Prof. Dr. Peter Kunkel
Fakultät für Mathematik und Informatik
Universität Leipzig SS 2022

Ergänzung zur Vorlesung
Verallgemeinerte Inverse

3 Anwendungen und Verallgemeinerungen

3.1 Elektronische Schaltkreise

Man kann sich einen elektronischen Schaltkreis vorstellen als einen (zusammenhängenden) Graphen, dessen Kanten die Bauteile und dessen Knoten die verbindenden Drähte repräsentieren. Wir betrachten hier nur Widerstandsnetzwerke mit zusätzlichen Gleichstrom- und Gleichspannungsquellen.

Sei

$$V = \{n_i \mid i = 1, \dots, m\} \tag{3.1}$$

die Menge der Knoten und

$$E = \{e_j \mid j = 1, \dots, n\} \subseteq V \times V \tag{3.2}$$

die Menge der Kanten des Graphen. Die Struktur des (gerichteten) Graphen kann dargestellt werden mittels der sog. Inzidenzmatrix $M \in \{-1, 0, +1\}^{m,n}$ definiert durch

$$m_{ij} = \begin{cases} +1 & \text{für } i = k, \\ -1 & \text{für } i = l, \\ \ \ 0 & \text{sonst,} \end{cases} \qquad e_j = (n_k, n_l). \tag{3.3}$$

Zu jedem Knoten n_i gehört ein Potential p_i und zu jeder Kante e_j ein Strom y_j. Weiter sei $x_j = p_k - p_l$ der Spannungsabfall längs der Kante $e_j = (n_k, n_l)$.

Es wird angenommen, daß das Netzwerk den beiden Kirchhoffschen Gesetzen genügt. Das erste besagt, daß die Summe aller Ströme in einem Knoten verschwindet, d. h.

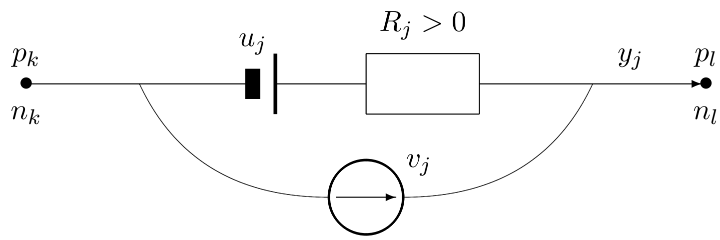
$$\sum_{j=1}^{n} m_{ij} y_j = 0 \quad \text{bzw.} \quad My = 0. \tag{3.4}$$

Das zweite besagt, daß die Summe der Spannungsabfälle in einer Masche verschwindet. In der vorliegenden Formulierung ist dies erfüllt durch die Definition von x_j über die Potentiale p_i der Knoten. Es gilt per Definition

$$\sum_{i=1}^{m} m_{ij} p_i = x_j \quad \text{bzw.} \quad x = M^T p. \tag{3.5}$$

Schließlich soll eine Kante $e_j = (n_k, n_l)$ das Bauteil

(3.6)



repräsentieren. Mit Hilfe des Ohmschen Gesetz bedeutet dies

$$R_j(y_j - v_j) = x_j - u_j. \tag{3.7}$$

Setzt man nun noch

$$\begin{array}{llll} \text{(a)} & a_j = R_j^{-1}, & & A = \operatorname{diag}(a_1, \ldots, a_n), \\ \text{(b)} & b_j = R_j^{-1} u_j - v_j, & & b = \operatorname{vec}(b_1, \ldots, b_n), \end{array} \tag{3.8}$$

so erhält man mit $L = M^T$ das Problem

$$Ax - y = b, \qquad \begin{array}{l} x \in \operatorname{bild} M^T = \operatorname{bild} L, \\ y \in \operatorname{kern} M = \operatorname{cobild} L, \end{array} \tag{3.9}$$

entsprechend zu (1.60).

Satz 3.1
Es gibt eine zum Problem (3.9) gehörige Bott-Duffin-Inverse A^L von A.

Beweis:
Nach Satz 1.44 ist zu zeigen, daß $ALL^+ + (I - LL^+)$ nichtsingulär ist. Sei ein z gegeben mit

$$(ALL^+ + (I - LL^+))z = 0.$$

Man kann z aufspalten gemäß $z = x + y$ mit $x \in \operatorname{bild} L$ und $y \in \operatorname{cobild} L$, d. h.

$$LL^+ x = x, \quad LL^+ y = 0.$$

Da $A = \operatorname{diag}(R_1^{-1}, \ldots, R_n^{-1})$ symmetrisch und positiv definit ist, erhält man

$$\begin{array}{rlllll} Ax + y = 0 & \Longrightarrow & Ax \in \operatorname{cobild} L & \Longrightarrow & L^T A x = 0 \\ & \Longrightarrow & x^T A L = 0 & \Longrightarrow & x^T A L L^+ x = 0 \\ & \Longrightarrow & x^T A x = 0 & \Longrightarrow & x = 0 \end{array}$$

sowie $y = -Ax = 0$, d. h. $z = 0$.

$\square$

Damit wird (3.9) gelöst durch

$$
\begin{aligned}
(3.10) \qquad & x = A^L b, \quad y = -(I - AA^L)b, \quad b = Au - v, \\
& u = \operatorname{vec}(u_1, \ldots, u_n), \quad v = \operatorname{vec}(v_1, \ldots, v_n).
\end{aligned}
$$

In der Physik interpretiert man gerne den Zustand, den ein System einnimmt, als den Zustand, der eine bestimmtes Energiefunktional minimiert. Im folgenden verwenden wir dazu M nur noch in der Bedeutung von Satz 1.44.

Satz 3.2
Die Größen (3.10) lösen die Minimierungsprobleme

$$
(3.11) \qquad
\begin{array}{ll}
\text{(a)} & p(x) = \frac{1}{2}(x-u)^T A(x-u) + v^T x = \min \ \text{s. t. } x \in \operatorname{bild} L, \\
\text{(a)} & q(y) = \frac{1}{2}(y-v)^T A^{-1}(y-v) + u^T y = \min \ \text{s. t. } y \in \operatorname{cobild} L.
\end{array}
$$

Beweis:
In der Notation von (1.11) für die zu L gehörigen Räume ist statt (3.11) zu lösen

$$\tilde{p}(w) = \tfrac{1}{2}(\hat{Z}w - u)^T A(\hat{Z}w - u) + v^T \hat{Z}w = \min$$

und

$$\tilde{q}(w) = \tfrac{1}{2}(Zw - v)^T A^{-1}(Zw - v) + u^T Zw = \min \ .$$

Für das erste Problem erhält man mit $x = \hat{Z}w$ in der Notation von Satz 1.44

$$
\begin{aligned}
\tilde{p}'(w) = (\hat{Z}w - u)^T A\hat{Z} + v^T\hat{Z} = 0 & \\
\Longrightarrow \quad & \hat{Z}^T A\hat{Z}w = \hat{Z}^T(Au - v) \\
\Longrightarrow \quad & \hat{Z}\hat{Z}^T A\hat{Z}w = \hat{Z}\hat{Z}^T(Au - v) \\
\Longrightarrow \quad & LL^+ALL^+\hat{Z}w = LL^+(Au - v) \\
\Longrightarrow \quad & \tilde{A}x = LL^+(Au - v) \\
\Longrightarrow \quad & x = A^L LL^+(Au - v) = A^L(Au - v),
\end{aligned}
$$

letzteres da A^L eine gleichungslösende Inverse von $\tilde{A}$ ist und die rechte Seite in $\operatorname{bild} \tilde{A}$ liegt, vgl. Satz 1.39. Wegen

$$
\begin{aligned}
&\hat{Z}^T(A(x-u) + v) = \\
&\quad = \hat{Z}^T(AA^L Au - AA^L v - Au + v) =^{\hat{Z}^T = \hat{Z}^T LL^+} \\
&\quad = \hat{Z}^T(\underbrace{LL^+AA^L}_{=LL^+} Au - LL^+Au - \underbrace{LL^+AA^L}_{=LL^+} v + LL^+v) = \\
&\quad = \hat{Z}^T(LL^+Au - LL^+Au - LL^+v + LL^+v) = 0
\end{aligned}
$$

ist $w = \hat{Z}^T x$ tatsächlich Nullstelle der Ableitung und eindeutiges Minimum, da $\hat{Z}^T A\hat{Z}$ positiv definit im quadratischen Funktional ist.

Das zweite Problem wird analog gelöst.

□

3.2 Markov-Ketten

Gegeben ist ein System, das sich in n verschiedenen Zuständen $s_1, \dots, s_n$ befinden kann. In einer Zeiteinheit geht Zustand s_j in Zustand s_i mit einer Wahrscheinlichkeit a_{ij} über. Man interessiert sich dann für das langfristige Verhalten einer so festgelegten sogenannten homogenen Markov-Kette.

Faßt man die a_{ij} in einer Matrix $A \in \mathbb{R}^{n,n}$ zusammen, so gilt

$$Ae = e, \quad A \geq 0, \quad e = (1, \dots, 1)^T. \tag{3.12}$$

Man nennt solche Matrizen auch stochastische Matrizen.

Satz 3.3
Es gilt

(a) Mit $A_1, \dots, A_k \in \mathbb{R}^{n,n}$ sind auch $A_1 \cdot \dots \cdot A_k$ und $\frac{1}{k}(A_1 + \dots + A_k)$ stochastisch.

(b) Ist A stochastisch, so ist $\varrho(A) = 1$ Eigenwert von A.

(c) Ist A stochastisch und λ Eigenwert von A mit $|\lambda| = 1$, so sind algebraische und geometrische Vielfachheit von λ gleich.

(d) Ist A stochastisch, so existiert der Grenzwert $\lim_{k\to\infty} \frac{1}{k}(A + A^2 + \dots + A^k)$ und ist ebenfalls stochastisch.

Beweis:

(a) Sind $A_1, \dots, A_k \in \mathbb{R}^{n,n}$ stochastisch, so ist offensichtlich $A_1 \cdot \dots \cdot A_k \geq 0$ und $\frac{1}{k}(A_1 + \dots + A_k) \geq 0$. Außerdem gilt

$$A_1 \cdot \dots \cdot A_k e = A_1 \cdot \dots \cdot A_{k-1} e = \dots = A_1 e = e$$

und

$$\tfrac{1}{k}(A_1 + \dots + A_k)e = \tfrac{1}{k}(A_1 e + \dots + A_k e) = \tfrac{1}{k} k e = e.$$

(b) Wegen $Ae = e$ ist $\lambda = 1$ Eigenwert von A mit Eigenvektor e, d. h. es gilt $\varrho(A) \geq 1$. Gemäß Lemma 1.28 gilt aber auch

$$\varrho(A) \leq \|A\|_\infty = \max_{i=1,\dots,n} \sum_{j=1}^{n} |a_{ij}| = \max_{i=1,\dots,n} (Ae)_i = 1$$

und somit $\varrho(A) = \lambda = 1$.

(c) Sei J eine Jordansche Normalform von A gemäß

$$T^{-1}AT = J = \operatorname{diag}(J_1, \dots, J_r),$$

$$J_i = \lambda_i I + N_i, \quad N_i = \begin{bmatrix} 0 & \varepsilon_1^{(i)} & & \\ & \ddots & \ddots & \\ & & \ddots & \varepsilon_{n_i}^{(i)} \\ & & & 0 \end{bmatrix}, \quad \varepsilon_m^{(i)} \in \{0, 1\},$$

wobei $\lambda_i \neq \lambda_j$ für $i \neq j$. Gibt es ein $k \in \{1, \ldots, r\}$ mit $|\lambda_k| = 1$ und $N_k \neq 0$, d. h. sind algebraische und geometrische Vielfachheit von λ_k verschieden, so gilt

$$J_k^l = (\lambda_k I + N_k)^l = \sum_{i=0}^{l} \binom{l}{i} \lambda_k^{l-i} N_k^i.$$

Wegen $N_k \neq 0$ enthält N_k mindestens eine 1. Damit folgt für beliebiges $l \in \mathbb{N}$

$$\|J_k^l\|_\infty \geq |\lambda_k|^l + l|\lambda_k|^{l-1} = l + 1$$

und somit

$$\begin{aligned} l + 1 \leq \|J_k^l\|_\infty \leq \|J^l\|_\infty = \|T^{-1} A^l T\|_\infty \leq \\ \leq \|T^{-1}\|_\infty \|A^l\|_\infty \|T\|_\infty = \operatorname{cond}_\infty T \end{aligned}$$

im Widerspruch zur Beschränktheit von $\operatorname{cond}_\infty T$.

(d) Durch Transformation auf Jordansche Normalform genügt es, einzelne Jordanblöcke zu untersuchen.

<u>1. Fall</u>: $\lambda = 1$
Nach Teil (c) gilt für den zugehörigen Jordanblock

$$J = \lambda I + N = I$$

und damit

$$\frac{1}{k} \sum_{i=1}^{k} J^i = \frac{1}{k} \sum_{i=1}^{k} I = I \stackrel{k \to \infty}{\to} I\,.$$

<u>2. Fall</u>: $|\lambda| = 1$, $\lambda \neq 1$
Nach Teil (c) gilt für den zugehörigen Jordanblock

$$J = \lambda I + N = \lambda I$$

und damit

$$\frac{1}{k} \sum_{i=1}^{k} J^i = \frac{1}{k} \sum_{i=1}^{k} \lambda^i I = \frac{1}{k} \frac{\lambda^{k+1} - 1}{\lambda - 1} I \stackrel{k \to \infty}{\to} 0\,,$$

letzteres wegen

$$\left| \frac{1}{k} \frac{\lambda^{k+1} - 1}{\lambda - 1} \right| \leq \frac{|\lambda|^{k+1} + 1}{k\,|\lambda - 1|} = \frac{2}{k\,|\lambda - 1|} \stackrel{k \to \infty}{\to} 0\,.$$

<u>3. Fall</u>: $|\lambda| < 1$
Für den zugehörigen Jordanblock gilt $\varrho(J) < 1$. Dies ist aber äquivalent zu $J^k \stackrel{k \to \infty}{\to} 0$. Da $I - J$ nichtsingulär ist, folgt die Behauptung aus

$$(I - J) \frac{1}{k} \sum_{i=1}^{k} J^i = \frac{1}{k} (J - J^{k+1}) \stackrel{k \to \infty}{\to} 0\,.$$

Also existiert der Grenzwert $\lim_{k \to \infty} \frac{1}{k}(A + A^2 + \cdots + A^k)$ und ist stochastisch, da jedes Folgeglied stochastisch ist.

□

Korollar 3.4
Sei $A \in \mathbb{R}^{n,n}$ stochastisch. Dann gilt

$$\lim_{k\to\infty} \frac{1}{k} \sum_{i=1}^{k} A^i = I - (I-A)(I-A)^{\#}. \tag{3.13}$$

Beweis:
Da algebraische und geometrische Vielfachheit des Eigenwertes $\lambda = 1$ von A nach Teil (c) von Satz 3.3 gleich sind, gibt es eine Jordansche Normalform von $I - A$ der Form

$$T^{-1}(I-A)T = \begin{bmatrix} C & 0 \\ 0 & 0 \end{bmatrix}$$

mit C nichtsingulär. Die Matrix $I - A$ besitzt also eine Gruppeninverse

$$(I-A)^{\#} = T \begin{bmatrix} C^{-1} & 0 \\ 0 & 0 \end{bmatrix} T^{-1}.$$

Da aber

$$T^{-1}AT = \begin{bmatrix} I-C & 0 \\ 0 & I \end{bmatrix}$$

ist und $I - C$ keinen Eigenwert 1 besitzt, folgt nach Teil (d) von Satz 3.3

$$\lim_{k\to\infty} T^{-1}\Big(\frac{1}{k}\sum_{i=1}^{k} A^i\Big)T = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}.$$

Die Behauptung folgt nun aus

$$I - T^{-1}(I-A)TT^{-1}(I-A)^{\#}T = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} - \begin{bmatrix} C & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} C^{-1} & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}.$$

$\square$

3.3 Differentiell-algebraische Gleichungen

Bei der Motivation der Drazin-Inversen wurde mit (2.12) schon die allgemeine Lösung der speziellen impliziten Differentialgleichung (2.11) angesprochen. Hier soll eine geschlossene Darstellung von Lösungen sogenannter linearer differentiell-algebraischen Gleichungen mit konstanten Koeffizienten

$$E\dot{x} = Ax + f(t), \tag{3.14}$$

wobei $E, A \in \mathbb{C}^{n,n}$ und $f \in C([a,b], \mathbb{C}^n)$, zumindest für den Fall, daß man durch Angabe eines Anfangswertes

$$x(t_0) = x_0, \quad t_0 \in [a,b], \quad x_0 \in \mathbb{C}^n \tag{3.15}$$

die Lösung eindeutig machen kann, angesprochen werden. Dies impliziert, daß das homogene Problem

$$E\dot{x} = Ax, \quad x(t_0) = 0 \tag{3.16}$$

nur die triviale Lösung besitzen soll.

Lemma 3.5
Seien $E, A \in \mathbb{C}^{n,n}$ und $\lambda \in \mathbb{C}$ derart, daß $\lambda E - A$ nichtsingulär ist. Weiter seien

$$\hat{E} = (\lambda E - A)^{-1}E, \quad \hat{A} = (\lambda E - A)^{-1}A\,. \tag{3.17}$$

Dann gilt

$$\hat{E}\hat{A} = \hat{A}\hat{E}\,. \tag{3.18}$$

Beweis:
Es gilt $\lambda\hat{E} - \hat{A} = I$ und somit

$$\hat{E}\hat{A} = \hat{E}(\lambda\hat{E} - I) = (\lambda\hat{E} - I)\hat{E} = \hat{A}\hat{E}\,.$$

□

Satz 3.6
Seien $E, A \in \mathbb{C}^{n,n}$. Das Problem (3.16) besitzt genau dann nur die triviale Lösung, wenn es ein $\lambda \in \mathbb{C}$ mit $\lambda E - A$ nichtsingulär gibt.

Beweis:
Sei $\lambda \in \mathbb{C}$ derart, daß $\lambda E - A$ nichtsingulär ist. Seien $\hat{E}, \hat{A}$ durch (3.17) festgelegt. Das Problem (3.16) hat die gleichen Lösungen wie

$$\hat{E}\dot{x} = \hat{A}x, \quad x(t_0) = 0.$$

Transformation von $\hat{E}$ auf Jordansche Normalform gemäß

$$T^{-1}\hat{E}T = \begin{bmatrix} C & 0 \\ 0 & N \end{bmatrix}$$

mit C nichtsingulär und N nilpotent liefert wegen $\hat{A} = \lambda\hat{E} - I$

$$T^{-1}\hat{A}T = \begin{bmatrix} \lambda C - I & 0 \\ 0 & \lambda N - I \end{bmatrix}.$$

Setzt man noch

$$T^{-1}x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

so kann man das transformierte Problem

$$\begin{aligned} C\dot{x}_1 &= (\lambda C - I)x_1, \quad & x_1(t_0) = 0, \\ N\dot{x}_2 &= (\lambda N - I)x_2, \quad & x_2(t_0) = 0 \end{aligned}$$

betrachten. Da C nichtsingulär ist, ist die erste Gleichung äquivalent zu einer (homogenen) gewöhnlichen Differentialgleichung. In diesem Fall folgt sofort $x_1 = 0$ als Lösung. Mit $k = \operatorname{ind} N = \operatorname{ind} \hat{E}$ folgt aus der zweiten Gleichung durch Multiplikation mit N^{k-1}, daß

$$0 = N^{k-1}(I - \lambda N)x_2 = N^{k-1}x_2$$

und damit auch $N^{k-1}\dot{x}_2 = 0$. Multiplikation mit N^{k-2} liefert dann

$$0 = N^{k-2}(I - \lambda N)x_2 = N^{k-2}x_2$$

usw. bis $x_2 = 0$. Also gibt es nur die triviale Lösung. Man beachte, daß die Anfangsbedingung $x_2(t_0) = 0$ nicht benutzt wurde.

Sei andererseits $\lambda E - A$ singulär für alle $\lambda \in \mathbb{C}$. Zu $n + 1$ paarweise verschiedenden $\lambda_i \in \mathbb{C}$ gibt es Vektoren $v_i \in \mathbb{C}^n \setminus \{0\}$ mit

$$(\lambda_i E - A)v_i = 0.$$

Da die Vektoren $v_1, \ldots, v_{n+1}$ linear abhängig sind, gibt es $\alpha_1, \ldots, \alpha_{n+1} \in \mathbb{C}$ mit

$$\sum_{i=1}^{n+1} \alpha_i v_i = 0$$

und o. E. $\alpha_{n+1} \neq 0$. Wählt man die Funktion x gemäß

$$x(t) = \sum_{i=1}^{n+1} \alpha_i e^{\lambda_i (t-t_0)} v_i,$$

so ist $x \neq 0$, außerdem $x(t_0) = 0$ und

$$E\dot{x}(t) - Ax(t) = \sum_{i=1}^{n+1} \alpha_i e^{\lambda_i (t-t_0)} (\lambda_i E - A)v_i = 0,$$

sodaß x eine nichttriviale Lösung von (3.16) ist.

□

Definition 3.7
Seien $E, A \in \mathbb{C}^{n,n}$. Man nennt das Paar (E, A) regulär genau dann, wenn das charakteristische Polynom

$$p(\lambda) = \det(\lambda E - A) \tag{3.19}$$

nicht identisch verschwindet.

□

Satz 3.8
Seien $E, A \in \mathbb{C}^{n,n}$ derart, daß (E, A) regulär ist, mit $\hat{E}, \hat{A}$ entsprechend Lemma 3.5 sowie $k = \operatorname{ind} \hat{E}$. Alle Lösungen des zu (3.14) gehörigen homogenen Problems haben die Form

$$x(t) = e^{\hat{E}^D \hat{A}(t-t_0)} \hat{E}^D \hat{E} c, \quad c \in \mathbb{C}^n. \tag{3.20}$$

Es ist x_0 ein konsistenter Anfagswert für das homogene Problem (d. h. es existiert eine Lösung des Anfangswertproblems) genau dann, wenn

$$x_0 \in \operatorname{bild} \hat{E}^k = \operatorname{bild}(\hat{E}^D \hat{E}). \tag{3.21}$$

Ist $f \in C^k([a, b], \mathbb{C}^n)$, so hat das inhomogene Problem (3.14) die spezielle Lösung

$$x(t) = \int_{t_0}^{t} e^{\hat{E}^D \hat{A}(t-s)} \hat{E}^D \hat{f}(s)\, ds - (I - \hat{E}^D \hat{E}) \sum_{i=0}^{k-1} (\hat{E}^D \hat{A})^i \hat{A}^D \hat{f}^{(i)}(t), \tag{3.22}$$

wobei $\hat{f} = (\lambda E - A)^{-1} f$. Es ist x_0 ein konsistenter Anfangswert für (3.14) genau dann, wenn

$$x_0 \in \hat{w} + \operatorname{bild}(\hat{E}^D \hat{E}) \tag{3.23}$$

mit

$$\hat{w} = -(I - \hat{E}^D \hat{E}) \sum_{i=0}^{k-1} (\hat{E}^D \hat{A})^i \hat{A}^D \hat{f}^{(i)}(t_0). \tag{3.24}$$

Die zugehörige eindeutige Lösung ist gegeben als Summe von (3.20) mit $c = x_0$ und (3.22).
Beweis:
Wie oben betrachten wir statt (3.14) die Gleichung

$$\hat{E}\dot{x} = \hat{A}x + \hat{f}(t),$$

die genau die gleichen Lösungen besitzt.

Transformation auf Jordansche Normalform von $\hat{E}$ wie im Beweis zu Satz 3.6 liefert als zu lösendes homogenes Problem

$$C\dot{x}_1 = (\lambda C - I)x_1, \quad N\dot{x}_2 = (\lambda N - I)x_2.$$

Wie dort schon gezeigt hat die zweite Gleichung nur die triviale Lösung. Die erste Gleichung hat die allgemeine Lösung

$$x_1(t) = e^{C^{-1}(\lambda C - I)(t-t_0)} c_1.$$

Zusammen kann man das in der Form

$$x(t) = T \exp\left(\begin{bmatrix} C^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \lambda C - I & 0 \\ 0 & \lambda N - I \end{bmatrix} (t - t_0) \right) \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \hat{c}$$

darstellen. Wegen $\exp(T^{-1}AT) = T^{-1}\exp(A)T$ folgt (3.20). Umgekehrt ist (3.20) wegen

$$\hat{E}\dot{x}(t) - \hat{A}x(t) = e^{\hat{E}^D\hat{A}(t-t_0)}(\hat{E}\hat{E}^D\hat{A}\hat{E}^D\hat{E} - \hat{A}\hat{E}^D\hat{E}) = 0$$

nach Lemma 2.4 Lösung des homogenen Problems. Mögliche Anfangswerte sind nun also solche $x_0 \in \mathbb{C}^n$ mit

$$x_0 = \hat{E}^D\hat{E}c, \quad c \in \mathbb{C}^n$$

und damit (3.21). Für (3.22) erhält man unter Verwendung von

$$\begin{aligned}(I - \hat{E}^D\hat{E})\hat{A}^D\hat{A} &= T\begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}\begin{bmatrix} * & 0 \\ 0 & (\lambda N - I)^{-1} \end{bmatrix}\begin{bmatrix} * & 0 \\ 0 & \lambda N - I \end{bmatrix}T^{-1} = \\ &= T\begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}T^{-1} = I - \hat{E}^D\hat{E}\end{aligned}$$

schließlich

$$\begin{aligned}\hat{E}\dot{x}(t) - \hat{A}x(t) = \hat{E}\Big[&\hat{E}^D\hat{f}(t) + \int_{t_0}^{t} e^{\hat{E}^D\hat{A}(t-s)}\hat{E}^D\hat{A}\hat{E}^D\hat{f}(s)\,ds - \\ &- (I - \hat{E}^D\hat{E})\sum_{i=0}^{k-1}(\hat{E}\hat{A}^D)^i\hat{A}^D\hat{f}^{(i+1)}(t)\Big] - \\ - \hat{A}\Big[&\int_{t_0}^{t} e^{\hat{E}^D\hat{A}(t-s)}\hat{E}^D\hat{f}(s)\,ds - \\ &- (I - \hat{E}^D\hat{E})\sum_{i=0}^{k-1}(\hat{E}\hat{A}^D)^i\hat{A}^D\hat{f}^{(i)}(t)\Big] = \\ = \hat{f}(t) &- \hat{E}(I - \hat{E}^D\hat{E})\sum_{i=0}^{k-2}(\hat{E}\hat{A}^D)^i\hat{A}^D\hat{f}^{(i+1)}(t) + \\ &+ (I - \hat{E}^D\hat{E})\sum_{i=0}^{k-1}(\hat{E}\hat{A}^D)^{i+1}\hat{A}^D\hat{f}^{(i+1)}(t) = \hat{f}(t).\end{aligned}$$

Letztendlich ergibt sich (3.23) entsprechend durch Auswerten von (3.22) gemäß $x(t_0) = \hat{w}$ dann zusammen mit (3.21).

□

3.4 Verteilungsfunktionen

Definition 3.9 (Wahrscheinlichkeitsraum)
Sei Ω eine nichtleere Menge und $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ eine σ-Algebra, d. h. gelte

$$(3.25) \qquad \begin{array}{ll} \text{(a)} & \Omega \in \mathcal{A}, \\ \text{(b)} & A \in \mathcal{A} \;\Rightarrow\; A^c = \Omega \setminus A \in \mathcal{A}, \\ \text{(c)} & A_i \in \mathcal{A},\; i \in I \subseteq \mathbb{N} \;\Rightarrow\; \bigcup_{i\in I} A_i \in \mathcal{A}, \end{array}$$

so heißt $(\Omega, \mathcal{A})$ ein meßbarer Raum. Ist weiter $P : \mathcal{A} \to [0,1]$ eine Funktion mit

$$(3.26) \qquad \begin{array}{ll} \text{(a)} & P(A) \geq 0 \text{ für alle } A \in \mathcal{A}, \\ \text{(b)} & P(\Omega) = 1, \\ \text{(c)} & A_i \in \mathcal{A},\ i \in I \subseteq \mathbb{N}, A_i \cap A_j = \emptyset \text{ für } i \neq j \\ & \qquad \Rightarrow\ P(\bigcup_{i\in I} A_i) = \sum_{i\in I} P(A_i), \end{array}$$

so heißt $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und P ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{A})$. Die Mengen $A \in \mathcal{A}$ heißen Ereignisse.

□

Wir beschränken uns hier auf den Fall $\Omega = \overline{\mathbb{R}}$, $\mathcal{A} = \mathcal{B}$, wobei $\mathcal{B}$ die sogenannte Borelsche σ-Algebra ist, d. h. die kleinste σ-Algebra, die alle Intervalle der Form $(a, b]$ enthält.

Definition 3.10 (Verteilungsfunktion)
Eine Funktion $F : \overline{\mathbb{R}} \to [0,1]$ heißt Verteilungsfunktion, wenn F rechtsstetig und monoton wachsend ist und $\lim_{x\to-\infty} F(x) = 0$ und $\lim_{x\to+\infty} F(x) = 1$ gilt.

□

Ist P ein Wahrscheinlichkeitsmaß auf $(\overline{\mathbb{R}}, \mathcal{B})$ mit $P(\mathbb{R}) = 1$, so erhält man durch

$$(3.27) \qquad F(x) = P((-\infty, x]) \text{ für } x \in \mathbb{R}$$

sowie $F(-\infty) = 0$ und $F(+\infty) = 1$ eine solche Verteilungsfunktion. Man beachte, daß man P aus F wegen

$$(3.28) \qquad P((-\infty, a]) + P((a, b]) = P((-\infty, b])$$

bzw.

$$(3.29) \qquad P((a, b]) = F(b) - F(a)$$

zurückerhalten kann.

Man würde nun gerne das Wahrscheinlichkeitsmaß P durch Transformation der Gleichverteilung auf $([0,1], \mathcal{B}_{[0,1]})$ realisieren.

Satz 3.11
Sei F eine Verteilungsfunktion. Durch

$$(3.30) \qquad F^-(y) = \inf\{x \in \overline{\mathbb{R}} \mid F(x) \geq y\}$$

wird eine (innere und äußere) Inverse $F^- : [0,1] \to \overline{\mathbb{R}}$ von F definiert. Diese ist linksstetig und monoton wachsend und hat die Eigenschaften

$$(3.31) \qquad \begin{array}{ll} \text{(a)} & F^-(F(x)) \leq x, \\ \text{(b)} & y \leq F(F^-(y)), \\ \text{(c)} & F(F^-(F(x))) = F(x), \\ \text{(d)} & F^-(F(F^-(y))) = F^-(y), \\ \text{(e)} & F^-(y) \leq x\ \Leftrightarrow\ y \leq F(x). \end{array}$$

Beweis:
Ist $y_1 \leq y_2$, so gilt

$$\begin{aligned} F^-(y_1) &= \inf\{x \in \overline{\mathbb{R}} \mid F(x) \geq y_1\} \leq \\ &= \inf\{x \in \overline{\mathbb{R}} \mid F(x) \geq y_2\} = F^-(y_2). \end{aligned}$$

Also ist F^- monoton wachsend.

Setzt man in (3.30) speziell $y = F(x)$ für $x \in \overline{\mathbb{R}}$, so ist x in der Menge und es gilt (3.31a).

Gegeben sei $y \in [0,1]$. Setzt man $x = F^-(y)$, so gibt es per Definition des Infimums eine Folge $x_n \downarrow x$ mit $F(x_n) \geq y$. Da F rechtsstetig ist, folgt $F(x) \geq y$ bzw. (3.31b).

Sei $x \in \overline{\mathbb{R}}$. Wegen der Monotonie von F folgt aus (3.31a)

$$F(F^-(F(x))) \leq F(x)$$

und für $y = F(x)$ in (3.31b)

$$F(x) \leq F(F^-(F(x))),$$

also (3.31c), analog (3.31d). Außerdem gilt (3.31e) wegen

$$\begin{aligned} F^-(y) \leq x \quad &\Longrightarrow \quad F(F^-(y)) \leq F(x) \\ &\Longrightarrow \quad y \leq F(x) \\ &\Longrightarrow \quad F^-(y) \leq F^-(F(x)) \\ &\Longrightarrow \quad F^-(y) \leq x. \end{aligned}$$

Zu zeigen bleibt die Linksstetigkeit von F^-. Sei dazu $y_n \uparrow y_0$, es gelte aber $F^-(y_n) \uparrow x_0' < x_0 = F^-(y_0)$. Dann folgt mit (3.31b)

$$y_n \leq F(F^-(y_n)) \leq F(x_0')$$

bzw. im Grenzwert

$$y_0 \leq F(x_0').$$

Mit (3.31a) ergibt sich daraus

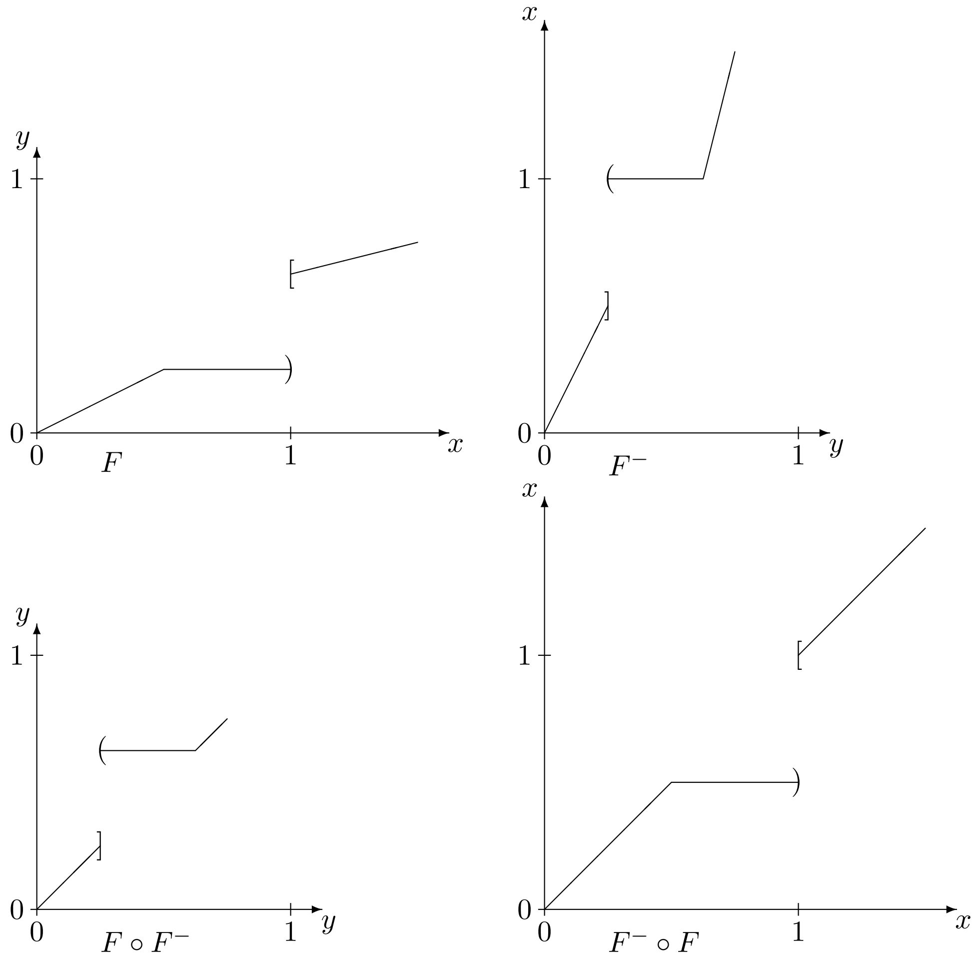
$$x_0 = F^-(y_0) \leq F^-(F(x_0')) \leq x_0'$$

im Widerspruch zur Annahme.

□

Man kann also F^- nach (3.30) auffassen als spezielle innere und äußere Inverse von F. Das spezielle dabei sind hier die Monotonie-Eigenschaften. Man beachte, daß im vorliegenden Fall zum ersten Mal eine verallgemeinerte Inverse einer nichtlinearen Abbildung behandelt wurde.

Beispiel 3.12



$\square$

Definition 3.13
Seien $(\Omega, \mathcal{A})$ und $(\Omega', \mathcal{A}')$ zwei meßbare Räume. Eine Funktion $X : \Omega \to \Omega'$ heißt meßbar genau dann, wenn $X^{-1}(A') \in \mathcal{A}$ für alle $A' \in \Omega'$. Eine Zufallsvariable ist eine auf der Stichprobenmenge Ω eines Wahrscheinlichkeitsraumes definierte meßbare Funktion. Ist $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und $X : \Omega \to \overline{\mathbb{R}}$ eine Zufallsvariable, so heißt die Funktion $F_X : \overline{\mathbb{R}} \to [0, 1]$ definiert durch

$$F_X(y) = P(X \leq y) = P(\{\omega \in \Omega \mid X(\omega) \leq y\}) \tag{3.32}$$

Verteilungsfunktion von X.

$\square$

Definition 3.14
Sei X eine Zufallsvariable bezogen auf die Räume $(\overline{\mathbb{R}}, \mathcal{B}, P)$ und $([0,1], \mathcal{B}_{[0,1]}, \lambda)$, wobei λ das Lebesgue-Maß ist. Man nennt X in $[0,1]$ gleichverteilt genau dann, wenn

$$P(X^{-1}(B)) = \lambda(B) \text{ für alle } B \in \mathcal{B}_{[0,1]}. \tag{3.33}$$

□

Satz 3.15
Sei $X : \overline{\mathbb{R}} \to [0,1]$ gleichverteilt in $[0,1]$. Sei F die zu P gehörige Verteilungsfunktion mit verallgemeinerter Inverser F^-. Dann hat die Zufallsvariable $F^-(X)$ ebenfalls F als Verteilungsfunktion.

Beweis:
Sei $Y = F^-(X)$. Dann gilt

$$\begin{aligned} F_Y(x) &= P(Y \leq x) = \\ &= P(F^-(X) \leq x) =^{(3.31e)} \\ &= P(X \leq F(x)) = \\ &= P(\{\omega \in \overline{\mathbb{R}} \mid X(\omega) \leq F(x)\}) = \\ &= P(X^{-1}(\{y \in [0,1] \mid y \leq F(x)\})) = \\ &= P(X^{-1}([0, F(x)])) =^{(3.33)} \\ &= \lambda([0, F(x)]) = F(x). \end{aligned}$$

□