Background 1 : Matrix exponentials

14/4/21
updated 18/6/21

Let $\mathbb{F}$ denote either $\mathbb{R}$ or $\mathbb{C}$. By a vector space we will always mean a vector space over $\mathbb{F}$. Given a finite-dimensional vector space V, an operator on V is a linear transformation $T: V \longrightarrow V$. We assume familiarity with normed vector spaces (see e.g. [MHS, L18, L19]) and the norm of a linear transformation.

Lemma B1-1 Any two norms $\|-\|_a, \|-\|_b$ on a finite-dimensional vector space V are Lipschitz equivalent, i.e. there exist $0 < C_1 \le C_2$ such that

$$C_1 \|x\|_a \le \|x\|_b \le C_2 \|x\|_a \qquad \forall x \in V.$$

Proof It is easy to check that this relation of Lipschitz equivalence $\sim$ is symmetric and transitive, so it suffices to prove that for any norm $\|-\|$ we have $\|-\| \sim \|-\|_1$ where the latter norm is defined by choosing basis $v_1, \dots, v_n$ for V and defining

$$\| \textstyle\sum_{i=1}^n a_i v_i \|_1 = \sum_{i=1}^n |a_i|. \qquad (1.1)$$

This is easily checked to be a norm (indeed it is the norm induced on V by the vector space isomorphism $V \cong \mathbb{F}^n$ and the L_1 norm on $\mathbb{F}^n$ [MHS, Thm L18-1]).

To prove $\|-\| \sim \|-\|_1$ we first prove that $\|-\| : V \longrightarrow \mathbb{F}$ is (uniformly) continuous when V is given the metric determined by $\|-\|_1$, or what is the same

$$\forall \varepsilon > 0 \; \exists \delta > 0 \left(\|x - x'\|_1 < \delta \Rightarrow \big| \|x\| - \|x'\| \big| < \varepsilon \right). \qquad (1.2)$$

To see this note by the reverse triangle inequality $\big| \|x\| - \|x'\| \big| \le \|x - x'\|$ and if $x = \sum_{i=1}^n a_i v_i$, $x' = \sum_{i=1}^n a_i' v_i$ then with $C = \sup\{ \|v_i\| \mid 1 \le i \le n \}$

$$\begin{aligned} \|x - x'\| &= \| \textstyle\sum_{i=1}^{n} (a_i - a_i') v_i \| \\ &\leq \textstyle\sum_{i=1}^{n} |a_i - a_i'| \, \|v_i\| \\ &\leq C \textstyle\sum_{i=1}^{n} |a_i - a_i'| \\ &= C \| x - x' \|_1 \end{aligned}$$

From this we easily deduce (1.2), proving that $\|-\| : (V, \|-\|_1) \longrightarrow (\mathbb{F}, |-|)$ is continuous. Now by construction V with the $\|-\|_1$-induced topology is homeomorphic to $\mathbb{R}^n$, and so $\{ v \in V \mid \|v\|_1 = 1 \}$ is compact since the corresponding set in $\mathbb{R}^n$ is closed and bounded [MHS, Theorem L10-3]. Hence by the extreme value theorem [MHS, Corollary L9-4] the continuous function $\|-\|$ attains its supremum and infimum on this unit $\|-\|_1$-sphere

$$\begin{aligned} C_1 &= \inf\{ \|v\| \mid \|v\|_1 = 1 \}, \\ C_2 &= \sup\{ \|v\| \mid \|v\|_1 = 1 \}. \end{aligned} \tag{2.1}$$

That is, $C_1 = \|v\|$ and $C_2 = \|w\|$ for some v, w with $\|v\|_1 = \|w\|_1 = 1$. In particular $v, w \neq 0$ and so $C_1, C_2 \neq 0$. We claim that for all $v \in V$

$$C_1 \|v\|_1 \leq \|v\| \leq C_2 \|v\|_1 \tag{2.2}$$

This immediate if $\|v\|_1 = 1$ or $v = 0$, and otherwise we may multiply by $\frac{1}{\|v\|_1}$ to reduce to this case. $\square$

Lemma B1-2 Let $(V, \|-\|_V)$, $(W, \|-\|_W)$ be normed spaces with V finite-dimensional. Then any linear transformation $T: V \longrightarrow W$ is bounded.

Proof A transformation is bounded with respect to a pair of norms iff. it is bounded with respect to any Lipschitz equivalent norms, so by Lemma B1-1 we may assume $\|-\|_V = \|-\|_1$ for some basis $v_1, \ldots, v_n$ of V. But then if $T: V \longrightarrow W$ is linear and $x = \sum_{i=1}^n a_i v_i$

$$\begin{aligned} \| Tx \|_W &= \| \textstyle\sum_{i=1}^n a_i T(v_i) \|_W \\ &\leq \textstyle\sum_{i=1}^n |a_i| \, \| T(v_i) \|_W \\ &\leq C \| x \|_1 \end{aligned}$$

where $C = \sup\{ \| T(v_i) \|_W \mid 1 \leq i \leq n \}$. Hence T is bounded. $\square$

Lemma B1-3 Let U, V, W be normed spaces and $S: V \longrightarrow W$, $T: U \longrightarrow V$ be bounded linear operators. Then $S \circ T$ is bounded and $\| S \circ T \| \leq \| S \| \| T \|$.

Proof Given $x \in U$ we have

$$\begin{aligned} \| (S \circ T)(x) \|_W &= \| S(T(x)) \|_W \\ &\leq \| S \| \, \| T(x) \|_V \\ &\leq \| S \| \, \| T \| \, \| x \|_U \end{aligned}$$

as claimed. $\square$

To introduce the matrix exponential we require some basic background in sequences, series and convergence in a normed space. Recall that the norm $\|-\|$ determines a metric $d(x,y) = \|x - y\|$ and we say a sequence is convergent or Cauchy, and a series is convergent, if these statements hold in the usual sense with respect to that metric:

- a sequence $(u_n)_{n=0}^{\infty}$ in a normed space $(V, \|-\|)$ <u>converges</u> to $u \in V$ if

$$\forall \varepsilon > 0 \, \exists N \in \mathbb{N} \, \forall n \in \mathbb{N} \, (n \geq N \Rightarrow \|u_n - u\| < \varepsilon)$$

- a sequence $(u_n)_{n=0}^{\infty}$ in a normed space $(V, \|-\|)$ is <u>Cauchy</u> if

$$\forall \varepsilon > 0 \, \exists N \in \mathbb{N} \, \forall m, n \in \mathbb{N} \, ((m \geq N \text{ and } n \geq N) \Rightarrow \|u_m - u_n\| < \varepsilon)$$

- a series $\sum_{n=0}^{\infty} u_n$ (which is actually the data of $(u_n)_{n=0}^{\infty}$) is said to <u>converge</u> in a normed space $(V, \|-\|)$ if the sequence $(\sum_{n=0}^{m} u_n)_{m=0}^{\infty}$ of partial sums converges, and we write $\sum_{n=0}^{\infty} u_n = \lim_{m \to \infty} \sum_{n=0}^{m} u_n$.

- a normed space $(V, \|-\|)$ is <u>complete</u> if every Cauchy sequence in V converges.

<u>Exercise B1-1</u> Use Lemma B1-1 to prove that every finite-dimensional normed space is complete (you may use that $\mathbb{R}$ is complete). Hence every finite-dimensional inner product space is a Hilbert space.

We can reduce checking convergence of a series in V to checking convergence of a series in $\mathbb{R}$:

<u>Lemma B1-4</u> Let $(V, \|-\|)$ be a complete normed vector space and $\sum_{n=0}^{\infty} u_n$ a series which has the property that $\sum_{n=0}^{\infty} \|u_n\|$ converges in $\mathbb{R}$ (such a series is called <u>absolutely convergent</u>). Then $\sum_{n=0}^{\infty} u_n$ converges.

<u>Proof</u> Set $a_m = \sum_{n=0}^{m} u_n$. Then for $m \geq m'$

$$\|a_m - a_{m'}\| = \left\| \sum_{n=m'+1}^{m} u_n \right\| \leq \sum_{n=m'+1}^{m} \|u_n\|. \qquad (4.1)$$

The sequence $\left(\sum_{n=0}^{m} \|u_n\| \right)_{m=0}^{\infty}$ is assumed to converge and is therefore Cauchy, so given $\varepsilon > 0$ we can choose N s.t. for $m \geq m' \geq N$ the RHS of (4.1) is less than ε. It follows that $(a_m)_{m=0}^{\infty}$ is Cauchy in V and hence converges. $\square$

We assume familiarity with the various tests for convergence for series in $\mathbb{R}$. It is easy to see that if linear transformations S, T are bounded so too are $S+T$ and λT for $\lambda \in \mathbb{F}$.

<u>Def^n</u> Given normed spaces $(V, \|-\|)$, $(W, \|-\|)$ we write $\mathcal{B}(V, W)$ for the vector space of bounded linear transformations $V \longrightarrow W$ with the pointwise operations

$$(T+S)(x) = T(x) + S(x) \qquad T, S \in \mathcal{B}(V), x \in V$$

$$(\lambda T)(x) = \lambda \cdot T(x) \qquad T \in \mathcal{B}(V), x \in V.$$

<u>Lemma B1-5</u> The operator norm makes $\mathcal{B}(V, W)$ a normed space.

<u>Proof</u> It is clear that $\|T\| \geq 0$ and that $\|T\| = 0$ implies $T = 0$. Given $\lambda \in \mathbb{F}$

$$\|\lambda T\| = \sup\left\{ \frac{\|\lambda T(x)\|}{\|x\|} \,\middle|\, x \neq 0 \right\}$$

$$= \sup\left\{ |\lambda| \frac{\|T(x)\|}{\|x\|} \,\middle|\, x \neq 0 \right\}$$

$$= |\lambda| \sup\left\{ \frac{\|T(x)\|}{\|x\|} \,\middle|\, x \neq 0 \right\} = |\lambda| \, \|T\|.$$

If S, T are bounded then

$$\|S + T\| = \sup\left\{ \frac{\|(S+T)(x)\|}{\|x\|} \,\middle|\, x \neq 0 \right\}$$

$$= \sup\left\{ \frac{\|Sx + Tx\|}{\|x\|} \,\middle|\, x \neq 0 \right\}$$

$$\leq \sup\left\{ \frac{\|Sx\|}{\|x\|} + \frac{\|Tx\|}{\|x\|} \,\middle|\, x \neq 0 \right\} \leq \|S\| + \|T\| . \ \square$$

Recall that a normed space is called <u>Banach</u> if it is complete (meaning all Cauchy sequences with respect to the induced metric converge).

<u>Lemma B1-6</u> If W is Banach then so is $\mathcal{B}(V,W)$ for any normed space V.

<u>Proof</u> Let $(T_n)_{n=0}^{\infty}$ be a Cauchy sequence in $\mathcal{B}(V,W)$, so each $T_n : V \to W$ is bounded and

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall m,n \geqslant N \left(\| T_m - T_n \| < \varepsilon \right). \qquad (6.1)$$

Given $x \in V$ we first claim $(T_n(x))_{n=0}^{\infty}$ is Cauchy in W. Given $\varepsilon > 0$ let N be as in (6.1) but for the positive real number $\varepsilon / \|x\|_V$ (if $x = 0$ the sequence is trivially Cauchy). Then for $m,n \geqslant N$

$$\| T_m(x) - T_n(x) \| \leqslant \| T_m - T_n \| \, \|x\|_V < \frac{\varepsilon}{\|x\|_V} \|x\|_V = \varepsilon \qquad (6.2)$$

as claimed. Since W is assumed complete we may set $T(x) := \lim_{n\to\infty} T_n(x)$. It remains to prove that thus defined T is linear and bounded, and that $T_n \to T$ in $\mathcal{B}(V,W)$.

To see that T is linear we use that any normed space is a topological vector space [MHS, Ex L18-10], that is, the operations are continuous (and therefore commute with limits [MHS, L8])

$$\begin{aligned} T(x+y) &= \lim_{n\to\infty} T_n(x+y) \\ &= \lim_{n\to\infty} \left(T_n(x) + T_n(y) \right) \\ &= \lim_{n\to\infty} (T_n(x)) + \lim_{n\to\infty} (T_n(y)) \\ &= T(x) + T(y) \end{aligned} \qquad (6.3)$$

$$\begin{aligned} T(\lambda x) &= \lim_{n\to\infty} T_n(\lambda x) \\ &= \lim_{n\to\infty} \lambda T_n(x) = \lambda \lim_{n\to\infty} T_n(x) = \lambda T(x) \end{aligned}$$

The norm on $\mathcal{B}(V,W)$ is uniformly continuous [MHS, Lemma L18-3] and the image of a Cauchy sequence under a uniformly continuous map is Cauchy, so $(\|T_n\|)_{n=0}^{\infty}$ is Cauchy in $\mathbb{R}$ and thus converges, say to α. We claim that $\|T(x)\| \le \alpha \|x\|$ for all $x \in V$.

To show this let $x \in V$ nonzero and $\varepsilon > 0$ be given. Let N be large enough that $\|T(x) - T_N(x)\| < \varepsilon/2$ and $|\,\|T_N\| - \alpha\,| < \varepsilon/2\|x\|$. Then

$$\begin{aligned}
\|T(x)\| &= \|T(x) - T_N(x) + T_N(x)\| \\
&\le \|T(x) - T_N(x)\| + \|T_N(x)\| \\
&< \varepsilon/2 + \|T_N\|\,\|x\| \qquad (7.1) \\
&< \varepsilon/2 + \left(\alpha + \varepsilon/2\|x\|\right)\|x\| \\
&= \varepsilon + \alpha\|x\|
\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary this proves $\|T(x)\| \le \alpha\|x\|$ and so T is bounded.

To prove $T_n \to T$ in $\mathcal{B}(V,W)$ is to prove that $\|T_n - T\| \to 0$ in $\mathbb{R}$. Given $\varepsilon > 0$ we may since $(T_n)_{n=0}^{\infty}$ is Cauchy find N such that $\|T_m - T_n\| < \varepsilon/2$ whenever $m, n \ge N$. Given $x \in V$ nonzero let N_x be such that $\|T_m(x) - T(x)\| < \frac{\varepsilon}{2}\|x\|$ whenever $m \ge N_x$. We may assume $N_x \ge N$. Then for any x nonzero and $n \ge N$

$$\begin{aligned}
\|T_n(x) - T(x)\| &\le \|T_n(x) - T_{N_x}(x) + T_{N_x}(x) - T(x)\| \\
&\le \|T_n(x) - T_{N_x}(x)\| + \|T_{N_x}(x) - T(x)\| \\
&\le \|T_n - T_{N_x}\|\,\|x\| + \tfrac{\varepsilon}{2}\|x\| < \tfrac{\varepsilon}{2}\|x\| + \tfrac{\varepsilon}{2}\|x\| = \varepsilon\|x\|
\end{aligned}$$

Hence $\|T_n - T\| \le \varepsilon$ for $n \ge N$ and hence $T_n \to T$ in $\mathcal{B}(V,W)$ as claimed. $\square$

We write $\mathcal{B}(V)$ for $\mathcal{B}(V,V)$ the normed space of bounded operators on V. Given $T \in \mathcal{B}(V)$ we write T^n for $T \circ \cdots \circ T$ the n-fold composition of T with itself.

Theorem B1-7 If $T: V \to V$ is a bounded operator on a Banach space V then the series

$$\exp(T) = \sum_{n=0}^{\infty} \frac{1}{n!} T^n \qquad (8.1)$$

converges absolutely in $\mathcal{B}(V)$.

Proof By Lemma B1-3 the operators $T^n = \overbrace{T \circ \cdots \circ T}^{n}$ are bounded, so the partial sums $S_m = \sum_{n=0}^{m} \frac{1}{n!} T^n$ are vectors in $\mathcal{B}(V)$. By Lemma B1-6 $\mathcal{B}(V)$ is a Banach space so by Lemma B1-4 to show that (8.1) converges it suffices to show that the series

$$\sum_{n=0}^{\infty} \frac{1}{n!} \| T^n \| \qquad (8.2)$$

converges in $\mathbb{R}$. But Lemma B1-3 this (positive) series is dominated by $\sum_{n=0}^{\infty} \frac{1}{n!} \|T\|^n$ which converges (to $\exp(\|T\|)$), hence (8.2) also converges. $\square$

Example B1-1 $V = \mathbb{F}^n$ with the $\|-\|_2$ norm, $\|x\|_2 = (|x_1|^2 + \cdots + |x_n|^2)^{1/2}$, any linear operator T on V is bounded (Lemma B1-2) and so $\exp(T)$ converges with respect to the operator norm on $\mathcal{B}(V) = \mathrm{End}_{\mathbb{F}}(V)$, the space of all linear operators. Note that since $\mathrm{End}_{\mathbb{F}}(V)$ is finite-dimensional the series $\sum_{n=0}^{\infty} \frac{1}{n!} T^n$ converges by Lemma B1-1 with respect to any norm on $\mathrm{End}_{\mathbb{F}}(V)$. In particular the series converges with respect to the Frobenius norm

$$\|S\|_F = \left(\sum_{i,j=1}^{n} |S_{ij}|^2 \right)^{1/2} \qquad (8.3)$$

which is just the norm induced by $\mathrm{End}_{\mathbb{F}}(V) \cong \mathbb{F}^{n^2}$ and the $\|-\|_2$ norm on $\mathbb{F}^{n^2}$.

Lemma B1-8 Let U, V, W be normed vector spaces. Then the composition map

$$\begin{aligned} \mathcal{B}(V,W) \times \mathcal{B}(U,V) &\longrightarrow \mathcal{B}(U,W) \\ (S,T) &\longmapsto S \circ T \end{aligned}$$

is continuous.

Proof Here we give $\mathcal{B}(V,W) \times \mathcal{B}(U,V)$ the product metric [MHS, Ex L13-8]

$$d\big((S_1,T_1),(S_2,T_2)\big) = \|S_1 - T_1\| + \|S_2 - T_2\|.$$

It suffices to prove that if $(S_n, T_n) \longrightarrow (S,T)$ in $\mathcal{B}(V,W) \times \mathcal{B}(U,V)$ then $S_n \circ T_n \longrightarrow S \circ T$ in $\mathcal{B}(U,W)$ [MHS, Lemma L8-4]. The projections from $\mathcal{B}(V,W) \times \mathcal{B}(U,V)$ to its two factors are continuous, so $S_n \longrightarrow S$ and $T_n \longrightarrow T$. Observe that

$$\begin{aligned} \| S_n \circ T_n - S \circ T \| &= \| S_n \circ T_n - S_n \circ T + S_n \circ T - S \circ T \| \\ &\leq \| S_n \| \, \| T_n - T \| + \| S_n - S \| \, \| T \| \end{aligned}$$

Since $\|-\|$ is continuous, $\|S_n\| \longrightarrow \|S\|$ and so

$$\begin{aligned} &\lim_{n \to \infty} \big(\|S_n\| \, \|T_n - T\| + \|S_n - S\| \, \|T\| \big) \\ &= \big(\lim_{n\to\infty} \|S_n\|\big)\big(\lim_{n\to\infty} \|T_n - T\|\big) + \big(\lim_{n\to\infty} \|S_n - S\|\big) \|T\| \\ &= \|S\| \cdot 0 + 0 \cdot \|T\| = 0 \end{aligned}$$

Hence also $\lim_{n \to \infty} \| S_n \circ T_n - S \circ T \| = 0$ as claimed. $\square$

<u>Lemma B1-9</u> If V is a Banach space and $\sum_{n=0}^{\infty} v_n$ converges absolutely, then any rearrangement $\sum_{n=0}^{\infty} v_{j(n)}$ converges absolutely and $\sum_{n=0}^{\infty} v_n = \sum_{n=0}^{\infty} v_{j(n)}$.

$j : \mathbb{N} \to \mathbb{N}$ a bijection

<u>Proof</u> From (absolute) convergence of $\sum_{n=0}^{\infty} \|v_n\|$ we know by the corresponding result for $\mathbb{R}$ (which we assume) that $\sum_{n=0}^{\infty} \|v_{j(n)}\|$ converges and to the same limit. Hence $\sum_{n=0}^{\infty} v_{j(n)}$ converges and we need only show $\sum_{n=0}^{\infty} v_n = \sum_{n=0}^{\infty} v_{j(n)}$.

Set $L_m = \sum_{n=0}^{m} v_n$, $R_m = \sum_{n=0}^{m} v_{j(n)}$. We show $\lim_{m\to\infty} \|L_m - R_m\| = 0$.

To this end let $\varepsilon > 0$ be given. We aim to show that there exists N such that for all $m \geqslant N$ we have $\|L_m - R_m\| < \varepsilon$.

Set $S_m = \sum_{n=0}^{m} \|v_n\|$. Since the sequence $(S_m)_{m=0}^{\infty}$ is Cauchy we can find N_1 such that for $m, m' \geqslant N_1$ $|S_m - S_{m'}| < \varepsilon/2$. Hence for $m' > m \geqslant N_1$

$$\sum_{i=m+1}^{m'} \|v_i\| = S_{m'} - S_m < \varepsilon/2$$

This implies that for $B \subseteq \mathbb{N} \setminus \{0, \dots, N_1\}$ finite $\sum_{i \in B} \|v_i\| < \varepsilon/2$.

Let $N_2 = \max\{j^{-1}(0), \dots, j^{-1}(N_1)\}$, and $N = \max\{N_1, N_2\}$. Then for $m > N$ we have $j(m) \notin \{1, \dots, N_1\}$. Hence with $A = \{j^{-1}(0), \dots, j^{-1}(N_1)\}$

$$L_m - R_m = \sum_{n=0}^{m} v_n - \sum_{n=0}^{m} v_{j(n)} = \sum_{n=N_1+1}^{m} v_n - \sum_{i \in \{0,\dots,m\} \setminus A} v_{j(i)}$$

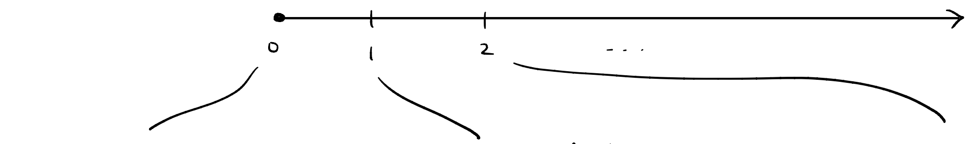
since $\sum_{n=0}^{N_1} v_n = \sum_{i \in A} v_{j(i)}$. But $\{N_1+1, \dots, m\}$, $j(\{0, \dots, m\} \setminus A)$ are both finite sets disjoint from $\{0, \dots, N_1\}$ so

$$\|L_m - R_m\| \leq \sum_{n=N_1+1}^{m} \|v_n| + \sum_{i \in \{0,\dots,m\} \setminus A} \|v_{j(i)}\|$$
$$< \varepsilon/2 + \varepsilon/2 = \varepsilon$$

as claimed. $\square$

<u>Lemma B1-10</u> If V is a Banach space and $\sum_{n=0}^{\infty} v_n$ converges absolutely then for any surjective map $j : \mathbb{N} \longrightarrow \mathbb{N}$ with the property that $j^{-1}(n)$ is finite for all $n \in \mathbb{N}$ the series $\sum_{n=0}^{\infty} \left(\sum_{i \in j^{-1}(n)} v_i \right)$ converges absolutely and
$\sum_{n=0}^{\infty} \left(\sum_{i \in j^{-1}(n)} v_i \right) = \sum_{n=0}^{\infty} v_n .$

<u>Proof</u> We define a bijection $J : \mathbb{N} \longrightarrow \mathbb{N}$ by enumerating $j^{-1}(0)$ then $j^{-1}(1)$ and so on, as in the diagram:



$$j^{-1}(0) = \{ J(0), J(1), \dots, J(a) \} \qquad j^{-1}(1) = \{ J(a+1), \dots, J(b) \} \qquad j^{-1}(2) = \{ J(b+1), \dots \} \tag{11.1}$$

number of elements

More formally, let $\alpha(i) = \sum_{a < j(i)} \# j^{-1}(a)$ and if $j^{-1}(j(i))$ arranged in ascending order contains $\beta(i)$ elements strictly less than i, define $J(i) = \alpha(i) + \beta(i)$. By Lemma B1-9 the series $\sum_{n=0}^{\infty} v_{J(n)}$ converges absolutely to $\sum_{n=0}^{\infty} v_n$. But with $s_m = \sum_{n=0}^{m} v_{J(n)}$ we see that $\sum_{n=0}^{m} \sum_{i \in j^{-1}(n)} v_i$ is a subsequence of $(s_m)_{m=0}^{\infty}$ and hence converges to the same limit. $\square$

<u>Theorem B1-11</u> Let S, T be bounded operators on a Banach space V

(i) $\exp(0) = 1_V$
(ii) if $ST = TS$ then $\exp(S)\exp(T) = \exp(S+T)$
(iii) if $\alpha, \beta \in \mathbb{F}$ then $\exp(\alpha S)\exp(\beta S) = \exp((\alpha+\beta)S)$.
(iv) $\exp(S)$ is invertible with inverse $\exp(-S)$.

<u>Proof</u> (i) is immediate from (8.1) and (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) so we need only prove (ii). By definition $\exp(S+T)$ is the limit of $a_m = \sum_{n=0}^{m} \frac{1}{n!} (S+T)^n$ and we claim that

$$(S+T)^n = \sum_{i=0}^{n} \binom{n}{i} S^{n-i} T^i \tag{11.2}$$

This is proven by induction on n, with $n=0$ and $n=1$ being trivial and the inductive step using $ST=TS$ as follows

$$
\begin{aligned}
(S+T)^{n+1} &= (S+T)\sum_{i=0}^{n}\binom{n}{i}S^{n-i}T^{i} \\
&= \sum_{i=0}^{n}\binom{n}{i}S^{n-i+1}T^{i} + \sum_{i=0}^{n}\binom{n}{i}S^{n-i}T^{i+1} \\
&= \sum_{i=0}^{n+1}\left[\binom{n}{i}+\binom{n}{i-1}\right]S^{n+1-i}T^{i} \qquad (12.1) \\
&= \sum_{i=0}^{n+1}\binom{n+1}{i}S^{n+1-i}T^{i}
\end{aligned}
$$

Hence

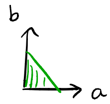
$$
\begin{aligned}
a_m &= \sum_{n=0}^{m}\frac{1}{n!}(S+T)^{n} \\
&= \sum_{n=0}^{m}\frac{1}{n!}\sum_{i=0}^{n}\frac{n!}{i!(n-i)!}S^{n-i}T^{i} \qquad (12.2) \\
&= \sum_{n=0}^{m}\sum_{i=0}^{n}\left(\frac{1}{(n-i)!}S^{n-i}\right)\left(\frac{1}{i!}T^{i}\right) \\
&= \sum_{n=0}^{m}\sum_{a+b=n}\left(\frac{1}{a!}S^{a}\right)\left(\frac{1}{b!}T^{b}\right)
\end{aligned}
$$

Now by definition $\exp(S) = \lim_{k\to\infty}\sum_{n=0}^{k}\frac{1}{n!}S^{n}$, $\exp(T) = \lim_{k\to\infty}\sum_{n=0}^{k}\frac{1}{n!}T^{n}$ so by Lemma B1-8 we have

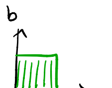
$$
\begin{aligned}
\exp(S)\exp(T) &= \lim_{k\to\infty}\left[\left(\sum_{n=0}^{k}\frac{1}{n!}S^{n}\right)\left(\sum_{n=0}^{k}\frac{1}{n!}T^{n}\right)\right] \qquad (12.3) \\
&= \lim_{k\to\infty}\sum_{a,b=0}^{k}\left(\frac{1}{a!}S^{a}\right)\left(\frac{1}{b!}T^{b}\right)
\end{aligned}
$$

If we write $X_{a,b} = \frac{1}{a!}S^{a}\frac{1}{b!}T^{b}$ then we have shown ($a,b \geq 0$ in the following)

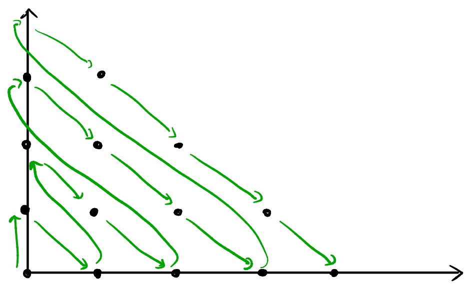
$$\exp(S+T) = \lim_{m\to\infty}\sum_{a+b\leq m}X_{a,b} \qquad (12.4)$$



$$\exp(S)\exp(T) = \lim_{m\to\infty}\sum_{a\leq m,\, b\leq m}X_{a,b}$$



Let $f : \mathbb{N} \longrightarrow \mathbb{N}\times\mathbb{N}$ be the bijection which enumerates $\mathbb{N}\times\mathbb{N}$ as below:



(13.1)

Set $(a_i, b_i) = f(i)$. Then $\sum_{i=0}^{\infty} \frac{1}{a_i!} \|S\|^{a_i} \frac{1}{b_i!} \|T\|^{b_i}$ converges since it is the limit of a sequence of partial sums which has $\sum_{a+b\leq m} \frac{1}{a!}\|S\|^a \frac{1}{b!}\|T\|^b$ as a subsequence (this sequence converges by an analogue of (12.2) to $\exp(\|S\|+\|T\|)$) and an increasing sequence with a converging subsequence is bounded, hence convergent.

Then the series $\sum_{i=0}^{\infty} X_{f(i)}$ converges absolutely since

$$\sum_{i=0}^{k} \|X_{f(i)}\| \leq \sum_{i=0}^{k} \frac{1}{a_i!}\|S\|^{a_i} \frac{1}{b_i!}\|T\|^{b_i}. \qquad (13.2)$$

Let $\Theta_0 \subset \Theta_1 \subset \cdots$ be any strictly ascending chain of nonempty finite subsets of $\mathbb{N}\times\mathbb{N}$ with $\bigcup_i \Theta_i = \mathbb{N}\times\mathbb{N}$, and define $j : \mathbb{N} \longrightarrow \mathbb{N}$ by $j(a) = \inf\{i \mid f(a) \in \Theta_i\}$. This is easily seen to be surjective, and hence by Lemma B1-10 the series $\sum_{n=0}^{\infty}\left(\sum_{i\in j^{-1}(n)} X_{f(i)}\right)$ converges absolutely to $\sum_{i=0}^{\infty} X_{f(i)}$. Taking alternatively

$$\begin{aligned} \Theta_m &= \{(a,b)\in\mathbb{N}\times\mathbb{N} \mid a+b\leq m\}, \\ \Theta_m &= \{(a,b)\in\mathbb{N}\times\mathbb{N} \mid a\leq m,\ b\leq m\} \end{aligned} \qquad (13.3)$$

shows that $\exp(S+T) = \exp(S)\exp(T)$ using (12.4). $\square$

Exercise B1-2 (i) Let V be a normed space and $v \in V$. Let $\eta_v : \mathbb{F} \longrightarrow V$ be the linear transformation $\eta_v(\lambda) = \lambda v$. Prove $\|\eta_v\| = \|v\|$, where we use the norm $|-|$ on $\mathbb{F}$.

(ii) Prove that there is a norm-preserving isomorphism of vector spaces $V \longrightarrow \mathcal{B}(\mathbb{F}, V)$ sending v to η_v.

(iii) Prove that for normed vector spaces V, W the function

$$\begin{aligned} \mathcal{B}(V, W) \times V &\longrightarrow W \\ (T, v) &\longmapsto T(v) \end{aligned}$$

is continuous.

(iv) Let $T: V \longrightarrow V$ be a bounded linear operator on a Banach space V. Prove that for $v \in V$ we have

$$\exp(T)(v) = \lim_{m \to \infty} \left(\sum_{n=0}^{m} \frac{1}{n!} T^n(v) \right)$$

Exercise B1-3 Let V, W be Banach spaces and $S: V \longrightarrow W$ a norm-preserving isomorphism of vector spaces. Let $T: V \longrightarrow V$ be a bounded linear operator. Prove that as operators on W,

$$S \exp(T) S^{-1} = \exp(S T S^{-1}).$$

Exercise B1-4 With $\mathbb{F} = \mathbb{R}$ consider $V = \mathbb{C}$ as a two-dimensional $\mathbb{R}$-vector space with basis $\beta = \{1, i\}$ and let $T: V \longrightarrow V$ be multiplication by i, so $T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ as a matrix. Prove that $e^{i\theta} = \cos\theta + i\sin\theta$, i.e. for $\theta \in \mathbb{R}$

$$\exp(\theta T) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Example B1-2 Suppose $\dim(V) = n < \infty$ and $T : V \to V$ is linear, hence bounded, over $\mathbb{F} = \mathbb{C}$. The existence of a Jordan normal form for T means that with respect to some basis β

$$[T]^{\beta}_{\beta} = \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_m \end{pmatrix} \qquad (15.1)$$

where each J_i is a Jordan block, i.e. a matrix of the form

$$\begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix} \qquad (15.2)$$

where λ is an eigenvalue of T. Note that this block is $\lambda I + N$ where N is nilpotent (some power is zero). A basis-free way of expressing (14.1) is to say that there is a direct sum decomposition

$$V = V_1 \oplus \cdots \oplus V_m \qquad (15.3)$$

with each V_i a T-invariant subspace (i.e. $T(V_i) \subseteq V_i$) and the restriction $T|_{V_i} : V_i \to V_i$ equal to $\lambda_i I + N_i$ for some $\lambda_i \in \mathbb{C}$ and nilpotent operator N_i. Then

$$\exp(T) = \exp(\lambda_1 I + N_1) \oplus \cdots \oplus \exp(\lambda_m I + N_m). \qquad (15.4)$$

Thus <u>in principle</u> to compute the exponential of any operator on a finite-dimensional $\mathbb{C}$-vector space, we need only compute exponentials $\exp(\lambda I + N)$ of Jordan blocks. But λI commutes with N, and so by Theorem B1-11 if $N^k = 0$,

$$\begin{aligned}\exp(\lambda I + N) &= \exp(\lambda I)\exp(N) \\ &= e^{\lambda}\left[1 + N + \tfrac{1}{2}N^2 + \cdots + \tfrac{1}{(k-1)!}N^{k-1}\right]\end{aligned} \tag{16.1}$$

Recall from (15.2) that N consists of 1's on the off-diagonal, and so

$$N^2 = \begin{pmatrix} 0 & 0 & 1 & & & \\ \vdots & 0 & 0 & 1 & & \\ \vdots & & 0 & \ddots & \ddots & \\ \vdots & & & \ddots & \ddots & 1 \\ \vdots & & & & \ddots & 0 \\ 0 & & & & & 0 \end{pmatrix} \tag{16.2}$$

and similarly for the other powers of N. Hence if $N^{k-1} \neq 0$ but $N^k = 0$,

$$\exp(\lambda I + N) = e^{\lambda}\begin{pmatrix} 1 & 1 & \frac{1}{2} & \cdots & \frac{1}{(k-1)!} \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \frac{1}{2} \\ & 0 & & \ddots & 1 \\ & & & & 1 \end{pmatrix} \tag{16.3}$$

This is however less useful than it appears, because we are often given operators (such as $\eta^{\hat{n}}$) defined abstractly for which we may not have a convenient way of determining the Jordan normal form.

Exercise B1-5 Prove that for any matrix $X \in M_n(\mathbb{C})$

$$\exp(\operatorname{tr} X) = \det(\exp(X)).$$

Exercise B1-6 Compute $\exp(\alpha X)$, $\exp(\alpha Y)$, $\exp(\alpha H)$ for $\alpha \in \mathbb{R}$ where

$$X = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

<u>Lemma B1-12</u> Let V be a Banach space. The function $\exp : \mathcal{B}(V) \longrightarrow \mathcal{B}(V)$ is continuous.

<u>Proof</u> By Lemma B1-8 the partial sums $a_m(X) = \sum_{n=0}^{m} \frac{1}{n!} X^n$ are continuous maps $\mathcal{B}(V) \longrightarrow \mathcal{B}(V)$. Given $R > 0$ let $S_R = \{ X \in \mathcal{B}(V) \mid \|X\| \le R \}$. The sequence

$$\left(a_m|_{S_R} : S_R \longrightarrow \mathcal{B}(V) \right)_{m=0}^{\infty} \tag{17.1}$$

converges <u>uniformly</u> since given $\varepsilon > 0$

$$\begin{aligned} \| a_m(X) - \exp X \| &= \| \textstyle\sum_{n=m+1}^{\infty} \frac{1}{n!} X^n \| \\ &= \| \lim_{h \to \infty} \textstyle\sum_{n=m+1}^{h} \frac{1}{n!} X^n \| \\ &\overset{\|-\| \text{ is cts}}{=} \lim_{h \to \infty} \| \textstyle\sum_{n=m+1}^{h} \frac{1}{n!} X^n \| \\ &\overset{\text{Lemma B1-3}}{\le} \lim_{h \to \infty} \textstyle\sum_{n=m+1}^{h} \frac{1}{n!} \|X\|^n \\ &\le \lim_{h \to \infty} \textstyle\sum_{n=m+1}^{h} \frac{1}{n!} R^n = \textstyle\sum_{n=m+1}^{\infty} \frac{1}{n!} R^n \end{aligned} \tag{17.2}$$

If $s_m = \sum_{n=0}^{m} \frac{1}{n!} R^n$ then since $s_m \to e^R$ we have $e^R - s_m \to 0$ and hence $\sum_{n=m+1}^{\infty} \frac{1}{n!} R^n \to 0$ as $m \to \infty$. Thus given $\varepsilon > 0$ we can find N such that for all $m \geq N$ $\sum_{n=m+1}^{\infty} \frac{1}{n!} R^n < \varepsilon$ and thus $\sup\{ \| a_m(X) - \exp X \| \mid X \in S_R \} \le \varepsilon$. This proves that (17.1) converges uniformly to $\exp(X)$ <u>as functions on S_R</u> and hence as the uniform limit of continuous functions, $\exp(X)|_{S_R}$ is continuous [MHS, Theorem L13-5]. But if $X \in \mathcal{B}(V)$ then there is an open neighbourhood of X contained in S_R for R sufficiently large, hence $\exp$ is continuous at X, and hence on all of $\mathcal{B}(V)$. $\square$

Logarithms

The following is taken from [H, §2.3]. We take as given that the powerseries

$$\log z = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(z-1)^m}{m} \tag{18.1}$$

converges absolutely for all $z \in \mathbb{C}$ with $|z-1|<1$, and that $e^{\log z} = z$ for such z. For all u with $|u| < \log 2$, we have $|e^u - 1| < 1$ and $\log e^u = u$.

<u>Lemma B1-13</u> If T is a bounded operator on a Banach space V and $\| T - 1_V \| < 1$ then the series

$$\log(T) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{(T-1_V)^n}{n} \tag{18.2}$$

converges absolutely in $\mathcal{B}(V)$. With $U = \{ T \mid \|T - 1_V\| < 1\}$ the function $\log : U \longrightarrow \mathcal{B}(V)$ is continuous.

<u>Proof</u> This is similar to Theorem B1-7. The operators $(T-1_V)^n$ are bounded by Lemma B1-3 and satisfy $\|(T-1_V)^n\| \leq \|T-1_V\|^n$. With $z = \|T - 1_V\|$ we have by (18.1) that

$$\log(1+z) = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{z^m}{m}$$

converges absolutely and hence

$$\sum_{n=1}^{\infty} \frac{\|(T-1_V)^n\|}{n} \leq \sum_{n=1}^{\infty} \frac{z^n}{n} < \infty$$

so that (18.2) converges absolutely and hence converges (Lemma B1-4). Continuity follows in the same way as for exp: for $R<1$ sequence $a_m = \sum_{n=1}^{m} (-1)^{n+1} \frac{(T-1_V)^n}{n}$ converges uniformly to $\log T$ on $S_R = \{ T \mid \|T-1_V\| \leq R\}$ and as a uniform limit of continuous functions log is continuous on S_R. $\square$

<u>Exercise B1-7</u> Let V, W be Banach spaces and $S : V \longrightarrow W$ a norm-preserving isomorphism of vector spaces. Let $T : V \longrightarrow V$ be a bounded linear operator satisfying $\| T - 1_V \| < 1$. Prove that $\| STS^{-1} - 1_W \| < 1$ and

$$S \log(T) S^{-1} = \log(STS^{-1}).$$

Some properties of the logarithm we will prove only for V finite-dimensional, since the general arguments are (as far as I can tell) more involved.

<u>Exercise B1-8</u> Prove that the set $D \subseteq M_n(\mathbb{C})$ of diagonalisable matrices is <u>dense</u>, i.e. for every $X \in M_n(\mathbb{C})$ there is a sequence $(Y_n)_{n \geq 0}$ in D with $\lim_{n \to \infty} Y_n = X$ (the limit being say with respect to the Frobenius norm $\| - \|_F$, but by Lemma B1-1 you can choose your norm) (<u>Hint</u>: an $n \times n$ matrix with n <u>distinct</u> eigenvalues is diagonalisable. Generically, n complex numbers are distinct!).

<u>Lemma B1-14</u> Let V be a finite-dimensional complex normed space. Then with

$$\begin{aligned} U &= \{ T \in \mathcal{B}(V) \mid \| T - 1_V \| < 1 \} \\ W &= \{ X \in \mathcal{B}(V) \mid \| X \| < \log 2 \} \end{aligned} \tag{19.1}$$

we have

(i) $\exp(X) \in U$ for all $X \in W$
(ii) $\exp(\log T) = T$ for all $T \in U$
(iii) $\log(\exp X) = X$ for all $X \in W$

<u>Proof</u> Since $\mathbb{F} = \mathbb{C}$ the diagonalisable matrices are dense in $\mathcal{B}(V) \cong M_n(\mathbb{C})$ (here $n = \dim V$). Suppose we can prove $\exp(\log A) = A$ for any <u>diagonalisable</u> A. Then with $(A_m)_{m=1}^{\infty}$ a sequence of diagonalisable matrices converging to T, we have

using continuity of exp, log (we may assume $\|A_m - 1_V\| < 1$ for all m)

$$\begin{aligned} \exp(\log T) &= \exp(\log(\lim_{m\to\infty} A_m)) \\ &= \exp(\lim_{m\to\infty} \log A_m) \\ &= \lim_{m\to\infty} \exp(\log A_m) \\ &= \lim_{m\to\infty} A_m = T \end{aligned} \qquad (20.1)$$

Hence we can reduce to checking that $\exp(\log A) = A$ for diagonalisable $A \in U$. If $S^{-1}AS = D$ for $D = (d_1, \dots, d_n)$ diagonal, then

$$\begin{aligned} \exp(\log A) &= \exp(\log(S^{-1}DS)) \\ &\overset{\text{Ex B1-7}}{=} \exp(S^{-1} \log D\, S) \\ &\overset{\text{Ex B1-3}}{=} S^{-1} \exp(\log D)\, S \qquad (20.2) \\ &= S^{-1} \exp \begin{pmatrix} \log d_1 & & \\ & \ddots & \\ & & \log d_n \end{pmatrix} S \\ &= S^{-1} \begin{pmatrix} \exp(\log d_1) & & \\ & \ddots & \\ & & \exp(\log d_n) \end{pmatrix} S = S^{-1} D S = A \end{aligned}$$

Note that A, D have the same eigenvalues $d_1, \dots, d_n$ and so $|d_i - 1| \le \|A - 1_V\| < 1$ for $1 \le i \le n$ (Ex B1-10) and hence $\log d_i$ is computed by (18.1) and $\exp(\log d_i) = d_i$.

Now suppose $X \in W$ that is $\|X\| < \log 2$. Then $\exp X \in U$ since

$$\|\exp X - 1_V\| = \|\textstyle\sum_{i=1}^{\infty} \frac{1}{i!} X^i\| \overset{\text{as in (17.2)}}{\le} \sum_{i=1}^{\infty} \frac{1}{i!} \|X\|^i = \exp\|X\| - 1 < 1$$

The proof that $\log(\exp X) = X$ follows the same pattern: we first reduce by continuity to the case where X is diagonalisable and then we use a calculation like (20.2) in this case. Note that if $X = \operatorname{diag}(d_1, \dots, d_n)$ is diagonal and $\|X\| < \log 2$ then $|d_i| < \log 2$ and hence $\log(\exp d_i) = d_i$ for all $1 \le i \le n$. $\square$

<u>Exercise B1-9</u> There is no exercise B1-9.

<u>Exercise B1-10</u> Let $T : V \to V$ be a bounded linear operator on a normed space V and $\lambda \in \mathbb{F}$ an eigenvalue of T. Then $|\lambda| \le \|T\|$.

<u>Exercise B1-11</u> Prove that if $\|-\|_1, \|-\|_2$ are Lipschitz equivalent norms on a vector space V, then they induce Lipschitz equivalent operator norms on $\mathcal{B}(V)$.

<u>Remark</u> Let $U = \{ T \in M_n(\mathbb{R}) \mid \|T - I_n\| < 1 \}$, $W = \{ X \in M_n(\mathbb{R}) \mid \|X\| < \log 2 \}$ where $\|-\|$ means the operator norm with respect to $\|-\|_2$ on $\mathbb{R}^n$. Let $U_{\mathbb{C}} \subseteq M_n(\mathbb{C})$, $W_{\mathbb{C}} \subseteq M_n(\mathbb{C})$ denote the analogous subsets for $(\mathbb{C}^n, \|-\|_2)$. Under the canonical inclusion $M_n(\mathbb{R}) \subseteq M_n(\mathbb{C})$ we have $U \subseteq U_{\mathbb{C}}$, $W \subseteq W_{\mathbb{C}}$ and the exponential and logarithm for complex matrices restrict to those for real matrices, so from Lemma B1-14 we deduce also for <u>real</u> matrices that

(i) $\exp(X) \in U$ for all $X \in W$

(ii) $\exp(\log T) = T$ for all $T \in U$

(iii) $\log(\exp X) = X$ for all $X \in W$

Lemma B1-15 Let T be a bounded operator on a Banach space V with $\|T\| < 1/2$. Then

$$\|\log(1_V + T) - T\| \leq c\|T\|^2 \qquad (22.1)$$

where c is a constant independent of $\|T\|$.

Proof Since

$$\log(1_V + T) - T = \sum_{m=2}^{\infty} (-1)^{m+1} \frac{T^m}{m} = T^2 \sum_{m=2}^{\infty} (-1)^{m+1} \frac{T^{m-2}}{m}$$

we have

$$\|\log(1_V + T) - T\| \leq \|T\|^2 \sum_{m=2}^{\infty} \frac{(1/2)^{m-2}}{m}$$

and since $c = \sum_{m=2}^{\infty} \frac{(1/2)^{m-2}}{m}$ converges we are done (by the ratio test) we are done. $\square$

Remark (Big O notation) One often sees (22.1) written as

$$\log(T + 1_V) = T + O(\|T\|^2). \qquad (22.2)$$

To interpret such statements you look inside the $O(-)$ to find the "variable", in this case T, and you rearrange to obtain a bounded operator which is a function of this variable, in this case $\log(T + 1_V) - T$, and to say this is $O(\|T\|^2)$ is to say that there exists $c > 0$ and $\varepsilon > 0$ such that whenever $\|T\| < \varepsilon$ we have $\|\log(T + 1_V) - T\| \leq c\|T\|^2$.

Theorem B1-16 (Lie Product formula) Let S, T be linear operators on a finite-dimensional complex normed space V. Then

$$\exp(S+T) = \lim_{m\to\infty}\left(\exp\left(\tfrac{S}{m}\right)\exp\left(\tfrac{T}{m}\right)\right)^m$$

Proof By Taylor's theorem (with the Lagrange remainder, see L6 p. 15) we have for any $t \in \mathbb{R}$

$$\exp(t) = 1 + t + \frac{\exp(b)}{2} t^2 \qquad (23.1)$$

where b depends on t and is between 0 and t. Hence

$$\left| \exp\left(\tfrac{\|S\|}{m}\right) - 1 - \tfrac{\|S\|}{m} \right| = \frac{\exp(b)}{2}\frac{\|S\|^2}{m^2} \qquad (23.2)$$

for some $0 \le b \le \frac{\|S\|}{m}$. That is, $\sum_{j=2}^{\infty} \frac{1}{j!}\left(\frac{\|S\|}{m}\right)^j = \frac{\exp(b)}{2}\frac{\|S\|^2}{m^2}$. But then

$$\left\| \exp\left(\tfrac{S}{m}\right) - 1 - \tfrac{S}{m} \right\| \le \sum_{j=2}^{\infty} \frac{1}{j!}\left(\frac{\|S\|}{m}\right)^j = \frac{\exp(b)}{2}\frac{\|S\|^2}{m^2} \qquad (23.3)$$

Similarly for some $0 \le a \le \frac{\|T\|}{m}$ we have

$$\left\| \exp\left(\tfrac{T}{m}\right) - 1 - \tfrac{T}{m} \right\| \le \sum_{j=2}^{\infty} \frac{1}{j!}\left(\frac{\|T\|}{m}\right)^j = \frac{\exp(a)}{2}\frac{\|T\|^2}{m^2} \qquad (23.4)$$

Note that $\exp\left(\frac{S}{m}\right) = 1 + \frac{S}{m} + O\left(\frac{1}{m^2}\right)$ in the following precise sense: if $\|S\| < \log 2$ then $\exp(b) \le \exp\left(\frac{\|S\|}{m}\right) \le \exp(\|S\|) < 2$ and so

$$\left\| \exp\left(\tfrac{S}{m}\right) - 1 - \tfrac{S}{m} \right\| \le (\log 2)^2 \frac{1}{m^2} \le \frac{1}{m^2} \qquad (23.5)$$

We claim that

$$\exp\left(\tfrac{S}{m}\right)\exp\left(\tfrac{T}{m}\right) = 1_V + \frac{S}{m} + \frac{T}{m} + O\left(\frac{1}{m^2}\right) \qquad (23.6)$$

in a sense to be made precise in a moment. Set $\mathcal{L} = \exp(\frac{S}{m}) - 1_V - \frac{S}{m}$ and $\mathcal{R} = \exp(\frac{T}{m}) - 1_V - \frac{T}{m}$.
Then

$$\begin{aligned} \exp(\tfrac{S}{m})\exp(\tfrac{T}{m}) &= \left(1_V + \tfrac{S}{m} + \mathcal{L}\right)\left(1_V + \tfrac{T}{m} + \mathcal{R}\right) \\ &= 1_V + \tfrac{T}{m} + \mathcal{R} + \tfrac{S}{m} + \tfrac{1}{m^2}ST + \tfrac{1}{m}S\mathcal{R} \\ &\quad + \mathcal{L} + \tfrac{1}{m}\mathcal{L}T + \mathcal{L}\mathcal{R}. \end{aligned} \tag{24.1}$$

Suppose $\|S\| < \log 2$, $\|T\| < \log 2$. Then by (23.5) $\|\mathcal{L}\| \leq (\log 2)^2 \frac{1}{m^2}$, $\|\mathcal{R}\| \leq (\log 2)^2 \frac{1}{m^2}$ and so for all $m \geq 1$

$$\left\| \exp(\tfrac{S}{m})\exp(\tfrac{T}{m}) - 1_V - \tfrac{S}{m} - \tfrac{T}{m} \right\|$$

$$= \left\| \mathcal{R} + \mathcal{L} + \tfrac{1}{m^2}ST + \tfrac{1}{m}S\mathcal{R} + \tfrac{1}{m}\mathcal{L}T + \mathcal{L}\mathcal{R} \right\|$$

$$\begin{aligned} &\leq \|\mathcal{R}\| + \|\mathcal{L}\| + \tfrac{1}{m^2}\|S\|\,\|T\| + \tfrac{1}{m}\|S\|\,\|\mathcal{R}\| + \\ &\quad \tfrac{1}{m}\|\mathcal{L}\|\,\|T\| + \|\mathcal{L}\|\,\|\mathcal{R}\| \end{aligned} \tag{24.2}$$

$$\begin{aligned} &\leq (\log 2)^2 \tfrac{1}{m^2} + (\log 2)^2 \tfrac{1}{m^2} + \tfrac{1}{m^2}(\log 2)^2 + \tfrac{1}{m^3}(\log 2)^3 \\ &\quad + \tfrac{1}{m^3}(\log 2)^3 + (\log 2)^4 \tfrac{1}{m^4} \end{aligned}$$

$$\leq 6(\log 2)^2 \frac{1}{m^2} \leq \frac{6}{m^2}$$

which proves (23.6). Since $\frac{S}{m} \to 0$, $\frac{T}{m} \to 0$ in $\mathcal{B}(V)$ as $m \to \infty$ and both the exponential and multiplication are continuous (Lemma B1-8 and Lemma B1-12) we have $\exp(\frac{S}{m})\exp(\frac{T}{m}) \to 1_V$ as $m \to \infty$ and hence for m sufficiently large this product is in the domain of the logarithm. Hence by (23-6) for sufficiently large m and $\|S\| < \log 2$, $\|T\| < \log 2$

$$\log\left(\exp(\tfrac{S}{m})\exp(\tfrac{T}{m})\right) = \log\left(1_V + \tfrac{S}{m} + \tfrac{T}{m} + O(\tfrac{1}{m^2})\right) \tag{24.3}$$

$\swarrow q = 2\log 2 + 6$

Now $\left\| \frac{S}{m} + \frac{T}{m} + O\left(\frac{1}{m^2}\right) \right\| \leq \frac{\|S\|}{m} + \frac{\|T\|}{m} + \frac{6}{m^2} \leq q\,\frac{1}{m}$ which is strictly less than $1/2$ for m sufficiently large. Hence for such m, by Lemma B1-15

$$\log\left(\exp\left(\frac{S}{m}\right)\exp\left(\frac{T}{m}\right)\right) = \frac{S}{m} + \frac{T}{m} + O\left(\left\| \frac{S}{m} + \frac{T}{m} + O\left(\frac{1}{m^2}\right)\right\|^2 \right) \qquad (25.1)$$

$$= \frac{S}{m} + \frac{T}{m} + O\left(\frac{1}{m^2}\right)$$

Hence by Lemma B1-14 applying exp to both sides yields

$$\exp\left(\frac{S}{m}\right)\exp\left(\frac{T}{m}\right) = \exp\left(\frac{S}{m} + \frac{T}{m} + O\left(\frac{1}{m^2}\right)\right) \qquad (25.2)$$

Hence by Lemma B1-11 (iii) for $\|S\| < \log 2$, $\|T\| < \log 2$ and $m \gg 0$

$$\left[\exp\left(\frac{S}{m}\right)\exp\left(\frac{T}{m}\right)\right]^m = \exp\left(S + T + O\left(\frac{1}{m}\right)\right).$$

By continuity of the exponential

$$\lim_{m\to\infty}\left[\exp\left(\frac{S}{m}\right)\exp\left(\frac{T}{m}\right)\right]^m = \exp\left(\lim_{m\to\infty}\left(S + T + O\left(\frac{1}{m}\right)\right)\right)$$

$$= \exp(S + T). \quad \square$$

Acknowledgements Thanks to Abraham Zhang and Brian Chan for pointing out an error in the original statement of Lemma B1-14.

References

[H] B. Hall "Lie Groups, Lie Algebras, and Representations" Springer GTM.