und erinnern uns an die Potenzreihe von $\log(1+x)$ aus Analysis I, nämlich

$$\log(1+x) \;=\; x - \frac{x^2}{2} + \frac{x^3}{3} - \ldots \;=\; \sum_{m\geq 1} (-1)^{m+1}\, \frac{x^m}{m}\,.$$

Diese Reihe konvergiert bekanntlich für alle $x \in \mathbb{C}$ mit $|x| < 1$. Ist $B \in \mathrm{Mat}(n, \mathbb{C})$ nun eine Matrix mit Spektralradius $\varphi(B) < 1$, so ist

$$\log(\mathbb{1} + B) \;=\; \sum_{m\geq 1} (-1)^{m+1}\, \frac{B^m}{m}$$

wohldefiniert und konvergent.

Lemma 7.15. *Ist M eine Markov-Matrix mit Spektrum $\sigma(M)$ und*

$$\varphi(M - \mathbb{1}) \;=\; \max_{\lambda\in\sigma(M)} |\lambda - 1| \;<\; 1\,,$$

so gibt es eine relle Matrix A mit verschwindenden Zeilensummen und $M = \mathrm{e}^A$.

Beweis. Wir setzen $A = \log(\mathbb{1} + (M - \mathbb{1}))$. Da die zugehörige Potenzreihe wegen der Annahme $\varphi(M - \mathbb{1}) < 1$ konvergiert, definiert dies eine reelle Matrix. Dabei ergibt sich $\sum_j A_{ij} = 0$, denn $M - \mathbb{1}$ ist selber ein Markov-Generator und wir können die Zeilensummen (analog zu früher) anhand der Potenzreihe ausrechnen. Weiter gilt dann:

$$\mathrm{e}^A \;=\; \mathrm{e}^{\log(\mathbb{1}+(M-\mathbb{1}))} \;=\; \mathrm{e}^{\log(M)} \;=\; M\,,$$

woraus sich die Behauptung konstruktiv ergibt. $\square$

Im letzten Beweis haben wir benutzt, dass der Logarithmus die Umkehrfunktion der Exponentialfunktion ist. Um zu sehen, dass dies auch für Matrizen so ist, erinnern wir kurz daran, dass für $|x| < 1$ folgende Rechnung (durch Einsetzen der Reihen ineinander) gilt:

$$\begin{aligned}
\mathrm{e}^{\log(1+x)} \;&=\; 1 + \bigl(x - \tfrac{1}{2}x^2 + \tfrac{1}{3}x^3 + \mathcal{O}(x^4)\bigr) \\
&\quad\; + \tfrac{1}{2}\bigl(x^2 - x^3 + \mathcal{O}(x^4)\bigr) \\
&\quad\; + \tfrac{1}{6}\bigl(x^3 + \mathcal{O}(x^4)\bigr) \;+\; \mathcal{O}(x^4) \\
&=\; 1 + x + \underbrace{\bigl(-\tfrac{1}{2} + \tfrac{1}{2}\bigr)}_{=0} x^2 + \underbrace{\bigl(\tfrac{1}{3} - \tfrac{1}{2} + \tfrac{1}{6}\bigr)}_{=0} x^3 + \mathcal{O}(x^4) \\
&=\; 1 + x + \mathcal{O}(x^4)\,,
\end{aligned}$$

und analog ergibt sich das für alle höheren Ordnungen, d.h.

$$\mathrm{e}^{\log(1+x)} \;=\; 1 + x + \mathcal{O}(x^n)$$

für alle $4 \leq n \in \mathbb{N}$ und damit insgeamt also $\mathrm{e}^{\log(1+x)} = 1+x$. Die Rechnung für Matrizen verläuft analog.

Nun wollen wir das Einbettungsproblem für $n = 2$ behandeln. Dazu betrachten wir eine Markov-Matrix der Form $M = \left(\begin{smallmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{smallmatrix}\right)$ mit den Parametern $0 \leq \alpha, \beta \leq 1$. Ein Eigenwert von M ist stets 1. Da die Spur der Matrix gleich der Summe der Eigenwerte ist, ergeben sich für obiges M also die zwei Eigenwerte $\lambda_1 = 1$ und $\lambda_2 = 1-\alpha-\beta$. Weiter lässt sich M als Summe zerlegen in

$$M \;=\; \mathbb{1} + \underbrace{\left(\begin{smallmatrix} -\alpha & \alpha \\ \beta & -\beta \end{smallmatrix}\right)}_{=M-\mathbb{1}},$$

wobei $M - \mathbb{1}$ ein Markov-Generator ist. Wenn nun $0 < \alpha, \beta < 1$ und $\alpha + \beta < 1$, so gilt auch $\varphi(M-\mathbb{1}) < 1$, und nach Lemma 7.15 läge es nahe, dass ein solches M darstellbar ist. Man beachte aber, dass im Lemma nur die Existenz einer Matrix A mit verschwindenden Zeilensummen attestiert wird. Wir müssen also noch klären, ob es dabei tatsächlich um einen Markov-Generator handelt. Dies kann man durch eine genaue Untersuchung der obigen Logarithmus-Reihe nachweisen (Details: Übung). Es gibt hier auch noch folgenden alternativen Zugang, der ohne den Logarithmus auskommt.

Der allgemeinste Markov-Generator für $n = 2$ lautet

$$A \;=\; \begin{pmatrix} -\alpha' & \alpha' \\ \beta' & -\beta' \end{pmatrix},$$

mit $\alpha', \beta' \geq 0$. Wir wollen jetzt die zugehörige Markov-Halbgruppe berechnen, also e^{tA} für $t \geq 0$. Falls $\alpha' = \beta' = 0$, bekommen wir $A = \mathbf{0}$ und $\mathrm{e}^{tA} \equiv \mathbb{1}$. Sei also $\alpha' + \beta' > 0$. Durch Induktion (Übung!) weist man nach, dass

$$A^n \;=\; (-1)^{n-1}\,(\alpha' + \beta')^{n-1}\,A$$

für alle $n \in \mathbb{N}$ gilt. Einsetzen in die allgemeine Formel für e^{tA} liefert dann

$$\begin{aligned}
\mathrm{e}^{tA} \;&=\; \mathbb{1} + \sum_{n\geq 1} (-1)^{n-1} (\alpha' + \beta')^{n-1} \frac{t^n}{n!} A \\
&=\; \mathbb{1} - \frac{1}{\alpha' + \beta'} \underbrace{\left(\sum_{n \geq 1} \frac{(-t)^n (\alpha' + \beta')^n}{n!} \right)}_{=\mathrm{e}^{-t(\alpha'+\beta')}-1} A \\
&=\; \mathbb{1} + \underbrace{\frac{1 - \mathrm{e}^{-t(\alpha'+\beta')}}{\alpha' + \beta'}}_{=:\varphi(t)} A \;=\; \begin{pmatrix} 1 - \varphi(t)\alpha' & \varphi(t)\alpha' \\ \varphi(t)\beta' & 1 - \varphi(t)\beta' \end{pmatrix}.
\end{aligned}$$

Dabei gilt $\varphi(0) = 0$ und $0 < (\alpha' + \beta')\varphi(t) = 1 - \mathrm{e}^{-t(\alpha'+\beta')} < 1$ für alle $t > 0$. Man kann sich nun überlegen, dass hierdurch genau die Matrizen aus obigem Lemma erfasst werden, wobei für festes $t > 0$ die Wahl von α' und β' jeweils eindeutig ist (Details: Übung). Damit erhalten wir den folgenden wichtigen Satz. Für weitere Überlegungen verweisen wir auf die Literatur, insbesondere auf [10].

Satz 7.16. *(Kendall)*
Die Markov-Matrix $M = \left(\begin{smallmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{smallmatrix}\right)$ *ist genau dann einbettbar, wenn* $\alpha, \beta \geq 0$ *gilt zusammen mit* $\alpha + \beta < 1$. *In diesem Fall gibt es genau einen Markov-Generator* A *mit* $M = \mathrm{e}^A$, *nämlich* $A = \log\big(\mathbb{1} + (M - \mathbb{1})\big)$.

Beweis. Falls $\alpha = \beta = 0$, ist $M = \mathbb{1} = \mathrm{e}^{\mathbf{0}}$. Sei also $\alpha + \beta > 0$. Dann ist aus unserer obigen Rechnung klar, dass wir Parameter α' und β' finden können, so dass M in der Form e^{tA} z.B. mit $t = 1$ geschrieben werden kann. Die Parameter α', β' sind dann eindeutig bestimmt durch $\alpha' = \alpha/\varphi(1)$ und $\beta' = \beta/\varphi(1)$. □

Beispiel 7.17. Wir betrachten noch einmal die Matrizen

$$M = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{pmatrix} \quad \text{und} \quad M^2 = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

aus Beispiel 7.13. Nach dem Satz von Kendall ist M^2 einbettbar, M aber nicht. Dabei ist also $M^2 = \mathrm{e}^A$, und man kann dann nachrechnen (Übung), dass

$$\exp\big(\tfrac{1}{2}A\big) = \begin{pmatrix} \frac{5}{6} & \frac{1}{6} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} =: M'$$

gilt, wobei natürlich $(M')^2 = M^2$ gilt. Wir sehen hier also das Phänomen, dass eine Matrix verschiedene Markov-Quadratwurzeln haben kann. Man kann sich rasch überlegen, dass es in diesem Fall keine weiteren gibt.

Bevor wir uns weiter mit dem Einbettungsproblem befassen, werfen wir noch einen kleinen Blick auf die Gleichgewichte von (eingebetteten) Markov-Matrizen. Dabei beschränken wir uns hier auf den Fall, dass es nur ein Gleichgewicht gibt. In diesem Fall ist die *stationäre Verteilung* eine wichtige Kenngröße des Markov-Prozesses, insbesondere für sein asymptotisches Verhalten. Mathematisch ist die stationäre Verteilung der (statistisch normierte) Links-Eigenvektor von einer Markov-Matrix (und damit, wie wir noch sehen werden, sogar von allen Matrizen) aus der Halbgruppe zum Eigenwert 1.

Beispiel 7.18. Für den Markov-Generator $A = \left(\begin{smallmatrix} -\alpha & \alpha \\ \beta & -\beta \end{smallmatrix}\right)$ mit $\alpha+\beta > 0$ ist die stationäre Verteilung durch

$$\left(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta}\right)$$

gegeben, wie man leicht nachrechnen kann.

Betrachten wir nun zunächst $\varphi(t)$ mit $\alpha + \beta > 0$ für $t \to \infty$, also

$$\varphi(t) = \frac{1 - \mathrm{e}^{-t(\alpha+\beta)}}{\alpha + \beta} \xrightarrow{t\to\infty} \frac{1}{\alpha+\beta},$$

so folgt daraus (wieder für $\alpha + \beta > 0$)

$$M(t) = \begin{pmatrix} 1 - \varphi(t)\alpha & \varphi(t)\alpha \\ \varphi(t)\beta & 1 - \varphi(t)\beta \end{pmatrix} \xrightarrow{t\to\infty} \begin{pmatrix} 1 - \frac{\alpha}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ \frac{\beta}{\alpha+\beta} & 1 - \frac{\beta}{\alpha+\beta} \end{pmatrix} = \begin{pmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \end{pmatrix}.$$

Dies ist das erwartete asymptotische Verhalten für eine primitive Markov-Matrix.

Nun wollen wir die stationäre Verteilung einer irreduziblen Markov-Halbgruppe etwas allgemeiner ansehen.

Proposition 7.19. *Sei A ein irreduzibler Markov-Generator, mit Dimension ≥ 2. Dann besitzt die Markov-Halbgruppe $\{\mathrm{e}^{tA} \mid t \geq 0\}$ einen simultanen PF-Eigenvektor. Dieser liegt dann im Kern von A, der wegen der Irreduzibilität eindimensional ist.*

Beweis. Es ist e^{tA} primitiv für alle $t > 0$. Sei nun ein $t > 0$ fixiert. Nach den Satz von Perron–Frobenius existiert nun ein eindeutiger, strikt positiver Wahrscheinlichkeitsvektor p mit $p\mathrm{e}^{tA} = p$. Weiter gilt für ein beliebiges $s \geq 0$

$$(p\mathrm{e}^{sA})\mathrm{e}^{tA} \;=\; p\mathrm{e}^{(t+s)A} \;=\; (p\mathrm{e}^{tA})\mathrm{e}^{sA} \;=\; p\mathrm{e}^{sA},$$

also ist auch $p\mathrm{e}^{sA}$ ein Eigenvektor von e^{tA} zum Eigenwert 1. Aufgrund der Markov-Eigenschaft von e^{tA} ist dies ebenfalls ein Wahrscheinlichkeitsvektor, also gilt $p\mathrm{e}^{sA} = p$ wegen der Eindeutigkeit des normierten Eigenvektors. Somit ist p simultaner Eigenvektor für die gesamte Markov-Halbgruppe.

Durch Differenzieren (nach t) bekommen wir nun

$$p\mathrm{e}^{tA}A \;=\; 0,$$

da p ja konstant ist. Einsetzen von $t = 0$ liefert $pA = 0$, also liegt p im Kern von A (genauer sollten wir sagen, im Links-Kern), und dieser ist eindimensional, weil ein linear unabhängiges Element im Kern die Existenz eines weiteren Eigenvektors von e^{tA} zum Eigenwert 1 bedeuten würde, was ja nicht möglich ist. □

Nun wollen wir noch eine allgemeine Formel angeben, um diesen Gleichgewichtszustand auszurechnen. Für einen Markov-Generator $A = (A_{ij})_{1 \leq i,j \leq n}$ setzen wir

$$v_i \;=\; \sum_{T \in \mathcal{R}^{(i)}} \; \prod_{\langle k,\ell \rangle \in T} A_{\ell k}\,,$$

wobei $\mathcal{R}^{(i)}$ die Menge der bezifferten Bäume mit n Knoten und Wurzel (i) sind und $\langle k, \ell \rangle$ eine gerichtete Kante von k nach ℓ bezeichnet. Man beachte, dass durch die Auszeichnung eines Knotens als Wurzel eine natürliche Richtung auf jeder Kante im Baum entsteht. Damit können wir nun folgendes Resultat formulieren.

Satz 7.20. *Ist A ein irreduzibler Markov-Generator, so gilt:*

(1) *$vA = 0$ mit $v = (v_1, \ldots, v_n)$ und v_i wie oben definiert. Dabei sind alle $v_i > 0$.*

(2) *$p_i = \frac{v_i}{\|v\|_1}$ definiert einen Wahrscheinlichkeitsvektor. Dieser ist die stationäre Verteilung von der Markov-Halbgruppe.*

(3) *$v_i = \det(-A_{(ii)})$. Dabei entsteht die Matrix $A_{(ii)}$ aus A durch Entfernen von Zeile i und Spalte i.*

Beweis-Bausteine. Die erste Aussage folgt aus dem Matrix-Baum-Satz, der auf Kirchhoff (1859) zurückgeht, zusammen mit dem Satz von Perron–Frobenius. Die zweite Aussage ist klar nach unserer obigen Rechnung für Proposition 7.19. Die dritte Aussage folgt aus der Cauchy–Binet-Formel, die eine Verallgemeinerung der Laplace-Entwicklung ist. □

Weiterführende Literatur zu diesem Thema ist [11].

Nun wollen wir zum Einbettungsproblem zurückkehren und noch ein wenig den Fall $n > 2$ betrachten, der deutlich komplizierter ist. In der Phylogenie benötigt man bekanntlich $n = 4$, weshalb diese Erweiterung notwendig ist.

Lemma 7.21. *Sei M eine $n \times n$ Markov-Matrix mit $\varphi(M - \mathbb{1}) < 1$. Dann existiert ein $B \in \mathrm{Mat}(n, \mathbb{R})$ mit $M = \mathrm{e}^B$ und $\sum_j B_{ij} = 0$ für alle $1 \le i \le n$. Für $n = 2$ ist B dann stets ein Markov-Generator.*

Beweis. $B = \log\bigl(\mathbb{1} + (M - \mathbb{1})\bigr)$ ist wegen $\varphi(M - \mathbb{1}) < 1$ eine konvergente Reihe, und es gilt $\mathrm{e}^B = \exp\bigl(\log(\mathbb{1} + (M - \mathbb{1}))\bigr) = M$. Dabei ist $A := M - \mathbb{1}$ ein Markov-Generator, also $A_{ij} \ge 0$ für alle $i \ne j$ und $\sum_j A_{ij} = 0$ für alle $1 \le i \le n$.

Damit bekommen wir aber auch (für i beliebig, aber fest):

$$\sum_{j=1}^{n} B_{ij} = \sum_{j=1}^{n} \sum_{k \ge 1} \frac{(-1)^{k-1}}{k} \bigl(A^k\bigr)_{ij} = \sum_{k \ge 1} \frac{(-1)^{k-1}}{k} \underbrace{\sum_{\ell=1}^{n} \bigl(A^{k-1}\bigr)_{i\ell} \underbrace{\Bigl(\sum_j A_{\ell j}\Bigr)}_{=0}}_{=0} = 0 .$$

Im Allgemeinen braucht $B_{ij} \ge 0$ für $i \ne j$ jedoch *nicht* zu gelten. Bei $n = 2$ stimmt das aber immer, wie wir aus der Herleitung des Satzes von Kendall bereits wissen. □

Beispiel 7.22. Sei $\beta > 0$ und $B = \left(\begin{smallmatrix} 0 & 0 \\ \beta & -\beta \end{smallmatrix}\right)$. Dann ist

$$\mathrm{e}^{tB} = \begin{pmatrix} 1 & 0 \\ 1 - \mathrm{e}^{-t\beta} & \mathrm{e}^{-t\beta} \end{pmatrix} = \begin{cases} \mathbb{1}, & \text{falls } t = 0, \\ \left(\begin{smallmatrix} 1 & 0 \\ 1 & 0 \end{smallmatrix}\right), & \text{im Grenzfall } t \to \infty. \end{cases}$$

Die 0 in der oberen rechten Ecke illustriert ein allgemeines Phänomen: Es gilt $\bigl(\mathrm{e}^{tB}\bigr)_{ij} = 0$ für $i \ne j$ und ein $t > 0$ genau dann, wenn dies für *alle* $t > 0$ stimmt. In dem Fall steht die 0 auch bereits an der entsprechenden Stelle im Generator.

Folgerung 7.23. *Ist B ein Markov-Generator, so gilt:*

$$\mathrm{e}^{tB} \textit{ ist irreduzibel} \iff \mathrm{e}^{tB} \textit{ ist primitiv} \iff \mathrm{e}^{tB} \textit{ ist total positiv.}$$

Wir sehen hieraus, dass der Zusammenhang zwischen den verschiedenen Irreduzibilitätsbegriffen bei Matrizen aus Markov-Halbgruppen deutlich einfacher ist als bei allgemeinen Markov-Matrizen.

Beispiel 7.24. Betrachten wir die Matrix $B = 2\pi \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, die kein Generator ist. Per Induktion sieht man rasch, dass $B^{2m} = (-1)^m (2\pi)^{2m} \mathbb{1}$ und $B^{2m+1} = (-1)^m (2\pi)^{2m} B$ gilt, mit $m \in \mathbb{N}_0$. Damit kann man nun e^B ausrechnen,

$$\begin{aligned} \mathrm{e}^B &= \sum_{m \geq 0} \frac{(-1)^m (2\pi)^{2m}}{(2m)!} \mathbb{1} + \sum_{m \geq 0} \frac{(-1)^m (2\pi)^{2m+1}}{(2m+1)!} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ &= \cos(2\pi)\, \mathbb{1} + \sin(2\pi) \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \mathbb{1}, \end{aligned}$$

was nichts anderes als eine matrixwertige Version der Relation $\mathrm{e}^{2\pi \mathrm{i}} = 1$ ist. Also ist dieses B ebenfalls ein 'Logarithmus' von $\mathbb{1}$, aber mit nicht-verschwindenden Zeilensummen.

Man kann mit einem einfachen Argument (Übung) zeigen, dass es bei $n = 2$ ausser $\mathbf{0}$ keine weitere Lösung von $\mathbb{1} = \mathrm{e}^B$ mit B reell und verschwindenden Zeilensummen gibt. Bereits bei $n = 3$ wird auch dies schon komplizierter.

Wenn eine Markov-Matrix M einbettbar ist, ergeben sich sofort mehrere Konsequenzen, die somit auch notwendige Eigenschaften für eine Einbettbarkeit sind, nämlich:

1. Nach dem *spektralen Abbildungssatz* (nachschlagen !) für die Eigenwerte von M und B in $M = \mathrm{e}^B$ muss gelten
$$\sigma(M) = \mathrm{e}^{\sigma(B)} = \{\mathrm{e}^\lambda \mid \lambda \in \sigma(B)\},$$
wobei dies auch für mehrfache Eigenwerte von M stimmen muss.

2. Es gilt $0 < \det(M) \leq 1$, also ist 0 niemals Eigenwert einer einbettbaren Markov-Matrix. Ausserdem gilt $\det(M) = 1$ nur für $M = \mathbb{1}$ (diese Eigenschaften folgen aus der Formel $\det(\mathrm{e}^B) = \mathrm{e}^{\mathrm{tr}(B)}$ zusammen mit $\mathrm{tr}(B) \leq 0$).

3. Falls $1 \neq \lambda \in \sigma(M)$, so muss $|\lambda| \leq 1$ gelten (Satz von Elfving).

4. $M_{ij} > 0$ und $M_{jk} > 0$ implizieren $M_{ik} > 0$.

5. M ist entweder reduzibel oder total positiv (und dann auch primitiv).

6. Ist $\lambda \in \sigma(M)$ reell und negativ, so ist die algebraische Vielfachheit von λ gerade.

Leider sind auch alle sechs Eigenschaften zusammen genommen immer noch nicht hinreichend für die Einbettbarkeit. Dabei hat die letzte Bedingung eine besondere Bedeutung. Eine etwas einfachere Frage ist, wann eine relle Matrix M einen reellen Logarithmus besitzt, also $M = \mathrm{e}^R$ mit R reell gilt. Die Antwort gibt das folgende Kriterium.

Satz 7.25 (Culver, 1966)**.** *Für $S \in \mathrm{GL}(n, \mathbb{R})$ besitzt $S = \mathrm{e}^R$ eine Lösung $R \in \mathrm{Mat}(n, \mathbb{R})$ genau dann wenn jeder elementare Jordan-Block zu einem negativen Eigenwert von S, in der komplexen Jordan-Normalform von S, geradzahlig oft vorkommt.*

Ist S diagonalisierbar, vereinfacht sich dies zu der Bedingung, dass jeder negative Eigenwert von S gerade algebraische Vielfachheit besitzen muss.

Beweis-Skizze. Ist $S = \mathrm{e}^R$ und T invertierbar, so gilt bekanntlich

$$TST^{-1} = T\mathrm{e}^R T^{-1} = \exp\bigl(TRT^{-1}\bigr).$$

Nun wähle man $T \in \mathrm{GL}(n, \mathbb{C})$ so, dass TST^{-1} in (komplexer) Jordan-Normalform vorliegt. Ist λ ein positiver Eigenwert von S, kann man den zugehörigen Block von TRT^{-1} ausrechnen. Ist λ ein genuin komplexer Eigenwert, also $\lambda \in \mathbb{C} \setminus \mathbb{R}$, muss es ein komplex konjugiertes Paar von Jordan-Blöcken geben, weil S ja reell ist und somit auch $\overline{\lambda}$ Eigenwert von S ist, mit denselben algebraischen und geometrischen Vielfachheiten wie λ. Dann kann man wieder eine Lösung ausrechnen, wenn man sich an

$$\mathrm{diag}(i, -i) \sim \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

erinnert, was wir oben (in Beispiel 7.24) schon einmal benutzt haben.

Der schwierige Fall ist ein Eigenwert $\lambda < 0$. Ist die Multiplizität der zugehörigen Jordan-Blöcke gleicher Größe gerade, kann man sich an Euler's Relation $\mathrm{e}^{\pi \mathrm{i}} = -1$ erinnern, die in Matrix-Form gerade

$$\exp\left(\pi \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right) = -\mathbb{1}$$

bedeutet, und wieder eine Lösung ausrechnen. Schließlich muss man noch nachweisen, dass es bei ungerader Vielfachheit nicht geht, weil immer ein Block übrigbleibt. □

Wegen der Struktur mit der Ähnlichkeitstransformation T kann man nun auch sehen, dass $\mathbb{1}$ für $n \geq 2$ sogar kontinuierlich viele reelle Logarithmen besitzt: Jedes T vertauscht mit $\mathbb{1}$, aber nicht unbedingt mit R; man vergleiche das mit Beispiel 7.24.

Die *Vertauschungseigenschaften* von Matrizen sind also (wieder einmal) entscheidend. Ist $M = \mathrm{e}^R$, so muss natürlich $[M, R] = \mathbf{0}$ gelten, also auch:

$$R \in \mathrm{cent}(M) = \{T \in \mathrm{Mat}(n, \mathbb{R}) \mid [M, T] = \mathbf{0}\}.$$

Schreiben wir wieder $M = \mathbb{1} + A$ mit $A = M - \mathbb{1}$, so gilt auch $\mathrm{cent}(M) = \mathrm{cent}(A)$. Dies hilft aufgrund des folgenden Satzes aus der (gehobenen) linearen Algebra.

Satz 7.26 (Frobenius)**.** *Sei $B \in \mathrm{Mat}(n, \mathbb{R})$. Dann sind äquivalent:*

(1) *Das charakteristische Polynom von B, also $\det(B - x\mathbb{1})$, stimmt mit dem Minimalpolynom von B überein.*

(2) *Es gilt $\dim(V_\lambda) = 1$ für jedes $\lambda \in \sigma(B)$, wobei V_λ der zugehörige Eigenraum ist.*

(3) *Die Matrix B ist* zyklisch. *Dies bedeutet: Es gibt einen Vektor $u \in \mathbb{R}^n$, so dass $\{u, Bu, B^2u, \ldots, B^{n-1}u\}$ eine Basis des $\mathbb{R}^n$ ist.*

(4) *Der Matrix-Ring $\mathrm{cent}(B)$ ist abelsch.*

(5) *Es gilt: $\mathrm{cent}(B) = \mathbb{R}[B]$; dieser Ring besteht also aus allen Polynomen in B mit reellen Koeffizienten.*

Einfache Matrizen, also solche mit paarweise verschiedenen Eigenwerten, sind zyklisch. Somit ist die Eigenschaft, zyklisch zu sein, auch eine *generische* Eigenschaft, und zwar im maßtheoretischen Sinne: Fast alle Matrizen sind zyklisch. Darum wollen wir nun diese Klasse noch etwas genauer ansehen. Dabei ist klar, dass $M = \mathbb{1} + A$ genau dann zyklisch ist, wenn das auch für A der Fall ist, wobei natürlich $\mathbb{R}[M] = \mathbb{R}[A]$ gilt.

Uns interessieren nur reelle Logarithmen von M mit Zeilensummen 0. Eine allgemeine Matrix $T \in \mathbb{R}[A]$ ist nun von der Form

$$T \,=\, a_0 \mathbb{1} + a_1 A + \ldots + a_m A^m$$

für ein $m \in \mathbb{N}_0$ und geeignete relle Koeffizienten a_i. Wegen $A = M - \mathbb{1}$ wissen wir aber schon, dass $\sum_j T_{ij} = a_0$ gilt, und zwar für alle i. Eine solche Matrix T hat also verschwindende Zeilensummen genau für $a_0 = 0$. Daher setzen wir jetzt

$$\mathrm{alg}(A) \,:=\, \langle A^m \mid m \in \mathbb{N} \rangle_{\mathbb{R}} \,=\, \langle A, A^2, \ldots, A^{n-1} \rangle_{\mathbb{R}} \subsetneq \mathbb{R}[A] \,,$$

wobei der zweite Schritt eine Konsequenz des Satzes von Cayley–Hamilton (nachschlagen !) ist. Für zyklische A gilt $\dim\bigl(\mathrm{alg}(A)\bigr) = n - 1$. Alle Elemente von $\mathrm{alg}(A)$ haben verschwindende Zeilensummen, sind aber i.A. keine Generatoren. Für zyklische Matrizen stellt sich dies nun gar nicht als Einschränkung heraus:

Lemma 7.27. *Ist M eine zyklische Markov-Matrix, so ist* 1 *ein einfacher Eigenwert von M. In diesem Fall besitzt jeder reelle Logarithmus von M automatisch Zeilensummen* 0*.*

Beweis. Für jede Markov-Matrix stimmt die algebraische Vielfachheit des Eigenwerts $\lambda = 1$ mit seine geometrischen Vielfachheit überein (Beweis per Jordan-Normalform, vgl. [2, Satz 13.10] für Details).

Ist die Matrix M zyklisch, muss 1 also ein einfacher Eigenwert sein, mit rechts-Eigenvektor $(1, \ldots, 1)^T$. Ist $M = \mathrm{e}^R$ mit R reell, gilt $R \in \mathbb{R}[A]$ mit $A = M - \mathbb{1}$. Ausserdem muss dann (nach dem spektralen Abbildungssatz) 0 ein einfacher Eigenwert von R sein, da dies die einzige relle Lösung von $\mathrm{e}^x = 1$ ist.

Sei nun u der zugehörige Eigenvektor von R zum Eigenwert 0. Dann gilt aber auch $Mu = \mathrm{e}^R u = u$, also folgt $u = \alpha(1, \ldots, 1)^T$ für ein $\alpha \neq 0$. Dies bedeutet aber gerade, dass in R alle Zeilensummen verschwinden. □

Aus dem Satz von Frobenius zusammen mit unserer obigen Überlegung zu den Vertauschungsrelationen ergibt sich nun unmittelbar die folgende Eigenschaft.

Folgerung 7.28. *Sei M eine zyklische $n \times n$ Markov-Matrix mit* $\det(M) \neq 0$*. Besitzt M einen reellen Logarithmus, $M = \mathrm{e}^R$, so ist $R \in \mathrm{alg}(A)$, mit $A = M - \mathbb{1}$.*

Damit können wir nun das Einbettungsproblem für eine wichtige Klasse von Matrizen voranbringen.

Satz 7.29. *Sei M eine zyklische Markov-Matrix mit reellem Spektrum. Dann besitzt M genau dann einen reellen Logarithmus, wenn alle Eigenwerte von M positiv sind.*

In diesem Fall, $M = \mathrm{e}^R$, ist R eindeutig, und erfüllt $R \in \mathrm{alg}(A)$ mit $A = M - \mathbb{1}$. Weiter ist M einbettbar genau wenn dieses R ein Markov-Generator ist.

Beweis-Skizze. Die erste Behauptung folgt auch dem Satz von Culver, und $R \in \mathrm{alg}(A)$ ist eine Konsequenz der voranstehenden Folgerung.

Die Eindeutigkeit bekommt man aus dem linearen Gleichungssystem, das man mittels des spektralen Abbildungssatz aus $\sigma(M) = \exp\bigl(\sigma(R)\bigr)$ ableitet, unter korrekter Verwendung der Jordan-Normalformen von M und R.

Die Generatoreigenschaft der Lösung bleibt offen, aber R ist der einzige Kandidat. □

Aus dem Beweis, der hier nicht im Detail geführt werden kann, ergibt sich noch, dass der Satz *konstruktiv* ist. Wir wissen, dass $R = \alpha_1 A + \ldots + \alpha_{n-1} A^{n-1}$ mit Koeffizienten $\alpha_i \in \mathbb{R}$ gelten muss, und diese Koeffizienten kann man aus dem im Beweis erwähnten linearen Gleichungssystem explizit berechnen. Die Generatoreigenschaft von R ist dann einfach nachzuweisen, indem man $R_{ij} \geq 0$ für alle $i \neq j$ überprüft.

Folgerung 7.30. *Sei M wie im vorigen Satz, mit positiven Eigenwerten. Dann ist die relle Matrix $R = \log(\mathbb{1} + A)$ der einzige Kandidat für die Einbettbarkeit von M.*

Beweis. Bei $\sigma(M) \subset \mathbb{R}_+$ gilt automatisch $\sigma(M) \subset (0, 1]$ nach dem Satz von Elfving. Dann ist aber der Spektralradius von A kleiner als 1, und $\log(\mathbb{1}+A)$ konvergiert. Aufgrund der Darstellung als Reihe in A ist dies eine relle Matrix, die auch in $\mathrm{alg}(A)$ liegt, und somit der einzige Kandidat für den gesuchten Markov-Generator ist. □

Nun bleibt noch festzuhalten, dass $\sigma(M)$ i.A. nicht reell sein muss. Beim Vorliegen von komplexen Eigenwerten (und dann von komplex-konjugierten Paaren) können Mehrdeutigkeiten auftreten, es kann also mehrere Lösungen für das Einbettungsproblem geben. Natürlich muss M nicht zyklisch sein. Wenn das passiert, wird das Einbettungsproblem noch einmal deutlich komplexer, und man benötigt Methoden aus der algebraischen Geometrie zu seiner Lösung. Bis $n = 4$ sind die möglichen Fälle inzwischen komplett studiert, darüberhinaus bleiben immer noch offene Fragen.