Z. Wahrscheinlichkeitstheorie verw. Gebiete
35, 39–43 (1976)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Bemerkungen zum Bang-Bang-Problem für stochastische Matrizen*

W. Hazod

Mathematisches Institut der Universität Tübingen, Auf der Morgenstelle 10, D-7400 Tübingen, Bundesrepublik Deutschland

A new proof for the Bang-Bang representation of imbeddable stochastic matrices (s. [2]) is given.

S. Johansen studierte in [1] das Einbettungsproblem für stochastische Matrizen und die damit eng verwandte Frage einer Beschreibung der zerlegbaren Wahrscheinlichkeitsmaße auf endlichen Halbgruppen. In [2] werden einbettbare Matrizen studiert, die eine Bang-Bang-Darstellung besitzen, d.h. als endliches Produkt von Poissonmatrizen darstellbar sind. In dieser Note wird ein neuer Beweis für die Bang-Bang-Darstellung angegeben, der auf der Lie-Struktur beruht und der auch eine Konstruktionsmöglichkeit angibt.

Anstelle des Browerschen Fixpunktsatzes, der in [2] als entscheidendes Hilfsmittel herangezogen wird, werden hier kanonische Koordinaten erster und zweiter Art in einer Lie-Gruppe betrachtet: Die Frage nach der Existenz einer „Bang-Bang-Darstellung" wird auf die Frage nach der Existenz gewisser Koordinatendarstellungen in Lie-Gruppen zurückgeführt. Diese Methode hat den Vorteil, daß sie unmittelbar auf stochastische Matrizen mit idempotentem Faktor sowie auf Wahrscheinlichkeitsmaße auf endlichen Halbgruppen angewandt werden kann.

I.

$\mathcal{G}$ sei eine Lie-Gruppe mit Lie-Algebra $\mathfrak{G}$, $\mathbf{X}_1, \ldots, \mathbf{X}_n$ sei eine Basis von $\mathfrak{G}$. $\pi_1, \pi_2\colon \mathfrak{G} \to \mathcal{G}$ seien die den kanonischen Koordinaten erster bzw. zweiter Art entsprechenden lokalen Diffeomorphismen

$$a_1 \mathbf{X}_1 + \cdots + a_n \mathbf{X}_n \xrightarrow{\pi_1} \exp\left(\sum_{i=1}^n a_i \mathbf{X}_i\right)$$

* Diese Arbeit entstand im Rahmen eines von der Deutschen Forschungsgemeinschaft finanzierten Forschungsvorhabens

bzw.

$$a_1\mathbf{X}_1+\cdots+a_n\mathbf{X}_n \xrightarrow{\pi_2} \prod_{i=1}^{n}\exp(a_i\mathbf{X}_i),$$

wobei das Produkt in einer festen Reihenfolge zu bilden ist.

Dann gibt es eine Null-Umgebung $U_\varepsilon=\{\sum a_i\mathbf{X}_i, \max|a_i|<\varepsilon\}$, so daß auf U_ε die Abbildungen $\pi_2^{-1}\pi_1$ und $\pi_1^{-1}\pi_2$ eindeutig definiert und diffeomorph sind. Ein solches $\varepsilon>0$ werde im folgenden festgehalten. Dann gilt:

Hilfssatz 1.1. *Zu jedem $\eta>0$ gibt es ein $\gamma(\eta,\varepsilon)>0$, so daß für alle t: $0<t<\gamma(\eta,\varepsilon)$ und alle a_i: $0<\eta<\min_i a_i\leqq\max a_i<\varepsilon$, $1\leqq i\leqq n$ gilt: Sei $A=\sum a_i\mathbf{X}_i$ und $B(t)=\pi_2^{-1}\pi_1(tA)=\sum b_i(ta_1,\dots,ta_n)\mathbf{X}_i$, dann ist $b_i(ta_1,\dots,ta_n)=b_i>0$, $1\leqq i\leqq n$.*

Beweis. Sei U_ε festgehalten und sei $V_\varepsilon=\pi_2^{-1}\pi_1(U_\varepsilon)$. Aus der Differenzierbarkeit von $f=\pi_2^{-1}\pi_1$ folgt: Sei g: $U_\varepsilon\to\mathfrak{G}$ definiert durch $g(A)=f(A)-A$, so ist für alle $A\in U_\varepsilon$:

$$\lim_{t\searrow 0}\frac{1}{t}\|g(tA)\|=0.$$

Das Differential der Abbildung f ist im Ursprung von $\mathfrak{G}$ ja gleich der Einheitsmatrix. Sei nun $A=\sum a_i\mathbf{X}_i$, $0<\eta<\min_i a_i$, dann ist $f(tA)=\sum b_i\mathbf{X}_i=tA+g(tA)=\sum(ta_i+g_i(tA))\mathbf{X}_i$. Wegen $g_i(tA)/t\to 0$ kann man für genügend kleines t, etwa $0<t<\gamma$ erreichen, daß $b_i=ta_i+g_i(tA)>0$, $1\leqq i\leqq n$.

Damit erhält man sofort das

Korollar 1.1. *Sei $\mathbb{P}$ der Kegel $\{\sum a_i\mathbf{X}_i, a_i\geqq 0\}$ in $\mathfrak{G}$, $\mathbb{P}^0$ sei der innere Kern von $\mathbb{P}$, weiter sei $\mathbb{E}=\{a_i\mathbf{X}_i, a_i\geqq 0, 1\leqq i\leqq n\}$ die Vereinigung der extremalen Strahlen.*

Dann gilt mit den Bezeichnungen des Hilfssatzes 1.1: Zu jedem $A\in\mathbb{P}^0$ gibt es ein $\gamma>0$, so daß für jedes t, $0<t<\gamma$, tA mittels $f=\pi_2^{-1}\pi_1$ in $\mathbb{P}^0$ abgebildet wird.

Auf $\mathbb{E}$ stimmt f mit der identischen Abbildung überein.

Beweis. Wegen $A\in\mathbb{P}^0$ ist $\min_i a_i>0$, andererseits ist γA für genügend kleines γ in U_ε, also ist der Hilfssatz 1.1 anwendbar.

Wenn man f durch $h=\pi_1^{-1}\pi_2$ ersetzt, so erhält man analog:

Hilfssatz 1.2. *Zu jedem $A\in\mathbb{P}^0\cap U_\varepsilon$ gibt es ein $\delta>0$, so daß $h(tA)\in\mathbb{P}^0$ für alle t: $0<t<\delta$.*

Satz 1.1. *Sei $A\in\mathbb{P}^0$, i.e. $A=\sum a_i\mathbf{X}_i$, $a_i>0$ für $1\leqq i\leqq n$. Dann läßt sich $\pi_1(A)=\exp(A)$ als endliches Produkt von Elementen aus $\pi_2(\mathbb{E})$ darstellen, genauer: Es gibt ein natürliches k und $g_i>0$, so daß $\exp\left(\sum_{i=1}^{n}a_i\mathbf{X}_i\right)=\left(\prod_{i=1}^{n}\exp(g_i\mathbf{X}_i)\right)^k$ (Bang-Bang-Darstellung, s. [2]). Zu jedem festen $A=\sum a_i\mathbf{X}_i\in\mathbb{P}^0$ gibt es ein kleinstes k, so daß diese Relation gilt und $g_i=g_i(k;a_1,\dots,a_n)$ eindeutig bestimmte analytische Funktionen sind.*

Beweis. Sei U_ε wie in Korollar 1.1 gewählt, dann sei k_1 die kleinste natürliche Zahl, so daß $\frac{1}{k_1}A\in\mathbb{P}^0\cap U_\varepsilon$. Nach Korollar 1.1 ist dann für alle t: $0<t<\gamma$,

$$\exp((t/k_1)A)=\prod\exp(g_i(t)X_i),\qquad g_i(t)>0.$$

Nun sei m die kleinste natürliche Zahl mit $1/m<\gamma$, dann erhält man

$$\exp(A)=\left(\exp\left(\frac{1}{mk_1}A\right)\right)^{mk_1}=\left(\prod \exp(g_i X_i)\right)^{mk_1}, \qquad g_i=g_i\left(\frac{1}{m}\right)>0.$$

Die Funktionen g_i sind in U_ε eindeutig und analytisch, da $f=\pi_2^{-1}\pi_1$ ein lokaler Diffeomorphismus ist, s. auch [3].

Nun sei $\mathfrak{Z}$ die Menge aller $u\in\mathscr{G}$, die eine Darstellung $u=v_1\dots v_N$ besitzen, wobei jedes u_i in der Menge $\exp(\mathbb{P}^0)$ liegt, i.e. $\mathfrak{Z}$ sei die von $\exp(\mathbb{P}^0)$ erzeugte Halbgruppe $\mathfrak{Z}=\langle\exp(\mathbb{P}^0)\rangle\subset\mathscr{G}$.

Korollar 1.3. *Jedes $u\in\mathfrak{Z}$ ist darstellbar in der Form $u=v_1'\dots v_M'$, wobei jedes v_i' von der Form* $\exp t\mathbf{X}_j$ *ist, i.e.* $\langle\exp t\mathbf{X}_j;\ t\geqq 0,\ 1\leqq j\leqq n\rangle=\langle\exp(\mathbb{P}^0\cup\mathbb{E})\rangle$.

Beweis. Sei $u=\exp A_1\dots\exp A_N$, $A_i\in\mathbb{P}^0\cup\mathbb{E}$. Auf jedes A_i wendet man nun den Satz 1.1 an und erhält die Behauptung.

II. Anwendung auf stochastische Matrizen

Sei $\mathscr{P}$ die Halbgruppe der stochastischen $n\times n$ Matrizen und $\mathscr{Q}$ die Menge der normierten Intensitätsmatrizen, i.e. $\mathscr{Q}=\{A-I, A\in\mathscr{P}\}$. Dann bildet der von $\mathscr{Q}$ aufgespannte Vektorraum $\mathscr{R}$ eine Lie-Algebra: Es ist:

$$[\mathbf{A}-\mathbf{I},\mathbf{B}-\mathbf{I}]=(\mathbf{A}-\mathbf{I})(\mathbf{B}-\mathbf{I})-(\mathbf{B}-\mathbf{I})(\mathbf{A}-\mathbf{I})=(\mathbf{AB}-\mathbf{I})-(\mathbf{BA}-\mathbf{I})\in\mathscr{R}.$$

Die Basis $\mathfrak{B}$ von $\mathscr{R}$ wähle man aus der Menge der extremalen Intensitätsmatrizen, $\mathfrak{B}\subseteq\mathscr{Q}_e$ (s. [1], 2.18). $\mathscr{P}_1$ sei die Menge aller Matrizen aus $\mathscr{P}$ mit nicht verschwindender Determinante, $\mathscr{P}_1$ kann daher als Halbgruppe in $Gl(n,R)$ aufgefaßt werden. Der Lie-Algebra $\mathscr{R}$ entspricht eine Lie-Untergruppe $\mathbf{K}\subseteq Gl(n,R)$, $\mathbf{K}$ enthält alle Matrizen der Form $\exp(t\mathbf{X})$, $t\geqq 0$, $\mathbf{X}\in\mathscr{Q}$, offensichtlich liegen diese Matrizen auch in $\mathscr{P}_1$. Die Matrizen der Form $\exp(t\mathbf{X})$, $t\geqq 0$, $\mathbf{X}\in\mathscr{Q}_e$, heißen nach Johansen „Poissonmatrizen“. Um Übereinstimmung mit der bei Wahrscheinlichkeitsmaßen üblichen Terminologie zu erreichen, wäre es besser, von „elementaren Poissonmatrizen“ zu sprechen. $\mathscr{A}$ sei die Menge der endlichen Produkte elementarer Poissonmatrizen, also $\mathscr{A}=\langle\exp t\mathbf{X};\ t\geqq 0,\ \mathbf{X}\in\mathscr{Q}_e\rangle$. Weiter sei

$$\mathscr{Q}_1=\{t\mathbf{X},\ t\geqq 0,\ \mathbf{X}\in\mathscr{Q}\}$$

und

$$\mathscr{Q}_1^0=\{\sum_{\mathfrak{B}} a_i\mathbf{X}_i;\ a_i>0,\ \{X_1,\dots,X_m\}=\mathfrak{B}\}.$$

Dann folgt aus Satz 1.1 und Korollar 1.3:

Satz 2.1. a) *Sei* $\mathbf{A}\in\mathscr{Q}_1^0$, *dann ist* $\exp\mathbf{A}\in\mathscr{A}$ (s. [2], *Korollar* 2.4).

b) *Sei $\mathbf{A}\in\mathscr{P}$ und zerlegbar in folgendem Sinne: Es gibt eine Darstellung $\mathbf{A}=\mathbf{A}_1\dots\mathbf{A}_N$ mit $\mathbf{A}_i\in\exp(\mathscr{Q}_1^0)$, $1\leqq i\leqq N$, dann ist $\mathbf{A}\in\mathscr{A}$, i.e.*

$$\mathscr{A}=\langle\exp(\mathscr{Q}_1^0\cup\{t\mathbf{X},\mathbf{X}\in\mathscr{Q}_e,\ t\geqq 0\})\rangle.$$

(*Dies ist eine schwache Version von* [2], *Theorem* 2.5.)

Ein allgemeinerer Satz folgt, wenn man anstelle der, bezüglich der Einheitsmatrix $\mathbf{I}$ invertierbaren stochastischen Matrizen die $\mathbf{J}$-invertierbaren Matrizen in $\mathscr{P}(\mathbf{J}) := \mathbf{J}\mathscr{P}\mathbf{J}$ betrachtet, wobei $\mathbf{J}$ eine idempotente stochastische Matrix bezeichnet. Die Lie-Algebra der von den $\mathbf{J}$-invertierbaren Elementen erzeugten Gruppe stimmt mit der Menge der reellen $\mathbf{J}$-invarianten Matrizen überein, die Exponentialabbildung ist definiert durch $\exp \mathbf{A} = \exp_{\mathbf{J}}(\mathbf{A}) = \mathbf{J} + \mathbf{A} + \frac{1}{2!}\mathbf{A}^2 + \cdots$. Es sei

$$\mathscr{Q}(\mathbf{J}) = \{\mathbf{J}\mathbf{X}\mathbf{J}, \mathbf{X} \in \mathscr{Q}\} = \{\mathbf{J}\mathbf{P}\mathbf{J} - \mathbf{J}, \mathbf{P} \in \mathscr{P}\}, \qquad \mathscr{Q}_e(\mathbf{J}) = \{\mathbf{J}\mathbf{X}\mathbf{J}, \mathbf{X} \in \mathscr{Q}_e\},$$

$$\mathscr{Q}_1(\mathbf{J}) = \{t\mathbf{X}, t \geqq 0, \mathbf{X} \in \mathscr{Q}(\mathbf{J})\}.$$

$\mathscr{R}(\mathbf{J})$ sei der von $\mathscr{Q}_1(\mathbf{J})$ erzeugte Vektorraum, und man wähle wieder eine Basis $\mathfrak{B}$ von $\mathscr{R}(\mathbf{J})$ $\mathfrak{B} \subset \mathscr{Q}_e(\mathbf{J})$. Es sei schließlich $\mathscr{Q}_1^0(\mathbf{J}) = \{\sum_{\mathfrak{B}} a_i \mathbf{X}_i, a_i > 0, \mathfrak{B} = \{\mathbf{X}_i\}\}$ sowie $\mathscr{A}(\mathbf{J}) = \langle \exp_{\mathbf{J}}(t\mathbf{X}), t > 0, \mathbf{X} \in \mathscr{Q}_e(\mathbf{J})\rangle$ und $\mathfrak{Z}'(\mathbf{J}) = \langle \exp_{\mathbf{J}} t\mathbf{A};\ \mathbf{A} \in \mathscr{Q}_1^0 \cup \mathscr{Q}_e(\mathbf{J}), t \geqq 0\rangle$. Mit diesen Bezeichnungen gilt

Satz 2.1′. *Es ist $\mathscr{A}(\mathbf{J}) = \mathfrak{Z}'(\mathbf{J})$ für jede idempotente Matrix in $\mathscr{P}$.* Genauer: Zu jedem $A \in \exp_{\mathbf{J}}(\mathscr{Q}_1^0(\mathbf{J}))$ gibt es eine kleinste natürliche Zahl k und Funktionen $g_1, \ldots, g_m$ ($m = \operatorname{card}(\mathfrak{B}) = \dim \mathscr{R}(\mathbf{J})$), die eindeutig bestimmt und analytisch sind, so daß $\exp_{\mathbf{J}}(t\mathbf{A}) = \left(\prod_{i=1}^{m} \exp_{\mathbf{J}}\left(g_i\left(t, \frac{1}{k}\mathbf{A}\right)\mathbf{X}_i\right)\right)^k$. Dabei ist $\mathfrak{B} = \{\mathbf{X}_1, \ldots, \mathbf{X}_m\}$.

III. Zerlegbare Maße auf endlichen Halbgruppen

Sei $\mathbb{X}$ eine lokalkompakte halbtopologische Halbgruppe, $M(\mathbb{X})$ bzw. $W(\mathbb{X})$ bezeichne die Menge der beschränkten Borel- bzw. der Wahrscheinlichkeitsmaße auf $\mathbb{X}$. Die Faltung von a und $b \in M(\mathbb{X})$ werde mit ab bezeichnet. Dann bildet der von $\{a - \delta_e, a \in W(\mathbb{X})\}$ aufgespannte Vektorraum eine Lie-Algebra, wenn man $[a, b] = ab - ba$ definiert. Wenn $\mathbb{X}$ endlich ist, dann ist auch die Lie-Algebra, die mit $\mathfrak{E}$ bezeichnet werde, endlich-dimensional, als Basis wähle man die Generatoren elementarer Poissonmaße, nämlich $\mathscr{Q}_e = \{\delta_x - \delta_e, x \in \mathbb{X}, x \neq e\}$. Die Maße in dem von $\mathscr{Q}_e$ aufgespannten positiven Kegel $\mathscr{Q} = \{t(a - \delta_e), a \in W(\mathbb{X}), t \geqq 0\}$ heißen Poissongeneratoren. Die Maße $\exp(tb)$, $t > 0$, $b \in \mathscr{Q}$, werden Poisson- bzw. elementare Poissonmaße genannt, wenn $b \in \mathscr{Q}$ bzw. $b \in \mathscr{Q}_e$. Allgemeiner: Es sei $\mathfrak{j} \in W(\mathbb{X})$ ein idempotentes Maß, mit $M(\mathfrak{j})$ werden die Maße der Gestalt $\mathfrak{j} a \mathfrak{j}$, $a \in M(\mathbb{X})$ bezeichnet, weiter sei $W(\mathfrak{j}) = W(\mathbb{X}) \cap M(\mathfrak{j})$ und schließlich $\mathscr{Q}(\mathfrak{j}) = \{t(\mathfrak{j} a \mathfrak{j} - \mathfrak{j}), t > 0, a \in W(\mathbb{X})\}$, $\mathscr{Q}_e(\mathfrak{j}) = \{\mathfrak{j}\delta_x \mathfrak{j} - \mathfrak{j}, x \in \mathbb{X}\}$. Dann bildet der von $\mathscr{Q}(\mathfrak{j})$ aufgespannte Vektorraum $\mathfrak{E}(\mathfrak{j})$ eine Lie-Algebra, $\mathscr{Q}_e(\mathfrak{j})$ ist eine Basis von $\mathfrak{E}(\mathfrak{j})$. Die Menge der bezüglich $\mathfrak{j}$ invertierbaren Elemente von $M(\mathbb{X})$ bildet eine Lie-Gruppe, die Exponentialabbildung ist durch $b \to \exp_{\mathfrak{j}}(b) = \mathfrak{j} + \sum_{k=1}^{\infty} \frac{b^k}{k!}$ gegeben. $G_{\mathfrak{j}}$ sei die zur Lie-Algebra $\mathfrak{E}(\mathfrak{j})$ gehörige Untergruppe. Bekanntlich ist jedes unendlich teilbare Maß in $W(\mathbb{X})$ von der Gestalt $\exp_{\mathfrak{j}}(b)$, $b \in \mathscr{Q}(\mathfrak{j})$.

$\mathscr{A}(\mathfrak{j})$ sei die Halbgruppe der endlichen Faltungsprodukte von Maßen der Gestalt $\exp_{\mathfrak{j}}(tb)$, $b \in \mathscr{Q}_e(\mathfrak{j})$, $t > 0$. (Im Falle $\mathfrak{j} = \delta_e$ sind das also die Faltungsprodukte elementarer Poissonmaße.)

Satz 3.1. a) *Sei* $a \in \mathcal{Q}(\mathbb{j})^0$, *also* $a = \sum_{x \in \mathbb{X}}' c_x(\mathbb{j}\,\delta_x \mathbb{j} - \mathbb{j})$, $c_x > 0$. (*Dabei ist die Summe über alle* $x \in \mathbb{X}$ *zu erstrecken, für die die Maße* $\mathbb{j}\,\delta_x \mathbb{j} - \mathbb{j}$ *verschieden sind.*) *Dann ist* $\exp_{\mathbb{j}}(a) \in \mathscr{A}(\mathbb{j})$.

b) *Sei* $a \in W(\mathbb{j})$ *zerlegbar in folgendem Sinne:*

a besitzt eine Darstellung $a = a_1 \ldots a_n$, $a_i \in \exp_{\mathbb{j}}(\mathcal{Q}(\mathbb{j})^0)$, *dann ist* $a \in \mathscr{A}(\mathbb{j})$.

Der *Beweis* folgt wiederum sofort aus Satz 1.1 und Korollar 1.3.

Es ist noch zu bemerken, daß nicht nur die Existenz einer „Bang-Bang-Darstellung", sondern auch eine Konstruktionsmöglichkeit aufgezeigt wurde. Ebenso lassen die Lösungen der mit dem Einbettungsproblem verbundenen Differentialgleichung $\frac{d^+}{dt} P(t) = P(t) Q(t)$, $P(.) \in \mathscr{P}$, $Q(.) \in \mathcal{Q}_1^0$ im Kleinen eine eindeutig bestimmte Bang-Bang-Darstellung zu, s. [3]. Die Bestimmung der Lösungen beruht im wesentlichen auf der Baker-Hausdorff-Formel

$$\exp(\mathbf{X})\, \mathbf{Y} \exp(-\mathbf{X}) = \exp(ad(\mathbf{X}))\, \mathbf{Y},$$

wobei $\mathbf{X}$ und $\mathbf{Y}$ Elemente der Lie-Algebra bezeichnen.

Bei den in III. betrachteten Anwendungen läßt sich die adjungierte Darstellung überschauen: Sei $\mathbb{X} = \{x_1, \ldots, x_n\}$ mit den Multiplikationsregeln $x_i x_j = x_{k(i,j)}$. Es sei $\mathbf{X}_i = \delta_{x_i} - \delta_e$, $1 \leqq i \leqq n$. Dann prüft man sofort nach, daß $ad(\mathbf{X}_i)\mathbf{X}_j = [\mathbf{X}_i, \mathbf{X}_j] = \mathbf{X}_{k(i,j)} - \mathbf{X}_{k(j,i)}$ für alle i, j gilt.

Literatur

1. Johansen, S.: A central Limit Theorem for Finite Semigroups and Its Application to the Imbedding Problem for finite State Markov Chains. Z. Wahrscheinlichkeitstheorie verw. Gebiete **26**, 171 – 190 (1973)
2. Johansen, S.: The Bang-Bang Problem for Stochastic Matrices. Z. Wahrscheinlichkeitstheorie verw. Gebiete **26**, 191 – 195 (1973)
3. Wei, J., Norman, E.: On Global Representations of the Solutions of Linear Differential Equations as a Product of Exponentials. Proc. Amer. Math. Soc. **15**, 327 – 334 (1965)

Eingegangen am 20. Juni 1974; in revidierter Form am 10. November 1975