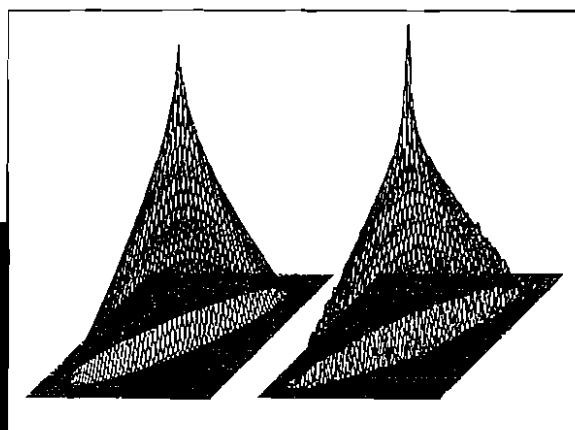


Studientexte zur Sprachkommunikation
Band 52

Arild Lacroix (Hrsg.)

**Beiträge zur Signaltheorie,
Signalverarbeitung, Sprachakustik
und Elektroakustik**

Dietrich Wolf zum 80. Geburtstag



TUD_{press}

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

Bibliografische Information der Deutschen Bibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet unter <<http://dnb.ddb.de>> abrufbar.

Bibliographic information published by Die Deutsche Bibliothek

Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at <<http://dnb.ddb.de>>

ISBN 978-3-941298-30-9

©TUDpress

Verlag der Wissenschaften GmbH

Bergstr. 70 | D-01069 Dresden

Tel.: 0351/47 96 97 20 | Fax: 0351/47 96 08 19

<http://www.tudpress.de>

Alle Rechte vorbehalten. All rights reserved.

Gedruckt aus den Manuskripten der Autoren.

Made in Germany.

Inhaltsverzeichnis

<i>A. Lacroix</i> Dietrich Wolf zum 80. Geburtstag	9
---	---

Signaltheorie

<i>P.-J. Chung und J. F. Böhme</i> Detecting the Number of Signals Using False Discovery Rate	11
--	----

<i>R. Eier</i> Zufallsfolgen mit zusammengefassten Markov-Zuständen	19
--	----

<i>A. Lacroix</i> Störabstände bei gleichförmiger und ungleichförmiger Quantisierung	35
---	----

<i>W. Mathis und T. Thiessen</i> Noise in Nonlinear Circuits and Non-Equilibrium Statistical Thermodynamics	47
--	----

<i>T. Munakata und W. Schwarz</i> Properties of Level Crossing Intervals of the Chaotic Process Generated by the Logistic Map	59
---	----

<i>E. Paulus</i> Über die bestmögliche Unterteilung der Periode bandbegrenzter, zeitdiskreter periodischer Signale in zwei zeit- und bandbegrenzte Anteile	67
--	----

<i>A. Reibiger</i> Zum Klemmenverhalten von Netzwerken	75
---	----

Signalverarbeitung

<i>M. Eichler und A. Lacroix</i> Superdirektives Beamforming: Vom Multipol-Ansatz zu einem allgemeinen Subraum-Verfahren	89
--	----

<i>F. Gollas und R. Tetzlaff</i> Verfahren zur raum-zeitlichen Analyse hirnelektrischer Aktivität bei Epilepsie mit zellularen nichtlinearen Netzwerken	106
---	-----

<i>K. Kroschel</i> Audio-Visual Analysis of Real-World Scenes	118
--	-----

<i>M. Möbius, M. Buschbeck, K. Kasper und H. Reininger</i> Klassifikation von EEG-Signalen mit Einzel- und Mehrfachereignissen durch Hidden Markov Modelle	126
--	-----

Sprachakustik

<i>S. Euler</i> Eine Experimentierumgebung zur automatischen Spracherkennung	134
<i>K. Fellbaum</i> Anwendungsgebiete der Sprachtechnologie für blinde und gehörlose Menschen	145
<i>R. Hoffmann</i> Johann Valentin Merbitz und sein sprechender Kopf – Ein Beitrag zur Frühgeschichte der multimedialen Kommunikation	154
<i>H. Höge und P. Setiawan</i> Shannon's Conditional Entropy and Error Rates on Phone Level	170
<i>D. Mehnert und R. Dietzel</i> Von Glyphen zu Tonhöhen und Intensitäten – das Boecksche Gestell, ein historisches Auswertegerät	198
<i>K. Schnell</i> Prädiktionsalgorithmen für die Sprachverarbeitung	209
<i>R. Vich</i> Komplexes Cepstrum in der Sprachsynthese	216
<i>J. Zinke</i> Unterstützung der Lehre in Audiotechnik durch Animationen, Simulationen und interaktive Laborfragen	224

Elektroakustik

<i>S. J. Rupitsch, J. Strobel und R. Lerch</i> Messung der Kavitationsaktivität mit neuartigem Ferroelektretsensor	234
<i>G. M. Sessler und J. Hillenbrand</i> New Developments in Acoustic Sensors	242

Dietrich Wolf – Daten und Dokumente

Curriculum Vitae	252
Schriftenverzeichnis	254
Laudationes und Kolloquiumsprogramm aus Anlass des 60. Geburtstags	268
Laudationes und Kolloquiumsprogramm aus Anlass des 70. Geburtstags	273
Autorenregister	276

Dietrich Wolf
geboren. Nach
dem Abitur an
1965 wurde er
einer Dissertati
Arbeit über sto
seinem Lehrer
Er hat zwei aus
Emeritierung in
80 Studierende
drei seiner Mit
ausgezeichnete
Fluktuationsph
gefassten Sinn
Rauschen in H
Fragen der Sp
Einwerbung v
Forschungsges
sich Dietrich
Selbstverwalt
Physik und er
sehr erfolgrei
Belange der U
Gesellschaft
und er hat vie
Technischen
Akademie de
der Technisc
intensiv weit
viele seiner I
zusammenge
Titel *Techni*
Wir wünsch
produktive I
Arild Lacro.

Studententexte zur Sprachkommunikation

Herausgegeben von Rüdiger Hoffmann

ISSN 0940-6832

Band 52:

**Arild Lacroix (Hrsg.): Beiträge zur Signaltheorie, Signalverarbeitung,
Sprachakustik und Elektroakustik**

ISBN 978-3-941298-30-9

Zum 80. Geburtstag des Physikers, Forschers und akademischen Lehrers Dietrich Wolf widmen ihm Schüler, Kollegen und langjährige Weggefährten diese Festschrift mit wissenschaftlichen Aufsätzen. Thematisch sind diese Aufsätze vier Gebieten zugeordnet: Signaltheorie, Signalverarbeitung, Sprachakustik und Elektroakustik. Diese Gebiete charakterisieren zugleich das breite Spektrum der wissenschaftlichen Interessen und Aktivitäten von Dietrich Wolf in den zurückliegenden Jahrzehnten seines beruflichen Wirkens.

Die Spannweite der in dieser Schrift zusammengefassten Arbeiten reicht von Fachaufsätzen auf einem begrenzten Gebiet bis hin zu Übersichtsartikeln, von grundlegenden theoretischen Analysen bis zu anwendungsorientierten Beiträgen und von wissenschaftshistorischen bis zu ausbildungsrelevanten Aufsätzen. Am Ende dieser Schrift findet der Leser in dem Abschnitt »Daten und Dokumente« detaillierte Angaben über den Lebensweg und die Lebensleistung von Dietrich Wolf.

TUDpress

ZUFALLSFOLGEN MIT ZUSAMMENGEFASSTEN MARKOV-ZUSTÄNDEN

Richard Eier

Institut für Computertechnik, Technische Universität Wien

E-mail eier@ict.tuwien.ac.at

Abstract: Von herkömmlichen Markov-Ketten ausgehend, wird die Komplexität der Zufallsfolgen durch ein planvolles Zusammenfassen von Markov-Zuständen erhöht. Es werden Gruppen von Zuständen gebildet, die Übergangsmatrizen der Ketten in Blöcken strukturiert und die Berechnungen der Statistik auf den Matrizenkalkül umgestellt. Die später verwendeten Modelle werden anhand von *entarteten* Übergangsmatrizen vorbereitet. Der prinzipielle Aufbau eines Markov-Generators wird gezeigt. Es werden die Lauflängen einer Kette mit zusammengefaßten Markov-Zuständen untersucht und an einem Beispiel eine negative-Binomialverteilung demonstriert. Um stationäre Markov-Ketten zu erhalten, müssen sie gleich von Anfang an mit ihren stationären Absolutwahrscheinlichkeiten gestartet werden. Es wird das Modell des gesteuerten Umschaltens zwischen einzelnen Markov-Ketten auch theoretisch erläutert, und es wird die Reflexion von Eigenwerten und stationären Absolutwahrscheinlichkeiten der Teilketten in der Domäne der Gesamtkette konkretisiert. Als Anwendung der abstrakten Überlegungen wird beispielhaft ein pragmatisches Vorgehen beim Entwurf einer hypothetischen Simulationsaufgabe überdacht. Am Ende werden noch ein paar extreme Überlegungen zum Design der behandelten Zufallsfolgen angestellt.

1 Einleitung und Widmung

Am Anfang gab es in den 60er Jahren den Forschungsschwerpunkt „Experimentale Stochastische Prozesse“, der damals von Günther Kraus am Institut für Niederfrequenztechnik der Technischen Hochschule Wien initiiert worden war [1] und dann über mehr als 10 Jahre Förderungen von mehreren namhaften Institutionen erhalten hatte. Bei den Arbeiten ging es um die Erzeugung und die einfache Beschreibung von speziellen Zufallsprozessen, mit denen Signale oder Störungen, die in der realen Welt immer typisch zufälligen Charakter haben, für Testzwecke möglichst wirklichkeitsnah und reproduzierbar nachzubilden sind. Die Wahl war damals auf volldigitale stufenförmige Zufallsprozesse gefallen, deren stochastischer Charakter mit Hilfe einer im Hintergrund ablaufenden steuernden Markov-Kette fast unbeschränkt variiert werden konnte.

Ich selbst war nur am Rande in dieses Projekt eingebunden. Komplementär zu den experimentellen Aspekten konzentrierte ich mich bald auf die Beschreibung und Festlegung derartiger Prozesse durch so verschiedenartige Parameter wie Verteilungsgrößen, Momente von Verteilungen, Korrelations- oder Leistungsgrößen. Anfangs waren für mich einschlägige Vorarbeiten am Institut zur Korrelationsanalyse solcher Prozesse natürlich hilfreich [2].

Dietrich Wolf kannte unsere Vorhaben. Er war mehrmals zu Besuch in Wien, und wir sprachen über die Fortschritte unserer Arbeiten. Er interessierte sich auch für meine Ergebnisse, und ich durfte darüber u. a. mehrfach an seinem Institut berichten.¹

Ich war damals auf eine vielversprechende Methode gestoßen [3], mit der ich zunächst systematisch Sätze von linear unabhängigen Prozeßparametern konfigurieren und daraus dann zur Nachbildung geeignete Markov-Ketten bestimmen konnte. In der Praxis mußte ich allerdings bald die Grenzen dieser Methode erkennen, wenn die Anzahl der Markov-Zustände sehr groß oder die vorgesehenen Prozeßparameter nur mit beschränkter Genauigkeit verfü-

¹ Nach meiner Emeritierung vor 6 Jahren habe ich heute leider nicht mehr detaillierte Unterlagen zur Verfügung.

bar waren. Deshalb habe ich dann aufgehört, die Feinheiten dieses an sich mächtigen Lösungsansatzes weiter zu verfolgen.

Ich begann alsdann in lockerer Folge Strukturprobleme bei den steuernden Markov-Ketten zu studieren, bei denen weder die Anzahl der Systemzustände noch die Genauigkeit der verlangten Systemparameter essentiell waren. Die behandelten Themen ergaben sich zwanglos neben meinen hauptamtlichen Verpflichtungen jeweils auf Grund von äußeren Anregungen. Beispiele dafür waren die Nachbildung von Rauschprozessen und Störungen oder die Analyse von Übertragungsprotokollen oder die Modellierung von codierten und modulierten digitalen Datensignalen u. a. m.

Auch Dietrich Wolf versorgte mich mit interessanteren Denkanstößen, die mein mit der Zeit erlahmendes Interesse an den Markov-Prozessen von Neuem anstachelten. An zwei Themenkreise erinnere ich mich im besonderen: Zum einen ging es um die Verteilungen von Lauflängen bei den Markov-Ketten und zum anderen um das Umschalten zwischen Subprozessen vom Markov-Typ unter Vermeiden von Einschwingvorgängen. Aus gegebenem Anlaß greife ich meine damaligen Überlegungen auf und widme dem Jubilar in großer Verbundenheit diese Arbeit, die als Übersichtsbeitrag gedacht ist. Mein Verständnis für die Zusammenhänge ist allerdings in den vergangenen Jahren so gereift, daß ich heute vieles anders darstelle und beurteile, als ich es seinerzeit gemacht habe.

Zugegebenermaßen habe ich in den letzten Jahren das Studium der einschlägigen Literatur arg vernachlässigt. Ich bemühe mich daher, komplexen Sätzen über Matrizen­theorie, stochastische Matrizen oder Wahrscheinlichkeitsrechnung so weit wie möglich aus dem Weg zu gehen bzw. sie im notwendigen Ausmaß selbst zu erklären. Falls dennoch der Wunsch nach einer Einführung in diese Gebiete bestehen sollte, führe ich 2 Werke [4], [5] an, die auch für mich in der Anfangsphase sehr hilfreich waren.

Neben den allerersten Quellen für meinen Einstieg in die Thematik referenziere ich noch zwei eigene Arbeiten [9], [10] mit der Absicht, die Weiterentwicklung meines Verständnisses für die *Prozesse mit zusammengefaßten Markov-Zuständen* zu zeigen. Indessen ist bin ich nicht daran interessiert, Prioritäten an Aussagen oder Beweisen zu postulieren, sondern es geht mir bloß darum, Sachverhalte und Methoden aus meiner Sicht zu veranschaulichen.

Obwohl ich mich schon lange nicht mehr aktiv darum bemüht habe, bin ich dennoch immer wieder auf Markov-Ketten gestoßen worden. Erst unlängst habe ich anläßlich eines Dissertationengesprächs² zwei Fachbücher der letzten Jahre über Markov-Ketten und deren Anwendungen [6], [7] durchgesehen und die Thematik mit dem Kandidaten diskutiert. Überrascht hat mich, daß in keinem der Bücher meine Vision von den zusammengefaßten Markov-Zuständen ein Thema ist. Dieser Umstand war für mich letztlich wohl auch zur Themenwahl für diesen Beitrag motiviert.

Natürlich weiß ich darüber Bescheid, daß in der Literatur den sogenannten Hidden-Markov-Modellen [7], [8] viel Raum gewidmet wird, weil sie sich im Zusammenhang mit der Mustererkennung bestens bewährt und durchgesetzt haben. Die Hidden-Markov-Modelle sind als 2-stufige stochastische Prozesse konzipiert, bei denen in der ersten Stufe eine konventionelle Markov-Kette abläuft, während in der zweiten Stufe ein neuer Zufallsprozeß für die Beobachtung (Ausgabe) generiert wird, der ausschließlich vom gerade aktuellen Markov-Zustand abhängt. Die etablierte Bedeutung dieser Prozesse anerkennend, sehe ich jedoch keinen direkten Bezug zu meinen Modellen mit den zusammengefaßten Markov-Zuständen.

Die Disposition der Arbeit ist:

~ **Im zweiten Abschnitt** werden auf der Basis von bekannten Grundlagen über die Markov-Ketten Zufallsfolgen mit zusammengefaßten Markov-Zuständen betrachtet. Es werden

² Herrn Dipl.Ing.(FH) Peter Göbel, PhD Student am ACIN (Automation and Control Institute), danke ich für anregende Gespräche und fachliche Hinweise.

Modelle von 'regulären' und 'entarteten' Markov-Ketten angesprochen und zum Entwurf von Zufallsfolgen mit speziellen Eigenschaften eingesetzt. Zum einen geht es um Zufallsfolgen mit nicht notwendigerweise geometrisch verteilten Lauflängen und zum anderen um das Umschalten zwischen Markov-Ketten im stationären Betrieb. Die Grundstruktur eines Markov-Generators wird gezeigt.

- ~ **Im dritten Abschnitt** geht es um die Verteilungen der Lauflängen bei Markov-Ketten mit und ohne zusammengefaßte Zustände. Es wird eine Zufallsfolge mit zusammengefaßten Zuständen konzipiert, deren Lauflängen prinzipiell negativ-binomialverteilt sind.
- ~ **Im vierten Abschnitt** werden die Eigenschaften der Zufallsfolgen mit zusammengefaßten Markov-Zuständen theoretisch beleuchtet. Als Anwendung wird beispielhaft eine einfache Simulationsaufgabe überlegt. Als Abschluß plaziere ich ein paar spekulative Gedanken über Langzeit- und Kurzzeiteigenschaften und über extreme Konsequenzen daraus.
- ~ **Im Resümee** werden die Ziele und Ergebnisse der Arbeit nochmals gegenübergestellt.

2 Zufallsprozesse mit Markov-Ketten

2.1 Grundsätzliches über das Beschreiben von Zufallsprozessen

Wir betrachten Zufallsprozesse, die zeitdiskret sind und außerdem nur endlich viele Zustände besitzen. Die Zeit- und Zustandsparameter werden durch Indizes gekennzeichnet. Wir indizieren die Zeitgrößen mit ganzen Zahlen und verwenden für die Zustandsgrößen die natürlichen Zahlen $\in \{1 \dots r\}$, wenn wir zwischen r verschiedenen Zuständen unterscheiden. Weiters sollen die Eigenschaften der Zufallsfolgen von den absoluten Beobachtungszeitpunkten *NICHT* abhängen, die Zeitabstände zwischen den Beobachtungen können im allgemeinen jedoch von Einfluß sein.

Über eine Zufallsfolge weiß man erst dann vollständig Bescheid, wenn man alle möglichen (denkbaren) Verteilungen von Zuständen zu den verschiedensten Beobachtungszeitpunkten kennt oder darüber Auskunft geben kann. In der Praxis trägt man dem dadurch Rechnung, daß man einen endlichen Satz von Bestimmungsstücken explizit vorgibt und für alle übrigen ein Bildungsgesetz zur Verfügung stellt.

Als Beispiel gehe ich gleich auf die **Markov-Ketten** (kurz 'M-Ketten') ein, die bekanntlich statistische Abhängigkeiten innerhalb einer Zufallsfolge modellieren lassen. Solche Abhängigkeiten können durch absoluten oder bedingte Wahrscheinlichkeiten ($\Pr \sim \text{Probabilities}$) von Zuständen ($S_\nu \sim \text{State im Zeitpunkt } \nu$) oder Zustandskombinationen ausgedrückt werden. Das Besondere an den M-Ketten ist, daß die Werte der bedingte Wahrscheinlichkeiten – auf Einfachheit bedacht - nur von dem jeweils jüngsten Ereignis einer beliebig langen Vorgeschichte beeinflußt werden. Die Umstände einer M-Kette lassen sich also folgendermaßen definieren

$$\Pr[S_{n+\nu} = j | S_n = i, \dots, \text{'Vorgeschichte'}] = \Pr[S_{n+\nu} = j | S_n = i] = p_{ij}^{(\nu)} \quad (1)$$

Man spricht hier kurzerhand von den Übergangswahrscheinlichkeiten (kurz *Ü-Wahrscheinlichkeiten*) in ν Schritten und unterstützt diese Vorstellung insbesondere in der Notation der Zustandsindizes, wo der Übergang von einem Zustand i als Bedingung nach j als Folgezustand innerhalb von ν Schritten detailliert ausgewiesen sind.

Aus der Definitionsgleichung der M-Ketten folgt nach wenigen Rekursionsschritten die an sich bekannte Produktregel (2) zur Berechnung der Verbundwahrscheinlichkeiten für alle möglichen Zustandskombinationen und Abstandskonstellationen.

$$\Pr[S_n = i, S_{n+\nu_1} = j_1, S_{n+\nu_1+\nu_2} = j_2, \dots, S_{n+\nu_1+\dots+\nu_t} = j] = \Pr[S_n = i] p_{ij_1}^{(\nu_1)} p_{j_1j_2}^{(\nu_2)} \dots p_{j_{t-1}j}^{(\nu_t)} \quad (2)$$

Schließlich weiß man von den M-Ketten - und ich werde in Abschnitt 2.3 noch darauf zu sprechen kommen -, daß sich die Ü-Wahrscheinlichkeiten in mehreren Schritten direkt aus

jenen in l Schritt ableiten lassen. Mit diesem Wissen kann ich resümieren, daß die Statistik einer M-Kette durch deren \ddot{U} -Wahrscheinlichkeiten in l Schritt gemeinsam mit den Zustandswahrscheinlichkeiten im Anfangszeitpunkt vollständig bestimmt wird.

2.2 Zufallsfolgen mit zusammengefaßten Markov-Zuständen

Im folgenden erweitere ich das Modell der ursprünglichen M-Ketten dahingehend, daß mehrere Zustände einer M-Kette zusammengefaßt werden können und dieser Umstand für einen unbedarften Beobachter nicht mehr auf den ersten Blick durchschaubar ist. Das Zusammenfassen von M-Zuständen erhöht natürlich die Komplexität der Folgen, trotzdem behalten die erzeugende M-Kette und damit auch die Produktregel (2) im wesentlichen ihre Relevanz für die Statistik der Folge. Letztere ist allerdings durch die Bildung von Randverteilungen für die zusammengefaßten Zustände zu adaptieren.

Im Matrizenkalkül ist diese Aufgabe bequem zu lösen, indem man die Indizes der Zustände, zwischen denen nicht mehr unterschieden werden soll, in eigene Teilmengen I, J_1, J_2, \dots, J verlagert, damit eigene Matrizen konfiguriert und die notwendigen Additionen von Zustandswahrscheinlichkeiten auf gewöhnliche Matrizenmultiplikationen überträgt. Das Ergebnis ist:

$$\begin{aligned} \Pr[S_n \in I, S_{n+v_1} \in J_1, S_{n+v_1+v_2} \in J_2, \dots, S_{n+v_1+v_1+\dots+v_l} \in J] = \\ = q_I^T(n) P_{IJ_1}^{(v_1)} P_{J_1J_2}^{(v_2)} \dots P_{J_{l-1}J}^{(v_l)} e_J \end{aligned} \quad (3)$$

- ~ Der Zeilenvektor³ $q_I^T(n)$ enthält die Absolutwahrscheinlichkeiten der Zustände aus der Teilmenge I .
- ~ Die $P_{X,Y}^{(\mu)}$ sind *Submatrizen* („Blöcke“) der \ddot{U} -Matrix, in denen die bedingten Wahrscheinlichkeiten für die Übergänge aus einem Zustand der Teilmenge X in einen Zustand aus der Teilmenge Y systematisch angeordnet sind. Der hochgestellte Parameter μ bezieht sich auf den aktuellen Abstand zwischen den Beobachtungszeitpunkten.
- ~ Das Kettenprodukt endet mit dem Spaltenvektor e_J , einem sogenannten *All-1-Vektor*, der so viele '1' enthält, wie die im Index genannte Teilmenge J Zustände besitzt. Er ist deshalb vonnöten, weil ja auch für die Elemente der *letzten* Teilmenge J die Randverteilung gebildet werden muß.

Wenn man die Indexnummern für die Zustände so vergibt, daß dann nur mehr Zustände mit benachbarten Nummern zusammengefaßt werden, wird der \ddot{U} -Matrix der M-Kette bereits von vornherein die Blockstruktur aufgeprägt, die das Kettenprodukt (3) direkt unterstützt.

2.3 Modelle mit Markov-Ketten

2.3.1 Die Übergangsmatrix einer Markov-Kette

Die *Ü-Wahrscheinlichkeiten in 1 Schritt* bestimmen die Eigenschaften einer M-Kette und sind auch die Grundlage für deren Realisierung (s. Abschn. 2.4). Sie werden üblicherweise in einer Matrix organisiert⁴, die als die Übergangsmatrix (kurz *Ü-Matrix*) der M-Kette genannt wird.

$$P = (p_{ij}) = \begin{pmatrix} p_{11} & \dots & p_{1r} \\ \dots & \dots & \dots \\ p_{r1} & \dots & p_{rr} \end{pmatrix}, \quad 0 \leq p_{ij} \leq 1, \quad P \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad e = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad Pe = e \quad (4)$$

Die Grundeigenschaften (4) der \ddot{U} -Matrizen sind unmittelbar einsichtig:

- ~ ihre Elemente sind bedingte Wahrscheinlichkeiten und haben daher nur Werte $\in [0, 1]$,

³ Vektoren haben in meiner Vorstellung grundsätzlich *Spaltenformat*. Zeilenvektoren werden als *transponierte Vektoren* ausgedrückt und als solche explizit mit einem hochgestellten T (kurz für *Transposition*) versehen.

⁴ Wir verwenden die Konvention, daß in den Übergangsmatrizen zeilenweise die Bedingungen der Übergänge und spaltenweise deren Folgezustände grundsätzlich invariant sind.

- ~ ihre Zeilensummen sind konstant und
- ~ sie sind zudem alle gleich 1, weil ja mit der Wahrscheinlichkeit 1 auf jeden Zustand i ein Zustand j folgt.

Matrizen dieser Art werden *stochastische Matrizen* genannt und sind als solche in der einschlägigen Literatur eingehend behandelt [4].

Weniger vertraut sind vielleicht die Aussagen über die Eigenwerte einer \ddot{U} -Matrix [4]:

- ~ Alle Eigenwerte der \ddot{U} -Matrix sind vom Betrag ≤ 1 .
- ~ Jede \ddot{U} -Matrix hat einen Eigenwert $\lambda_l = 1$ (s. auch (4)).
- ~ Im regulären Fall ist $\lambda_l = 1$ ein einfacher Eigenwert und der einzige Eigenwert vom Betrag $= 1$; alle übrigen Eigenwerte haben einen Betrag < 1 .
- ~ In entarteten Fällen gibt es mehrere Eigenwerte vom Betrag $= 1$. 2 typische Fälle sind:
 - $\lambda_l = 1$ ist mehrfacher Eigenwert. Diese Situation ist typisch für die sogenannten *zerfallenden M-Ketten* (s. Abschn. 2.3.4.1).
 - Am Einheitskreis gibt es insgesamt s Eigenwerte ($\lambda_l = 1$ inkludiert). Sie sind einfach, gleich den s -ten Einheitswurzeln und stehen immer im Zusammenhang mit *periodischen Anteilen in der M-Kette* (s. Abschn. 2.3.4.2).

Zu jedem einfachen Eigenwert gibt es einen rechten und einen linken Eigenvektor. Der Name drückt aus, ob der betreffende Eigenvektor von rechts oder von links mit der Matrix multipliziert wird. Auf Grund ihrer Positionen im Produkt mit der Matrix muß der rechte Eigenvektor grundsätzlich ein Spaltenvektor und sein linkes Pendant ein Zeilenvektor sein. Ich interessiere mich insbesondere für die Eigenvektoren zum einfachen Eigenwert $\lambda_l = 1$.

$$\begin{aligned} \text{rechter Eigenvektor:} & \quad (\text{siehe Gl. (6)}) & \quad P e = 1 e \\ \text{linker Eigenvektor:} & \quad q^T = (q_1 \cdots q_r), & \quad q^T P = 1 q^T, \quad q^T e = 1 \end{aligned} \quad (5)$$

Den rechten Eigenvektor haben wir bereits in (4) kennengelernt. Er ist der *All-1-Vektor* e , den ich in diesem Kontext ohne Index schreibe. Seine Existenz ergibt sich aus der Zeilensummenkonstanz der \ddot{U} -Matrix und steht daher für alle \ddot{U} -Matrizen von vornherein fest.

Den linken Eigenvektor, den ich pro forma q^T benenne, führe ich zunächst aus Symmetriegründen an. Da wir später noch zweimal auf ihn stoßen werden, kommentiere ich den Namen vorerst nicht weiter. Die mathematische Aufgabe, ihn zu bestimmen, möchte ich jedoch umreißen, da er anders als sein rechtes Pendant in jedem Fall extra berechnet werden muß. Im Prinzip steht dafür das homogene Gleichungssystem der Eigenwertaufgabe in (5) zur Verfügung, wobei am Ende noch die Normierung von q^T in Hinblick auf die Orthogonalität mit dem rechten Eigenvektor e einbezogen werden muß.

Wenn $\lambda_l = 1$ ein mehrfacher Eigenwert ist - es liegt der Typus einer *zerfallenden M-Kette* vor -, gibt es noch weitere Paare von linear unabhängigen Eigenvektoren (vgl. 'Eigenräume');

2.3.2 Die \ddot{U} -Matrizen in mehreren Schritten

Die \ddot{U} -Matrix in μ Schritten erhält man als μ -te Potenz der Matrix in 1 Schritt.

$$P^{(\mu)} = (P^{(1)})^\mu = P^\mu, \quad \lim_{\mu \rightarrow \infty} P^\mu = P^\infty = \begin{pmatrix} q^T \\ \vdots \\ q^T \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} q^T = e q^T, \quad P^\infty P = P^\infty \quad (6)$$

Diese Potenzregel ist im Grunde genommen so zu verstehen, daß in den Produkten aus mehreren 1-schrittigen \ddot{U} -Wahrscheinlichkeiten die Randverteilungen für sämtliche *Zwischenzustände* gebildet werden.

Im Regelfall, wo abgesehen vom 'einfachen' Eigenwert $\lambda_l = 1$ alle übrigen Eigenwerte der \ddot{U} -Matrix vom Betrag < 1 sind, konvergiert der Grenzübergang in (6) und liefert am Ende eine Matrix, die durchwegs identische Zeilen hat. Es ist wohl begründet, daß die hier mit q^T bezei-

chneten Zeilen der Grenzmatrix mit dem besagten linken Eigenvektor q^T der \ddot{U} -Matrix übereinstimmen, der damit eine erste alternative Deutung erfährt.

Der Aufbau dieser Grenzmatrix zeigt, daß die Wahrscheinlichkeitsverteilung der Zustände nach unendlich vielen Schritten vom Anfangszustand des Grenzübergangs unabhängig wird. Der eingeschwungene Zustand ist erreicht. Aus diesem Grund werden die Zustandswahrscheinlichkeiten in q^T auch als die *ergodischen Grenzwahrscheinlichkeiten* der M-Kette bezeichnet.

In den entarteten Fällen, wo es mehrere Eigenwerte vom Betrag = 1 gibt, hängt das Ergebnis des Grenzübergangs entweder vom Anfangszustand oder auch von den Schrittweiten beim Grenzübergang ab. Das Ergebnis des Grenzübergangs ist jedenfalls problematisch. Mathematisch gesehen heißt das kurz, daß der Grenzübergang nicht konvergiert.

Formelmäßig ist die Struktur der Grenzmatrix ein spezielles Vektorprodukt, in dem der *All-1 Vektor* e des linken Faktors im Zuge der Multiplikation den Zeilenvektor q^T des rechten Faktors in jeder Zeile wiederholt. Wir werden diese Schreibweise auch in anderen Fällen anwenden, wenn es darum geht, Matrizen mit identischen Zeilen zu deklarieren.

2.3.3 Die absoluten Zustandswahrscheinlichkeiten und Einschwingvorgänge

Die absoluten Wahrscheinlichkeiten der Zustände einer M-Kette hängen im allgemeinen vom Zustand zum Startzeitpunkt ab. Im Zeilenvektor $q^T(m)$ zusammengefaßt, erfahren sie im Prinzip die gleichen Transformationen wie die bedingten Wahrscheinlichkeiten in der \ddot{U} -Matrix. Ihre Entwicklung (s. (7)) kann entweder Schritt für Schritt rekursiv oder en bloc mit den Potenzen der \ddot{U} -Matrix dargestellt werden.

$$q^T(m) P = q^T(m+1), \quad q^T(0) P^m = q^T(m), \quad q^T(\infty) P = q^T(\infty), \quad q^T P = q^T \quad (7)$$

Als einfaches Beispiel überlege man eine M-Kette, die prinzipiell im Zustand '1', also mit der Anfangsverteilung $q^T(0) = (1, 0, \dots, 0)$, gestartet wird. Die Verteilungen in den folgenden Schritten sind dann der Reihe nach die jeweils ersten Zeilen der Potenzen der \ddot{U} -Matrix, die in diesem Fall offenkundig den Einschwingvorgang der Zustandverteilungen repräsentieren.

Wenn schließlich die Anzahl der Schritte über alle Grenzen wächst, gehen offenbar die absoluten Zustandswahrscheinlichkeiten in die bereits erklärten *ergodischen Grenzwahrscheinlichkeiten* über, die dann unverändert für alle Zeiten gültig bleiben. Am Ende ist also ein *stationärer Betrieb* erreicht worden. Mit diesem Bezug ist es daher üblich, hier von den *stationären Absolutwahrscheinlichkeiten* zu sprechen. Jedenfalls haben wir damit eine 2. Interpretation für den linken Eigenvektor der \ddot{U} -Matrix zum Eigenwert $\lambda_1 = 1$ gefunden.

Die 3 Begriffe von *ergodischen Grenzwahrscheinlichkeiten* und *stationären Absolutwahrscheinlichkeiten* und vom *linken Eigenvektor der \ddot{U} -Matrix* sind im Prinzip äquivalent. Der erste sieht vordergründig die dynamische Entwicklung der Zustandsvektoren, der zweite betont die Sicht der stationären Abläufe und der dritte verweist auf die Methode, mit der diese Verteilungen numerisch bestimmt werden können.

Aus diesen Überlegungen ziehe ich folgende Lehren:

- ~ Im Ablauf einer M-Kette ergeben sich bei den Verteilungen der absoluten Zustandswahrscheinlichkeiten im allgemeinen zeitabhängige Einschwingvorgänge, die letztlich jedoch gegen die (*ergodischen*) *Grenzwahrscheinlichkeiten* konvergieren.
- ~ Eine stationäre M-Kette ohne Einschwingvorgänge entsteht genau dann, wenn sie sofort mit den Grenzwahrscheinlichkeiten ihrer \ddot{U} -Matrix gestartet wird.

Wir werden darauf noch in den folgenden Abschnitten zurückkommen.

2.3.4 Ausgefallene M-Kette und Modifikationen

Es ist wert, auch einen Blick auf jene M-Ketten bzw. \ddot{U} -Matrizen zu werfen, die nicht zu den Regelfällen gehören, und sie dann durch Linearkombinationen mit anderen M-Ketten in Regelfälle überzuführen.

2.3.4.1 Zerfallende Markov-Ketten

Die Einheitsmatrix in (8) erfüllt offensichtlich die Bedingungen (4) für die \ddot{U} -Matrizen. Sie besitzt aber $\lambda_I = 1$ als 3-fachen Eigenwert und beschreibt konkret eine zerfallende M-Kette, die einmal in einem Zustand begonnen, diesen für alle Zukunft nie mehr verläßt.

$$P_{zerf} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad P_{mod} = \left(\alpha \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} \pi_1 & \pi_2 & \pi_3 \end{pmatrix} \right) \quad (8)$$

mit $\alpha + \beta = 1$

Als Eigenvektor zum Eigenwerte $\lambda_I = 1$ kommt im Prinzip jeder beliebige Vektor sowohl von rechts als auch von links in Frage (vgl. 'Eigenräume'), und darunter ist natürlich auch der *All-I-Vektor* e , der bei den \ddot{U} -Matrizen seine immanente Bedeutung hat.

Ich kombiniere die zerfallende M-Kette mit einer statistisch unabhängigen Kette, indem ich deren \ddot{U} -Matrizen unter Wahrung Zeilensummen = 1 linear kombiniere. Mit der neuen \ddot{U} -Matrix P_{mod} wird jeder Zustand mit der Wahrscheinlichkeit $(1 - \alpha)$ verlassen und mit der Wahrscheinlichkeit β in einen anderen Zustand übergeführt. Mit den Werten von α und β kann man beeinflussen, inwieweit letztlich das *träge* Verharren in den einzelnen Zuständen gegenüber den statistisch unabhängigen Zustandswechseln dominieren soll. Das Ergebnis dieser Kombination ist für $\alpha < 1$ jedenfalls eine reguläre M-Kette.

2.3.4.2 Zufallsfolgen mit zufälligen Wechseln zwischen mehreren stationären M-Ketten

Ich verallgemeinere die zerfallenden M-Ketten von (8), indem ich die bisherigen Einträge in den \ddot{U} -Matrizen durch mehrreihige Matrizen von Teilprozessen mit eigenen Zustandsräumen, stochastischen \ddot{U} -Matrizen und Grenzwahrscheinlichkeiten ersetze. Als Beispiel fasse ich 3 Subprozesse ins Auge, die ich willkürlich 'I' (für 'In'), 'O' (für 'Out') und 'X' (für 'eXtra'), nenne und damit entsprechende Indizes verwalte.

$$P_{zerf1} = \begin{pmatrix} P_O & 0 & 0 \\ 0 & P_I & 0 \\ 0 & 0 & P_X \end{pmatrix} \quad P_{mod1} = \left(\begin{array}{c|c|c} \alpha_1 P_O & \beta_1 (e_O)(q_I)^T & \gamma_1 (e_O)(q_X)^T \\ \hline \alpha_2 (e_I)(q_O)^T & \beta_2 P_I & \gamma_2 (e_I)(q_X)^T \\ \hline \alpha_3 (e_X)(q_O)^T & \beta_3 (e_X)(q_I)^T & \gamma_3 P_X \end{array} \right) \quad (9)$$

mit $\alpha_z + \beta_z + \gamma_z = 1$ für $z = 1, 2, 3$

Offenbar beschreibt P_{zerf1} nach wie vor eine zerfallende M-Kette, weil eine Teilfolge, die einmal durch einen zufälligen Anfangszustand gestartet worden ist, ihren Zustandsraum nie mehr verlassen kann.

In der \ddot{U} -Matrix P_{MOD1} kombiniere ich jetzt die zerfallende M-Kette P_{zerf1} mit statistisch unabhängigen Folgen auf der Basis von Blockmatrizen mit folgenden Auflagen:

- ~ Es sollen prinzipiell Übergänge zwischen allen Teilfolgen möglich sein.
 \Rightarrow Alle Koeffizienten in der Hauptdiagonale der Linearkombination müssen < 1 sein!
- ~ Jede Teilfolge soll jedes Mal nach ihrem Start sofort in ihrer stationären (eingeschwungenen) Phase zu laufen beginnen, so als wäre sie schon unendlich lange unterwegs.
 \Rightarrow Der Wahrscheinlichkeitsvektor, mit dem eine neue Teilkette begonnen wird, muß proportional zu den Grenzwahrscheinlichkeiten dieser Kette sein.

Die \ddot{U} -Matrix P_{MOD1} in (9) ist bereits in entsprechende Blöcke strukturiert und erfüllt diese Anforderungen:

- ~ Die quadratischen Blöcke der Hauptdiagonale enthalten die \ddot{U} -Matrizen der Teilprozesse, die in der Gesamtfolge integriert werden sollen. Der jeweils eingesetzte Koeffizient entscheidet darüber, mit welcher Wahrscheinlichkeit der Ablauf in einem Teilprozeß verweilt oder komplementär dazu zu einem anderen Teilprozeß umgeschaltet wird.
- ~ Die i. a. rechteckigen Blöcke außerhalb der Hauptdiagonale beschreiben die Bedingungen, unter denen ein Zustand aus einer Kette X_1 in einen Zustand aus einer anderen Kette X_2 übergeht.

Jeder dieser Blöcke repräsentiert eine statistisch unabhängige Kette in der Weise, daß zufolge des eingesetzten *All-1-Vektors* alle Zeilen innerhalb eines Blocks identisch sind. Mit den Koeffizienten der Linearkombination wird darüber entschieden, mit welchen Wahrscheinlichkeiten das Umschalten zwischen den einzelnen Teilketten erfolgen oder vielleicht auch unterbleiben soll.

Diese phänomenologische Beschreibung des Umschaltens zwischen den stationären M-Ketten wird in Abschn. 4 noch vertieft.

2.3.4.3 Markov-Folgen mit periodischem Charakter

Eine 3-reihige zyklische Matrix P_{zykl} hat als Eigenwerte die 3. Einheitswurzeln. Sie liegen alle am Einheitskreis, sind einfach und schließen natürlich den obligaten Eigenwert $\lambda_1 = 1$ ein.

Da $\lambda_1 = 1$ ein einfacher Eigenwert ist, sind seine beiden Eigenvektoren eindeutig. Der rechte steht bereits von vornherein mit e fest. Aber auch der linke Eigenvektor ist wegen der Spaltensummenkonstanz der Matrix proportional zum *All-1-Vektor* e^T (diesmal allerdings als Zeile). Mit der notwendigen Normierung folgt daraus, daß in der zyklischen Kette die *stationären Absolutwahrscheinlichkeiten* q^T der Zustände gleichverteilt sind.

$$P_{\text{zykl}} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad P_{\text{per}} = \alpha \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} + \beta \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \beta & \alpha & 0 \\ 0 & \beta & \alpha \\ \alpha & 0 & \beta \end{pmatrix} \quad \text{mit } \alpha + \beta = 1 \quad (10)$$

Dieser Fall ist ausgefallen, weil alle Potenzen der \ddot{U} -Matrix, deren Exponenten ein Vielfaches der *Periodenlänge*, hier also 3, sind, in jedem Fall die Einheitsmatrix ergeben. In Abständen der *Periodenlänge* entstehen also eigenständige zerfallende Subketten. Auf Details dazu möchte ich aber nicht mehr eingehen.

Den Ablauf der Kette kann man sich so vorstellen, daß die Indizes der Zustände der Reihe nach hochgezählt und nach Ausschöpfen des Vorrats wiederum auf 1 zurückgesetzt werden. Die Abstände zwischen den '1'-Zuständen sind gleich der Reihenanzahl der \ddot{U} -Matrix und bestimmen die Periode der Kette.

Hier kombiniere ich die zyklische \ddot{U} -Matrix mit einer zerfallenden Matrix im Verhältnis $\alpha : (\beta = (1 - \alpha))$. Wenn $\alpha \neq 0$ oder 1 ist, gibt es die kritischen Eigenwerte am Einheitskreis nicht mehr, so daß die entstehende Kette *regulär* wird. Die Werte in der Hauptdiagonale bedeuten konkret, daß der Zählerstand mit einer *geringen* Wahrscheinlichkeit $\beta = (1 - \alpha)$ einmal *NICHT* geändert wird und damit das normale zyklische Weiterzählen entfällt. Auf diese Weise werden die Abstände zwischen den Zeitpunkten der Rückkehr in den Zustand '1' zu einer statistischen Größe. Es entsteht eine neue Zufallsvariable, auf die ich aber erst in Abschn. 3.3 näher eingehen will. Später werde ich auf dieses Prozeßmodell als eine '*gestörte periodische Zustandskette*' verweisen.

2.4 Erzeugung von Zufallsprozesse mit M-Ketten

Es stehe bereits ein Zufallsgenerator für statistisch unabhängige Ereignisse zur Verfügung, bei dem in einem Zeilenvektor π^T die Auftrittswahrscheinlichkeiten der Ereignisse als Parameter eingestellt werden können und der als Output eine natürlichen Zahlen $\in \{1 \dots r\}$ liefert.

Für einen Markov-Generator (kurz M-Generator genannt) benötigen wir noch einen zeilenorganisierten Speicher, in dem wir die \ddot{U} -Matrix P der M-Kette und den Anfangsvektor $q^T(0)$ der Zustandsverteilung für das Anlaufen der Kette vorrätig halten können. Die Adressierung des Speichers erfolgt über die Indexnummern der Zustände und über die sonst nicht verwendete Adresse 0 , auf der ja die *Anfangsverteilung* $q^T(0)$ für die M-Kette abgelegt ist.

Weiters brauchen wir eine Speicherzelle für den neuen Zustandsindex, der gerade aktuell ist, und eine Abbildungsfunktion, um am Ausgang die abstrakten Indexnummern der Zustände in 'Objekte der realen Welt' umzusetzen. Bei dieser Gelegenheit läßt sich dann auch das Konzept mit den 'zusammengefaßten Markov-Zuständen' gleich einbeziehen.

Für einen Markov-Generator (kurz M-Generator genannt) benötigen wir noch einen zeilenorganisierten Speicher, in dem wir die \ddot{U} -Matrix P der M-Kette und den Anfangsvektor $q^T(0)$ der Zustandsverteilung für das Anlaufen der Kette vorrätig halten können. Die Adressierung des Speichers erfolgt über die Indexnummern der Zustände und über die sonst nicht verwendete Adresse 0 , auf der ja die *Anfangsverteilung* $q^T(0)$ für die M-Kette abgelegt ist.

Weiters brauchen wir eine Speicherzelle für den neuen Zustandsindex, der gerade aktuell ist, und eine Abbildungsfunktion, um am Ausgang die abstrakten Indexnummern der Zustände in 'Objekte der realen Welt' umzusetzen. Bei dieser Gelegenheit läßt sich dann auch das Konzept mit den 'zusammengefaßten Markov-Zuständen' gleich einbeziehen.

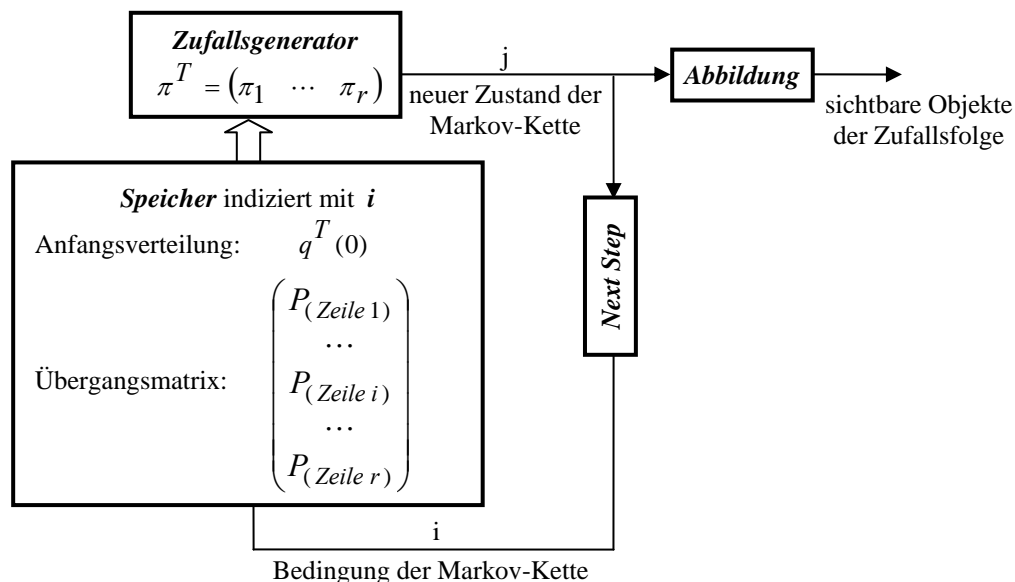


Abbildung 1 – Prinzipielle Struktur eines Markov-Generators

Als Beispiel zeigt Abbildung 1 den Aufbau eines M-Generators mit diesen Funktionsblöcken:

- ~ Der *Next-Step-Speicher* liefert den Index des aktuellen Zustands und damit die Adresse für die nächste Zeile der \ddot{U} -Matrix, die in $\pi^T(n)$ an den Zufallsgenerator übergeben wird. Dieser liefert darauf ('unter der Bedingung $\pi^T(n)$ ') den nächsten Zustand der Kette.
- ~ Der so gefundene Zustand wird in der *Next-Step-Einheit* für den nächsten Arbeitsschritt gespeichert und über die Abbildungsfunktion am Ausgang verfügbar gemacht.
- ~ Zum Start einer M-Kette liefert die *Next-Step-Einheit* die Adresse '0', so daß dem Zufallsgenerator $q^T(0)$ als Anfangsverteilung für den ersten Zustand der M-Kette übergeben wird.

Man beachte: Wenn man, wie Abschn. 2.3.4.2 angenommen, im besonderen einen Generator für *stationäre Markov-Ketten* benötigt, muß man auf dem Speicherplatz '0' ausdrücklich die (ergodischen) Grenzwahrscheinlichkeiten q^T der M-Ketten vorrätig halten!!!

3 Sequenzen von gleichen Zuständen in Markov-Ketten

3.1 Lauflängenverteilungen

Im folgenden betrachte ich Sequenzen, bei denen Zustände mit derselben Indexnummer oder mit Indizes derselben Teilmenge unmittelbar aufeinander folgen, und deren Lauflängen (kurz *LL* genannt). Die Wahrscheinlichkeiten für solche *LL* können wir direkt mit Hilfe der Produktregeln für die Verbundwahrscheinlichkeiten aus (4) bzw. (3) angeben.

Da es hier ausschließlich um Zustände im Abstand von *l* Schritt geht, lasse ich in den Formeln die darauf verweisenden Parameter ν_* weg. Außerdem teile ich den Zustandsraum der Einfachheit halber bloß in zwei Teile, in die Teilmenge 'I' mit den in besonderer Weise interessanten Zuständen, die schließlich auch detaillierter betrachtet werden, und in die Restmenge 'O', deren Umstände und Umfang so einfach wie möglich gestaltet sein sollen. Eine Sequenz von exakt *t* Elementen aus 'I' beschreibe ich formal mit $\{\{ 'I' \}^t \{ 'O' \} \}$, spreche dann von der 'Lauflänge' *t* und komme zur sogenannten 'RunLengthDistribution' *RLD*[*t*].

Ich beginne mit dem einfachen Fall einer gewöhnlichen M-Kette. Sie besitzt eine 2-reihige Ü-Matrix mit skalaren Ü-Wahrscheinlichkeiten, und die *RLD*[*t*] ergibt sich direkt aus (4).

$$P = \left(\begin{array}{c|c} p_{00} & p_{01} \\ \hline p_{10} & p_{11} \end{array} \right) \quad RLD[t] = \Pr[\{ 'I' \}^t \{ 'O' \}] = q_1(0) (p_{11})^{t-1} p_{10} \quad (11)$$

Diese *RLD*[*t*] ist eine geometrische Folge, die mit dem Quotienten $p_{11} < 1$ offensichtlich monoton abnimmt. Es ist genau diese Eigenschaft der *LL*, die man den M-Ketten vorwirft [8]. Für mich ist der Knackpunkt dagegen, daß zwischen den *RDL*[*t*]-Werten eine *einstufige Rekursion* existiert, so daß sich der Wert für die nächste Länge aus dem des Vorgängers einfach durch Multiplikation mit dem konstanten *Quotienten* p_{11} der Folge berechnen läßt.

Fachkundig sagt man, daß die *LL* der M-Ketten eine geometrische Verteilung haben [5].

Im aufwendigeren Fall, wo M-Zuständen zusammengefaßt sind, ist die Ü-Matrix aus Blöcken aufgebaut, und die *RLD*[*t*] ist daher nach (3) zu berechnen:

$$P = \left(\begin{array}{c|c} P_{OO} & P_{OI} \\ \hline P_{IO} & P_{II} \end{array} \right) \quad RLD[t] = \Pr[\{ 'I' \}^t \{ 'O' \}] = q_I^T(0) (P_{II})^{t-1} P_{IO} e_O \quad (12)$$

Für den Zeilenvektor $q_I^T(0)$ ist die Anfangsverteilung einzusetzen, mit der der erste Zustand aus der Teilmenge 'I' ausgewählt wird. Ansonsten ist diese *RLD*[*t*] ebenfalls eine geometrische Folge, diesmal sind allerdings die Potenzen einer Matrix, nämlich des Blocks P_{II} , im Spiel. Die Potenzen einer Matrix besitzen im allgemeinen jedoch eine höhere Komplexität als jene von Skalaren.

Es ist üblich, die Komplexität einer Folge anhand der minimalen Länge der Rekursion, die zwischen den Elementen der Folge gefunden werden kann, zu messen [4]. Wegen der bekannten linearen Abhängigkeiten zwischen den Matrixpotenzen à la Charakteristisches Polynom oder Minimalpolynom ist die Komplexität der Matrixpotenzen höchstens gleich der Reihenanzahl der Matrizen. Ich werde im nächsten Abschn.3.3 ein einschlägiges Beispiel behandeln.

3.2 Kriterium für einfache und zusammengefaßte Markov-Zustände

Ohne auf die theoretische Hintergründe weiter einzugehen, stelle ich hier fest:

- ~ Die Zustände, deren Lauflängen eine einstellige Rekursion besitzen und daher geometrisch verteilt sind, können als *elementare M-Zustände* einer Kette verstanden werden.
- ~ Wenn die Lauflängen eines Zustandes allerdings eine mehrstellige Rekursion erfordern, sind sie notwendigerweise mit *zusammengefaßten M-Zuständen* verbunden. Die Anzahl der zusammengefaßten Zustände ist (mindestens) gleich der Länge der Rekursion.

3.3 Zufallsfolgen mit 'negativ-binomialverteilten' Lauflängen

Als aussichtsreicher Kandidat für eine M-Kette mit 'nicht-geometrischverteilten' Lauflängen von Zuständen kommt das Modell der *gestörten periodischen M-Kette* à la (10) aus Abschn. 2.3.4.3 in Frage. Mit der Reihenzahl der \ddot{U} -Matrix kann die Periodenlänge der Grundkette ad libitum angepaßt werden.

Ich bleibe aber bei der in (10) vorbereiteten *gestörten periodischen M-Kette* mit 3 Zuständen und deklariere nun die beiden Teilmengen 'I' = {1, 2} und 'O' = {3}. Außerdem soll der Eintritt in die Teilmenge 'I' immer, d.h. mit der Wahrscheinlichkeit = 1, im Zustand 'I' erfolgen.

$$P = \begin{pmatrix} \alpha & \beta & 0 \\ 0 & \alpha & \beta \\ \beta & 0 & \alpha \end{pmatrix} \Rightarrow P_{II} = \begin{pmatrix} \alpha & \beta \\ 0 & \alpha \end{pmatrix}, \quad P_{OI} = (\beta \ 0), \quad P_{IO} = \begin{pmatrix} 0 \\ \beta \end{pmatrix}, \quad P_{OO} = (\alpha) \quad (13)$$

$$\alpha + \beta = 1 \qquad q_I^T = (1 \ 0) \qquad e_O = 1$$

Die *Lauflängenverteilung* $RLD_I[t]$ für die Zustände $\in 'I'$ erhält man dann direkt aus Gl. (12):

$$RLD_I(t) = (1 \ 0) \begin{pmatrix} \alpha & (1-\alpha) \\ 0 & \alpha \end{pmatrix}^{t-1} \begin{pmatrix} 0 \\ (1-\alpha) \end{pmatrix} = \binom{t-1}{\nu-1} (1-\alpha)^{t-\nu} \alpha^\nu \quad \text{mit } \nu = 2 \quad (14)$$

und der Rekursion $RLD_I(t+2) = 2(1-\alpha)^1 RLD_I(t+1) - 1(1-\alpha)^2 RLD_I(t)$

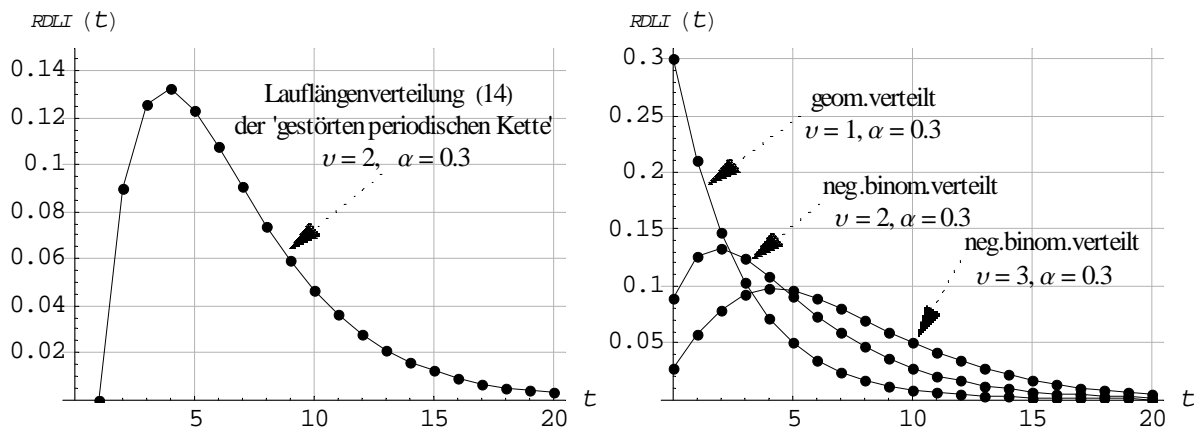


Abbildung 2 – Beispiele von Lauflängenverteilung

Ich kommentiere noch kurz dieses Ergebnis:

- ~ Der Wert $RLD_I[1] = 0$ bedeutet, daß es hier keine 'I'-Sequenz mit der $LL = 1$ gibt. Wegen des angenommenen q_I^T in (13) ist klar, daß 'I' immer im Zustand 1 betreten wird und nicht vor dem Zustand 2 wieder verlassen werden kann.
- ~ Ansonsten ist (14) im Prinzip eine sogenannte 'negative Binomialverteilung' [5]. Sie unterscheidet sich von einer einfachen geometrischen Verteilung dadurch, daß für sie eine zwei-stufige Rekursion gilt und ihre Wahrscheinlichkeitswerte ein deutliches Maximum zeigen.
- ~ Der Parameter ν in Gl. (14), der gleich die Anzahl der zusammengefaßten Zustände in 'I' anzeigt, ist dazu da, um das behandelte Beispiel direkt auf größere Teilmengen 'I' und somit auf 'negative Binomialverteilungen höherer Ordnung' erweitern zu können.
- ~ Das Beispiel läßt noch weitere Optionen offen.
 - Es kann z. B. in (13) eine andere Anfangsverteilung q_I^T eingesetzt und damit das verzögerte Einsetzen der resultierenden Verteilung übersprungen werden.
 - Man kann aber auch die bisher nicht beachtete Teilmenge 'O' mit ihren Blockmatrizen P_{OI} und P_{OO} in die Betrachtungen einbeziehen und damit andere, vielleicht gewünschte Eigenschaften für die Zufallsfolgen aktivieren.

Ich resümiere: Die *gestörten periodischen Ketten* machen es möglich, gezielt Ereignisse zu generieren, deren Abstände einer beliebigen *negative Binomialverteilung* gehorchen.

4 Umschalten zwischen stationären Markov-Ketten

4.1 Eigenschaften der geschalteten stationären Markov-Ketten

Die Grundidee, zwischen stationären M-Ketten umzuschalten, ist in Abschn. 2.4.3.2 plausibel gemacht worden. Ich möchte nun den theoretischen Hintergrund beleuchten und zeigen, wie sich die Eigenschaften der Teilketten in der Gesamtfolge widerspiegeln.

Die Herleitung erfordert allerdings eine gewisse Übung im Umgang mit Matrizen, deren Elemente nicht nur Skalare, sondern auch Vektoren oder abermals Matrizen sein können. Trotzdem möchte ich für Eingeweihte die Eckpunkte der Überlegungen skizzieren, die Ergebnisse sind dann aber wieder für jedermann brauchbar.

Die Bausteine der globalen Zufallsfolge (mit der \ddot{U} -Matrix P_{mod1} aus (9)) sind einerseits die Teilfolgen O, I, X, \dots mit ihren Zustandsräumen, \ddot{U} -Matrizen und Grenzwahrscheinlichkeiten:

$$\{P_{O|IX}, q_{O|IX}^T\} \Rightarrow P_{O|IX}^\infty = e_{O|IX} q_{O|IX}^T \quad (15)$$

Andererseits gibt es noch die Koeffizienten zur Linearkombination der Teilfolgen, die in sinnvoller Anordnung die Standards (4) von stochastischen Matrizen erfüllen. Wegen ihrer Funktion spreche ich von der 'Koppelmatrix' und schreibe für sie formal K . Sie besitzt einen eigenen Satz von Zustandsraum, \ddot{U} -Matrix und Grenzwahrscheinlichkeiten:

$$\{K = (k_{ij}), \kappa^T\} \Rightarrow K^\infty = e_3 \kappa^T \quad (16)$$

Sie enthält die bedingten Wahrscheinlichkeiten für das zufällige Umschalten zwischen den sonst voneinander unabhängigen Teilfolgen. Man sagt, daß hier das Umschalten zwischen den Teilfolgen von einer M-Kette mit der \ddot{U} -Matrix K gesteuert wird.

Im Matrizenkalkül ist es (allerdings mit Einschränkungen bezüglich der Ausführbarkeit der Rechenoperationen) erlaubt, mit Vektoren und Submatrizen genauso umzugehen wie mit Skalaren. So kann man bei der \ddot{U} -Matrix P_{mod1} in (9) Spaltenvektoren links und Zeilenvektoren rechts als gemeinsame Faktoren von Zeilen und Spalten herausheben und in Diagonalmatrizen voranstellen bzw. hinten anfügen. Wenn man das mit der Matrix von (9) konsequent macht, erhält man für die \ddot{U} -Matrix der Gesamtfolge die interessante Darstellung als Summe:

$$P_{mod1} = \text{Diag}[e_O, e_I, e_X] K \text{Diag}[q_O^T, q_I^T, q_X^T] + \\ + \text{Diag}[k_{11}(P_O - P_O^\infty), k_{22}(P_I - P_I^\infty), k_{33}(P_X - P_X^\infty)] \quad (17)$$

Der 1. Summand ist eine 'aufgeblasene' quadratische Matrix, die als Kern die Koppelmatrix K enthält. Der zweite Summand ist eine Diagonalform aus quadratischen Submatrizen, die für jede Teilfolge die Differenz zwischen ihrer \ddot{U} -Matrix und ihrer Grenzmatrix bewertet mit dem Koppelkoeffizienten k_{ii} aus der Hauptdiagonale der Koppelmatrix enthält. Man beachte, daß die Submatrizen durchwegs die Zeilensummen = 0 haben und sowohl positive als auch negative Werte enthalten können; sie sind also wohlgeformt keine stochastischen Matrizen.

Die theoretische Bedeutung des Splittings in (17) liegt darin, daß das *Matrizenprodukt* der beiden Summanden die *Null-Matrix* ergibt. Sie impliziert dadurch eine *Orthogonalität* der Summanden, die erst eine geschlossene Darstellung der Potenzen von P_{mod1} möglich macht..

$$(P_{mod1})^m = \text{Diag}[e_O, e_I, e_X] K^m \text{Diag}[q_O^T, q_I^T, q_X^T] + \\ + \text{Diag}[k_{11}^m (P_O^m - P_O^\infty), k_{22}^m (P_I^m - P_I^\infty), k_{33}^m (P_X^m - P_X^\infty)] \quad (18)$$

Aus den Matrixpotenzen folgen die Aussagen (19) über die Eigenwerte, und aus dem Grenzübergang (20) ergeben sich die Grenzwahrscheinlichkeiten von P_{mod1} .

$$EW(P_{mod1}) = \{EW(K)\} \cup \\ \cup k_{11} \{EW(P_O) \setminus 1\} \cup k_{22} \{EW(P_I) \setminus 1\} \cup k_{33} \{EW(P_X) \setminus 1\}, \quad (19)$$

$$\begin{aligned} \lim_{m \rightarrow \infty} (P_{\text{mod1}})^m &= \text{Diag}[e_O, e_I, e_X] e_3 \kappa^T \text{Diag}[q_O^T, q_I^T, q_X^T] = \\ &= e \left(\kappa_1 q_O^T, \kappa_2 q_I^T, \kappa_3 q_X^T \right) \end{aligned} \quad (20)$$

Damit bin ich am Ende des theoretischen Exkurses und resümiere nun für die Allgemeinheit:

- ~ Die Eigenwerte der Teilfolgen werden in der Gesamtfolge in abgeschwächter Form sichtbar; die Abschwächung erfolgt mit dem jeweiligen Koppelfaktor $k_{i,i}$ aus der Hauptdiagonale von K . Je größer $k_{i,i}$ ist, um so geringer ist die Abschwächung. Es kommen aber noch die Eigenwerte des Koppelprozesses hinzu.
- ~ Die stationären Absolutwahrscheinlichkeiten oder Grenzwahrscheinlichkeiten sind erwartungsgemäß proportional zu den Grenzwahrscheinlichkeiten der Teilfolgen. Als Gewichte treten die Absolutwahrscheinlichkeiten der Koppelmatrix auf.

Zur Abrundung nehme ich das Prozeßmodell nochmals differenzierter unter die Lupe:

Vom Umschalten zwischen M-Ketten kann man sprechen, wenn in der Hauptdiagonale der globalen \ddot{U} -Matrix Blöcke stehen, die selbst bereits Vielfache von stochastischen Matrizen sind. Die Blöcke außerhalb der Hauptdiagonale regeln die Übergänge zwischen den Teilketten. Wenn sie gleiche Zeilen mit Vielfachen von den Grenzwahrscheinlichkeiten der nächsten Teilkette haben, gewährleisten sie nicht nur das stationäre Ablaufen der Teilketten, sondern auch die Darstellung der resultierenden \ddot{U} -Matrix nach Gl. (18) mit allen Konsequenzen.

Die Koppelmatrix ist für die Statistik der Umschaltungen verantwortlich. Sie bestimmt einerseits mit ihren Hauptdiagonalelementen $k_{i,i}$ die Verweildauern der zugewiesenen Teilketten im eigenen Zustandsraum und steuert andererseits mit den übrigen Elementen das Wechselspiel der Übergänge zwischen den Zuständen in verschiedenen Teilketten.

Die Verweilzeiten einer Teilkette im eigenen Zustandsraum sind gleich den Lauflängen der Zustände in der steuernden M-Kette, und von diesen wissen wir, daß sie mit $k_{i,i}$ geometrisch verteilt sind (s. Abschn.3.1, Gl. (11)) und nach [5] den Erwartungswert $1/(1 - k_{i,i})$ haben.

$$K = (k_{i,j}), \quad EW[k_{i,i}] = \text{Erwartungswert}[(k_{i,i})^{t-1}(1 - k_{i,i})] = 1/(1 - k_{i,i}) \quad (21)$$

Wenn die Abschnitte m der Beobachtungen kleiner als die erwartete Verweildauer einer Teilkette in ihrem Zustandsraum sind, also $m < EW[k_{i,i}]$ zutrifft, wird man sozusagen in einer 'Kurzzeitanalyse' die Eigenwerte und Grenzwahrscheinlichkeiten der Teilkette in vollem Umfang registrieren. Im komplementären Fall, wenn also $m \gg EW[k_{i,i}]$ ist, paßt es, von 'Langzeitanalysen' zu sprechen. Es kommt zur Mittelung, bei der die individuellen Eigenschaften der einzelnen Teilketten nur mehr als Mix eingehen. Die Details dazu sind die Inhalte von (19) und (20).

Wenn ein Diagonalelement in der Koppelmatrix gegen '1' geht (vgl. etwa Abschn.2.3.4.1, Gl. (8) mit $\beta \approx 0$), tendiert die Verweildauer, die die betreffende Teilkette in ihrem Zustandsraum verbringt, gegen Unendlich, so daß schließlich Kurzzeit- und Langzeiteigenschaften dieser Teilkette ineinander übergehen.

Zum Abschluß gebe ich noch dem Splitting der \ddot{U} -Matrix in (17) eine pragmatische Deutung: Im 1. Term enthält (17) quasi ein Grundgerüst von statistisch unabhängigen Subprozessen auf der Basis der Grenzwahrscheinlichkeiten der Teilprozesse, zwischen denen nach den Regeln der Koppelmatrix K umgeschaltet wird. Der 2. Term stellt mit einer Diagonalform von quadratischen Submatrizen individuell für jede Teilkette die Anpassungen her an die aktuellen Umstände. Mit größer werdenden Abständen m geht dieses Korrektiv offensichtlich gegen 0. Daß die Elemente der Diagonalmatrix keine stochastischen Matrizen sind, ist für diese Interpretation unerheblich.

4.2 Entwurfsbeispiel für eine Zufallsfolge mit stationären Markov-Ketten

Anhand eines hypothetischen Entwurfs zeige ich in groben Zügen, wie die besprochenen Modelle angewendet werden könnten. Ich konzipiere die Simulation eines sprachähnlichen Prozesses bestehend aus stimmhaften Lauten (Teilkette 'I'), stimmlosen Lauten (Teilkette 'O') und Pausen (Teilkette 'X'). Die Dauer der einzelnen Teilprozesse und das Umschalten zwischen den Teilprozessen sollen der Detailplanung vorbehalten bleiben.

Zur Nachbildung der stimmhaften Laute wähle ich eine leicht 'gestörte periodische' oder auch 'periodische' \ddot{U} -Matrix à la P_{per} in Abschn. 2.3.4.3. Die Grenzwahrscheinlichkeiten dieser \ddot{U} -Matrizen sind in jedem Fall gleichverteilt. (Die Anzahl der Stufen entscheidet über die Periodenlänge und damit über die Tonhöhe. Als Designparameter lasse ich sie aber vorerst unbestimmt. Das gleiche gilt für den bereits in (10) vorgesehenen Parameter $\alpha \approx 1$ bzw. $\beta \approx 0$. Als Realisierung stellen ich mir ein AM-Signal vor, in dem die Indexnummern der Kette direkt auf die Amplitudenwerte abgebildet sind.)

Zur Nachbildung der stimmlosen Laute verwende ich einen rauschähnlichen Prozeß, den ich als binäre M-Kette mit statistisch unabhängigen, gleichverteilten Zuständen realisieren kann. (Auch hier wäre eine Designvariante denkbar, bei der ich zwar eine symmetrische \ddot{U} -Matrix auswähle, deren Hauptdiagonalelemente jedoch etwas unter $\frac{1}{2}$ ansetze. Psychoakustische Untersuchungen in den 70-er Jahren haben gezeigt, daß derartige Ketten von Probanden als rauschähnlicher beurteilt worden sind.)

Zur Nachbildung der Pausen genüge an sich 1 M-Zustand, trotzdem würde ich mehrere zusammengefaßte M-Zustände vorziehen, um mit ihnen im Sinne von Abschn.3.3 die Pausenlängen lokal gestalten zu können.

Zur Modellierung der Ton- und Pausenlängen steht nach Abschn. 4.1 Gl. (21) im Prinzip die Koppelmatrix zur Verfügung. Man kann damit aber auch festlegen, ob z. B. der eine oder andere Übergang bevorzugt oder gar ausgeschlossen werden soll.

Weitere Möglichkeiten dazu möchte ich aber hier nicht mehr anschnitten.

Damit sind im Prinzip alle Parameter für die Entwurfsaufgabe besprochen. Vom Anschreiben der kompletten \ddot{U} -Matrix des Gesamtprozesses nehme ich Abstand.

4.3 Perspektiven für fortgeschrittene Anwendungen

Das Aufspalten der \ddot{U} -Matrix, das in (17) bzw. (18) für die globale Zufallsfolge gezeigt worden ist, gilt natürlich auch für den einzelnen Teilprozeß. Stellvertretend betrachte ich als Beispiel die 2-reihige \ddot{U} -Matrix einer binären M-Kette:

$$\begin{aligned} (P_{\text{binär}})^m &= \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}^m = e_X q_X^T + (P_X^m - P_X^\infty) = \\ &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} q_1 & q_2 \end{pmatrix} + \begin{pmatrix} 1 & -q_2 \\ -1 & q_1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & \lambda_2^m \end{pmatrix} \begin{pmatrix} q_1 & q_2 \\ -1 & 1 \end{pmatrix} \quad -1 \leq \lambda_2 \leq 1 \end{aligned} \quad (22)$$

Hier ist es wohl anschaulicher zu argumentieren als mit dem allgemeinen Ausdruck in (18):

- Für sehr große Abstände, also mit einer *Langzeitwirkung*, dominieren natürlich die Grenzwahrscheinlichkeiten.
- Im 2. Term stecken die Abweichungen von der Grenzverteilung, die von den Abständen m der Beobachtungen abhängen. Damit diese Abweichungen über größere Abstände m , also in einer *Kurzzeitvision*, noch nennenswerte Beiträge liefern – man spricht in diesem Zusammenhang von einer 'Korrelation' -, muß $|\lambda_2|$ einigermassen groß sein. Gemessen werden diese unscharfen Aussagen an der Haltezeit, die der Teilprozeß im eigenen Zustandsraum während des Ablaufs der Gesamtfolge verbringt (s. Gl. (21)).
- Diese Aussagen gelten im Prinzip auch für mehrstufige M-Ketten, wobei dann die betragsgrößten Eigenwerte nach $\lambda_1 = 1$ die Stellung von λ_2 einnehmen.
- Je näher $|\lambda_2|$ am Einheitskreis liegt, um so nachhaltiger ist seine Wirkung über langes m .

Damit komme ich zum letzten Schritt meiner spekulativen Überlegungen:

Wenn man schon auf betragsgroße Eigenwerte bei den \ddot{U} -Matrizen Wert legt, dann verschiebe man sie doch gleich auf den Einheitskreis à la Abschn.2.3.4.3; die Teilprozesse entarten dann a priori zu periodischen Prozessen, die des Markov-Konzepts eigentlich nicht mehr bedürfen. Das Zufällige an der Gesamtfolge würde nur mehr von der steuernden M-Kette beigetragen, die damit eine von den Teilketten völlig unabhängige Bedeutung erhält. Unter diesen Umständen wäre es ohne weiteres möglich, die Haltezeiten für die Teilprozesse komplexer als geometrisch verteilt auszustatten, wobei ich da z. B. an die negativen Binomialverteilungen aus Abschn. 3.3 denke.

Das angedachte Design läuft typisch auf ein verteiltes System von deterministischen Teilprozessen hinaus, in dem die Auswahl der Teilprozesse über eine steuernde M-Kette vielleicht sogar asynchron getriggert wird. Da ich diese Visionen nicht mehr weiterspinne, möchte ich als Abschluß nur noch auf verwandte Architekturkonzepte aus anderen Bereichen aufmerksam machen. Für mich liegen da natürlich die Modelle von codierten und modulierten Datensignale besonders nahe, aber auch eine gewisse Affinität zu den schon erwähnten Hidden-Markov-Quellen [7], [8] scheint mir nicht mehr ausgeschlossen.

5 Resümee

Ich habe mich vornehmlich mit 2 Entwurfsaufgaben auseinandergesetzt, mit der Erzeugung von negativ binomialverteilten Lauflängen und mit dem Umschalten zwischen stationären M-Ketten. Die gemeinsame Methode ist, daß die M-Zustände in Gruppen zusammengefaßt, die \ddot{U} -Matrizen in entsprechende Blöcke eingeteilt und die Berechnungen der Statistik auf den Matrizenkalkül umgestellt werden.

Als Musterfall habe ich die Lauflängen einer 3-stufigen M-Kette behandelt, die sich ohne weiteres auf eine größere (beliebige) Anzahl von Zuständen ausbauen läßt. Das Ergebnis sind Folgen, deren Lauflängen negativ-binomialverteilt sind.

Für das Umschalten zwischen stationären M-Ketten konnte die Reflexion von Eigenwerten und stationären Absolutwahrscheinlichkeiten in der Domäne der globalen Zufallsfolge aufgezeigt und die Sichtweisen von Kurzzeit- und Langzeitanalysen begründet werden.

Die Überlegungen zu einer hypothetischen Entwurfsaufgabe bringen die abstrakte Gedankenwelt der Herleitung mit den realen Anliegen einer Anwendung in Beziehung. Die Abschweifung zu den verteilten Systemen soll darauf aufmerksam machen, daß noch etliche Fragen bei dem besprochenen Modell auf Antwort warten.

Dem Jubilar gratuliere ich zu seinem 80. Geburtstag mit den besten Wünschen.

Literatur

- [1] Kraus, G. (1967): Experimentelle stochastische Prozesse. Arch.elekt.Übertr. 21 (1967), S. 19-22.
- [2] Wehrmann, W. (1967): Korrelationsanalyse stochastischer, durch stationäre Markoffsche Ketten bestimmter Prozesse. Arch.elekt.Übertr. 21 (1967), S. 31-44.
- [3] Eier, R. (1973): Analyse und Synthese von diskreten Zufallsprozessen mit Hilfe von Markoffschen Ketten. Verband d. wissenschaftl. Gesellschaften Österreichs Verlag, Wien 1973
- [4] Gantmacher, F.R. (1966): Matrizenrechnung, Bd. 2: Spezielle Fragen und Anwendungen. 2. Aufl.. VEB Deutscher Verlag der Wissenschaften, Berlin 1966
- [5] Fisz, M. (1958): Wahrscheinlichkeitsrechnung und mathematische Statistik. VEB Deutscher Verlag der Wissenschaften, Berlin 1958
- [6] Waldmann, K.-H., Stocker U.M. (2003): Stochastische Modelle: Eine anwendungsorientierte Einführung. In: Medienreihe zur angewandten Statistik Springer, 2003

- [7] Fink G.A. (2003): Mustererkennung mit Markov-Modellen: Theorie, Praxis, Anwendungsgebiete. Teubner, 2003.
- [8] Rabiner L.R., Juang B.H. (1986): An Introduction to Hidden Markov Models. IEEE ASSP Mag., vol. 3, pp. 4-16, 1986.
- [9] Eier, R. (1979): Modellprozesse mit verschachtelten Markoff-Ketten. 3. Aachener Kolloquium Stochastische Signale. Kurzfassung S. 219-223, 1979
- [10] Eier, R. (1999): Mapped Markov Chains for Modelling Non-Markovian Processes. Proceedings of SCI 1999, p. 348 – 355, Orlando, Florida, 31st July-4th August 1999