\frac{M}{\ln(\rho)}\|x\|_1.$$

Moreover, since $x = e^{-tB}e^{tB}x$, we have

$$\|x\|_1 \leq \|e^{-tB}\|_1\|e^{tB}x\|_1 \leq e^{t\|B\|_1}\|e^{tB}x\|_1$$

and hence

$$\|e^{tB}x\|_1 \geq e^{-t\|B\|_1}\|x\|_1.$$

It follows that

$$\|x\|_2 \geq \left(\int_0^\infty e^{-\tau\|B\|_1}d\tau\right)\|x\|_1 = \frac{1}{\|B\|_1}\|x\|_1.$$

Combining the two estimates above, we obtain

$$\frac{1}{\|B\|_1}\|x\|_1 \leq \|x\|_2 \leq -\frac{M}{\ln(\rho)}\|x\|_1,$$

meaning that $\|\cdot\|_2$ is equivalent to $\|\cdot\|_1$.

Finally, we compute

$$\begin{aligned}
\|e^{tA}x\|_2 &= \int_0^\infty \|e^{\tau B}e^{tA}x\|_1 d\tau \;=\; \int_0^\infty \|e^{-\omega\tau}e^{(t+\tau)A}x\|_1 d\tau \\
&= \int_t^\infty \|e^{-\omega(\theta-t)}e^{\theta A}x\|_1 d\theta \;=\; e^{\omega t}\int_t^\infty \|e^{-\omega\theta}e^{\theta A}x\|_1 d\theta \\
&\le e^{\omega t}\int_0^\infty \|e^{-\omega\theta}e^{\theta A}x\|_1 d\theta \;\le\; e^{\omega t}\int_0^\infty \|e^{\theta B}x\|_1 d\theta \;=\; e^{\omega t}\|x\|_2,
\end{aligned}$$

from which it follows that $\|e^{tA}\|_2 \le e^{\omega t}$ for all $t \ge 0$. □

Definition 2.20 *A norm $\|\cdot\|_2$ in $\mathbb{R}^n$ is called a* ***Lyapunov norm*** *for a linear system $\dot{x} = Ax$ if*

$$\|e^{tA}x\|_2 \le e^{-\alpha t}\|x\|_2, \quad t \ge 0,$$

for some $\alpha > 0$.

Theorem 2.21 *If all eigenvalues of a matrix A have negative real part, then there is an equivalent Lyapunov norm $\|\cdot\|_2$ on $\mathbb{R}^n$ for the linear system $\dot{x} = Ax$ and the origin is a globally asymptotically stable equilibrium of this system.*

Proof: When all eigenvalues of A have negative real part, we have by definition $s(A) < 0$. Thus, there exists $\omega < 0$ such that $s(A) < \omega < 0$. Theorem 2.19 implies that $\|\cdot\|_2$ is the Lyapunov norm with $\alpha = -\omega > 0$. The global asymptotical stability of $x = 0$ in this case is obvious. □

As in the discrete time case one can show that stability can be fully characterized in terms of eigenvalue conditions, see Exercise 2.5.11 for details.

By time reversal, we obtain the following theorem.

Theorem 2.22 *Assume that* $\mathrm{Re}(\lambda) > \tilde{\omega}$ *for all $\lambda \in \sigma(A)$. Then there exists $M \ge 1$ such that*

$$\|e^{-tA}\| \le Me^{-\tilde{\omega}t}, \quad t \ge 0,$$

and there exists an equivalent norm $\|\cdot\|_2$ on $\mathbb{R}^n$ such that

$$\|e^{-tA}\|_2 \le e^{-\tilde{\omega}t}, \quad t \ge 0.$$

2.2.3 Hyperbolic linear ODEs

Definition 2.23 *A linear ODE $\dot{x} = Ax$ is called* hyperbolic *if A has no eigenvalue λ with* $\mathrm{Re}\ \lambda = 0$.

The stable and unstable invariant subspaces $T^{s,u}$ of a linear hyperbolic ODE are defined like those of a linear hyperbolic map. Namely, we can write as before

$$\sigma(A) = \sigma_s(A) \cup \sigma_u(A),$$

where now

$$\sigma_s(A) = \{\lambda \in \sigma(A) : \mathrm{Re}\ \lambda < 0\} \;\text{ and }\; \sigma_u(A) = \{\lambda \in \sigma(A) : \mathrm{Re}\ \lambda > 0\}.$$

Then T^s and T^u will be spanned by all eigenvectors and generalized eigenvectors of A corresponding to $\sigma_s(A)$ and $\sigma_u(A)$, respectively. Spectral projectors can also be computed as before, with γ_i in the formula (2.20) replaced by any contour encircling $\sigma_s(A)$.

The dynamics on T^s is described by Theorem 2.21. There is no need to study the dynamics on T^u separately, since the results can be obtained from those about T^s by reversal of time, i.e. the substitution $t \mapsto -t$. Thus, we arrive at the following theorem (we leave it to the reader to provide the details of the proof).

Theorem 2.24 *For any hyperbolic system $\dot{x} = Ax$, there is an equivalent norm $\|\cdot\|_2$ on $\mathbb{R}^n$ such that for some $\alpha > 0$ we have for all $t \geq 0$*

$$\|e^{tA}x\|_2 \leq e^{-\alpha t}\|x\|_2, \quad x \in T^s,$$

and

$$\|e^{-tA}x\|_2 \leq e^{-\alpha t}\|x\|_2, \quad x \in T^u. \quad \square$$

2.3 Topological classification of linear systems

2.3.1 Topological equivalence of hyperbolic linear ODEs

We consider first the problem of topological classification of linear ODEs. We advise the reader to make Exercises 1.5.13 and 1.5.14 before proceeding.

Theorem 2.25 *The linear system*

$$\dot{x} = Ax, \quad x \in \mathbb{R}^n, \tag{2.28}$$

where all eigenvalues of A have negative real part, is topologically conjugate to the system

$$\dot{y} = -y, \quad y \in \mathbb{R}^n. \tag{2.29}$$

Proof: Let $\|\cdot\|_2$ be a Lyapunov norm: $\|e^{tA}x\|_2 \leq e^{-\alpha t}\|x\|_2, \quad \alpha > 0$. Any nontrivial orbit of (2.28) crosses

$$\Sigma^{n-1} = \{x \in \mathbb{R}^n : \|x\|_2 = 1\}$$

exactly once. Denote by $H : \Sigma^{n-1} \to S^{n-1}$ the map that assigns to the point of intersection of the ray $\{sv : s > 0\}$ with Σ^{n-1} the point of intersection of the same ray with S^{n-1} (see Figure 2.6), i.e.

$$H(x) = \frac{x}{\|x\|}$$

where the standard (Euclidian) norm $\|\cdot\|$ in $\mathbb{R}^n$ is used.

Define now the map $h : \mathbb{R}^n \to \mathbb{R}^n$ as follows. For any $x \neq 0$, let $\tau(x) \neq 0$ be the (positive or negative) time such that

$$x = e^{\tau(x)A}\tilde{x} \quad \text{with} \quad \tilde{x} \in \Sigma^{n-1}.$$

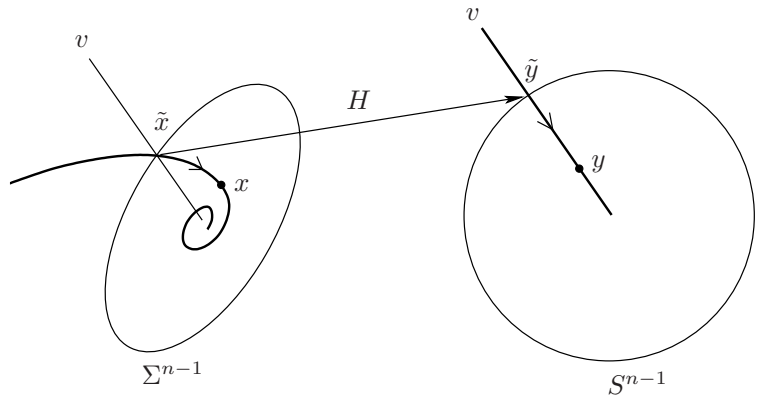

Figure 2.6: Construction of the homeomorphism $x \mapsto y = h(x)$.

We have $\tilde{x} = e^{-\tau(x)A}x$. Map $\tilde{x} \mapsto \tilde{y} = H(\tilde{x})$ and then set

$$y = h(x) = e^{-\tau(x)}\tilde{y} = e^{-\tau(x)}H(e^{-\tau(x)A}x).$$

Let $h(0) = 0$. The map $h : \mathbb{R}^n \to \mathbb{R}^n$ thus constructed has the following properties:

(i) h is a homeomorphism on $\mathbb{R}^n$ (a detailed check of this is left to the reader);

(ii) h maps orbits of (2.23) onto orbits of (2.29), preserving time parametrization.

For the proof of the last statement use the fact that $\tau(e^{sA}x) = \tau(x) + s$. $\square$

Applying Theorem 2.25 in the stable eigenspace directly and after time reversal in the unstable eigenspace, we arrive at the following result.

Theorem 2.26 *Two hyperbolic linear systems in $\mathbb{R}^n$, $\dot{x} = Ax$ and $\dot{y} = By$, are topologically conjugate if the matrices A and B have the same number of eigenvalues with negative real part.*

Remark:

Actually, the inverse result is also valid, i.e. one can write "if and only if" in Theorem 2.26. However, this fact is only of theoretical interest, since it is rarely known a priori that two linear ODE systems are topologically conjugate. $\diamond$

2.3.2 Topological equivalence of hyperbolic linear maps

When are two hyperbolic linear maps topologically equivalent ? We begin with an example that illustrates that that we can use essentially the same idea as used in the continuous time case, but, because now we deal with "jumps" rather than with a continuous orbit, we need to replace the boundary of a ball by a set with non-empty interior.

Example 2.27 (Topologically equivalent scalar linear maps)

Consider two linear scalar maps:

$$f : x \mapsto \frac{1}{2}x \quad \text{and} \quad g : y \mapsto \frac{1}{3}y.$$

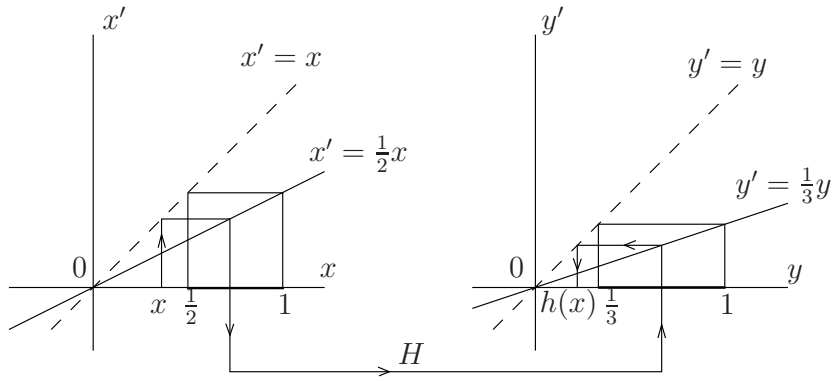

Figure 2.7: The maps $f : x \mapsto \frac{1}{2}x$ and $g : y \mapsto \frac{1}{3}y$ are topologically conjugate.

These maps are not smoothly conjugate, since they have different eigenvalues, i.e. $\frac{1}{2}$ and $\frac{1}{3}$ (see Exercise 1.5.12). However, they are topologically conjugate. Indeed, the following construction (see Figure 2.7) gives a conjugating homeomorphism $h : \mathbb{R} \to \mathbb{R}$.

Take any, for example linear, homeomorphism

$$H : [\tfrac{1}{2}\,, 1] \to [\tfrac{1}{3}\,, 1],$$

satisfying $H(\frac{1}{2}\,) = \frac{1}{3}\,,\ H(1) = 1$. The intervals are called the *fundamental domains*.

For $x \in (0, \frac{1}{2}\,)$ take

$$h(x) = g^{k(x)}(H(f^{-k(x)}(x))),$$

where $k(x)$ is the minimal[2] positive integer k, such that

$$f^{-k}(x) \in [\tfrac{1}{2}\,, 1].$$

For $x > 1$ take

$$h(x) = g^{-l(x)}(H(f^{l(x)}(x))),$$

where $l(x)$ is the minimal positive integer l, such that

$$f^{l}(x) \in [\tfrac{1}{2}\,, 1].$$

Make the analogous construction for $x < 0$. Set $h(0) = 0$. The map h constructed in this manner has the following properties:

(i) $h : \mathbb{R} \to \mathbb{R}$ is a homeomorphism; if H is linear, h is piecewise-linear.

(ii) h sends orbits of f to orbits of g, i.e. $f = h^{-1} \circ g \circ h$.

This means that $f \sim g$.

Now consider another map g defined by

$$g : y \mapsto -\frac{1}{3}y.$$

[2]The adjective "minimal" is added to be specific but it is, in fact, superfluous. Ambiguity can only arise when $x = 2^{-n}$ but then both choices $k(x) = n - 1$ and $k(x) = n$ yield $h(x) = 3^{-n}$ so, after all, the ambiguity is harmless.

This map is not topologically equivalent to f. Indeed, suppose that there exists a continuous map $h : \mathbb{R} \to \mathbb{R}$ sending orbits of f to orbits of g. In particular, it should map the orbit of f

$$\ldots, 1, \frac{1}{2}, \frac{1}{4}, \ldots$$

to the orbit of g

$$\ldots, a, -\frac{a}{3}, \frac{a}{9}, \ldots,$$

where $a \neq 0$. Therefore, we must have

$$h(1) = a, \ h\left(\frac{1}{2}\right) = -\frac{a}{3}, \ h\left(\frac{1}{4}\right) = \frac{a}{9},$$

showing that h is not monotone. Therefore, h is not invertible. $\diamond$

Lemma 2.28 *Let*

$$A : x \mapsto Ax, \quad x \in \mathbb{R}^n, \tag{2.30}$$

be an invertible linear map which is a contraction with respect to a norm $\|\cdot\|$ in $\mathbb{R}^n$. Then any linear map B,

$$B : y \mapsto By, \quad y \in \mathbb{R}^n, \tag{2.31}$$

with $\|B - A\|$ sufficiently small, is topologically conjugate to A.

Proof: We have

$$\|Ax\| \leq \rho \|x\|, \quad x \in \mathbb{R}^n,$$

for some $\rho < 1$. Take the unit sphere

$$S^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$$

and consider its image $AS^{n-1} \subset \mathrm{Int} S^{n-1}$. Define a *fundamental domain* for A:

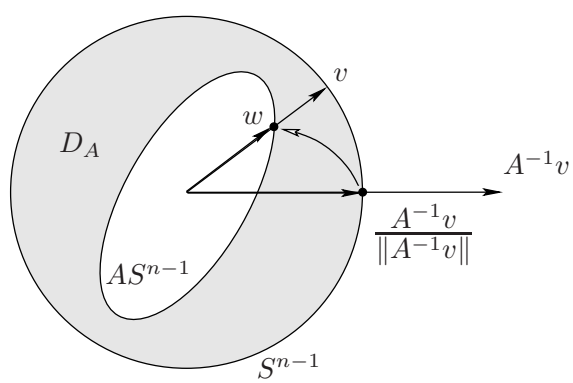

Figure 2.8: Fundamental domain for the contraction A.

$$D_A = S^{n-1} \cup (\mathrm{Int}\ S^{n-1} \setminus \mathrm{Int}\ AS^{n-1}).$$

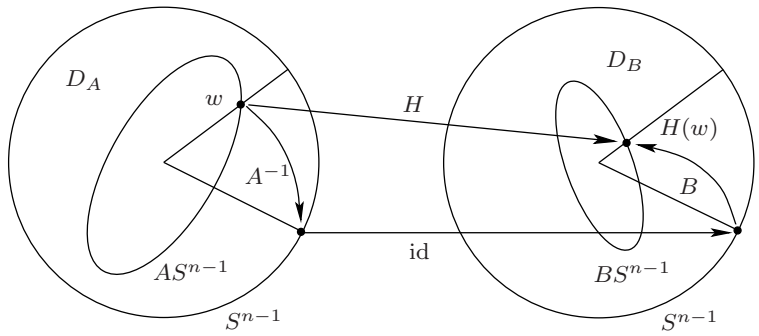

Figure 2.9: Constraints on the map $H : H|_{S^{n-1}} = \mathrm{id},\ H|_{AS^{n-1}} = BA^{-1}$.

This is a closed annulus between S^{n-1} and AS^{n-1} (see Figure 2.8).

If $\|B - A\| \leq \varepsilon$ with sufficiently small ε, we obtain by the triangle inequality

$$\|B\| = \|B - A + A\| \leq \|B - A\| + \|A\| \leq \varepsilon + \rho =: \rho_1 < 1$$

and therefore

$$\|By\| \leq \rho_1 \|y\|, \quad y \in \mathbb{R}^n.$$

Define the fundamental domain for B by

$$D_B = S^{n-1} \cup (\mathrm{Int}\ S^{n-1} \setminus \mathrm{Int}\ BS^{n-1}).$$

Suppose that we can construct a homeomorphism

$$H : D_A \to D_B,$$

which maps the corresponding boundaries of the fundamental domains D_A and D_B into each other and which "agrees" with the maps A and B on them, in the sense that

$$H|_{S^{n-1}} = \mathrm{id}. \tag{2.32}$$

and

$$H|_{AS^{n-1}} = BA^{-1} \tag{2.33}$$

(see Figure 2.9). Then we can define a map

$$h : \mathbb{R}^n \to \mathbb{R}^n$$

by setting first $h|_{D_A} = H$ and then defining for $x \in \mathrm{Int} AS^{n-1}, x \neq 0$,

$$h(x) = B^{k(x)} H(A^{-k(x)} x),$$

where $k(x)$ is the minimal integer $k > 0$, such that $A^{-k}x \in D_A$. Similarly define $h(x)$ for $x \in \mathrm{Ext} S^{n-1}$, and finally set $h(0) = 0$. A careful reasoning using (2.32) and (2.33) shows that the resulting map $h : \mathbb{R}^n \to \mathbb{R}^n$ is a homeomorphism with the property

$$Ax = h^{-1}(Bh(x)), \quad x \in \mathbb{R}^n,$$

meaning that B is topologically conjugate to A. Therefore, the lemma will be proved if we construct a homeomorphism $H : D_A \to D_B$, satisfying (2.32) and (2.33). The rest of the proof is devoted to this construction.

First, we notice that any ray

$$\{x \in \mathbb{R}^n : x = sv,\ s \in (0, \infty)\}$$

for some given $v \in \mathbb{R}^n$ with $\|v\| = 1$, intersects AS^{n-1} at a unique point

$$w = A\left(\frac{A^{-1}v}{\|A^{-1}v\|}\right) = \frac{v}{\|A^{-1}v\|}$$

(see Figure 2.8). This allows us to introduce new coordinates (v, t) in D_A. Namely, identify $x \in D_A$ with a point $(v, t) \in S^{n-1} \times [0, 1]$ as follows. For each $x \in D_A$, take

$$v = \frac{x}{\|x\|} \in S^{n-1},$$

compute

$$a = \frac{1}{\|A^{-1}v\|} = \frac{\|x\|}{\|A^{-1}x\|} \in \mathbb{R},$$

so that $w = av$, and then set

$$t = \frac{a - \|x\|}{a - 1} \in [0, 1].$$

(In Exercise 2.5.16 below you are asked to compute the inverse of the map $x \mapsto (v, t)$.) Make a similar construction for D_B, identifying each point $y \in D_B$ with $(\eta, \tau) \in S^{n-1} \times [0, 1]$.

The boundary constraints (2.32) and (2.33) motivate us to define two homeomorphisms

$$\delta_{1,0} : S^{n-1} \to S^{n-1},$$

by $\delta_1 = \mathrm{id}$ and

$$\delta_0(v) = \frac{BA^{-1}v}{\|BA^{-1}v\|}.$$

From the fact that B is close to A and, thus, BA^{-1} is close to the unit matrix, we infer that δ_0 is close to the identity map. The map

$$\delta^t(v) = \frac{(1-t)\delta_0(v) + t\delta_1(v)}{\|(1-t)\delta_0(v) + t\delta_1(v)\|},$$

is a homeomorphism on S^{n-1} for all $t \in [0, 1]$ (Prove this!). Therefore, the map $\Delta : S^{n-1} \times [0, 1] \to S^{n-1} \times [0, 1]$ defined by the formula

$$\Delta(v, t) = (\delta^t(v), t)$$

is a homeomorphism. Finally, the homeomorphism Δ induces, via the above constructed coordinates, a homeomorphism

$$H : D_A \to D_B$$

with the required properties (2.32) and (2.33). Indeed, the property (2.32) is obvious, while the property (2.33) can be verified as follows. Take $v \in S^{n-1}$ and consider the point

$$w = \frac{v}{\|A^{-1}v\|} \in AS^{n-1}$$

that has coordinates $(v, 0) \in S^{n-1} \times [0, 1]$ in D_A. Note that any point in AS^{n-1} can be uniquely represented in this way. Then

$$BA^{-1}w = BA^{-1}\left(\frac{v}{\|A^{-1}v\|}\right) = \frac{BA^{-1}v}{\|A^{-1}v\|} \in BS^{n-1}.$$

On the other hand,

$$\delta_0(v) = \frac{BA^{-1}v}{\|BA^{-1}v\|} \in S^n,$$

which implies

$$\eta = \delta^0(v) = \delta_0(v) = \frac{BA^{-1}v}{\|BA^{-1}v\|} \in S^n.$$

The point with coordinates $(\eta, 0) \in S^{n-1} \times [0, 1]$ corresponds to the point

$$H(w) = \frac{\eta}{\|B^{-1}\eta\|} = \frac{BA^{-1}v}{\|A^{-1}v\|} \in BS^{n-1},$$

from which we conclude that $H(w) = BA^{-1}w$ for $w \in AS^{n-1}$. $\square$

Theorem 2.29 *Two linear invertible contractions in $\mathbb{R}^n$ that either both preserve or both reverse the orientation of $\mathbb{R}^n$ are topologically conjugate.*

Proof: Let $x \mapsto Ax$ and $x \mapsto Bx$ be two such maps. By the assumption, the determinants of the corresponding matrices A and B are both positive or negative. Therefore, there exists a continuous family of matrices $M(t), t \in [0, 1]$, such that $M(0) = A, M(1) = B$, and

$$\det M(t) \neq 0, \quad r(M(t)) < 1$$

for all $t \in [0, 1]$ (see Exercise 2.5.17). By Lemma 2.28, any two linear maps $M(t')$ and $M(t'')$ are topologically conjugate if $|t'' - t'|$ is sufficiently small. Since $[0, 1]$ can be covered with finitely many such intervals, A is topologically conjugate to B. $\square$

Suppose that

$$f_1 : \mathbb{R}^{n_1} \to \mathbb{R}^{n_1}, \quad g_1 : \mathbb{R}^{n_1} \to \mathbb{R}^{n_1},$$

are two topologically equivalent maps, i.e. $f_1 = h_1^{-1} \circ g_1 \circ h_1$ for some homeomorphism $h_1 : \mathbb{R}^{n_1} \to \mathbb{R}^{n_1}$. Suppose also that

$$f_2 : \mathbb{R}^{n_2} \to \mathbb{R}^{n_2}, \quad g_2 : \mathbb{R}^{n_2} \to \mathbb{R}^{n_2},$$

are two other topologically equivalent maps, i.e. $f_2 = h_2^{-1} \circ g_2 \circ h_2$ for some homeomorphism $h_2 : \mathbb{R}^{n_2} \to \mathbb{R}^{n_2}$. Then obviously the direct product maps $f, g : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ defined coordinate-wise,

$$f : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} f_1(x_1) \\ f_2(x_2) \end{pmatrix}, \quad g : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} g_1(x_1) \\ g_2(x_2) \end{pmatrix},$$

are topologically equivalent. Indeed,

$$f = h^{-1} \circ g \circ h,$$

where $h : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ is a homeomorphism defined by

$$h : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} h_1(x_1) \\ h_2(x_2) \end{pmatrix}.$$

The topology in the product space can be induced, for example, by the norm

$$\|x\| = \|x_1\| + \|x_2\|.$$

Let $n_s(M)$ and $n_u(M)$ denote the number of eigenvalues of the matrix M inside and outside the unit circle, respectively (taking multiplicities into account). Further, let $p_s(M)$ and $p_u(M)$ denote the products of the eigenvalues inside and outside the unit circle, respectively. Recall that topological conjugacy of two (in our case linear) invertible maps is equivalent to the conjugacy of their inverses (with the same conjugating homeomorphism). Applying Theorem 2.29 to the restriction of a hyperbolic linear map to its stable eigenspace, and to the inverse of the restriction to the unstable eigenspace, we obtain the following theorem.

Theorem 2.30 *Two linear hyperbolic maps in $\mathbb{R}^n$, $x \mapsto Ax$, $y \mapsto By$, are topologically conjugate if the following properties hold simultaneously:*

$$n_s(A) = n_s(B), \; p_s(A)p_s(B) > 0, \; p_u(A)p_u(B) > 0. \quad \square$$

Remark:

As in the continuous time case, the inverse for Theorem 2.30 is also valid but again of minor interest. $\diamond$

2.4 References

There are many good books, where linear autonomous ODEs are treated. The standard references here are [Arnol'd 1973, Hirsch & Smale 1974]. The analysis of linear maps can be done using similar techniques, see, for example [Kato 1980, Chapter I]. The construction of spectral projectors using the partial fraction decomposition is inspired by unpublished notes of Joop Kolk (Utrecht University).

The fundamental domains are used in [Katok & Hasselblatt 1995] to prove the local topological equivalence of a map with a stable hyperbolic linear part to its linearization.

2.5 Exercises

E 2.5.1 (Exceptional cases of real simple eigenvalue)

Describe the dynamics generated by the linear map A on the line X_λ when
(i) $\lambda = 1$;
(ii) $\lambda = -1$;
(iii) $\lambda = 0$.

E 2.5.2 (Linear algebra refresher)

(a) Let λ be a nonreal eigenvalue of A with eigenvector v. Show that:
(i) v and $\bar{v}$ are linearly independent in $\mathbb{C}^n$;
(ii) Re v and Im v are linearly independent in $\mathbb{R}^n$.
(b) Let λ and μ be eigenvalues of A with eigenvectors v and w, respectively. Show that $\lambda \neq \mu$ implies that these vectors are linearly independent.
(c) Let $Av^{(i)} = \lambda_i v^{(i)}$ with $v^{(i)} \neq 0$ and $\lambda_i \neq \lambda_j$ for $i \neq j$, where $i, j = 1, 2, \ldots, n$. Prove that the vectors $v^{(i)}$ are linearly independent.

E 2.5.3 (Spectral projection)

Consider

$$A = \begin{pmatrix} \frac{11}{2} & -3 \\ 6 & -\frac{7}{2} \end{pmatrix}$$

as a generator of a discrete-time dynamical system in $\mathbb{R}^2$. Compute the projectors corresponding to the decomposition $\mathbb{R}^2 = T^s \oplus T^u$.

E 2.5.4 (Stability conditions for linear maps in $\mathbb{R}^2$ and $\mathbb{R}^3$)

(a) Prove all statements formulated in Example 2.3.
(b) Consider a three-dimensional linear map

$$x \mapsto Ax, \quad x \in \mathbb{R}^3$$

and write the characteristic equation $\det(\lambda I - A) = 0$ in the form

$$\lambda^3 + a_0\lambda^2 + a_1\lambda + a_2 = 0.$$

Show that all its roots lie strictly inside the unit circle if and only if the following inequalities hold simultaneously:

$$1 + a_0 + a_1 + a_2 > 0, \quad 1 - a_0 + a_1 - a_2 > 0, \quad |a_2| < 1, \quad \text{and} \quad 1 - a_2^2 > a_1 - a_0 a_2.$$

Also verify that $\lambda_1 = 1$ is a root if $1 + a_0 + a_1 + a_2 = 0$, that $\lambda_1 = -1$ is a root if $1 - a_0 + a_1 - a_2 = 0$, and that there are two roots $\lambda_{1,2}$ with $\lambda_1\lambda_2 = 1$ if

$$1 - a_2^2 = a_1 - a_0 a_2.$$

Show that in the latter case $\lambda_{1,2} = e^{\pm i\theta}$ with $0 < \theta < \pi$ where $\cos\theta = \frac{1}{2}(a_2 - a_0)$, provided that $|a_2 - a_0| < 2$.

E 2.5.5 (Operator norm of matrices)

Let $\|x\|$ be the standard Euclidian norm of $x \in \mathbb{R}^n$. Consider a $n \times n$ matrix A with real elements and define its *operator norm* $\|A\|$ by the formula (2.9) with Ax understood as the matrix-vector product.

(a) Prove that $\|A\| = \sqrt{\mu_{\max}}$, where $\mu_{\max}$ is the largest eigenvalue of the symmetric matrix $A^T A$. (*Hint*: $\|Ax\|^2 = \langle Ax, Ax\rangle = \langle A^T Ax, x\rangle$, where $\langle x, y\rangle = x^T y$, the standard scalar product of $x, y \in \mathbb{R}^n$. Recall that a symmetric matrix has an orthonormal basis consisting of eigenvectors.)

(b) Using (a), show that for a (2×2)-matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad a, b, c, d \in \mathbb{R},$$

the operator norm satisfies

$$\|A\|^2 = \frac{1}{2}\left[a^2 + b^2 + c^2 + d^2 + \sqrt{[(a-d)^2 + (b+c)^2][(a+d)^2 + (b-c)^2]}\right].$$

(c) Apply the formula derived in (b) to prove that

$$\|A^{-1}\| = \frac{1}{|\det A|}\|A\|$$

for any nonsingular 2×2 real matrix A. Is this result valid for $n > 2$?

E 2.5.6 (Fibonacci numbers)

Define a sequence of integers by the recurrence relation

$$a_k = a_{k-1} + a_{k-2}, \quad k \geq 2,$$

with initial condition $a_0 = a_1 = 1$. Find an explicit expression for a_k.

Hint: Write

$$\begin{pmatrix} a_k \\ a_{k+1} \end{pmatrix} = A \begin{pmatrix} a_{k-1} \\ a_k \end{pmatrix},$$

where

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix},$$

and consider a discrete-time dynamical system generated by the linear map $x \mapsto Ax$ in $\mathbb{R}^2$ (cf. Exercise 1.5.6 in Chapter 1). Find eigenvalues λ and μ and corresponding eigenvectors v and w of A and compute the associated projection operator(s).

Answer :

$$a_k = \frac{1}{2^{k+1}\sqrt{5}}\left[(1+\sqrt{5})^{k+1} - (1-\sqrt{5})^{k+1}\right].$$

E 2.5.7 (Quantum oscillations)

In Quantum Mechanics, a system with two observable states is characterized by a vector

$$\psi = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \in \mathbb{C}^2,$$

where a_k are complex numbers called *amplitudes*, satisfying the condition

$$|a_1|^2 + |a_2|^2 = 1.$$

The probability of finding the system in the kth state is equal to $p_k = |a_k|^2, k = 1, 2$.

The behaviour of the system between observations is governed by the *Heisenberg equation*:

$$i\hbar\frac{d\psi}{dt} = H\psi,$$

where the real symmetric matrix

$$H = \begin{pmatrix} E_0 & -A \\ -A & E_0 \end{pmatrix}, \quad E_0, A > 0,$$

is the Hamiltonian matrix of the system and $\hbar$ is Planck's constant divided by 2π.

(i) Write the Heisenberg equation as a system of two linear complex differential equations for the amplitudes.

(ii) Integrate the obtained system and show that the probability p_k of finding the system in the kth state oscillates periodically in time.

(iii) How does $p_1 + p_2$ behave?

E 2.5.8 (Jordan blocks)

Write the Jordan block from (2.6) as $J = \lambda I + N$ with the matrix

$$N = \begin{pmatrix} 0 & 1 & & & 0 \\ & 0 & 1 & & \\ & & & & \\ & & & 0 & 1 \\ 0 & & & & 0 \end{pmatrix}$$

(nilpotent since $N^{n_j - 1} = 0$) and use the binomial formula to find expressions for J^k and $\exp(tJ) = \sum_{k=0}^{\infty} \frac{(tJ)^k}{k!}$.

E 2.5.9 (Jordan decomposition)

It is known[3] that each $n \times n$ complex matrix A can be uniquely written as

$$A = S + N,$$

where S is diagonal in some basis (*semisimple*) and $N^m = 0$ for some $m < n$ (*nilpotent*), and such that

$$SN = NS$$

(Jordan decomposition).

Let P_i be the spectral projectors onto the generalized eigenspaces corresponding to the distinct eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_d$ of A.

(i) Prove that $S = \sum_{i=1}^{d} \lambda_i P_i$ is semisimple.

(ii) Define $N = A - S$ and prove that it is nilpotent.

(iii) Using the definition of $\exp(tJ)$ from Exercise 2.5.8, prove that

$$\exp(tS) = \sum_{i=1}^{d} e^{\lambda_i t} P_i, \quad \text{and} \quad \exp(tN) = \sum_{j=1}^{m-1} \frac{t^j}{j!} N^j \ .$$

(iv) Show that

$$\exp(tA) = \sum_{i=1}^{d} e^{\lambda_i t} \left(\sum_{j=1}^{m-1} \frac{t^j}{j!} N^j \right) P_i.$$

(v) Combine the formula from step (iv) with the partial fraction decomposition method to compute P_i and obtain an efficient algorithm for finding flows of linear autonomous systems.

E 2.5.10 (Stability and eigenvalues for maps)

[3]See, for example, Section 6.2 in Hirsch, M.W. and Smale, S. *Differential Equations, Dynamical Systems, and Linear Algebra.* Academic Press, New York-London, 1974.

(i) Extend Theorem 2.9 by showing that asymptotic stability of the origin implies that all eigenvalues of A lie strictly inside the unit circle.
(ii) Prove that for a map $x \mapsto Ax$ the following are equivalent
(a) The origin is stable;
(b) Each eigenvalue either lies strictly inside the unit circle or lies on the unit circle but has no generalized eigenvector;
(c) There exists a norm $\|\cdot\|_1$ in $\mathbb{R}^n$ such that $\|A\|_1 \leq 1$.

Hint: (a) $\Rightarrow$ (b) follows from Theorem 2.2 and for (b)$\Rightarrow$ (c) prove the boundedness of $A^k (k \geq 0)$ and define $\|x\|_1 = \sup_{k\geq 0} \|A^k x\|$.

E 2.5.11 (Stability and eigenvalues for flows)

Make an analog of Exercise 2.5.10 for ODE's $\dot{x} = Ax$.
(i) Asymptotic stability of the origin implies that all eigenvalues of A have negative real part.
(ii)Lyapunov stability holds if and only if all eigenvalues either have negative real part or have zero real part but no generalized eigenvectors. Another equivalent condition says that there exists an equivalent norm $\|\cdot\|_1$ in $\mathbb{R}^n$ such that $\|e^{tA}\|_1 \leq 1$ for all $t \geq 0$.

Hint: For the construction of a proper norm use $\|x\|_1 = \sup_{t\geq 0} \|e^{tA}x\|$.

E 2.5.12 (Asymptotic stability and contraction)

Consider two maps in $\mathbb{R}^2$ generated by the matrices

$$A = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{4} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ 1 & \frac{1}{4} \end{pmatrix}.$$

(i) Show that the origin is an asymptotically stable fixed points for both maps.

(ii) Which of these maps is a linear contraction with respect to the standard norm $\|x\| = \sqrt{x_1^2 + x_2^2}$ in $\mathbb{R}^2$?

Hint: Consider the images AU and BU of the unit disk

$$U = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}$$

under the maps A and B, respectively. Show that AU is located strictly inside U, while BU does not have this property (see Figure 2.10).

(iii) Construct several first elements of the sequences of the images

$$U, AU, A^2U, A^3U, \ldots$$

and

$$U, BU, B^2U, B^3U, \ldots$$

and convince yourself that both sequences contract towards the origin.

E 2.5.13 (Asymptotic stability and monotonicity)

Consider a linear ODE

$$\dot{x} = Ax, \quad x \in \mathbb{R}^2, \tag{2.34}$$

defined by the matrix

$$A = \begin{pmatrix} -\frac{2}{5} & 2 \\ -\frac{1}{2} & -\frac{2}{5} \end{pmatrix}.$$

(i) Prove that $x = 0$ is an asymptotically stable equilibrium of (2.34). What kind of equilibrium is it?

(ii) By numerical integration, show that the orbit starting at $x_0 = (1, 0)^T$ crosses the unit circle $\|x\|^2 = x_1^2 + x_2^2 = 1$ several times (see Figure 2.11 (a)). Prove this analytically.

(iii) Define $\|x\|_2^2 = x_1^2 + 4x_2^2$. Prove that $\|\cdot\|_2$ is a norm in $\mathbb{R}^2$ that is equivalent to $\|\cdot\|$.

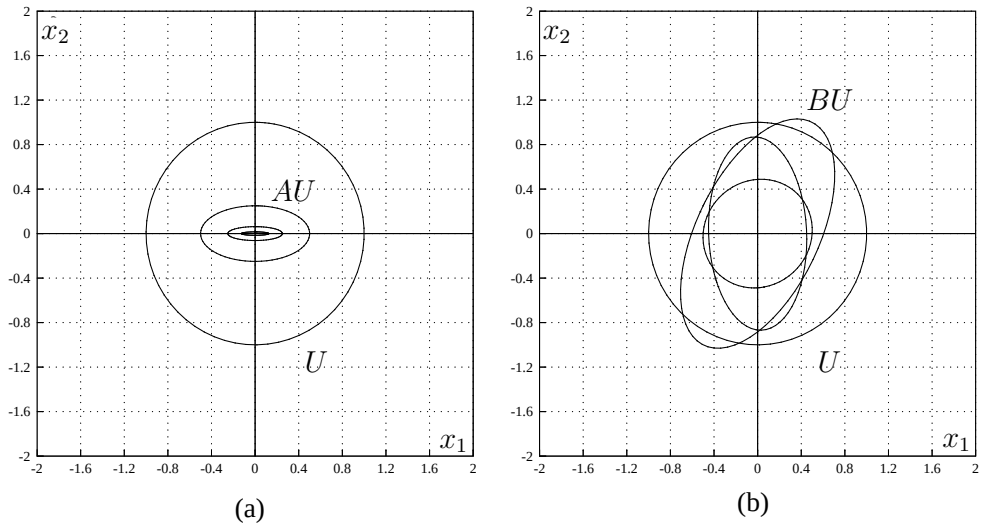

Figure 2.10: Asymptotically stable fixed points: (a) a linear contraction; (b) not a contraction.

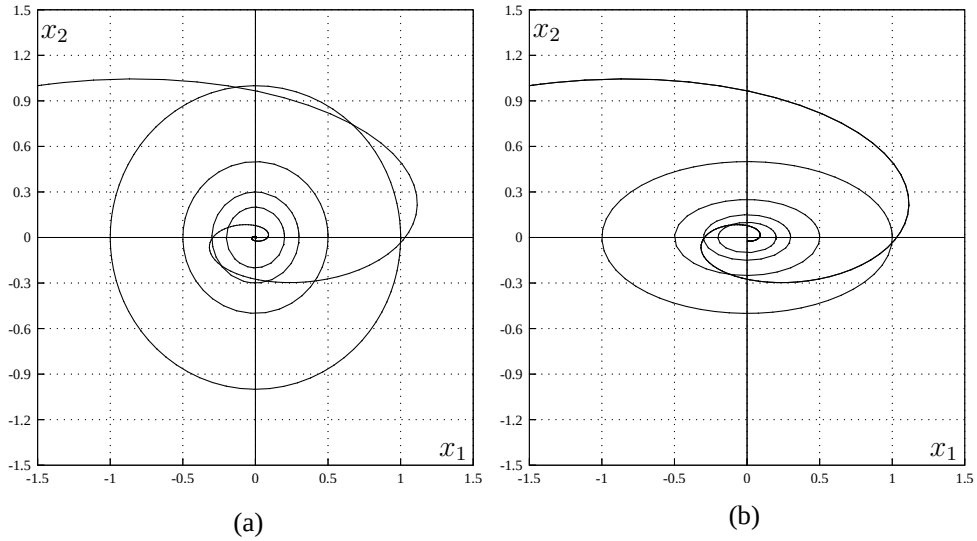

Figure 2.11: Nonmonotone (a) and monotone (b) convergence.

(iv) Prove that any "circle" $\|x\|_2 = R_0 > 0$ is crossed by any orbit of the system only once (see Figure 2.11 (b)).

Hint: Show that

$$\frac{d}{dt}\|x(t)\|_2^2 = -\frac{4}{5}\|x(t)\|_2^2$$

along any nonequilibrium solution of (2.34). Find $R(t) = \|x(t)\|_2$ by integrating this differential equation and demonstrate that it converges to zero monotonously. Is $\|\cdot\|_2$ a Lyapunov norm for (2.34)?

E 2.5.14 (Orientation and phase portraits)

In the two dimensional focus case derive a criterion for the matrix coefficients that characterizes whether orbits are rotating clockwise or counterclockwise. Compare with the phase portraits from the previous exercise.

E 2.5.15 (Explicit conjugating homeomorphism)

Verify that map $h : \mathbb{R} \to \mathbb{R}$ defined by the formula

$$h(x) = \begin{cases} x^\nu, & x \geq 0, \\ -|x|^\nu, & x < 0, \end{cases}$$

defines a homeomorphism conjugating the maps f and g from Example 2.27 for a suitable ν (which one ?).

E 2.5.16 (Inverse map in Lemma 2.28) *Compute the inverse of the map $x \mapsto (v, t)$ defined in the proof of Lemma 2.28.*

Answer:

$$(v, t) \mapsto \left(t\left(1 - \frac{1}{\|A^{-1}v\|}\right) + \frac{1}{\|A^{-1}v\|}\right) v.$$

E 2.5.17 (Connectedness of contractions)

Show that the set of contracting and positively oriented matrices $M = \{A : \det(A) > 0, r(A) < 1\}$ is connected, i.e. for any two $A, B \in M$ there is a continuous path $M(t) \in M, 0 \leq t \leq 1$ such that $M(0) = A$ and $M(1) = B$.

Hint: It is sufficient to connect any A to the special matrix $B = \frac{1}{2}I \in M$. If A has no negative eigenvalues, then take the line $tB + (1-t)A$. If A has negative eigenvalues, they must occur in pairs. Isolate the invariant subspaces belonging to them and move the eigenvalues to the positive real axis by a suitable rotation.