Cases of Surjectivity of the Matrix Exponential into Lie Groups: $\mathfrak{so}(3) \to \mathbf{SO}(3)$ and $M_{n\times n}(\mathbb{C}) \to \mathbf{GL}_n(\mathbb{C})$

Larsen Bier

MATH 110B Winter 2025

UCLA

1 Introduction

In this short paper, we prove some properties of the matrix exponential and logarithm, guiding us to two examples of when the exponential map is surjective: $\mathfrak{so}(3) \to \mathrm{SO}(3)$ and $M_{n\times n}(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})$.

2 Matrix Lie Groups

Let $M_{n\times m}(\Omega)$ denote the set of $n \times m$ matrices over the set Ω. Following the convention of [1], we define a matrix lie group as a topologically closed subgroup of $\mathrm{GL}_n(\mathbb{C})$. All norms induce the same open sets in finite-dimensional vector spaces (see §1.4 of [2]), so we choose the operator norm: $\|\mathbf{X}\|_{\mathrm{op}} := \sup\{\|\mathbf{Xv}\|_2 : \|\mathbf{v}\|_2 = 1\}$, where $\|\cdot\|_2$ is the Euclidean norm on $\mathbb{C}^n$.

3 The Matrix Exponential

Given a matrix $\mathbf{A} \in M_{n\times n}(\mathbb{C})$, we define the matrix exponential as

$$\exp(\mathbf{A}) := e^{\mathbf{A}} := \sum_{k=0}^{\infty} \frac{\mathbf{X}^k}{k!}.$$

This extends the familiar definition of exponentiation from $\mathbb{C}$ to square matrices over $\mathbb{C}$. Before we can prove that this power series converges, we need two facts. The first is Theorem 1.3.9 of *An Introduction to Banach Space Theory* [3]. We restate the theorem below but skip the proof for brevity.

Proposition 1. *Consider any series $\mathbf{s}_m = \sum_{k=0}^{m} \mathbf{x}_k$ in a complete normed vector space $(X, \|\cdot\|)$. If $\sum_{k=0}^{\infty} \|\mathbf{s}_k\| < \infty$ (i.e. the series converges absolutely), then $\mathbf{x}_m$ converges in $(X, \|\cdot\|)$.*

This allows us to prove the convergence of $e^{\mathbf{A}}$ by showing that it is absolutely convergent. Absolute convergence follows from our next proposition.

Proposition 2. *If $\mathbf{X}, \mathbf{Y}$ are any two elements of $M_{n\times n}(\mathbb{C})$, then $\|\mathbf{XY}\|_{op} \le \|\mathbf{X}\|_{op}\|\mathbf{Y}\|_{op}$.*

Proof: Recall that $\|\mathbf{XY}\|_{\mathrm{op}} = \sup\{\|\mathbf{XYv}\|_2 : \|\mathbf{v}\|_2 = 1\}$. Let $\mathbf{u} = \mathbf{Yv}$. Then $\|\mathbf{u}\|_2 \le \|\mathbf{Y}\|_{\mathrm{op}}\|\mathbf{v}\|_2 = \|\mathbf{Y}\|_{\mathrm{op}}$ by the definition of the operator norm. Likewise, $\|\mathbf{Xu}\|_2 \le \|\mathbf{X}\|_{\mathrm{op}}\|\mathbf{u}\|_2 \le \|\mathbf{X}\|_{\mathrm{op}}\|\mathbf{Y}\|_{\mathrm{op}}$. Combining the previous two inequalities yields $\|\mathbf{XYv}\|_2 \le \|\mathbf{X}\|_{\mathrm{op}}\|\mathbf{Y}\|_{\mathrm{op}}$ for all $\mathbf{v} \in \mathbb{C}^n$ with unit norm. Thus, the supremum over all such $\mathbf{v}$ will be less than or equal to $\|\mathbf{X}\|_{\mathrm{op}}\|\mathbf{Y}\|_{\mathrm{op}}$. $\square$

We can now easily show that $e^{\mathbf{A}}$ is convergent for any square matrix over $\mathbb{C}$. Note that $\|\mathbf{A}^m\|_{\mathrm{op}} \le \|\mathbf{A}\|_{\mathrm{op}}^m$ by Proposition 2, so

$$\sum_{k=0}^{\infty} \Big\|\frac{\mathbf{A}^k}{k!}\Big\|_{\mathrm{op}} \le \sum_{k=0}^{\infty} \frac{1}{k!}\|\mathbf{A}\|_{\mathrm{op}}^k = e^{\|\mathbf{A}\|_{\mathrm{op}}} < \infty.$$

So $e^{\mathbf{A}}$ converges by Proposition 1. Furthermore, this function is continuous since for any closed ball of radius $r > 0$ centered around $\mathbf{0}$ in $M_{n\times n}(\mathbb{C})$,

$$\Big\|e^{\mathbf{X}} - \sum_{k=0}^{n} \frac{\mathbf{X}^k}{k!}\Big\|_{\mathrm{op}} = \Big\|\sum_{k=0}^{\infty} \frac{\mathbf{X}^k}{k!} - \sum_{k=0}^{n} \frac{\mathbf{X}^k}{k!}\Big\|_{\mathrm{op}} = \Big\|\sum_{k=n+1}^{\infty} \frac{\mathbf{X}^k}{k!}\Big\|_{\mathrm{op}} \le \sum_{k=n+1}^{\infty} \frac{1}{k!}\|\mathbf{X}\|_{\mathrm{op}} \le \sum_{k=n+1}^{\infty} \frac{r^k}{k!}.$$

Since $\sum_{k=0}^{\infty} \frac{r^k}{k!}$ converges to e^r, the tail end of the partial sums must approach zero, meaning that for any $\varepsilon > 0$, we can choose an n_ε large enough that $\sum_{k=n_\varepsilon+1}^{\infty} \left(\frac{r^k}{k!}\right) < \varepsilon$. Since this is true for all $\mathbf{X}$ with $\|\mathbf{X}\|_{\text{op}} \leq r$, the partial sums converge uniformly to $e^{\mathbf{X}}$ on this set. Since any compact set in $M_{n\times n}(\mathbb{C})$ can be covered by a closed ball of large enough radius, this means the series converges uniformly on all compact subsets of $M_{n\times n}(\mathbb{C})$. That means that the series is normally convergent, so the fact that the partial sums are continuous means that the limit function is continuous (See §5.2 of [4]).

Proposition 3. *For any $\mathbf{X}, \mathbf{Y} \in M_{n\times n}(\mathbb{C})$, the following facts hold.*

(a) *if $\mathbf{X}$ and $\mathbf{Y}$ commute, then $e^{\mathbf{X}+\mathbf{Y}} = e^{\mathbf{X}}e^{\mathbf{Y}}$.*

(b) *if $\mathbf{C}$ is invertible, then $e^{\mathbf{C}\mathbf{X}\mathbf{C}^{-1}} = \mathbf{C}e^{\mathbf{X}}\mathbf{C}^{-1}$.*

Proof of part (a): Because the matrix exponential converges absolutely, we can multiply the series term by term according to Merten's Theorem (Theorem 3.50 of [5]), so

$$e^{\mathbf{X}}e^{\mathbf{Y}} = (\mathbf{I} + \mathbf{X} + \frac{\mathbf{X}^2}{2!} + \frac{\mathbf{X}^3}{3!} + \cdots)(\mathbf{I} + \mathbf{Y} + \frac{\mathbf{Y}^2}{2!} + \frac{\mathbf{Y}^3}{3!} + \cdots).$$

The terms of combined degree m are the product of an $\mathbf{X}$ term of degree $k \leq m$ and a $\mathbf{Y}$ term of degree $m-k$, so the Cauchy product simplifies to

$$e^{\mathbf{X}}e^{\mathbf{Y}} = \sum_{m=0}^{\infty}\sum_{k=0}^{m} \frac{\mathbf{X}^k\mathbf{Y}^{m-k}}{k!(m-k)!} = \sum_{m=0}^{\infty} \frac{1}{m!}\sum_{k=0}^{m} \frac{m!\mathbf{X}^k\mathbf{Y}^{m-k}}{k!(m-k!)} = \sum_{m=0}^{\infty} \frac{1}{m!}\sum_{k=0}^{m} \binom{m}{k}\mathbf{X}^k\mathbf{Y}^{m-k}. \tag{3.1}$$

In (3.1) we have what looks like the binomial theorem. Note that the binomial coefficients assume that we can choose products in any order, i.e. that $\mathbf{X}$ and $\mathbf{Y}$ commute. Luckily, they do, so

$$e^{\mathbf{X}}e^{\mathbf{Y}} = \sum_{m=0}^{\infty} \frac{1}{m!}(\mathbf{X}+\mathbf{Y})^m = e^{\mathbf{X}+\mathbf{Y}}.$$

Proof of part (b): The key is to recognize that $(\mathbf{C}\mathbf{X}\mathbf{C}^{-1})^n = \mathbf{C}\mathbf{X}\mathbf{C}^{-1}\cdots\mathbf{C}\mathbf{X}\mathbf{C}^{-1} = \mathbf{C}\mathbf{X}^n\mathbf{C}^{-1}$. Thus,

$$\begin{aligned} e^{\mathbf{C}\mathbf{X}\mathbf{C}^{-1}} &= \lim_{m\to\infty}\sum_{n=0}^{m} \frac{1}{n!}(\mathbf{C}\mathbf{X}\mathbf{C}^{-1})^n \\ &= \lim_{m\to\infty}\sum_{n=0}^{m} \frac{1}{n!}\mathbf{C}\mathbf{X}^n\mathbf{C}^{-1} \\ &= \mathbf{C}\Big(\lim_{m\to\infty}\sum_{n=0}^{m} \frac{1}{n!}\mathbf{X}^n\Big)\mathbf{C}^{-1} \\ &= \mathbf{C}e^{\mathbf{X}}\mathbf{C}^{-1}. \quad \square \end{aligned} \tag{3.2}$$

Proposition 3 is a powerful tool for computing exponentials and will aid us in proving the following results.

4 The Matrix Logarithm

Just like for the exponential, we can use our existing definitions from the complex numbers to extend the logarithm to $M_{n\times n}(\mathbb{C})$. For $\mathbf{A} \in M_{n\times n}(\mathbb{C})$, we define the matrix logarithm as

$$\log(\mathbf{A}) = \sum_{m=1}^{\infty} (-1)^{m+1}\frac{(\mathbf{A}-\mathbf{I})^m}{m}.$$

If we consider the absolute convergence of $\log \mathbf{A}$, we can see that

$$\sum_{m=0}^{\infty} \Big\|(-1)^{m+1}\frac{(\mathbf{A}-\mathbf{I})^m}{m}\Big\|_{\text{op}} = \sum_{m=0}^{\infty} \frac{1}{m}\big\|(\mathbf{A}-\mathbf{I})^m\big\|_{\text{op}} \leq \sum_{m=0}^{\infty} \|\mathbf{A}-\mathbf{I}\|_{\text{op}}^m. \tag{4.1}$$

Since $\|\mathbf{A}-\mathbf{I}\|_{\text{op}} < 1$, (4.1) converges by the geometric series test. So $\log \mathbf{A}$ converges by Proposition 1.

Proposition 4. *log is continuous on the set* $\{\mathbf{A} \in M_{n\times n}(\mathbb{C}) : \|\mathbf{A}-\mathbf{I}\|_{op} < 1\}$.

Proof: Let $\mathbf{X} = \mathbf{A} - \mathbf{I}$. Then by assumption, $\|\mathbf{X}\|_{\text{op}} < 1$, and

$$\log \mathbf{A} = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}(\mathbf{A}-\mathbf{I})^n = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}(\mathbf{X})^n.$$

Let $R \in (0,1)$. Let $\|\mathbf{X}\|_{\text{op}} < R$. Then

$$\begin{aligned}
\Big\| \log \mathbf{A} - \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}\mathbf{X}^n \Big\|_{\text{op}} &= \Big\| \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}\mathbf{X}^n - \sum_{n=1}^{m} \frac{(-1)^{n+1}}{n}\mathbf{X}^n \Big\|_{\text{op}} \\
&= \Big\| \sum_{n=m+1}^{\infty} \frac{(-1)^{n+1}}{n}\mathbf{X}^n \Big\|_{\text{op}} \\
&\le \sum_{n=m+1}^{\infty} \Big| \frac{(-1)^{n+1}}{n} \Big| \|\mathbf{X}\|_{\text{op}}^n \\
&\le \sum_{n=m+1}^{\infty} \Big| \frac{(-1)^{n+1}}{n} \Big| R^n.
\end{aligned} \tag{4.2}$$

Since $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} R^n = \log(R+1)$, we know that the tail of the sum approaches 0, so as $m \to \infty$, the last line of (4.1) gets arbitrarily small. That is, for any $\varepsilon > 0$, there is some positive integer M such that $\sum_{n=m+1}^{\infty} \frac{(-1)^n+1}{n} R^n < \varepsilon$. Since this bound does not depend on which $\mathbf{X}$ we choose, the series converges uniformly to $\log \mathbf{A}$ on $\{\|\mathbf{A}-\mathbf{I}\|_{\text{op}} < 1\}$. This is true for all $R \in (0,1)$, so the series converges normally on $\{\|\mathbf{A}-\mathbf{I}\|_{\text{op}} < 1\}$. Since each of the partial sums is a continuous polynomial and the convergence is normal, the limit ($\log \mathbf{A}$) is continuous on $\{\|\mathbf{A}-\mathbf{I}\|_{\text{op}} < 1\}$. $\square$

Proposition 5. *For all* $\mathbf{A}, \mathbf{B} \in M_{n\times n}(\mathbb{C})$*:*

(a) *If* $\|\mathbf{A}-\mathbf{I}\|_{op} < 1$*, then* $\exp(\log \mathbf{A}) = \mathbf{A}$.

(b) *If* $\mathbf{B}$ *has* $\|\mathbf{B}\|_{op} < \log 2$*, then* $\log(\exp \mathbf{B}) = \mathbf{B}$.

Proof of part (a): Let $\mathbf{X} = \mathbf{A} - \mathbf{I}$. Then $\exp(\log \mathbf{A}) = \mathbf{A}$ if and only if $\exp(\log(\mathbf{I}+\mathbf{X})) = \mathbf{I}+\mathbf{X}$. We break this proof down into two cases.

Case 1: X is diagonalizable. Then $\mathbf{X} = \mathbf{C}\mathbf{D}\mathbf{C}^{-1}$ where $\mathbf{D}$ is a diagonal matrix whose entries are the (not necessarily distinct) eigenvalues of $\mathbf{A}$. Since $\|\mathbf{X}\|_{\text{op}} < 1$ by assumption, each eigenvalue λ_i has modulus less than 1 (since otherwise $\mathbf{X}$ scales some vector by more than $\|\mathbf{X}\|_{\text{op}}$, which is impossible). Thus, $\lambda_i + 1$ is contained in the ball of radius 1 around 1 in $\mathbb{C}$, so $\log(\lambda_i + 1)$ converges. So

$$\begin{aligned}
\log(\mathbf{X}+\mathbf{I}) &= \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m}\mathbf{X}^m \\
&= \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m}(\mathbf{C}\mathbf{D}\mathbf{C}^{-1})^m \\
&= \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m}\mathbf{C}\begin{pmatrix} \lambda_1^m & 0 & \dots & 0 \\ 0 & \lambda_2^m & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^m \end{pmatrix}\mathbf{C}^{-1} \\
&= \mathbf{C}\begin{pmatrix} \log(\lambda_1+1) & 0 & \dots & 0 \\ 0 & \log(\lambda_2+1) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \log(\lambda_n+1) \end{pmatrix}\mathbf{C}^{-1}.
\end{aligned} \tag{4.3}$$

Since $|(\lambda_i+1)-1|<1$, $e^{\log(\lambda_i+1)}=\lambda_i+1$. Thus,

$$\begin{aligned}
\exp(\log(\mathbf{X}+\mathbf{I})) &= \mathbf{C}\begin{pmatrix}\lambda_1+1 & 0 & \dots & 0\\ 0 & \lambda_2+1 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & \lambda_n+1\end{pmatrix}\mathbf{C}^{-1}\\
&= \mathbf{C}(\mathbf{D}+\mathbf{I})\mathbf{C}^{-1}\\
&= \mathbf{C}\mathbf{D}\mathbf{C}^{-1}+\mathbf{I}\\
&= \mathbf{X}+\mathbf{I}.
\end{aligned} \tag{4.4}$$

Case 2: X is not diagonalizable. Then since the characteristic polynomial of $\mathbf{X}$ splits over $\mathbb{C}$, it has a Jordan normal decomposition $\mathbf{X}=\mathbf{B}\mathbf{J}\mathbf{B}^{-1}$, where $\mathbf{J}$ has the eigenvalues of $\mathbf{X}$ on its diagonal. Let $\mathbf{D}$ be the diagonal matrix with diagonal entries $1,2,\dots,n$. Then let $\mathbf{X}_k=\mathbf{B}(\mathbf{J}+\frac{1}{k}\mathbf{D})\mathbf{B}^{-1}$ define a sequence that converges to $\mathbf{X}$. $\mathbf{X}_k$ is written in Jordan normal form (since the addition of $\frac{1}{k}\mathbf{D}$ only changes the diagonal entries). Thus, the i^{th} eigenvalue of $\mathbf{X}_k$ is $\mathbf{J}_{ii}+\frac{i}{k}$. For sufficiently large k, each of these values will be distinct, meaning $\mathbf{X}_k$ will be diagonalizable. Furthermore, since $\mathbf{X}_k\to\mathbf{X}$ and $\|\mathbf{X}\|_{\text{op}}<1$, eventually $\|\mathbf{X}_k\|_{\text{op}}<1$. By the continuity of log when $\|\mathbf{X}\|_{\text{op}}<1$ and the global continuity of exp, we can swap limits for this sequence, so

$$\exp(\log(\mathbf{X}+\mathbf{I}))=\exp(\log(\lim_{k\to\infty}[\mathbf{X}_k+\mathbf{I}]))=\lim_{k\to\infty}\exp(\log(\mathbf{X}_k+\mathbf{I}))=\lim_{k\to\infty}\mathbf{X}_k+\mathbf{I}=\mathbf{X}+\mathbf{I}. \tag{4.5}$$

Where we were able to cancel exp and log in the fourth expression using the diagonalizable case.

Proof of part (b): First we note that if $\|\mathbf{B}\|_{\text{op}}<\log 2$, then

$$\begin{aligned}
\|e^{\mathbf{B}}-\mathbf{I}\|_{\text{op}} &= \|(\mathbf{I}+\mathbf{B}+\frac{\mathbf{B}^2}{2!}+\frac{\mathbf{B}^3}{3!}+\cdots)-\mathbf{I}\|_{\text{op}}\\
&\le \|\mathbf{B}\|_{\text{op}}+\left\|\frac{\mathbf{B}^2}{2!}\right\|_{\text{op}}+\left\|\frac{\mathbf{B}^3}{3!}\right\|_{\text{op}}+\cdots\\
&= e^{\|\mathbf{B}\|_{\text{op}}}-1\\
&< e^{\log 2}-1\\
&= 1.
\end{aligned} \tag{4.6}$$

Again, if $\mathbf{B}$ is diagonalizable, then $\mathbf{B}=\mathbf{C}\mathbf{D}\mathbf{C}^{-1}$, where $\mathbf{D}$ is a diagonal matrix of the eigenvalues $\lambda_1,\dots,\lambda_n$ of $\mathbf{B}$ and $\mathbf{C}$ is invertible. Since the exponential of a diagonal matrix is taken entry-wise, $e^{\mathbf{D}}_{ii}=e^{\lambda_i}$. Furthermore, these are the eigenvalues of $e^{\mathbf{D}}$ because $e^{\mathbf{B}}=\mathbf{C}e^{\mathbf{B}}\mathbf{C}^{-1}$ is a diagonalization of $\mathbf{B}$.To get a bound on the modulus of these eigenvalues, consider that subtracting $\mathbf{I}$ from a matrix decreases all its eigenvalues by 1, so the eigenvalues of $e^{\mathbf{B}}-\mathbf{I}$ are $e^{\lambda_i}-1$. Since $\|e^{\mathbf{B}}-\mathbf{I}\|_{\text{op}}<1$, each eigenvalue must have modulus less than 1 as shown in case 1 of part (a). Thus, $|e^{\lambda_i}-1|<1$, so $\log e^{\lambda_i}=\lambda_i$. Since the log of a diagonal matrix is also taken entry-wise [as we showed in equation (4.2)], that means $\log(e^{\mathbf{D}})=\mathbf{D}$. In (4.2), we also showed that $\log(\mathbf{C}\mathbf{D}\mathbf{C}^{-1})=\mathbf{C}\log(\mathbf{A})\mathbf{C}^{-1}$ for invertible $\mathbf{C}$ and diagonal $\mathbf{A}$. Thus,

$$\begin{aligned}
\log e^{\mathbf{B}} &= \log(\mathbf{C}e^{\mathbf{D}}\mathbf{C}^{-1})\\
&= \mathbf{C}\log(e^{\mathbf{D}})\mathbf{C}^{-1}\\
&= \mathbf{C}\mathbf{D}\mathbf{C}^{-1}\\
&= \mathbf{B}.
\end{aligned} \tag{4.7}$$

In the case that $\mathbf{B}$ is not diagonalizable, we can approximate it by a sequence of diagonalizable matrices like in part (a), using the continuity of log and exp to get the desired result. $\square$

5 Cases of Surjectivity of The Exponential

We now have the tools needed to investigate cases where the exponential map is surjective.

5.1 Surjectivity of $\exp : M_{n\times n}(\mathbb{C}) \to \mathbf{GL}_n(\mathbb{C})$

To prove that this map is surjective, we follow the outline of exercises 2.8 and 2.9 from *Lie Groups, Lie Algebras, and Representations* [1]. First we require some definitions. A square matrix $\mathbf{N}$ is said to be nilpotent if there exists a positive integer k such that $\mathbf{N}^k = \mathbf{0}$. A square matrix $\mathbf{U}$ is unipotent if $\mathbf{U} - \mathbf{I}$ is nilpotent.

Proposition 6. *If* $\mathbf{U}$ *is unipotent, then* $\log(\mathbf{U})$ *is nilpotent, and if* $\mathbf{N}$ *is nilpotent, then* $\exp(\mathbf{N})$ *is unipotent.*

Proof: We first must show that finite sums of commuting nilpotent matrices are nilpotent. Let $\mathbf{N}, \mathbf{M}$ be nilpotent matrices that commute with eachother. Then there exists some $k \in \mathbb{N}$ such that $\mathbf{N}^k = \mathbf{M}^k = \mathbf{0}$. By the binomial theorem: $(\mathbf{N}+\mathbf{M})^p = \sum_{j=1}^{p} \binom{p}{j} \mathbf{N}^j \mathbf{M}^{p-j}$ (this is where we needed commuting matrices). If we choose $p \geq 2k$, then $p-j \geq k$ or $j \geq k$ since $(p-j)+j = p \geq 2k$. Thus, for every term in $(\mathbf{N}+\mathbf{M})^p$, either $\mathbf{N}^j$ or $\mathbf{M}^{p-j}$ is $\mathbf{0}$, meaning the whole sum is $\mathbf{0}$. So, $\mathbf{N}+\mathbf{M}$ is nilpotent. Continuing inductively, we can see this is true for any finite sum (since any term in the sum will commute).

Now we can prove the proposition. Since $\mathbf{U}$ is unipotent, $\mathbf{U} - \mathbf{I}$ is nilpontent, so there are finitely many $k \in \mathbb{N}$ such that $(\mathbf{U} - \mathbf{I})^k \neq \mathbf{0}$. The power series of $\log(\mathbf{U})$ becomes

$$\log(\mathbf{U}) = \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} (\mathbf{U} - \mathbf{I})^m = \sum_{m=1}^{k} \frac{(-1)^{m+1}}{m} (\mathbf{U} - \mathbf{I})^m. \tag{5.1}$$

Where for the last equality we used the fact that eventually every summand is $\mathbf{0}$. Since each term in the sum is a scaled positive power of $\mathbf{U} - \mathbf{I}$, the terms all commute. Thus, this sum is nilpotent, proving part one of the proposition.

For part two, since $\mathbf{N}$ is nilpotent, there are finitely many $k \in \mathbb{N}$ such that $\mathbf{N}^k \neq \mathbf{0}$. The power series expansion of $\exp(\mathbf{N})$ can be rewritten as

$$\exp(\mathbf{N}) = \sum_{m=0}^{\infty} \frac{1}{m!} \mathbf{N}^m = \sum_{m=0}^{k} \frac{1}{m!} \mathbf{N}^m = \mathbf{I} + \sum_{m=1}^{k} \frac{1}{m!} \mathbf{N}^m. \tag{5.2}$$

Since the terms of the summation are all scaled positive powers of $\mathbf{N}$, they commute, so the summation is nilpotent. Thus, $\exp(\mathbf{N})$ is unipotent. $\square$

These facts help us expand the cases where log is the inverse of exp, as the next proposition will show.

Proposition 7. *If* $\mathbf{U}$ *is unipotent, then* $\exp(\log \mathbf{U}) = \mathbf{U}$*, and if* $\mathbf{N}$ *is nilpotent, then* $\log(\exp \mathbf{N}) = \mathbf{N}$*.*

Proof:. Let $\mathbf{U}(t) = \mathbf{I} + t(\mathbf{U} - \mathbf{I})$ be a smooth curve in $M_{n\times n}(\mathbb{C})$ parameterized by $t \in \mathbb{R}$. By Proposition 6, since $\mathbf{U}$ is unipotent, the power series of $\log \mathbf{U}$ is finite, hence a polynomial of t. Since $\log \mathbf{U}$ is nilpotent, the power series of $\exp \log(\mathbf{U})$ is finite by Proposition 6, so it is a polynomial over t. Furthermore, $\|\mathbf{U}(t) - \mathbf{I}\|_{\text{op}} = \|t(\mathbf{U} - \mathbf{I})\|_{\text{op}}$, which approaches zero as t does. Thus, we can pick some $\varepsilon > 0$ such that for all $t \in [0, \varepsilon] : \|\mathbf{U}(t) - \mathbf{I}\|_{\text{op}} < 1$. Thus, $\exp(\log \mathbf{U}(t)) = \mathbf{U}(t)$ on $[0, \varepsilon]$ by Proposition 5. If we consider $[0, \varepsilon]$ to be a subset of $\mathbb{C}$, then for the i^{th} row and j^{th} column of our matrices, $\exp \log(\mathbf{U}(t))_{i,j}$ and $\mathbf{U}(t)_{i,j}$ are two polynomials over $\mathbb{C}$ that agree on $[0, \varepsilon]$, which has a nonisolated point. By the identity principle (See §5.7 of [4]), this means $\exp \log(\mathbf{U}(t))_{i,j} = \mathbf{U}(t)_{i,j}$ on all of $\mathbb{C}$. Since all entries agree, $\exp(\log(\mathbf{U}(t)) = \mathbf{U}(t)$ on all of $\mathbb{C}$, namely at $t = 1$. Since $\mathbf{U}(1) = \mathbf{U}$, we have $\exp(\log \mathbf{U}) = \mathbf{U}$ for all unipotent matrices.

The proof proceeds similarly for $\mathbf{N}$. If we let $\mathbf{N}(t) = t\mathbf{N}$, then as a consequence of Proposition 7, $\log(\exp \mathbf{N}(t))$ depends polynomially on t. Since $\|\mathbf{N}(t)\|_{\text{op}} = |t| \|\mathbf{N}\|_{\text{op}}$, we can again choose a small $\varepsilon > 0$ such that for all $t \in [0, \varepsilon] : \|\mathbf{N}(t)\|_{\text{op}} < \log(2)$. Thus, on $[0, \varepsilon]$, $\log(\exp \mathbf{N}(t)) = \mathbf{N}(t)$. Applying the identity principle element-wise as done above will yield the desired result. $\square$

With Propositions 5 and 6 under our belts, we are ready to prove $\exp : M_{n\times n}(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})$ is surjective. From the fundamental theorem of algebra, the characteristic polynomial of any matrix $\mathbf{A} \in \mathrm{GL}_n(\mathbb{C})$ will split over $\mathbb{C}$. Thus, there is a Jordan normal decomposition of the form $\mathbf{A} = \mathbf{CJC}^{-1}$ where $\mathbf{C}$ is an invertible matrix and $\mathbf{J}$ is a Jordan block matrix (see §7 of [6]). Jordan block matrices are of the form

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{J}_k \end{bmatrix}. \tag{5.3}$$

Where each $\mathbf{J}_i = \lambda_i\mathbf{I}+\mathbf{N}_i$ for a nilpotent matrix $\mathbf{N}_i$ and a nonzero complex number λ_i (the λ_i are nonzero since $\mathbf{A}$ is invertible). We first show that any Jordan block is in the image of the exponential, then show that combining these preimages results in a preimage for $\mathbf{J}$. Let $\mathbf{B} = \lambda\mathbf{I}+\mathbf{N} = \lambda\mathbf{I}(\mathbf{I}+\mathbf{N}/\lambda)$ be any Jordan block. Since $\lambda \neq 0$, there exists some $z \in \mathbb{C}$ such that $\lambda = e^z$. Thus, $\lambda\mathbf{I} = \exp(z\mathbf{I})$ (This is because exponentials of diagonal matrices are taken entry-wise). So we have $\mathbf{B} = \exp(z\mathbf{I})(\mathbf{I}+\mathbf{N}/\lambda)$. Since $(\mathbf{I}+\mathbf{N}/\lambda)$ is unipotent, $\exp(\log(\mathbf{I}+\mathbf{N}/\lambda)) = \mathbf{I}+\mathbf{N}/\lambda$ by Proposition 7. Thus, $\mathbf{B} = \exp(z\mathbf{I})\exp(\log(\mathbf{I}+\mathbf{N}/\lambda))$. Scalar multiples of the identity matrix commute with any other matrix , so $\mathbf{B} = \exp(z\mathbf{I}+\log(\mathbf{I}+\mathbf{N}/\lambda))$ by Proposition 3. Thus, any invertible Jordan block is in the image of the exponential.

If we let $\mathbf{L}_i = z_i\mathbf{I}+\log(\mathbf{I}+\mathbf{N}_i/\lambda_i)$ be the preimage of $\mathbf{J}_i$, then substituting $\mathbf{L}_i$ into the position of $\mathbf{J}_i$ for each $i \in \{1,\dots,k\}$ yields a matrix $\mathbf{L}$ whose image under exp is $\mathbf{J}$. This is because every power of $\mathbf{L}$ looks like

$$\mathbf{L}^m = \begin{bmatrix} \mathbf{L}_1^m & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_2^m & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{L}_k^m \end{bmatrix}. \tag{5.4}$$

Taking the limit of the partial sums of $\exp(\mathbf{L})$ yields

$$\exp(\mathbf{L}) = \lim_{m\to\infty} \begin{bmatrix} \mathbf{L}_1^m/m! & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_2^m/m! & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{L}_k^m/m! \end{bmatrix} = \begin{bmatrix} \exp(\mathbf{L}_1) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \exp(\mathbf{L}_2) & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \exp(\mathbf{L}_k) \end{bmatrix}. \tag{5.5}$$

Where we used the fact that limits of matrices are taken element wise. Since $\exp(\mathbf{L_i}) = \mathbf{J}_i$, this matrix is equal to $\mathbf{J}$. Finally, since $\mathbf{C}$ is invertible, $\exp(\mathbf{CLC}^{-1}) = \mathbf{C}\exp(\mathbf{L})\mathbf{C}^{-1} = \mathbf{CJC}^{-1} = \mathbf{A}$ by Proposition 3, concluding the proof that $\exp : M_{n\times n}(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})$ is surjective $\square$

5.2 Surjectivity of $\exp : \mathfrak{so}(3) \to \mathbf{SO}(3)$

The following definition is taken from §2.5 of *Lie Groups, Lie Algebras, and Representations* [1]. Given a matrix Lie Group G, its Lie algebra, denoted $\mathfrak{g}$, is the set of matrices $\mathbf{X}$ such that for all $t \in \mathbb{R}$: $e^{t\mathbf{X}} \in G$. The Lie algebra is of special interest in group theory because it is a vector space under the bracket operation, enabling us to study Lie groups via their Lie algebras using the tools of linear algebra [1]. While we won't explore Lie algebras deeply, we *are* concerned with the image of the exponential on the lie algebra of SO(3), which we denote $\mathfrak{so}(3)$. As shown in section 5.4 of [7], the lie algebra of $\mathfrak{so}(3)$ is the set of skew-symmetric real matrices.

Proposition 8. *The map* $\exp : \mathfrak{so}(3) \to SO(3)$ *is surjective.*

Proof: In the standard basis, a rotation through the x-axis by angle θ can be represented in the form

$$\mathbf{R}_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix}. \tag{5.6}$$

By Euler's Rotation Theorem (Euler 1775, [8]), $\mathbf{R}$ is equivalent to a rotation about an axis spanned by a unit vector $\mathbf{v} \in \mathbb{R}^3$. If we consider any set of linearly independent vectors where $\mathbf{v}$ is our first vector, then the Graham Schmidt process will generate an orthonormal basis $\{\mathbf{v},\mathbf{u},\mathbf{w}\}$. In this basis, $\mathbf{v}$ remains fixed under transformation by $\mathbf{R}$, and $\mathbf{u},\mathbf{v}$ span the plane which $\mathbf{R}$ induces a rotation of angle θ

through. Thus, in our new basis, we may represent $\mathbf{R}$ in the form of (5.6). In the language of similarity, $\mathbf{R} = \mathbf{B}\mathbf{R}_x\mathbf{B}^{-1}$ for an invertible change-of-basis matrix $\mathbf{B}$. We would now like to show that $\mathbf{R}_x$ is in the image. Consider the matrix

$$\mathbf{X} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\theta \\ 0 & \theta & 0 \end{pmatrix}.$$

The powers of $\mathbf{X}$ are cyclic, since

$$\mathbf{X}^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\theta^2 & 0 \\ 0 & 0 & -\theta^2 \end{pmatrix}$$
$$\mathbf{X}^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \theta^3 \\ 0 & -\theta^3 & 0 \end{pmatrix}$$
$$\mathbf{X}^4 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \theta^4 & 0 \\ 0 & 0 & \theta^4 \end{pmatrix}$$
$$\mathbf{X}^5 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \theta^5 \\ 0 & \theta^5 & 0 \end{pmatrix} = \theta\mathbf{X}.$$

Using this fact we can show that the explicit formula for $e^{\mathbf{X}}$ is

$$e^{\mathbf{X}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \alpha & \beta \\ 0 & \gamma & \delta \end{pmatrix}. \tag{5.7}$$

Where the 1 in the first entry comes from the identity term $\mathbf{X}^0 = \mathbf{I}$, and the remaining terms come from inspecting the entries of successive powers of $\mathbf{X}$. Computing them yields

$$\begin{aligned}
\alpha &= 1 + \frac{0}{1!} - \frac{\theta^2}{2!} + \frac{0}{3!} + \frac{\theta^4}{4!} + \cdots = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \cdots = \cos(\theta), \\
\beta &= 0 - \frac{\theta}{1!} + \frac{0}{2!} + \frac{\theta^3}{3!} + \frac{0}{4!} + \cdots = -\theta + \frac{\theta^3}{3!} - \frac{\theta^5}{5!} + \frac{\theta^7}{7!} + \cdots = -\sin(\theta), \\
\gamma &= 0 + \frac{\theta}{1!} + \frac{0}{2!} - \frac{\theta^3}{3!} + \frac{0}{4!} + \cdots = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!} + \cdots = \sin(\theta), \\
\delta &= 1 + \frac{0}{1!} - \frac{\theta^2}{2!} + \frac{0}{3!} + \frac{\theta^4}{4!} + \cdots = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \cdots = \cos(\theta).
\end{aligned} \tag{5.8}$$

Which shows that $e^{\mathbf{X}} = \mathbf{R}_x$. Thus,

$$e^{(\mathbf{B}\mathbf{X}\mathbf{B}^{-1})} = \mathbf{B}e^{\mathbf{X}}\mathbf{B}^{-1} = \mathbf{B}\mathbf{R}_x\mathbf{B} = \mathbf{R}. \tag{5.9}$$

If we can show that $\mathbf{B}\mathbf{X}\mathbf{B}^{-1}$ is skew-symmetric, then we are done. Recall that $\mathbf{B}$ is the change of basis matrix from the standard basis to the orthonormal basis $\{\mathbf{v}, \mathbf{u}, \mathbf{w}\}$. These are two orthonormal bases, so $\mathbf{B}$ is orthogonal (see §14.3 of [9]). So $\mathbf{B}^\intercal = \mathbf{B}^{-1}$ and

$$\begin{aligned}
\mathbf{B}\mathbf{X}\mathbf{B}^{-1} &= \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\theta \\ 0 & \theta & 0 \end{pmatrix} \begin{pmatrix} a & d & g \\ b & e & h \\ c & f & i \end{pmatrix} \\
&= \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ -\theta c & -\theta f & -\theta i \\ \theta b & \theta e & \theta h \end{pmatrix} \\
&= \begin{pmatrix} \theta bc + \theta bc & -\theta bf + \theta ec & -\theta bi + \theta ch \\ -\theta ec + \theta bf & -\theta ef + \theta ef & -\theta ei + \theta fh \\ -\theta ch + \theta bi & -\theta fh + \theta ei & -\theta hi + \theta hi \end{pmatrix} \\
&= \begin{pmatrix} 0 & -\theta bf + \theta ec & -\theta bi + \theta ch \\ -(-\theta bf + \theta ec) & 0 & -\theta ei + \theta fh \\ -(-\theta bc + \theta ch) & -(\theta ei + \theta fh)ei & 0 \end{pmatrix}.
\end{aligned} \tag{5.10}$$

Is skew-symmetric, concluding the proof that $\exp : \mathfrak{so}(3) \to \mathrm{SO}(3)$ is surjective. $\square$

References

[1] Brian C. Hall, *Lie Groups, Lie Algebras, and Representations*, Springer-Verlag New York Inc, 2010, https://doi.org/10.1007/978-0-387-21554-9.

[2] Amol Sasane, *A Friendly Approach to Functional Analysis*, World Scientific Publishing Europe Ltd, 2017.

[3] Robert E. Megginson, *An Introduction to Banach Space Theory*, Springer-Verlag New York Inc, 1998.

[4] Theodore W. Gamelin, *Complex Analysis*, Undergraduate Texts in Mathematics, 2001.

[5] Walter Rudin, *Principles of Mathematical Analysis*, MCGraw-Hill, 1976.

[6] Stephen H. Friedberg, *Linear Algebra*, Pearson Education, Inc, 2019.

[7] Kristopher Tapp, *Matrix Groups for Undergraduates*, American Mathematical Society, 2005.

[8] Leonard Euler, General formulas for any translation of rigid bodies, https://www.17centurymaths.com/contents/euler/e478tr.pdf, 1775. Accessed: 2025-03-11

[9] David Cherney, *Linear Algebra*, UCDavis, 2013.