LECTURE 10

Date of Lecture: February 3, 2019

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. The Exponential Map

1.1. Fix $n \in \mathbf{N}$. For $R > 0$, let $B_R = \{T \in L(\mathbf{K}^n) \mid \|T\|_\circ \leq R\}$. Let $T \in B_R$. Let

$$S_k(T) = \sum_{i=0}^{k} \frac{T^m}{m!}.$$

[1]See §§2.1 of Lecture 5 of ANA2.

For $0 < l \le k$ we have

$$\begin{aligned}\|S_k(T) - S_l(T)\|_\circ &= \Big\| \sum_{m=l+1}^{n} \frac{T^m}{m!} \Big\|_\circ \\ (\dagger) \qquad &\le \sum_{m=l+1}^{n} \Big\| \frac{T^m}{m!} \Big\|_\circ \\ &\le \sum_{m=l+1}^{n} \frac{R^m}{m!}\end{aligned}$$

Given $\epsilon > 0$, we can find N such that for $N \le l \le k$, $\sum_{m=l+1}^{n} \frac{R^m}{m!} < \epsilon$. Since N does not depend upon T, it follows from ($\dagger$) that $\{S_m\}$ is unformly Cauchy on B_R. Since $L(\mathbf{K}^n)$ is complete (being a finite dimensional $\mathbf{K}$-vector space), this means $\{S_m(T)\}$ converges and the convergence is uniform on B_R. Another way of saying this is that the the series $\sum_{m=0}^{\infty} \frac{T^m}{m!}$ converges uniformly on compact subsets of $L(\mathbf{K}^n)$. This allows to make the following definition.

Definition 1.1.1. For $T \in L(\mathbf{K}^n)$ the *exponential* e^T *of* T is the element of $L(\mathbf{K}^n)$ given by the formula

$$e^T = \sum_{m=0}^{\infty} \frac{T^m}{m!}.$$

Theorem 1.1.2. *The exponential series* $\sum_{m=0}^{\infty} T^m/m!$ *converges uniformly on compact subsets of* $L(\mathbf{K}^n)$. *Moreover, if* S *and* T *are elements of* $L(\mathbf{K}^n)$ *such that* $ST = TS$ *then*

$$e^{T+S} = e^T e^S.$$

Proof. We have already seen that the exponential series converges uniformly on compact sets. Now suppose S and T are commuting linear operators on $\mathbf{K}^n$. Since S and T commute, the binomial theorem applies to $(S+T)^m$ and we have

$$\sum_{i+j=m} \frac{S^i}{i!}\frac{T^j}{j!} = \frac{(S+T)^m}{m!}.$$

It follows that **(check this!)**

$$\|e^{T+S} - S_k(S)S_k(T)\|_\circ \le 2 \sum_{m=k+1}^{\infty} \frac{(\|S\|_\circ + \|T\|_\circ)^m}{m!}$$

and the sum on the right can be made as small as we wish by choosing k large. It follows that $S_k(S)S_k(T) \to e^{T+S}$ as $k \to \infty$. By **8** in §1 of Lecture 7 of ANA2 we are done. □

Example 1.1.3. Let $T \in L(\mathbf{R}^n)$. Then $\{e^{tT}\}$ is a one-parameter group of linear transformations. We already know that $e^{(t+s)T} = e^{tT}e^{sT}$ for all $s, t \in \mathbf{R}$. We thus only have to check that the map

$$g\colon \mathbf{R} \times \mathbf{R}^n \to \mathbf{R}^n, \quad (t, \boldsymbol{x}) \mapsto e^{tT}\boldsymbol{x}$$

is a $\mathscr{C}^r$. It is in fact $\mathscr{C}^\infty$ since it is given by a convergent power series which converges uniformly when t varies in a compact set. Moreover,

$$\frac{de^{tT}}{dt}\Big|_{t=0} = T.$$

Indeed

$$\frac{de^{tT}}{dt}\Big|_{t=0} - T = \lim_{h\to 0}\frac{1}{h}(e^{hT} - I_n) - T$$
$$= \lim_{h\to 0}\frac{1}{h}\Big(\sum_{m=0}^{\infty}\frac{h^m T^m}{m!} - I_n\Big) - T$$
$$= \lim_{h\to 0}\sum_{m=2}^{\infty}\frac{h^{m-1}T^m}{m!}$$

Since we wish to take a limit as $h \to 0$, we may assume $|h| \le 1$, whence $|h^{m-1}| \le |h|$ for $m \ge 2$. Hence

$$\left\|\frac{de^{tT}}{dt}\Big|_{t=0} - T\right\|_{\circ} \le \lim_{h\to 0}|h|\sum_{m=2}^{\infty}\frac{\|T\|_{\circ}^m}{m!} \le \lim_{h\to 0}|h|e^{\|T\|_{\circ}} = 0.$$

Thus the phase velocity vector of $\{e^{tT}\}$ at $\boldsymbol{x}$ is $\boldsymbol{v}(\boldsymbol{x}) = T\boldsymbol{x}$.

Theorem 1.1.4. *Let $A \in M_n(\mathbf{R})$, $\boldsymbol{a} \in R^n$ and $\boldsymbol{\varphi}\colon \mathbf{R} \to \mathbf{R}^n$ the map given by $\boldsymbol{\varphi}(t) = e^{tA}\boldsymbol{a}$. $g^t = e^{tA}$ for $t \in \mathbf{R}$. Then $\boldsymbol{\varphi}$ is the unique solution to the IVP $\dot{\boldsymbol{x}} = A\boldsymbol{x}$, $\boldsymbol{x}(0) = \boldsymbol{a}$.*

Proof. $\{e^{tA}\}$ is a one-parameter groups of linear transformations on $\mathbf{R}^n$ with phase velocity given by $\boldsymbol{v}(\boldsymbol{x}) = A\boldsymbol{x}$ (see Example 1.1.3 above). By Theorem 1.3.2 of Lecture 9 we are done. □

Corollary 1.1.5. *All one-parameter groups of linear transformations on $\mathbf{R}^n$ are of the form $\{e^{tA}\}$, $A \in L(\mathbf{R}^n)$.*

Proof. Let $\{g^t\}$ be a one-parameter group of linear transformations. From §§ 1.4 of Lecture 9 the phase velocity vector of $\{g^t\}$ at $\boldsymbol{x}$ is $A\boldsymbol{x}$ where $A = \frac{\mathrm{d}g^t}{\mathrm{d}t}|_{t=0}$. It follows from Theorem 1.3.2 of Lecture 9 that the map $\boldsymbol{\psi}\colon \mathbf{R} \to \mathbf{R}^n$ given by $\boldsymbol{\psi}(t) = g^t\boldsymbol{a}$ is the unique solution to the IVP $\dot{\boldsymbol{x}} = A\boldsymbol{x}$, $\boldsymbol{x}(0) = \boldsymbol{a}$. On the other hand, from Theorem 1.1.4 so is $\boldsymbol{\varphi}\colon \mathbf{R} \to \mathbf{R}^n$ given by $\boldsymbol{\varphi}(t) = e^{tA}\boldsymbol{a}$. It follows that $\boldsymbol{\psi} = \boldsymbol{\varphi}$, whence $g^t = e^{tA}$, $t \in \mathbf{R}$. □

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*, translated by Roger Cooke, Universitext, Springer-Verlag, Berlin, 2006.