LECTURES 28 AND 29

Dates of the Lecture: April 21 and 23, 2021

As always, $\mathbf{K} \in \{\mathbf{R},\, \mathbf{C}\}$.

The symbol ☡ is for flagging a cautionary comment or a tricky argument. It occurs in the margins and is Knuth's version of Bourbaki's "dangerous bend symbol".

An n-tuple $(x_1, \ldots, x_n)$ of symbols (x_i not necessarily real or complex numbers) will also be written as a column vector when convenient. Thus

$$(x_1, \ldots, x_n) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

A map $\boldsymbol{f}$ from a set S to a product set $T_1 \times \cdots \times T_n$ will often be written as an n-tuple $\boldsymbol{f} = (f_1, \ldots, f_n)$, with f_i a map from S to T_i, and hence, by the above convention, as a column vector

$$\boldsymbol{f} = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}.$$

(See Remark 2.2.2 of Lecture 5 of ANA2.)

The default norm on Euclidean spaces of the form $\mathbf{R}^n$ is the Euclidean norm $\|\ \|_2$ and we will simply denote it as $\|\ \|$. The space of $\mathbf{K}$-linear transformations from $\mathbf{K}^n$ to $\mathbf{K}^m$ will be denoted $\mathrm{Hom}_{\mathbf{R}}(\mathbf{K}^n, \mathbf{K}^m)$ and will be identified in the standard way with the space of $m \times n$ matrices $M_{m,n}(\mathbf{K})$ and the operator norm[1] on both spaces will be denoted $\|\ \|_\circ$. If $m = n$, we write $M_n(\mathbf{R})$ for $M_{m,n}(\mathbf{R})$, and $L(\mathbf{K}^n)$ for $\mathrm{Hom}_{\mathbf{K}}(\mathbf{K}^n,\, \mathbf{K}^n)$.

Note that $(x_1, \ldots, x_n) \neq [x_1 \ldots x_n]$. Each side is the transpose of the other.

1. Canonical forms again

1.1. **Basic Decomposition.** Let $A\colon E \to E$ be an linear endomorphism on a finite dimensional real vector space E. We are interested in establishing the basic decompositions (1.1.5) and (1.1.6) below for A.

Let $\dim_{\mathbf{R}} E = n$. We can find an oredered basis $\mathscr{B} = \{\boldsymbol{u}_i\}_{i=1}^n$ of E such that the matrix M_A of A with respect to $\mathscr{B}$ is in a real canonical form (see Lecture 11). Let $\sigma(A) \subset \mathbf{C}$ be the set of eigenvalues of A (including non-real eigenvalues). Then each column of the matrix M_A is associated to a member λ of $\sigma(A)$. In greater

[1] See §§2.1 of Lecture 5 of ANA2.

detail

$$M_A = \begin{bmatrix} J_1 & 0 & & & 0 \\ & J_2 & 0 & & 0 \\ & & J_3 & \ddots & \\ & & & \ddots & 0 \\ & & & & J_t \end{bmatrix} \tag{1.1.1}$$

where J_i is either of the form

$$J = \begin{bmatrix} M & I_2 & & & 0 \\ & M & I_2 & & \\ & & M & \ddots & \\ & & & \ddots & I_2 \\ & & & & M \end{bmatrix} \tag{1.1.2}$$

with $M = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, $b > 0$, and $\lambda = a + ib$ an eigenvalue of A, or J_i is of the form

$$J = \begin{bmatrix} \lambda & 1 & & & 0 \\ & \lambda & 1 & & \\ & & \lambda & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda \end{bmatrix}. \tag{1.1.3}$$

with $\lambda \in \mathbf{R}$ an eigenvalue of A. Moreover, every eigenvalue of A has some J associated with it.

Let $i \in \{1, \ldots, n\}$. The i^{th} column of M_A meets exactly one J_μ in the representation (1.1.1). If J_μ is of the form (1.1.2) then let $e(i) = a + ib$. If J_μ is of the form (1.1.3), then set $e(i) = \lambda$. We therefore have a surjective map $e\colon \{1, \ldots, n\} \to \sigma(A)$. Then $\mathscr{B}$ is the disjoint union of $\mathscr{B}^+$ and $\mathscr{B}^-$ where $\mathscr{B}^+ = \{\boldsymbol{u}_i \in \mathscr{B} \mid \operatorname{Re} e(i) > 0\}$ and $\mathscr{B}^- = \mathscr{B} \smallsetminus \mathscr{B}^+$. Let

$$E^+ = \operatorname{Span}\langle \boldsymbol{u} \mid \boldsymbol{u} \in \mathscr{B}^+\rangle \quad \text{and} \quad E^- = \operatorname{Span}\langle \boldsymbol{u} \mid \boldsymbol{u} \in \mathscr{B}^-\rangle. \tag{1.1.4}$$

Then

$$E = E^+ \oplus E^- \tag{1.1.5}$$

It is clear that $A(E^+) \subset E^+$ and $A(E^-) \subset E^-$. We therefore have $\mathbf{R}$-linear endomorphisms

$$A^+\colon E^+ \longrightarrow E^+ \quad \text{and} \quad A^-\colon E^- \to E^- \tag{1.1.6}$$

with $A^+ = A|_{E^+}$ and $A^- = A|_{E^-}$.

1.1.7. Exercise. Consider the matrix J in (1.1.2). It is of size $2k \times 2k$. Let $\mathbf{e}_1, \mathbf{e}'_1, \mathbf{e}_2, \mathbf{e}'_2, \ldots, \mathbf{e}_k, \mathbf{e}'_k$ be the basis on a $2k$ dimensional $\mathbf{R}$-vector space V such that J is the matrix of a linear transformation $T\colon V \to V$. Let ε be any positive real number. For $i = 1, \ldots, k$ set $\boldsymbol{v}_i = \frac{1}{\varepsilon^{k-i}}\mathbf{e}_i$ and $\boldsymbol{v}'_i = \frac{1}{\varepsilon^{k-i}}\mathbf{e}'_i$. Show that with respect to the basis $\boldsymbol{v}_1, \boldsymbol{v}'_1, \ldots, \boldsymbol{v}_k, \boldsymbol{v}'_k$ of V, the matrix of T is

$$\begin{bmatrix} M & \varepsilon I_2 & & & 0 \\ & M & \varepsilon I_2 & & \\ & & M & \ddots & \\ & & & \ddots & \varepsilon I_2 \\ & & & & M \end{bmatrix} \tag{1.1.7.1}$$

Next, consider the matrix J in (1.1.3). Say it is $k \times k$. Let $\mathbf{e}_1, \ldots, \mathbf{e}_k$ be a basis of an $\mathbf{R}$-vector space V such that J is the matrix of a linear transformation $T\colon V \to V$ with respect to this basis. For $i = 1, \ldots, k$ set $\boldsymbol{v}_i = \frac{1}{\varepsilon^{k-i}}\mathbf{e}_i$. Show that the matrix of T with respect to the basis $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ is

$$\begin{bmatrix} \lambda & \varepsilon & & & 0 \\ & \lambda & \varepsilon & & \\ & & \lambda & \ddots & \\ & & & \ddots & \varepsilon \\ & & & & \lambda \end{bmatrix}. \tag{1.1.7.2}$$

Finally show that if $A\colon E \to E$ is the linear transformation we are considering, there is a basis for E such that the matrix of A is of the form:

$$M_A(\varepsilon) = \begin{bmatrix} J_{1,\,\varepsilon} & 0 & & & 0 \\ & J_{2,\,\varepsilon} & 0 & & 0 \\ & & J_{3,\,\varepsilon} & \ddots & \\ & & & \ddots & 0 \\ & & & & J_{t,\,\varepsilon} \end{bmatrix} \tag{1.1.7.3}$$

where the $J_{i,\,\varepsilon}$ are of the form (1.1.7.1) or of the form (1.1.7.2).

1.1.8. The matrix $M_A(\varepsilon)$ in (1.1.7.3) can be re-written as

$$M_A(\varepsilon) = \begin{bmatrix} B_1 & 0 & & & 0 \\ & B_2 & 0 & & 0 \\ & & B_3 & \ddots & \\ & & & \ddots & 0 \\ & & & & B_\mu \end{bmatrix} + \varepsilon \begin{bmatrix} 0 & C_1 & & & 0 \\ & 0 & C_2 & & \\ & & 0 & \ddots & \\ & & & \ddots & C_{\mu-1} \\ & & & & 0 \end{bmatrix} \tag{1.1.8.1}$$

where B_i and C_i are of the same size, and B_i and C_i are either 2×2 matrices or 1×1 matrices. If, for an i, B_i and C_i are 2×2 then B_i is of the form $\left[\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right]$, $b > 0$, and C_i is either 0 or the identity matrix I_2. In this case $\lambda = a + ib$ is an eigenvalue of A. If B_i is 1×1, say $B_i = [\lambda]$, then C_i is either 0 or the 1×1 identity matrix [1]. In this case, clearly λ is a real eigenvalue of A. Every eigenvalue of A appears in this process. We re-write the above decomposition as

$$M_A(\varepsilon) = B + \varepsilon C. \tag{1.1.8.2}$$

1.2. **Crucial Lemma.** Let $A\colon E \to E$ be as in §§ 1.1. As before, let $\sigma(A)$ be the set of eigenvalues of A (including non-real eigenvalues).

Lemma 1.2.1. *Let $c < \min\{\mathrm{Re}(\lambda) \mid \lambda \in \sigma(A)\}$. Then there exists an inner product $\langle -, - \rangle$ on E such that*

$$\langle \boldsymbol{x}, A\boldsymbol{x}\rangle \geq c\|\boldsymbol{x}\|^2 \quad (\boldsymbol{x} \in E).$$

The norm appearing on the right side is the norm given by $\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x}\rangle}$.

Proof. First note that

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = a(x_1^2 + x_2^2).$$

Choose a basis $\{\boldsymbol{w}_i\}$ so that the matrix M_A of A with respect to this basis is of the form (1.1.8.1). Use this basis to define the inner product $\langle -, - \rangle$ on E. In other words, $\langle -, - \rangle$ is such that $\langle \boldsymbol{w}_i, \boldsymbol{w}_j\rangle = \delta_{ij}$. Let the norm on E be induced by this inner prduct.

Let $K = \|C\|_\circ$. Then using the above equality we see that

$$\begin{aligned} \langle \boldsymbol{x}, A\boldsymbol{x}\rangle &= \langle \boldsymbol{x}, B\boldsymbol{x}\rangle + \varepsilon\langle \boldsymbol{x}, C\boldsymbol{x}\rangle \\ &\geq \min_{\lambda\in\sigma(A)} \mathrm{Re}(\lambda)\|\boldsymbol{x}\|^2 + \varepsilon\langle \boldsymbol{x}, C\boldsymbol{x}\rangle \\ &\geq \min_{\lambda\in\sigma(A)} \mathrm{Re}(\lambda)\|\boldsymbol{x}\|^2 - \varepsilon\|C\|_\circ\|\boldsymbol{x}\|^2 \\ &> c\|\boldsymbol{x}\|^2 \end{aligned}$$

for $\varepsilon > 0$ so small that $\varepsilon\|C\|_\circ < \min_{\lambda\in\sigma(A)} \mathrm{Re}(\lambda) - c$. ☐

Corollary 1.2.2. *If $c > \max\{\mathrm{Re}(\lambda) \mid \lambda \in \sigma(A)\}$, then there exists an inner product $\langle -, - \rangle$ on E such that*

$$\langle \boldsymbol{x}, A\boldsymbol{x}\rangle \leq c\|\boldsymbol{x}\|^2 \quad (\boldsymbol{x} \in E).$$

where, as before, the norm appearing on the right is the norm induced by $\langle -, - \rangle$.

2. Linearisation and Stability

2.1. **Linearisation.** Let $\boldsymbol{v}$ be an $\mathscr{C}^1$ vector field on an open subset U of $\mathbf{R}^n$. Let $\boldsymbol{x}_0$ be an equilibrium point for $\boldsymbol{v}$. Let $D\boldsymbol{v}\colon U \to L(\mathbf{R}^n)$ denote the derivative of $\boldsymbol{v}$, and finally set

$$A = (D\boldsymbol{v})(\boldsymbol{x}_0).$$

Then the differential equation

$$\dot{\boldsymbol{x}} = A(\boldsymbol{x} - \boldsymbol{x}_0) \tag{2.1.1}$$

is called the linearisation of

$$\dot{\boldsymbol{x}} = \boldsymbol{v}(\boldsymbol{x}). \tag{2.1.2}$$

2.2. **Stability and linearisation.** Let $\boldsymbol{v}$ be $\mathscr{C}^2$ (notice the stronger smoothness condition than what we had in §§2.1), and let $\boldsymbol{x}_0$, A, etc. be as above. We wish to understand the solutions of (2.1.2) in terms of solutions of (2.1.1), at least when the initial phase $\boldsymbol{a}$ is near $\boldsymbol{x}_0$.

Theorem 2.2.1. *Suppose $\boldsymbol{v}$ is $\mathscr{C}^2$, and A has has eigenvalue whose real part is positive. Then $\boldsymbol{x}_0$ is an unstable equilibrium point of $\boldsymbol{v}$.*

Proof. Without loss of generality, let us assume $\boldsymbol{x}_0 = \mathbf{0}$. Then (2.1.1) is a homogeneous linear differential equation with constant coefficients. Using A, from (1.1.5) we have a decomposition $\mathbf{R}^n = E^+ \oplus E^-$ such that $A(E^+) \subset E^+$ and $A(E^-) \subset E^-$. As in (1.1.6) we therefore have $A^+ \colon E^+ \to E^+$ and $A^- \colon E^- \to E^-$ with $A = \begin{bmatrix} A^+ & 0 \\ 0 & A^- \end{bmatrix}$. Let, as before, $\sigma(A^\pm)$ denote the set of eigenvalues of $A^\pm$, According to our hypothesis, $E^+ \neq \mathbf{0}$. In particular $\min\{\mathrm{Re}(\lambda) \mid \lambda \in \sigma(A^+)\} > 0$. Choose positive numbers a and b such that

$$\min\{\mathrm{Re}(\lambda) \mid \lambda \in \sigma(A^+)\} > a > b > 0.$$

We point out that

$$\max\{\mathrm{Re}(\lambda) \mid \lambda \in \sigma(A^-)\} < b.$$

By Lemma 1.2.1, we can find inner products $\langle -, - \rangle_+$ on E^+ and $\langle -, - \rangle_-$ on E^- such that $\langle \boldsymbol{y}, A^+\boldsymbol{y} \rangle \geq a\|\boldsymbol{y}\|_+^2$ and $\langle \boldsymbol{z}, A^-\boldsymbol{z} \rangle \leq b\|\boldsymbol{z}\|_-^2$, where $\|\ \|_\pm$ are the norms on $E^\pm$ induced by $\langle -, - \rangle_\pm$. Let $\langle -, - \rangle$ be the unique inner product on $\mathbf{R}^n$ which when restricted tp E^+ is $\langle -, - \rangle_+$, and when restricted to E^- is $\langle -, - \rangle_-$ and such that the decomposition $\mathbf{R}^n = E^+ \oplus E^-$ in (1.1.5) is an orthogonal decomposition.

Every element $\boldsymbol{x} \in \mathbf{R}^n$ can be written as $\boldsymbol{x} = (\boldsymbol{y}, \boldsymbol{z})$ in a unique way via the decomposition $\mathbf{R}^n = E^+ \oplus E^-$. Similarly $\boldsymbol{v}$ breaks up into $\boldsymbol{v} = (\boldsymbol{u}, \boldsymbol{w})$ In other words $\boldsymbol{v}(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{u}(\boldsymbol{a}, \boldsymbol{b}), \boldsymbol{w}(\boldsymbol{a}, \boldsymbol{b}))$ for $(\boldsymbol{a}, \boldsymbol{b}) \in U$. Since $\boldsymbol{v}$ is $\mathscr{C}^2$, by Taylor's theorem we have in a neighbourhood of $\boldsymbol{x}_0 = \mathbf{0}$

$$\boldsymbol{v}(x) = A\boldsymbol{x} + \boldsymbol{R}(\boldsymbol{x})$$

with

$$\lim_{\boldsymbol{x} \to \mathbf{0}} \frac{\boldsymbol{R}(\boldsymbol{x})}{\|\boldsymbol{x}\|} = \mathbf{0}.$$

Now $\boldsymbol{R} = (\boldsymbol{R}^+, \boldsymbol{R}^-)$ with $\boldsymbol{R}^\pm(\boldsymbol{x})/\|\boldsymbol{x}\| \to \mathbf{0}$ as $\boldsymbol{x} \to \mathbf{0}$.

Write $\{g^t\}$ for the flow on U given by the vector field $\boldsymbol{v}$. Note that g^t can be further broken up as $g^t = (f^t, h^t)$ under the decomposition $\mathbf{R}^n = E^+ \oplus E^-$. We will use this notation. Note that

$$\boldsymbol{u}(g^t\boldsymbol{x}) = A^+(f^t\boldsymbol{x}) + R^+(\boldsymbol{x}),$$

and

$$\boldsymbol{w}(g^t\boldsymbol{x}) = A^-(h^t\boldsymbol{x}) + r^-(\boldsymbol{x}).$$

Pick $\eta > 0$ such that $a - b - 2\sqrt{2}\eta > 0$. Let $\varepsilon > 0$ be such that

$$\|\boldsymbol{R}^\pm(\boldsymbol{x})\| \leq \eta\|\boldsymbol{x}\| \qquad (\boldsymbol{x} \in B(\mathbf{0}, \varepsilon)).$$

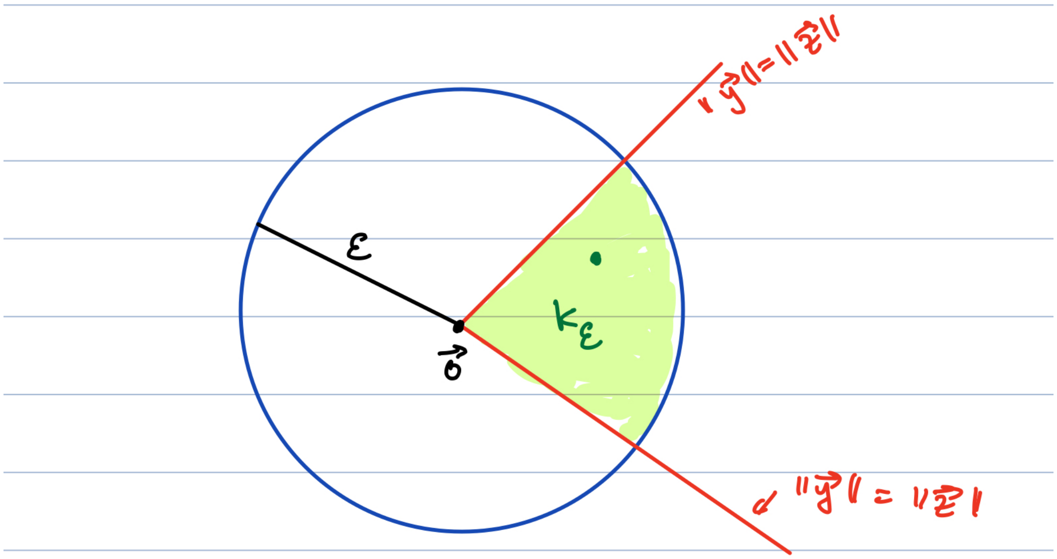
Consider the restricted cone:

$$K_\varepsilon = \{(\boldsymbol{y}, \boldsymbol{z}) \in B(\mathbf{0}, \varepsilon) \mid \|\boldsymbol{y}\| > \|\boldsymbol{z}\|\}.$$

Let $\boldsymbol{x} \in K_\varepsilon$ and let $T = T_{\boldsymbol{x}}$ be the first instant t that $g^t\boldsymbol{x}$ exits the restricted cone K_ε. In other words,

$$T_{\boldsymbol{x}} = \inf\{t \in J_{\geqslant 0}(\boldsymbol{x}) \mid g^t\boldsymbol{x} \notin K_\varepsilon\}.$$

Here $J_{\geqslant \geq 0}(\boldsymbol{x}) = [0, \infty) \cap J(\boldsymbol{x})$, where, for $\boldsymbol{a} \in U$, $J(\boldsymbol{a})$ is the maximal interval of existence of the solution $t \mapsto g^t\boldsymbol{x}$ of $\dot{\boldsymbol{x}} = \boldsymbol{v}(\boldsymbol{x})$, $\boldsymbol{x}(0) = \boldsymbol{a}$. If the set $\{t \in J_{\geqslant 0}(\boldsymbol{x}) \mid g^t\boldsymbol{x} \notin K_\varepsilon\}$ is empty, then we set $T_{\boldsymbol{x}} = \infty$, consistent with the convention that the infimum of the empty set is ∞. The following picture may be useful in following the rest of the proof.



We claim that $T_{\boldsymbol{x}} < \infty$ and that $g^{T_{\boldsymbol{x}}}\boldsymbol{x} \notin B(\boldsymbol{0},\, \varepsilon)$ (recall, we have picked $\boldsymbol{x} \in K_\varepsilon$). For $0 \le t < T_{\boldsymbol{x}}$ we have (using Lemma 1.2.1 to get $\langle \boldsymbol{y},\, A^+\boldsymbol{y}\rangle \ge a\|\boldsymbol{y}\|^2$ for all $\boldsymbol{y} \in E^+$)

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\|f^t\boldsymbol{x}\|^2 &= 2\langle f^t\boldsymbol{x},\, \tfrac{\mathrm{d}}{\mathrm{d}t}(f^t\boldsymbol{x})\rangle \\
&= 2\langle f^t\boldsymbol{x},\, \boldsymbol{u}(g^t\boldsymbol{x})\rangle \\
&= 2\langle f^t\boldsymbol{x},\, A^+(f^t\boldsymbol{x})\rangle + 2\langle f^t\boldsymbol{x},\, \boldsymbol{R}^+(g^t\boldsymbol{x})\rangle \\
&\ge 2a\|f^t\boldsymbol{x}\|^2 - 2\|f^t\boldsymbol{x}\|\|\boldsymbol{R}^+(g^t\boldsymbol{x})\| \\
&\ge 2a\|f^t\boldsymbol{x}\|^2 - 2\eta\|f^t\boldsymbol{x}\|(\|g^t\boldsymbol{x}\|) \\
&= 2a\|f^t\boldsymbol{x}\|^2 - 2\eta\|f^t\boldsymbol{x}\|\Big(\|f^t\boldsymbol{x}\|^2 + \|h^t\boldsymbol{x}\|^2\Big)^{\frac{1}{2}} \\
&\ge 2a\|f^t\boldsymbol{x}\|^2 - 2\sqrt{2}\eta\|f^t\boldsymbol{x}\|^2 \\
&= 2(a - \sqrt{2}\eta)\|f^t\boldsymbol{x}\|^2 \\
&> 0,
\end{aligned}$$

the last inequality following from the fact that $a - \sqrt{2}\eta > b + \sqrt{2}\eta > 0$. This shows that for $t \in [0,\, T_{\boldsymbol{x}})$, we have $\frac{\mathrm{d}}{\mathrm{d}t}\|f^t\boldsymbol{x}\|^2 > 2a\|f^t\boldsymbol{x}\|^2 - 2\sqrt{2}\eta\|f^t\boldsymbol{x}\|^2$. In other words, $t \mapsto \|f^t\boldsymbol{x}\|$ has exponential growth in $[0,\, T_{\boldsymbol{x}})$. It follows that $T_{\boldsymbol{x}} < \infty$, for, if $T_{\boldsymbol{x}} = \infty$ then for all $t \ge 0$ we have $\|f^t\boldsymbol{x}\| \le \|g^t\boldsymbol{x}\| < \varepsilon$, contradicting the exponential growth of $\|f^t\boldsymbol{x}\|$. Similarly we have (again using Lemma 1.2.1 to conclude that $\langle z,\, A^{-}\boldsymbol{z}\rangle \leq b\|\boldsymbol{z}\|^2$ for all $\boldsymbol{z} \in E^{-}$)

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\|h^t\boldsymbol{x}\|^2 &= 2\langle h^t\boldsymbol{x},\, \tfrac{\mathrm{d}}{\mathrm{d}t}(h^t\boldsymbol{x})\rangle \\
&= 2\langle h^t\boldsymbol{x},\, \boldsymbol{w}(g^t\boldsymbol{x})\rangle \\
&= 2\langle h^t\boldsymbol{x},\, A^{-}(h^t\boldsymbol{x})\rangle + 2\langle h^t\boldsymbol{x},\, \boldsymbol{R}^{-}(g^t\boldsymbol{x})\rangle \\
&\leq 2b\|h^t\boldsymbol{x}\|^2 + 2\|h^t\boldsymbol{x}\|\|\boldsymbol{R}^{-}(g^t\boldsymbol{x})\| \\
&\leq 2b\|h^t\boldsymbol{x}\|^2 + 2\eta\|h^t\boldsymbol{x}\|(\|g^t\boldsymbol{x}\|) \\
&\leq 2b\|h^t\boldsymbol{x}\|^2 + 2\eta\|f^t\boldsymbol{x}\|\Big(\|f^t\boldsymbol{x}\|^2 + \|h^t\boldsymbol{x}\|^2\Big)^{\frac{1}{2}} \\
&\leq 2b\|f^t\boldsymbol{x}\|^2 + 2\sqrt{2}\eta\|f^t\boldsymbol{x}\|^2 \\
&= 2(b+\sqrt{2}\eta)\|f^t\boldsymbol{x}\|^2.
\end{aligned}$$

Let $F\colon \mathbf{R}^n = E^{+} \oplus E^{-} \longrightarrow \mathbf{R}$ be the map given by the formula

$$F(\boldsymbol{a}, \boldsymbol{b}) = \|\boldsymbol{a}\|^2 - \|\boldsymbol{b}\|^2.$$

Let L be the lateral "surface" in $B(\mathbf{0},\, \varepsilon)$,

$$L = \{(\boldsymbol{a},\, \boldsymbol{b}) \in B(\mathbf{0},\, \varepsilon) \;\mid\; \|\boldsymbol{a}\| = \|\boldsymbol{b}\|\}.$$

For $t \in [0,\, T_{\boldsymbol{x}})$ we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}F(g^t\boldsymbol{x}) = \frac{\mathrm{d}}{\mathrm{d}t}F(f^t\boldsymbol{x},\, h^t\boldsymbol{x}) &= \frac{\mathrm{d}}{\mathrm{d}t}\|f^t\boldsymbol{x}\|^2 - \frac{\mathrm{d}}{\mathrm{d}t}\|h^t\boldsymbol{x}\|^2 \\
&\geq 2(a - b - \sqrt{2}\eta)\|f^t\boldsymbol{x}\|^2 \\
&> 0.
\end{aligned}$$

Thus $t \mapsto F(g^t\boldsymbol{x})$ is an increasing function on $[0,\, T_{\boldsymbol{x}})$, and at $t = 0$ it is positive. Now $F|_L \equiv 0$. It follows that that $g^t\boldsymbol{x}$ must exit K_ϵ from the sphere $S(\mathbf{0},\, \varepsilon)$ and not via the L.

Since $\boldsymbol{x} \in K_\varepsilon$ can be chosen to have arbitrarily small norm, this shows that $\boldsymbol{x}_0$ is an unstable equilibrium point. □

2.3. **Hyperbolic equilibrium points.** The qualitative behaviour of the flows some $\mathscr{C}^2$ vector fields near an equilibrium point mimic the behaviour of the flows of the corresponding linear differential equation near that equilibrium point. Specifically, the flows of vector fields near a so-called *hyperbolic* equilibrium point are qualitatively identical to the flows of the corresponding linearised problem, via the Hartman-Grobman theorem. We will not have time to prove the Hartman-Grobman theorem, but we can give evidence of the philosophy underpinning the Hartman-Grobman theorem by proving results about stability in a neighbourhood of a hyperbolic equilibrium point. But first we need to define a hyperbolic equilibrium point.

Definition 2.3.1. Let $\boldsymbol{v}$ be a $\mathscr{C}^2$ vector field on an open subset U of $\mathbf{R}^n$ and let $\boldsymbol{x}_0$ be an equilibrium point of the corresponding dynamical system. Let $A = (D\boldsymbol{v})(\boldsymbol{x}_0)$. The equilibrium point $\boldsymbol{x}_0$ of $\boldsymbol{v}$ is said to be *hyperbolic* if none of the roots of the characteristic polynomial of A lie on the imaginary axis.

The important result is the following.

Theorem 2.3.2. *Let $\boldsymbol{v}$ be a $\mathscr{C}^2$ vector field on an open set U of $\mathbf{R}^n$ and $\boldsymbol{x}_0$ a hyperbolic equilibrium point of $\boldsymbol{v}$. Let $A = (D\boldsymbol{v})(\boldsymbol{x}_0)$ and $\sigma(A)$ the set of roots of the characteristic polynomial of A. If $\mathrm{Re}(\lambda) < 0$ for every $\lambda \in \sigma(A)$ then $\boldsymbol{x}_0$ is an asymptotically stable equilibrium point of $\boldsymbol{v}$. Otherwise, $\boldsymbol{x}_0$ is an unstable equilibrium point.*

Proof. The "otherwise" part of the statement is taken care of by Theorem 2.2.1. So now suppose all the characteristic roots of A have negative real parts. Without loss of generality, assume $\boldsymbol{x}_0 = \mathbf{0}$. Let $c > 0$ be such that

$$\max\{\mathrm{Re}(\lambda) \mid \lambda \in \sigma(A)\} < -c.$$

Here, as before, $\sigma(A)$ is the set of characteristic values of A. By Lemma 1.2.1 we have an inner product $\langle -, - \rangle$ on $\mathbf{R}^n$ such that

$$\langle \boldsymbol{x}, A\boldsymbol{x} \rangle \le -c\|\boldsymbol{x}\|^2 \quad (\boldsymbol{x} \in \mathbf{R}^n).$$

The norm on the right side is the norm induced by this (possibly new) inner product. For the rest of this proof we work with this inner product, and all norms, distances, open and closed balls, etc., will be with respect to this inner product. By Taylor's theorem

$$\boldsymbol{v}(x) = A\boldsymbol{x} + \boldsymbol{R}(\boldsymbol{x})$$

with $\|\boldsymbol{R}(\boldsymbol{x})\|/\|\boldsymbol{x}\| \to 0$ as $\boldsymbol{x} \to \mathbf{0}$. Let $\delta > 0$ be chosen so that $B(\mathbf{0}, \delta) \subset U$ and $\|\boldsymbol{R}(\boldsymbol{x})\| \le \frac{c}{2}\|\boldsymbol{x}\|$ for $\boldsymbol{x} \in B(\mathbf{0}, \delta)$. Our candidate for a strict Lyaponuv function is $Q\colon B(\mathbf{0}, \delta) \to \mathbf{R}$ given by

$$Q(\boldsymbol{x}) = \|x\|^2 \qquad (\boldsymbol{x} \in B(\mathbf{0}, \delta)).$$

Then for $\boldsymbol{x} \in B(\mathbf{0}, \delta)$ we have,

$$\begin{aligned} \dot{Q}(\boldsymbol{x}) &= \langle 2\boldsymbol{x}, \boldsymbol{v}(\boldsymbol{x}) \rangle \\ &= 2c\langle \boldsymbol{x}, A\boldsymbol{x} \rangle + 2\langle \boldsymbol{x}, \boldsymbol{R}(\boldsymbol{x}) \rangle \\ &\le -2c\|\boldsymbol{x}\|^2 + 2\|\boldsymbol{R}(\boldsymbol{x})\|\|\boldsymbol{x}\| \\ &\le -2c\|\boldsymbol{x}\|^2 + 2(\tfrac{c}{2})\|\boldsymbol{x}\|^2 \\ &\le -c\|\boldsymbol{x}\|^2. \end{aligned}$$

It follows easily that Q is a strict Lyapunov function, whence $\boldsymbol{x}_0$ is asymptotically stable. ☐

2.3.3. Pictures of flows For pictures of flows of linear differential equations with $\mathbf{R}^2$ as the phase space see [G, pp.46–50] or see the iPad version of the notes of iPad version of the notes of Lecture 29. If time permits, I will transfer these pictures to this file later.

2.4. **Further reading.** The above theorem is the last result of this course. However, it is worth drawing your attention to a related important theorem in topological dynamics. This is *not* part of the course. We begin with a definition.

Definition 2.4.1. Let $\boldsymbol{v}\colon U \to \mathbf{R}^n$ be a $\mathscr{C}^1$ vector field on an open set in $\mathbf{R}^n$ such that $\mathbf{0} \in U$ and $\boldsymbol{v}(\mathbf{0}) = \mathbf{0}$. Let $A = (D\boldsymbol{v}(\mathbf{0})$, and let $\boldsymbol{w}\colon \mathbf{R}^n \to \mathbf{R}^n$ be the vector field $\boldsymbol{x} \mapsto A\boldsymbol{x}$.[2] The vector field $\boldsymbol{v}$ is A $\mathscr{C}^1$ vector field $\boldsymbol{v}$ on an open subset U of $\mathbf{R}^n$ containing the is said to be *locally topologically conjugate* to its linearisation $\boldsymbol{w}$

[2] In other words, $\boldsymbol{w} = A$, but we write $\boldsymbol{w}$ because A does not look typographically like a vector field.

at the origin if there is a homeomorphism $h\colon N \to V$ of neighbourhoods of $\mathbf{0}$ such that

$$h(g^t \boldsymbol{x}) = e^{tA} h(\boldsymbol{x})$$

where g^t denotes the flow for $\boldsymbol{v}$ for $\boldsymbol{x} \in N$.

We often say, in the above situation, that $\boldsymbol{v}$ is locally topologically conjugate to A at the origin.

Theorem 2.4.2. [Hartman-Grobman Theorem] *Let $\boldsymbol{v}$ be a $\mathscr{C}^1$ vector field on an open subset U of $\mathbf{R}^n$ containing $\mathbf{0}$ such that $\boldsymbol{v}(\mathbf{0}) = \mathbf{0}$. If the linearisation $A = (D\boldsymbol{v})(\mathbf{0})$ of $\boldsymbol{v}$ is such that every root of the characteristic polynomial of A has a negative real part, then $\boldsymbol{v}$ is locally topologically conjugate to A at the origin.*

Proof. See [C, Theorem 4.14, pp.359–362]

References

[A1] V. I. Arnold, *Ordinary Differential Equations*, translated by Richard A. Silverman, MIT Press (also Prentice-Hall, India), Cambridge,MA, U.S.A., 1973.

[A2] V. I. Arnold, *Ordinary Differential Equations*, Third Edition, translated by Roger Cooke, Universitext, Springer-Verlag, Berlin, 2006.

[B] J. Burke, Notes of course given at University of Washington, http://sites.math.washington.edu/~burke/crs/555/555_notes/continuity.pdf.

[C] C. Chicone, *Ordinary Differential Equations with Applications*, Springer, 2006.

[CL] E.A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, 1955.

[G] C.P. Grant, *Theory of Ordinary Differential Equations*. https://www.math.utah.edu/~treiberg/GrantTodes2008.pdf, Brigham Young University.