

# Symbolische Dynamik, Punktprozesse und Markov-Ketten

Jürgen Grieser

08.01.1999

# Inhaltsverzeichnis

<b>1</b>	<b>Zwischenbetrachtung</b>	<b>2</b>
1.1	Wertdiskrete Zeitreihen . . . . .	2
1.2	Extremwerte . . . . .	3
1.2.1	Definition . . . . .	3
1.2.2	Wiederkehrzeit und Risiko . . . . .	4
<b>2</b>	<b>Punktprozesse</b>	<b>6</b>
2.1	Der Poisson-Prozeß . . . . .	7
2.1.1	Poisson-Prozeß in kontinuierlicher Zeit . . . . .	7
2.1.2	Betrachtung in diskreten Zeitschritten . . . . .	9
2.2	Nichtkonstante Ereigniswahrscheinlichkeitsdichten . . . . .	11
2.2.1	Allgemeine Formulierung . . . . .	11
2.2.2	Alterungsprozesse . . . . .	13
2.2.3	Zyklische Punktprozesse und stochastische Resonanz . . . . .	15
2.2.4	Change-Point Punktprozesse . . . . .	16
2.2.5	Betrachtung bei diskreten Zeitschritten . . . . .	16
<b>3</b>	<b>Markovketten</b>	<b>18</b>
3.1	Homogene Markovketten und Chapman-Kolmogoroff-Gleichung . . . . .	19
3.2	Beispiele . . . . .	20

# Kapitel 1

## Zwischenbetrachtung

### 1.1 Wertdiskrete Zeitreihen

Eine Zeitreihe kann theoretisch entweder einen kontinuierlichen Wertebereich oder einen diskreten Wertebereich umfassen. In der Realität umfaßt jede Beobachtungsreihe und jede daraus abgeleitete Reihe nur einen, durch die Meßgenauigkeit vorgegebenen, diskreten Wertebereich. Eine Temperaturreihe, die eine Spanne von zehn Kelvin umfaßt und mit einer Genauigkeit von 0.1 Kelvin vorliegt, kann nur 100 verschiedene Werte annehmen und umfaßt somit einen diskreten Wertebereich. Zweifelsohne ist dies jedoch nur eine Approximation einer Variable, die einen kontinuierlichen Wertebereich ausschöpft. Zur Beschreibung solcher Variablen ist eine Dynamik auf einem kontinuierlichen Wertebereich adäquat. Demgegenüber existieren Zeitreihen, die nur sehr wenige diskrete Werte annehmen können. Dies sind insbesondere Zeitreihen von Ereignissen. Solche Ereignisse können z.B. Extremereignisse sein. Erstellt man eine Zeitreihe indem man jedem Jahr, in dem mindestens ein Extremereignis aufgetreten ist, eine 1 zuordnet und allen anderen Jahren eine 0, so erhält man eine Zeitreihe, die nur diese beiden diskreten Zustände annehmen kann. Eine solche Zeitreihe muß als Realisation einer symbolischen Dynamik mit zwei Symbolen interpretiert werden. Auch Zeitreihen mit mehr als zwei Symbolen kommen durchaus nicht selten vor. So ist z.B. die Ausbruchsstärke explosiver Vulkanausbrüche beim Volcanic Explosivity Index (VEI) in Form von ganzen Zahlen gegeben. Man hat damit eine symbolische Dynamik vorgegeben mit den Werten 1 bis 7. Oft liegen lange Zeitreihen nicht mit gleichbleibender Genauigkeit vor. So sind z.B. die Wasserstände des gelben Flusses in China seit 1400 registriert. Änderungen der Meßmethode, aber auch der Abflußcharakteristika, machen die Zahlenwerte über diesen langen Zeitraum aber unvergleichbar. In einem solchen Fall ist es sinnvoll, den Wertebereich zu diskretisieren, um die vorliegende Ungenauigkeit zu minimieren. Man könnte dann unterscheiden zwischen normalen Jahren, Dürrejahren und Überflutungsjahren. Wie man diese drei Zustände sinnvoll voneinander abgrenzt ist eine Frage, auf die hier nicht eingegangen wird. Eine der traditionellen und wesentlichen Anwendungen der Symbolischen Dynamik liegt in der Bearbeitung der Frage, wie wahr-

scheinlich eine kontinuierliche Variable einen vorgegebenen Grenzwert überschreitet. Dies ist von großer Bedeutung für die Untersuchung von Extremwerten. Grundsätzlich kann jede auf einem quasi-kontinuierlichen Wertebereich definierte Zeitreihe in eine symbolische Zeitreihe konvertiert werden. Dies ist z.B. durch eine Klasseneinteilung möglich. Dabei kann man entweder feste Klassengrenzen vorgeben, oder Quantile als Klassengrenzen verwenden. So teilen z.B. Quartile den Wertebereich in vier Klassen, von denen jede gleich oft realisiert wird.

Eine weitere interessante Anwendung tritt auf, wenn man Zeitreihen von Parametern schätzt. Dies ist z.B. der Fall, wenn man den zeitlichen Verlauf der räumlichen Korrelation zwischen zwei Feldern untersucht. Diese kann entweder unsignifikant, signifikant positiv oder signifikant negativ sein. Man hat dann eine Zeitreihe aus drei Symbolen.

Will man die zeitliche Struktur von wertdiskreten Zeitreihen untersuchen, so kann man nicht die traditionellen Methoden der Zeitreihenanalyse anwenden. Dies liegt daran, daß die Residuen nicht gaußverteilt sein können, was für die traditionellen Verfahren i.A. Voraussetzung ist. Es treten nun verschiedene Fragen auf, die anhand vorliegender Zeitreihen beantwortet werden sollen:

1. Wie wahrscheinlich ist das  $k$ -malige Auftreten eines bestimmten Symbols in einer Zeitreihe der Länge  $N$ ?
2. Wie hoch ist die Wahrscheinlichkeit für das Auftreten eines bestimmten Symbols an einer bestimmten Stelle in der Zeitreihe?
3. Wie lange muß man im Mittel auf das erneute Auftreten eines Symbols warten?
4. Wie wahrscheinlich ist es, daß man  $m$  Zeitschritte warten muß, bis das Symbol erneut auftritt?

Um diese Fragen zu bearbeiten, muß die wertdiskrete Zeitreihe als Realisation einer symbolischen Dynamik aufgefaßt werden. Im einfachsten Fall existiert dabei ein Symbol, daß eintreten kann, oder nicht. Solche Prozesse, bei denen zu bestimmten Zeitpunkten Ereignisse stattfinden, heißen Punktprozesse. Falls die Eintrittswahrscheinlichkeit konstant ist, handelt es sich um einen Poisson-Prozeß, dessen zeitdiskrete Version in Anlehnung an das Bernoulli-Experiment auch Bernoulli-Prozeß genannt wird.

## 1.2 Extremwerte

### 1.2.1 Definition

Man kann einen Extremwert  $x_e$  dadurch definieren, daß es ein Wert aus einer kontinuierlichen oder diskreten Variable  $X$  ist, der eine vorgegebene Schwelle  $x_c$  überschreitet (bzw.

unterschreitet) und für den die Eintrittswahrscheinlichkeit  $p_c = p(x \geq x_c)$  klein ist. Kennt man die Wahrscheinlichkeitsdichte  $f(x, t)$  der Variablen  $X$  zur Zeit  $t$  so folgt für  $p_c$

$$p_c(t) = 1 - \int_{-\infty}^{x_c} f(x, t) dx. \quad (1.1)$$

Falls  $X$  eine zeitkontinuierliche Variable ist, ist  $p_c(t)dt$  die Wahrscheinlichkeitsdichte dafür, daß der Extremwert im Zeitintervall  $t$  bis  $t + dt$  eintritt. Bei diskreter Zeitbetrachtung, gibt  $p_c(t)$  die Wahrscheinlichkeit für das Eintreten eines Extremwertes im Zeitintervall von  $t$  bis  $t + \Delta t$  an. Man beachte, daß im zeitkontinuierlichen Fall erst das Integral über  $p_c(t)$  wieder eine Wahrscheinlichkeit ist. Diese Unterscheidung wird uns noch öfter begegnen. Gleichung (1.1) vereinfacht sich drastisch für stationäre Prozesse ohne Gedächtnis (weisses Rauschen). Für den Fall von Gaußschem weißen Rauschen mit der Standardabweichung 1 und dem Mittelwert 0 folgt

$$\begin{aligned} p_c(t) &= 1 - \int_{-\infty}^{x_c} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{\sigma^2}\right) dx \\ &= .5 \left[1 - \operatorname{erf}\left(\frac{x_c - \mu}{\sigma}\right)\right] \quad \text{für } x_c \geq \mu \\ &= .5 \left[1 + \operatorname{erf}\left(\frac{x_c - \mu}{\sigma}\right)\right] \quad \text{für } x_c \leq \mu. \end{aligned} \quad (1.2)$$

Obwohl diese Gleichung mit Hilfe der Numerical Recipies leicht lösbar ist, muß man auch in diesem einfachen Fall aufpassen, ob man sich für Maxima, Minima oder beides interessiert.

Falls die Variable  $X$  einer einfachen theoretisch beschreibbaren Verteilung wie der eben verwendeten Gaußverteilung genügt, weiß, aber instationär ist, in dem Sinn, daß Lage- und Streuparameter als Funktionen der Zeit dargestellt werden können, kann die Integration von Gleichung (1.1) immernoch durch Funktionen der Numerical Recipies genähert werden. Es gilt dann z.B. für daß varianz- und mittelwertinstationäre gaußsche weiße Rauschen

$$p_c(t) = 1 - \int_{-\infty}^{x_c} \frac{1}{\sqrt{2\pi\sigma^2(t)}} \exp\left(-\frac{(x - \mu(t))^2}{\sigma^2(t)}\right) dx. \quad (1.3)$$

Falls ein Prozeß Autokorrelation besitzt, kann er nicht mehr auf diese Weise formuliert werden. Man muß dann Information über die Vorgängerwerte berücksichtigen, denn dann ist  $f(x, t)$  eine Funktion der Verteilungen zu früheren Zeiten. Für wertdiskrete Reihen kann dies durch eine Markov-Kette beschrieben werden.

## 1.2.2 Wiederkehrzeit und Risiko

Wiederkehrzeit und Risiko sind die wesentlichen Begriffe in der klassischen Extremwertstatistik. Da das Eintreten eines Extremwertes ein zufälliges seltenes Ereignis ist, ist die Zeit bis zum erneuten Eintreten eines Extremwertes (Wartezeit) auch eine Zufallsvariable. Die Wiederkehrzeit  $\tau$  ist nun der Erwartungswert  $\tau = E(g(t))$  der Wartezeitverteilung  $g(t)$ .

Zusätzlich zum Erwartungswert als Lageparameter der Wartezeit, interessiert auch dessen Verteilung. In vielen Fällen ist die Wartezeitverteilung nur abschätzbar. Dies führt, wie in vielen anderen Bereichen, in denen Verteilungen nicht formal einfach angegeben werden können, dazu, daß man Quartile angibt. So kann man z.B. die Wahrscheinlichkeit dafür angeben, daß sich ein Extremwert innerhalb von  $k$  Jahren wiederholt. Diese Wahrscheinlichkeit heißt Risiko und folgt aus der Wartezeitverteilung  $g(t)$  als

$$R(k) = \sum_{t=1}^k g(t). \quad (1.4)$$

Das Risiko für das Eintreten eines Extremwertes innerhalb der nächsten  $k$  Zeiteinheiten folgt also in diesem diskreten Fall aus der Summe der Wahrscheinlichkeiten für alle Wartezeiten bis  $k$ .

Wenn die Eintrittswahrscheinlichkeit für Extremwerte stationär ist, dann hängen weder die Wiederkehrzeit, noch das Risiko explizit von der Zeit ab, und sind analytisch ausdrückbar. Im zeitdiskreten Fall handelt es sich dann um einen Bernoulli-Prozeß, der im folgenden Kapitel beschrieben wird.

# Kapitel 2

## Punktprozesse

Punktprozesse sind Prozesse, bei denen zu bestimmten Zeitpunkten Ereignisse stattfinden. Im einfachsten Fall erlaubt man nur eine Art von Ereignis, das entweder stattfindet (Wertzuweisung 1), oder nicht (Wertzuweisung 0). Im allgemeinen findet ein solcher Prozeß in kontinuierlicher Zeit statt. Dann ist die Wahrscheinlichkeit für das Auftreten eines Ereignisses im Zeitintervall von  $t$  bis  $t + dt$  gegeben durch die Wahrscheinlichkeitsfunktion  $f(t)dt$ . Analog können auch gemeinsame (Mehrzeiten-) Wahrscheinlichkeitsfunktionen definiert werden. Die Wahrscheinlichkeit für das Auftreten eines Ereignisses im Intervall  $[t_1, t_1 + dt]$  und eines Ereignisses im Intervall  $[t_2, t_2 + dt]$  ist dann gegeben durch die zweidimensionale Wahrscheinlichkeitsfunktion  $f(t_1, t_2) dt_1 dt_2$ . In analoger Weise kann man auch die höherdimensionalen Wahrscheinlichkeitsfunktionen definieren. Verschiedene Arten von Punktprozessen sind durch verschiedene Wahrscheinlichkeitsfunktionen definierbar. Der einfachste Punktprozeß ist der Poisson-Prozeß, der im nächsten Abschnitt vorgestellt wird.

Die Verallgemeinerung auf Punktprozesse mit mehreren unterscheidbaren Ereignissen kann leicht angegeben werden. Geht man davon aus, daß  $k = 1, 2, \dots, K$  verschiedene Ereignisse möglich sind, so ist die Wahrscheinlichkeitsfunktion für das Auftreten des Ereignisses  $i$  im Zeitintervall  $[t, t + dt]$  gegeben durch  $f_i(t)dt$ . Analog ist die Wahrscheinlichkeitsfunktion dafür, daß im Zeitintervall  $[t_1, t_1 + dt]$  ein Ereignis der Sorte  $i$  und im Zeitintervall  $[t_2, t_2 + dt]$  ein Ereignis der Sorte  $j$  auftritt, gegeben durch  $f_{i,j}(t_1, t_2) dt_1 dt_2$ .

Beobachtungen liegen i.A. zu diskreten Zeiten (bzw. über Zeitintervallen  $\Delta t$ ) vor. Dadurch kann in der Praxis i.A. von den  $dt$  zu  $\Delta t$  übergegangen werden. Dabei dürfen Ereignisse innerhalb eines Zeitschrittes nur einfach vorkommen. Um das Problem des mehrfachen Auftretens eines Ereignisses innerhalb eines Zeitschrittes zu handhaben, kann man mehrfaches Auftreten eines Ereignisses innerhalb eines Zeitschrittes als ein anderes Ereignis bezeichnen. Hat man z.B. als Zeitschritt ein Jahr, so kann man die Ereignisse *kein Hochwasser*, *ein Hochwasser* und *mehr als ein Hochwasser* unterscheiden.

## 2.1 Der Poisson-Prozeß

### 2.1.1 Poisson-Prozeß in kontinuierlicher Zeit

Der Poisson-Prozeß ist der einfachste Punktprozeß. Er findet in kontinuierlicher Zeit statt. Die Eintrittswahrscheinlichkeit eines Ereignisses im Zeitintervall  $[t, t + dt]$  ist dabei eine Konstante  $f(t)dt = \lambda dt$ . Dann ist die Wahrscheinlichkeit dafür, daß ein solches Ereignis im Zeitintervall  $[t_1, t_1 + dt]$  und im Zeitintervall  $[t_2, t_2 + dt]$  auftritt, gegeben durch  $f(t_1, t_2) dt_1 dt_2 = \lambda^2 dt_1 dt_2$ . Die gemeinsame zweidimensionale Verteilung faktorisiert demnach in die eindimensionale Verteilung, und diese ist eine Konstante. Auch alle höherdimensionalen Wahrscheinlichkeitsfunktionen faktorisieren, so daß gilt  $f(t_1, t_2, \dots, t_m) dt_1 dt_2 \dots dt_m = \lambda^m$ .

Wie hoch ist nun die Wahrscheinlichkeit  $p_k(t)$  dafür, daß im Zeitraum  $[0, t]$  genau  $k$  Ereignisse stattfinden? Um diese Frage (und alle damit zusammenhängenden) lösen zu können, muß die Master-Gleichung für den Prozeß aufgestellt werden. Die Master-Gleichung ist eine Differentialgleichung für  $p_k(t)$ . Auf ihre allgemeine Herleitung muß hier verzichtet werden, um nicht zu weit vom eigentlichen Ziel abzurücken. Die Wahrscheinlichkeit zur Zeit  $t$  genau  $k$  Ereignisse vorgefunden zu haben, ändert sich durch das Eintreten eines Ereignisses zur Zeit  $t$ . Falls also zur Zeit  $t$   $k-1$  Ereignisse eingetreten waren, so ist die Wahrscheinlichkeit, daß sich die Anzahl auf  $k$  erhöht, gerade  $p_{k-1}(t) \cdot \lambda$ . Andererseits können zur Zeit  $t$  auch schon  $k$  Ereignisse vorliegen. Dann führt das Eintreten eines Ereignisses zur Zeit  $t$  zu einer Verringerung der Wahrscheinlichkeit  $k$  Ereignisse vorzufinden. Für die Mastergleichung des Poisson-Prozesses folgt daher

$$\dot{p}_k(t) = \lambda p_{k-1}(t) - \lambda p_k(t). \quad (2.1)$$

Die Lösung dieser Gleichung ist

$$p_k(t) = \frac{(\lambda t)^k}{k!} \exp(-\lambda t). \quad (2.2)$$

Diese Lösung ist leicht zu verifizieren, indem man sie nach  $t$  ableitet und in die linke Seite von Gleichung (2.1) einsetzt.

Spätestens jetzt sollte dem interessierten Leser auffallen, warum der Prozeß Poisson-Prozeß heißt. Die Anzahl der in einem Zeitintervall der Länge  $t$  auftretenden Ereignisse ist Poisson-verteilt. Das heißt insbesondere, daß diese Anzahl für jedes beliebige Intervall nicht von der Vorgeschichte abhängt. Aus Gleichung (2.2) kann man leicht berechnen, wie wahrscheinlich es ist, daß über einen Zeitraum der Länge  $\Theta$  kein Ereignis eintritt. Diese Wahrscheinlichkeit ist gegeben durch  $p_0(\Theta)$ . Kennt man die Verteilung kann man deren Erwartungswert und Varianz berechnen. Für den Erwartungswert  $\langle k(t) \rangle$  gilt:

$$\langle k(t) \rangle = \sum_{k=0}^{\infty} k p_k(t). \quad (2.3)$$



Da der Summand für  $k = 0$  verschwindet, gilt weiter

$$\begin{aligned} \langle k(t) \rangle &= \sum_{k=1}^{\infty} k p_k(t) = \sum_{k=1}^{\infty} k \frac{(\lambda t)^k}{k!} \exp(-\lambda t) = \lambda t \sum_{k=1}^{\infty} \frac{(\lambda t)^{k-1}}{(k-1)!} \exp(-\lambda t) \\ &= \lambda t \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} \exp(-\lambda t) = \lambda t. \end{aligned} \quad (2.4)$$

Für die Varianz als zweites zentrales Moment gilt

$$\begin{aligned} \langle\langle k(t) \rangle\rangle &= \sum_{k=0}^{\infty} (k - \langle k \rangle)^2 p_k(t) \\ &= \sum_{k=0}^{\infty} k^2 p_k(t) + \sum_{k=0}^{\infty} \langle k \rangle^2 p_k(t) - 2 \sum_{k=0}^{\infty} \langle k \rangle k p_k(t) \\ &= \langle k^2 \rangle + \langle k \rangle^2 - 2 \langle k \rangle \langle k \rangle = \langle k^2 \rangle - \langle k \rangle^2. \end{aligned} \quad (2.5)$$

Für das zweite (nichtzentrierte) Moment  $\langle k^2 \rangle$  folgt

$$\begin{aligned} \langle k^2 \rangle &= \sum_{k=0}^{\infty} k^2 \frac{(\lambda t)^k}{k!} \exp(-\lambda t) = \sum_{k=1}^{\infty} k^2 \frac{(\lambda t)^k}{k!} \exp(-\lambda t) \\ &= \lambda t \sum_{k=1}^{\infty} \frac{k (\lambda t)^{k-1}}{(k-1)!} \exp(-\lambda t) = \lambda t \sum_{k=1}^{\infty} \frac{(k-1+1)(\lambda t)^{k-1}}{(k-1)!} \exp(-\lambda t) \\ &= \lambda t \left( \sum_{k=1}^{\infty} \frac{(k-1)(\lambda t)^{k-1}}{(k-1)!} \exp(-\lambda t) + \sum_{k=1}^{\infty} \frac{(\lambda t)^{k-1}}{(k-1)!} \exp(-\lambda t) \right) \\ &= \lambda t \left( \sum_{k=2}^{\infty} \frac{(\lambda t)^{k-1}}{(k-2)!} \exp(-\lambda t) + 1 \right) \\ &= \lambda t \left( \lambda t \sum_{k=2}^{\infty} \frac{(\lambda t)^{k-2}}{(k-2)!} \exp(-\lambda t) + 1 \right) = \lambda t (\lambda t + 1). \end{aligned} \quad (2.6)$$

Setzt man diesen Ausdruck in die Gleichung für die Varianz ein, so erhält man

$$\langle\langle k(t) \rangle\rangle = \lambda t \quad (2.7)$$

Nun stellt sich die Frage, wie lange man warten muß, bis ein Ereignis eintritt. Man beginnt also zu einem beliebigen Zeitpunkt  $t_0$  und fragt nach dem Zeitraum  $\Theta$ , in dem kein Ereignis auftritt, während zwischen  $\Theta$  und  $\Theta + d\Theta$  ein Ereignis auftritt. Dieser Zeitraum ist naturgemäß selbst eine Zufallsvariable, deren Verteilung  $g(t_0, \Theta) d\Theta$  zu bestimmen ist. Dazu sind zwei Wahrscheinlichkeiten zu betrachten. Einerseits, die Wahrscheinlichkeit, daß im Intervall  $[t_0, t_0 + \Theta]$  kein Ereignis stattfindet und andererseits soll dann im Intervall  $[t_0 + \Theta, t_0 + \Theta + d\Theta]$  mindestens ein Ereignis stattfinden. Letzteres hat die Wahrscheinlichkeit  $\lambda d\Theta$ . Ersteres hat die Wahrscheinlichkeit  $p_{k=0}(t_0, t_0 + \Theta)$ . Demnach ist  $g(t_0, \Theta) d\Theta$  gegeben durch

$$g(t_0, \Theta) d\Theta = \frac{(\lambda (t_0 + \Theta - t_0))^0}{0!} \exp(-\lambda \Theta) \cdot \lambda d\Theta. \quad (2.8)$$

Damit folgt, daß die Wartezeit auf das nächste Ereignis exponentialverteilt ist, mit

$$g(t_0, \Theta) = \lambda \exp(-\lambda \Theta) = g(\Theta). \quad (2.9)$$

Für den Poisson-Prozeß hängt die Wartezeit auf das nächste Ereignis also nicht vom Startzeitpunkt ab. Daraus folgt insbesondere auch, daß die Wartezeit zwischen zwei Ereignissen genau so verteilt ist, wie die Wartezeit bei beliebigem Startzeitpunkt. Desweiteren sieht man daraus, daß die Wartezeit nicht von der Wartezeit auf vorherige Ereignisse abhängt. Man beachte auch die Eigenschaft der Exponentialverteilung, daß kürzeste Wartezeiten die höchste Wahrscheinlichkeitsdichte haben. Dies mag erschrecken, zumindest aber verwundern, denn Ereignisse können z.B. Katastrophen sein. Es ist, falls diese einem Poisson-Prozeß entstammen, dann wahrscheinlicher, daß sie morgen eintreten als erst übermorgen. Andererseits kann man sich dieses Verhalten am Beispiel des Würfels verdeutlichen. Die Wahrscheinlichkeit bei einem Wurf eine sechs zu würfeln ist ein Sechstel. Die Wahrscheinlichkeit beim ersten Wurf keine sechs zu würfeln, aber beim zweiten Wurf, ist fünf Sechstel mal ein Sechstel und damit geringer.

Interessant ist nun der Erwartungswert und die Varianz der Wartezeit. Der Erwartungswert (der in der Extremwertstatistik auch Wiederkehrzeit genannt wird) ist gegeben durch

$$\begin{aligned} \langle \Theta \rangle &= \int_0^{\infty} \Theta g(\Theta) d\Theta = \lambda \int_0^{\infty} \Theta \exp(-\lambda \Theta) d\Theta. \\ &= \lambda \left[ \frac{\exp(-\lambda \Theta)}{(-\lambda)^2} (-\lambda \Theta - 1) \right]_0^{\infty} \\ &= \frac{1}{\lambda} \end{aligned} \quad (2.10)$$

Für den Poissonprozeß ist demnach der Erwartungswert für die Wartezeit auf ein Ereignis gleich dem Reziproken der Eintrittswahrscheinlichkeit. Für die Varianz folgt

$$\langle\langle \Theta^2 \rangle\rangle = \langle \Theta^2 \rangle - \langle \Theta \rangle^2 = \lambda \int_0^{\infty} \Theta^2 \exp(-\lambda \Theta) d\Theta - \langle \Theta \rangle^2. \quad (2.11)$$

Mit

$$\lambda \int_0^{\infty} \Theta^2 \exp(-\lambda \Theta) d\Theta = \left[ \exp(-\lambda \Theta) \left( \frac{\Theta^2}{-\lambda} - \frac{2\Theta}{\lambda^2} + \frac{2}{-\lambda^3} \right) \right]_0^{\infty} = \frac{2}{\lambda^2} \quad (2.12)$$

folgt daraus

$$\langle\langle \Theta^2 \rangle\rangle = \frac{1}{\lambda^2}. \quad (2.13)$$

## 2.1.2 Betrachtung in diskreten Zeitschritten

Oft betrachtet man diskrete Zeitschritte und berücksichtigt nur, ob innerhalb eines Zeitschrittes ein Ereignis eingetreten ist, oder nicht. In diesem Fall, in dem man von  $dt$  zu  $\Delta t$  übergeht, verändert sich die Beschreibung. Betrachtet man nun die Wahrscheinlichkeit  $p$  für das Auftreten eines Ereignisses innerhalb eines Zeitintervalls  $\Delta t$ , so kann der Prozeß als Bernoulli-Experiment beschrieben werden und man hat einen leichten Zugang zu den Antworten auf verschiedene Fragen.

Wie wahrscheinlich ist es, daß das Ereignis  $k$  mal in einer Zeitreihe der Länge  $N$  auftritt? Das Ereignis soll also  $k$  mal auftreten und  $N - k$  mal nicht. Dabei soll die Reihenfolge keine

Rolle spielen. Demnach folgt für die Wahrscheinlichkeit für das Auftreten von  $k$  Ereignissen bei  $N$  Zeitschritten

$$\binom{N}{k} (1-p)^{N-k} p^k. \quad (2.14)$$

Dies ist die Binomialverteilung, die bekannterweise für kleine  $p$  in die Poissonverteilung übergeht. Damit sieht man, daß wir wegen der Verwendung von makroskopischen Zeitintervallen  $\Delta t$ , auf denen wir jetzt den Prozeß definieren, vom Poisson-Prozeß abgewichen sind.

Wie hoch ist die Wahrscheinlichkeit für das  $k$ -malige Hintereinanderauftreten des Ereignisses? Aus dem eben Besprochenen folgt die Antwort einfach zu  $p^k$ , denn nach wie vor ist die Eintrittswahrscheinlichkeit eine Konstante.

Wie hoch ist die Wahrscheinlichkeit dafür, daß man  $k$  Zeitschritte warten muß bis das nächste Ereignis eintritt? In diesem Fall darf das Ereignis  $k-1$  mal nicht eintreten um danach einzutreten. Die Wahrscheinlichkeit, das es in einem Zeitschritt nicht eintritt, ist gegeben durch  $1-p$ . Demnach folgt für die Wahrscheinlichkeit, daß erst im  $k$ -ten Zeitschritt ein Ereignis eintritt  $p_k(1-p)^{k-1} p$ .

Wie lange muß man im Mittel warten, bis das Ereignis eintritt? Die Antwort auf diese Frage kann nicht leicht aus  $\sum_{k=0}^{\infty} k(1-p)^{k-1} p$  berechnet werden. Wesentlich einfacher geht dies mit Hilfe der erzeugenden Funktionen, die hier aber nicht zur Verfügung gestellt werden sollen. Man erhält dann als mittlere Wartezeit  $\tau = 1/p$  und für deren Varianz  $(1-p)/p^2 \approx 1/p^2$ . Die mittlere Wartezeit ist die Zeit, die in der Extremwertstatistik Wiederkehrzeit heißt.

Wie hoch ist die Wahrscheinlichkeit  $R_n$ , daß in den nächsten  $n$  Zeitschritten mindestens ein Ereignis stattfindet? Die Antwort auf diese Frage ist die Summe über die Wahrscheinlichkeiten dafür, daß man mindestens  $k$  Zeitschritte warten muß:

$$\begin{aligned} R_n &= \sum_{k=1}^n p_k \\ &= \sum_{k=1}^n p_k (1-p)^{k-1} p \\ &= 1 - (1-p)^n. \end{aligned} \quad (2.15)$$

Diese Größe wird in der Extremwertstatistik als Risiko bezeichnet. Sie ist mit der Wiederkehrzeit wie folgt verknüpft:

$$\tau = \frac{1}{1 - (1-R)^{1/n}} \approx \frac{-n}{\ln(1-R)} \approx \frac{n}{R}. \quad (2.16)$$

Will man z.B. mit einem Risiko von 10 % in Kauf nehmen, daß ein Ereignis innerhalb der nächsten einhundert Jahre stattfindet, dann muß die Wiederkehrzeit des Ereignisses 950 Jahre betragen, falls man es mit einem stationären Poisson-Prozeß (bzw. einem Bernoulli-Experiment als dessen zeitdiskrete Form) zu tun hat.

## 2.2 Nichtkonstante Ereigniswahrscheinlichkeitsdichten

### 2.2.1 Allgemeine Formulierung

Falls die Ereigniswahrscheinlichkeitsdichte  $f(t)$  keine Konstante ist, bekommt die Zeit eine explizite Bedeutung. Man kann zwischen stetigen und unstetigen  $f(t)$  unterscheiden. Für stetige  $f(t)$  können theoretisch immer alle Fragen, die für den Poisson-Prozeß lösbar sind, auch gelöst werden. In der Praxis kann das sehr aufwendig sein, so daß man numerische Näherungen bevorzugen könnte. Hier sollen nun allgemeine Ausdrücke für die Wahrscheinlichkeit  $k$  Ereignisse im Intervall  $t_0$  bis  $t_1$  und für die Wartezeit auf das nächste Ereignis formuliert werden. In den Abschnitten 2.2.2 bis 2.2.4 werden dann Spezialfälle besprochen.

Um die Wahrscheinlichkeit zu finden, daß  $k$  Ereignisse im Intervall von  $t_0$  bis  $t_1$  stattfinden, benötigen wir wieder die Lösungen der Mastergleichung, deren allgemeine Formulierung gegeben ist durch

$$\frac{dp_k}{dt} = \begin{cases} -f(t)p_0 & , \text{ falls } k = 0 \\ f(t)p_{k-1} - f(t)p_k & , \text{ sonst.} \end{cases} \quad (2.17)$$

Für stetige  $f(t)$  hat dieses Differentialgleichungssystem immer eindeutige Lösungen. Die allgemeinen Lösungen der inhomogenen Gleichungen folgen als Summe aus deren allgemeinen homogenen Lösungen und jeweils einer speziellen inhomogenen Lösung, die durch Variation der Konstanten aus den jeweiligen allgemeinen Lösungen der homogenen Gleichungen gefunden werden können. Das bedeutet, daß man im obigen Fall zunächst die allgemeine Lösung der homogenen Gleichung für  $p_0$

$$p_0(t) = \exp(C_0) \cdot \exp\left(-\int f(t) dt\right) = \exp(C) \cdot \exp(-F(t)). \quad (2.18)$$

angibt und die Anfangsbedingung

$$\lim_{t \rightarrow 0} p_0(t) = 1 \quad (2.19)$$

einflechtet. Diese Anfangsbedingung bedeutet, daß die Wahrscheinlichkeit, kein Ereignis im Intervall von 0 bis  $t$  vorzufinden gegen 1 geht, wenn die Intervalllänge gegen 0 geht. Dann ist  $C_0 = F(0)$  und es folgt die Lösung der homogenen Differentialgleichung unter Berücksichtigung der Anfangsbedingungen zu

$$p_0(t) = \exp(F(0) - F(t)). \quad (2.20)$$

Für  $p_1(t)$  gilt dann nach Gleichung (2.17) die inhomogene Differentialgleichung

$$\frac{dp_1(t)}{dt} = -f(t)p_1(t) + f(t)p_0(t). \quad (2.21)$$

Dabei ist der zweite Term auf der rechten Seite der inhomogene Anteil der DGL. An dieser Stelle muß die Lösung der DGL für  $p_0$  eingesetzt werden. Die allgemeine Lösung des

homogenen Anteils von Gleichung (2.21) ist analog zu Gleichung (2.20) gegeben durch

$$p_1(t) = C_1 \cdot \exp\left(\int f(t) dt\right). \quad (2.22)$$

Aus der Variation der Konstanten (d.h. setze Gleichung (2.22) mit dem Ansatz  $C_1 = C_1(t)$  in die inhomogene Differentialgleichung ein) folgt dann die Differentialgleichung für  $C_1(t)$

$$\frac{dC_1(t)}{dt} = f(t) p_0(t) \exp\left(-\int f(t) dt\right). \quad (2.23)$$

Daraus ist  $C_1(t)$  durch Integration bestimmbar. Es ist leicht einsichtig, daß das Lösungsverfahren sehr langatmig und für große  $k$  auch sehr aufwendig werden kann, denn zur Lösung der  $k$ -ten DGL muß die Lösung der  $k-1$ -ten DGL bekannt sein und als inhomogener Term eingesetzt werden. Deshalb werden solche Probleme gerne mit Hilfe von Markov-Ketten gelöst.

Einfacher ist die Frage nach der Wartezeitverteilung, d.h. nach der Wahrscheinlichkeit, das man wenn man ab dem Zeitpunkt  $t_0$  wartet, bis  $t_0 + \Theta$  warten muß, bis das erste Ereignis stattfindet. Dazu darf erstens im Zeitintervall  $t_0$  bis  $t_0 + \Theta$  kein Ereignis stattfinden, während zweitens im Intervall  $t_0 + \Theta$  bis  $t_0 + \Theta + d\Theta$  ein Ereignis stattfinden muß. Die Wartezeitwahrscheinlichkeitsdichte  $g(t_0, \Theta) d\Theta$  ist dann das Produkt der Wahrscheinlichkeit, daß zwischen  $t_0$  und  $t_0 + \Theta$  kein Ereignis stattfindet ( $p_0(t_0, t_0 + \Theta)$  aus Gleichung (2.20)) mit der Wahrscheinlichkeitsdichte dafür, daß zwischen  $t_0 + \Theta$  und  $t_0 + \Theta + d\Theta$  ein Ereignis stattfindet ( $f(t_0 + \Theta) d\Theta$ ). Damit gilt für die Wartezeitverteilungsdichte

$$g(t_0, \Theta) d\Theta = \exp(F(t_0) - F(t_0 + \Theta)) \times f(t_0 + \Theta) d\Theta. \quad (2.24)$$

Ein Vergleich von Gleichung (2.24) und den Gleichungen (2.22) und (2.23) zeigt, daß die Wartezeitverteilung eine wesentlich einfacher zugängliche Größe ist, als die Wahrscheinlichkeitsverteilung für beliebige  $k$  Ereignisse im Zeitraum von  $t_0$  bis  $t_1$ . Die Wahrscheinlichkeit, daß man, wenn man ab dem Zeitpunkt  $t_0$  auf ein Ereignis wartet, höchstens die Zeit  $T$  warten muß, folgt dann zu

$$p_g(\Theta \leq T) = \int_{t_0}^{t_0+T} g(t_0, \Theta) d\Theta. \quad (2.25)$$

Diese Größe wird in der Extremwertstatistik Risiko genannt. Sie gibt an, wie groß das Risiko ist, in der ab  $t_0$  folgenden Zeitspanne  $T$  ein Ereignis zu haben.  $1 - p_g(\Theta \leq T)$  ist dann die Wahrscheinlichkeit dafür, daß man mindestens so lange warten muß.

In den folgenden Abschnitten werden einige einfache aber nützliche Beispiele explizit besprochen.

## 2.2.2 Alterungsprozesse

Ändert sich die Ereigniswahrscheinlichkeitsdichte  $f(t)$  monoton mit der Zeit, so spricht man von einem Alterungsprozeß. Prozesse dieser Art beschreiben z.B. die Ausfallwahrscheinlichkeit von Geräten in Abhängigkeit von ihrem Alter. In der Klimatologie sind solche Beschreibungen interessant, da ein Trend im Systemzustand womöglich zu einer Alterung der Eintrittswahrscheinlichkeit besonderer (extremer) Ereignisse führen kann. Man kann zwei (einfache) Arten von Alterungsprozessen nach dem formalen Charakter der Funktion  $f(t)$  unterscheiden. Einerseits kann man  $f(t)$  als Potenz von der Zeit ausdrücken. Andererseits kann das System exponentiell altern. Diese beiden einfachen Beispiele sollen im Folgenden diskutiert werden.

### Exponentielles Altern

Für das exponentielle Altern kann man die Ereigniswahrscheinlichkeitsdichte

$$f(t) = a \exp(bt) \quad (2.26)$$

wählen. Gefragt ist nun nach der Wahrscheinlichkeitsdichte dafür, eine Zeit  $\Theta$  auf ein Ereignis warten zu müssen, wenn man zur Zeit  $t_0$  mit der Beobachtung beginnt. Dazu müssen wir die Wahrscheinlichkeitsdichte  $f(t)$  umformulieren für den Fall, daß ein Ereignis zur Zeit  $t = t_0 + \Theta$  stattfindet. Es gilt

$$f(t_0 + \Theta) = a \exp(bt_0) \exp(b\Theta) = A(t_0) \exp(b\Theta). \quad (2.27)$$

Die Formulierung ist demnach vom Anfangszeitpunkt der Beobachtung unabhängig, der Koeffizient vor der Exponentialfunktion hingegen nicht. Die Stammfunktion von  $f(t)$  ist gegeben durch

$$F(t_0 + \Theta) = \frac{A}{b} \exp b\Theta. \quad (2.28)$$

Damit folgt für die Wartezeitverteilungsdichte aus Gleichung (2.24)

$$\begin{aligned} g(t_0, \Theta)d\Theta &= \exp(F(t_0) - F(t_0 + \Theta)) \times f(t_0 + \Theta) d\Theta \\ &= A(t_0) \exp \left[ b\Theta - \frac{A(t_0)}{b} [\exp(b\Theta) - 1] \right] d\Theta. \end{aligned} \quad (2.29)$$

Nun kann der Begriff des Alterns auf zwei Arten verstanden werden. Einerseits kann die Wahrscheinlichkeit für das Eintreten eines Ereignisses zwischen zwei Ereignissen altern, um dann, wenn das Ereignis eingetreten ist wieder auf seinen Ursprungszustand zurückzufallen. Andererseits kann sich die Eintrittswahrscheinlichkeit unbeeindruckt davon, ob ein Ereignis stattfindet weiter entsprechend  $f(t)$  verhalten. Die erste Interpretation ist z.B. im Zusammenhang mit Vulkanen und Erdbeben von Bedeutung. Man kann mit einer gewissen Begründung annehmen, daß sich in einem Vulkan etwas anstaut, bis dieser ausbricht, wonach der Prozeß Zyklus erneut beginnt. Auf die gleiche Weise kann man sich vorstellen,

daß die Schubspannung an Kontinentalplatten kontinuierlich zunimmt, bis sie eine Schwelle überschreitet, wodurch es zu einem Erdbeben kommt. Danach ist die Situation wieder völlig entspannt. Diese Interpretation von Gleichung (2.29) vereinfacht die Situation, für den Fall, daß man nach der Wahrscheinlichkeitsdichte für die Zeit zwischen zwei Ereignissen fragt. Dann ist  $t_0$  nämlich immer null und  $A(t_0) = a$ . Gleichung (2.29) reduziert sich dann zur Gompertz-Verteilung

$$g(t_0, \Theta)d\Theta = a \exp \left[ b \Theta - \frac{a}{b} [\exp(b \Theta) - 1] \right] d\Theta \quad (2.30)$$

die gelegentlich auch als Gumbel-Verteilung bekannt ist.

Bei der zweiten Interpretation von Gleichung (2.29) unterstellt man ein stochastisches System, das einen langfristigen exponentiellen Trend aufweist, der auf die Ereigniswahrscheinlichkeit wirkt. Dies kann z.B. durch einen exponentiellen externen Antrieb realisiert sein. Man könnte dies als Modellvorstellung für die mögliche Zunahme von Stürmen (und anderen extremen Ereignissen) durch den exponentiell ansteigenden anthropogenen Einfluß auf das Klima unterstellen.

### Altern nach dem Potenzgesetz

Für das Altern nach dem Potenzgesetz bietet sich der Ansatz

$$f(t) = \frac{a t^{a-1}}{b^a} \text{ mit } a > 0, b > 0 \quad (2.31)$$

an. Will man die Wartezeit wieder ab einem beliebigen Startpunkt  $t_0$  messen, so folgt

$$f(t_0 + \Theta) = \frac{a(t_0 + \Theta)^{a-1}}{b^a}. \quad (2.32)$$

Für die Stammfunktion gilt dann

$$F(t_0 + \Theta) = \left( \frac{t_0 + \Theta}{b} \right)^a. \quad (2.33)$$

Dies kann man in Gleichung (2.24) einsetzen und erhält die Wartezeitverteilungsdichte

$$g(t_0 + \Theta)d\Theta = \frac{a}{b} \left( \frac{t_0 + \Theta}{b} \right)^{a-1} \exp \left[ \left( \frac{t_0}{b} \right)^a - \left( \frac{t_0 + \Theta}{b} \right)^a \right]. \quad (2.34)$$

Für den Fall, das man sich nur für die Wartezeit zwischen zwei Ereignissen (also unter klar vorgegebenen wenigen Startpunkten) interessiert und davon ausgeht daß, nach jedem Ereignis die Zeit wieder bei 0 losläuft, d.h.  $t_0 = 0$  folgt die Weibull-Verteilung

$$g(\Theta)d\Theta = \frac{a}{b} \left( \frac{\Theta}{b} \right)^{a-1} \exp \left[ - \left( \frac{\Theta}{b} \right)^a \right]. \quad (2.35)$$

Die Wahrscheinlichkeit dafür, daß zwischen zwei Ereignissen weniger als die Zeit  $T$  vergeht, ist bei einem solchen Prozeß gegeben durch

$$p(\Theta \leq T) = \int_{t_0}^{t_0+T} g(t_0, \Theta) d\Theta = 1 - \exp \left[ - \left( \frac{\Theta}{b} \right)^a \right]. \quad (2.36)$$

Für  $a = 1$  ist  $f(t)$  eine Konstante und somit geht die Betrachtung in den Poisson-Prozeß über. Die Weibull-Verteilung muß dann in die Exponentialverteilung übergehen. Das tut sie, wie man leicht sieht.

### 2.2.3 Zyklische Punktprozesse und stochastische Resonanz

Bei einem zyklischen Punktprozeß hängt die Ereigniswahrscheinlichkeitsdichte  $f(t)$  zyklisch von der Zeit ab, d.h. es gilt

$$f(t) = \lambda[1 + A \cos(2\pi\omega t)]. \quad (2.37)$$

Für beliebige Anfangszeitpunkte schreiben wir wieder

$$f(t_0 + \Theta) = \lambda + \lambda A \cos[2\pi\omega(t_0 + \Theta)] dt. \quad (2.38)$$

Dann gilt für die Stammfunktion

$$F(t_0 + \Theta) = \lambda(t_0 + \Theta) + \frac{A\lambda}{2\pi\omega} \sin[2\pi\omega(t_0 + \Theta)]. \quad (2.39)$$

Damit folgt für die Wartezeitverteilungsdichte aus Gleichung (2.24)

$$g(t_0 + \Theta)d\Theta = \exp \left\{ -\lambda\Theta + \frac{A\lambda}{2\pi\omega} \{ \sin 2\pi\omega t_0 - \sin[2\pi\omega(t_0 + \Theta)] \} \right\} \lambda \{ 1 + A \cos[2\pi\omega(t_0 + \Theta)] \} d\Theta. \quad (2.40)$$

Man kann leicht sehen, daß dieser Ausdruck für  $A \rightarrow 0$  gegen die Exponentialverteilung konvergiert. Dies ist im Einklang damit, daß  $f(t)$  dann eine Konstante ist.

Modelle dieser Art kamen Anfang der achtziger Jahre in der Klimatologie auf, als man sich fragte, warum die Eiszeiten nicht kontinuierlich den Milankovic-Zyklen folgten. Mit Modellen dieser Art kann man diese Frage leichter angehen. Die zyklische Komponente sagt dann nur etwas über die Wahrscheinlichkeit des Übergangs von einem Eiszeit- in einen Nichteiszeitzustand aus. Von den Physikern ist dieser Gedanke eines System mit zyklischen Übergangswahrscheinlichkeiten dankbar aufgenommen worden. Wenn das System sensibel auf extreme Werte reagiert, und die Wahrscheinlichkeit für diese zyklisch ist, kann man, wenn das Rauschen in einem günstigen Verhältnis zur zyklischen Komponente steht, darüber zyklische Komponenten erkennen, die man im mittleren Verhalten nicht erkennen kann. Deshalb hat man diesen Effekt stochastische Resonanz genannt.



## 2.2.4 Change-Point Punktprozesse

Punktprozesse mit unstetigem  $f(t)$  sind nur schwer zugänglich. Andererseits gehören sie auch nicht in das Zentrum des Alltags der stochastischen Prozesse. Trotzdem könnte man sich ein bistabiles System vorstellen, das nun einige Zeit in dem einen attraktiven Zustand verweilt, bevor es wieder in den anderen Zustand übergeht. Wenn das System Ereignisse gemäß eines Punktprozesses hervorruft, so kann man sich vorstellen, daß diese sprunghaft anders sind, wenn das System von dem einen Attraktorbereich auf den anderen springt. Allgemein könnte man bei Systemen mit einem Sprung zur Zeit  $t_s$  folgenden Ansatz machen

$$f(t) = \begin{cases} \lambda_1, & \text{falls } t \leq t_s \\ \lambda_2, & \text{falls } t > t_s. \end{cases} \quad (2.41)$$

Falls  $t_s$  vorgegeben ist, könnte man z.B. vor und nach  $t_s$  vorliegende Daten getrennt analysieren. Falls  $t_s$  nicht bekannt ist, steht man vor einem Problem.

## 2.2.5 Betrachtung bei diskreten Zeitschritten

Wir betrachten nun die Zeit in diskreten Schritten  $t = 1, 2, 3, \dots$ . Die Wahrscheinlichkeit für das Eintreten eines Ereignisses zur Zeit  $n$  ist dann  $p_n$ . Das Ereignis tritt dann im Zeitschritt  $n$  mit der Wahrscheinlichkeit  $q_n = 1 - p_n$  nicht ein. Wir fragen zunächst wieder allgemein nach der Wahrscheinlichkeit dafür, daß  $k$  Ereignisse in den  $m + 1$  Zeitschritten von  $n_1$  bis  $n_1 + m$  stattfinden und gehen dazu sukzessive vor. Die Wahrscheinlichkeit kein Ereignis vorzufinden ist offensichtlich

$$f_0(n_1, n_1 + m) = q_{n_1} q_{n_1+1} \cdots q_{n_1+m} = \prod_{i=0}^m q_{n_1+i}. \quad (2.42)$$

Die Wahrscheinlichkeit ein Ereignis vorzufinden ergibt sich daraus, daß an einer beliebigen Stelle  $q_{n_1+i}$  durch  $p_{n_1+i}$  ersetzt werden muß. Da jede Stelle im Intervall  $n_1$  bis  $n_1 + m$  dafür in Frage kommt, entstehen  $m + 1$  Terme. Es folgt also

$$\begin{aligned} f_1(n_1, n_1 + m) &= p_{n_1} q_{n_1+1} \cdots q_{n_1+m} \\ &+ q_{n_1} p_{n_1+1} \cdots q_{n_1+m} \\ &+ \cdots \\ &+ q_{n_1} q_{n_1+1} \cdots p_{n_1+m} \\ &= \sum_{j=0}^m \left[ p_{n_1+j} \prod_{\substack{i=0 \\ i \neq j}}^m q_{n_1+i} \right]. \end{aligned} \quad (2.43)$$

Für den Spezialfall, daß die Eintrittswahrscheinlichkeiten alle gleich sind ( $p_t = p$ ), folgt daraus wieder das Ergebnis des Bernoulli-Experiments (s. Gleichung (2.14)) für den Spezialfall  $k = 1$ :

$$f_1^{\text{Bernoulli}}(m) = m p (1 - p)^{m-1}. \quad (2.44)$$

Man kann sich leicht vorstellen, daß die allgemeinen Ausdrücke für  $f_k(n_1, n_1 + m)$  schon für kleine  $k$  recht kompliziert werden. Andererseits kann man sich bestimmte Strukturen von  $q_{n_1+i}$  vorstellen, für die sich wiederum vereinfachte Ausdrücke ergeben. So folgt z.B. aus der exponentiellen Abnahme der Wahrscheinlichkeit des Nichteintretens des Ereignisses (das ist die zeitdiskrete Form des exponentiellen Alterns)  $q_{n_1+i} = a * q_{n_1}^i$

$$f_0(n_1, n_1 + m) = \prod_{i=0}^m q_{n_1+i} = \prod_{i=0}^m a q_{n_1}^i = a^{m+1} \prod_{i=0}^m q_{n_1}^i = a^{m+1} q_{n_1}^{\sum_{i=0}^m i} = a^{m+1} q_{n_1}^{m(m+1)/2}. \quad (2.45)$$

Wir beenden nun die Betrachtungen von  $f_i(n_1, n_1 + m)$  und wenden uns der einfacheren Frage zu, wie lange man warten muß bis das erste Ereignis eintritt. Die Wartezeit  $g(n_1, m)$  vom Zeitschritt  $n_1$  bis zum Zeitschritt mit dem ersten Ereignis  $n_1 + m$ , ist gegeben durch das Produkt der Wahrscheinlichkeit dafür, daß in der Zeit von  $n_1$  bis  $n_1 + m - 1$  kein Ereignis stattgefunden hat und im Zeitschritt  $n_1 + m$  ein Ereignis stattfindet:

$$g(n_1, m) = f_0(n_1, n_1 + m - 1) p_{n_1+m} = \prod_{i=0}^{m-1} q_{n_1+i} p_{n_1+m}. \quad (2.46)$$

Für den Fall  $p_t = p$  folgt daraus wieder die Wartezeitverteilung des Bernoulli-Experiments  $g_{Bernoulli}(m) = (1-p)^{m-1} p$  mit der mittleren Wartezeit (= Wiederkehrzeit)  $\tau_{Bernoulli} = 1/p$ . Für das zeitdiskrete exponentielle Altern  $q_{n_1+i} = a q_{n_1}^i$  folgt durch Einsetzen von Gleichung (2.45) in Gleichung (2.46)

$$g(n_1, m) = p_{n_1+m} a^m q_{n_1}^{m(m-1)/2}. \quad (2.47)$$

Die Wiederkehrzeit in Abhängigkeit von  $n_1$  und  $m$  anzugeben, wird hier erst garnicht versucht, sondern scheint eher ein Fall für numerische Experimente zu sein.

Falls  $p_{n_1+i}$  für alle  $i$  bekannt ist, kann im Prinzip die Wartezeitverteilung für jedes beliebige  $n_1$  und  $m$  angegeben werden.

Ein Spezialfall soll hier noch vorgestellt werden, da er in der Klimatologie von Bedeutung ist. Eine Variable sei Gauß-verteilt und stationär um einen zeitabhängigen Mittelwert  $\mu(t)$ . Gesucht ist nun die Wartezeitverteilung ab einem bestimmten Zeitpunkt  $n_1$  für das Überschreiten einer Schwelle  $x_c$ . Wegen

$$p(x \geq x_c, n_1 + i) = 1 - \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x_c} \exp\left(-\frac{(x - \mu(t))^2}{\sigma^2}\right) dx \quad (2.48)$$

folgt für die Wartezeitverteilung  $g(n_1, m) = \prod_{i=0}^{m-1} q_{n_1+i} p_{n_1+m}$  ein Produkt aus  $m + 1$  Errorfunktionen.

# Kapitel 3

## Markovketten

Markov-Ketten sind wertdiskrete Zufallsprozesse, die entweder zeitkontinuierlich oder zeitdiskret sein können. Der Einfachheit halber und weil sie für die Anwendung auf Zeitreihen völlig ausreichend sind, werden hier nur zeitdiskrete Markov-Ketten besprochen. Die wesentliche Eigenschaft von Markov-Ketten heißt Markov-Eigenschaft und liegt darin, daß eine Markov-Kette ihre Vergangenheit vergißt. Man kann Markov-Ketten verschiedener Ordnung definieren, je nachdem wie weit das Gedächtnis des Prozesses in die Vergangenheit reichen soll. Umfaßt der Wertebereich des Prozesses die Werte  $X_1, X_2, \dots, X_K$ , so kann man fragen, mit welcher Wahrscheinlichkeit die Zufallsvariable  $x$ , die die Realisation des Prozesses darstellt, zum Zeitschritt  $n$  den Wert  $X_j$  annimmt, wenn zu vorhergehenden Zeiten bestimmte andere Zustände angenommen wurden. Allgemein kann diese bedingte Wahrscheinlichkeit  $p(x_n = X_j | x_{n-1} = X_k, x_{n-2} = X_l, \dots)$  von beliebig vielen vergangenen Realisationen abhängen. Für eine Markov-Kette  $r$ -ter Ordnung hängt die Wahrscheinlichkeit für eine bestimmte Realisation zur Zeit  $n$  nur von den vorhergehenden  $r$  Realisationen ab. Oft wird mit dem Begriff Markov-Kette nur die Markov-Kette erster Ordnung gemeint, für die die Wahrscheinlichkeit der Realisation des Wertes  $X_j$  zur Zeit  $n$  nur von der Wahrscheinlichkeit der Realisationen aller möglichen Werte  $X_k$  zur Zeit  $n - 1$  und zugeordneten Übergangswahrscheinlichkeiten  $p_{k,j}(n - 1)$  zum Zeitpunkt  $n - 1$  abhängt. Es gilt dann für die Markov-Kette erster Ordnung

$$p(x_n = X_j) = \sum_{k=1}^K p(x_{n-1} = X_k) \cdot p_{k,j}(n - 1). \quad (3.1)$$

Wenn die Übergangswahrscheinlichkeiten  $p_{k,j}(n)$  nicht von der Zeit  $n$  abhängen, spricht man von einer zeitlich homogenen Markov-Kette.

Für die weitere Untersuchung von Markov-Ketten bietet sich die vereinfachte formelle Schreibweise  $p(x_n = X_j) = p_j(n)$  an. Damit kann Gleichung (3.1) in der handlicheren Form

$$p_j(n + 1) = \sum_{k=1}^K p_k(n) \cdot p_{k,j}(n) \quad (3.2)$$

geschrieben werden. Man kann nun für jeden Zeitpunkt einen Wahrscheinlichkeitsvektor  $\vec{p}(n)$  definieren, dessen  $K$  Elemente die Wahrscheinlichkeiten für die Realisation der  $K$  möglichen Zustände sind:

$$\vec{p}(n) = (p_1(n), p_2(n), \dots, p_K(n)). \quad (3.3)$$

Die Übergangswahrscheinlichkeiten  $p_{ij}(n)$  können zu einer Übergangsmatrix  $P(n)$  mit

$$P(n) = \begin{pmatrix} p_{11}(n) & p_{12}(n) & \cdots & p_{1k}(n) \\ p_{21}(n) & p_{22}(n) & \cdots & p_{2k}(n) \\ \dots & \dots & \dots & \dots \\ p_{k1}(n) & p_{k2}(n) & \cdots & p_{kk}(n) \end{pmatrix} \quad (3.4)$$

zusammengefaßt werden. Solche Matrizen, mit Wahrscheinlichkeiten als Elementen heißen stochastische Matrizen. Sie sind quadratisch, alle ihre Elemente erfüllen die Eigenschaft  $p_{i,j}(n) \geq 0$  und alle Zeilensummen sind gleich 1. Die letzte Eigenschaft ist eine Erhaltungseigenschaft, die gewährleistet, daß sich die Realisation immer in einem Zustand befindet. Mit Hilfe der Übergangsmatrix kann die zeitlichen Entwicklung des Wahrscheinlichkeitsvektors geschrieben werden als

$$\vec{p}(n+1) = P(n) \cdot \vec{p}(n). \quad (3.5)$$

Eine häufige Anwendung dieser Gleichung liegt darin, daß man damit Wahrscheinlichkeitsprognosen durchführen kann. Meist ist ein Anfangszustand gegeben, d.h.  $\vec{p}(n=0)$  hat an einer Stelle eine 1 und sonst überall 0. Dann kann man mit Hilfe von Gleichung (3.5) prognostizieren, mit welcher Wahrscheinlichkeit welche Zustände in der Zukunft eingenommen werden.

### 3.1 Homogene Markovketten und Chapman-Kolmogoroff-Gleichung

Markov-Ketten heißen zeitlich homogen (oft einfach *homogen* genannt), wenn die Übergangswahrscheinlichkeiten nicht von der Zeit  $n$  abhängen. Dann hängt auch die Übergangsmatrix nicht von der Zeit ab. Die Elemente  $p_{i,j}$  sind dann konstant. Mit  $p_{i,j}^{(n)}$  wird nun die Wahrscheinlichkeit bezeichnet, in  $n$  Schritten vom Zustand  $i$  in den Zustand  $j$  zu gelangen. Dann gilt:

$$\begin{aligned} p_{i,j}^{(0)} &= \delta_{i,j} = \begin{cases} 0, & \text{falls } i \neq j \\ 1, & \text{falls } i = j \end{cases} \\ p_{i,j}^{(1)} &= p_{i,j} \\ p_{i,j}^{(2)} &= \sum_k p_{i,k} p_{k,j} \\ p_{i,j}^{(n)} &= \sum_k p_{i,k} p_{k,j}^{(n-1)} = \sum_k p_{i,k}^{(n-1)} p_{k,j}. \end{aligned} \quad (3.6)$$

Benutzt man wieder die Matrixschreibweise, so kann man  $p_{i,j}^{(n)}$  leicht berechnen, denn die  $p_{i,j}^{(n)}$  sind die Elemente der  $n$ -ten Potenz der Austauschmatrix. Es gilt also allgemein für die  $n + m$ -stufige Übergangsmatrix

$$P^{n+m} = P^n P^m \quad (3.7)$$

und daher auch für die Elemente

$$p_{i,j}^{(n+m)} = \sum_k p_{i,k}^{(n)} p_{k,j}^{(m)}. \quad (3.8)$$

Dies ist die Chapman-Kolmogoroff-Gleichung. Ein Anwendungsbeispiel soll deren Nutzen verdeutlichen. Ein System befinde sich zur Zeit  $n = 2$  im Zustand  $i$ . Wie hoch ist dann die Wahrscheinlichkeit dafür, daß es sich zur Zeit  $n = 4$  im Zustand  $j$ , zur Zeit  $n = 6$  im Zustand  $k$  und zur Zeit  $n = 9$  im Zustand  $l$  befindet. Es spielt dabei keine Rolle in welchem Zustand es sich zu den Zeiten  $n = 0, 1, 3, 5, 7, 8$  befindet. Die gesuchte Wahrscheinlichkeit folgt direkt aus der allgemeinen Chapman-Kolmogoroff-Gleichung:

$$p(x(n = 4) = X_j, x(n = 6) = X_k, x(n = 9) = X_l | x(2) = X_i) = p_{i,j}^{(2)} p_{j,k}^{(2)} p_{k,l}^{(3)}.$$

## 3.2 Beispiele

Im folgenden werden einige einfache Beispiele von Markov-Ketten vorgestellt.

### 1. Irrfahrt (Diffusion)

Der Irrfahrer in einem unbegrenzten eindimensionalen Raum ist das Paradebeispiel für eine Markov-Kette. Der diskrete Wertebereich kann durch die Menge der ganzen Zahlen ausgedrückt werden. In jedem Zeitschritt wird der Irrfahrer mit der Wahrscheinlichkeit  $p$  um 1 nach rechts und mit der Wahrscheinlichkeit  $q$  um eins nach links wandern. Eine Trajektorie als Realisation der Markov-Kette beginnt also bei einer vorgegebenen ganzen Zahl als Anfangsbedingung und kann mit der Zeit immer weiter von diesem Anfangsort abdriften. Andererseits kann eine Trajektorie auch sehr lange in der Nähe des Anfangs bleiben. Falls  $p$  recht groß ist (dann ist  $q = 1 - p$  recht klein), wird der Irrfahrer wahrscheinlicher zu höheren Zahlen abdriften, als zu niedrigeren. Wie bei jeder Markov-Kette kann man die Wahrscheinlichkeit den Irrfahrer zu einem bestimmten Zeitpunkt an einem bestimmten Ort zu treffen, mit Hilfe der Übergangsmatrix und deren Potenzen berechnen. Die Elemente der Übergangsmatrix folgen aus der Bedingung, daß der Irrfahrer von einem Zustand  $X_i$  nur in die beiden benachbarten Zustände  $j = i \pm 1$  kommen kann. Damit hat die Übergangsmatrix nur in den Nebendiagonalen von 0 verschiedene Elemente. Es gilt

$$p_{i,j} = \begin{cases} p, & \text{falls } j = i + 1 \\ q, & \text{falls } j = i - 1 \\ 0, & \text{sonst} \end{cases} \quad (3.9)$$

Für die unendlich große Übergangsmatrix folgt dann

$$P = \begin{pmatrix} \cdots & & & & & & \cdots \\ \cdots & 0 & q & 0 & p & 0 & \cdots & \cdots \\ \cdots & \cdots & 0 & q & 0 & p & 0 & \cdots \\ \cdots & & & & & & & \cdots \end{pmatrix}. \quad (3.10)$$

Für praktische Berechnungen benötigt man diese unendlich große Matrix nicht. Für den ersten Zeitschritt wird zum Beispiel nur ein Ausschnitt mit  $3 \cdot 3 = 9$  Elementen benötigt. Für  $n$  Zeitschritte wird eine Matrix der Größe  $(1 + 2n)^2$  benötigt. Nach  $n$  Zeitschritten kann der Irrfahrer maximal  $n$  Orte von seiner Startposition entfernt sein. Am Beispiel des Irrfahrers läßt sich sehr viel veranschaulichen, so daß wir darauf zurückkommen werden.

## 2. Irrfahrt mit einem absorbierenden Rand (Ruinproblem)

In diesem Fall steht dem Irrfahrer kein unbegrenzter Raum in beiden Richtungen zur Verfügung. Stattdessen soll bei 0 ein Rand sein. Dadurch wird der Wertebereich auf die Menge der natürlichen Zahlen eingeschränkt. Falls der Irrfahrer nun bei 0 vorbeikommt wird er dort gefangen werden, d.h.  $p_{0,0} = 1$  und  $p_{0,j} = 0$  für alle  $j$ . Da man den Anfangszustand als Kapital zu Beginn einer Reihe von aufeinanderfolgenden Einzelspielen mit jeweils gleicher Gewinnwahrscheinlichkeit ansehen kann, wird diese Form der Irrfahrt in der Spieltheorie als Ruinproblem bezeichnet. Falls man einmal kein Kapital mehr besitzt, ist das Spiel verloren. Die Übergangsmatrix ist jetzt nur noch in einer Richtung unendlich weit ausgedehnt:

$$P = \begin{pmatrix} 1 & 0 & \cdots & & & & \\ q & 0 & p & 0 & \cdots & & \\ 0 & q & 0 & p & 0 & \cdots & \\ 0 & 0 & q & 0 & p & 0 & \\ \cdots & & & & & & \cdots \end{pmatrix}. \quad (3.11)$$

## 3. Irrfahrt mit zwei absorbierenden Rändern

Bei der Irrfahrt mit zwei Rändern ist der Wertebereich endlich. Wir können den kleinsten Wert 0 und den größten Wert  $c$  nennen. Dann gilt zusätzlich zu  $p_{0,0} = 1$  auch noch  $p_{c,c} = 1$  und deshalb auch  $p_{c,j} = 0$  für alle  $j$ . Wenn die Zustände 0 oder  $c$  einmal eingenommen sind, ist ein stationärer Zustand erreicht. Eine Solche Markov-Kette kann als Nullsummenspiel zweier Spieler interpretiert werden. Diese machen fortlaufend Spiele gegeneinander mit gleichbleibender Gewinnwahrscheinlichkeit. Gerechtfertigt wäre dabei die Anfangsbedingung, daß jeder  $c/2$  Mark hat. Der Verlierer eines Spiels zahlt dem Gewinner eine Mark. Die Summe des Geldes ist dabei konstant (deshalb der Name Nullsummenspiel). Versetzt man sich in die Position des einen Spielers, so hat er gewonnen, wenn er  $c$  Mark hat. Er hat verloren, wenn er kein Geld mehr besitzt. Die Wahrscheinlichkeit, mit der er den Abend gewinnt, hängt

von der Anfangsverteilung des Geldes unter den beiden Spielern und seiner Gewinnwahrscheinlichkeit bei den Einzelspielen ab. Die Übergangsmatrix, mit der er seine Gewinnwahrscheinlichkeit berechnen kann ist eine endliche quadratische Matrix der Ordnung  $c + 1$ . Für  $c = 4$  folgt zum Beispiel

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 \\ 0 & q & 0 & p & 0 \\ 0 & 0 & q & 0 & p \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \quad (3.12)$$

#### 4. Irrfahrt mit zwei reflektierenden Rändern

Die Irrfahrt mit reflektierenden Rändern ist als Diffusion in einem begrenzten Gebiet interpretierbar. Spieltheoretisch ist dieser Fall nicht besonders interessant, da immer wenn ein Spieler verlieren würde, dieser wieder ins Spiel gebracht wird. Die Randbedingungen reflektierender Ränder bei 0 und  $c$  sind gegeben durch  $p_{0,0} = 0$ ,  $p_{0,1} = 1$ ,  $p_{c,c} = 0$  und  $p_{c,c-1} = 1$ . Bei einer Kette mit den 5 Zuständen 0 bis 4 folgt dann für die Übergangsmatrix

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 \\ 0 & q & 0 & p & 0 \\ 0 & 0 & q & 0 & p \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}. \quad (3.13)$$

#### 5. Irrfahrt mit Pausen (Unentschiedene Einzelspiele)

Falls der Irrfahrer nicht in jedem Zeitschritt um eine Einheit wandern muß, sondern mit einer gewissen Wahrscheinlichkeit  $r$  in dem Zustand bleibt, in dem er gerade ist, so sind auch die Hauptdiagonalen der Übergangsmatrix gefüllt. Für die Übergangswahrscheinlichkeiten  $p_{ij}$  folgt dann

$$p_{ij} = \begin{cases} p, & \text{falls } i = j + 1 \\ q, & \text{falls } i = j - 1 \\ r, & \text{falls } i = j \\ 0, & \text{sonst.} \end{cases} \quad (3.14)$$

Auch hier ist die Übergangsmatrix von der Ordnung unendlich.

#### 6. Nichtlokale Markov-Ketten

Alle bis jetzt explizit besprochenen Markov-Ketten waren lokale Markov-Ketten, da in jedem Zeitschritt nur direkt benachbarte Zustände erreicht werden konnten. Dies ist keinesfalls eine notwendige Eigenschaft. Die Formulierung durch Übergangsmatrizen entfaltet ihren Sinn erst dann klar, wenn nicht nur die Diagonalen und ersten Nebendiagonalen mit Elementen ungleich 0 gefüllt sind. Dann sind nicht nur Übergänge zu direkten Nachbarn möglich, und man spricht von nichtlokalen Ketten.

## 7. Räumlich und zeitlich homogene Markov-Ketten

Die bisher betrachteten Beispiele waren alle zeitlich homogen und räumlich homogen, da  $p_{ij}$  nicht explizit von  $i$  oder  $j$ , sondern nur von der Differenz  $j - i$  abhängen. Allgemein gilt für räumlich homogene Markov-Ketten

$$p_{ij} = a_{j-i}. \quad (3.15)$$

Die Übergangswahrscheinlichkeit ist demnach für jeden Abstand eine Konstante. Mit  $a_{-1} = q$ ,  $a_0 = r$  und  $a_1 = p$  erhält man die Irrfahrt mit Pausen.

## 8. Zustandsabhängige Markovketten

Zustandsabhängige Markov-Ketten sind räumlich nicht homogene Markov-Ketten. Das bedeutet, dass bei diesen Markov-Ketten die Übergangswahrscheinlichkeiten explizit vom Zustand des Systems abhängen. Es gilt dann z.B. für den zustandsabhängigen Irrfahrer mit Pausen

$$p_{ij} = \begin{cases} p_i, & \text{falls } i = j + 1 \\ q_i, & \text{falls } i = j - 1 \\ r_i, & \text{falls } i = j \\ 0, & \text{sonst.} \end{cases} \quad (3.16)$$

## 9. Ehrenfest-Modell

Das Ehrenfest-Modell ist ein Spezialfall der zustandsabhängigen Markov-Ketten. Es hat als Wertebereich die Werte 0 bis  $c$ . Die Übergangswahrscheinlichkeiten hängen nun vom Ort ab, und zwar in der Form  $p_{i,i+1} = \frac{c-i}{c}$  und  $p_{i,i-1} = \frac{i}{c}$ . Das bedeutet, dass sich eine Realisation eher von den Rändern entfernt, als diesen nahe zu kommen. Für  $i = c/2$  befindet sich die Realisation in der Mitte des Wertebereichs. Die Übergangswahrscheinlichkeiten für beide Richtungen sind dann gleich groß. Für ein Ehrenfest-Modell mit den 5 Zuständen 0 bis 4 folgt dann die Übergangsmatrix

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1/4 & 0 & 3/4 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 3/4 & 0 & 1/4 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}. \quad (3.17)$$

## 10. Wartezeit-Modell

Da Wartezeiten nicht kleiner als 0 sein sollten, ist der Wertebereich des Wartezeit-Modells durch die natürlichen Zahlen inkl. 0 gegeben. Es werden nun die Wahrscheinlichkeiten von einem Ort  $i$  auf 0 zu springen oder einen Wert höher (also auf  $i + 1$ ) zu springen vorgegeben:

$$\begin{aligned} p_{i,0} &= p_i \\ p_{i,i+1} &= 1 - p_i \end{aligned} \quad (3.18)$$



Das bedeutet, daß die Realisation der Markov-Kette entweder um einen Wert höher springt, oder auf 0 zurückgeht. Der Sprung auf 0 soll nun immer ein Ereignis darstellen. Dann ist der Wert  $i$  die Zeit, die seit dem letzten Eintreten des Ereignisses vergangen ist. In einem solchen Wartezeit-Modell ist es möglich, die Wahrscheinlichkeit für das Eintreten des Ereignisses beliebig von der Zeit seit dem letzten Eintreten abhängig zu machen. Betrachtet man den Spezialfall, daß die Wahrscheinlichkeit für das Eintreten des Ereignisses eine Konstante ist ( $p_{i,0} = p$ ), so ist die Wahrscheinlichkeit dafür, daß erst im  $n$ -ten Schritt auf 0 gesprungen wird, gegeben durch  $p(x(\nu) \neq 0, \text{ für } 1 \leq \nu < n, x_n = 0 | x_0 = 0) = p(1-p)^{n-1}$ . Das sollte dem interessierten Leser bekannt vorkommen.

#### 11. Zweidimensionale Irrfahrt

Markov-Ketten können natürlich auch mehrdimensional sein. Das einfachste Beispiel einer mehrdimensionalen Markov-Kette ist der zweidimensionale Irrfahrer. Man kann ihn als Zufallswanderer in den Straßen einer sehr symmetrischen Stadt interpretieren. An jeder Kreuzung hat er vier Möglichkeiten, von denen jede mit einer bestimmten Wahrscheinlichkeit  $p_k$  gewählt wird. Wenn der Wanderer keine Pausen machen darf, muß gelten  $p_1 + p_2 + p_3 + p_4 = 1$ . Die beiden Dimensionen werden nun dadurch formal unterschieden, daß die zweite Dimension durch gestrichene Größen dargestellt wird. Damit geht der Zufallswanderer in jedem Zeitschritt vom Zustand  $i, i'$  in den Zustand  $j, j'$  über. Der Zufallswanderer wird aber immer nur in eine Richtung pro Zeitschritt wandern können. Es ergibt sich deshalb für die zweidimensionale Übergangswahrscheinlichkeit

$$p(i, i')(j, j') = \begin{cases} p_1 & \text{für } j = i + 1, \quad j' = i' \\ p_2 & \text{für } j = i - 1, \quad j' = i' \\ p_3 & \text{für } j = i, \quad j' = i' + 1 \\ p_4 & \text{für } j = i, \quad j' = i' - 1. \end{cases} \quad (3.19)$$