5. Matrix exponentials and Von Neumann's theorem

5.1. The matrix exponential. For an $n \times n$ matrix X we define

$$e^X = \exp X = I + X + \frac{X^2}{2!} + \dots = \sum_{n \geq 0} \frac{X^n}{n!}.$$

We assume that the entries are complex so that exp is well defined on $\mathcal{A}$, the algebra of $n \times n$ matrices. We denote by $|\cdot|$ a norm on $\mathcal{A}$ with the property that $|XY| \leq |X||Y|$. Such norms are easy to construct: if $|\cdot|$ is a norm on $\mathbf{C}^n$ we can take

$$|X| = \sup_{u \in \mathbf{C}^n, |u| \leq 1} |Xu|.$$

Since $|X^n| \leq |X|^n$, the series for $\exp X$ is majorized in norm by the numerical series for $e^{|X|}$. This shows that the series for $\exp X$ is absolutely convergent everywhere and uniformly on compact (=bounded in norm) subsets of $\mathcal{A}$. Hence $\exp X$ is a holomorphic matrix valued function on $\mathcal{A}$. Its properties resemble closely those of the ordinary exponential function.

(i) $\exp 0 = I$

(ii) $\exp(X + Y) = \exp X \exp Y$ if X and Y commute

(iii) $\exp X \exp -X = I$. In particular exp takes values in $\mathrm{GL}(n, \mathbf{C})$.

(iv) $\frac{d}{dt} \exp tX = X \exp tX = (\exp tX)X$.

(v) If $X_n \to X$, then

$$\exp X = \lim_{n \to \infty} \left(I + \frac{X_n}{n} \right)^n .$$

The proofs of these are very similar to the corresponding proofs in the scalar case except that (ii) requires a little more care. In fact, it is only when X and Y commute that we can write

$$(X + Y)^n = \sum_{0 \leq r \leq n} \binom{n}{r} X^r Y^{n-r}$$

so that

$$\frac{(X+Y)^n}{n!} = \sum_{0\le r\le n} \frac{X^r}{r!}\frac{Y^{n-r}}{(n-r)!}.$$

Then, with X and Y commuting,

$$\exp(X+Y) = \sum_{n\ge 0, 0\le r\le n} \frac{X^r}{r!}\frac{Y^{n-r}}{(n-r)!} = \sum_r \frac{X^r}{r1}\sum_s \frac{Y^s}{s!} = \exp X \exp Y.$$

For (v) we proceed as in the scalar case and write

$$\left(I + \frac{X_n}{n}\right)^n = \sum_{0\le r\le n} \binom{n}{r}\frac{X_n^r}{n^r} = \sum_{r\ge 0} u_r(n)$$

where

$$u_r(n) = \begin{cases} \frac{X_n^r}{r!}(1)(1-\frac{1}{n})(1-\frac{2}{n})\dots(1-\frac{r-1}{n}) & \text{if } r\le n \\ 0 & \text{if } r > n. \end{cases}$$

We may assume that $|X_n| \le C$ for some constant C for all n; then we have the estimate

$$|u_r(n)| \le \frac{C^r}{r!}$$

for all r *uniformly in* n, so that

$$\lim_{n\to\infty}\left(I + \frac{X_n}{n}\right)^n = \lim_{n\to\infty}\sum_{r\ge 0} u_r(n) = \sum_{r\ge 0}\lim_{n\to\infty} u_r(n) = \sum_{r\ge 0}\frac{X^r}{r!} = \exp X.$$

Besides these we have two less obvious limit formulae. The first one is a special case of the *Trotter product formula* valid in vastly greater generality, in the setting of Hilbert spaces and exponentials of unbounded self adjoint operators.

Proposition 1. *We have the following.*

(*i*) $\exp(X+Y) = \lim_{n\to\infty}\left(\exp(\frac{X}{n})\exp(\frac{Y}{n})\right)^n$

(*ii*) $\exp[X,Y] = \lim_{n\to\infty}\left(\exp(\frac{X}{n})\exp(\frac{Y}{n})\exp(-\frac{X}{n})\exp(-\frac{Y}{n})\right)^{n^2}$

Proof. For (i) we use $\exp(\frac{X}{n}) = I + \frac{X}{n} + O(\frac{1}{n^2})$ to find that

$$\left(\exp\left(\frac{X}{n}\right)\exp\left(\frac{Y}{n}\right)\right)^n = \left(I + \frac{X+Y}{n} + O\left(\frac{1}{n^2}\right)\right)^n$$

and the limit of the right side is $\exp(X+Y)$. For (ii) we need to expand the exponentials to the third order. We have

$$\exp\left(\frac{X}{n}\right) = I + \frac{X}{n} + \frac{X^2}{2n^2} + O\left(\frac{1}{n^3}\right).$$

It is then an easy calculation to find that

$$\exp\left(\frac{X}{n}\right)\exp\left(\frac{Y}{n}\right)\exp\left(-\frac{X}{n}\right)\exp\left(-\frac{Y}{n}\right) = I + \frac{[X,Y]}{n^2} + O\left(\frac{1}{n^3}\right)$$

from which (ii) follows at once.

Remark. All the above results are true if we replace $\mathbf{C}$ by $\mathbf{R}$.

5.2. Proof of Von Neumann's theorem. Von Neumann's theorem is the following.

Theorem (Von Neumann). *Let G be a closed subgroup of* $\mathrm{GL}(n,\mathbf{R})$. *Then G is a submanifold whose connected components all have the same dimension. In particular G is a Lie group.*

Proof. We introduce

$$\mathfrak{g} = \left\{X \in \mathfrak{gl}(n,\mathbf{R}) \;\middle|\; \exp tX \in G \text{ for all } t \in \mathbf{R}\right\}.$$

Select a matrix norm $|\cdot|$ on $\mathfrak{gl}(n,\mathbf{R})$. It is immediate from Proposition 1 that if $X, Y \in \mathfrak{g}$, then $cX, X+Y, [X,Y]$ are all in $\mathfrak{g}$ for $c \in \mathbf{R}$. Hence $\mathfrak{g}$ is a Lie algebra. We select a linear space $\mathfrak{a} \subset \mathfrak{gl}(n,\mathbf{R})$ such that $\mathfrak{g} \oplus \mathfrak{a} = \mathfrak{gl}(n,\mathbf{R})$. Let E be the map $\mathfrak{g} \times \mathfrak{a} \longrightarrow \mathrm{GL}(n,\mathbf{R})$ defined by

$$E(X,Y) = \exp X \exp Y.$$

It is immediate that E is analytic and its differential is bijective at $(0,0)$. In fact

$$dE_{(0,0)}(X,Y) = dE_{(0,0)}(X,0) + dE_{(0,0)}(0,Y) = X + Y$$

so that $dE_{(0,0)}$ is surjective, hence bijective. Hence E is a diffeomorphism at $(0,0)$. So there is a number $a > 0$ such that E maps $\mathfrak{g}_a \times \mathfrak{a}_a$ diffeomorphically onto an open neighborhood G_a of I in $\mathrm{GL}(n, \mathbf{R})$; here, for any subspace $\mathfrak{m} \subset \mathfrak{gl}(n, \mathbf{R})$ we write $\mathfrak{m}_a$ for the open ball of center 0 and radius a in $\mathfrak{m}$. Thus any element $x \in G_a$ can be written uniquely a $x = \exp A \exp B$ where $A \in \mathfrak{g}_a, B\mathfrak{a}_a$; if $x \to 1$, then $A, B \to 0$.

We claim that for some $b > 0$ with $0 < b < a$, E maps $\mathfrak{g}_b$ onto $G_b \cap G$. If this were not true, we can find $x_n \in G_a, x_n \to 1$ but $x_n = \exp A_n \exp B_n$ where $A_n \in \mathfrak{g}_a, B_n \in \mathfrak{a}_a$ with $A_n, B_n \to 0$ and $B_n \neq 0$ for all n. If $y_n = \exp(-A_n)x_n$, then $y_n \in G, y_n \to 1, y_n = \exp B_n$ where $B_n \in \mathfrak{a}, B_n \neq 0, B_n \to 0$. The B_n are very small and so we want to blow them up to look more closely at them. Since $B_n \neq 0$ we can find an integer $r_n \geq 1$ such that

$$|r_n B_n| \leq 1, \quad (r_n + 1)B_n| > 1.$$

The sequence $(r_n B_n)$ must have a convergent subsequence, and so, replacing it by this subsequence we may assume that

$$X = \lim_{n\to\infty} r_n B_n$$

exists. Clearly $|X| \leq 1$; on the other hand, $|r_n B_n| \geq |(r_n+1)B_n| - |B_n| \geq 1 - |B_n|$ and so, letting $n \to \infty$, we have $|X| \geq 1$ also. Hence $|X| = 1$, in particular, $X \neq 0$.

We claim that $\exp tX \in G$ for all $t \in \mathbf{R}$. It is enough to show this for all *rational* $t > 0$. Writing $t = c/k$ where c, k are integers ≥ 1, it is enough to show that $\exp((1/k)X) \in G$ for all integers $k \geq 1$. We use the argument that if $y_n^{m_n}$ has a limit where the m_n are integers ≥ 1, then this limit must be in G. Certainly $\exp(r_n B_n) = y_n^{r_n} \in G$ for all n and so $\exp X = \lim_{n\to\infty} y_n^{r_n} \in G$. Write $r_n = ks_n + t_n$ where $0 \leq t_n < k$. Then

$$\exp\left(\frac{r_n}{k} B_n\right) = \exp(s_n B_n) \exp\left(\left(\frac{t_n}{k}\right) B_n\right).$$

Since $|(t_n/k)B_n| \leq |B_n| \to 0$, it follows that

$$\lim_{n\to\infty} \exp(s_n B_n) = \lim_{n\to\infty} y_n^{s_n} = \exp\left(\frac{1}{k} X\right)$$

and so $\exp\left((1/k)X\right) \in G$ as we wanted.

E is thus a diffeomorphism of $\mathfrak{g}_b \times \mathfrak{a}_b$ with G_b and we have in addition that $G_b \cap G$ corresponds to $\mathfrak{g}_b$ under E. It is immediate that $G_b \cap G$ is a submanifold of G_b. This finishes the proof of the theorem.

Remark. The result is false for *complex* groups; just consider the unitary groups $\mathrm{U}(n) \subset \mathrm{GL}(n, \mathbf{C})$. But if we can prove that $\mathfrak{g}$ as defined above is closed under multiplication by $i = \sqrt{-1}$, then the proof will go through and establish that G is a complex submanifold, hence a complex Lie group.

We call $\mathfrak{g}$ the *Lie algebra of* G and denote it by $\mathrm{Lie}(G)$.

Problems

1. Determine $\mathrm{Lie}(G)$ for the classical groups.