Capítulo 11
Tópicos de Álgebra Linear. II

Sumário

O presente capítulo diferencia-se do anterior por explorar aspectos mais topológicos de álgebras de matrizes. Portanto, uma certa familiaridade com as noções básicas de espaços métricos (vide Capítulo 25, página 1410) é útil. Discutiremos a definição de funções analíticas de matrizes, em particular, a exponencial e o logaritmo. Nosso principal objetivo, porém, é provar as seguintes relações: para matrizes $A, B \in \mathrm{Mat}(\mathbb{C}, n)$, valem:

- *Fórmula de Lie-Trotter*[1]:
$$\exp(A+B) \;=\; \lim_{m\to\infty}\left[\exp\left(\frac{1}{m}A\right)\exp\left(\frac{1}{m}B\right)\right]^m . \tag{11.1}$$

- *Fórmula do comutador*:
$$\exp\big([A,\, B]\big) \;=\; \lim_{m\to\infty}\left[\exp\left(\frac{1}{m}A\right)\exp\left(\frac{1}{m}B\right)\exp\left(-\frac{1}{m}A\right)\exp\left(-\frac{1}{m}B\right)\right]^{m^2} . \tag{11.2}$$

- *Série de Lie*:
$$\exp(B)A\exp(-B) \;=\; A + \sum_{m=1}^{\infty}\frac{1}{m!}\Big[\underbrace{B,\, \big[B,\, \ldots,\, [B}_{m\ \text{vezes}},\, A]\cdots\big]\Big] . \tag{11.3}$$

- *Fórmula de Baker-Campbell-Hausdorff*[2] (sobre a convergência, vide comentário adiante):
$$\exp(A)\exp(B) \;=\; \exp\left(A+B+\frac{1}{2}[A,\, B]+\frac{1}{12}\big[A,\, [A,\, B]\big]+\frac{1}{12}\big[B,\, [B,\, A]\big]+\cdots\right) . \tag{11.4}$$

- *Fórmula de Duhamel*[3]:
$$\exp(A+B) \;=\; \exp(A)+\int_0^1 \exp\big((1-s)(A+B)\big)\, B\, \exp\big(sA\big)\, ds\, , \tag{11.5}$$

  da qual se obtem a *série de Duhamel*:
$$e^{t(A+B)} \;=\; e^{tA}\left[\mathbb{1}+\int_0^t e^{-t_1A}Be^{t_1A}dt_1+\sum_{m=2}^{\infty}\int_0^t\int_0^{t_1}\!\cdots\int_0^{t_{m-1}}\prod_{k=1}^{m}\big(e^{-t_kA}Be^{t_kA}\big)\, dt_m\cdots dt_1\right] . \tag{11.6}$$

[1] Marius Sophus Lie (1842–1899). Hale Freeman Trotter (1931–).
[2] Henry Frederick Baker (1866–1956). John Edward Campbell (1862–1924). Felix Hausdorff (1868–1942).
[3] Jean Marie Constant Duhamel (1797–1872).

Uma outra relação útil que obteremos é a chamada *Fórmula de Duhamel para derivadas de exponenciais*: se $A(\lambda)$ for uma família de matrizes em Mat (ℂ, n) que depende contínua e diferenciavelmente de um parâmetro λ, então vale
$$\frac{d}{d\lambda}\left(e^{A(\lambda)}\right) \;=\; \int_0^1 e^{(1-s)A(\lambda)}\left(\frac{d}{d\lambda}A(\lambda)\right)e^{sA(\lambda)}\, ds\, . \tag{11.7}$$

A série dentro da exponencial no lado direito de (11.4) é um tanto complexa, mas envolve apenas comutadores múltiplos de ordem cada vez maior de A e B. A expressão completa encontra-se em (11.68), página 682. Ao contrário das fórmulas que lhe precedem e sucedem, a fórmula de Baker-Campbell-Hausdorff não é válida para quaisquer matrizes A e B pois, no caso geral, a convergência da série do lado direito só pode ser estabelecida para matrizes suficientemente "pequenas", a saber, tais que $\|A\|_{\mathbb{C}}$ e $\|B\|_{\mathbb{C}}$ sejam ambas menores que $\frac{1}{2}\ln\left(2-\frac{\sqrt{2}}{2}\right) \approx 0,12844\ldots$ (a definição da norma operatorial $\|\cdot\|_{\mathbb{C}}$ de matrizes será apresentada adiante). Claro é que, nos casos felizes em que os comutadores múltiplos das matrizes A e B se anulam a partir de uma certa ordem, a série do lado direito será finita e, portanto, convergente.

Comentamos ao leitor mais avançado que as expressões acima (e, *mutatis mutandis*, suas demonstrações abaixo) valem não apenas para álgebras de matrizes, mas também no contexto mais geral de álgebras-$*$ de Banach com unidade.

As fórmulas acima são empregadas em várias áreas da Física (como na Mecânica Quântica, na Mecânica Estatística e na Teoria Quântica de Campos) e da Matemática (como na Teoria de Grupos). Faremos uso delas, por exemplo, nos Capítulos 21 e 22. Suas provas serão apresentadas, pela ordem, na Proposição 11.13, página 669, na Proposição 11.15, página 678, no Teorema 11.1 da Seção 11.5, página 682 e na Seção 11.6, página 687. A única demonstração que se pode classificar como complexa é a da fórmula de Baker-Campbell-Hausdorff, as demais são simples. No correr das páginas seguintes outras identidades úteis, não listadas acima, serão obtidas.

11.1 Uma Topologia Métrica em Mat (ℂ, n)

Discutiremos nesta seção uma topologia métrica natural em Mat (ℂ, n) a qual usaremos na Seção 11.2 para definir certas funções analíticas de matrizes, tais como a exponencial e o logaritmo.

Recordando, Mat (ℂ, n) é o conjunto de todas as matrizes complexas $n\times n$ e $\mathrm{GL}(\mathbb{C},\, n) \subset \mathrm{Mat}(\mathbb{C},\, n)$ é o conjunto de todas as matrizes complexas $n \times n$ *inversíveis*. Como já observamos, GL(ℂ, n) é um grupo.

• Normas de matrizes. A norma operatorial

Seja V um espaço vetorial de dimensão finita, como $\mathbb{C}^n$ ou $\mathbb{R}^n$, dotado de uma norma $\|\cdot\|_V$. Para $\mathbb{C}^n \ni u = (u_1,\, \ldots, u_n)$, por exemplo, podemos adotar $\|u\|_{\mathbb{C}^n} := \sqrt{|u_1|^2+\cdots+|u_n|^2}$. Vamos denotar por $L(V)$ o conjunto de todas as aplicações lineares de V em V. É bem sabido que $L(V)$ é igualmente um espaço vetorial. Por exemplo, $L(\mathbb{C}^n) = \mathrm{Mat}(\mathbb{C},\, n)$ e $L(\mathbb{R}^n) = \mathrm{Mat}(\mathbb{R},\, n)$.

Com uso da norma de V é possível definir uma norma também em $L(V)$. Para $A \in L(V)$ define-se
$$\|A\|_{L(V)} \;:=\; \sup_{\substack{u\in V\\ u\neq 0}}\frac{\|Au\|_V}{\|u\|_V} .$$

E. 11.1 *Exercício*. Mostre que $\|\cdot\|_{L(V)}$ assim definida é, de fato, uma norma no espaço vetorial $L(V)$. ♣

Observações. Note que
$$\|A\|_{L(V)} \;=\; \sup_{\substack{u\in V\\ \|u\|_V=1}}\|Au\|_V .$$
Para $A \in L(V)$, a norma $\|A\|_{L(V)}$ definida acima é denominada *norma operatorial* induzida pela norma $\|\cdot\|_V$. Como comentaremos abaixo, há outras normas em $L(\mathbb{C}^n)$ e $L(\mathbb{R}^n)$ que não a norma operatorial, mas que são equivalentes àquela. É uma consequência imediata da definição de norma operatorial que
$$\|Au\|_V \;\leq\; \|A\|_{L(V)}\, \|u\|_V\, , \tag{11.8}$$
para todo vetor $u \in V$. ♣

A norma operatorial tem a seguinte propriedade importante: para $A,\, B \in L(V)$ quaisquer, tem-se
$$\|AB\|_{L(V)} \;\leq\; \|A\|_{L(V)}\, \|B\|_{L(V)} .$$

Essa propriedade é denominada *submultiplicatividade* da norma $\|\cdot\|_{L(V)}$. Nem toda norma em $\text{Mat}(\mathbb{C},\ n)$ possui essa propriedade.

E. 11.2 *Exercício importante.* Mostre isso. Sugestão: use (11.8). ✠

Observação. Em $\text{Mat}(\mathbb{C},\ n)$ é possível provar que $\|A^*\|_{\text{Mat}(\mathbb{C},\ n)} = \|A\|_{\text{Mat}(\mathbb{C},\ n)}$ e que $\|A\|^2_{\text{Mat}(\mathbb{C},\ n)} = \|A^*A\|_{\text{Mat}(\mathbb{C},\ n)}$ (propriedade C*). Vide Teorema 42.11, página 2363. ♣

É importante comentar que o procedimento de construção de normas em $L(V)$ pode ser repetido. Como $L(V)$ é igualmente um espaço vetorial normado e de dimensão finita, podemos definir uma norma em $L(L(V))$ (o conjunto de todas as aplicações lineares de $L(V)$ em $L(V)$) definindo para $\mathcal{A} \in L(L(V))$

$$\|\mathcal{A}\|_{L(L(V))} := \sup_{\substack{A\in L(V)\\ A\neq 0}} \frac{\|\mathcal{A}A\|_{L(V)}}{\|A\|_{L(V)}} .$$

E assim por diante para todos os espaços de aplicações $L(L(\cdots L(V))\cdots)$.

Vamos a um exemplo. Tomemos $V = \mathbb{C}^n$, $L(V) = \text{Mat}(\mathbb{C},\ n)$. Seja uma matriz $X \in \text{Mat}(\mathbb{C},\ n)$ fixa. Com ela poderemos definir um elemento denotado por $\mathbf{ad}[X]$ de $L(\text{Mat}(\mathbb{C},\ n))$ por

$$\mathbf{ad}[X]A := [X,\ A] = XA - AX, \qquad A \in \text{Mat}(\mathbb{C},\ n) .$$

É evidente que $\mathbf{ad}[X]$ é uma aplicação linear de $\text{Mat}(\mathbb{C},\ n)$ em $\text{Mat}(\mathbb{C},\ n)$, ou seja, um elemento de $L(\text{Mat}(\mathbb{C},\ n))$. Note-se que

$$\|\mathbf{ad}[X]\|_{L(\text{Mat}(\mathbb{C},\ n))} = \sup_{\substack{A\in L(V)\\ A\neq 0}} \frac{\|XA-AX\|_{\text{Mat}(\mathbb{C},\ n)}}{\|A\|_{\text{Mat}(\mathbb{C},\ n)}} \leq \sup_{\substack{A\in L(V)\\ A\neq 0}} \frac{\|XA\|_{\text{Mat}(\mathbb{C},\ n)}+\|AX\|_{\text{Mat}(\mathbb{C},\ n)}}{\|A\|_{\text{Mat}(\mathbb{C},\ n)}} \leq 2\|X\|_{\text{Mat}(\mathbb{C},\ n)} .$$

Daqui para a frente denotaremos a norma operatorial de matrizes em $\mathbb{C}^n$ por $\|\cdot\|_{\mathbb{C}}$ ou simplesmente por $\|\cdot\|$. Além da norma operatorial, há outras normas que podem ser definidas em $L(\mathbb{C}^n)$. Para $A \in \text{Mat}(\mathbb{C},\ n)$ podemos, por exemplo, definir as seguintes normas:

$$\|A\|_\infty := \max_{a,\ b=1,\ldots,n} |A_{ab}|, \tag{11.9}$$

$$\|A\|_1 := \sum_{a=1}^n\sum_{b=1}^n |A_{ab}| , \tag{11.10}$$

$$\|A\|_2 := \left(\sum_{a=1}^n\sum_{b=1}^n |A_{ab}|^2\right)^{1/2}, \tag{11.11}$$

$$\|A\|_p := \left(\sum_{a=1}^n\sum_{b=1}^n |A_{ab}|^p\right)^{1/p}, \qquad \text{com } p\geq 1 . \tag{11.12}$$

A expressão (11.12) generaliza (11.10) e (11.11). A norma $\|A\|_2$ é por vezes denominada a *norma de Frobenius*[4] da matriz A.

E. 11.3 *Exercício.* Mostre que (11.9)-(11.12) de fato definem normas em $\text{Mat}(\mathbb{C},\ n)$. (Note que (11.10)-(11.11) são casos particulares de (11.12)). Use a desigualdade de Minkowski (página 1449) para (11.12). ✠

E. 11.4 *Exercício.* A norma de Frobenius (11.11) tem uma interpretação interessante. Mostre que,

$$\langle A,\ B\rangle = \text{Tr}(A^*B) = \sum_{a=1}^n\sum_{b=1}^n \overline{A_{ab}}B_{ab} , \tag{11.13}$$

[4]Ferdinand Georg Frobenius (1849–1917).

$A,\ B \in \text{Mat}(\mathbb{C},\ n)$, define um produto escalar em $\text{Mat}(\mathbb{C},\ n)$. Mostre que (11.11) é a norma associada a esse produto escalar, ou seja, $\|A\|_2 = \sqrt{\langle A,\ A\rangle} = \sqrt{\text{Tr}(A^*A)}$. ✠

Observação. É importante lembrar o Teorema 3.2, página 278, que afirma que em espaços vetoriais de dimensão finita todas as normas são equivalentes. Assim, em $\text{Mat}(\mathbb{C},\ n)$ a norma operatorial $\|A\|_{\mathbb{C}}$ e as normas $\|A\|_\infty$ e $\|A\|_p$ com $p\geq 1$ são todas equivalentes. Note-se, porém, que a propriedade de submultiplicatividade $\|AB\|_{\mathbb{C}} \leq \|A\|_{\mathbb{C}}\|B\|_{\mathbb{C}}$ da norma operatorial não é necessariamente compartilhada por outras normas. Devido à equivalência de todas as normas matriciais, tem-se em geral $\|AB\| \leq c\|A\|\,\|B\|$ para alguma constante $c > 0$. ♣

E. 11.5 *Exercício.* Seja $D \in \text{Mat}(\mathbb{C},\ n)$ uma matriz diagonal: $D = \text{diag}(d_1,\ \ldots,\ d_n)$ com $d_k \in \mathbb{C}$. Mostre que $\|D\|_{\mathbb{C}} = \max\{|d_1|,\ \ldots,\ |d_n|\}$, ou seja, para matrizes diagonais $\|D\|_{\mathbb{C}} = \|D\|_\infty$. ✠

• Equivalência entre normas matriciais

Aqui denotaremos a norma operatorial de uma matriz A por $\|A\|$.

Sejam $\mathbf{e}_i$, $i = 1,\ \ldots,\ n$ os vetores da base canônica de $\mathbb{C}^n$, ou seja, os vetores cuja j-ésima componente é $(\mathbf{e}_i)_j = \delta_{ij}$. Se $A \in \text{Mat}(\mathbb{C},\ n)$, é claro que a i-ésima componente do vetor $A\mathbf{e}_j$ é $(A\mathbf{e}_j)_i = A_{ij}$. Daí,

$$\frac{\|A\mathbf{e}_j\|^2_{\mathbb{C}}}{\|\mathbf{e}_j\|^2_{\mathbb{C}}} = \sum_{i=1}^n |A_{ij}|^2 .$$

Logo, para todo j,

$$\|A\|^2 := \sup_{\substack{v\in\mathbb{C}^n\\ v\neq 0}} \frac{\|Av\|^2_{\mathbb{C}}}{\|v\|^2_{\mathbb{C}}} \geq \max_{j=1,\ldots,n} \frac{\|A\mathbf{e}_j\|^2_{\mathbb{C}}}{\|\mathbf{e}_j\|^2_{\mathbb{C}}} = \max_{j=1,\ldots,n}\left\{\sum_{i=1}^n |A_{ij}|^2\right\} . \tag{11.14}$$

Tem-se também o seguinte. Para qualquer vetor $v \in \mathbb{C}^n$, vale $(Av)_i = \sum_{j=1}^n A_{ij}v_j$. Assim, pela desigualdade de Cauchy-Schwarz (3.20), página 276,

$$|(Av)_i|^2 \leq \left(\sum_{j=1}^n |A_{ij}|^2\right)\left(\sum_{k=1}^n |v_k|^2\right) = \left(\sum_{j=1}^n |A_{ij}|^2\right)\|v\|^2_{\mathbb{C}} .$$

Daí,

$$\|Av\|^2_{\mathbb{C}} = \sum_{i=1}^n |(Av)_i|^2 \leq \left(\sum_{i=1}^n\sum_{j=1}^n |A_{ij}|^2\right)\|v\|^2_{\mathbb{C}} .$$

Logo,

$$\|A\|^2 := \sup_{\substack{v\in\mathbb{C}^n\\ v\neq 0}} \frac{\|Av\|^2_{\mathbb{C}}}{\|v\|^2_{\mathbb{C}}} \leq \sum_{i=1}^n\sum_{j=1}^n |A_{ij}|^2 . \tag{11.15}$$

Como $\sum_{i=1}^n |A_{ij}|^2 \geq \max_{i=1,\ldots,n}\left\{|A_{ij}|^2\right\}$, segue de (11.14) que

$$\|A\|^2 \geq \max_{j=1,\ldots,n}\ \max_{i=1,\ldots,n} |A_{ij}|^2 .$$

Logo, para todo $i,\ j$ vale $|A_{ij}| \leq \|A\|$, ou seja,

$$\|A\|_\infty \leq \|A\| .$$

De (11.15) vemos também que

$$\|A\|^2 \leq \sum_{i=1}^n\sum_{j=1}^n |A_{ij}|^2 \leq \sum_{i=1}^n\sum_{j=1}^n \|A\|^2_\infty = n^2\|A\|^2_\infty .$$

Concluímos assim que em $\text{Mat}(\mathbb{C},\ n)$

$$\|A\|_\infty \leq \|A\| \leq n\|A\|_\infty . \tag{11.16}$$

A expressão (11.16) mostra-nos que caso tenhamos uma sequência de matrizes A_m com $\|A_m\| \to 0$ quando $m \to \infty$, então cada elemento de matriz $(A_m)_{ij}$ também converge a zero quando $m \to \infty$. E vice-versa: Se $(A_m)_{ij} \to 0$ para todos ij quando $m \to \infty$, então $\|A_m\| \to 0$ quando $m \to \infty$.

Nota. Antes de prosseguirmos, comentemos também que as duas desigualdades (11.16) são optimais, ou seja, não podem ser melhoradas para matrizes genéricas. Por exemplo, é evidente que $\|\mathbb{1}\|_\infty = 1$ e que $\|\mathbb{1}\| = 1$. Assim, pelo menos nesse caso tem-se a igualdade na primeira desigualdade de (11.16). Há também um caso em que se tem a igualdade na segunda desigualdade de (11.16). Considere-se a matriz M cujos elementos de matriz são todos iguais a 1, ou seja, $M_{ij} = 1$ para todos i, j. Seja o vetor u de $\mathbb{C}^n$ cujas componentes são todas iguais a 1, ou seja, $u_i = 1$ para todo i. É elementar ver que $Mu = nu$. Logo $\frac{\|Mu\|_{\mathbb{C}}}{\|u\|_{\mathbb{C}}} = n$. Portanto, $\|M\| \geq n$ e $\|M\|_\infty = 1$. Assim, $\|M\| \geq n\|M\|_\infty$ e, da segunda desigualdade de (11.16), concluímos que, nesse caso, $\|M\| = n\|M\|_\infty$. ♣

A desigualdade (11.15) significa que $\|A\| \leq \|A\|_2$. Ao mesmo tempo, a desigualdade (11.14) mostra que

$$n\|A\|^2 \;=\; \sum_{j=1}^n \|A\|^2 \;\geq\; \sum_{j=1}^n \sum_{i=1}^n |A_{ij}|^2 \;=\; \|A\|_2^2 \, .$$

Logo, concluímos que em $\mathrm{Mat}(\mathbb{C}, n)$

$$\frac{1}{\sqrt{n}}\|A\|_2 \;\leq\; \|A\| \;\leq\; \|A\|_2 \, . \tag{11.17}$$

E. 11.6 *Exercício*. Mostre que em $\mathrm{Mat}(\mathbb{C}, n)$

$$\frac{1}{n^2}\|A\|_1 \;\leq\; \|A\| \;\leq\; n\|A\|_1 \, . \tag{11.18}$$

Sugestão: Mostre primeiro que $\|A\|_\infty \leq \sum_{i,\, j=1}^n |A_{ij}| \leq n^2\|A\|_\infty$ ou seja

$$\|A\|_\infty \;\leq\; \|A\|_1 \;\leq\; n^2\|A\|_\infty \, . \tag{11.19}$$

e, então, use (11.16). ✣

E. 11.7 *Exercício*. Mostre que as desigualdades (11.19) também não podem ser melhoradas. ✣

Nota. As expressões (11.16), (11.17), (11.18) e (11.19) mostram-nos de modo explícito que em $\mathrm{Mat}(\mathbb{C}, n)$ as normas $\|\cdot\|$, $\|\cdot\|_\infty$, $\|\cdot\|_1$ e $\|\cdot\|_2$ são equivalentes (vide definição à página 278). Como já mencionamos, em espaços de dimensão finita todas as normas matriciais são equivalentes (Teorema 3.2, página 278). ♣

*

A importância de se introduzir uma norma em $L(V)$ é que podemos dessa forma introduzir uma noção de distância entre elementos desse conjunto, ou seja, podemos definir uma métrica em $L(V)$ por $d(A, B) = \|A - B\|$. Deixamos para o leitor a tarefa de demonstrar que isso de fato define uma métrica em $L(V)$. Com isso, fazemos de $L(V)$ um espaço dotado de uma topologia métrica. Fora isso, o importante Teorema 42.2 demonstrado à página 2343 afirma que $L(V)$ será um espaço métrico completo se V o for. Logo, como $\mathbb{C}^n$ e $\mathbb{R}^n$ são sabidamente espaços vetoriais completos, assim o serão $\mathrm{Mat}(\mathbb{C}, n)$, $\mathrm{Mat}(\mathbb{R}, n)$, assim como $L(\mathrm{Mat}(\mathbb{C}, n))$ etc. É possível dessa forma falar de convergência de sequências e séries de matrizes de $\mathrm{Mat}(\mathbb{C}, n)$, $\mathrm{Mat}(\mathbb{R}, n)$, assim como de elementos de $L(\mathrm{Mat}(\mathbb{C}, n))$ etc. Abaixo faremos uso repetido desse fato fundamental.

11.2 Exponenciais, Logaritmos e Funções Analíticas de Matrizes

No estudo da teoria de grupos e em outras áreas é muito conveniente definir certas funções de operadores lineares, tais como exponenciais, logaritmos etc. Já abordamos a definição da exponenciação de matrizes nos capítulos 10 e 14. Vamos aqui tentar uma abordagem mais geral.

• **Séries de potências de matrizes**

Seja $A \in \mathrm{Mat}(\mathbb{C}, n)$ uma matriz $n \times n$ complexa e seja $\{a_m \; m \in \mathbb{N}\}$ uma sequência de números complexos. A expressão

$$\sum_{m=0}^\infty a_m A^m \;=\; \lim_{N\to\infty} \sum_{m=0}^N a_m A^m \;=\; a_0\mathbb{1} + a_1 A + a_2 A^2 + a_3 A^3 + \cdots$$

é dita ser uma série de potências convergente, caso o limite acima exista em $\mathrm{Mat}(\mathbb{C}, n)$.

Nota. Adotaremos sempre a convenção que $A^0 = \mathbb{1}$. ♣

A seguinte proposição é fundamental:

Proposição 11.1 *A séria de potências* $\sum_{m=0}^\infty a_m A^m$ *é convergente se* $\sum_{m=0}^\infty |a_m| \, \|A\|_{\mathbb{C}}^m < \infty$. □

A importância dessa proposição reside no fato que $\sum_{m=0}^\infty |a_m| \|A\|_{\mathbb{C}}^m$ é uma série numérica e, portanto, mais simples de lidar.

Prova. Sejam as somas parciais $S_N := \sum_{m=0}^N a_m A^m$. Teremos para $M < N$,

$$\|S_N - S_M\|_{\mathbb{C}} \;=\; \left\| \sum_{m=M+1}^N a_m A^m \right\|_{\mathbb{C}} \;\leq\; \sum_{m=M+1}^N |a_m| \, \|A\|_{\mathbb{C}}^m \, .$$

Agora, como a série numérica $\sum_{m=0}^\infty |a_m| \, \|A\|_{\mathbb{C}}^m$ converge, $s_N := \sum_{m=0}^N |a_m| \, \|A\|_{\mathbb{C}}^m$ é uma sequência de Cauchy. Logo $\sum_{m=M+1}^N |a_m| \, \|A\|_{\mathbb{C}}^m$ pode ser feito menor que qualquer $\epsilon > 0$ dado, desde que escolhamos M e N grandes o suficiente. Logo S_N é também uma sequência de Cauchy no espaço métrico completo $\mathrm{Mat}(\mathbb{C}, n)$. Portanto, S_N converge em $\mathrm{Mat}(\mathbb{C}, n)$ quando $N \to \infty$. ■

• **Funções analíticas de matrizes**

A Proposição 11.1 conduz à seguinte definição. Seja $r > 0$ e $D_r = \{z \in \mathbb{C} | \; |z| < r\}$ o disco aberto de raio r centrado em 0 no plano complexo. Seja $f : D_r \to \mathbb{C}$ uma função analítica em D_r. Como bem sabemos, f pode ser expressa em termos de uma série de potências (série de Taylor centrada em $z_0 = 0$): $f(z) = \sum_{m=0}^\infty f_m z^m$, onde $f_m = f^{(m)}(0)/m!$. É bem sabido também que essa série é absolutamente convergente em D_r: $\sum_{m=0}^\infty |f_m| \, |z|^m < \infty$, se $|z| < r$. Podemos então definir

$$f(A) \;:=\; \sum_{m=0}^\infty f_m A^m$$

para toda a matriz A com $\|A\|_{\mathbb{C}} < r$, pois a proposição acima garante que a série de matrizes do lado direito converge a alguma matriz de $\mathrm{Mat}(\mathbb{C}, n)$, que denotamos por $f(A)$, fazendo uma analogia óbvia com a função numérica f.

A seguinte proposição sobre essas funções de matrizes será frequentemente usada no que seguirá.

Proposição 11.2 I. *Sejam f e g duas funções analíticas no mesmo domínio D_r. Definamos $(f+g)(z) := f(z) + g(z)$ e $(fg)(z) := f(z)g(z)$, $z \in D_r$. Então, para $A \in \mathrm{Mat}(\mathbb{C}, n)$ com $\|A\|_{\mathbb{C}} < r$ teremos $f(A) + g(A) = (f+g)(A)$ e $f(A)g(A) = g(A)f(A) = (fg)(A)$.*

II. *Sejam f e g duas funções analíticas, com domínios D_{r_f} e D_{r_g}, respectivamente, e tais que a imagem de g esteja contida no domínio de f. Podemos então definir $f \circ g(z) := f(g(z))$. Então, para $A \in \mathrm{Mat}(\mathbb{C}, n)$ com $\|A\|_{\mathbb{C}} < r_g$ teremos $f(g(A)) = f \circ g(A)$.* □

Prova. $\longleftrightarrow$ *Exercício*. ■

Note-se que a parte **I** da proposição acima afirma que existe um homomorfismo da álgebra das funções analíticas em um domínio $D_r \subset \mathbb{C}$ e $\mathrm{Mat}(\mathbb{C}, n)$.

Vamos mais adiante usar o seguinte resultado, que essencialmente afirma que as matrizes $f(A)$ definidas acima, com f analítica em um domínio $D_r \subset \mathbb{C}$, dependem continuamente de A.

Proposição 11.3 *Seja f função complexa analítica em um domínio $D_r \subset \mathbb{C}$, com f tendo a série de Taylor absolutamente convergente $f(z) = \sum_{k=0}^{\infty} f_k z^k$, $|z| < r$. Seja também B_m, $m \in \mathbb{N}$, uma sequência de matrizes de $\mathrm{Mat}(\mathbb{C},\ n)$ tais que $\lim_{m\to\infty} \|B_m\|_{\mathbb{C}} = 0$. Então, para todo $A \in \mathrm{Mat}(\mathbb{C},\ n)$ com $\|A\|_{\mathbb{C}} < r$ tem-se*

$$\lim_{m\to\infty} f(A + B_m) \;=\; f(A)\ .$$

□

Prova. Comecemos com um comentário sobre o enunciado do teorema. Para que $f(A + B_m)$ esteja definido é necessário que $\|A + B_m\|_C < r$. Como $\|A + B_m\|_C \leq \|A\|_{\mathbb{C}} + \|B_m\|_{\mathbb{C}}$ e $\|A\|_{\mathbb{C}} < r$, a condição é satisfeita para m grande o suficiente, pois $\lim_{m\to\infty} \|B_m\|_{\mathbb{C}} = 0$. Assim, estaremos supondo que m é grande o suficiente de modo que $\|B_m\|_{\mathbb{C}} < \epsilon$ para algum ϵ tal que $\|A\|_{\mathbb{C}} + \epsilon < r$. Feita essa ressalva, passemos à demonstração.

A prova da proposição segue das duas observações seguintes. A primeira é que para quaisquer matrizes $X,\ Y \in \mathrm{Mat}(\mathbb{C},\ n)$ e qualquer k inteiro positivo tem-se a seguinte identidade algébrica, denominada *soma telescópica* ou *identidade telescópica*:

$$X^k - Y^k \;=\; \sum_{p=0}^{k-1} X^p\,(X - Y)\,Y^{k-1-p}\ . \tag{11.20}$$

Para provar isso, basta expandir a soma do lado direito e mostrar, após alguns cancelamentos, que obtem-se o lado esquerdo (faça!).

A segunda observação é que se f é analítica em D_r, sua derivada também o é. Assim, $f'(z) = \sum_{k=0}^{\infty} k f_k\, z^{k-1}$ converge absolutamente para $|z| < r$, ou seja, $\sum_{k=0}^{\infty} k|f_k|\,|z|^{k-1} < \infty$ sempre que $|z| < r$.

Assim,

$$f(A + B_m) - f(A) \;=\; \sum_{k=0}^{\infty} f_k\left[(A + B_m)^k - A^k\right]\ .$$

Usando (11.20) com $X = A + B_m$ e $Y = A$, teremos

$$f(A + B_m) - f(A) \;=\; \sum_{k=0}^{\infty} f_k \sum_{p=0}^{k-1} (A + B_m)^p\, B_m\, A^{k-1-p}.$$

Logo,

$$\|f(A + B_m) - f(A)\|_{\mathbb{C}} \;\leq\; \|B_m\|_{\mathbb{C}} \sum_{k=0}^{\infty} |f_k| \sum_{p=0}^{k-1} \|A + B_m\|_{\mathbb{C}}^p\, \|A\|_{\mathbb{C}}^{k-1-p}\ .$$

Agora, como dissemos, $\|A + B_m\|_{\mathbb{C}} < \|A\|_{\mathbb{C}} + \epsilon < r$ e, obviamente, $\|A\|_{\mathbb{C}} < \|A\|_{\mathbb{C}} + \epsilon < r$. Portanto,

$$\|f(A + B_m) - f(A)\|_{\mathbb{C}} \;\leq\; \|B_m\|_{\mathbb{C}} \sum_{k=0}^{\infty} |f_k| \sum_{p=0}^{k-1} (\|A\|_{\mathbb{C}} + \epsilon)^{k-1} \;=\; \|B_m\|_{\mathbb{C}} \sum_{k=0}^{\infty} k|f_k|\,(\|A\|_{\mathbb{C}} + \epsilon)^{k-1}.$$

Como comentamos acima, a soma do lado direito é finita. Como, porém, $\|B_m\|_{\mathbb{C}} \to 0$ para $m \to \infty$, teremos $\lim_{m\to\infty} \|f(A + B_m) - f(A)\|_{\mathbb{C}} = 0$, que é o que queríamos provar. ■

• Exponenciais e logaritmos de matrizes

Com as definições apresentadas acima, podemos definir exponenciais e logaritmos de matrizes. Temos,

$$\exp(A) \;\equiv\; e^A \;:=\; \sum_{m=0}^{\infty} \frac{1}{m!} A^m \tag{11.21}$$

para toda matriz $A \in \mathrm{Mat}(\mathbb{C},\ n)$, pois a série de Taylor da função exponencial converge absolutamente em todo o plano complexo.

Analogamente, podemos definir

$$\ln(\mathbb{1} + A) \;=\; \sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{m} A^m \tag{11.22}$$

para toda matriz $A \in \mathrm{Mat}(\mathbb{C},\ n)$ com $\|A\|_{\mathbb{C}} < 1$, pois a série de Taylor da função $\ln(1 + z)$ converge absolutamente em D_1.

Nota. Para $\|A - \mathbb{1}\|_{\mathbb{C}} < 1$ podemos definir $\ln(A)$ por $\ln(A) := \ln\big(\mathbb{1} + (A - \mathbb{1})\big)$. ♣

E. 11.8 *Exercício*. Usando a Proposição 11.2, mostre que $(\exp(A))^m = \exp(mA)$ para toda matriz $A \in \mathrm{Mat}(\mathbb{C},\ n)$ e todo $m \in \mathbb{Z}$. Mostre também que

$$\exp\big(\ln(\mathbb{1} + A)\big) \;=\; \mathbb{1} + A$$

para toda matriz $A \in \mathrm{Mat}(\mathbb{C},\ n)$ com $\|A\|_{\mathbb{C}} < 1$ e que

$$\ln\big(\exp(B)\big) \;=\; B$$

para toda matriz $B \in \mathrm{Mat}(\mathbb{C},\ n)$ com $\|\exp(B) - \mathbb{1}\|_{\mathbb{C}} < 1$.

Note que

$$\|\exp(B) - \mathbb{1}\|_{\mathbb{C}} \;=\; \left\|\sum_{m=1}^{\infty} \frac{1}{m!} B^m\right\|_{\mathbb{C}} \;\leq\; \sum_{m=1}^{\infty} \frac{1}{m!} \|B\|_{\mathbb{C}}^m \;=\; e^{\|B\|_{\mathbb{C}}} - 1\ .$$

Assim, a condição $\|\exp(B) - \mathbb{1}\|_{\mathbb{C}} < 1$ é satisfeita se $\|B\|_{\mathbb{C}} < \ln 2$. ♣

Sobre a exponencial de matrizes temos o seguinte:

Proposição 11.4 *Existe uma bola aberta $B_r(0)$ de raio $r > 0$ centrada em 0 em $\mathrm{Mat}(\mathbb{C},\ n)$ tal que a aplicação $\exp : \mathrm{Mat}(\mathbb{C},\ n) \to \mathrm{Mat}(\mathbb{C},\ n)$ definida acima é um homeomorfismo (em verdade, um difeomorfismo) entre $B_r(0)$ e sua imagem, $\exp(B_r(0))$, a qual é uma vizinhança aberta da matriz identidade $\mathbb{1}$.* □

Prova. Temos que, para todo $A \in \mathrm{Mat}(\mathbb{C},\ n)$, $\exp(A) - \mathbb{1} = A + \varphi(A)$, onde $\varphi(A) := \sum_{m=2}^{\infty} \frac{1}{m!} A^m$. É fácil ver que $\frac{\|\varphi(A)\|}{\|A\|} \to 0$ para $\|A\| \to 0$. $\exp(A) - \mathbb{1}$ é contínua e diferenciável em uma vizinhança de 0 (em verdade, em toda parte) e sua derivada em 0 é a identidade. A afirmação da Proposição 11.4 segue então do bem conhecido Teorema da Aplicação Inversa (vide, por exemplo, [340]). ■

Junto com o último exercício, isso prova a seguinte proposição:

Proposição 11.5 *Para toda matriz $A \in \mathrm{Mat}(\mathbb{C},\ n)$ com $\|A - \mathbb{1}\|_{\mathbb{C}} < 1$ tem-se*

$$\exp\big(\ln(A)\big) \;=\; A\ .$$

Para toda matriz $B \in \mathrm{Mat}(\mathbb{C},\ n)$ com $\|B\|_{\mathbb{C}} < \ln 2$ tem-se

$$\ln\big(\exp(B)\big) \;=\; B\ . \tag{11.23}$$

□

• Exponenciais de matrizes diagonalizáveis e o Teorema Espectral

Se $A \in \mathrm{Mat}(\mathbb{C},\ n)$ é diagonalizável, o Teorema Espectral (Teorema 10.7, página 562) e o Cálculo Funcional (Teorema 10.8, página 564) permitem obter expressões simples para a exponencial $\exp(A)$ em termos dos autovalores e dos projetores espectrais de A. De fato, seja $A = \sum_{k=1}^{r} \alpha_k E_k$ a decomposição espectral de A, com $\{\alpha_1,\ \ldots,\ \alpha_r\}$ sendo seus autovalores distintos ($1 \leq r \leq n$) e E_k sendo seus projetores espectrais (dados, por exemplo, como em (10.63), página 565). Pelo Cálculo Funcional, Teorema 10.8, página 564, temos para o polinômio de Taylor $p_n(x) = \sum_{a=0}^{n} \frac{1}{a!} x^a$ que

$$p_n(A) \; = \; \sum_{k=1}^{r} p_n\big(\alpha_k\big) E_k \; .$$

Tomando-se o limite $n \to \infty$, segue facilmente que

$$e^A \; = \; \sum_{k=1}^{r} e^{\alpha_k} E_k \; ,$$

expressão essa de grande utilidade na determinação explícita da exponencial de matrizes diagonalizáveis.

E. 11.9 *Exercício*. Usando também o Teorema Espectral e o Cálculo Funcional, obtenha expressões para o logaritmo de matrizes diagonalizáveis (em situações nas quais ele esteja definido). ✣

• Exponenciais de matrizes. Comutatividade

Para dois números complexos z e w é bem conhecida a validade da propriedade $\exp(z)\exp(w) = \exp(z+w)$ da função exponencial. Podemos nos perguntar: será essa propriedade válida também para matrizes? A resposta é que em geral tal relação **não** é válida, apenas em certos casos especiais. A questão de determinar o produto de exponenciais de matrizes tem grande importância em várias manipulações algébricas e muito do que seguirá abordará esse problema.

Lembremos a primeiramente a seguinte proposição.

Proposição 11.6 *Se $A, B \in \mathrm{Mat}(\mathbb{C}, n)$ são duas matrizes que comutam, ou seja, $AB = BA$, então*

$$e^{A+B} \; = \; e^A e^B \; = \; e^B e^A \; . \tag{11.24}$$

□

A propriedade (11.24) é familiar quando A e B são números, mas não é óbvia quando A e B são matrizes. De fato a relação acima é geralmente *falsa* caso A e B sejam matrizes que não comutam. No caso em que A e B não comutam o produto $e^A e^B$ pode ser computado com uso da fórmula de Baker-Campbell-Hausdorff, discutida na Seção 11.5, página 682.

Prova da Proposição 11.6. Pela definição

$$e^{A+B} \; = \; \mathbb{1} + \sum_{m=1}^{\infty} \frac{1}{m!}(A+B)^m \; = \; \sum_{m=0}^{\infty} \frac{1}{m!}(A+B)^m \; ,$$

onde convencionamos que $(A+B)^0 = \mathbb{1}$. Como A e B comutam, vale a regra do binômio de Newton[5]

$$(A+B)^m \; = \; \sum_{p=0}^{m} \binom{m}{p} A^p B^{m-p} \; .$$

E. 11.10 *Exercício*. Por quê? Vale a regra do binômio de Newton no caso de A e B não comutarem? Teste alguns exemplos. ✣

Assim,

$$e^{A+B} \; = \; \sum_{m=0}^{\infty}\sum_{p=0}^{m} \frac{1}{m!}\binom{m}{p} A^p B^{m-p} \; = \; \sum_{m=0}^{\infty}\sum_{p=0}^{m} \frac{1}{(m-p)!p!} A^p B^{m-p} \; .$$

Agora, vale a seguinte regra de mudança de ordem de somas:

$$\sum_{m=0}^{\infty}\sum_{p=0}^{m}(\cdots) \; = \; \sum_{p=0}^{\infty}\sum_{m=p}^{\infty}(\cdots) \; .$$

[5] Sir Isaac Newton (1643–1727).

E. 11.11 *Exercício*. Por quê? ✣

Logo,

$$e^{A+B} \; = \; \sum_{p=0}^{\infty}\sum_{m=p}^{\infty} \frac{1}{(m-p)!p!} A^p B^{m-p} \; = \; \sum_{p=0}^{\infty} \frac{1}{p!} A^p \left(\sum_{m=p}^{\infty} \frac{1}{(m-p)!} B^{m-p} \right) \; .$$

Agora, com a mudança de variável $l = m - p$,

$$\sum_{m=p}^{\infty} \frac{1}{(m-p)!} B^{m-p} \; = \; \sum_{l=0}^{\infty} \frac{1}{l!} B^l \; = \; e^B \; .$$

Assim,

$$e^{A+B} \; = \; \sum_{p=0}^{\infty} \frac{1}{p!} A^p e^B \; = \; e^A e^B \; .$$

Analogamente se prova que $e^{A+B} = e^B e^A$. ■

*

Podemos nos perguntar: o que ocorre se A e B não comutarem? Há alguma maneira de calcular $\exp(A+B)$ em termos de produtos de $\exp(A)$ e $\exp(B)$ nesse caso? A resposta a essas questões é dada por três fórmulas muito importantes, a fórmula de Lie-Trotter, a fórmula do comutador e a fórmula de Baker-Campbell-Hausdorff, das quais trataremos mais adiante.

• Algumas propriedades de funções analíticas de matrizes

Os exercícios seguintes, os quais são muito simples de provar, apresentam afirmativas frequentemente usadas sobre funções analíticas de matrizes.

E. 11.12 *Exercício*. Usando a definição (11.21), mostre que

$$P^{-1}\exp(A)P \; = \; \exp\left(P^{-1}AP\right) \tag{11.25}$$

para matrizes $n \times n$ reais ou complexas A e P, sendo P inversível. ✣

E. 11.13 *Exercício*. Usando a definição (11.21), mostre que

$$\exp(A)^T \; = \; \exp\left(A^T\right) \qquad \text{e que} \qquad \exp(A)^* \; = \; \exp\left(A^*\right)$$

para $A \in \mathrm{Mat}(\mathbb{C}, n)$ ou $A \in \mathrm{Mat}(\mathbb{R}, n)$. ✣

Os exercícios acima podem ser facilmente generalizados:

E. 11.14 *Exercício*. Seja $f(z) := \sum_{m=0}^{\infty} f_m z^m$ uma série de potências convergente para $|z| < r_0$ para algum $r_0 > 0$. Então para $A \in \mathrm{Mat}(\mathbb{C}, n)$ com $\|A\| < r_0$ tem-se

$$\left(\sum_{m=0}^{\infty} f_m A^m\right)^T \; = \; \sum_{m=0}^{\infty} f_m \left(A^T\right)^m \qquad \text{e} \qquad \left(\sum_{m=0}^{\infty} f_m A^m\right)^* \; = \; \sum_{m=0}^{\infty} \overline{f_m}\left(A^*\right)^m \; ,$$

ou seja, $f(A)^T = f\left(A^T\right)$ e $f(A)^* = \overline{f}\left(A^*\right)$, onde $\overline{f}(z) := \sum_{m=0}^{\infty} \overline{f_m} z^m = \overline{f\left(\overline{z}\right)}$. Prove essas afirmativas. Prove também que

$$P^{-1}\left(\sum_{m=0}^{\infty} f_m A^m\right) P \; = \; \sum_{m=0}^{\infty} f_m \left(P^{-1}AP\right)^m \; ,$$

ou seja, $P^{-1}f(A)P = f\left(P^{-1}AP\right)$. ✣

Também muito útil é a afirmação contida no seguinte exercício:

E. 11.15 *Exercício*. Sejam $f(z) = \sum_{m=0}^{\infty} f_m z^m$ e $g(z) = \sum_{m=0}^{\infty} g_m z^m$ duas séries de potências convergentes em $|z| < r_1$ e $|z| < r_2$, respectivamente. Sejam A e $B \in \mathrm{Mat}(\mathbb{C},\ n)$ duas matrizes com $\|A\| < r_1$ e $\|B\| < r_2$ tais que $AB = BA$. Então $f(A)g(B) = g(B)f(A)$. Prove isso. ✠

• **O determinante de exponenciais de matrizes**

O Teorema de Decomposição de Jordan (Teorema 10.22, página 598) permite-nos demonstrar o resultado a seguir, muito útil, sobre o determinante de exponenciais de matrizes. Uma primeira demonstração do mesmo foi apresentada na Proposição 10.14, página 548.

Proposição 11.7 *Seja $A \in \mathrm{Mat}(\mathbb{C},\ n)$ ou $A \in \mathrm{Mat}(\mathbb{R},\ n)$. Então vale que*

$$\det\left(e^A\right) \;=\; e^{\mathrm{Tr}(A)} \;. \tag{11.26}$$

□

É suficiente que provemos (11.26) para matrizes complexas primeiro, pois matrizes reais podem ser obtidas de matrizes complexas do limite quando a parte imaginária dos elementos de matriz vai a zero e a continuidade, tanto do lado direito quanto do lado esquerdo de (11.26) em relação aos elementos de matriz de A, garante a validade daquela expressão para matrizes reais também.

Para a prova precisamos de um lema preparatório simples.

Lema 11.1 *Se $D \in \mathrm{Mat}(\mathbb{C},\ n)$ é uma matriz diagonal complexa $n \times n$, então*

$$\det\left(e^D\right) \;=\; e^{\mathrm{Tr}(D)} \;.$$

Igualmente, se $N \in \mathrm{Mat}(\mathbb{C},\ n)$ é uma matriz nilpotente complexa $n \times n$, então

$$\det\left(e^N\right) \;=\; e^{\mathrm{Tr}(N)} \;=\; 1 \;.$$

□

Prova. A parte referente à matriz diagonal é a mais fácil. Suponhamos que D é a matriz diagonal $D = \mathrm{diag}(d_1,\ \ldots,\ d_n)$, sendo que os elementos da diagonal são os autovalores de D. Segue que e^D é a matriz diagonal $D = \mathrm{diag}\left(e^{d_1},\ \ldots,\ e^{d_n}\right)$. Assim, pela Proposição 10.4, página 541, $\det\left(e^D\right) = e^{d_1+\cdots+d_n} = e^{\mathrm{Tr}(D)}$.

Tratemos agora da parte referente à matriz nilpotente N. Iremos provar provar que se N é nilpotente todos os autovalores de e^N são iguais a 1. Pela Proposição 10.31, página 594, os autovalores de N são todos nulos, Assim, se ϕ é um autovetor de N teremos $e^N\phi = \phi$, ou seja, ϕ é autovetor de e^N com autovalor 1. Infelizmente, isso não nos permite concluir diretamente que todos os demais autovetores de e^N têm a mesma propriedade mas, como veremos, isso é verdade.

Vamos supor que o índice de N seja k, ou seja, $N^{k+1} = 0$. Assim, $e^N = \mathbb{1} + \sum_{m=1}^{k} \frac{1}{m!} N^m$. Seja $\psi \neq 0$ um autovetor de e^N com autovalor λ e suponhamos que $\lambda \neq 1$. De $e^N\psi = \lambda\psi$ tem-se

$$(\lambda - 1)\psi \;=\; \sum_{m=1}^{k} \frac{1}{m!} N^m \psi \tag{11.27}$$

e, assim, aplicando N^k a ambos os lados, concluímos que $(\lambda - 1)N^k\psi = 0$, já que no lado direito aparecem potências como $N^{k+1}\psi$, $N^{k+2}\psi$ etc., todas nulas. Como $\lambda \neq 1$, devemos ter $N^k\psi = 0$. Retornando a (11.27), podemos re-escrevê-la como

$$(\lambda - 1)\psi \;=\; \sum_{m=1}^{k-1} \frac{1}{m!} N^m \psi$$

eliminando o termo com $N^k\psi$. Aplicando N^{k-1} a ambos os lados, concluímos que $(\lambda - 1)N^{k-1}\psi = 0$, já que no lado direito aparecem potências como $N^k\psi$, $N^{k+1}\psi$ etc., todas nulas. Como $\lambda \neq 1$, devemos ter $N^{k-1}\psi = 0$. Prosseguindo dessa forma concluiremos, por fim, que $N\psi = 0$. Assim, $e^N\psi = \mathbb{1}\psi = \psi$, provando que $\lambda = 1$, uma contradição.

A conclusão é que todos os autovalores de e^N são iguais a 1, e pela Proposição 10.4, página 541, $\det\left(e^N\right) = 1$. Notemos que, pela Proposição 10.31, página 594, os autovalores de N são todos nulos e, assim, $\mathrm{Tr}(N) = 0$. Logo, $\det\left(e^N\right) = 1 = e^{\mathrm{Tr}(N)}$. Isso completa a prova do lema. ■

Prova da Proposição 11.7. Pelo Teorema de Decomposição de Jordan, na forma do Teorema 10.22, página 598, existe uma matriz inversível T tal que $A = T^{-1}(D + N)T$, onde D é diagonal, N é nilpotente e $DN = ND$. Logo,

$$e^A \;=\; \exp\left(T^{-1}(D+N)T\right) \;=\; T^{-1}\exp(D+N)T \;=\; T^{-1}\exp(D)\exp(N)T \;.$$

Portanto,

$$\det\left(e^A\right) \;=\; \det\left(T^{-1}e^De^NT\right) \;=\; \det\left(T^{-1}\right)\det\left(e^D\right)\det\left(e^N\right)\det\left(T\right) \;=\; \det\left(e^D\right)\det\left(e^N\right) \;,$$

pois $\det\left(T^{-1}\right) = 1/\det(T)$. Assim, pelo Lema 11.1, pela Proposição 10.11 e pela propriedade (10.33),

$$\det\left(e^A\right) \;=\; e^{\mathrm{Tr}(D)}e^{\mathrm{Tr}(N)} \;=\; e^{\mathrm{Tr}(D+N)} \;=\; e^{\mathrm{Tr}\left(T^{-1}(D+N)T\right)} \;=\; e^{\mathrm{Tr}(A)} \;,$$

completando a prova. ■

11.2.1 Exponencial de Matrizes Como Limite de Potências

É um fato bem conhecido que para $z \in \mathbb{C}$ vale a relação (devida a Euler)

$$e^z \;=\; \lim_{n\to\infty}\left(1 + \frac{z}{n}\right)^n \;. \tag{11.28}$$

Demonstrações para o caso em que z é real podem ser encontradas em textos sobre Cálculo (vide, *e.g.*, [504]). Apresentemos aqui uma prova para z complexo. Provemos primeiramente que

$$\lim_{n\to\infty} n\,\ln\left(1 + \frac{z}{n}\right) \;=\; z \;. \tag{11.29}$$

Para $r > 0$ seja Ω_r a região aberta do plano complexo que consiste dos números com valor absoluto menor que r: $\Omega_r := \left\{z \in \mathbb{C}|\ |z| < r\right\}$. Para $n > r$ temos $|z/n| < 1$ para todo $z \in \Omega_r$ e, portanto, a função $\ln(1 + z/n)$ é unívoca (e analítica) em Ω_r. Para $z \in \Omega_r$ e n grande, temos, portanto, fazendo uso da série de Taylor da função logaritmo, que

$$n\ln\left(1 + \frac{z}{n}\right) \;=\; z + n\sum_{m=1}^{\infty}\frac{(-1)^m}{m+1}\left(\frac{z}{n}\right)^{m+1} \;,$$

que converge em Ω_r. Assim, para $z \in \Omega_r$,

$$\left|n\ln\left(1 + \frac{z}{n}\right) - z\right| \;\leq\; n\sum_{m=1}^{\infty}\frac{1}{m+1}\left|\frac{z}{n}\right|^{m+1} \;<\; n\sum_{m=1}^{\infty}\left(\frac{r}{n}\right)^{m+1} \;=\; n\frac{\frac{r^2}{n^2}}{1-\frac{r}{n}} \;=\; \frac{r^2}{n-r} \;,$$

onde, na penúltima igualdade, usamos a fórmula da progressão geométrica infinita. Isso claramente mostra que

$$\lim_{n\to\infty}\left|n\ln\left(1 + \frac{z}{n}\right) - z\right| \;=\; 0 \;,$$

o que estabelece (11.29), independentemente de $r > 0$ e, portanto, para todo $z \in \mathbb{C}$. Segue de (11.29) e da continuidade da função exponencial que

$$e^z \;=\; \exp\left(\lim_{n\to\infty} n\ln\left(1 + \frac{z}{n}\right)\right) \;=\; \lim_{n\to\infty}\exp\left(n\ln\left(1 + \frac{z}{n}\right)\right) \;=\; \lim_{n\to\infty}\left(1 + \frac{z}{n}\right)^n \;,$$

estabelecendo (11.28).

Vamos agora provar que uma relação equivalente a (11.28) vale também para exponenciais de matrizes. Especificamente, temos:

Proposição 11.8 *Seja* $A \in \mathrm{Mat}(\mathbb{C},\ m)$. *Então,*

$$e^A \ = \ \lim_{n\to\infty}\left(\mathbb{1}+\frac{1}{n}A\right)^n . \tag{11.30}$$

□

Demonstração. Pelo Teorema da Decomposição de Jordan, na forma do Teorema 10.22, página 598, existe $T \in \mathrm{Mat}(\mathbb{C},\ m)$ inversível tal que $T^{-1}AT = D + N$, onde D é uma matriz diagonal, N é nilpotente e $DN = ND$. Como já vimos, isso implica que $T^{-1}e^AT = e^De^N = e^{D+N}$. Escrevamos D na sua decomposição espectral: $D = \sum_{a=1}^{b} d_aP_a$, onde $d_1,\ \ldots,\ d_b$ são os autovalores distintos de D (e, portanto, de A) e $P_a \in \mathrm{Mat}(\mathbb{C},\ m)$ são seus respectivos projetores espectrais. Como também já vimos, temos

$$e^D \ = \ \sum_{a=1}^{b} e^{d_a}P_a \ \stackrel{(11.28)}{=} \ \lim_{n\to\infty}\sum_{a=1}^{b}\left(1+\frac{d_a}{n}\right)^n P_a .$$

Agora, pelo Cálculo Funcional, Teorema 10.8, página 564, (especialmente, por (10.61), página 564), tem-se

$$\sum_{a=1}^{b}\left(1+\frac{d_a}{n}\right)^n P_a \ = \ \left(\mathbb{1}+\sum_{a=1}^{b}\frac{d_a}{n}P_a\right)^n \ = \ \left(\mathbb{1}+\frac{1}{n}D\right)^n$$

e assim estabelecemos que

$$e^D \ = \ \lim_{n\to\infty}\left(\mathbb{1}+\frac{1}{n}D\right)^n . \tag{11.31}$$

Podemos igualmente, mas com outra argumentação, provar que $e^N = \lim_{n\to\infty}\left(\mathbb{1}+\frac{1}{n}N\right)^n$. Para tal observemos que, como N é nilpotente, para algum $k \in \mathbb{N}$ vale $N^{k+1} = 0$ e, portanto,

$$e^N \ = \ \mathbb{1}+\sum_{j=1}^{k}\frac{1}{j!}N^j . \tag{11.32}$$

Analogamente, tem-se

$$\left(\mathbb{1}+\frac{1}{n}N\right)^n \ = \ \sum_{p=0}^{n}\binom{n}{p}n^{-p}N^p \ = \ \sum_{p=0}^{k}\frac{n!}{(n-p)!n^p}\frac{1}{p!}N^p$$

a última igualdade sendo válida se escolhermos $n > k$. Como os limites da somatória do lado direito independem de n (para $n > k$), podemos escrever

$$\lim_{n\to\infty}\left(\mathbb{1}+\frac{1}{n}N\right)^n \ = \ \sum_{p=0}^{k}\left(\lim_{n\to\infty}\frac{n!}{(n-p)!n^p}\right)\frac{1}{p!}N^p . \tag{11.33}$$

Vamos agora nos concentrar no termo entre parênteses. Para p fixo e n grande, podemos usar a aproximação de Stirling (7.80), página 433, e substituir

$$\frac{n!}{(n-p)!}\frac{1}{n^p} \ \longrightarrow \ \frac{n^{n+\frac{1}{2}}e^{-n}}{(n-p)^{n-p+\frac{1}{2}}e^{-n+p}}\frac{1}{n^p} \ = \ \frac{1}{\left(1-\frac{p}{n}\right)^{n+\frac{1}{2}}}\left(1-\frac{p}{n}\right)^p\frac{1}{e^p} .$$

Verifique! Agora, no limite quando $n \to \infty$ o fator $\left(1-\frac{p}{n}\right)^n$ converge a e^{-p}, devido a (11.28). O fator $\left(1-\frac{p}{n}\right)^{1/2}$ converge a 1 e o fator $\left(1-\frac{p}{n}\right)^p$ converge igualmente a 1. Assim, concluímos que

$$\lim_{n\to\infty}\frac{n!}{(n-p)!n^p} \ = \ 1 . \tag{11.34}$$

Retornando a (11.33), temos que

$$\lim_{n\to\infty}\left(\mathbb{1}+\frac{1}{n}N\right)^n \ = \ \sum_{p=0}^{k}\frac{1}{p!}N^p \ \stackrel{(11.32)}{=} \ e^N ,$$

como desejado.

Com os fatos colhidos até o momento, obtivemos

$$T^{-1}e^AT \ = \ \left[\lim_{n\to\infty}\left(\mathbb{1}+\frac{1}{n}D\right)^n\right]\left[\lim_{n\to\infty}\left(\mathbb{1}+\frac{1}{n}N\right)^n\right] .$$

Podemos ainda escrever isso como

$$T^{-1}e^AT \ = \ \lim_{n\to\infty}\left[\left(\mathbb{1}+\frac{1}{n}D\right)\left(\mathbb{1}+\frac{1}{n}N\right)\right]^n .$$

Para ver isso, defina-se

$$F_n \ := \ \mathbb{1}+\frac{1}{n}D \qquad \text{e} \qquad G_n \ := \ \mathbb{1}+\frac{1}{n}N .$$

Como D e N comutam entre si, temos

$$e^De^N-(F_n)^n(G_n)^n \ = \ \frac{1}{2}\Big(e^D-(F_n)^n\Big)\Big(e^N+(G_n)^n\Big)+\frac{1}{2}\Big(e^N-(G_n)^n\Big)\Big(e^D+(F_n)^n\Big) .$$

Verifique! Do fato já estabelecido que, $\lim_{n\to\infty}(F_n)^n = e^D$ e $\lim_{n\to\infty}(G_n)^n = e^N$, segue, portanto, que

$$e^De^N \ = \ \lim_{n\to\infty}(F_n)^n(G_n)^n \ = \ \lim_{n\to\infty}\left(F_nG_n\right)^n ,$$

como desejado, com a última igualdade sendo uma mera decorrência do fato de D e N comutarem.

Cabe-nos agora provar que

$$\lim_{n\to\infty}\left[\left(\mathbb{1}+\frac{1}{n}D\right)\left(\mathbb{1}+\frac{1}{n}N\right)\right]^n \ = \ \lim_{n\to\infty}\left(\mathbb{1}+\frac{1}{n}(D+N)\right)^n . \tag{11.35}$$

Para tal, defina-se

$$S_n \ := \ \left(\mathbb{1}+\frac{1}{n}D\right)\left(\mathbb{1}+\frac{1}{n}N\right) \qquad \text{e} \qquad T_n \ := \ \mathbb{1}+\frac{1}{n}(D+N) .$$

Demostrar (11.35) consiste em provar que

$$\lim_{n\to\infty}\Big((S_n)^n-(T_n)^n\Big) \ = \ 0 .$$

Pela soma telescópica, equação (11.20), página 658, temos

$$(S_n)^n-(T_n)^n \ = \ \sum_{p=0}^{n-1}(S_n)^p\left(S_n-T_n\right)(T_n)^{n-1-p} .$$

Disso decorre que

$$\Big\|(S_n)^n-(T_n)^n\Big\| \ \leq \ \sum_{p=0}^{n-1}\|S_n\|^p\,\big\|S_n-T_n\big\|\,\|T_n\|^{n-1-p} .$$

Agora, é claro que

$$\begin{aligned}
\|S_n\| \ &\leq \ \left(1+\frac{\|D\|}{n}\right)\left(1+\frac{\|N\|}{n}\right) ,\\
\|T_n\| \ &\leq \ \left(1+\frac{\|D\|+\|N\|}{n}\right) \ \leq \ \left(1+\frac{\|D\|}{n}\right)\left(1+\frac{\|N\|}{n}\right) \quad \text{e}\\
\big\|S_n-T_n\big\| \ &= \ \left\|\frac{DN}{n^2}\right\| \ \leq \ \frac{\|D\|\,\|N\|}{n^2} .
\end{aligned}$$

Com isso, podemos escrever (verifique!)

$$\left\|(S_n)^n-(T_n)^n\right\| \;\leq\; \sum_{p=0}^{n-1}\left[\left(1+\frac{\|D\|}{n}\right)\left(1+\frac{\|N\|}{n}\right)\right]^{n-1}\frac{\|D\|\,\|N\|}{n^2}$$

$$= \left(1+\frac{\|D\|}{n}\right)^{n-1}\left(1+\frac{\|N\|}{n}\right)^{n-1}\frac{\|D\|\,\|N\|}{n} ,$$

Novamente evocando (11.28), o fator $\left(1+\frac{\|D\|}{n}\right)^{n-1}\left(1+\frac{\|N\|}{n}\right)^{n-1}$ converge a $e^{\|D\|+\|N\|}$ quando $n\to\infty$, restando o fator $\frac{\|D\|\,\|N\|}{n}$ que converge a 0 no mesmo limite.

Isso estabeleceu que $\lim_{n\to\infty}\left\|(S_n)^n-(T_n)^n\right\|=0$ e, portanto, podemos afirmar que

$$T^{-1}e^AT \;=\; = \lim_{n\to\infty}(S_n)^n \;=\; \lim_{n\to\infty}(T_n)^n \;=\; \lim_{n\to\infty}\left(\mathbb{1}+\frac{1}{n}(D+N)\right)^n .$$

Consequentemente,

$$e^A \;=\; \lim_{n\to\infty}\left(\mathbb{1}+\frac{1}{n}T(D+N)T^{-1}\right)^n \;=\; \lim_{n\to\infty}\left(\mathbb{1}+\frac{1}{n}A\right)^n ,$$

como desejávamos demonstrar. ■

11.2.2 A Exponenciação de Matrizes e os Grupos GL($\mathbb{C}$, n) e GL($\mathbb{R}$, n)

Recordemos que GL($\mathbb{C}$, n) (respectivamente, GL($\mathbb{R}$, n)) designa o grupo das matrizes inversíveis complexas (reais) $n\times n$. Aqui discutiremos a relação entre a exponenciação de matrizes e esses grupos. Essa discussão terá um papel mais relevante quando tratarmos da teoria dos grupos de Lie e álgebras de Lie nos Capítulos 21 e 22.

Em primeiro lugar, tem-se a seguinte proposição elementar:

Proposição 11.9 *A aplicação* exp *definida em (11.21) é uma aplicação de* Mat ($\mathbb{C}$, n) *em* GL($\mathbb{C}$, n) *(ou, correspondentemente, de* Mat ($\mathbb{R}$, n) *em* GL($\mathbb{R}$, n)*).* □

Prova. É evidente pela definição (11.21) que $\exp(0)=\mathbb{1}$. Tudo o que se deseja provar é que para qualquer $A\in$ Mat ($\mathbb{C}$, n) então $\exp(A)$ é inversível. Ora, por (11.24), é elementar constatar que $\exp(A)^{-1}=\exp(-A)$. ■

Tem-se também o seguinte:

Proposição 11.10 *Para $n\geq 2$ as aplicações* exp : Mat ($\mathbb{C}$, n) $\to$ GL($\mathbb{C}$, n) *e* exp : Mat ($\mathbb{R}$, n) $\to$ GL($\mathbb{R}$, n) *não são injetoras.* □

Prova. Para matrizes complexas, basta constatar que, no exemplo das matrizes diagonais na forma

$$D \;=\; \operatorname{diag}(2\pi k_1 i,\;\ldots,\;2\pi k_n i) ,$$

com $k_l\in\mathbb{Z}$, tem-se $\exp(D)=\mathbb{1}$.

Para matrizes reais, considere-se a matriz real $A(\alpha):=\alpha J$ onde $J:=\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}$, $\alpha\in\mathbb{R}$. Como facilmente se vê, tem-se para $m\in\mathbb{N}$, $A(\alpha)^{2m}=(-1)^m(\alpha)^{2m}\mathbb{1}$ e $A(\alpha)^{2m+1}=(-1)^m(\alpha)^{2m+1}J$. Daí, como facilmente se verifica por (11.21),

$$\exp(A(\alpha)) \;=\; \cos(\alpha)\mathbb{1}+\operatorname{sen}(\alpha)J \;=\; \begin{pmatrix}\cos\alpha & \operatorname{sen}\alpha\\ -\operatorname{sen}\alpha & \cos\alpha\end{pmatrix} .$$

Logo, $\exp(A(2\pi k))=\mathbb{1}$ para todo $k\in\mathbb{Z}$. Assim a exponenciação de matrizes reais 2×2 não pode ser injetora. É fácil, a partir desse exemplo, construir outros para matrizes reais $n\times n$ com $n\geq 2$. ■

Agora demonstraremos duas proposições nas quais as matrizes reais e complexas se diferenciam.

Proposição 11.11 *As aplicações* exp : Mat ($\mathbb{R}$, n) $\to$ GL($\mathbb{R}$, n), $n\geq 1$, *não são sobrejetoras.* □

Proposição 11.12 *As aplicações* exp : Mat ($\mathbb{C}$, n) $\to$ GL($\mathbb{C}$, n), $n\geq 1$, *são sobrejetoras.* □

Prova da Prop. 11.11. Pela Proposição 11.26, o determinante da exponencial de qualquer matriz real é positivo. Ora, existem em GL($\mathbb{R}$, n) matrizes com determinante negativo. Logo, a exponenciação de matrizes reais não pode ser sobrejetora. ■

À página 668 fazemos alguns comentários adicionais sobre a Proposição 11.11.

Prova da Prop. 11.12. A Proposição 11.12 afirma que toda matriz complexa inversível $n\times n$ pode ser escrita como exponencial de outra matriz complexa $n\times n$. Provemos isso. Seja $A\in$ GL($\mathbb{C}$, n). Pelo Teorema da Decomposição de Jordan (Teorema 10.22, página 598) existe uma matriz inversível P tal que $P^{-1}AP=D+N$ com D diagonal, N nilpotente, $DN=ND$, sendo que D tem na diagonal principal os autovalores da matriz A. Esse último fato diz-nos que D não tem autovalores nulos e, portanto, é também inversível.

Podemos assim escrever $D+N=D(\mathbb{1}+D^{-1}N)$. O que faremos agora é provar os seguintes fatos:

1. D pode ser escrita como $D=e^F$ para alguma matriz F conveniente.

2. $\mathbb{1}+D^{-1}N$ pode ser escrita como $\mathbb{1}+D^{-1}N=e^G$ para alguma matriz G conveniente.

3. Podemos escolher F e G de modo que $FG=GF$.

Desses três fatos concluímos que $P^{-1}AP=\exp(F+G)$ e, portanto, $A=\exp(M)$, onde $M=P(F+G)P^{-1}$, provando o que desejamos.

Prova de 1. Sejam $\alpha_1,\;\ldots,\;\alpha_l$ os autovalores distintos de D. Pelo Teorema Espectral (vide Teorema 10.7, página 562, ou Teorema 10.9, página 567) podemos escrever $D=\sum_{j=1}^{l}\alpha_jE_j$, onde as matrizes E_j satisfazem (10.71) e (10.72) e, de acordo com (10.73), podem ser expressas como polinômios em D (um fato que será usado mais abaixo): $E_j=\frac{1}{m_j(\alpha_j)}m_j(D)$. (Os polinômios m_j foram definidos na demonstração do Teorema 10.9). Seja, para cada j, um número complexo f_j escolhido de forma que $\exp(f_j)=\alpha_j$. Encontrar tais f_j's sempre é possível pois os α_j's são não-nulos, já que D é inversível. Se definirmos

$$F \;:=\; \sum_{j=1}^{l}f_j\,E_j$$

é fácil constatar por (10.71) e (10.72) que $\exp(F)=D$ (faça!). Isso prova *1.* Note que, pelo que comentamos acima, vale

$$F \;=\; \sum_{j=1}^{l}\frac{f_j}{m_j(\alpha_j)}\;m_j(D) , \tag{11.36}$$

ou seja, F pode ser expressa como um polinômio em D.

Prova de 2. Como D^{-1} e N comutam (por quê?), segue que $D^{-1}N$ é nilpotente de ordem, digamos, k, ou seja $\left(D^{-1}N\right)^{k+1} = 0$. Assim, para $z \in \mathbb{C}$ escolhido de modo que $\|zD^{-1}N\| < 1$, o logaritmo de $\mathbb{1} + zD^{-1}N$ está bem definido e vale (vide (11.22))

$$G(z) \;=\; -\sum_{m=1}^{k} \frac{(-z)^m}{m} \left(D^{-1}N\right)^m \;. \tag{11.37}$$

Sabemos pela Proposição 11.5 que nesse caso em que $\|zD^{-1}N\| < 1$, ou seja, $|z| < 1/\|D^{-1}N\|$, temos

$$\exp(G(z)) \;=\; \mathbb{1} + zD^{-1}N \;. \tag{11.38}$$

Queremos agora provar que essa igualdade vale para <u>todo</u> z. Usando novamente o fato que as matrizes D^{-1} e N comutam entre si, o fato que $\left(D^{-1}N\right)^{k+1} = 0$ e o fato que a soma em (11.37) é finita, teremos

$$\exp(G(z)) \;=\; \exp\left(-\sum_{m=1}^{k} \frac{(-z)^m}{m} \left(D^{-1}N\right)^m\right) \;=\; \prod_{m=1}^{k} \exp\left(-\frac{(-z)^m}{m} \left(D^{-1}N\right)^m\right)$$

$$= \prod_{m=1}^{k} \left[\mathbb{1} + \sum_{l=1}^{k} \frac{(-1)^l}{l!} \frac{(-z)^{ml}}{m^l} \left(D^{-1}N\right)^{ml}\right] \;.$$

Como as somas a produtos acima são finitos (consequência da nilpotência de $D^{-1}N$), constatamos que $\exp(G(z))$ é um polinômio em z para todo $z \in \mathbb{C}$. Ora, já verificamos acima que, quando $|z|$ é pequeno, $\exp(G(z))$ é igual ao polinômio em z dado por $\mathbb{1} + zD^{-1}N$. Como polinômios são funções analíticas em toda parte isso implica que $\exp(G(z)) = \mathbb{1} + zD^{-1}N$ para <u>todo</u> $z \in \mathbb{C}$. Em particular, para $z = 1$, o que significa que $\mathbb{1} + D^{-1}N = \exp(G)$, onde

$$G \;\equiv\; G(1) \;=\; \sum_{m=1}^{k} \frac{(-1)^{m+1}}{m} \;\left(D^{-1}N\right)^m \;. \tag{11.39}$$

E. 11.16 <u>*Exercício*</u>. Usando a definição (11.39), prove explicitamente que $\exp(G) = \mathbb{1} + D^{-1}N$. ✣

Prova de 3. Por (11.36), F é um polinômio em D. Assim, F comuta com D^{-1} e com N. Logo, por (11.39), F comuta com G. Isso é o que queríamos provar e, assim, a prova da Proposição 11.12 está completa. ■

• **<u>Comentários sobre a Proposição 11.11</u>**

Sobre matrizes reais é possível dizer mais que o enunciado da Proposição 11.11 e sua prova. Em verdade, não são apenas as matrizes com determinante negativo que estão fora da imagem da exponenciação de matrizes reais. Há algumas com determinante positivo que também estão fora. Se M é uma matriz real inversível, então seus autovalores são as raízes do polinômio característico $p(x) = \det(x\mathbb{1} - M)$. Como M é real, esse polinômio tem coeficientes reais e, como é bem sabido, as raízes de polinômios com coeficientes reais ou são números reais ou são pares de números complexos complexo-conjugados uns dos outros. Por exemplo, as raízes do polinômio característico da matriz $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ são $\pm i$. De qualquer forma, uma matriz com determinante positivo pode, digamos, ter duas raízes negativas distintas <u>simples</u>, como é, por exemplo, o caso da matriz

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{pmatrix} . \tag{11.40}$$

Isso posto, estudemos os autovalores das matrizes da forma e^A com A real. Esses são as raízes do polinômio característico $p(x) = \det(x\mathbb{1} - e^A)$. Como toda matriz real é também membro de $\mathrm{Mat}\,(\mathbb{C},\, n)$ podemos aplicar o Teorema da Decomposição de Jordan (Teorema 10.22, página 598) e afirmar que existe uma matriz inversível complexa P tal que $P^{-1}AP = D + N$ com D diagonal, N nilpotente, $DN = ND$, sendo que D tem na diagonal os autovalores da matriz real A. Assim, pela propriedade do determinante,

$$p(x) \;=\; \det(x\mathbb{1} - e^A) \;=\; \det\left(P^{-1}(x\mathbb{1} - e^A)P\right) \;=\; \det(x\mathbb{1} - e^D e^N) \;.$$

É fácil de ver daí[6] que os autovalores de e^A são os elementos da diagonal da matriz diagonal e^D, que são, como comentamos acima, exponenciais dos autovalores da matriz real A. Podemos nos perguntar: podem os elementos da diagonal de e^D serem números negativos? A resposta é sim, mas para isso é necessário que A tenha um autovalor complexo cuja parte imaginária seja da forma $(2k+1)\pi$, com k inteiro. Ora, como A é real, existe pelo que comentamos acima, um outro autovalor complexo de A cuja parte imaginária é da forma $-(2k+1)\pi$, pois os autovalores complexos aparecem em pares complexo-conjugados. Isso diz-nos que os autovalores negativos de e^A têm multiplicidade <u>par</u>! Ora, isso nem sempre é o caso para matrizes inversíveis, como mostra o exemplo do último parágrafo. Assim, matrizes reais com determinante positivo e com pelo menos um autovalor negativo com multiplicidade ímpar não estão na imagem da exponencial de nenhuma matriz real. Tal é o caso da matriz de (11.40). Em verdade, mesmo matrizes com determinante positivo e com autovalores negativos com multiplicidade par podem não estar na imagem da exponencial. Tal é o caso das matrizes $\left(\begin{smallmatrix} -1 & a \\ 0 & -1 \end{smallmatrix}\right)$ com $a \neq 0$ (mostre isso).

11.3 A Fórmula de Lie-Trotter e a Fórmula do Comutador

Há duas expressões envolvendo produtos de exponenciais de matrizes que são bastante úteis. São as fórmulas conhecidas como *fórmula de Lie-Trotter*[7] e *fórmula do comutador*. A fórmula de Lie-Trotter é importante não apenas no estudo de grupos de Lie matriciais mas também na Mecânica Estatística e na Mecânica Quântica, onde é frequentemente empregada. A fórmula de Lie-Trotter, por exemplo, é usada na Mecânica Estatística para relacionar sistemas quânticos de spin a sistemas clássicos de spin.

Proposição 11.13 *Para quaisquer matrizes $A,\ B \in \mathrm{Mat}\,(\mathbb{C},\ n)$ valem:*

Fórmula de Lie-Trotter:

$$\exp\left(A + B\right) \;=\; \lim_{m\to\infty} \left[\exp\left(\frac{1}{m}A\right)\exp\left(\frac{1}{m}B\right)\right]^m \;. \tag{11.41}$$

Fórmula do Comutador:

$$\exp\left([A,\ B]\right) \;=\; \lim_{m\to\infty} \left[\exp\left(\frac{1}{m}A\right)\exp\left(\frac{1}{m}B\right)\exp\left(-\frac{1}{m}A\right)\exp\left(-\frac{1}{m}B\right)\right]^{m^2} \;. \tag{11.42}$$

□

Prova. Vamos primeiramente provar a fórmula de Lie-Trotter[8] e posteriormente passar à fórmula do comutador. Começamos definindo, para $m \in \mathbb{N}$,

$$S_m \;:=\; \exp\left(\frac{1}{m}A\right)\exp\left(\frac{1}{m}B\right) \qquad \text{e} \qquad T_m \;:=\; \exp\left(\frac{1}{m}(A+B)\right) \;.$$

Note-se que $(T_m)^m = \exp\left(A+B\right)$ e que tudo o que desejamos é provar que $(S_m)^m$ converge a $exp\left(A+B\right)$, ou seja,

$$\lim_{m\to\infty} \|(S_m)^m - (T_m)^m\|_{\mathbb{C}} \;=\; 0 \;.$$

[6]Pois numa base conveniente a matriz $e^D e^N$ é uma matriz triangular superior, tendo na diagonal principal os elementos da diagonal de e^D.

[7]A fórmula de Lie-Trotter foi originalmente demonstrada por Lie (Marius Sophus Lie (1842–1899)) e posteriormente generalizada por vários autores, entre eles Trotter (Hale Freeman Trotter (1931–)) em "On the Product of Semi-Groups of Operators". Proc. Amer. Math. Soc. **10**, 545–551 (1959). O leitor poderá encontrar várias dessas generalizações (por exemplo para operadores autoadjuntos não-limitados agindo em espaços de Hilbert) em [438]. O assunto é ainda hoje objeto de pesquisa.

[8]Para a fórmula de Lie-Trotter seguiremos aqui a demonstração de [438].

Precisamos, portanto, estudar $(S_m)^m - (T_m)^m$. Para isso, é útil empregarmos a identidade algébrica (11.20), página 658 ("soma telescópica"). Daquela relação e das propriedades da norma operatorial, segue que

$$\|(S_m)^m - (T_m)^m\|_{\mathbb{C}} \;\leq\; \sum_{p=0}^{m-1} \|S_m\|_{\mathbb{C}}^p \,\|S_m - T_m\|_{\mathbb{C}}\, \|T_m\|_{\mathbb{C}}^{m-1-p} \,. \tag{11.43}$$

Pela definição, temos para qualquer matriz $M \in \mathrm{Mat}\,(\mathbb{C},\ n)$

$$\|\exp{(M)}\|_{\mathbb{C}} \;=\; \left\|\sum_{k=0}^{\infty}\frac{1}{k!}M^k\right\|_{\mathbb{C}} \;\leq\; \sum_{k=0}^{\infty}\frac{1}{k!}\|M\|_{\mathbb{C}}^k \;=\; e^{\|M\|_{\mathbb{C}}}\,.$$

Assim,

$$\|S_m\|_{\mathbb{C}} \;\leq\; \left\|\exp\left(\frac{1}{m}A\right)\right\|_{\mathbb{C}} \left\|\exp\left(\frac{1}{m}B\right)\right\|_{\mathbb{C}} \;\leq\; e^{(\|A\|_{\mathbb{C}}+\|B\|_{\mathbb{C}})/m}$$

e $\|T_m\|_{\mathbb{C}} \leq e^{(\|A\|_{\mathbb{C}}+\|B\|_{\mathbb{C}})/m}$. Retornando a (11.43), teremos

$$\|(S_m)^m - (T_m)^m\|_{\mathbb{C}} \;\leq\; e^{(\|A\|_{\mathbb{C}}+\|B\|_{\mathbb{C}})(m-1)/m}\sum_{p=0}^{m-1}\|S_m - T_m\|_{\mathbb{C}} \;\leq\; m\|S_m - T_m\|_{\mathbb{C}}\, e^{(\|A\|_{\mathbb{C}}+\|B\|_{\mathbb{C}})}\,.$$

Na última desigualdade usamos que $(m-1)/m < 1$ e que $\|S_m - T_m\|_{\mathbb{C}}$ não depende de p.

Como se vê da última expressão, tudo o que temos de fazer para mostrar que $\|(S_m)^m - (T_m)^m\|_{\mathbb{C}}$ vai a zero quando $m \to \infty$ é provar que $\|S_m - T_m\|_{\mathbb{C}}$ vai a zero com $1/m^2$ quando m cresce. Isso é feito escrevendo as expressões explícitas para S_m e T_m em termos da série de Taylor da função exponencial:

$$\begin{aligned} S_m - T_m \;&=\; \exp\left(\frac{1}{m}A\right)\exp\left(\frac{1}{m}B\right) - \exp\left(\frac{1}{m}(A+B)\right) \\ &=\; \left[\mathbb{1} + \frac{1}{m}A + \sum_{k=2}^{\infty}\frac{m^{-k}}{k!}A^k\right]\left[\mathbb{1} + \frac{1}{m}B + \sum_{k=2}^{\infty}\frac{m^{-k}}{k!}B^k\right] - \left[\mathbb{1} + \frac{1}{m}(A+B) + \sum_{k=2}^{\infty}\frac{m^{-k}}{k!}(A+B)^k\right]. \end{aligned}$$

Expandindo-se a última linha, e identificando os termos em $1/m$, é fácil constatar que

$$S_m - T_m \;=\; \mathbb{1} + \frac{1}{m}A + \frac{1}{m}B - \mathbb{1} - \frac{1}{m}(A+B) + \frac{1}{m^2}\mathcal{S}_m \;=\; \frac{1}{m^2}\mathcal{S}_m\,,$$

onde $\mathcal{S}_m$ é uma série, um tanto complicada, mas convergente em norma e tal que $\lim_{m\to\infty}\|\mathcal{S}_m\|_{\mathbb{C}} =$ finito. Assim, $m\|S_m - T_m\|_{\mathbb{C}} \leq \dfrac{1}{m}\|\mathcal{S}_m\|_{\mathbb{C}}$ e, portanto, $\lim\limits_{m\to\infty}\|(S_m)^m - (T_m)^m\|_{\mathbb{C}} = 0$. Isso demonstrou a fórmula de Lie-Trotter. O estudante mais avançado pode facilmente convencer-se que precisamente a mesma demonstração se aplica ao contexto de operadores limitados agindo em espaços de Banach.

Para a fórmula do comutador usaremos outro procedimento. Definimos

$$U_m \;:=\; \exp\left(\frac{1}{m}A\right)\exp\left(\frac{1}{m}B\right)\exp\left(-\frac{1}{m}A\right)\exp\left(-\frac{1}{m}B\right)$$

e teremos

$$\begin{aligned} U_m \;=\; &\left[\mathbb{1} + \frac{1}{m}A + \frac{1}{2m^2}A^2 + \sum_{k=3}^{\infty}\frac{m^{-k}}{k!}A^k\right]\left[\mathbb{1} + \frac{1}{m}B + \frac{1}{2m^2}B^2 + \sum_{k=3}^{\infty}\frac{m^{-k}}{k!}B^k\right] \\ &\times\left[\mathbb{1} - \frac{1}{m}A + \frac{1}{2m^2}A^2 + \sum_{k=3}^{\infty}\frac{(-m)^{-k}}{k!}A^k\right]\left[\mathbb{1} - \frac{1}{m}B + \frac{1}{2m^2}B^2 + \sum_{k=3}^{\infty}\frac{(-m)^{-k}}{k!}B^k\right]. \end{aligned}$$

Com um pouco de paciência podemos expandir o produto dos quatro fatores do lado direito e constatar (faça!) que os termos envolvendo $1/m$ se cancelam e o termo proporcional a $1/m^2$ é $AB - BA$ (outros termos como $(1/m^2)A^2$ e $(1/m^2)B^2$ também se cancelam. Verifique!). Ou seja, ficamos com

$$U_m \;=\; \mathbb{1} + \frac{1}{m^2}(AB - BA) + \frac{1}{m^3}\mathcal{R}_m\,, \tag{11.44}$$

onde $\frac{1}{m^3}\mathcal{R}_m$ são os termos restantes da expansão. $\mathcal{R}_m$ é uma expressão complicada, mas envolvendo séries convergentes e de tal forma que $\lim_{m\to\infty}\|\mathcal{R}_m\|_{\mathbb{C}}$ é finito.

Isso diz que para m grande o suficiente a norma de $U_m - \mathbb{1}$ é pequena e, assim, podemos tomar o logaritmo de U_m, definido por $\ln(U_m) = \ln(\mathbb{1} + (U_m - \mathbb{1}))$. Por (11.44) e pela expansão do logaritmo teremos

$$\ln(U_m) \;=\; \ln\big(\mathbb{1} + (U_m - \mathbb{1})\big) \;=\; \ln\left(\mathbb{1} + \frac{1}{m^2}(AB - BA) + \frac{1}{m^3}\mathcal{R}_m\right) \;=\; \frac{1}{m^2}(AB - BA) + \frac{1}{m^3}\mathcal{R}'_m\,,$$

ou seja,

$$m^2\ln(U_m) \;=\; [A,\ B] + \frac{1}{m}\mathcal{R}'_m\,, \tag{11.45}$$

onde $\mathcal{R}'_m$ é novamente uma expressão complicada, mas envolvendo séries convergentes e de tal forma que $\lim_{m\to\infty}\|\mathcal{R}'_m\|_{\mathbb{C}}$ é finito. Como $\lim_{m\to\infty}\frac{1}{m}\mathcal{R}'_m = 0$ podemos escrever, pela Proposição 11.3,

$$\exp\big([A,\ B]\big) \;=\; \lim_{m\to\infty}\exp\left([A,\ B] + \frac{1}{m}\mathcal{R}'_m\right).$$

Agora, por (11.45),

$$\exp\left([A,\ B] + \frac{1}{m}\mathcal{R}'_m\right) \;=\; \exp\big(m^2\ln(U_m)\big) \;=\; \Big(\exp\big(\ln(U_m)\big)\Big)^{m^2} \;=\; (U_m)^{m^2}\,.$$

Logo,

$$\exp\big([A,\ B]\big) \;=\; \lim_{m\to\infty}(U_m)^{m^2}\,.$$

Isso é o que desejávamos provar[9]. ■

A demonstração que apresentamos da fórmula do comutador pode ser usada para obter-se uma outra demonstração da fórmula de Lie-Trotter. Isso é o conteúdo exercício que segue.

E. 11.17 *Exercício*. Demonstre a fórmula de Lie-Trotter usando as ideias da prova da fórmula do comutador, exibida acima. ✣

• Uma segunda versão da fórmula de Lie-Trotter

A fórmula da Lie-Trotter é por vezes evocada (notadamente na Mecânica Estatística) em uma forma ligeiramente diferente:

$$\exp{(A+B)} \;=\; \lim_{m\to\infty}\left[\exp\left(\frac{1}{2m}A\right)\exp\left(\frac{1}{m}B\right)\exp\left(\frac{1}{2m}A\right)\right]^m\,. \tag{11.46}$$

E. 11.18 *Exercício*. Demonstre (11.46) a partir de (11.41). ***Sugestão:*** verifique primeiramente que, para todo $m \in \mathbb{N}$, vale

$$\left[\exp\left(\frac{1}{2m}A\right)\exp\left(\frac{1}{m}B\right)\exp\left(\frac{1}{2m}A\right)\right]^m \;=\; \exp\left(\frac{1}{2m}A\right)\left[\exp\left(\frac{1}{m}B\right)\exp\left(\frac{1}{m}A\right)\right]^m\exp\left(-\frac{1}{2m}A\right)$$

e, em seguida, use (11.41), tomando adequadamente o limite $m \to \infty$. ✣

A vantagem de (11.46) sobre (11.41) reside no fato de que $\exp\left(\frac{1}{2m}A\right)\exp\left(\frac{1}{m}B\right)\exp\left(\frac{1}{2m}A\right)$ é autoadjunta se A e B o forem, enquanto que $\exp\left(\frac{1}{m}A\right)\exp\left(\frac{1}{m}B\right)$ não é. Em certas aplicações (notadamente na Mecânica Estatística) é importante preservar a autoadjunção dos aproximantes de $\exp(A+B)$ usados na fórmula da Lie-Trotter.

[9] O estudante pode estar curioso (ou perplexo) sobre o por quê de não finalizamos a demonstração partindo de (11.45), escrevendo $m^2\ln(U_m) = \ln((U_m)^{m^2})$ e tomando diretamente daí o limite $m \to \infty$. A razão é que o fato de U_m ser próximo de $\mathbb{1}$ em norma não garante que $(U_m)^{m^2}$ também o seja. Assim, o logaritmo de $(U_m)^{m^2}$ pode não fazer sentido. Para evitar esse transtorno lógico é mais conveniente finalizar a demonstração com uso da função exponencial de matrizes, para a qual tais problemas de definição não ocorrem.

11.4 Aplicações Lineares em $\mathrm{Mat}(\mathbb{C},\ n)$

O conjunto de matrizes $\mathrm{Mat}(\mathbb{C},\ n)$ é naturalmente um espaço vetorial complexo de dimensão finita n^2, pois combinações lineares de matrizes complexas $n \times n$ são novamente matrizes complexas $n \times n$, com a matriz nula fazendo o papel de vetor nulo. Em áreas relacionadas à Teoria de Grupos e à Mecânica Quântica (Informação Quântica) há interesse no estudo de aplicações lineares agindo no espaço vetorial $\mathrm{Mat}(\mathbb{C},\ n)$. Na Seção 11.4.1 apresentaremos alguns fatos gerais sobre tais aplicações lineares e na Seção 11.4.2, página 677, vamos exibir e estudar algumas dessas aplicações lineares de interesse específico e discutir suas relações. Os resultados aos quais chegaremos têm interesse por si só, mas nossa intenção é também a de preparar a demonstração da fórmula de Baker-Campbell-Hausdorff, a ser realizada na Seção 11.5, página 682.

11.4.1 Alguns Fatos Gerais sobre Aplicações Lineares em $\mathrm{Mat}(\mathbb{C},\ n)$

Como espaço vetorial complexo, $\mathrm{Mat}(\mathbb{C},\ n)$ pode ser dotado de diversos produtos escalares. O mais relevante, talvez, e que empregaremos no que segue, é aquele definido em (11.13): $\langle A,\ B\rangle := \mathrm{Tr}(A^*B) = \sum_{i=1}^{n}\sum_{j=1}^{n}\overline{A_{ij}}\,B_{ij}$, com $A,\ B \in \mathrm{Mat}(\mathbb{C},\ n)$.

Dizemos que uma coleção $\{F^\alpha,\ \alpha \in \{1,\ \ldots,\ n^2\}\}$ de n^2 elementos linearmente independentes de $\mathrm{Mat}(\mathbb{C},\ n)$ é uma base ortonormal em $\mathrm{Mat}(\mathbb{C},\ n)$, se valer $\langle F^\alpha,\ F^\beta\rangle = \mathrm{Tr}\Big(\big(F^\alpha\big)^* F^\beta\Big) = \delta_{\alpha\beta}$.

$\mathrm{Mat}(\mathbb{C},\ n)$ possui uma base natural de vetores que, coincidentemente, é uma base ortonormal em relação ao produto escalar acima. Trata-se da base composta pelas n^2 matrizes $E^{a,\,b}$, com $a,\ b \in \{1,\ \ldots,\ n\}$, onde E^{ab} é a matriz cujo elemento ij é nulo a menos que $i = a$ e que $j = b$, em cujo caso $(E^{a,\,b})_{ij} = 1$. Em símbolos,

$$\big(E^{a,\,b}\big)_{ij} \;=\; \delta_{ia}\,\delta_{jb}\ .$$

Note-se que $\big(E^{a,\,b}\big)^* = E^{b,\,a}$. Claro está que toda matriz $A \in \mathrm{Mat}(\mathbb{C},\ n)$ pode ser escrita na forma

$$A \;=\; \sum_{a=1}^{n}\sum_{b=1}^{n} A_{ab}\,E^{a,\,b}$$

e que

$$\big\langle E^{a,\,b},\ E^{c,\,d}\big\rangle \;=\; \sum_{i=1}^{n}\sum_{j=1}^{n}(E^{a,\,b})_{ij}(E^{c,\,d})_{ij} \;=\; \left(\sum_{i=1}^{n}\delta_{ia}\,\delta_{ic}\right)\left(\sum_{j=1}^{n}\delta_{jb}\,\delta_{jd}\right) \;=\; \delta_{ac}\delta_{bd}\ ,$$

mostrando que $\big\{E^{a,\,b},\ a,\ b \in \{1,\ \ldots,\ n\}\big\}$ é uma base ortonormal em $\mathrm{Mat}(\mathbb{C},\ n)$.

• **O espaço $\mathcal{L}\big(\mathrm{Mat}(\mathbb{C},\ n)\big)$ das aplicações lineares de $\mathrm{Mat}(\mathbb{C},\ n)$ em si mesmo**

Uma aplicação $\mathsf{L} : \mathrm{Mat}(\mathbb{C},\ n) \to \mathrm{Mat}(\mathbb{C},\ n)$ é dita ser uma aplicação linear se satisfizer $\mathsf{L}(zA+wB) = z\mathsf{L}(A)+w\mathsf{L}(B)$ para todos $z,\ w \in \mathbb{C}$ e todas $A,\ B \in \mathrm{Mat}(\mathbb{C},\ n)$. Denotaremos por $\mathcal{L}\big(\mathrm{Mat}(\mathbb{C},\ n)\big)$ o conjunto de todas as aplicações lineares de $\mathrm{Mat}(\mathbb{C},\ n)$ em si mesmo. É bastante claro que $\mathcal{L}\big(\mathrm{Mat}(\mathbb{C},\ n)\big)$ é também um espaço vetorial complexo, pois se $\mathsf{L},\ \mathsf{M} \in \mathcal{L}\big(\mathrm{Mat}(\mathbb{C},\ n)\big)$ e $z,\ w \in \mathbb{C}$, definimos $z\mathsf{L}+w\mathsf{M}$ como o elemento de $\mathcal{L}\big(\mathrm{Mat}(\mathbb{C},\ n)\big)$ dado por $\big(z\mathsf{L}+w\mathsf{M}\big)(A) := z\mathsf{L}(A)+w\mathsf{M}(A)$ para todo $A \in \mathrm{Mat}(\mathbb{C},\ n)$.

Podemos dotar $\mathcal{L}\big(\mathrm{Mat}(\mathbb{C},\ n)\big)$ de um produto escalar por meio do seguinte procedimento. Seja $\big\{F^\alpha,\ \alpha \in \{1,\ \ldots,\ n^2\}\big\}$ uma base ortonormal em $\mathrm{Mat}(\mathbb{C},\ n)$. Definimos para L, M a expressão.

$$\big\langle \mathsf{L},\ \mathsf{M}\big\rangle \;:=\; \sum_{\alpha=1}^{n^2}\mathrm{Tr}\Big(\mathsf{L}(F^\alpha)^*\ \mathsf{M}(F^\alpha)\Big)\ . \tag{11.47}$$

É evidente que trata-se de uma forma sesquilinear e é fácil ver que é uma forma sesquilinear Hermitiana, pois

$$\overline{\big\langle \mathsf{L},\ \mathsf{M}\big\rangle} \;=\; \sum_{\alpha=1}^{n^2}\overline{\mathrm{Tr}\Big(\mathsf{L}(F^\alpha)^*\ \mathsf{M}(F^\alpha)\Big)} \;=\; \sum_{\alpha=1}^{n^2}\mathrm{Tr}\Big(\mathsf{M}(F^\alpha)^*\mathsf{L}(F^\alpha)\Big) \;=\; \big\langle \mathsf{M},\ \mathsf{L}\big\rangle\ ,$$

onde usamos que $\overline{\mathrm{Tr}(A)} = \mathrm{Tr}(A^*)$ para toda $A \in \mathrm{Mat}(\mathbb{C},\ n)$. É também claro que

$$\big\langle \mathsf{L},\ \mathsf{L}\big\rangle \;:=\; \sum_{\alpha=1}^{n^2}\mathrm{Tr}\Big(\mathsf{L}(F^\alpha)^*\mathsf{L}(F^\alpha)\Big) \;\geq\; 0$$

para todo $\mathsf{L} \in \mathcal{L}\big(\mathrm{Mat}(\mathbb{C},\ n)\big)$. Tem-se também que $\big\langle \mathsf{L},\ \mathsf{L}\big\rangle = 0$ implica que $\mathrm{Tr}\Big(\mathsf{L}(F^\alpha)^*\mathsf{L}(F^\alpha)\Big) = 0$ para todo α, o que implica que $\mathsf{L}(F^\alpha) = 0$ para todo α, o que, por sua vez implica que $\mathsf{L} = 0$, pois os F^α compõem uma base em $\mathrm{Mat}(\mathbb{C},\ n)$. Isso estabeleceu que (11.47) é, de fato, um produto escalar em $\mathcal{L}\big(\mathrm{Mat}(\mathbb{C},\ n)\big)$. É interessante ainda mostrar que (11.47) independe da particular base ortonormal adotada em $\mathrm{Mat}(\mathbb{C},\ n)$. Para ver isso, seja $\{G^\alpha,\ \alpha \in \{1,\ \ldots,\ n^2\}\}$ uma outra base ortonormal em $\mathrm{Mat}(\mathbb{C},\ n)$ e escrevamos

$$F^\alpha \;=\; \sum_{\beta=1}^{n^2}\mathrm{Tr}\big((G^\beta)^*F^\alpha\big)G^\beta\ .$$

Teremos

$$\big\langle \mathsf{L},\ \mathsf{M}\big\rangle \;=\; \sum_{\alpha=1}^{n^2}\sum_{\beta=1}^{n^2}\sum_{\gamma=1}^{n^2}\overline{\mathrm{Tr}\Big((G^\beta)^*F^\alpha\Big)}\ \mathrm{Tr}\Big((G^\gamma)^*F^\alpha\Big)\ \mathrm{Tr}\Big(\mathsf{L}(G^\beta)^*\mathsf{M}(G^\gamma)\Big)\ . \tag{11.48}$$

Agora,

$$\sum_{\alpha=1}^{n^2}\overline{\mathrm{Tr}\Big((G^\beta)^*F^\alpha\Big)}\mathrm{Tr}\Big((G^\gamma)^*F^\alpha\Big) \;=\; \sum_{\alpha=1}^{n^2}\mathrm{Tr}\Big((F^\alpha)^*G^\beta\Big)\ \mathrm{Tr}\Big((G^\gamma)^*F^\alpha\Big)$$

$$=\; \mathrm{Tr}\left((G^\gamma)^*\left(\sum_{\alpha=1}^{n^2}\mathrm{Tr}\big((F^\alpha)^*G^\beta\big)F^\alpha\right)\right) \;=\; \mathrm{Tr}\Big((G^\gamma)^*G^\beta\Big) \;=\; \delta_{\gamma\beta}\ .$$

Retornando com isso a (11.48), obtemos

$$\big\langle \mathsf{L},\ \mathsf{M}\big\rangle \;=\; \sum_{\beta=1}^{n^2}\sum_{\gamma=1}^{n^2}\delta_{\gamma\beta}\mathrm{Tr}\Big(\mathsf{L}(G^\beta)^*\mathsf{M}(G^\gamma)\Big) \;=\; \sum_{\beta=1}^{n^2}\mathrm{Tr}\Big(\mathsf{L}(G^\beta)^*\ \mathsf{M}(G^\beta)\Big)\ ,$$

estabelecendo a independência que desejávamos provar.

• **Uma identidade para o traço de elementos de $\mathrm{Mat}(\mathbb{C},\ n)$**

A seguinte identidade é importante e será empregada adiante: para toda $A \in \mathrm{Mat}(\mathbb{C},\ n)$ vale

$$\mathrm{Tr}(A)\mathbb{1} \;=\; \sum_{a=1}^{n}\sum_{b=1}^{n}\big(E^{a,\,b}\big)^*AE^{a,\,b}\ . \tag{11.49}$$

Para demonstrá-la, determinemos o elemento ij da matriz $\big(E^{a,\,b}\big)^*AE^{a,\,b}$. Pela regra de produto de matrizes, temos

$$\Big(\big(E^{a,\,b}\big)^*AE^{a,\,b}\Big)_{ij} \;=\; \sum_{k=1}^{n}\sum_{l=1}^{n}\big(E^{a,\,b}\big)^*_{ik}A_{kl}\big(E^{a,\,b}\big)_{lj} \;=\; \sum_{k=1}^{n}\sum_{l=1}^{n}\big(E^{a,\,b}\big)_{ki}A_{kl}\big(E^{a,\,b}\big)_{lj}$$

$$=\; \sum_{k=1}^{n}\sum_{l=1}^{n}\delta_{ka}\,\delta_{ib}A_{kl}\delta_{la}\,\delta_{jb} \;=\; A_{aa}\delta_{ib}\,\delta_{jb}\ .$$

Logo, o elemento ij da matriz do lado direito de (11.49) vale

$$\sum_{a=1}^{n}\sum_{b=1}^{n}A_{aa}\delta_{ib}\,\delta_{jb} \;=\; \left(\sum_{a=1}^{n}A_{aa}\right)\sum_{b=1}^{n}\delta_{ib}\,\delta_{jb} \;=\; \mathrm{Tr}(A)\,\delta_{ij}\ ,$$

que é o elemento ij da matriz $\mathrm{Tr}(A)\mathbb{1}$, provando (11.49).

Uma das razões pelas quais a relação (11.49) é relevante é que a mesma pode ser estendida para outras bases orgonormais em $\mathrm{Mat}(\mathbb{C},\ n)$. Seja $\{F^\alpha,\ \alpha \in \{1,\ \ldots,\ n^2\}\}$ uma base ortonormal em $\mathrm{Mat}(\mathbb{C},\ n)$, ou seja, tal que $\langle F^\alpha,\ F^\beta\rangle = \mathrm{Tr}\Big(\big(F^\alpha\big)^* F^\beta\Big) = \delta_{\alpha\beta}$. Afirmamos que vale

$$\mathrm{Tr}(A)\mathbb{1} \;=\; \sum_{\alpha=1}^{n^2} \big(F^\alpha\big)^* A F^\alpha \;. \tag{11.50}$$

Para a demonstração, observemos que podemos escrever

$$E^{a,\,b} \;=\; \sum_{\alpha=1}^{n^2} \mathcal{G}^{ab}{}_\alpha F^\alpha \;,$$

para certas constantes $\mathcal{G}^{ab}{}_\alpha \in \mathbb{C}$, pois as matrizes F^α formam uma base. Usando a ortonormalidade dessas matrizes, segue facilmente também que

$$\mathcal{G}^{ab}{}_\alpha \;=\; \mathrm{Tr}\Big(\big(F^\alpha\big)^* E^{a,\,b}\Big) \;=\; \sum_{k=1}^{n}\sum_{l=1}^{n} \overline{\big(F^\alpha\big)_{kl}}\big(E^{a,\,b}\big)_{kl} \;=\; \sum_{k=1}^{n}\sum_{l=1}^{n} \overline{\big(F^\alpha\big)_{kl}}\,\delta_{ka}\,\delta_{lb} \;=\; \overline{\big(F^\alpha\big)_{ab}} \;. \tag{11.51}$$

Assim, temos por (11.49) que

$$\mathrm{Tr}(A)\mathbb{1} \;=\; \sum_{a=1}^{n}\sum_{b=1}^{n}\left(\sum_{\alpha=1}^{n^2}\mathcal{G}^{ab}{}_\alpha F^\alpha\right)^* A \left(\sum_{\beta=1}^{n^2}\mathcal{G}^{ab}{}_\beta F^\beta\right) \;=\; \sum_{a=1}^{n}\sum_{b=1}^{n}\sum_{\alpha=1}^{n^2}\sum_{\beta=1}^{n^2} \big(F^\alpha\big)_{ab}\,\overline{\big(F^\beta\big)_{ab}}\,\big(F^\alpha\big)^* A F^\beta \tag{11.52}$$

Agora,

$$\sum_{a=1}^{n}\sum_{b=1}^{n} \big(F^\alpha\big)_{ab}\,\overline{\big(F^\beta\big)_{ab}} \;=\; \mathrm{Tr}\big(\big(F^\beta\big)^* F^\alpha\big) \;=\; \delta_{\alpha\beta} \;.$$

Portanto,

$$\mathrm{Tr}(A)\mathbb{1} \;=\; \sum_{\alpha=1}^{n^2}\sum_{\beta=1}^{n^2} \delta_{\alpha\beta}\big(F^\alpha\big)^* A F^\beta \;=\; \sum_{\alpha=1}^{n^2} \big(F^\alpha\big)^* A F^\alpha \;,$$

como queríamos provar.

Segue de (11.50), tomando-se $A = \mathbb{1}$, que

$$\sum_{\alpha=1}^{n^2} \big(F^\alpha\big)^* F^\alpha \;=\; n\mathbb{1} \;. \tag{11.53}$$

De (11.50) vamos extrair uma importante conclusão sobre a forma geral de aplicações lineares de $\mathrm{Mat}(\mathbb{C},\ n)$ em si mesmo.

• A forma geral de elementos de $\mathcal{L}\big(\mathrm{Mat}(\mathbb{C},\ n)\big)$

Afirmamos que se $\big\{F^\alpha,\ \alpha \in \{1,\ \ldots,\ n^2\}\big\}$ é uma base ortonormal em $\mathrm{Mat}(\mathbb{C},\ n)$, então L pode ser escrita na forma

$$\mathsf{L}(A) \;=\; \sum_{\alpha=1}^{n^2}\sum_{\beta=1}^{n^2} \ell_{\alpha\beta}\ \big(F^\alpha\big)^* A F^\beta \;, \qquad \forall A \in \mathrm{Mat}(\mathbb{C},\ n) \;, \tag{11.54}$$

para certas constantes $\ell_{\alpha\beta} \in \mathbb{C}$ independentes de A. Demonstremos essa afirmação. Como as matrizes F^α compõem uma base ortonormal e $\mathsf{L}(A) \in \mathrm{Mat}(\mathbb{C},\ n)$, podemos escrever

$$A \;=\; \sum_{\alpha=1}^{n^2} \mathrm{Tr}\Big(\big(F^\alpha\big)^* A\Big) F^\alpha \;.$$

Logo,

$$\mathsf{L}(A) \;=\; \sum_{\alpha=1}^{n^2} \mathrm{Tr}\Big(\big(F^\alpha\big)^* A\Big)\mathsf{L}(F^\alpha) \;=\; \sum_{\alpha=1}^{n^2} \mathsf{L}(F^\alpha)\mathrm{Tr}\Big(\big(F^\alpha\big)^* A\Big) \;\overset{(11.50)}{=}\; \sum_{\alpha=1}^{n^2}\sum_{\gamma=1}^{n^2} \mathsf{L}(F^\alpha)\big(F^\gamma\big)^*\big(F^\alpha\big)^* A F^\gamma \;=\; \sum_{\gamma=1}^{n^2} M^\gamma A F^\gamma \;,$$

onde

$$M^\gamma \;\equiv\; \sum_{\alpha=1}^{n^2} \mathsf{L}(F^\alpha)\big(F^\gamma\big)^*\big(F^\alpha\big)^* \;.$$

Como $M^\gamma \in \mathrm{Mat}(\mathbb{C},\ n)$, podemos escrever $M^\gamma = \sum_{\beta=1}^{n^2} \ell_{\beta\gamma}\ \big(F^\beta\big)^*$ (já que $\big\{\big(F^\beta\big)^*,\ \beta \in \{1,\ \ldots,\ n^2\}\big\}$ é também uma base ortonormal em $\mathrm{Mat}(\mathbb{C},\ n)$) com

$$\ell_{\beta\gamma} \;=\; \sum_{\alpha=1}^{n^2} \mathrm{Tr}\Big(\mathsf{L}(F^\alpha)\big(F^\gamma\big)^*\big(F^\alpha\big)^* F^\beta\Big) \;. \tag{11.55}$$

Portanto,

$$\mathsf{L}(A) \;=\; \sum_{\beta=1}^{n^2}\sum_{\gamma=1}^{n^2} \ell_{\beta\gamma}\ \big(F^\beta\big)^* A F^\gamma \;,$$

como queríamos mostrar.

• A forma geral de elementos de $\mathcal{L}\big(\mathrm{Mat}(\mathbb{C},\ n)\big)$. Uma segunda abordagem

Há uma segunda demonstração da forma geral (11.54), a qual é, talvez, mais elegante e instrutiva. Para $\alpha,\ \beta \in \{1,\ \ldots,\ n^2\}$, seja $\mathsf{T}^{\alpha\beta} \in \mathcal{L}\big(\mathrm{Mat}(\mathbb{C},\ n)\big)$ definido por

$$\mathsf{T}^{\alpha\beta}(A) \;:=\; \big(F^\alpha\big)^* A F^\beta$$

para toda $A \in \mathrm{Mat}(\mathbb{C},\ n)$. Vamos mostrar que a coleção $\big\{\mathsf{T}^{\alpha\beta},\ \alpha,\ \beta \in \{1,\ \ldots,\ n^2\}\big\}$ é ortonormal em relação ao produto escalar (11.47). De fato,

$$\begin{aligned}
\big\langle \mathsf{T}^{\alpha\beta},\ \mathsf{T}^{\gamma\delta}\big\rangle \;:=\; \sum_{\omega=1}^{n^2} \mathrm{Tr}\Big(\mathsf{T}^{\alpha\beta}(F^\omega)^*\ \mathsf{T}^{\gamma\delta}(F^\omega)\Big) \;&=\; \sum_{\omega=1}^{n^2} \mathrm{Tr}\Big(\big(F^\beta\big)^*(F^\omega)^* F^\alpha \big(F^\gamma\big)^* F^\omega F^\delta\Big) \\
&=\; \mathrm{Tr}\left(\big(F^\beta\big)^*\left(\sum_{\omega=1}^{n^2}(F^\omega)^* F^\alpha\big(F^\gamma\big)^* F^\omega\right) F^\delta\right) \;\overset{(11.50)}{=}\; \mathrm{Tr}\Big(\big(F^\beta\big)^*\ \mathrm{Tr}\big(F^\alpha\big(F^\gamma\big)^*\big)\ F^\delta\Big) \\
&=\; \mathrm{Tr}\Big(\big(F^\beta\big)^* F^\delta\Big)\ \mathrm{Tr}\Big(F^\alpha\big(F^\gamma\big)^*\Big) \;=\; \delta_{\alpha\gamma}\delta_{\beta\delta} \;,
\end{aligned}$$

como desejávamos estabelecer.

Com isso, vemos que $\big\{\mathsf{T}^{\alpha\beta},\ \alpha,\ \beta \in \{1,\ \ldots,\ n^2\}\big\}$ é uma base ortonormal em $\mathcal{L}\big(\mathrm{Mat}(\mathbb{C},\ n)\big)$ e que todo elemento $\mathsf{L} \in \mathcal{L}\big(\mathrm{Mat}(\mathbb{C},\ n)\big)$ pode ser univocamente escrito na forma

$$\mathsf{L} \;=\; \sum_{\beta=1}^{n^2}\sum_{\gamma=1}^{n^2} \ell_{\beta\gamma}\ \mathsf{T}^{\beta\gamma} \;, \tag{11.56}$$

para certas constantes $\ell_{\beta\gamma} \in \mathbb{C}$, as quais são dadas por

$$\ell_{\beta\gamma} \;=\; \big\langle \mathsf{T}^{\beta\gamma},\ \mathsf{L}\big\rangle \;=\; \sum_{\alpha=1}^{n^2} \mathrm{Tr}\Big(\mathsf{T}^{\beta\gamma}(F^\alpha)^*\ \mathsf{L}(F^\alpha)\Big) \;=\; \sum_{\alpha=1}^{n^2} \mathrm{Tr}\Big((F^\gamma)^*(F^\alpha)^* F^\beta\ \mathsf{L}(F^\alpha)\Big) \;,$$

tal como em (11.55). A relação (11.56) é precisamente (11.54). Da independência dos $\mathsf{T}^{\beta\gamma}$ vemos que a representação (11.56) e (11.54) determina L univocamente.

• Operações Lineares em $\mathrm{Mat}\,(\mathbb{C},\ n)$ que preservam autoadjunticidade

Importante no contexto da Física Quântica é a identificação de quais elementos de $\mathcal{L}\big(\mathrm{Mat}\,(\mathbb{C},\ n)\big)$ levam matrizes autoadjuntas em matrizes autoadjuntas. O resultado a seguir fornece a resposta a essa questão.

Proposição 11.14 *Uma aplicação $\mathsf{L}\in\mathcal{L}\big(\mathrm{Mat}\,(\mathbb{C},\ n)\big)$ leva matrizes autoadjuntas em matrizes autoadjuntas se e somente se satisfizer $\mathsf{L}(A)^*=\mathsf{L}(A^*)$ para toda $A\in\mathrm{Mat}\,(\mathbb{C},\ n)$.*

Se $\mathsf{L}\in\mathcal{L}\big(\mathrm{Mat}\,(\mathbb{C},\ n)\big)$ for escrita na forma geral (11.54) ou (11.56),

$$\mathsf{L}\ =\ \sum_{\beta=1}^{n^2}\sum_{\gamma=1}^{n^2}\ell_{\beta\gamma}\,\mathsf{T}^{\beta\gamma}\ , \tag{11.57}$$

uma condição necessária e suficiente para que tenhamos $\mathsf{L}(A)^*=\mathsf{L}(A^*)$ para toda $A\in\mathrm{Mat}\,(\mathbb{C},\ n)$ é que valha $\ell_{\beta\gamma}=\overline{\ell_{\gamma\beta}}$ para todos $\beta,\ \gamma\in\{1,\ \ldots,\ n^2\}$. Por fim, uma condição necessária e suficiente para que isso se dê é que existam constantes reais $d_\alpha\in\mathbb{R}$, $\alpha\in\{1,\ \ldots,\ n^2\}$ e matrizes $M^\alpha\in\mathrm{Mat}\,(\mathbb{C},\ n)$, $\alpha\in\{1,\ \ldots,\ n^2\}$ tais que

$$\mathsf{L}(A)\ =\ \sum_{\alpha=1}^{n^2}d_\alpha\,\big(M^\alpha\big)^*A\,M^\alpha\ , \tag{11.58}$$

para toda $A\in\mathrm{Mat}\,(\mathbb{C},\ n)$. As matrizes M^α podem ser escolhidas ortonormais:

$$\mathrm{Tr}\big((M^\alpha)^*M^\gamma\big)\ =\ \delta_{\alpha\gamma}\ ,$$

para toodos $\alpha,\ \gamma\in\{1,\ \ldots,\ n^2\}$. □

Em outras palavras, essa proposição estabelece que $\mathsf{L}\in\mathcal{L}\big(\mathrm{Mat}\,(\mathbb{C},\ n)\big)$ preserva a propriedade de autoadjunticidade de matrizes de $\mathrm{Mat}\,(\mathbb{C},\ n)$ se e somente se existir uma base ortonormal $M^\alpha\in\mathrm{Mat}\,(\mathbb{C},\ n)$, $\alpha\in\{1,\ \ldots,\ n^2\}$ e números reais $d_\alpha\in\mathbb{R}$, $\alpha\in\{1,\ \ldots,\ n^2\}$ tais que $\mathsf{L}(A)=\sum_{\alpha=1}^{n^2}d_\alpha\,\big(M^\alpha\big)^*A\,M^\alpha$ para toda $A\in\mathrm{Mat}\,(\mathbb{C},\ n)$.

Prova da Proposição 11.14. Seja $\mathsf{L}\in\mathcal{L}\big(\mathrm{Mat}\,(\mathbb{C},\ n)\big)$ dotada da propriedade que $L(B)^*=L(B)$ para toda $B\in\mathrm{Mat}\,(\mathbb{C},\ n)$ satisfazendo $B^*=B$. Se $A\in\mathrm{Mat}\,(\mathbb{C},\ n)$, podemos escrever $A=\mathrm{Re}\,(A)+i\mathrm{Im}\,(A)$ com $\mathrm{Re}\,(A)$ e $\mathrm{Im}\,(A)$ sendo as matrizes autoadjuntas definidas por $\mathrm{Re}\,(A):=\frac{1}{2}(A+A^*)$ e $\mathrm{Im}\,(A):=\frac{1}{2i}(A-A^*)$. Teremos, $\mathsf{L}(A)=\mathsf{L}\big(\mathrm{Re}\,(A)\big)+i\mathsf{L}\big(\mathrm{Im}\,(A)\big)$. Logo, como $\mathsf{L}\big(\mathrm{Re}\,(A)\big)$ e $\mathsf{L}\big(\mathrm{Im}\,(A)\big)$ são, por hipótese, autoadjuntas, segue que $\mathsf{L}(A)^*=\mathsf{L}\big(\mathrm{Re}\,(A)\big)-i\mathsf{L}\big(\mathrm{Im}\,(A)\big)=\mathsf{L}\big(\mathrm{Re}\,(A)-i\mathrm{Im}\,(A)\big)=\mathsf{L}(A^*)$, como desejavamos constatar.

Vamos agora supor, reciprocamente, que $\mathsf{L}(A)^*=\mathsf{L}(A^*)$ para toda $A\in\mathrm{Mat}\,(\mathbb{C},\ n)$. Se $B\in\mathrm{Mat}\,(\mathbb{C},\ n)$ satisfaz $B^*=B$, teremos $\mathsf{L}(B)^*=\mathsf{L}(B^*)=\mathsf{L}(B)$, provando que $\mathsf{L}(B)$ é autoadjunta.

Se $\mathsf{L}\in\mathcal{L}\big(\mathrm{Mat}\,(\mathbb{C},\ n)\big)$ satisfaz $\mathsf{L}(A)^*=\mathsf{L}(A^*)$, ou seja, $\mathsf{L}(A)=\mathsf{L}\big(A^*\big)^*$, para toda $A\in\mathrm{Mat}\,(\mathbb{C},\ n)$, temos, pela fórmula geral (11.54), que

$$\sum_{\alpha=1}^{n^2}\sum_{\beta=1}^{n^2}\ell_{\alpha\beta}\,\big(F^\alpha\big)^*AF^\beta\ =\ \mathsf{L}(A)\ =\ \mathsf{L}\big(A^*\big)^*\ =\ \sum_{\alpha=1}^{n^2}\sum_{\beta=1}^{n^2}\overline{\ell_{\alpha\beta}}\,\big(F^\beta\big)^*AF^\alpha\ ,$$

ou seja,

$$\sum_{\beta=1}^{n^2}\sum_{\gamma=1}^{n^2}\ell_{\beta\gamma}\,\mathsf{T}^{\beta\gamma}\ =\ \sum_{\beta=1}^{n^2}\sum_{\gamma=1}^{n^2}\overline{\ell_{\gamma\beta}}\,\mathsf{T}^{\beta\gamma}\ .$$

Pela unicidade da representação (11.56), concluímos que $\ell_{\beta\gamma}=\overline{\ell_{\gamma\beta}}$ para todos $\beta,\ \gamma\in\{1,\ \ldots,\ n^2\}$. A recíproca é evidente.

Se $\mathsf{L}\in\mathcal{L}\big(\mathrm{Mat}\,(\mathbb{C},\ n)\big)$ é da forma (11.58), é evidente que $\mathsf{L}(A)^*=\mathsf{L}(A^*)$ para todaa $A\in\mathrm{Mat}\,(\mathbb{C},\ n)$. Seja agora $\mathsf{L}\in\mathcal{L}\big(\mathrm{Mat}\,(\mathbb{C},\ n)\big)$ da forma geral (11.54) ou (11.56), com $\ell_{\beta\gamma}=\overline{\ell_{\gamma\beta}}$ para todos $\beta,\ \gamma\in\{1,\ \ldots,\ n^2\}$. Isso diz-nos

que a matriz $\ell\in\mathrm{Mat}\,(\mathbb{C},\ n^2)$, cujos elementos de matriz são $\ell_{\beta\gamma}$, é uma matriz autoadjunta e, portanto, pode ser diagonalizada por uma matriz unitária (Teorema 10.15, página 578). Assim, existe $u\in\mathrm{Mat}\,(\mathbb{C},\ n^2)$ com $\ell=u^*du$, com $d=\mathrm{diag}\,\big(d_1,\ \ldots,\ d_\alpha\big)$, sendo d_α os autovalores reais de ℓ Assim, escrevemos, para toda $A\in\mathrm{Mat}\,(\mathbb{C},\ n)$,

$$\begin{aligned}\mathsf{L}(A)\ &=\ \sum_{\beta=1}^{n^2}\sum_{\gamma=1}^{n^2}\ell_{\beta\gamma}\,\big(F^\beta\big)^*AF^\gamma\ =\ \sum_{\beta=1}^{n^2}\sum_{\gamma=1}^{n^2}\sum_{\alpha=1}^{n^2}\overline{u_{\alpha\beta}}d_\alpha u_{\alpha\gamma}\big(F^\beta\big)^*AF^\gamma\\ &=\ \sum_{\alpha=1}^{n^2}d_\alpha\left(\sum_{\beta=1}^{n^2}u_{\alpha\beta}F^\beta\right)^*A\left(\sum_{\gamma=1}^{n^2}u_{\alpha\gamma}F^\gamma\right)\ =\ \sum_{\alpha=1}^{n^2}d_\alpha\,\big(M^\alpha\big)^*A\,M^\alpha\ ,\end{aligned}$$

onde

$$M^\alpha\ :=\ \sum_{\beta=1}^{n^2}u_{\alpha\beta}F^\beta\ ,$$

provando (11.58). Note-se que, como $u\in\mathrm{Mat}\,(\mathbb{C},\ n^2)$ é unitária, vale

$$\mathrm{Tr}\big((M^\alpha)^*M^\gamma\big)\ =\ \sum_{\beta=1}^{n^2}\sum_{\delta=1}^{n^2}\overline{u_{\alpha\beta}}\,u_{\gamma\delta}\,\underbrace{\mathrm{Tr}\big((F^\beta)^*F^\delta\big)}_{\delta_{\beta\delta}}\ =\ \sum_{\beta=1}^{n^2}u_{\gamma\beta}(u^*)_{\beta\alpha}\ =\ \delta_{\gamma\alpha}\ .$$

■

11.4.2 Alguns Exemplos Específicos de Aplicações Lineares em $\mathrm{Mat}\,(\mathbb{C},\ n)$

Nesta seção apresentaremos alguns elementos de $\mathcal{L}\big(\mathrm{Mat}\,(\mathbb{C},\ n)\big)$ dotados de interesse especial e estudaremos suas propriedades, tendo como objetivo maior a demonstração da fórmula de Baker, Campbell e Hausdorff na Seção 11.5, página 682. Alguns dos resultados que obteremos, porém, são de utilidade na Teoria de Grupos, na Mecânica Quântica e outas áreas.

• As aplicações ad

Dada uma matriz $X\in\mathrm{Mat}\,(\mathbb{C},\ n)$ fixa podemos definir uma aplicação linear $\mathbf{ad}[X]$ em $\mathrm{Mat}\,(\mathbb{C},\ n)$, $\mathbf{ad}[X]:\mathrm{Mat}\,(\mathbb{C},\ n)\to\mathrm{Mat}\,(\mathbb{C},\ n)$ por

$$\mathbf{ad}[X](A)\ :=\ [X,\ A]\ =\ XA-AX\ .$$

para toda matriz $A\in\mathrm{Mat}\,(\mathbb{C},\ n)$.

• As aplicações Ad

Analogamente, seja $G\in\mathrm{GL}(\mathbb{C},\ n)$ uma matriz inversível fixa. Podemos definir uma aplicação linear $\mathbf{Ad}[G]$ em $\mathrm{Mat}\,(\mathbb{C},\ n)$, $\mathbf{Ad}[G]:\mathrm{Mat}\,(\mathbb{C},\ n)\to\mathrm{Mat}\,(\mathbb{C},\ n)$ por

$$\mathbf{Ad}[G](A)\ :=\ GAG^{-1}\ .$$

• Definindo a exponenciação de ad

Denotaremos por $(\mathbf{ad}[X])^p$ ou $\mathbf{ad}[X]^p$ a p-ésima potência de $\mathbf{ad}[X]$:

$$\mathbf{ad}[X]^p(A)\ =\ \Big[\underbrace{X,\ \big[X,\ \ldots,\ [X,}_{p\text{ vezes}}\ A]\cdots\big]\Big]\ .$$

Aqui, $p=1,\ 2,\ \ldots$. Para facilitar a notação em aplicações futuras, convencionaremos que $\mathbf{ad}[X]^0(A)=A$ para toda matriz $A\in\mathrm{Mat}\,(\mathbb{C},\ n)$.

Dado que $\mathbf{ad}[X]$ é uma aplicação linear em um espaço vetorial de dimensão finita, sua exponencial é bem definida. Definimos $\mathrm{Exp}\,[\mathbf{ad}[X]]$ como sendo a aplicação linear no espaço das matrizes complexas $n\times n$, $\mathrm{Exp}\,[\mathbf{ad}[X]] : \mathrm{Mat}\,(\mathbb{C},\ n) \to \mathrm{Mat}\,(\mathbb{C},\ n)$ dada por

$$\begin{aligned}\mathrm{Exp}\,\big[\mathbf{ad}[X]\big](A) \ := \ \sum_{m=0}^{\infty}\frac{1}{m!}\,\big(\mathbf{ad}[X]\big)^m(A) \ &:= \ A+\sum_{m=1}^{\infty}\frac{1}{m!}\,\big(\mathbf{ad}[X]\big)^m(A)\ , \\ &= \ A+\sum_{m=1}^{\infty}\frac{1}{m!}\,\Big[\underbrace{X,\ \big[X,\ \ldots,\ [X}_{m \text{ vezes}},\ A]\cdots\big]\Big]\ ,\end{aligned}$$

para toda $A\in\mathrm{Mat}\,(\mathbb{C},\ n)$. A convergência da série é automaticamente garantida pelas observações da Seção 11.2.

• A relação entre ad e Ad

Há uma relação elegante entre as aplicações **ad** e **Ad**, a qual se expressa na seguinte proposição:

Proposição 11.15 *Seja $X\in\mathrm{Mat}\,(\mathbb{C},\ n)$ qualquer. Então*

$$\mathbf{Ad}\big[\exp(X)\big] \ = \ \mathrm{Exp}\,\big[\mathbf{ad}[X]\big]\ , \tag{11.59}$$

ou seja, para toda matriz $A\in\mathrm{Mat}\,(\mathbb{C},\ n)$ vale

$$\exp(X)A\exp(-X) \ = \ A+\sum_{m=1}^{\infty}\frac{1}{m!}\,\big(\mathbf{ad}[X]\big)^m(A)\ , \tag{11.60}$$

ou seja,

$$\begin{aligned}\exp(X)A\exp(-X) \ &= \ A+\sum_{m=1}^{\infty}\frac{1}{m!}\,\Big[\underbrace{X,\ \big[X,\ \ldots,\ [X}_{m \text{ vezes}},\ A]\cdots\big]\Big] \\ &= \ A+[X,\ A]+\frac{1}{2!}\big[X,\ [X,\ A]\big]+\frac{1}{3!}\Big[X,\ \big[X,\ [X,\ A]\big]\Big]+\cdots\ .\end{aligned} \tag{11.61}$$

□

Comentários. A expressão (11.60) ou (11.61) é comummente denominada *série de Lie*, mas alguns autores também a denominam *fórmula de Baker-Campbell-Hausdorff*. Reservaremos esse nome apenas para a expressão (11.68), adiante.

As expressões (11.60) e (11.61) são empregadas de várias formas na Mecânica Quântica, na Mecânica Estatística Quântica e na Teoria Quântica de Campos, especialmente na Teoria de Perturbações e nas Teorias de Calibre. ♣

Prova. Seja $t\in\mathbb{R}$ e sejam A e X matrizes complexas $n\times n$ fixas quaisquer. Definamos

$$\Gamma_1(t) \ := \ \mathrm{Exp}\,\big[\mathbf{ad}[tX]\big](A) \ = \ A+\sum_{m=1}^{\infty}\frac{t^m}{m!}\,\big(\mathbf{ad}[X]\big)^m(A)$$

e

$$\Gamma_2(t) \ := \ \mathbf{Ad}\big[\exp(tX)\big](A) \ = \ \exp(tX)A\exp(-tX)\ .$$

Vamos mostrar que $\Gamma_1(t)=\Gamma_2(t)$ para todo t provando para isso que ambas satisfazem a mesma equação diferencial linear com a mesma condição inicial.

É trivial constatar que $\Gamma_1(0)=\Gamma_2(0)=A$. Pela definição tem-se

$$\begin{aligned}\frac{d}{dt}\Gamma_1(t) \ &= \ \sum_{m=1}^{\infty}\frac{t^{m-1}}{(m-1)!}\,\big(\mathbf{ad}[X]\big)^m(A) \\ &= \ \mathbf{ad}[X]\left(\sum_{m=1}^{\infty}\frac{t^{m-1}}{(m-1)!}\,\big(\mathbf{ad}[X]\big)^{m-1}(A)\right) \\ &= \ \mathbf{ad}[X]\left(\sum_{m=0}^{\infty}\frac{t^{m}}{m!}\,\big(\mathbf{ad}[X]\big)^{m}(A)\right) \\ &= \ \mathbf{ad}[X]\,\big(\mathrm{Exp}\,\big[\mathbf{ad}[tX]\big](A)\big) \\ &= \ \mathbf{ad}[X]\big(\Gamma_1(t)\big)\ .\end{aligned}$$

Em resumo, $\Gamma_1(t)$ satisfaz

$$\frac{d}{dt}\Gamma_1(t) \ = \ \mathbf{ad}[X]\big(\Gamma_1(t)\big)\ .$$

Analogamente, calculemos $\frac{d}{dt}\Gamma_2(t)$. Aplicando a regra de Leibniz[10],

$$\begin{aligned}\frac{d}{dt}\Gamma_2(t) \ &= \ \frac{d}{dt}\,(\exp(tX)A\exp(-tX)) \\ &= \ X\exp(tX)A\exp(-tX)-\exp(tX)A\exp(-tX)X \\ &= \ \mathbf{ad}[X]\big(\exp(tX)A\exp(-tX)\big) \\ &= \ \mathbf{ad}[X]\big(\Gamma_2(t)\big)\ .\end{aligned}$$

Em resumo, $\Gamma_2(t)$ satisfaz

$$\frac{d}{dt}\Gamma_2(t) \ = \ \mathbf{ad}[X]\big(\Gamma_2(t)\big)\ .$$

Constatamos assim que $\Gamma_1(t)$ e $\Gamma_2(t)$ satisfazem a mesma equação diferencial com a mesma condição inicial. Pelo Teorema de existência e unicidade de soluções de sistemas de equações diferenciais lineares com coeficientes constantes discutido na Seção 14.2, isso implica que $\Gamma_1(t)=\Gamma_2(t)$ para todo $t\in\mathbb{R}$ e, em particular para $t=1$, que é a afirmação do teorema. ■

Comentário. O teorema acima e sua demonstração exemplificam uma situação não muito incomum, onde apresenta-se um resultado que é muito difícil de ser provado por um procedimento mas muito fácil de ser demonstrado por outro. Tente o leitor demonstrar a identidade (11.60) expandindo as exponenciais do lado direito em suas séries de Taylor, ou seja, escrevendo

$$\exp(X)A\exp(-X) \ = \ \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\frac{(-1)^l}{k!l!}X^kAX^l$$

e reordenando as somas de modo a obter o lado esquerdo de (11.60)! Ainda que seja possível provar (11.60) dessa forma, um tal procedimento é muitíssimo mais complexo que aquele que empregamos, e que faz apenas uso de um fato básico bem conhecido da teoria das equações diferenciais. ♣

E. 11.19 *Exercício*. Tenha a ideia certa antes de tentar resolver qualquer problema. ✠

• A aplicação diferencial exponencial dexp

Seja $F(t)$ uma matriz complexa $n\times n$ cujos elementos de matriz $(F(t))_{ij}$ são funções diferenciáveis em relação a t. Seja também $F'(t)$ a matriz cujo elemento ij é $\frac{d}{dt}(F(t))_{ij}$. Em palavras, $F'(t)$ é obtida diferenciando cada elemento de matriz de $F(t)$.

[10] Gottfried Wilhelm von Leibniz (1646–1716).

Vamos nos colocar o seguinte problema: como calcular $\frac{d}{dt}\exp(F(t))$? O estudante apressado poderia imaginar que $\frac{d}{dt}\exp(F(t)) = \exp(F(t))F'(t)$. Isso é, todavia, em geral falso, pois essa regra de derivação não vale para matrizes! Isso é assim, pois a matriz $F'(t)$ não necessariamente comuta com a matriz $F(t)$. Tem-se, em verdade, que para todo $m = 1, 2, 3, \ldots$,

$$\frac{d}{dt}(F(t))^m \;=\; \frac{d}{dt}\left(\underbrace{F(t)\cdots F(t)}_{m \text{ vezes}}\right) \;=\; \sum_{k=0}^{m-1} F(t)^k F'(t) F(t)^{m-k-1} \;.$$

Consequentemente,

$$\frac{d}{dt}\exp\big(F(t)\big) \;=\; \sum_{n=1}^{\infty}\sum_{k=0}^{n-1}\frac{1}{n!}F(t)^k F'(t) F(t)^{n-k-1} \;. \tag{11.62}$$

Isso motiva a seguinte definição. Para $X \in \mathrm{Mat}(\mathbb{C}, n)$ fixo, definimos uma aplicação linear $\mathbf{dexp}[X] : \mathrm{Mat}(\mathbb{C}, n) \to \mathrm{Mat}(\mathbb{C}, n)$, denominada *aplicação diferencial exponencial*, por

$$\mathbf{dexp}[X](A) \;:=\; \sum_{n=1}^{\infty}\sum_{k=0}^{n-1}\frac{1}{n!}X^k A X^{n-k-1} \;, \tag{11.63}$$

para todo $A \in \mathrm{Mat}(\mathbb{C}, n)$.

E. 11.20 *Exercício*. Mostre que a série do lado direito está bem definida, ou seja, que é convergente para todos X e A. ✠

Com essa definição podemos, por (11.62), escrever

$$\frac{d}{dt}\exp\big(F(t)\big) \;=\; \mathbf{dexp}\big[F(t)\big]\big(F'(t)\big) \;. \tag{11.64}$$

Para uma expressão alternativa para a derivada da exponencial de uma matriz dependente de um parâmetro, vide equação (11.86), página 687.

Por razões que ficarão claras adiante quando provarmos a fórmula de Baker, Campbell e Hausdorff, é conveniente expressar $\mathbf{dexp}[X]$ em termos de $\mathbf{ad}[X]$. Como veremos, é possível fazer isso e o resultado está expresso na Proposição 11.16 que apresentaremos e demonstraremos a seguir.

Antes, porém, duas definições. Para $z \in \mathbb{C}$ definimos a função complexa $\phi(z)$ por

$$\phi(z) \;:=\; \frac{1-e^{-z}}{z} \;=\; \sum_{m=0}^{\infty}\frac{(-1)^m}{(m+1)!}z^m \;. \tag{11.65}$$

Como a série de Taylor do lado direito converge para todo $z \in \mathbb{C}$, $\phi(z)$ é uma função inteira, ou seja, é analítica em toda parte.

Pelos nossos comentários da Seção 11.2, podemos definir para todo $X \in \mathrm{Mat}(\mathbb{C}, n)$ uma aplicação linear $\Phi[X] : \mathrm{Mat}(\mathbb{C}, n) \to \mathrm{Mat}(\mathbb{C}, n)$ dada por

$$\Phi[X] \;:=\; \phi(\mathbf{ad}[X]) \;, \tag{11.66}$$

ou seja, $\Phi[X]$ é a aplicação que a todo $A \in \mathrm{Mat}(\mathbb{C}, n)$ associa a matriz $\Phi[X](A)$ dada por

$$\Phi[X](A) \;=\; \sum_{m=0}^{\infty}\frac{(-1)^m}{(m+1)!}\mathbf{ad}[X]^m(A) \;. \tag{11.67}$$

Pelos comentários da Seção 11.2 a série do lado direito converge para todos $X, A \in \mathrm{Mat}(\mathbb{C}, n)$.

Proposição 11.16 *Com as definições apresentadas acima, vale para todos $A, X \in \mathrm{Mat}(\mathbb{C}, n)$ a expressão*

$$\mathbf{dexp}[X](A) \;=\; \exp(X)\;\Phi[\mathbf{ad}[X]](A) \;,$$

ou seja,

$$\mathbf{dexp}[X](A) \;=\; \exp(X)\left(\sum_{m=0}^{\infty}\frac{(-1)^m}{(m+1)!}\mathbf{ad}[X]^m(A)\right) \;.$$

□

Também como comentado acima, é inútil tentar provar a proposição partindo de (11.63) e aplicando força-bruta. A demonstração usará uma série de truques elegantes.

Prova. Vamos definir, para $A, X \in \mathrm{Mat}(\mathbb{C}, n)$ fixas e $t \in \mathbb{R}$,

$$H(t) \;:=\; t\,\mathbf{dexp}[tX](A) \;.$$

A ideia é descobrir uma equação diferencial que $H(t)$ satisfaz e, em seguida, resolvê-la. Note-se que, pela definição, $H(0) = 0$. Como veremos, resolver a equação diferencial é tarefa relativamente fácil. Um pouco mais trabalhoso é encontrar a equação diferencial. Para isso temos que calcular a derivada de $H(t)$ em relação a t.

Pela definição de $H(t)$ e de $\mathbf{dexp}[tX](A)$ em (11.63), tem-se

$$\begin{aligned}
\frac{d}{dt}H(t) &= \frac{d}{dt}\big(t\,\mathbf{dexp}[tX](A)\big) \;=\; \frac{d}{dt}\left(\sum_{n=1}^{\infty}\sum_{k=0}^{n-1}\frac{t^n}{n!}X^k A X^{n-k-1}\right)\\
&= \sum_{n=1}^{\infty}\sum_{k=0}^{n-1}\frac{t^{n-1}}{(n-1)!}X^k A X^{n-k-1} \;=\; \sum_{n=0}^{\infty}\sum_{k=0}^{n}\frac{t^n}{n!}X^k A X^{n-k}\\
&= A + \sum_{n=1}^{\infty}\sum_{k=0}^{n}\frac{t^n}{n!}X^k A X^{n-k} \;=\; A + \sum_{n=1}^{\infty}\frac{t^n}{n!}AX^n + \sum_{n=1}^{\infty}\sum_{k=1}^{n}\frac{t^n}{n!}X^k A X^{n-k}\\
&= A\left(\mathbb{1} + \sum_{n=1}^{\infty}\frac{t^n}{n!}X^n\right) + \sum_{n=1}^{\infty}\sum_{k=1}^{n}\frac{t^n}{n!}X^k A X^{n-k} \;=\; A\exp(tX) + \sum_{n=1}^{\infty}\sum_{k=1}^{n}\frac{t^n}{n!}X^k A X^{n-k}\\
&= A\exp(tX) + tX\left(\sum_{n=1}^{\infty}\sum_{k=1}^{n}\frac{t^{n-1}}{n!}X^{k-1} A X^{n-k}\right)\\
&= A\exp(tX) + tX\left(\sum_{n=1}^{\infty}\sum_{k=0}^{n-1}\frac{t^{n-1}}{n!}X^{k} A X^{n-k-1}\right)\\
&= A\exp(tX) + X\big(t\,\mathbf{dexp}[tX](A)\big) \;=\; A\exp(tX) + XH(t) \;.
\end{aligned}$$

Em resumo, $H(t)$ satisfaz a equação diferencial

$$\frac{d}{dt}H(t) \;=\; XH(t) + A\exp(tX) \;,$$

com a condição inicial $H(0) = 0$.

Como estudamos à página 751 da Seção 14.2.2, a solução geral da equação matricial

$$\frac{d}{dt}\mathcal{M}(t) \;=\; X\mathcal{M}(t) + \mathcal{G}(t) \qquad \text{é} \qquad \mathcal{M}(t) \;=\; \exp(tX)\mathcal{M}(0) + \int_0^t \exp\big((t-s)X\big)\,\mathcal{G}(s)ds \;.$$

Assim, como $H(0) = 0$ e $\mathcal{G}(t) = A\exp(tX)$, teremos

$$
\begin{aligned}
H(t) &= \int_0^t \exp\big((t-s)X\big) A \exp(sX)\, ds \\
&= \exp(tX)\int_0^t \exp(-sX) A \exp(sX)\, ds \;=\; \exp(tX)\int_0^t \mathbf{Ad}\big[\exp(-sX)\big](A)\, ds \\
&\overset{(11.59)}{=} \exp(tX)\int_0^t \mathrm{Exp}\big[-\mathbf{ad}[sX]\big](A)\, ds \;=\; \exp(tX)\int_0^t \sum_{m=0}^{\infty}\frac{(-s)^m}{m!}\mathbf{ad}[X]^m(A)\, ds \\
&= \exp(tX)\sum_{m=0}^{\infty}\frac{(-1)^m}{m!}\mathbf{ad}[X]^m(A)\int_0^t s^m\, ds \;=\; \exp(tX)\sum_{m=0}^{\infty}\frac{(-1)^m t^{m+1}}{(m+1)!}\mathbf{ad}[X]^m(A) \\
&= t\exp(tX)\sum_{m=0}^{\infty}\frac{(-1)^m t^m}{(m+1)!}\mathbf{ad}[X]^m(A) \\
&\overset{(11.67)}{=} t\,\exp(tX)\,\Phi[tX](A)\,.
\end{aligned}
$$

Essa expressão vale para todo $t \in \mathbb{R}$. Tomando $t = 1$, teremos $H(1) = \exp(X)\Phi[X](A)$, ou seja,

$$\mathbf{dexp}\,[X](A) \;=\; \exp(X)\;\Phi[X](A)\;,$$

que é o que queríamos provar. ■

Reunindo todos esses resultados, estamos agora preparados para provar a fórmula de Baker, Campbell e Hausdorff.

11.5 A Fórmula de Baker, Campbell e Hausdorff

A presente seção é dedicada à demonstração da célebre *Fórmula de Baker-Campbell-Hausdorff*. Seguiremos com diversas modificações o tratamento de [244]-[245]. O resultado principal que desejamos provar encontra-se expresso no seguinte teorema:

Teorema 11.1 (Fórmula de Baker-Campbell-Hausdorff) *Para $A,\ B \in \mathrm{Mat}\,(\mathbb{C},\ n)$ tais que $\|A\|_{\mathbb{C}}$ e $\|B\|_{\mathbb{C}}$ sejam ambas menores que $\frac{1}{2}\ln\left(2-\frac{\sqrt{2}}{2}\right) \approx 0,12844\ldots$, vale*

$$\exp(A)\exp(B) \;=\; \exp(A*B)\;,$$

com

$$A*B \;:=\; A+B+\sum_{\substack{k,\ l\geq 0\\ k+l>0}}\ \sum_{\substack{a_1,\ b_1\geq 0\\ a_1+b_1>0}}\cdots\sum_{\substack{a_k,\ b_k\geq 0\\ a_k+b_k>0}}\ \frac{(-1)^k}{l!(k+1)(b_1+\cdots+b_k+1)}\left(\prod_{i=1}^{k}\frac{1}{a_i!b_i!}\right)$$
$$\times\,\mathbf{ad}[A]^{a_1}\mathbf{ad}[B]^{b_1}\cdots\mathbf{ad}[A]^{a_k}\mathbf{ad}[B]^{b_k}\mathbf{ad}[A]^{l}(B)\,. \tag{11.68}$$

Os primeiros termos de (11.68) são

$$A*B \;=\; A+B+\frac{1}{2}[A,\ B]+\frac{1}{12}\big[A,\ [A,\ B]\big]+\frac{1}{12}\big[B,\ [B,\ A]\big]+\cdots\,. \tag{11.69}$$

□

<u>*Comentário*</u>. A expressão (11.68) é a célebre fórmula de Baker[11], Campbell[12] e Hausdorff[13], que desempenha um papel importante no estudo de grupos de Lie e outras áreas. Advertimos que, devido à sua complexidade e devido à restrição quanto à norma das matrizes A e B, a fórmula de Baker-Campbell-Hausdorff tem um escopo de aplicações relativamente limitado no que concerne a cômputos de produtos de exponenciais. A mesma fórmula, porém, presta-se à demonstração de vários teoremas, especialmente na teoria dos grupos de Lie. Uma situação interessante na qual a fórmula de Baker-Campbell-Hausdorff pode ser empregada é aquela na qual comutadores de ordem suficientemente grande das matrizes A e B se anulam, pois aí o lado direito de (11.68) ou (11.69) tem um número <u>finito</u> de termos (e, portanto, não há nesse caso restrição quanto à norma das matrizes). Tal ocorre nas chamadas *álgebras de Lie nilpotentes*. O leitor que procura um exemplo simples do uso de (11.69) pode interessar-se em ler sobre o chamado *grupo de Heisenberg* na Seção 21.2.2, página 1126. ♣

Prova do Teorema 11.1. A estratégia que empregaremos para provar a fórmula de Baker, Campbell e Hausdorff é muito semelhante àquela empregada na demonstração da Proposição 11.16. Seja, para $A,\ B \in \mathrm{Mat}\,(\mathbb{C},\ n)$ fixas tais que $\|A\|_{\mathbb{C}} < \ln(2)/2$ e $\|B\|_{\mathbb{C}} < \ln(2)/2$, a matriz[14]

$$G(t) \;:=\; \ln\Big(\exp(A)\exp(tB)\Big)\;, \tag{11.70}$$

para $t \in [-1,\ 1]$. Vamos identificar uma equação diferencial satisfeita por $G(t)$ e, em seguida, resolvê-la.

Comecemos procurando calcular a derivada de $G(t)$ em relação a t. Isso é uma tarefa mais difícil do que parece e procederemos de modo indireto. É conveniente calcular primeiro a derivada de $\exp(G(t))$. Por um lado, temos que

$$\exp(G(t)) \;=\; \exp(A)\exp(tB)$$

e, portanto,

$$\frac{d}{dt}\exp(G(t)) \;=\; \exp(A)\frac{d}{dt}\exp(tB) \;=\; \exp(A)\exp(tB)B\,.$$

Por outro tem-se, pela definição da aplicação **dexp**, que

$$\frac{d}{dt}\exp(G(t)) \;=\; \mathbf{dexp}\,\big[G(t)\big]\big(G'(t)\big)\,.$$

Portanto,

$$\mathbf{dexp}\,\big[G(t)\big]\big(G'(t)\big) \;=\; \exp(A)\exp(tB)B\,.$$

Usando a Proposição 11.16, página 680, essa última igualdade pode ser escrita como

$$\exp\big(G(t)\big)\,\Phi\big[G(t)\big]\big(G'(t)\big) \;=\; \exp(A)\exp(tB)B\;,$$

o que implica que

$$\Phi\big[G(t)\big]\big(G'(t)\big) \;=\; \exp(-G(t))\exp(A)\exp(tB)B \;=\; \exp(-tB)\exp(-A)\exp(A)\exp(tB)B \;=\; B\,.$$

Resumindo, tem-se

$$\Phi\big[G(t)\big]\big(G'(t)\big) \;=\; B\,. \tag{11.71}$$

A ideia que agora perseguiremos é tentar inverter essa expressão de modo a obter $G'(t)$ (que aparece no argumento de Φ no lado esquerdo). Para isso faremos uso do seguinte lema:

Lema 11.2 *Sejam as funções complexas*

$$\phi(z) \;:=\; \frac{1-e^{-z}}{z}\;, \qquad z \in \mathbb{C}\;,$$

e

$$\psi(z) \;:=\; \frac{z\ln(z)}{z-1}, \qquad |z-1| < 1\;,$$

sendo que a primeira já fora definida em (11.65). Então vale

$$\psi(e^z)\phi(z) \;=\; 1 \tag{11.72}$$

para todo $z \in \mathbb{C}$ tal que $|z| < \ln 2$. □

[11] Henry Frederick Baker (1866–1956).
[12] John Edward Campbell (1862–1924).
[13] Felix Hausdorff (1868–1942).
[14] A condição $\|A\|_{\mathbb{C}} < \ln(2)/2$ e $\|B\|_{\mathbb{C}} < \ln(2)/2$ garante que $\|\exp(A)\exp(tB) - \mathbb{1}\|_{\mathbb{C}} < 1$ para todo $t \in [-1,\ 1]$. Assim, o logaritmo de $\exp(A)\exp(tB)$ em (11.70) está definido.

Prova. Usando a expansão em série de Taylor da função ln, podemos escrever

$$\psi(z) \;:=\; z\frac{\ln(z)}{z-1} \;=\; z\frac{\ln(1+(z-1))}{z-1} \;=\; z\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}\,(z-1)^{k-1}\;, \tag{11.73}$$

mostrando que $\psi(z)$ é analítica na região $|z-1|<1$. Agora, se $|z|<\ln 2$, tem-se $|e^z-1|<1$, pois $e^z-1=\sum_{m=1}^{\infty}\frac{1}{m!}z^m$ e

$$|e^z-1| \;\leq\; \sum_{m=1}^{\infty}\frac{1}{m!}|z|^m \;<\; \sum_{m=1}^{\infty}\frac{1}{m!}(\ln 2)^m \;=\; e^{\ln 2}-1 \;=\; 1\;.$$

Assim, e^z está dentro da região onde ψ é analítica, onde vale que

$$\psi(e^z)\phi(z) \;=\; \left(\frac{e^z z}{e^z-1}\right)\left(\frac{1-e^{-z}}{z}\right) \;=\; 1\;,$$

que é o que queríamos provar. ■

O uso que faremos desse lema é o seguinte. Seja $X\in\mathrm{Mat}(\mathbb{C},\ n)$ qualquer. Por analogia com a definição de $\Phi[X]$ em (11.66), definimos

$$\Psi[X] \;:=\; \psi\Big(\mathrm{Exp}\,\big[\mathbf{ad}[X]\big]\Big) \;=\; \psi\Big(\mathbf{Ad}\big[\exp(X)\big]\Big)\;.$$

Assim, se $\|\mathbf{ad}[X]\|<\ln 2$ (para ficarmos no domínio de validade de (11.72)), teremos

$$\Psi[X]\Phi[X] \;:=\; \psi\Big(\mathrm{Exp}\,\big[\mathbf{ad}[X]\big]\Big)\phi\big(\mathbf{ad}[X]\big) \;=\; \mathbf{id}\;,$$

onde $\mathbf{id}$ é a aplicação identidade: $\mathbf{id}\,(A):=A$, para toda $A\in\mathrm{Mat}(\mathbb{C},\ n)$. Portanto, assumindo que $\|\mathbf{ad}[G(t)]\|<\ln 2$ teremos, aplicando $\Psi[G(t)]$ a (11.71),

$$G'(t) \;=\; \Psi\big[G(t)\big](B)\;. \tag{11.74}$$

Essa é a equação diferencial procurada e que é satisfeita por $G(t)$, com a condição inicial $G(0)=A$.

Note-se que para que as manipulações acima sejam válidas é necessário (para ficarmos no domínio de validade de (11.72)) que $\|\mathbf{ad}G(t)\|<\ln 2$. Afirmamos que, para tal, é suficiente ter-se

$$\|A\|_{\mathrm{C}},\ \|B\|_{\mathrm{C}} \;<\; \frac{1}{2}\ln\left(2-\frac{\sqrt{2}}{2}\right) \;<\; \frac{\ln 2}{2}\;. \tag{11.75}$$

De fato, para que se tenha $\|\mathbf{ad}G(t)\|<\ln 2$ é suficiente que $\|G(t)\|_{\mathrm{C}}<\ln(2)/2$. Se $Z(t):=\exp(A)\exp(tB)$, então $G(t)=\ln(Z(t))$ e teremos

$$\|G(t)\|_{\mathrm{C}} \;=\; \|\ln(Z(t))\|_{\mathrm{C}} \;=\; \|\ln\big(\mathbb{1}+(Z(t)-\mathbb{1})\big)\|_{\mathrm{C}} \;\leq\; \sum_{k=1}^{\infty}\frac{1}{k}\|Z(t)-\mathbb{1}\|_{\mathrm{C}}^k \;=\; \ln\left(\frac{1}{1-\|Z(t)-\mathbb{1}\|_{\mathrm{C}}}\right)\;. \tag{11.76}$$

Agora,

$$\begin{aligned}
\|Z(t)-\mathbb{1}\|_{\mathrm{C}} \;=\; \Big\|\exp(A)\exp(tB)-\mathbb{1}\Big\|_{\mathrm{C}} \;&=\; \Big\|\big(\exp(A)-\mathbb{1}\big)\big(\exp(tB)-\mathbb{1}\big)+\big(\exp(A)-\mathbb{1}\big)+\big(\exp(tB)-\mathbb{1}\big)\Big\|_{\mathrm{C}}\\
&\leq\; \|\exp(A)-\mathbb{1}\|_{\mathrm{C}}\,\|\exp(tB)-\mathbb{1}\|_{\mathrm{C}}+\|\exp(A)-\mathbb{1}\|_{\mathrm{C}}+\|\exp(tB)-\mathbb{1}\|_{\mathrm{C}}\\
&\leq\; \Big(e^{\|A\|_{\mathrm{C}}}-1\Big)\Big(e^{\|B\|_{\mathrm{C}}}-1\Big)+\Big(e^{\|A\|_{\mathrm{C}}}-1\Big)+\Big(e^{\|B\|_{\mathrm{C}}}-1\Big)\\
&\leq\; e^{\|A\|_{\mathrm{C}}+\|B\|_{\mathrm{C}}}-1 \;\overset{(11.75)}{<}\; 2-\frac{\sqrt{2}}{2}-1 \;=\; 1-\frac{\sqrt{2}}{2}\;.
\end{aligned}$$

Logo, por (11.76),

$$\|G(t)\|_{\mathrm{C}} \;<\; \ln\left(\frac{1}{1-1+\frac{\sqrt{2}}{2}}\right) \;=\; \frac{\ln 2}{2}\;,$$

como desejamos.

Para prosseguir devemos escrever (11.74) de forma mais conveniente. Pela definição da aplicação $\mathbf{Ad}$, é bem fácil ver que

$$\mathbf{Ad}\Big[e^Xe^Y\Big] \;=\; \mathbf{Ad}\Big[e^X\Big]\,\mathbf{Ad}\Big[e^Y\Big]\;.$$

E. 11.21 *Exercício*. Verifique. ✠

Assim,

$$\begin{aligned}
\Psi\big[G(t)\big] \;&=\; \psi\Big(\mathrm{Exp}\,\big[\mathbf{ad}[G(t)]\big]\Big) \;=\; \psi\Big(\mathbf{Ad}\big[\exp\big(G(t)\big)\big]\Big) \;=\; \psi\Big(\mathbf{Ad}\big[\exp(A)\exp(tB)\big]\Big)\\
&=\; \psi\Big(\mathbf{Ad}\big[\exp(A)\big]\,\mathbf{Ad}\big[\exp(tB)\big]\Big) \;=\; \psi\Big(\mathrm{Exp}\,\big[\mathbf{ad}[A]\big]\ \mathrm{Exp}\,\big[\mathbf{ad}[tB]\big]\Big)\;.
\end{aligned}$$

A equação diferencial (11.74) para $G(t)$ assume, portanto, a forma

$$G'(t) \;=\; \psi\Big(\mathrm{Exp}\,\big[\mathbf{ad}[A]\big]\ \mathrm{Exp}\,\big[\mathbf{ad}[tB]\big]\Big)(B)\;, \tag{11.77}$$

com $G(0)=A$ como condição inicial. Isto posto, nossa tarefa agora é resolver (11.77), o que pode ser feito por uma simples integração. Teremos, portanto,

$$G(t)-G(0) \;=\; \int_0^t G'(s)\,ds \;=\; \int_0^t \psi\Big(\mathrm{Exp}\,\big[\mathbf{ad}[A]\big]\ \mathrm{Exp}\,\big[\mathbf{ad}[sB]\big]\Big)(B)\,ds\;.$$

Tomando-se $t=1$ teremos

$$\ln\big(e^A\,e^B\big) \;=\; A+\int_0^1 \psi\Big(\mathrm{Exp}\,\big[\mathbf{ad}[A]\big]\ \mathrm{Exp}\,\big[\mathbf{ad}[sB]\big]\Big)(B)\,ds\;. \tag{11.78}$$

Estando já na reta final, resta-nos calcular a integral do lado direito, o que pode ser feito com o uso da expansão em série de ψ dada em (11.73) e um pouco de paciência. É o que faremos. Por (11.73), teremos

$$\begin{aligned}
&\psi\Big(\mathrm{Exp}\,\big[\mathbf{ad}[A]\big]\ \mathrm{Exp}\,\big[\mathbf{ad}[sB]\big]\Big)(B)\\
&\quad=\; \Big(\mathrm{Exp}\,\big[\mathbf{ad}[A]\big]\ \mathrm{Exp}\,\big[\mathbf{ad}[sB]\big]\Big)\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}\Big(\mathrm{Exp}\,\big[\mathbf{ad}[A]\big]\ \mathrm{Exp}\,\big[\mathbf{ad}[sB]\big]-\mathbf{id}\Big)^{k-1}(B)\\
&\quad=\; \left[\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}\Big(\mathrm{Exp}\,\big[\mathbf{ad}[A]\big]\ \mathrm{Exp}\,\big[\mathbf{ad}[sB]\big]-\mathbf{id}\Big)^{k-1}\right]\mathrm{Exp}\,\big[\mathbf{ad}[A]\big]\ \mathrm{Exp}\,\big[\mathbf{ad}[sB]\big](B)\\
&\quad=\; \left[\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}\Big(\mathrm{Exp}\,\big[\mathbf{ad}[A]\big]\ \mathrm{Exp}\,\big[\mathbf{ad}[sB]\big]-\mathbf{id}\Big)^{k-1}\right]\mathrm{Exp}\,\big[\mathbf{ad}[A]\big](B)\;,
\end{aligned} \tag{11.79}$$

onde, na última passagem, usamos o fato óbvio que

$$\mathrm{Exp}\,\big[\mathbf{ad}[sB]\big](B) \;=\; \mathbf{Ad}\big[\exp(sB)\big](B) \;=\; \exp(sB)B\exp(-sB) \;=\; B\;.$$

Desejamos escrever esta última expressão diretamente em termos das aplicações $\mathbf{ad}[A]$ e $\mathbf{ad}[sB]$. O último fator, $\mathrm{Exp}\,\big[\mathbf{ad}[A]\big]$, é simplesmente

$$\mathrm{Exp}\,\big[\mathbf{ad}[A]\big] \;=\; \sum_{l=0}^{\infty}\frac{1}{l!}\mathbf{ad}[A]^l\;. \tag{11.80}$$

Fora isso,

$$\mathrm{Exp}\left[\mathbf{ad}[A]\right]\,\mathrm{Exp}\left[\mathbf{ad}[sB]\right]-\mathbf{id} \;=\; \sum_{a=0}^{\infty}\sum_{b=0}^{\infty}\frac{1}{a!b!}\,\mathbf{ad}[A]^a\,\mathbf{ad}[sB]^b-\mathbf{id} \;=\; \sum_{\substack{a,\,b\geq 0\\ a+b>0}} s^b\,\frac{1}{a!b!}\,\mathbf{ad}[A]^a\,\mathbf{ad}[B]^b\,.$$

Com isso,

$$\Big(\mathrm{Exp}\left[\mathbf{ad}[A]\right]\,\mathrm{Exp}\left[\mathbf{ad}[sB]\right]-\mathbf{id}\Big)^{k-1}$$

$$= \sum_{\substack{a_1,\,b_1\geq 0\\ a_1+b_1>0}}\cdots\sum_{\substack{a_{k-1},\,b_{k-1}\geq 0\\ a_{k-1}+b_{k-1}>0}}\frac{s^{b_1+\cdots+s_{k-1}}}{a_1!b_1!\cdots a_{k-1}!b_{k-1}!}\,\mathbf{ad}[A]^{a_1}\,\mathbf{ad}[B]^{b_1}\cdots\mathbf{ad}[A]^{a_{k-1}}\,\mathbf{ad}[B]^{b_{k-1}}\,. \qquad (11.81)$$

Inserindo-se (11.80) e (11.81) em (11.79), tem-se

$$\int_0^1\psi\Big(\mathrm{Exp}\left[\mathbf{ad}[A]\right]\,\mathrm{Exp}\left[\mathbf{ad}[sB]\right]\Big)(B)\,ds \;=$$

$$\int_0^1\sum_{k=1}^{\infty}\sum_{l=0}^{\infty}\sum_{\substack{a_1,\,b_1\geq 0\\ a_1+b_1>0}}\cdots\sum_{\substack{a_{k-1},\,b_{k-1}\geq 0\\ a_{k-1}+b_{k-1}>0}}\frac{(-1)^{k-1}s^{b_1+\cdots+b_{k-1}}}{l!k}\left(\prod_{i=1}^{k-1}\frac{1}{a_i!b_i!}\right)\mathbf{ad}[A]^{a_1}\,\mathbf{ad}[B]^{b_1}\cdots\mathbf{ad}[A]^{a_{k-1}}\,\mathbf{ad}[B]^{b_{k-1}}\,\mathbf{ad}[A]^{l}(B)\,ds\,.$$

Trocando-se a integral pelas somas e usando que $\int_0^1 s^{b_1+\cdots+b_{k-1}}\,ds=(b_1+\cdots+b_{k-1}+1)^{-1}$, temos

$$\int_0^1\psi\Big(\mathrm{Exp}\left[\mathbf{ad}[A]\right]\,\mathrm{Exp}\left[\mathbf{ad}[sB]\right]\Big)(B)\,ds \;=$$

$$\sum_{k=1}^{\infty}\sum_{l=0}^{\infty}\sum_{\substack{a_1,\,b_1\geq 0\\ a_1+b_1>0}}\cdots\sum_{\substack{a_{k-1},\,b_{k-1}\geq 0\\ a_{k-1}+b_{k-1}>0}}\frac{(-1)^{k-1}}{l!k(b_1+\cdots+b_{k-1}+1)}\left(\prod_{i=1}^{k-1}\frac{1}{a_i!b_i!}\right)\mathbf{ad}[A]^{a_1}\,\mathbf{ad}[B]^{b_1}\cdots\mathbf{ad}[A]^{a_{k-1}}\,\mathbf{ad}[B]^{b_{k-1}}\,\mathbf{ad}[A]^{l}(B)$$

$$=\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{\substack{a_1,\,b_1\geq 0\\ a_1+b_1>0}}\cdots\sum_{\substack{a_k,\,b_k\geq 0\\ a_k+b_k>0}}\frac{(-1)^{k}}{l!(k+1)(b_1+\cdots+b_k+1)}\left(\prod_{i=1}^{k}\frac{1}{a_i!b_i!}\right)\mathbf{ad}[A]^{a_1}\,\mathbf{ad}[B]^{b_1}\cdots\mathbf{ad}[A]^{a_k}\,\mathbf{ad}[B]^{b_k}\,\mathbf{ad}[A]^{l}(B)\,. \qquad (11.82)$$

Na última igualdade fizemos apenas a mudança de variáveis $k\to k+1$.

Retornando a (11.78), temos então

$$\ln\left(e^A\,e^B\right) \;=\; A*B\,,$$

onde

$$A*B \;:=\; A+\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{\substack{a_1,\,b_1\geq 0\\ a_1+b_1>0}}\cdots\sum_{\substack{a_k,\,b_k\geq 0\\ a_k+b_k>0}}\frac{(-1)^{k}}{l!(k+1)(b_1+\cdots+b_k+1)}\left(\prod_{i=1}^{k}\frac{1}{a_i!b_i!}\right)$$

$$\times\,\mathbf{ad}[A]^{a_1}\,\mathbf{ad}[B]^{b_1}\cdots\mathbf{ad}[A]^{a_k}\,\mathbf{ad}[B]^{b_k}\,\mathbf{ad}[A]^{l}(B)\,. \qquad (11.83)$$

É fácil ver que o termo com $k=l=0$ nas somas do lado direito é igual a B. Com essa identificação, finalmente chega-se a (11.68). Como já comentamos, a convergência é garantida se $\|A\|_{\mathbb{C}}$ e $\|B\|_{\mathbb{C}}$ forem ambas menores que $\frac{1}{2}\ln\left(2-\frac{\sqrt{2}}{2}\right)\approx 0,12844\ldots$. ■

E. 11.22 *Exercício importante*. Colecionando os termos com $a_1+b_1+\cdots+a_k+b_k+l\leq 2$ em (11.68), mostre que os primeiros termos de $A*B$ são aqueles dados em (11.69), página 682. ♣

*

Comentário. Um comentário que adiantamos é que, como discutiremos melhor no Capítulo 22, página 1280 (vide, em especial, a Proposição 22.10, página 1312), o produto "$*$" expresso em (11.68), define uma estrutura de grupo em subálgebras de Lie nilpotentes de $\mathrm{Mat}\,(\mathbb{C},\,n)$. De fato, é possível provar que "$*$" é um produto associativo (pois o produto de exponenciais de matrizes é associativo) e é fácil ver que $A*0=A$ e que $A*(-A)=0$ para toda matriz A. Com isso, a matriz nula é o elemento neutro do grupo e $-A$ é a inversa de A. Isso também mostra que é por vezes possível construir um produto associativo a partir de outro não-associativo, como o comutador de matrizes. ♣

11.6 A Fórmula de Duhamel e Algumas de suas Consequências

Nesta seção demonstraremos a *Fórmula de Duhamel*[15]:

$$\exp(A+B) \;=\; \exp(A)+\int_0^1\exp\big((1-s)(A+B)\big)\,B\,\exp(sA)\,ds\,, \qquad (11.84)$$

válida para quaisquer matrizes $A,\,B\in\mathrm{Mat}\,(\mathbb{C}.\,n)$, e estudaremos algumas de suas consequências. A demonstração é simples. Diferenciando-se $e^{s(A+B)}e^{-sA}$ em relação a s, tem-se

$$\begin{aligned}\frac{d}{ds}\left(e^{s(A+B)}e^{-sA}\right) \;&=\; \left(\frac{d}{ds}e^{s(A+B)}\right)e^{-sA}+e^{s(A+B)}\left(\frac{d}{ds}e^{-sA}\right)\\ &=\; \left(e^{s(A+B)}\,(A+B)\right)e^{-sA}+e^{s(A+B)}\left((-A)\,e^{-sA}\right)\\ &=\; e^{s(A+B)}\,B\,e^{-sA}\,.\end{aligned}$$

Integrando-se ambos os lados entre 0 e t, obtem-se

$$e^{t(A+B)}e^{-tA}-\mathbb{1} \;=\; \int_0^t e^{s(A+B)}\,B\,e^{-sA}\,ds\,,$$

de onde segue que

$$e^{t(A+B)} \;=\; e^{tA}+\int_0^t e^{s(A+B)}\,B\,e^{-(s-t)A}\,ds\,,$$

A mudança de variável de integração $s\to t-s$ conduz a

$$e^{t(A+B)} \;=\; e^{tA}+\int_0^t e^{(t-s)(A+B)}\,B\,e^{sA}\,ds\,. \qquad (11.85)$$

Para $t=1$, isso reduz-se a (11.84), que é o que queríamos provar. De (11.85) podem ser extraídas várias relações úteis, que trataremos agora.

• **Derivada de uma exponencial em relação a um parâmetro**

Uma das consequências mais úteis da fórmula de Duhamel é uma relação para a derivada da exponencial de uma matriz que depende de um parâmetro. Seja $A(\lambda)\in\mathrm{Mat}\,(\mathbb{C}.\,n)$ uma matriz que depende contínua e diferenciavelmente de um parâmetro λ. Então vale

$$\frac{d}{d\lambda}\left(e^{A(\lambda)}\right) \;=\; \int_0^1 e^{(1-s)A(\lambda)}\left(\frac{d}{d\lambda}A(\lambda)\right)e^{sA(\lambda)}\,ds\,. \qquad (11.86)$$

[15] Jean Marie Constant Duhamel (1797–1872).

Essa relação tem aplicações em equações diferenciais e na Mecânica Estatística (dentro e fora do equilíbrio). Alguns autores também denominam-na *fórmula de Duhamel*. O leitor deve compará-la à expressão alternativa (11.64). Passemos à demonstração.

Sendo $A(\lambda)$ diferenciável, vale, para todo ϵ suficientemente pequeno,

$$A(\lambda+\epsilon) \;=\; A(\lambda)+\epsilon\frac{d}{d\lambda}A(\lambda)+R(\lambda,\;\epsilon)\;, \tag{11.87}$$

onde

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}R(\lambda,\;\epsilon)\;=\;0\;. \tag{11.88}$$

Tem-se, então,

$$\begin{aligned}
\frac{d}{d\lambda}\exp(A(\lambda)) \quad &\stackrel{\text{def.}}{=} \quad \lim_{\epsilon\to 0}\frac{1}{\epsilon}\Big[\exp(A(\lambda+\epsilon))-\exp(A(\lambda))\Big] \\
&\stackrel{(11.87)}{=} \quad \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\exp\left(A(\lambda)+\epsilon\frac{d}{d\lambda}A(\lambda)+R(\lambda,\;\epsilon)\right)-\exp\left(A(\lambda)\right)\right] \\
&\stackrel{(11.84)}{=} \quad \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[e^{A(\lambda)}+\int_0^1 e^{(1-s)\left(A(\lambda)+\epsilon\frac{dA}{d\lambda}(\lambda)+R(\lambda,\;\epsilon)\right)}\left(\epsilon\frac{dA}{d\lambda}(\lambda)+R(\lambda,\;\epsilon)\right)e^{sA(\lambda)}\,ds-e^{A(\lambda)}\right] \\
&= \quad \lim_{\epsilon\to 0}\left[\int_0^1 e^{(1-s)\left(A(\lambda)+\epsilon\frac{dA}{d\lambda}(\lambda)+R(\lambda,\;\epsilon)\right)}\left(\frac{dA}{d\lambda}(\lambda)\right)e^{sA(\lambda)}\,ds\right] \\
&\qquad + \lim_{\epsilon\to 0}\left[\int_0^1 e^{(1-s)\left(A(\lambda)+\epsilon\frac{dA}{d\lambda}(\lambda)+R(\lambda,\;\epsilon)\right)}\left(\frac{1}{\epsilon}R(\lambda,\;\epsilon)\right)e^{sA(\lambda)}\,ds\right] \\
&= \quad \int_0^1 e^{(1-s)A(\lambda)}\left(\frac{dA}{d\lambda}(\lambda)\right)e^{sA(\lambda)}\,ds+\left[\int_0^1 e^{(1-s)A(\lambda)}\left(\lim_{\epsilon\to 0}\frac{1}{\epsilon}R(\lambda,\;\epsilon)\right)e^{sA(\lambda)}\,ds\right] \\
&\stackrel{(11.88)}{=} \quad \int_0^1 e^{(1-s)A(\lambda)}\left(\frac{dA}{d\lambda}(\lambda)\right)e^{sA(\lambda)}\,ds\;,
\end{aligned}$$

como queríamos demonstrar.

• **Iterando a fórmula de Duhamel**

Na expressão (11.85) exponenciais do tipo $e^{\lambda(A+B)}$ aparecem em ambos os lados. Isso sugere que podemos inserir iterativamente (11.85) dentro de si mesma de modo a obter outras expressões recorrentes, como apresentado nas passagens

autoexplicativas abaixo. Partindo de (11.85) e repetindo a iteração duas vezes, tem-se

$$\begin{aligned}
e^{t(A+B)} \quad &= \quad e^{tA}+\int_0^t e^{(t-s_1)(A+B)}\,B\,e^{s_1A}\,ds_1 \\
&= \quad e^{tA}+\int_0^t\left(e^{(t-s_1)A}+\int_0^{t-s_1}e^{(t-s_1-s_2)(A+B)}\,B\,e^{s_2A}\,ds_2\right)B\,e^{s_1A}\,ds_1 \\
&= \quad e^{tA}+\int_0^t e^{(t-s_1)A}\,B\,e^{s_1A}\,ds_1+\int_0^t\int_0^{t-s_1}e^{(t-s_1-s_2)(A+B)}\,B\,e^{s_2A}\,B\,e^{s_1A}\,ds_2\,ds_1 \\
&= \quad e^{tA}+\int_0^t e^{(t-s_1)A}\,B\,e^{s_1A}\,ds_1+ \\
&\qquad \int_0^t\int_0^{t-s_1}\left(e^{(t-s_1-s_2)A}+\int_0^{t-s_1-s_2}e^{(t-s_1-s_2-s_3)(A+B)}\,B\,e^{s_3A}\,ds_3\right)B\,e^{s_2A}\,B\,e^{s_1A}\,ds_2\,ds_1 \\
&= \quad e^{tA}+\int_0^t e^{(t-s_1)A}\,B\,e^{s_1A}\,ds_1+\int_0^t\int_0^{t-s_1}e^{(t-s_1-s_2)A}\,B\,e^{s_2A}\,B\,e^{s_1A}\,ds_2\,ds_1 \\
&\qquad +\int_0^t\int_0^{t-s_1}\int_0^{t-s_1-s_2}e^{(t-s_1-s_2-s_3)(A+B)}\,B\,e^{s_3A}\,B\,e^{s_2A}\,B\,e^{s_1A}\,ds_3\,ds_2\,ds_1\;.
\end{aligned}$$

Repetindo-se N vezes o procedimento, teremos

$$\begin{aligned}
e^{t(A+B)} \quad &= \quad e^{tA}\left[\mathbb{1}+\int_0^t e^{-s_1A}\,B\,e^{s_1A}\,ds_1+\sum_{m=2}^{N}\int_0^t\int_0^{t-s_1}\cdots\int_0^{t-s_1-\cdots-s_{m-1}}e^{-(s_1+\cdots+s_m)A}\prod_{k=0}^{m-1}\left(B\,e^{s_{m-k}A}\right)ds_m\,\cdots\,ds_1\right] \\
&\qquad +\int_0^t\int_0^{t-s_1}\cdots\int_0^{t-s_1-\cdots-s_m}e^{(t-s_1-\cdots-s_{m+1})(A+B)}\prod_{k=0}^{m}\left(B\,e^{s_{m+1-k}A}\right)ds_{m+1}\,\cdots\,ds_1\;,
\end{aligned} \tag{11.89}$$

para todo $N\in\mathbb{N}$, $N\geq 2$, sendo que convencionamos definir a produtória de matrizes da esquerda para a direita, ou seja, na forma $\prod_{k=1}^{L}M_k=M_1\cdots M_L$ (é necessário fixar uma convenção devido à não-comutatividade do produto de matrizes). Com as mudanças de variáveis

$$\begin{aligned}
t_1 \;&=\; t-s_1\;, & s_1 \;&=\; t-t_1\;, \\
t_2 \;&=\; t-(s_1+s_2)\;, & s_2 \;&=\; t_1-t_2\;, \\
&\;\;\vdots & &\;\;\vdots \\
t_m \;&=\; t-(s_1+\cdots+s_m)\;, & s_m \;&=\; t_{m-1}-t_m\;,
\end{aligned}$$

podemos re-escrever as integrais entre colchetes acima na forma

$$\begin{aligned}
e^{t(A+B)} \quad &= \quad \left[\mathbb{1}+\int_0^t e^{t_1A}\,B\,e^{-t_1A}\,dt_1+\sum_{m=2}^{N}\int_0^t\int_0^{t_1}\cdots\int_0^{t_{m-1}}\prod_{k=0}^{m-1}\left(e^{t_{m-k}A}\,B\,e^{-t_{m-k}A}\right)dt_m\,\cdots\,dt_1\right]e^{tA} \\
&\qquad +\int_0^t\int_0^{t-s_1}\cdots\int_0^{t-s_1-\cdots-s_m}e^{(t-s_1-\cdots-s_{m+1})(A+B)}\prod_{k=0}^{m}\left(B\,e^{s_{m+1-k}A}\right)ds_{m+1}\,\cdots\,ds_1\;.
\end{aligned} \tag{11.90}$$

E. 11.23 *Exercício*. Verifique! ✣

Substituindo $A \to A^*$ e $B \to B^*$ na expressão acima, tomando a adjunta da expressão resultante e usando o fato que, para qualquer matriz $M \in \mathrm{Mat}(\mathbb{C},\ n)$, vale $(\exp(M^*))^* = \exp(M)$, obtem-se

$$e^{t(A+B)} \;=\; e^{tA}\left[\mathbb{1} + \int_0^t e^{-t_1 A}\, B\, e^{t_1 A}\, dt_1 + \sum_{m=2}^{N} \int_0^t \int_0^{t_1} \cdots \int_0^{t_{m-1}} \prod_{k=1}^{m} \left(e^{-t_k A}\, B\, e^{t_k A}\right) dt_m \cdots dt_1\right]$$
$$+ \int_0^t \int_0^{t-s_1} \cdots \int_0^{t-s_1-\cdots-s_m} \left[\prod_{k=1}^{m+1} \left(e^{s_k A}\, B\right)\right] e^{(t-s_1-\cdots-s_{m+1})(A+B)} ds_{m+1} \cdots ds_1 \,. \tag{11.91}$$

E. 11.24 *Exercício*. Verifique! ✣

Para matrizes ou elementos de uma álgebra-$*$ de Banach é possível tomar o limite $N \to \infty$ nas expressões (11.89)-(11.91), como na proposição que segue.

Proposição 11.17 *Sejam matrizes $A,\ B \in \mathrm{Mat}(\mathbb{C},\ n)$. Então,*

$$e^{t(A+B)} \;=\; e^{tA}\left[\mathbb{1} + \int_0^t e^{-s_1 A}\, B\, e^{s_1 A}\, ds_1\right.$$
$$\left. + \sum_{m=2}^{\infty} \int_0^t \int_0^{t-s_1} \cdots \int_0^{t-s_1-\cdots-s_{m-1}} e^{-(s_1+\cdots+s_m)A} \prod_{k=0}^{m-1} \left(B\, e^{s_{m-k} A}\right) ds_m \cdots ds_1\right] , \tag{11.92}$$

ou, equivalentemente,

$$e^{t(A+B)} \;=\; e^{tA}\left[\mathbb{1} + \int_0^t e^{-t_1 A}\, B\, e^{t_1 A}\, dt_1 + \sum_{m=2}^{\infty} \int_0^t \int_0^{t_1} \cdots \int_0^{t_{m-1}} \prod_{k=1}^{m} \left(e^{-t_k A}\, B\, e^{t_k A}\right) dt_m \cdots dt_1\right] , \tag{11.93}$$

para todo $t \in \mathbb{R}$, a convergência sendo uniforme para t em compactos. As expansões em série acima são denominadas séries de Duhamel. □

Prova. A prova consiste em mostrar que o limite $N \to \infty$ de (11.89) ou (11.91) existe. Tomemos provisoriamente $t \in [-T,\ T]$ para algum $T > 0$. Para $\tau \in [-T,\ T]$, tem-se $\|e^{\tau A}\| \le e^{|\tau|\|A\|} \le e^{T\|A\|}$. Seja $M := \max\left(e^{T\|A\|},\ e^{T\|A+B\|}\right)$. Tem-se

$$\left\|\int_0^t \int_0^{t_1} \cdots \int_0^{t_{m-1}} \prod_{k=1}^{m} \left(e^{-t_k A}\, B\, e^{t_k A}\right) dt_m \cdots dt_1\right\| \;\le\; M^{2m}\|B\|^m \int_0^t \int_0^{t_1} \cdots \int_0^{t_{m-1}} dt_m \cdots dt_1 \;=\; \frac{\left(M^2\|B\||t|\right)^m}{m!}$$

e, analogamente,

$$\left\|\int_0^t \int_0^{t-s_1} \cdots \int_0^{t-s_1-\cdots-s_m} e^{t-(s_1+\cdots+s_{m+1})(A+B)} \prod_{k=0}^{m} \left(B\, e^{s_{m+1-k} A}\right) ds_{m+1} \cdots ds_1\right\| \qquad \le \qquad M\frac{(M\|B\||t|)^{m+1}}{(m+1)!} \qquad .$$

As duas desigualdades provam a convergência uniforme para $t \in [-T,\ T]$. Como T é arbitrário, a convergência se dá para todo $t \in \mathbb{R}$. ■

Na Seção 14.4, página 761, apresentamos uma generalização da expressão (11.93), a chamada série de Dyson para da teoria de perturbações (vide, em particular, a expressão (14.35)). Vide tabém Exercício E. 14.8, página 763.

• **Outros resultados análogos**

O método de demonstração da fórmula de Duhamel apresentado acima pode ser empregado na obtenção de outros resultados. Sejam novamente matrizes $A,\ B \in \mathrm{Mat}(\mathbb{C},\ n)$. Então, vale

$$[A,\ e^{tB}] \;=\; \int_0^t e^{(t-s)B}[A,\ B]e^{sB}\, ds\,. \tag{11.94}$$

Para a prova, observamos que $\frac{d}{ds}\left(e^{-sB}Ae^{sB}\right) = e^{-sB}[A,\ B]e^{sB}$ (justifique!). Integrando-se ambos os lados de 0 a t, obtem-se

$$e^{-tB}Ae^{tB} - A \;=\; \int_0^t e^{-sB}[A,\ B]e^{sB} ds\,. \tag{11.95}$$

Multiplicando-se à esquerda por e^{tB} chega-se à expressão (11.94). Expressões como (11.94) são empregadas na teoria de perturbações na Mecânica Quântica.

11.7 Continuidade do Determinante

Com uso da desigualdade de Hadamard obtida no Teorema 10.40, página 644, podemos estabelecer a continuidade da função definida pelo determinante em $\mathrm{Mat}(\mathbb{C},\ n)$. No processo estabeleceremos uma desigualdade relevante sobre a diferença entre dois determinantes, a desigualdade (11.96), abaixo.

Proposição 11.18 *A função* $\det : \mathrm{Mat}(\mathbb{C},\ n) \to \mathbb{C}$, *que fornece o determinante de matrizes, é contínua quando* $\mathrm{Mat}(\mathbb{C},\ n)$ *é dotado de uma topologia métrica definida por qualquer uma de suas normas. Especificamente, tem-se*

$$\left|\det A - \det A'\right| \;\le\; \sqrt{n}\max\left\{\|A\|_2,\ \|A'\|_2\right\}^{n-1}\|A - A'\|_2\,, \tag{11.96}$$

para quaisquer $A,\ A' \in \mathrm{Mat}(\mathbb{C},\ n)$ onde, para $B \in \mathrm{Mat}(\mathbb{C},\ n)$, temos a norma $\|B\|_2 := \sqrt{\mathrm{Tr}(B^*B)}$. □

Demonstração. Seja $A \in \mathrm{Mat}(\mathbb{C},\ n)$ e seja $\mathbf{a}_k \in \mathbb{C}^n$, $1 \le k \le n$, sua k-ésima coluna, de sorte que, pela notação introduzida em (10.9), página 531, valha $A = [\![\mathbf{a}_1,\ \ldots,\ \mathbf{a}_n]\!]$. A desigualdade de Hadamard (10.216), página 644, afirma que

$$|\det A| \;\le\; \prod_{j=1}^{n} \|\mathbf{a}_j\|_2\,, \tag{11.97}$$

onde, para $v \in \mathbb{C}^n$ com componentes $v_1,\ \ldots,\ v_n \in \mathbb{C}$, definimos $\|v\|_2 := \sqrt{\sum_{i=1}^n |v_i|^2}$, a chamada norma Euclidiana em $\mathbb{C}^n$.

Sabemos por definição (vide (3.10), página 268) que $\det A$ é uma forma multilinear nos vetores-coluna de A: $\det A = \omega_{det}(\mathbf{a}_1,\ \ldots,\ \mathbf{a}_n)$. Seja uma segunda matriz $A' \in \mathrm{Mat}(\mathbb{C},\ n)$ com $A' = [\![\mathbf{a}'_1,\ \ldots,\ \mathbf{a}'_n]\!]$. Podemos escrever

$$\det A - \det A' \;=\; \omega_{det}(\mathbf{a}_1,\ \ldots,\ \mathbf{a}_n) - \omega_{det}(\mathbf{a}'_1,\ \ldots,\ \mathbf{a}'_n) \;=\; \sum_{k=1}^{n} \omega_{det}(\mathbf{a}'_1,\ \ldots, \mathbf{a}'_{k-1},\ \mathbf{a}_k - \mathbf{a}'_k,\ \mathbf{a}_{k+1},\ \ldots,\ \mathbf{a}_n)\,, \tag{11.98}$$

expressão essa que é denominada *soma telescópica*. De fato, se a escrevermos explicitamente, teremos

$$\sum_{k=1}^{n} \omega_{det}(\mathbf{a}'_1,\ \ldots, \mathbf{a}'_{k-1},\ \mathbf{a}_k - \mathbf{a}'_k,\ \mathbf{a}_{k+1},\ \ldots,\ \mathbf{a}_n)$$
$$= \omega_{det}(\mathbf{a}_1 - \mathbf{a}'_1,\ \mathbf{a}_2,\ \ldots, \mathbf{a}_n) + \omega_{det}(\mathbf{a}'_1,\ \mathbf{a}_2 - \mathbf{a}'_2,\ \mathbf{a}_3,\ \ldots, \mathbf{a}_n) + \omega_{det}(\mathbf{a}'_1,\ \mathbf{a}'_2,\ \mathbf{a}_3 - \mathbf{a}'_3,\ \mathbf{a}_4,\ \ldots, \mathbf{a}_n)$$
$$+ \cdots + \omega_{det}(\mathbf{a}'_1,\ \mathbf{a}'_2,\ \ldots,\ \mathbf{a}'_{n-1} - \mathbf{a}'_n,\ \mathbf{a}_n)$$

e o leitor poderá se convencer que, por meio de cancelamentos de termos sucessivos, resulta disso precisamente a diferença $\omega_{det}(\mathbf{a}_1, \ldots, \mathbf{a}_n) - \omega_{det}(\mathbf{a'}_1, \ldots, \mathbf{a'}_n)$.

Com isso, usando (11.97), temos

$$\left|\det A - \det A'\right| \leq \sum_{k=1}^{n} \left|\omega_{det}(\mathbf{a'}_1, \ldots, \mathbf{a'}_{k-1}, \mathbf{a}_k - \mathbf{a'}_k, \mathbf{a}_{k+1}, \ldots, \mathbf{a}_n)\right| \overset{(11.97)}{\leq} \sum_{k=1}^{n} \left(\prod_{l=1}^{k-1} \|\mathbf{a'}_l\|_2\right) \|\mathbf{a}_k - \mathbf{a'}_k\|_2 \left(\prod_{m=k+1}^{n} \|\mathbf{a}_m\|_2\right) .$$

Se

$$\mathcal{A} := \max\left\{\|\mathbf{a}_1\|_2, \ldots, \|\mathbf{a}_n\|_2, \|\mathbf{a'}_1\|_2, \ldots, \|\mathbf{a'}_n\|_2, \right\} , \tag{11.99}$$

podemos escrever, portanto,

$$\left|\det A - \det A'\right| \leq \mathcal{A}^{n-1} \sum_{k=1}^{n} \|\mathbf{a}_k - \mathbf{a'}_k\|_2 .$$

Usando a desigualdade de Cauchy-Schwarz, temos

$$\sum_{k=1}^{n} \|\mathbf{a}_k - \mathbf{a'}_k\|_2 = \sum_{k=1}^{n} 1 \times \|\mathbf{a}_k - \mathbf{a'}_k\|_2 \leq \sqrt{n}\sqrt{\sum_{k=1}^{n} \|\mathbf{a}_k - \mathbf{a'}_k\|_2^2} .$$

Assim, temos

$$\left|\det A - \det A'\right| \leq \sqrt{n}\mathcal{A}^{n-1} \sqrt{\sum_{k=1}^{n} \|\mathbf{a}_k - \mathbf{a'}_k\|_2^2} .$$

A soma do lado direito é igual a $\sum_{i,\, j=1}^{n} \left|A_{ij} - A'_{ij}\right|^2$, que por sua vez vale $\|A - A'\|_2^2 := \mathrm{Tr}\big((A - A')^*(A - A')\big)$, onde $\|\cdot\|_2$ é a chamada norma 2 da matriz $A - A'$. Temos, assim,

$$\left|\det A - \det A'\right| \leq \sqrt{n}\mathcal{A}^{n-1}\|A - A'\|_2 ,$$

com $\mathcal{A}$ definido em (11.99)

Vale ainda observar que $\|A\|_2^2 = \|\mathbf{a}_1\|_2^2 + \ldots + \|\mathbf{a}_n\|_2^2$ e, portanto,

$$\mathcal{A}^2 = \max\left\{\|\mathbf{a}_1\|_2^2, \ldots, \|\mathbf{a}_n\|_2^2, \|\mathbf{a'}_1\|_2^2, \ldots, \|\mathbf{a'}_n\|_2^2\right\} \leq \max\left\{\|A\|_2^2, \|A'\|_2^2\right\}$$

Assim, temos ainda

$$\left|\det A - \det A'\right| \leq \sqrt{n}\max\left\{\|A\|_2, \|A'\|_2\right\}^{n-1}\|A - A'\|_2 , \tag{11.100}$$

que é nossa expressão final.

Dela concluímos que se para $A \in \mathrm{Mat}(\mathbb{C}, n)$ houver uma sequência $A_n \in \mathrm{Mat}(\mathbb{C}, n)$, $n \in \mathbb{N}$, com $\|A - A_n\|_2 \to 0$ para $n \to \infty$, então $\left|\det A - \det A_n\right| \to 0$. Isso estabelece a continuidade da função $\det : \mathrm{Mat}(\mathbb{C}, n) \to \mathbb{C}$ em relação à norma $\|\cdot\|_2$. Como $\mathrm{Mat}(\mathbb{C}, n)$ é um espaço vetorial de dimensão finita, todas as normas lá definidas são equivalentes[16] e, portanto, estabelecemos também a continuidade do determinante para qualquer norma definida em $\mathrm{Mat}(\mathbb{C}, n)$. ■

[16] Vide Teorema 3.2, página 278, cuja demonstração encontra-se no no Apêndice 3.A, página 312.

11.8 Exercícios Adicionais

E. 11.25 *Exercício*. Seja A uma matriz $n \times n$ diagonalizável e seja

$$A = \sum_{k=1}^{r} \alpha_k E_k$$

sua representação espectral, onde $\alpha_1, \ldots, \alpha_r$ são seus r autovalores distintos ($1 \leq r \leq n$) e E_k são seus projetores espectrais, satisfazendo $E_a E_b = \delta_{a,\, b} E_a$ e $\mathbb{1} = \sum_{k=1}^{r} E_k$.

a) Mostre que

$$\exp(A) = \sum_{k=1}^{r} e^{\alpha_k} E_k . \tag{11.101}$$

b) Usando esse fato calcule $\exp(tA_1)$ e $\exp(tA_2)$ para as matrizes A_1 e A_2 dadas por

$$A_1 = \begin{pmatrix} 2 & 0 \\ -9i & 1-6i \end{pmatrix} , \qquad A_2 = \begin{pmatrix} -2i & 1+5i \\ 3-8i & 9 \end{pmatrix} .$$

✠

E. 11.26 *Exercício*. As chamadas *matrizes de Pauli* são definidas por

$$\sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \quad \sigma_2 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \text{e} \quad \sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} .$$

a) Mostre que as mesmas satisfazem as seguintes relações algébricas: para todos a, $b = 1, 2, 3$ valem

$$[\sigma_a, \sigma_b] := \sigma_a\sigma_b - \sigma_b\sigma_a = 2i\sum_{c=1}^{3} \varepsilon_{abc}\sigma_c ,$$

$$\{\sigma_a, \sigma_b\} := \sigma_a\sigma_b + \sigma_b\sigma_a = 2\delta_{ab}\mathbb{1} ,$$

$$\sigma_a\sigma_b = \delta_{ab}\mathbb{1} + i\sum_{c=1}^{3} \varepsilon_{abc}\sigma_c .$$

b) Mostre que as quatro matrizes $\mathbb{1}$, σ_1, σ_2, σ_3 formam uma base em $\mathrm{Mat}(\mathbb{C}, 2)$: toda matriz complexa 2×2 pode ser escrita como uma combinação linear das mesmas.

c) Mostre que as matrizes $\mathbb{1}$, σ_1, σ_2, σ_3 são ortonormais em relação ao seguinte produto escalar definido em $\mathrm{Mat}(\mathbb{C}, 2)$: $\langle A, B\rangle := \frac{1}{2}\mathrm{Tr}(A^*B)$.

d) Obtenha a representação espectral das matrizes de Pauli.

e) Seja $\vec{\eta} := (\eta_1, \eta_2, \eta_3)$ um vetor de comprimento 1 de $\mathbb{R}^3$, ou seja, $\|\vec{\eta}\| = 1$. Seja, $\vec{\eta}\cdot\vec{\sigma} := \eta_1\sigma_1 + \eta_2\sigma_2 + \eta_3\sigma_3$, onde σ_k são as matrizes de Pauli, definidas acima. Prove que

$$\exp(i\theta\vec{\eta}\cdot\vec{\sigma}) = \cos(\theta)\,\mathbb{1} + i\,\mathrm{sen}(\theta)\left(\vec{\eta}\cdot\vec{\sigma}\right) .$$

Sugestão: Obtenha a decomposição espectral de $\vec{\eta}\cdot\vec{\sigma}$ e use (11.101).

Várias das expressões acima obtidas são empregadas na Mecânica Quântica. ✠

E. 11.27 *Exercício*. Sabemos pelo Exercício E. 11.26, página 693, que as matrizes $\mathbb{1}$, σ_1, σ_2, e σ_3 formam uma base em $\mathrm{Mat}(\mathbb{C}, 2)$. Sejam A, $B \in \mathrm{Mat}(\mathbb{C}, 2)$. Escrevendo-se $A = a_0\mathbb{1} + \vec{a}\cdot\vec{\sigma}$ e $B = b_0\mathbb{1} + \vec{b}\cdot\vec{\sigma}$, prove que valem as relações

$$AB - BA = 2i\big(\vec{a}\times\vec{b}\big)\cdot\vec{\sigma} \tag{11.102}$$

e

$$(AB-BA)^2 \;=\; -4\big\|\vec{a}\times\vec{b}\big\|^2\mathbb{1}\,. \tag{11.103}$$

Prove a partir disso que

$$e^{[A,\,B]} \;=\; \cos\Big(2\big\|\vec{a}\times\vec{b}\big\|\Big)\,\mathbb{1}+i\frac{\operatorname{sen}\Big(2\big\|\vec{a}\times\vec{b}\big\|\Big)}{\big\|\vec{a}\times\vec{b}\big\|}\,\Big(\big(\vec{a}\times\vec{b}\big)\cdot\vec{\sigma}\Big)\,. \tag{11.104}$$

para o caso em que $\vec{a}\times\vec{b}\neq\vec{0}$ e $e^{[A,\,B]}=\mathbb{1}$ se $\vec{a}\times\vec{b}=\vec{0}$.

Sugestão: Use os fatos sobre matrizes de Pauli provados no Exercício E. 11.26, página 693.

O fato (vide (11.103)) de a matriz $(AB-BA)^2$ ser sempre um múltiplo da matriz identidade para quaisquer matrizes A e B é específico de duas dimensões e não é geralmente válido em mais de duas dimensões. Encontre exemplos. ✠

E. 11.28 *Exercício*. Este é um exercício sobre exponenciação de matrizes nilpotentes com algumas observações sobre o grupo de Lorentz em $2+1$-dimensões. O grupo de Lorentz é discutido na Seção 21.6, página 1200.

Mostre que $N := \left(\begin{smallmatrix}0&1&0\\1&0&-1\\0&1&0\end{smallmatrix}\right)$ é uma matriz nilpotente de índice 3, ou seja, que $N^2\neq 0$, mas $N^3=0$. Mostre, com isso, que $D(q):=e^{qN}$, com $q\in\mathbb{R}$, é dada por

$$D(q) \;=\; \begin{pmatrix}1+\frac{q^2}{2} & q & -\frac{q^2}{2}\\[6pt] q & 1 & -q\\[6pt] \frac{q^2}{2} & q & 1-\frac{q^2}{2}\end{pmatrix}. \tag{11.105}$$

Verifique explicitamente, usando o lado esquerdo de (11.105), que valem $D(0)=\mathbb{1}$ e $D(q_1)D(q_2)=D(q_1+q_2)$ para todos $q_1,\ q_2\in\mathbb{R}$. As matrizes $D(q)$ representam, assim, o que se chama de um subgrupo uniparamétrico, cujo gerador é N.

As matrizes $D(q)$ desempenham um papel no estudo do grupo de Lorentz (no caso, em $2+1$-dimensões), estando relacionadas às chamadas *translações horosféricas*: um tipo de transformação de Lorentz que mantém invariante um vetor tipo luz dado. No caso, o vetor tipo luz é o vetor $\ell := \left(\begin{smallmatrix}1\\0\\1\end{smallmatrix}\right)$. De fato, $N\ell=0$ e, portanto, $D(q)\ell=\ell$ para todo $q\in\mathbb{R}$. Verifique!

A transformação de Lorentz $D(q)$ é obtida por uma combinação de "boosts" e rotações, no senguinte sentido. A matriz N pode ser escrita na forma $N=L+M_1$, onde $L := \left(\begin{smallmatrix}0&0&0\\0&0&-1\\0&1&0\end{smallmatrix}\right)$ e $M_1 := \left(\begin{smallmatrix}0&1&0\\1&0&0\\0&0&0\end{smallmatrix}\right)$. A matriz L é o gerador das rotações espaciais e a matriz M_1 é o gerador de "boosts" na direção 1. Pela fórmuta de Trotter, vale

$$D(q) \;=\; e^{qL+qM_1} \;=\; \lim_{n\to\infty}\left[\exp\left(\frac{q}{n}L\right)\exp\left(\frac{q}{n}M_1\right)\right]^n\,.$$

A matriz $\exp\left(\frac{q}{n}L\right)$ representa uma rotação espacial de $\frac{q}{n}$ e matriz $\exp\left(\frac{q}{n}M_1\right)$ representa um "boost" na direção 1 com rapidez $\frac{q}{n}$. Assim, podemos dizer informalmente que a transformação de Lorentz $D(q)$ é obtida por uma sucessão de rotações infinitesimais intercaladas a "boosts" infinitesimais na direção 1.

É curioso notar que as componentes espaciais de ℓ apontam na direção 2, mas para que se mantenha esse vetor tipo luz invariante por uma transformação de Lorentz é preciso combinar "boosts" na direção 1 com rotações espaciais, na forma acima esclarecida. ✠

JCABarata. Notas de Aula. Versão de 9 de maio de 2025. Capítulo 11 695/3018

Parte IV

Equações Diferenciais