

Skript zur Vorlesung

## **Kommunikationsnetze**

# **Analyse und Leistungsbewertung**

Prof. Dr. Rudolf Mathar  
Rheinisch-Westfälische Technische Hochschule Aachen  
Lehrstuhl für Theoretische Informationstechnik  
Sommerfeldstraße 24  
52056 Aachen



# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Erzeugende Funktionen und Laplace-Transformierte</b>	<b>3</b>
<b>3</b>	<b>Markov-Ketten und Markov-Prozesse</b>	<b>9</b>
<b>4</b>	<b>Markovsche Bediensysteme</b>	<b>31</b>
4.1	<i>M/M/1</i> ( $K = \infty$ , unendliche Wartekapazität) . . . . .	32
4.2	<i>M/M/s/0</i> . . . . .	35
4.3	<i>M/M/<math>\infty</math></i> . . . . .	38
4.4	<i>M/M/s/<math>\infty</math></i> . . . . .	39
4.5	<i>M/M/s/K</i> . . . . .	40
<b>5</b>	<b>Warteschlangennetze</b>	<b>41</b>
5.1	Offene Jackson-Netze . . . . .	44
5.2	Geschlossene Jackson-Netze . . . . .	48
<b>6</b>	<b>Allgemeine Bedienzeitverteilung, <i>M/G/1</i></b>	<b>53</b>
6.1	<i>M/G/1</i> – Systemzustand im Gleichgewicht . . . . .	54
6.2	Wartezeitverteilungen bei <i>M/G/1</i> . . . . .	56



# 1 Einleitung

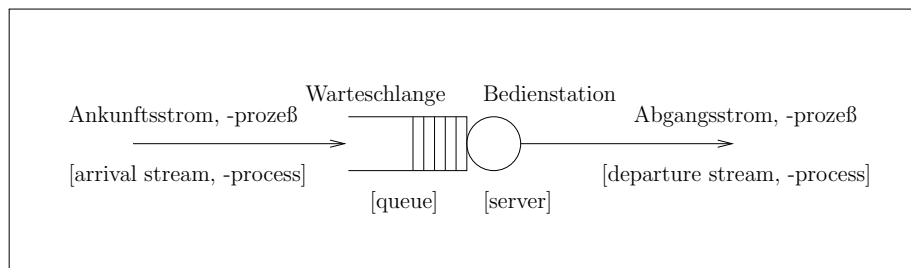
Gegenstand der Vorlesung:

Modellbasierte<1>, quantitative<2> Leistungsanalyse<3> von informationsverarbeitenden Systemen und Kommunikationssystemen. [Model based, quantitative performance evaluation of information processing systems and communication systems.]

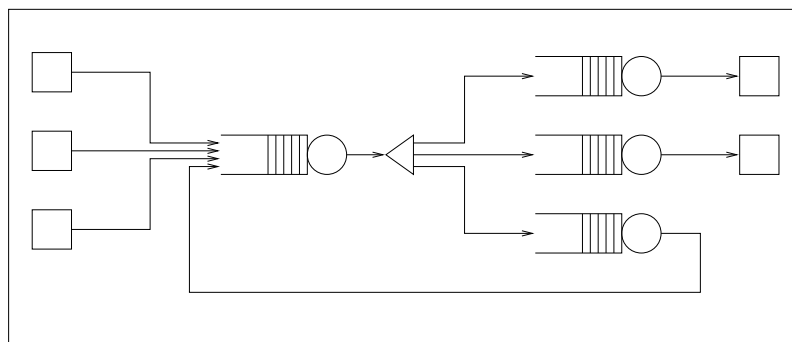
Blockschaltbilder für Prozesse der Informationsverarbeitung.

**Beispiel 1.1.** (nach Abstraktion auf wesentliche Komponenten)

Ein einfaches Beispiel ist wie folgt.



Ein komplexeres Beispiel ist dies.



Ad <1> Prozess der Modellbildung:

- Festlegung der Systemstruktur und -komponenten
- Festlegung von Bedienstrategien
- Vernachlässigung unwesentlicher Details

- Definition von Eingabedaten

Ad <2> Mathematische, überwiegend stochastische Methoden aus der Verkehrstheorie = Bedientheorie = Warteschlangentheorie [Queueing theory]

Ad <3> Berechnung des Systemverhaltens, Auswertung der Ergebnisse, eventuell Validierung durch Vergleich mit Simulation.

Kenngrößen aus

- Benutzersicht: Wartezeit, Antwortzeit, Verzögerung [waiting time / queue time / sojourn time, response time, delay]
- Systemsicht: Auslastung, Durchsatz, Verlustwahrscheinlichkeit [utilization, throughput, loss probability]

Ziel der Vorlesung: Behandlung und Verständnis

- ausgewählter Aspekte der Verkehrstheorie und Leistungsbewertung basierend auf
- mathematischen, insbesondere stochastischen Modellen und Methoden.

## 2 Erzeugende Funktionen und Laplace-Transformierte

$X$  sei diskrete Zufallsvariable mit Träger  $\mathbb{N}_0$  und Zähldichte

$$p_k = P(X = k), \quad k = 0, 1, \dots$$

**Definition 2.1.** Die Funktion  $G_X(z) : [0, 1] \rightarrow \mathbb{R}$ , definiert durch

$$G_X(z) = E(z^X) = \sum_{k=0}^{\infty} p_k z^k$$

heißt (wahrscheinlichkeits-) *erzeugende Funktion* oder  *$z$ -Transformierte* von  $X$  (bzw. der Verteilung  $P^X$ ) [probability generating function,  $z$ -transform].

Beachte:

- $G_X(z)$  ist eine Potenzreihe mit Koeffizienten  $p_k$ .
- $G_X(z)$  kann mit  $z \in \mathbb{C}$  ins Komplexe fortgesetzt werden.
- Da  $p_k \geq 0, \sum_k p_k = 1$ , konvergiert  $\sum_k p_k z^k \quad \forall |z| \leq 1$ . (Es existiert eine konvergente Majorante.) Der Konvergenzradius kann auch größer Eins sein.

Einige Beispiele für erzeugende Funktionen sind in der nachstehenden Tabelle aufgeführt.

$X$	$\sim \text{Bin}(n, p)$	$\sim \text{Geo}(p)$	$\sim \text{Poi}(\lambda)$	$\sim U(\{0, \dots, n\})$
$G_X(z)$	$(zp + 1 - p)^n$	$\frac{p}{1 - (1-p)z}$	$e^{-\lambda(1-z)}$	$\frac{1}{n+1} \frac{1-z^{n+1}}{1-z}$

### Eigenschaften der erzeugenden Funktion

**Proposition 2.2.**  $X, Y$  seien Zufallsvariablen wie in Definition 2.1.

- a) Die erzeugende Funktion bestimmt die Verteilung eindeutig, d.h.

$$G_X(z) = G_Y(z) \Rightarrow P^X = P^Y,$$

ferner gilt

$$p_k = P(X = k) = \frac{1}{k!} G^{(k)}(0).$$

- b) Sind  $X, Y$  stochastisch unabhängig mit erzeugenden Funktionen  $G_X(z), G_Y(z)$ , so gilt

$$G_{X+Y}(z) = G_X(z) \cdot G_Y(z).$$

c) Berechnung von faktoriellen Momenten:

$$\begin{aligned} E(X) &= G'_X(1), \\ E[X(X-1)\dots(X-k+1)] &= G_X^{(k)}(1), \\ V(X) &= G''_X(1) + G'_X(1) - [G'_X(1)]^2. \end{aligned}$$

Beachte:  $G_X^{(k)}(1)$ , Ableitungen als linksseitiger Grenzwert, wenn Konvergenzradius gleich Eins.

**Beispiel 2.3.** (Poissonverteilung)

$$X \sim \text{Poi}(\lambda), \lambda > 0: \quad P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, \dots$$

Poissonverteilungen sind ein Modell für die Anzahl ankommender Aufträge  $N_t$  in einem Intervall der Länge  $t$ :  $N_t \sim \text{Poi}(\lambda t)$ .

Es gilt

$$\begin{aligned} G_X(z) &= \sum_{k=0}^{\infty} z^k e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda z)^k}{k!} = e^{-\lambda} e^{\lambda z} = e^{-\lambda(1-z)}, \quad z \in \mathbb{R} \\ G'_X(z) &= \lambda e^{-\lambda(1-z)}, \quad G'_X(1) = \lambda \\ G''_X(z) &= \lambda^2 e^{-\lambda(1-z)}, \quad G''_X(1) = \lambda^2 \end{aligned}$$

Es folgt:  $E(X) = \lambda, V(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$

Sind  $X, Y$  stochastisch unabhängig,  $X \sim \text{Poi}(\lambda), Y \sim \text{Poi}(\mu)$ , so gilt

$$G_{X+Y}(z) = e^{-\lambda(1-z)} e^{-\mu(1-z)} = e^{-(\lambda+\mu)(1-z)}$$

Also wegen Proposition 2.2 a):  $X + Y \sim \text{Poi}(\lambda + \mu)$ .

Vorteile/Anwendungen erzeugender Funktionen:

- Berechnung von Momenten, ohne die Verteilung explizit zu kennen.
- Erzeugende Funktion der Faltung ist das Produkt der beteiligten erzeugenden Funktionen.

Laplace-Transformierte sind das Analogon zu erzeugenden Funktionen für absolut-stetige Zufallsvariablen.

**Definition 2.4.**  $X$  sei absolut-stetige Zufallsvariable mit Dichte  $f, f(x) = 0$ , falls  $x \leq 0$ .

$$L_X(s) = E(e^{-sX}) = \int_0^{\infty} e^{-sx} f(x) dx, \quad s \geq 0,$$

heißt *Laplace-Transformierte* [Laplace transform] von  $X$  (der Verteilung  $P^X$ ).

Beachte: Die Existenz von  $E(e^{-sX})$  ist wegen „ $f(x) = 0, x \leq 0$ “ sichergestellt.  $E(e^{-sX})$  kann auch für  $s \leq 0$  und Zufallsvariablen mit negativem Träger existieren.



## Eigenschaften der Laplace-Transformierten analog zu erzeugenden Funktionen

**Proposition 2.5.** a) Die Laplace-Transformierte bestimmt die Verteilung eindeutig.

b)  $L_{X+Y}(s) = L_X(s) \cdot L_Y(s)$ , falls  $X, Y$  stochastisch unabhängig.

c)  $E(X) = -L'_X(0)$ ,  $E(X^k) = (-1)^k L_X^{(k)}(0)$ ,  $k \in \mathbb{N}$ ,  $V(X) = L''_X(0) - (L'_X(0))^2$ .

d)  $\lim_{s \rightarrow 0} L_X(s) = 1$ .

**Beispiel 2.6.** (Wichtige Verteilungen und ihre Laplace-Transformierten)

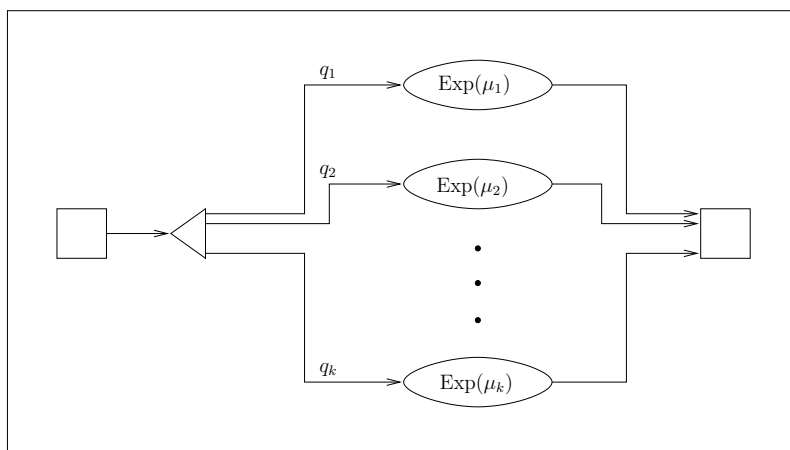
a)  $X \sim \text{Exp}(\lambda)$ . Die zugehörige Laplace-Transformierte lautet

$$L_X(s) = \int_0^{\infty} e^{-sx} \lambda e^{-\lambda x} dx = \frac{\lambda}{\lambda + s}, \quad s \geq 0.$$

b) *Hyperexponentialverteilung* [hyperexponential distribution],  $X \sim \text{HrExp}(\mu_1, \dots, \mu_k, q_1, \dots, q_k)$  mit

$$\mu_i > 0, q_i \geq 0, \sum_{i=1}^k q_i = 1.$$

Anforderungen werden mit Wahrscheinlichkeit  $q_i$  auf Server  $i$  mit  $\text{Exp}(\mu_i)$ -verteilter Bedienzeit geleitet.



$\mu_i > 0, q_i \geq 0, \sum_{i=1}^k q_i = 1$ . Frage: Welche Verteilung hat die Bedienzeit?

„Mischung“ aus Exponentialverteilungen.

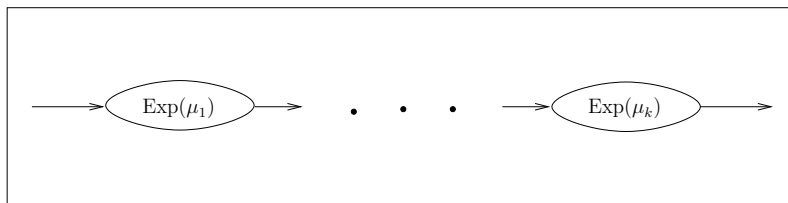
Dichte:  $f_X(x) = \sum_{i=1}^k q_i \mu_i e^{-\mu_i x}$ ,  $x \geq 0$

Verteilungsfunktion:  $F_X(x) = 1 - \sum_{i=1}^k q_i e^{-\mu_i x}$ ,  $x \geq 0$

Laplace-Transformierte:

$$\begin{aligned}
 L_X(s) &= \int_0^\infty e^{-sx} \sum_{i=1}^k q_i \mu_i e^{-\mu_i x} dx \\
 &= \sum_{i=1}^k q_i \int_0^\infty e^{-sx} \mu_i e^{-\mu_i x} dx \\
 &= \sum_{i=1}^k q_i \frac{\mu_i}{\mu_i + s}, \quad s \geq 0.
 \end{aligned}$$

c) *Hypoexponentialverteilung* [hypoexponential distribution],  $X \sim \text{HoExp}(\mu_1, \dots, \mu_k)$



Die Anforderungen werden nacheinander mit  $\text{Exp}(\mu_i)$ -verteilten Bedienzeiten bedient. Es gilt

$$X = \sum_{i=1}^k X_i, \quad X_i \sim \text{Exp}(\mu_i) \text{ s.u.}$$

Falls  $\mu_i \neq \mu_j \forall i \neq j$ , lautet eine Dichte:

$$f_X(x) = \sum_{i=1}^k \left( \prod_{j=1, j \neq i}^k \frac{\mu_j}{\mu_j - \mu_i} \right) \mu_i e^{-\mu_i x}, \quad x \geq 0.$$

Die Laplace-Transformierte ist

$$L_X(s) = \prod_{i=1}^k \frac{\mu_i}{\mu_i + s}, \quad s \geq 0.$$

Falls  $\mu_1 = \mu_2 = \dots = \mu_k = \mu$ , ergibt sich die *Erlangverteilung* [Erlang distribution]

$$X \sim \text{Erl}(k, \mu)$$

mit Dichte

$$f_X(x) = \frac{\mu^k}{(k-1)!} x^{k-1} e^{-\mu x}, \quad x \geq 0$$

und Verteilungsfunktion

$$F_X(x) = 1 - \left( \sum_{i=0}^{k-1} \frac{(\mu x)^i}{i!} \right) e^{-\mu x}, \quad x \geq 0.$$

Die Laplace-Transformierte der Erlangverteilung ist

$$L_X(s) = \left( \frac{\mu}{\mu + s} \right)^k.$$

Die Erlangverteilung ist ein Spezialfall der  $\Gamma$ -Verteilung,  $\Gamma(\alpha, \lambda)$ , mit Dichte

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \quad x \geq 0, \lambda > 0$$

und der Laplace-Transformierten

$$L(s) = \left( \frac{\lambda}{\lambda + s} \right)^\alpha, \quad s \geq 0.$$

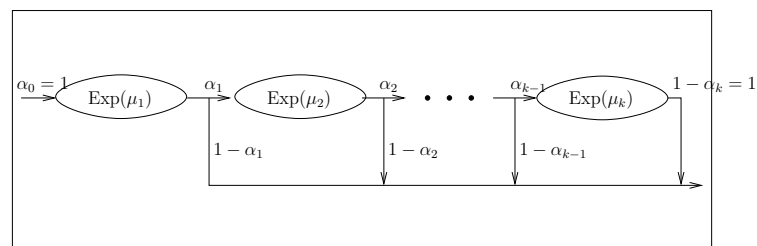
Beachte:  $\Gamma(\alpha)$  bezeichnet die *Gamma-Funktion* [Gamma function]. Es gilt

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$$

$$\Gamma(n) = (n-1)!, \quad n \in \mathbb{N}$$

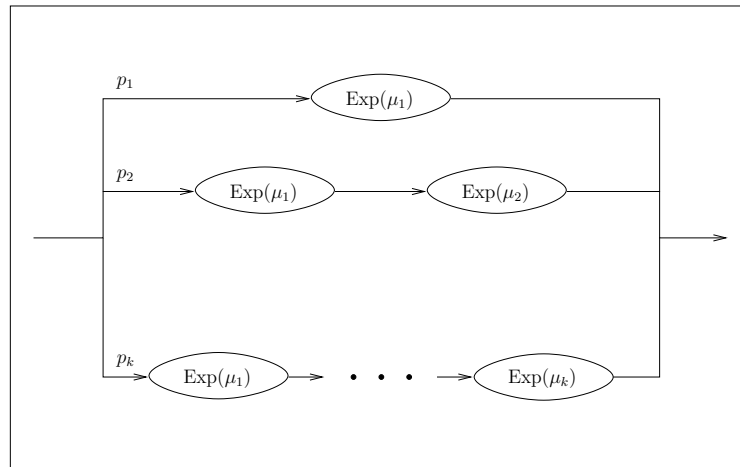
d) *Cox-Verteilungen* (Branching Erlang, Phase-type distribution)

In Verallgemeinerung von den Ansätzen oben wird ein System des folgenden Typs betrachtet:



In Teilservers  $i$  ist die Bedienzeit  $\text{Exp}(\mu_i)$ -verteilt. Nach Teilservers  $i$  gehen Anforderungen mit Wahrscheinlichkeit  $\alpha_i$  in Teilservers  $i + 1$ , mit Wahrscheinlichkeit  $1 - \alpha_i$  verlassen sie das System. Alle Zufallsvariablen sind stochastisch unabhängig.

Eine äquivalente Darstellung mit  $p_i = (1 - \alpha_i) \prod_{j=1}^{i-1} \alpha_j$  ist:



Die Verteilung der Antwortzeit im obigen System heißt Cox-Verteilung mit Parametern  $0 \leq \alpha_1, \dots, \alpha_{k-1} \leq 1, \mu_1, \dots, \mu_k > 0$ .

Bezeichnung:  $\text{Cox}(\alpha_1, \dots, \alpha_{k-1}; \mu_1, \dots, \mu_k)$ . Sie ist eine Mischung aus Hypoexponentialverteilungen, d.h. eine Verteilung der Form:

$$\underbrace{(1 - \alpha_1)}_{=p_1} \text{Exp}(\mu_1) + \underbrace{\alpha_1(1 - \alpha_2)}_{=p_2} \text{Exp}(\mu_1) * \text{Exp}(\mu_2) + \dots + \underbrace{\alpha_1 \dots \alpha_{k-1}}_{=p_k} \text{Exp}(\mu_1) * \dots * \text{Exp}(\mu_k).$$

(„\*“ bedeutet hierbei die Faltung von Verteilungen.)

Die Laplace-Transformierte einer Zufallsvariablen  $X \sim \text{Cox}(\alpha_1, \dots, \alpha_{k-1}; \mu_1, \dots, \mu_k)$  lautet:

$$L_X(s) = \sum_{i=1}^k p_i \prod_{j=1}^i \frac{\mu_j}{\mu_j + s}, \quad s \geq 0.$$

Jede beliebige Verteilung mit positivem Träger kann durch eine Folge von Cox-Verteilungen approximiert werden. Genauer: Bezeichne

$$\mathcal{C} = \{ \text{Cox}(\alpha_1, \dots, \alpha_{k-1}; \mu_1, \dots, \mu_k) \mid 0 \leq \alpha_i \leq 1, \mu_i > 0, k \in \mathbb{N} \}$$

die Menge aller Cox-Verteilungen. Für jede Verteilungsfunktion  $G$  mit  $G(0) = 0$  existiert eine Folge  $C_n \in \mathcal{C}$  mit

$$F_{C_n}(x) \rightarrow G(x) \quad (n \rightarrow \infty) \quad \forall \text{ Stetigkeitspunkte } x \text{ von } G.$$

$F_{C_n}$  bezeichnet die zu  $C_n$  gehörige Verteilungsfunktion,  $F_{C_n}(x) = C_n((-\infty, x])$ . Obige Konvergenz heißt *schwache Konvergenz* [weak convergence] oder *Verteilungskonvergenz* [convergence in distribution].

Also: Jede Bedienzeitverteilung kann durch Cox-Verteilungen approximiert werden. Eine approximative Darstellung des Systems durch Markov-Prozesse ist möglich (siehe später).

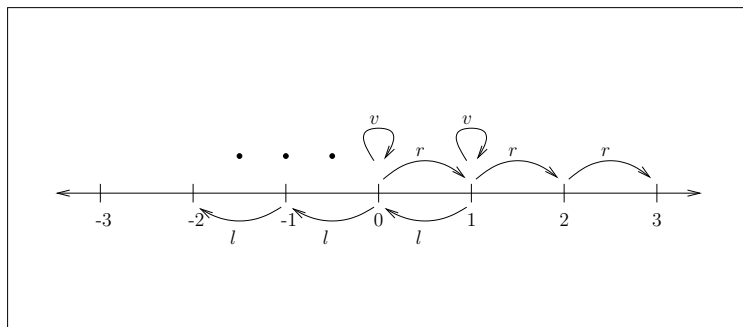
# 3 Markov-Ketten und Markov-Prozesse

Modelliert werden Systeme mit Gedächtnis, allerdings nur für den aktuellen Zustand, nicht für die Vergangenheit. Zunächst gibt es nur diskrete Zeitpunkte  $n = 0, 1, 2, \dots$ . Betrachte hierzu eine Folge von diskreten Zufallsvariablen  $\{X_n\}_{n \in \mathbb{N}_0}$  (auf Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, \mathcal{P})$ ), alle mit demselben Träger  $\mathcal{S} = \{s_1, s_2, \dots\}$ .

Hiermit erfolgt die Modellierung eines zeitabhängigen Systems mit der Interpretation

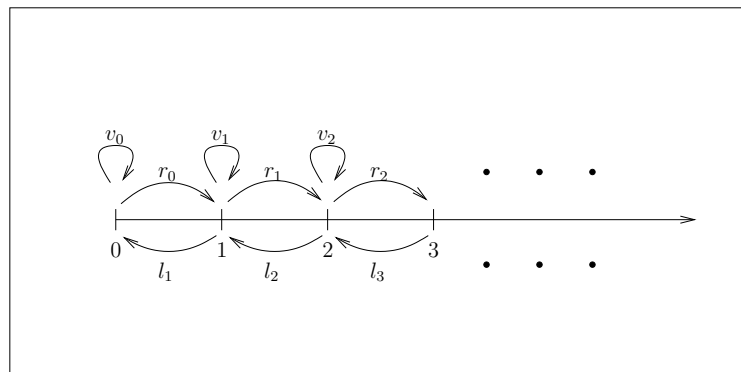
- $\mathcal{S}$  : Zustandsraum
- $\{X_n = s\}$  : System befindet sich zum aktuellen Zeitpunkt im Zustand  $s \in \mathcal{S}$ .

**Beispiel 3.1.** a) **Irrfahrt** (random walk) auf  $\mathbb{Z}$ : Ein Teilchen (Partikel) bewegt sich zu getakteten Zeitpunkten mit Wahrscheinlichkeit  $l$  nach links,  $r$  nach rechts,  $v$  gar nicht, wobei  $l + r + v = 1$ . Veranschaulichung:



$X_n$  bezeichne die Position zum Zeitpunkt  $n$ ,  $X_0$  die Startposition. Es gelte  $P(X_0 = 0) = 1$  (Start im Nullpunkt). Frage: Mit welcher Wahrscheinlichkeit befindet sich das System zum Zeitpunkt  $n$  im Zustand  $k$ , d.h. wie lässt sich  $P(X_n = k)$  berechnen?

b) **Irrfahrt auf  $\mathbb{N}_0$** : Wie a), jedoch mit  $\mathcal{S} = \mathbb{N}_0$ . Zusätzlich hängen die Übergangswahrscheinlichkeiten vom Zustand ab.

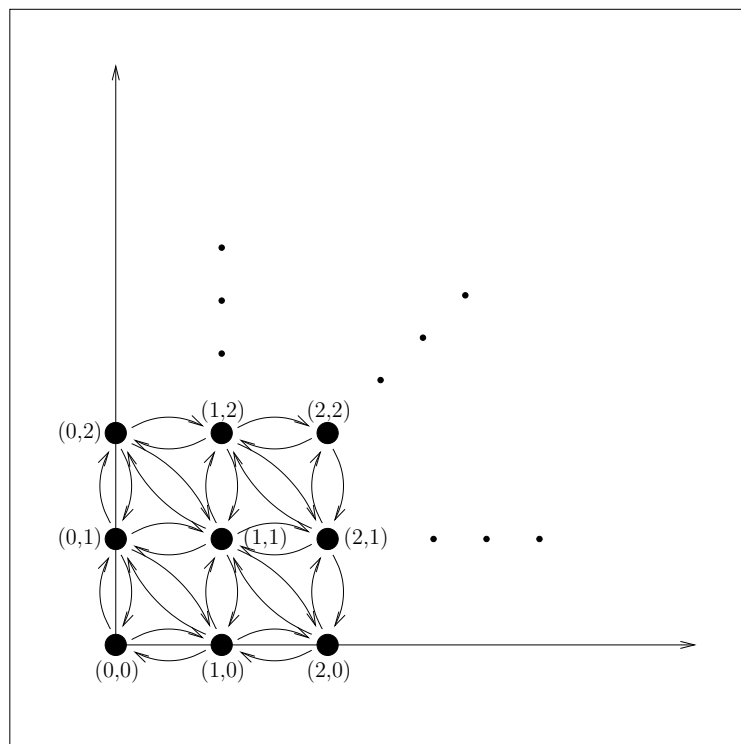


Interpretation als Wartesystem, beobachtet zu diskreten Zeitpunkten. Der Zustandsraum ist  $\mathcal{S} = \mathbb{N}_0$ , Anzahl der Anforderungen im System.

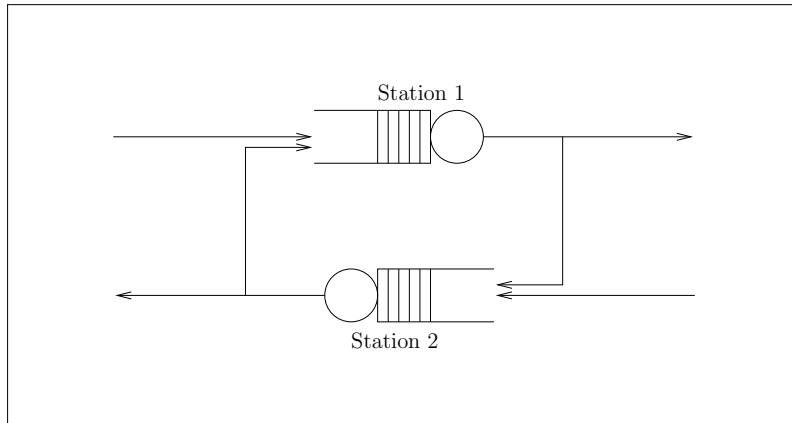
$X_n$  bezeichne die Anzahl der Anforderungen im System zum Zeitpunkt  $n$ . Es gilt:

$$P(X_{n+1} = j \mid X_n = i) = \begin{cases} r_i & j = i + 1 \\ l_i & j = i - 1 \\ v_i & j = i \end{cases} \quad (\text{wobei } l_0 = 0 \text{ gesetzt})$$

c) Auch mehrdimensional, z.B. Irrfahrt auf  $\mathbb{N}_0^2$ :



Die Pfeile bedeuten mögliche positive Übergangswahrscheinlichkeiten.



Interpretation als Warteschlangennetz mit zwei Stationen.

Zustandsraum ist  $\mathcal{S} = \mathbb{N}_0^2$ , die Anzahl der Anforderungen in Station 1 bzw. Station 2.

$P(X_n = (i, j))$  bezeichne die Wahrscheinlichkeit, dass sich zum Zeitpunkt  $n$  genau  $i$  Anforderungen in Station 1 und  $j$  Anforderungen in Station 2 befinden. Zugänge, Abgänge und Wechsel sind jeweils nur für eine Anforderung zwischen den Zeittakten möglich.

**Definition 3.2.** Eine Folge von Zufallsvariablen  $\{X_n\}_{n \in \mathbb{N}_0}$ , alle mit demselben Träger  $\mathcal{S} = \{s_1, s_2, \dots\}$  heißt *Markov-Kette* (MK) [Markov chain], wenn

$$P(X_n = x_n \mid X_{n-1} = x_{n-1}, \dots, X_0 = x_0) = P(X_n = x_n \mid X_{n-1} = x_{n-1})$$

$$\forall n \in \mathbb{N}, x_0, \dots, x_n \in \mathcal{S} \text{ mit } P(X_{n-1} = x_{n-1}, \dots, X_0 = x_0) > 0.$$

$P(X_n = y \mid X_{n-1} = x)$  heißt *Übergangswahrscheinlichkeit* [transition probability] im  $n$ -ten Schritt von  $x$  nach  $y$ .

Eine Markov-Kette heißt *homogen* [homogeneous], wenn

$$P(X_n = y \mid X_{n-1} = x) = P(X_m = y \mid X_{m-1} = x)$$

$$\forall x, y \in \mathcal{S}, m, n \in \mathbb{N} \text{ mit } P(X_{n-1} = x), P(X_{m-1} = x) > 0.$$

Bei homogenen Markov-Ketten sind die Übergangswahrscheinlichkeiten unabhängig von der Zeit. Nur diese werden im weiteren betrachtet.

### Bezeichnungen und Konventionen

Ist  $\mathcal{S}$  endlich, so sei  $\mathcal{S} = \{1, \dots, r\}$  oder  $\mathcal{S} = \{0, \dots, r\}$ . Ist  $\mathcal{S}$  unendlich, so sei  $\mathcal{S} = \mathbb{N}$  (oder  $\mathbb{N}_0$  oder  $\mathbb{Z}$ ).

Setze  $P(X_n = j \mid X_{n-1} = i) = p_{ij}$  für  $i, j \in \mathcal{S}$ .

Die Übergangswahrscheinlichkeiten bilden eine (eventuell unendlich-dimensionale) Matrix

$$\mathbf{\Pi} = \begin{pmatrix} p_{11} & p_{12} & p_{13} & \cdots \\ p_{21} & p_{22} & p_{23} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} = (p_{ij})_{i,j \in \mathcal{S}},$$

die *Übergangsmatrix* [transition matrix]. Übergangsmatrizen sind *stochastische Matrizen* [stochastic matrix], d.h. es gilt

$$p_{ij} \geq 0, \sum_{j \in \mathcal{S}} p_{ij} = 1 \quad \forall i, j \in \mathcal{S}.$$

Die Verteilung von  $X_n$  wird beschrieben durch den *stochastischen Vektor* [stochastic vector]

$$\mathbf{p}(n) = (p_1(n), p_2(n), \dots) = (p_i(n))_{i \in \mathcal{S}} = (P(X_n = 1), P(X_n = 2), \dots).$$

$\mathbf{p}(0) = (p_i(0))_{i \in \mathcal{S}} = (P(X_0 = i))_{i \in \mathcal{S}}$  heißt *Anfangsverteilung* [initial distribution] der Markov-Kette.

### Multiplikation stochastischer Vektoren und Matrizen

$\mathbf{\Pi} = (p_{ij})_{i,j \in \mathcal{S}}$ ,  $\mathbf{\Phi} = (q_{ij})_{i,j \in \mathcal{S}}$  seien stochastische Matrizen,  $\mathbf{p} = (p_i)_{i \in \mathcal{S}}$  ein stochastischer Vektor.

Definiere:

$$\begin{aligned} (\mathbf{p} \mathbf{\Pi})_j &= \sum_{l \in \mathcal{S}} p_l p_{lj}, \quad j \in \mathcal{S} \\ (\mathbf{\Pi} \mathbf{\Phi})_{ij} &= \sum_{l \in \mathcal{S}} p_{il} q_{lj}, \quad i, j \in \mathcal{S} \end{aligned}$$

Beachte:

- Ist  $|\mathcal{S}| < \infty$ , so liegt die übliche Matrix- und Vektormultiplikation vor.
- Ist  $|\mathcal{S}| = \infty$ , so konvergieren obige Reihen.
- $\mathbf{p} \mathbf{\Pi}$  ist ein stochastischer Vektor,  $\mathbf{\Pi} \mathbf{\Phi}$  eine stochastische Matrix.

Gegeben sei eine homogene Markov-Kette mit Anfangsverteilung  $\mathbf{p}(0)$  und Übergangsmatrix  $\mathbf{\Pi}$ . Mögliche Fragen:

- Verteilung von  $X_n$ , dem Zustand zum Zeitpunkt  $n$
- gemeinsame Verteilung von  $(X_k, \dots, X_n)$ ,  $k \leq n$
- Übergangswahrscheinlichkeiten höherer Stufe  $P(X_n = j \mid X_0 = i)$

**Proposition 3.3.** (Verteilungen bei homogenen Markov-Ketten)

$\{X_n\}_{n \in \mathbb{N}_0}$  sei eine homogene Markov-Kette mit Anfangsverteilung  $\mathbf{p}(0)$  und Übergangsmatrix  $\mathbf{\Pi}$ . Es bezeichne  $p_{ij}^{(n)} = P(X_n = j \mid X_0 = i)$ ,  $i, j \in \mathcal{S}$ ,  $n \in \mathbb{N}$  die  $n$ -Schritt-Übergangswahrscheinlichkeiten. Dann gilt

- $\mathbf{p}(n) = \mathbf{p}(n-1)\mathbf{\Pi} = \mathbf{p}(n-2)\mathbf{\Pi}^2 = \dots = \mathbf{p}(0)\mathbf{\Pi}^n$ ,
- $P(X_k = i_k, \dots, X_n = i_n) = p_{i_k}(k) \prod_{j=k+1}^n p_{i_{j-1}i_j} \quad \forall i_k, \dots, i_n \in \mathcal{S}$ ,



$$c) p_{ij}^{(n)} = (\mathbf{\Pi}^n)_{ij} \quad \forall i, j \in \mathcal{S}, n \in \mathbb{N},$$

$$d) p_{ij}^{(n+m)} = \sum_{l \in \mathcal{S}} p_{il}^{(n)} p_{lj}^{(m)} \quad (\text{Chapman-Kolmogorov-Gleichung}).$$

Teil a) zeigt, wie man die Verteilung des Zustandes zum Zeitpunkt  $n$  berechnet. Beachte, dass  $\mathbf{p}(n) = (P(X_n = i))_{i \in \mathcal{S}}$  ein stochastischer Vektor ist. Unterpunkt b) berechnet die Wahrscheinlichkeit eines „Pfads“  $(i_k, \dots, i_n)$ , also die gemeinsame Verteilung von  $(X_k, \dots, X_n)$ . Teil c) zeigt, dass die  $n$ -Schritt-Übergangswahrscheinlichkeit von  $i$  nach  $j$  der  $(i, j)$ -te Eintrag der Matrix  $\mathbf{\Pi}^n$  ist. Die Wahrscheinlichkeit, in  $n + m$  Schritten von  $i$  nach  $j$  zu springen, ergibt sich wie in Unterpunkt d) angegeben.

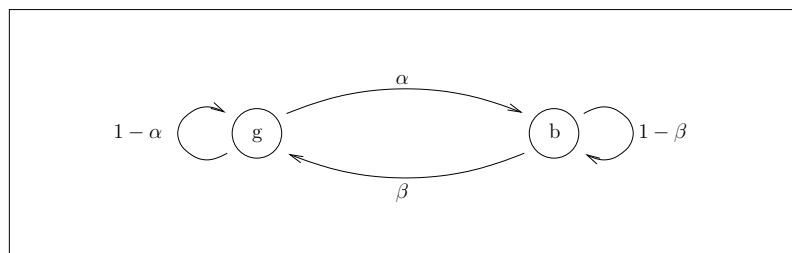
**Beispiel 3.4.** (Modellierung eines Satellitenkanals, Elliot 1963)

Ein Satellitenkanal befindet sich in einem von zwei Zuständen:

g: „good“, kleine Bitfehlerrate

b: „bad“, große Bitfehlerrate

Durch Wettereinflüsse wechselt er zwischen diesen Zuständen mit Wahrscheinlichkeiten  $0 \leq \alpha, \beta \leq 1$  wie folgt:



Die Modellierung als Markov-Kette erfolgt mit der Übergangsmatrix

$$\mathbf{\Pi} = \begin{pmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{pmatrix}.$$

Es gilt (falls  $\alpha + \beta > 0$ )

$$\mathbf{\Pi}^n = \frac{1}{\alpha + \beta} \begin{pmatrix} \beta & \alpha \\ \beta & \alpha \end{pmatrix} + \frac{(1 - \alpha - \beta)^n}{\alpha + \beta} \begin{pmatrix} \alpha & -\alpha \\ -\beta & \beta \end{pmatrix}.$$

Mit Proposition 3.3 a) und der Anfangsverteilung  $\mathbf{p}(0) = (p_1, p_2)$  ergibt sich

$$\mathbf{p}(n) = \mathbf{p}(0)\mathbf{\Pi}^n = \frac{1}{\alpha + \beta}(\beta, \alpha) + \frac{(1 - \alpha - \beta)^n(p_1\alpha - p_2\beta)}{\alpha + \beta}(1, -1).$$

Falls  $0 < \alpha + \beta < 1$  ist, so folgt wegen  $\lim_{n \rightarrow \infty} (1 - \alpha - \beta)^n = 0$

$$\lim_{n \rightarrow \infty} \mathbf{p}(n) = \frac{1}{\alpha + \beta}(\beta, \alpha) \quad \text{und}$$

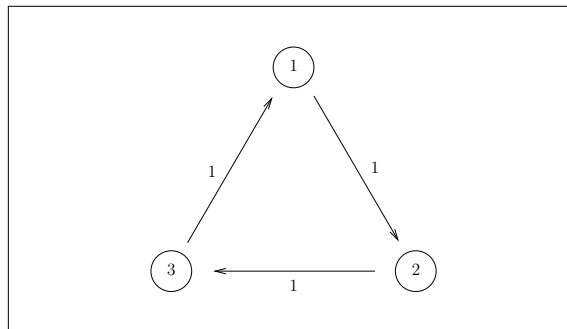
$$\lim_{n \rightarrow \infty} \mathbf{\Pi}^n = \frac{1}{\alpha + \beta} \begin{pmatrix} \beta & \alpha \\ \beta & \alpha \end{pmatrix}.$$

$\mathbf{p}^* = \lim_{n \rightarrow \infty} \mathbf{p}(n)$  beschreibt die Grenzverteilung, also das asymptotische Verhalten des Kanals. Interessant ist, dass  $\mathbf{p}^*$  als Zeilen der Matrix  $\lim_{n \rightarrow \infty} \mathbf{\Pi}^n$  auftritt. Dies wird in bestimmten Fällen allgemein auftreten.

Wichtig ist im Folgenden das asymptotische Verhalten von Systemen, beschrieben durch  $\lim_{n \rightarrow \infty} \mathbf{p}(n)$  und  $\lim_{n \rightarrow \infty} \mathbf{\Pi}^n$ . Zunächst einige Vorbetrachtungen an einfachen Beispielen.

**Beispiel 3.5.** Gegeben sei eine homogene Markov-Kette mit Zustandsraum  $\mathcal{S} = \{1, 2, 3\}$ . Die Darstellung erfolgt durch Übergangsgraphen.

a)



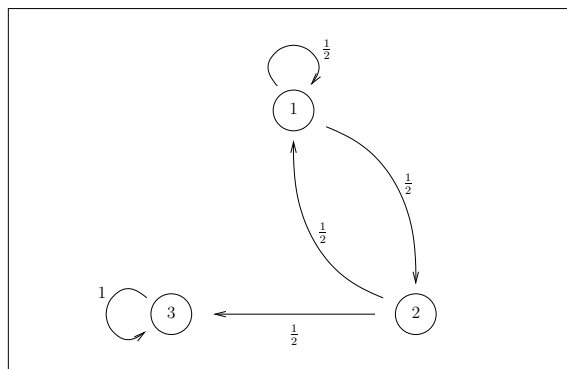
$$\mathbf{\Pi} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad \mathbf{\Pi}^2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad \mathbf{\Pi}^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{\Pi}^4 = \mathbf{\Pi}, \dots$$

Bemerkungen:

- $\mathbf{\Pi}$  ist *periodisch* [periodic], das Verhalten der Markov-Kette ist zyklisch.
- $\mathbf{\Pi}^n$  konvergiert nicht, es gibt drei Häufungspunkte  $\mathbf{\Pi}$ ,  $\mathbf{\Pi}^2$  und  $\mathbf{\Pi}^3$ .
- Aber für  $\mathbf{p}(0) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  gilt

$$\mathbf{p}(n) = \mathbf{p}(0)\mathbf{\Pi}^n = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right), \text{ also } \lim_{n \rightarrow \infty} \mathbf{p}(n) = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right).$$

b)



$$\mathbf{\Pi} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{\Pi}^2 = \frac{1}{4} \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 2 \\ 0 & 0 & 4 \end{pmatrix},$$

$$\mathbf{\Pi}^4 = \frac{1}{16} \begin{pmatrix} 5 & 3 & 8 \\ 3 & 2 & 11 \\ 0 & 0 & 16 \end{pmatrix}, \quad \mathbf{\Pi}^{16} = \frac{1}{256} \begin{pmatrix} 34 & 21 & 201 \\ 21 & 13 & 222 \\ 0 & 0 & 256 \end{pmatrix}$$

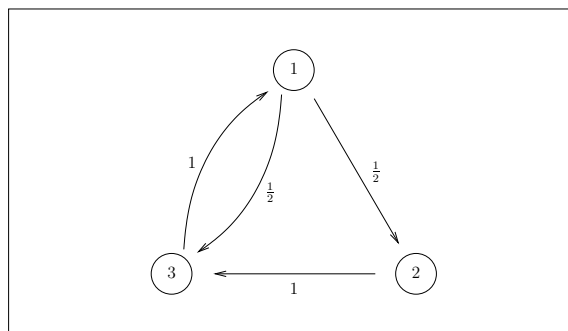
Es gilt

$$\lim_{n \rightarrow \infty} \mathbf{\Pi}^n = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix},$$

also  $\lim_{n \rightarrow \infty} \mathbf{p}(0)\mathbf{\Pi}^n = (0, 0, 1)$ .

Bemerkung:  $\mathbf{\Pi}$  ist nicht irreduzibel. Vom Zustand 3 kann kein anderer Zustand mehr erreicht werden.

c)



$$\mathbf{\Pi} = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad \mathbf{\Pi}^{16} = \frac{1}{256} \begin{pmatrix} 103 & 51 & 102 \\ 102 & 52 & 102 \\ 102 & 51 & 103 \end{pmatrix}$$

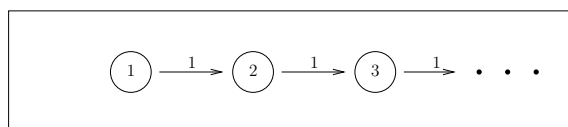
Es gilt

$$\lim_{n \rightarrow \infty} \mathbf{\Pi}^n = \begin{pmatrix} 0.4 & 0.2 & 0.4 \\ 0.4 & 0.2 & 0.4 \\ 0.4 & 0.2 & 0.4 \end{pmatrix},$$

also ist  $\lim_{n \rightarrow \infty} \mathbf{p}(0)\mathbf{\Pi}^n = (0.4, 0.2, 0.4)$ , unabhängig von der Anfangsverteilung  $\mathbf{p}(0)$ .

Bemerkung:  $\mathbf{\Pi}$  ist irreduzibel und aperiodisch.

d) Hier sei nun der Zustandsraum  $\mathcal{S} = \mathbb{N}$ .



$$\mathbf{\Pi} = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Es gilt

$$\begin{aligned} \mathbf{\Pi}^2 &= \begin{pmatrix} 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, & \mathbf{\Pi}^3 &= \begin{pmatrix} 0 & 0 & 0 & 1 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \\ \mathbf{\Pi}^n &= \begin{pmatrix} 0 & \dots & 0 & 0 & \dots \\ 0 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \xrightarrow{(n \rightarrow \infty)} 0 \text{ (Nullmatrix)}. \end{aligned}$$

Daraus folgt  $\lim_{n \rightarrow \infty} \mathbf{p}(0)\mathbf{\Pi}^n = (0, 0, \dots)$ . Der Nullvektor ist kein stochastischer Vektor, es existiert also keine Grenzverteilung.

Bemerkung:  $\mathbf{\Pi}$  ist periodisch und nicht irreduzibel.

Im Folgenden wird das asymptotische Verhalten von Markov-Ketten untersucht, d.h. uns interessiert  $\lim_{n \rightarrow \infty} \mathbf{p}(n)$ . Die Konvergenz ist hierbei komponentenweise definiert, d.h.

$$\lim_{n \rightarrow \infty} \mathbf{p}(n) = \mathbf{p}^* \Leftrightarrow \lim_{n \rightarrow \infty} p_i(n) = p_i^* \quad \forall i \in \mathcal{S}.$$

Gilt  $\mathbf{p}(0)\mathbf{\Pi} = \mathbf{p}(0)$ , so folgt mit Proposition 3.3 a)

$$\mathbf{p}(n) = \mathbf{p}(0)\mathbf{\Pi}^n = \mathbf{p}(0) \quad \forall n \in \mathbb{N},$$

d.h.  $\mathbf{p}(0)$  ist bereits die Grenzverteilung.

Daher:

**Definition 3.6.** Ein stochastischer Vektor  $\mathbf{p}$  heißt *stationär* [stationary], wenn

$$\mathbf{p}\mathbf{\Pi} = \mathbf{p}.$$

Beachte: Ist  $|\mathcal{S}| < \infty$ , so ist  $\mathbf{p}\mathbf{\Pi} = \mathbf{p}$  äquivalent zu  $(\mathbf{I} - \mathbf{\Pi}')\mathbf{p}' = \mathbf{0}$ , dies ist ein homogenes lineares Gleichungssystem.

Fragen: Existiert stets eine stationäre Verteilung und Grenzverteilung, insbesondere unabhängig von der Anfangsverteilung? Zur Klärung dient die folgende Klassifikation von Markov-Ketten.

Seien  $i \in \mathcal{S}$ ,  $m \in \mathbb{N}_0$  so, dass  $P(X_m = i) > 0$ . Setze

$$f_{ii}^{(n)} = P(X_{m+n} = i, X_{m+n-1} \neq i, \dots, X_{m+1} \neq i \mid X_m = i), \quad n \in \mathbb{N}.$$

$f_{ii}^{(n)}$  ist die Wahrscheinlichkeit, nach  $n$  Schritten erstmalig wieder nach  $i$  zurückzukehren. Wegen der Homogenität ist  $f_{ii}^{(n)}$  unabhängig von  $m$ .

---

Setze

$$\mu_i = \sum_{n=1}^{\infty} n f_{ii}^{(n)}.$$

$\mu_i$  ist der Erwartungswert der ersten Rückkehrzeit, sofern  $f_{ii}^{(n)}$ ,  $n \in \mathbb{N}$ , eine Wahrscheinlichkeitsverteilung repräsentiert.

**Definition 3.7.** Ein Zustand  $i \in \mathcal{S}$  heißt *rekurrent* [recurrent, persistent], wenn  $\sum_{n=1}^{\infty} f_{ii}^{(n)} = 1$  gilt. Andernfalls (d.h. wenn  $\sum_{n=1}^{\infty} f_{ii}^{(n)} < 1$  gilt) heißt  $i$  *transient* [transient]. Ein rekurrenter Zustand  $i \in \mathcal{S}$  heißt *positiv-rekurrent* [positive recurrent, non-null-recurrent], wenn  $\mu_i < \infty$  ist. Er heißt *null-rekurrent* [null-recurrent], wenn  $\mu_i = \infty$  gilt.

Rekurrenz oder Transienz sind Eigenschaften aller Zustände, wenn die Markov-Kette irreduzibel im folgenden Sinn ist.

**Definition 3.8.** Eine Markov-Kette bzw. die zugehörige Übergangsmatrix  $\mathbf{\Pi} = (p_{ij})_{i,j \in \mathcal{S}}$  heißt *irreduzibel* [irreducible], wenn für alle  $i, j \in \mathcal{S}$  ein  $n \in \mathbb{N}$  sowie Zwischenzustände  $i_1, \dots, i_{n-1} \in \mathcal{S}$  existieren mit  $p_{ii_1} p_{i_1 i_2} \dots p_{i_{n-1} j} > 0$ .

Irreduzibel bedeutet, dass man von jedem Zustand aus jeden anderen Zustand mit positiver Wahrscheinlichkeit in endlich vielen Schritten erreichen kann.

**Proposition 3.9.** Für irreduzible Markov-Ketten gilt genau eine der folgenden Aussagen.

a) Alle  $i \in \mathcal{S}$  sind transient und es gilt

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty \quad \forall i, j \in \mathcal{S}.$$

b) Alle  $i \in \mathcal{S}$  sind null-rekurrent und es gilt

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty, \quad \lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0 \quad \forall i, j \in \mathcal{S}.$$

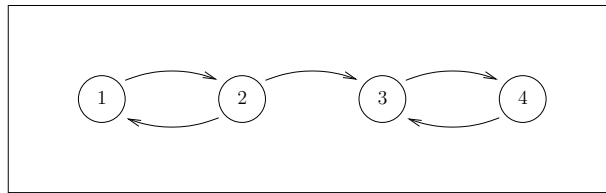
c) Alle  $i \in \mathcal{S}$  sind positiv-rekurrent und es gilt

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty, \quad \limsup_{n \rightarrow \infty} p_{ij}^{(n)} > 0 \quad \forall i, j \in \mathcal{S}.$$

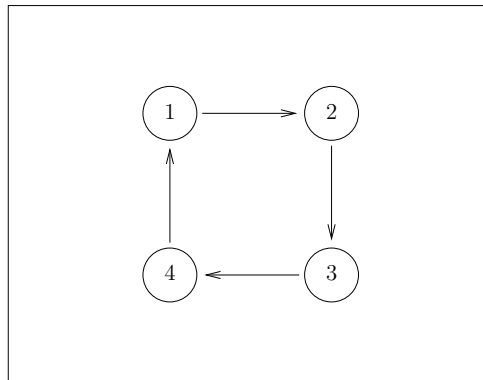
Zyklisches Verhalten kann die Konvergenz stören, siehe Beispiel 3.5 a). Dieses wird ausgeschlossen durch den Begriff „aperiodisch“.

**Definition 3.10.**  $d(i) = \text{ggT}\{n \in \mathbb{N} \mid p_{ii}^{(n)} > 0\}$  heißt *Periode* [period] des Zustands  $i \in \mathcal{S}$ . Ist  $\mathbf{\Pi}$  irreduzibel, so haben alle Zustände dieselbe Periode  $d(i) = d$ . Ist  $d = 1$ , so heißt  $\mathbf{\Pi}$  *aperiodisch* [aperiodic].

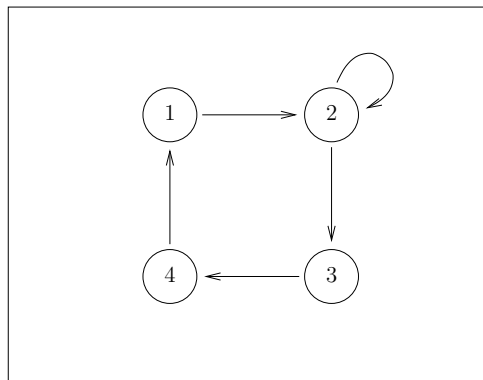
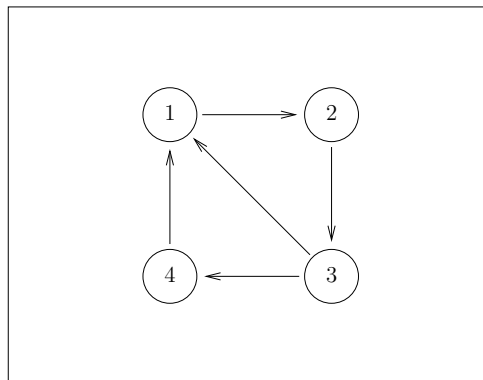
**Beispiel 3.11.** (Positive Übergangswahrscheinlichkeiten sind durch Pfeile markiert.)



Diese Markov-Kette ist nicht irreduzibel, beispielsweise kommt man nicht von Zustand 3 nach Zustand 1 oder von Zustand 4 nach Zustand 2.



Diese Markov-Kette ist irreduzibel, aber nicht aperiodisch.



Diese beiden Markov-Ketten sind aperiodisch.

Es folgt das Hauptergebnis über die Existenz stationärer Verteilungen und Grenzverteilungen.

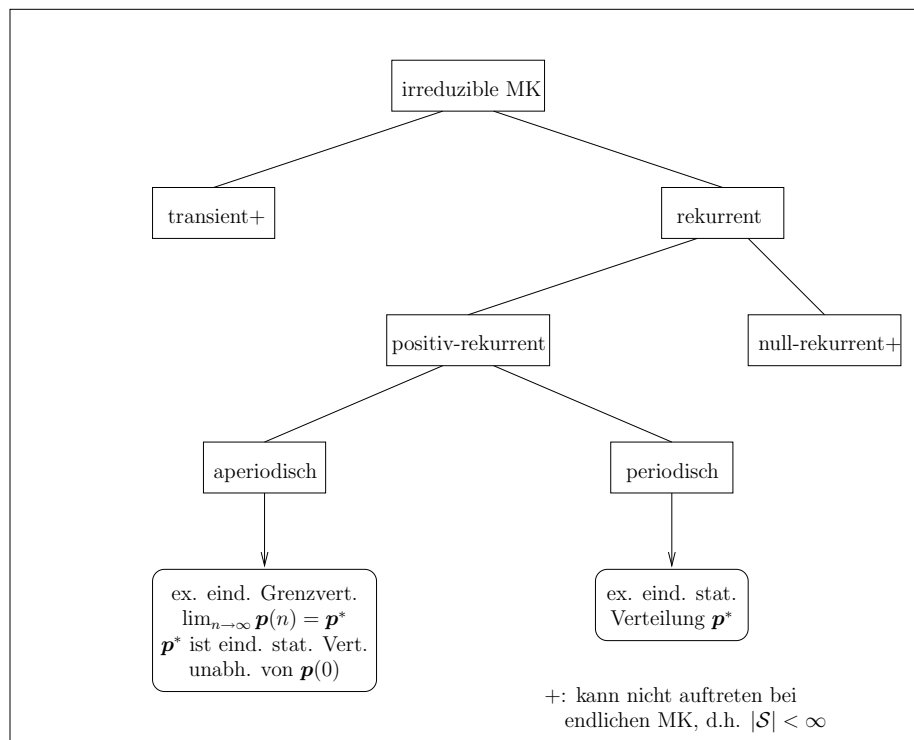
**Theorem 3.12.**  $\{X_n\}_{n \in \mathbb{N}_0}$  sei eine homogene Markov-Kette mit irreduzibler Übergangsmatrix  $\mathbf{\Pi}$ .

- a) Es existiert eine stationäre Verteilung  $\mathbf{p}^*$  genau dann, wenn alle  $i \in \mathcal{S}$  positiv-rekurrent sind.  $\mathbf{p}^*$  ist eindeutig bestimmt, und es gilt  $p_i^* = \frac{1}{\mu_i}$  für alle  $i \in \mathcal{S}$ .
- b) Ist  $\mathbf{\Pi}$  irreduzibel und aperiodisch, so existiert  $\lim_{n \rightarrow \infty} p_{ij}^{(n)}$  für alle  $i, j \in \mathcal{S}$ . Ist die Markov-Kette
  - transient oder null-rekurrent, so gilt  $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$  für alle  $i, j \in \mathcal{S}$ ,
  - positiv-rekurrent, so ist  $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \frac{1}{\mu_j} > 0$  für alle  $i, j \in \mathcal{S}$ .  $\mathbf{p}^* = \left(\frac{1}{\mu_i}\right)_{i \in \mathcal{S}}$  ist die eindeutig bestimmte stationäre Verteilung, und es gilt

$$\lim_{n \rightarrow \infty} \mathbf{p}(n) = \lim_{n \rightarrow \infty} \mathbf{p}(0)\mathbf{\Pi}^n = \mathbf{p}^*$$

unabhängig von der Anfangsverteilung.

### Klassifikation von irreduziblen Markov-Ketten



Zur Bestimmung einer stationären Verteilung löse  $\mathbf{p} = \mathbf{p}\mathbf{\Pi}$  in der Menge der stochastischen Vektoren. Die zugehörigen Gleichungen

$$p_j = \sum_{i \in \mathcal{S}} p_i p_{ij}, \quad j \in \mathcal{S},$$

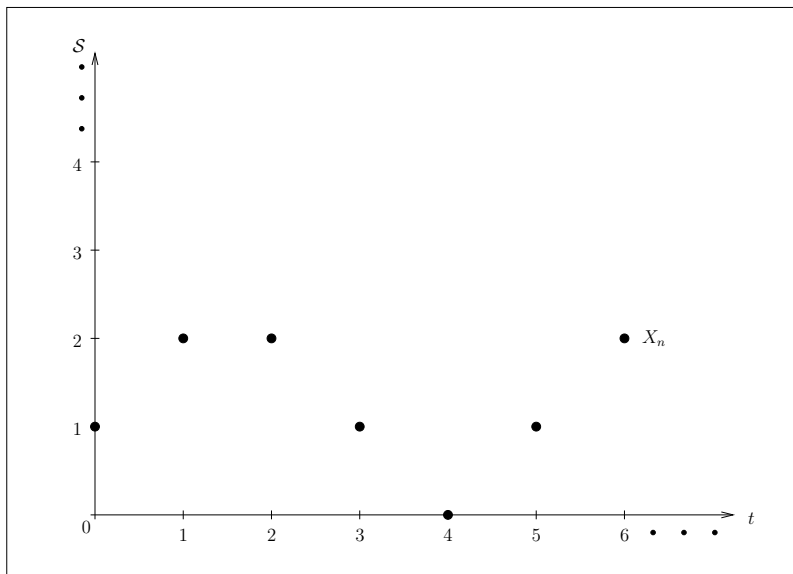
heißen *globale Gleichgewichtsgleichungen* [global balance equations]. Wenn eine Lösung existiert, erhält man die Aussage von Theorem 3.12. Stationäre Verteilungen beschreiben das System im *Gleichgewicht* [steady state, equilibrium], d.h. die Verteilung des Auftretens der Zustände ändert sich nicht beim Fortschreiten.

### Markov-Prozesse

Synonym: Markov-Ketten in stetiger Zeit [Continuous time Markov chains].

Bisher betrachtet: Markov-Ketten  $\{X_n\}_{n \in \mathbb{N}_0}$  mit diskreten Zeitpunkten.

Anschaulich:



(z.B. Anzahl der Anforderungen in einem Wartesystem zu diskreten Zeitpunkten)

Jetzt: Familie von Zufallsvariablen (auf Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, \mathcal{P})$ ) mit stetiger Zeit.

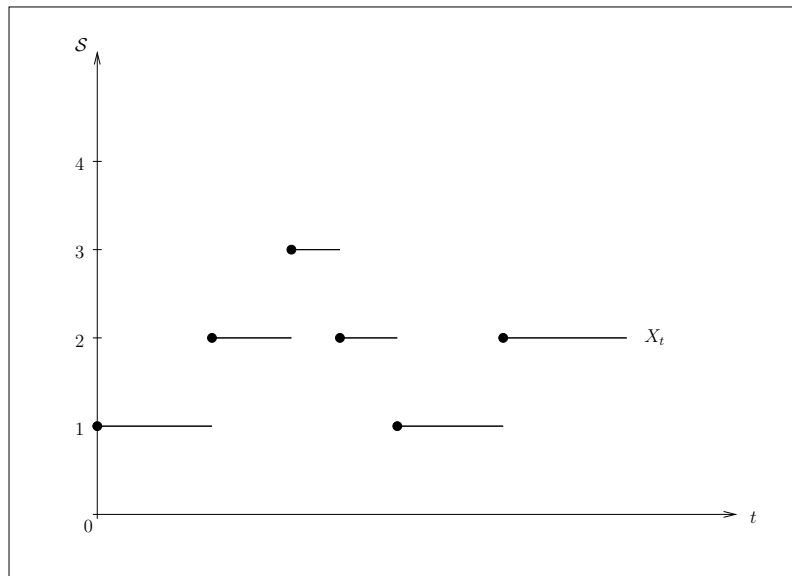
$$X(t) = X_t : (\Omega, \mathcal{A}, \mathcal{P}) \rightarrow (\mathcal{S}, \mathcal{P}(\mathcal{S})), \quad t \geq 0$$

$\mathcal{S}$  ist dabei der Zustandsraum, wie auch bei Markov-Ketten ist er höchstens abzählbar.



---

Anschaulich:



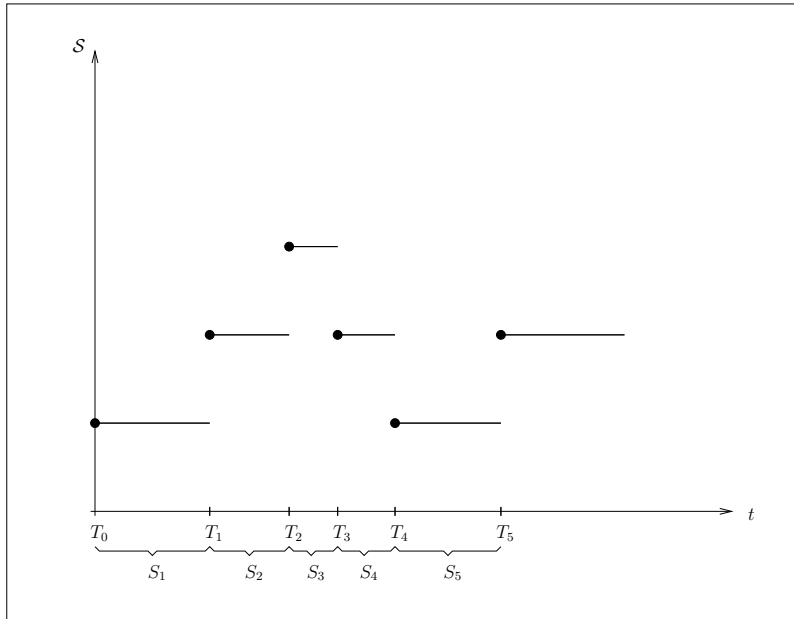
(z.B. Anzahl der Anforderungen in einem Wartesystem zum Zeitpunkt  $t$ )

$X_t = X_t(\omega)$  als Funktion von  $t$  bei festem  $\omega \in \Omega$  heißt *Pfad* [path]. Zustandsübergänge können zu beliebigen Zeitpunkten erfolgen.

Ein analoges Konzept für die Abhängigkeit nur vom aktuellen Zustand enthält die folgende Definition.

**Definition 3.13.** Eine Familie von Zufallsvariablen  $\{X_t\}_{t \geq 0}$  mit Träger  $\mathcal{S}$ ,  $\mathcal{S}$  höchstens abzählbar, heißt *Markov-Prozess* (MP) [Markov process], wenn für alle  $0 \leq t_0 < t_1 < t_2 < \dots$  die Folge  $\{X_{t_n}\}_{n \in \mathbb{N}_0}$  eine Markov-Kette ist. Ein Markov-Prozess heißt *homogen*, wenn für alle  $0 \leq t_0 < t_1 < t_2 < \dots$  mit  $t_{n+1} - t_n = \text{const}$ ,  $n \in \mathbb{N}_0$ , die Markov-Kette  $\{X_{t_n}\}_{n \in \mathbb{N}_0}$  homogen ist.

Wichtige Größen bei Markov-Prozessen (unter der Annahme rechtsstetiger Pfade):



$$T_0 = 0, \quad T_{n+1} = \inf\{t > T_n \mid X_t \neq X_{T_n}\}, \quad n \in \mathbb{N}_0$$

heißen *Sprungzeiten* [transition/jump times].

$$S_0 = 0, \quad S_n = T_n - T_{n-1}, \quad n \in \mathbb{N}$$

heißen *Verweilzeiten* [sojourn/dwell times].

Im Folgenden werden nur homogene Markov-Prozesse betrachtet. Definiere:

*Übergangswahrscheinlichkeiten* [transition probabilities]:

$$p_{ij}(t) = P(X_{s+t} = j \mid X_s = i), \quad t \geq 0, \quad i, j \in \mathcal{S} \quad (\text{unabhängig von } s)$$

*Übergangsmatrizen* [transition matrices]:

$$\mathbf{\Pi}(t) = (p_{ij}(t))_{i,j \in \mathcal{S}}$$

Setze:

$$\mathbf{\Pi}(0) = \mathbf{I} = (\delta_{ij})_{i,j \in \mathcal{S}}, \quad \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad (\text{Einheitsmatrix})$$

Verteilung von  $X_t$ :

$$\mathbf{p}(t) = (p_i(t))_{i \in \mathcal{S}} = (P(X_t = s_1), P(X_t = s_2), \dots)$$

Wie bei Markov-Ketten gilt für alle  $t \geq 0$ :

$$\mathbf{p}(t) = \mathbf{p}(0)\mathbf{\Pi}(t)$$

Ferner gilt die *Chapman-Kolmogorov-Gleichung* (in Matrixform)

$$\mathbf{\Pi}(s+t) = \mathbf{\Pi}(s) \cdot \mathbf{\Pi}(t), \quad s, t \geq 0.$$

Die Übergangsmatrizen  $\mathbf{\Pi}(t)$ ,  $t \geq 0$ , beschreiben das Übergangsverhalten des Systems vollständig, erfordern aber großen technischen Aufwand. Eine kompaktere Darstellung und Interpretation liefern die folgenden so genannten Übergangsraten.

Im Folgenden gilt die Annahme  $\lim_{t \rightarrow 0+} p_{ij}(t) = \delta_{ij}$  für alle  $i, j \in \mathcal{S}$ , in Matrixform  $\lim_{t \rightarrow 0+} \mathbf{\Pi}(t) = \mathbf{I}$ , d.h.  $\mathbf{\Pi}(t)$  bildet eine *stetige Übergangshalbgruppe* [continuous transition semigroup].

**Proposition 3.14.** Für alle  $i, j \in \mathcal{S}$  existieren

$$q_{ij} = \lim_{t \rightarrow 0+} \frac{p_{ij}(t) - \delta_{ij}}{t}, \quad i \neq j, \quad \text{und} \quad q_{ii} = \lim_{t \rightarrow 0+} \frac{p_{ii}(t) - 1}{t}.$$

Die Matrix

$$\mathbf{Q} = (q_{ij})_{i,j \in \mathcal{S}} = \lim_{t \rightarrow 0+} \frac{1}{t}(\mathbf{\Pi}(t) - \mathbf{I}) = \left. \frac{d}{dt} \mathbf{\Pi}(t) \right|_{t=0+}$$

heißt *Intensitätsmatrix* (IM) [(infinitesimal) generator, transition rate matrix].

Es gilt

$$q_{ij} \geq 0, \quad i \neq j, \quad q_{ii} \leq 0, \quad \sum_{j \in \mathcal{S}} q_{ij} = 0 \quad \forall i \in \mathcal{S}.$$

Interpretation:  $q_{ij}$ ,  $i \neq j$ , ist die Rate, mit der das System Zustand  $i$  verlässt und nach Zustand  $j$  springt.  $-q_{ii}$  ist die Gesamtrate, mit der Zustand  $i$  verlassen wird.

Mit Hilfe der Chapman-Kolmogorov-Gleichung gilt für alle  $t, s \geq 0$

$$\mathbf{\Pi}(t+s) - \mathbf{\Pi}(t) = \mathbf{\Pi}(t)[\mathbf{\Pi}(s) - \mathbf{\Pi}(0)],$$

also

$$\lim_{s \rightarrow 0+} \frac{1}{s}[\mathbf{\Pi}(t+s) - \mathbf{\Pi}(t)] = \mathbf{\Pi}(t) \underbrace{\lim_{s \rightarrow 0+} \frac{1}{s}[\mathbf{\Pi}(s) - \mathbf{I}]}_{=\mathbf{Q}}.$$

Es folgt die *Kolmogorov-Vorwärts-Differentialgleichung* [Kolmogorov forward differential equation]

$$\frac{d}{dt} \mathbf{\Pi}(t) = \mathbf{\Pi}(t) \mathbf{Q}, \quad t \geq 0.$$

Analog hergeleitet wird die *Kolmogorov-Rückwärts-Differentialgleichung* [Kolmogorov backward differential equation]

$$\frac{d}{dt} \mathbf{\Pi}(t) = \mathbf{Q} \mathbf{\Pi}(t), \quad t \geq 0.$$

Mit Hilfe dieser Differentialgleichungen können die Übergangswahrscheinlichkeiten  $\mathbf{\Pi}(t)$  aus der Intensitätsmatrix  $\mathbf{Q}$  bestimmt werden.

Analog zu den entsprechenden Begriffen bei Markov-Ketten wird definiert:

**Definition 3.15.** Ein Markov-Prozess mit Übergangswahrscheinlichkeiten  $\mathbf{\Pi}(t)$  heißt *irreduzibel* [irreducible], wenn für alle  $i, j \in \mathcal{S}$ ,  $i \neq j$ , ein  $t > 0$  existiert mit  $p_{ij}(t) > 0$ .

Interpretation: Jeder Zustand  $j$  ist von jedem Zustand  $i$  aus erreichbar.

**Definition 3.16.** Ein stochastischer Vektor  $\mathbf{p}$  heißt *stationäre Verteilung* [stationary, steady-state distribution] eines Markov-Prozesses mit Übergangswahrscheinlichkeiten  $\mathbf{\Pi}(t)$ , wenn

$$\mathbf{p}\mathbf{\Pi}(t) = \mathbf{p} \quad \forall t \geq 0.$$

Interpretation: Die Verteilung  $\mathbf{p}$  ändert sich nicht beim Fortschreiten des Systems.

**Berechnung der stationären Verteilung:**

Es gilt  $\mathbf{p}(t) = \mathbf{p}(0)\mathbf{\Pi}(t)$  für alle  $t \geq 0$ , also durch komponentenweise Differentiation, als Vektor geschrieben

$$\frac{d}{dt}\mathbf{p}(t) = \mathbf{p}(0)\frac{d}{dt}\mathbf{\Pi}(t) = \mathbf{p}(0)\mathbf{\Pi}(t)\mathbf{Q} = \mathbf{p}(t)\mathbf{Q}.$$

Ist  $\mathbf{p}(t) = \mathbf{p}^*$  stationär, also unabhängig von  $t$ , so gilt  $\frac{d}{dt}\mathbf{p}(t) = \mathbf{0}$ . Insgesamt ergibt sich: Ist  $\mathbf{p}^*$  die stationäre Verteilung, so gilt

$$\mathbf{0} = \mathbf{p}^*\mathbf{Q}.$$

Zur Bestimmung der stationären Verteilung löse also das homogene lineare Gleichungssystem  $\mathbf{0} = \mathbf{p}^*\mathbf{Q}$ . Komponentenweise:

$$\sum_{i \in \mathcal{S}} p_i^* q_{ij} = 0 \quad \forall j \in \mathcal{S}$$

Mit  $-q_{jj} = \sum_{i \neq j} q_{ji}$  folgt

$$\sum_{i \neq j} p_i^* q_{ij} = p_j^* \sum_{l \neq j} q_{jl} \quad \forall j \in \mathcal{S}.$$

Diese Gleichungen heißen *globale Gleichgewichtsgleichungen* [global balance equations].

Interpretation: Die linke Seite repräsentiert den gesamten Fluss (Wahrscheinlichkeit der anderen Zustände multipliziert mit der entsprechenden Rate) aus allen anderen Zuständen in den Zustand  $j$ . Die rechte Seite entspricht dem gesamten Fluss aus dem Zustand  $j$  in alle anderen Zustände. Im Gleichgewichtszustand stimmen beide Seiten überein.

Zusammenfassend besitzen Markov-Prozesse die folgenden Eigenschaften:

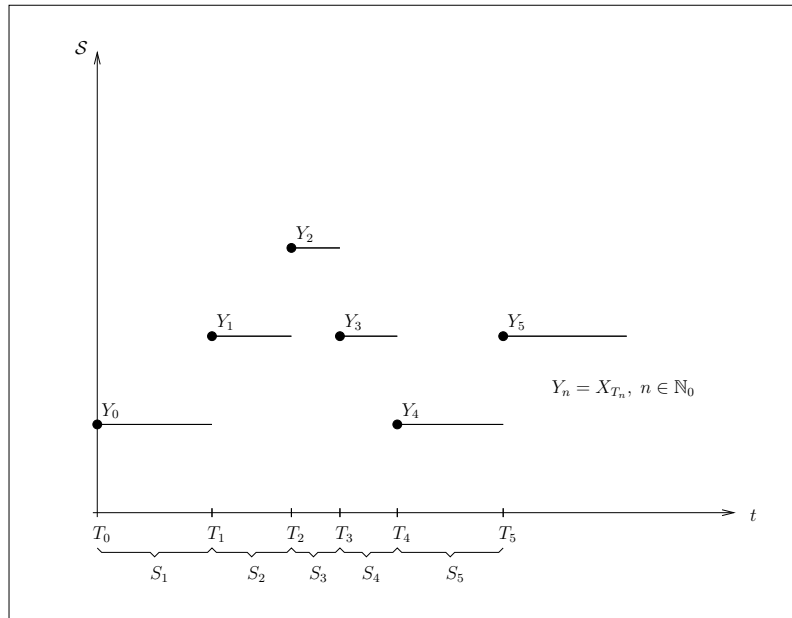
**Theorem 3.17.** Für irreduzible homogene Markov-Prozesse mit Übergangswahrscheinlichkeiten  $\mathbf{\Pi}(t)$  gilt:

- a) Falls eine stationäre Verteilung  $\mathbf{p}^*$  existiert, so ist
  - $\mathbf{p}^*$  eindeutig,
  - $\lim_{t \rightarrow \infty} p_{ij}(t) = p_j^* \quad \forall i, j \in \mathcal{S}$ ,
  - $\lim_{t \rightarrow \infty} \mathbf{p}(t) = \mathbf{p}(0) \lim_{t \rightarrow \infty} \mathbf{\Pi}(t) = \mathbf{p}^*$  unabhängig von der Anfangsverteilung.
  - $\mathbf{p}^*$  ist die eindeutige Lösung von  $\mathbf{0} = \mathbf{p}^*\mathbf{Q}$  in der Menge der stochastischen Vektoren.
- b) Falls keine stationäre Verteilung existiert, so gilt  $\lim_{t \rightarrow \infty} p_{ij}(t) = 0$  für alle  $i, j \in \mathcal{S}$ .

## Strukturelle Eigenschaften von Markov-Prozessen

Beantwortet werden folgende Fragen: Welche Verteilung hat die Verweilzeit in den einzelnen Zuständen? Nach welchem Prinzip wechselt ein Markov-Prozess die Zustände?

Graphisch:



Betrachte einen Markov-Prozess  $X_t, t \geq 0$ , mit Intensitätsmatrix  $\mathbf{Q} = (q_{ij})_{i,j \in \mathcal{S}}$ . Sei

$$p_{ij} = \begin{cases} \frac{q_{ij}}{-q_{ii}} & i \neq j \\ 0 & i = j \end{cases} \quad \text{falls } -q_{ii} > 0,$$

$$p_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \quad \text{falls } -q_{ii} = 0.$$

Setze  $Y_n = X_{T_n}$  (der Zustand sofort nach dem Wechsel),  $S_n, n \in \mathbb{N}$ , sind die Verweilzeiten. Es gelten die folgenden Zusammenhänge.

**Theorem 3.18.**  $\{X_t\}_{t \geq 0}$  sei ein homogener Markov-Prozess mit Intensitätsmatrix  $\mathbf{Q}$ . Dann gilt:

- $\{Y_n\}_{n \in \mathbb{N}_0}$  ist eine homogene Markov-Kette mit Übergangsmatrix  $\mathbf{\Pi} = (p_{ij})_{i,j \in \mathcal{S}}$ .
- $\{S_n\}_{n \in \mathbb{N}}$  ist bedingt stochastisch unabhängig, gegeben  $\{Y_n = i_n\}, n \in \mathbb{N}$ , mit

$$P^{S_{n+1}|Y_n=i} = \text{Exp}(-q_{ii}).$$

$\{Y_n\}_{n \in \mathbb{N}_0}$  heißt *eingebettete Markov-Kette* [embedded Markov chain].

Also: Die Folge der Zustände zu den Sprungzeiten bildet eine Markov-Kette, die eingebettete Markov-Kette. Die Verweilzeit im Zustand  $i \in \mathcal{S}$  ist  $\text{Exp}(-q_{ii})$ -verteilt. Wenn die Zustände bekannt sind, sind die Verweilzeiten stochastisch unabhängig. Umgekehrt können nach diesem Verfahren Markov-Prozesse generiert bzw. simuliert werden:

- Erzeuge den Anfangszustand  $i_0$  nach der Verteilung  $\mathbf{p}_0$ .

Iteriere:

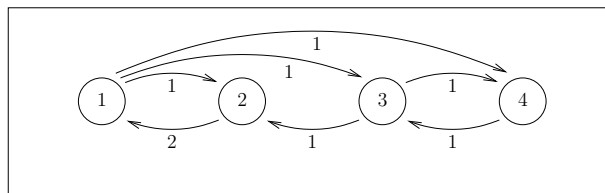
- Erzeuge eine  $\text{Exp}(-q_{i_n i_n})$ -verteilte Zufallszeit (die Verweilzeit im Zustand  $i_n$ ).
- Erzeuge einen neuen Zustand  $i_{n+1}$  gemäß der Verteilung  $(p_{i_n j})_{j \in S}$ , gehe zurück zum vorherigen Schritt.

Das stochastische Verhalten von Markov-Prozessen wird also durch die Intensitätsmatrix  $\mathbf{Q}$  charakterisiert. Eine graphische Darstellung erfolgt analog zu der von Markov-Ketten mit Hilfe des *Intensitätsgraphen* [intensity graph].

**Beispiel 3.19.**

$$\mathbf{Q} = \begin{pmatrix} -3 & 1 & 1 & 1 \\ 2 & -2 & 0 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

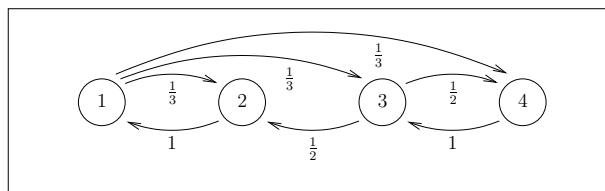
Intensitätsgraph:



Eingebettete Markov-Kette  $\{Y_n\}_{n \in \mathbb{N}_0}$  mit

$$\mathbf{\Pi} = \begin{pmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Übergangsgraph:

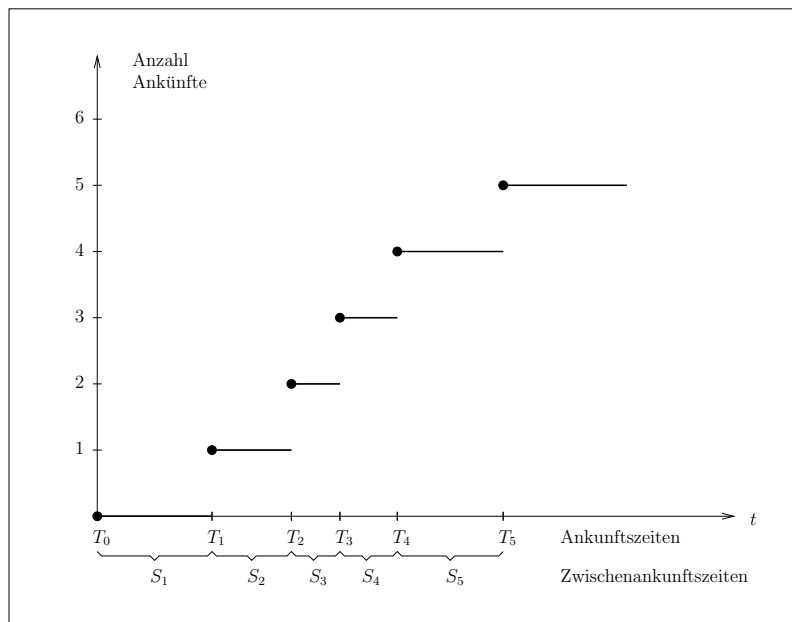


Die Verweilzeit im Zustand 1 ist  $\text{Exp}(3)$ -verteilt, in den Zuständen 2 und 3 ist die Verteilung  $\text{Exp}(2)$  und im Zustand 4  $\text{Exp}(1)$ .

## Der Poisson-Prozess

Poisson-Prozesse sind ein Modell für zufällige Ankünfte von Anforderungen (Kunden, Datenpaketen, Jobs) über der Zeit. Es wird gezählt, wieviele Anforderungen bis zum Zeitpunkt  $t$  angekommen sind. Poisson-Prozesse sind also spezielle *Zählprozesse*.

Graphisch:

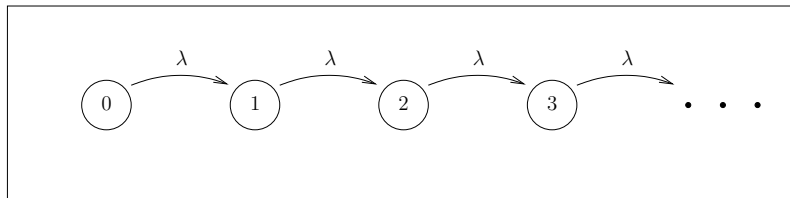


**Definition 3.20.** Ein homogener Markov-Prozess  $N_t = N(t)$ ,  $t \geq 0$ , mit Zustandsraum  $\mathcal{S} = \mathbb{N}_0$  und Intensitäten

$$q_{ij} = \begin{cases} \lambda & j = i + 1 \\ -\lambda & j = i \\ 0 & \text{sonst} \end{cases} \quad i, j \in \mathbb{N}_0, \lambda > 0,$$

heißt homogener *Poisson-Prozess* mit Intensität  $\lambda$ , kurz  $PP(\lambda)$ .

Intensitätsgraph:



Intensitätsmatrix:

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & \dots \\ 0 & -\lambda & \lambda & 0 & \dots \\ 0 & 0 & -\lambda & \lambda & \dots \\ 0 & 0 & 0 & -\lambda & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Aus Theorem 3.18 folgt: Die Zwischenankunftszeiten [interarrival times]  $S_n$ ,  $n \in \mathbb{N}$ , sind stochastisch unabhängig, jeweils  $\text{Exp}(\lambda)$ -verteilt.

Umgekehrt besitzen Poisson-Prozesse die folgende Darstellung:  $\{S_n\}_{n \in \mathbb{N}}$  sei eine Folge von stochastisch unabhängig identisch verteilten Zufallsvariablen, jeweils  $\text{Exp}(\lambda)$ -verteilt. Setze  $T_0 = 0$ ,  $T_n = \sum_{i=1}^n S_i$ ,  $n \in \mathbb{N}$ . Dann ist

$$N_t = |\{n \in \mathbb{N} \mid T_n \leq t\}| = \sup\{n \in \mathbb{N}_0 \mid T_n \leq t\}$$

ein Poisson-Prozess mit Intensität  $\lambda$ .

Betrachte die Anzahl neu hinzukommender Ankünfte im Intervall  $(s, t]$ .

$$N_{(s,t]} = N_t - N_s, \quad 0 \leq s < t$$

heißt *Zuwachs* [increment] im Intervall  $(s, t]$ .

Die Zuwächse in disjunkten Intervallen sind stochastisch unabhängige, poissonverteilte Zufallsvariablen. Genauer:

**Proposition 3.21.** Sei  $\{N_t\}_{t \geq 0}$  ein Poissonprozess mit Intensität  $\lambda$ ,  $0 = t_0 < t_1 < t_2 < \dots$ . Die Zuwächse  $N_{(t_n, t_{n+1}]}$ ,  $n \in \mathbb{N}_0$ , sind stochastisch unabhängig. Ferner gilt

$$N_{(s,t]} \sim \text{Poi}(\lambda(t-s)) \quad \forall 0 \leq s < t.$$

Insbesondere folgt für alle  $i \leq j \in \mathbb{N}_0$ ,  $0 \leq s < t$ :

$$P(N_t = j \mid N_s = i) = P(N_{(s,t]} = j - i) = e^{-\lambda(t-s)} \frac{(\lambda(t-s))^{j-i}}{(j-i)!}$$

Zur Erinnerung:  $X$  heißt poissonverteilt mit Parameter  $\mu$ ,  $X \sim \text{Poi}(\mu)$ , wenn  $P(X = k) = e^{-\mu} \frac{\mu^k}{k!}$  für  $k \in \mathbb{N}_0$ .

**Beispiel 3.22.**  $\{N_t\}_{t \geq 0}$  sei ein Poissonprozess mit Intensität  $\lambda$ . Bestimme die Wahrscheinlichkeit, dass im Intervall  $(0, 2]$  genau  $k$  Anforderungen und in  $(2, 3]$  genau  $l$  Anforderungen ankommen,  $k, l \in \mathbb{N}_0$ .

$$\begin{aligned} P(N_{(0,2]} = k, N_{(2,3]} = l) &= P(N_{(0,2]} = k)P(N_{(2,3]} = l) \\ &= e^{-2\lambda} \frac{(2\lambda)^k}{k!} e^{-\lambda} \frac{\lambda^l}{l!} = e^{-3\lambda} \frac{2^k \lambda^{k+l}}{k!l!} \end{aligned}$$

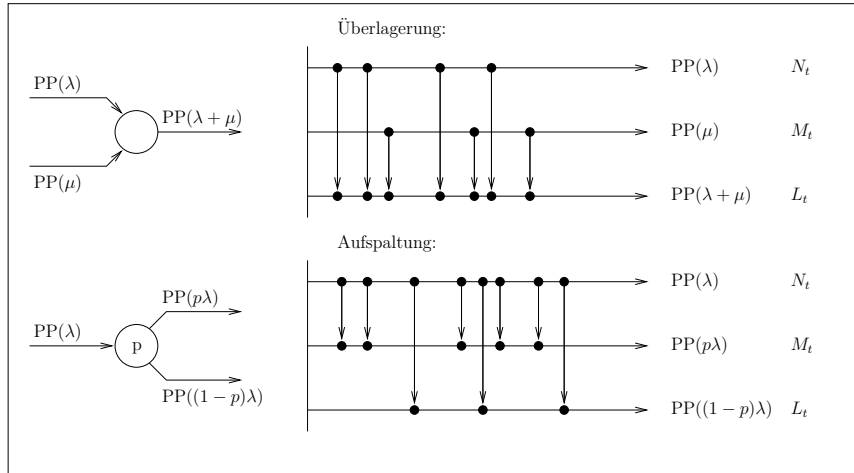
Bestimme für  $k \leq m \in \mathbb{N}_0$  die Wahrscheinlichkeit, dass bis zum Zeitpunkt 3 insgesamt  $m$  Anforderungen eingetroffen sind, unter der Bedingung, dass bis zum Zeitpunkt 2 bereits  $k$  Anforderungen eintrafen:

$$P(N_3 = m \mid N_2 = k) = e^{-\lambda} \frac{\lambda^{m-k}}{(m-k)!}$$

Poissonprozesse beschreiben Ankunftsströme. Eine wichtige Eigenschaft ist, dass die Überlagerung [superposition, merging] und die Aufspaltung [splitting] von Poissonprozessen wieder Poissonprozesse liefern.

Graphisch:





Bei jeder Ankunft wird ein unabhängiges Zufallsexperiment ausgeführt, mit Wahrscheinlichkeit  $p$  für „oben“ und Wahrscheinlichkeit  $1 - p$  für „unten“.

Ist  $\{N_t\}_{t \geq 0}$  ein Poissonprozess mit Intensität  $\lambda$  und  $\{M_t\}_{t \geq 0}$  ein Poissonprozess mit Intensität  $\mu$ , so hat die Überlagerung die Darstellung  $L_t = N_t + M_t$ ,  $t \geq 0$ . Zur Aufspaltung wird eine Folge  $\{I_n\}_{n \in \mathbb{N}}$  von stochastisch unabhängigen, jeweils  $\text{Bin}(1, p)$ -verteilten Zufallsvariablen benötigt. Die Teilprozesse haben dann die Darstellung

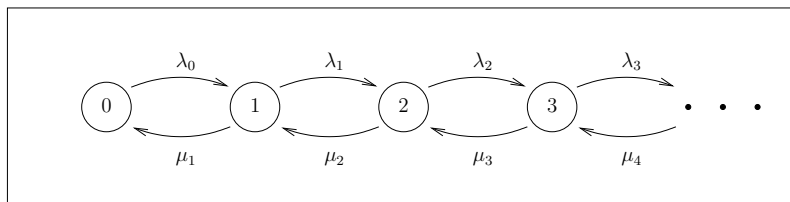
$$M_t = \sum_{j=1}^{N_t} I_j, \quad L_t = \sum_{j=1}^{N_t} (1 - I_j).$$

### Geburts- und Todesprozesse

Beim Poissonprozess gibt es Zuwächse von  $i$  nach  $i + 1$  mit konstanter Intensität  $\lambda$ . Dieses Konzept wird nun wie folgt erweitert:

- Die Zuwächse erfolgen mit variabler Intensität.
- Auch Abgänge sind möglich.

Der Intensitätsgraph hat die folgende Form:



**Definition 3.23.** Ein homogener Markov-Prozess  $\{X_t\}_{t \geq 0}$  mit Zustandsraum  $\mathcal{S} = \mathbb{N}_0$  oder  $\mathcal{S} = \{0, 1, \dots, k\}$  und Intensitätsmatrix

$$Q = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \dots \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & 0 & \dots \\ 0 & \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & \dots \\ 0 & 0 & \mu_3 & -(\lambda_3 + \mu_3) & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad \lambda_i, \mu_i \geq 0,$$

heißt *Geburts- und Todesprozess* (GTP) [birth and death process].

Mit Theorem 3.18 folgt: Die Verweilzeiten im Zustand  $j$  sind  $\text{Exp}(\lambda_j + \mu_j)$ -verteilt,  $j \in \mathbb{N}_0$ , wobei  $\mu_0 = 0$  gesetzt wird. Für die eingebettete Markov-Kette  $\{Y_n\}_{n \in \mathbb{N}_0}$  (siehe Theorem 3.18) gilt

$$P(Y_{n+1} = j \mid Y_n = i) = \begin{cases} \frac{\lambda_i}{\lambda_i + \mu_i} & j = i + 1, i \in \mathbb{N}_0 \\ \frac{\mu_i}{\lambda_i + \mu_i} & j = i - 1, i \in \mathbb{N} \\ 0 & \text{sonst} \end{cases}.$$

Im weiteren wird behandelt, ob Geburts- und Todesprozesse eine stationäre Verteilung besitzen, also im Grenzwert ein statistisches Gleichgewicht erreichen.

**Proposition 3.24.**  $\{X_t\}_{t \geq 0}$  sei ein Geburts- und Todesprozess mit  $\lambda_i, \mu_i > 0$ . Setze

$$\vartheta_0 = 1, \quad \vartheta_j = \frac{\lambda_0 \cdot \dots \cdot \lambda_{j-1}}{\mu_1 \cdot \dots \cdot \mu_j}, \quad j \geq 1.$$

Es existiert eine stationäre Verteilung  $\mathbf{p}^* = (p_i^*)_{i \in \mathcal{S}}$  genau dann, wenn  $\sum_{j \in \mathcal{S}} \vartheta_j < \infty$ . In diesem Fall gilt

$$p_0^* = \frac{1}{\sum_{j \in \mathcal{S}} \vartheta_j}, \quad p_j^* = \vartheta_j p_0^*, \quad j \in \mathcal{S}.$$

Denn: Das oben angegebene  $\mathbf{p}^*$  ist die stationäre Verteilung, wie man mit Hilfe der Gleichung  $\mathbf{p}\mathbf{Q} = \mathbf{0}$ , also

$$(p_0, p_1, p_2, \dots) \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & \dots \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & \dots \\ 0 & \mu_2 & -(\lambda_2 + \mu_2) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} = (0, 0, 0, \dots),$$

nachprüfen kann. Elementweise ergeben sich die Gleichungen

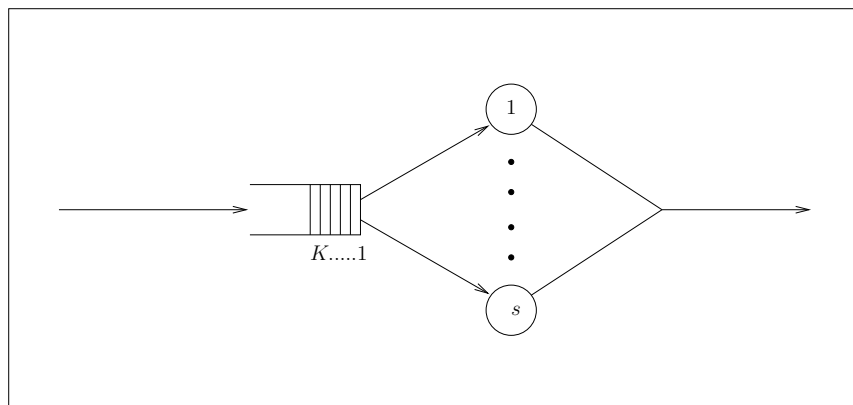
$$\begin{aligned} -\lambda_0 p_0 + \mu_1 p_1 &= 0, \\ \lambda_{j-1} p_{j-1} - (\lambda_j + \mu_j) p_j + \mu_{j+1} p_{j+1} &= 0, \quad j \in \mathcal{S}. \end{aligned}$$

Durch Einsetzen von  $\mathbf{p}^*$  und Division durch  $p_0^*$  ergibt sich

$$\begin{aligned} -\lambda_0 + \mu_1 \frac{\lambda_0}{\mu_1} &= 0, \\ \lambda_{j-1} \frac{\lambda_0 \cdot \dots \cdot \lambda_{j-2}}{\mu_1 \cdot \dots \cdot \mu_{j-1}} - (\lambda_j + \mu_j) \frac{\lambda_0 \cdot \dots \cdot \lambda_{j-1}}{\mu_1 \cdot \dots \cdot \mu_j} + \mu_{j+1} \frac{\lambda_0 \cdot \dots \cdot \lambda_j}{\mu_1 \cdot \dots \cdot \mu_{j+1}} &= 0. \end{aligned}$$

# 4 Markovsche Bediensysteme

Betrachtet werden einfache Bediensysteme des Typs:



Die Bedienzeiten sind voneinander und vom Ankunftsprozess stochastisch unabhängig. Freie Server werden sofort genutzt.

Diese Systeme werden allgemein beschrieben durch die Kendallsche Notation:

$$A/B/s/K - \text{Warteschlangendisziplin}$$

$A$ : Zwischenankunftszeiten

$B$ : Bedienzeiten

$s$ : Anzahl der Server

$K$ : Anzahl der Warteschlangenplätze ( $K = \infty$  wird weggelassen)

Für  $A$  und  $B$  werden folgende Symbole verwendet:

$M$ : Exponentialverteilung (Markov-Eigenschaft, Memoryless)

$E_k$ : Erlang-Verteilung mit  $k$  Phasen

$H_k$ : Hyperexponentialverteilung mit  $k$  Phasen

$C_k$ : Cox-Verteilung mit  $k$  Phasen

$D$ : Deterministisch, konstant

$G$ : Allgemeine Verteilung (General, auch  $GI$ )

Als Warteschlangendisziplin kommen in Frage:

- FCFS=FIFO: First Come First Served, First In First Out
- LCFS=LIFO: Last Come First Served, Last In First Out
- SIRO: Service In Random Ord
- RR: Round Robin (Zyklische Bearbeitung in Zeitscheiben)
- PS: Processor Sharing (RR mit infinitesimal kleinen Zeitscheiben, alle werden gleichzeitig bedient)

Behandelt werden jetzt Bediensysteme des Typs  $M/M/s/K$ . Der Ankunftsstrom ist ein Poisson-Prozess mit Intensität  $\lambda$ , die Zwischenankunftszeiten sind also stochastisch unabhängig identisch  $\text{Exp}(\lambda)$ -verteilt. Die Bedienzeiten sind stochastisch unabhängig identisch  $\text{Exp}(\mu)$ -verteilt, unabhängig vom Ankunftsstrom.

Untersucht werden soll  $X_t$ , die Anzahl der Anforderungen im System zum Zeitpunkt  $t \geq 0$ . Dieses Modell ist wohlbegründet:  $X_t, t \geq 0$ , ist ein Geburts- und Todesprozess mit Zustandsraum  $\mathcal{S} \subseteq \mathbb{N}_0$ .

Die oben aufgeführten Eigenschaften folgen bei Geburts- und Todesprozessen mit geeigneten Parametern aus den strukturellen Eigenschaften von Markov-Prozessen (siehe Kapitel 3).

#### 4.1 $M/M/1$ ( $K = \infty$ , unendliche Wartekapazität)

$X_t, t \geq 0$ , ist ein Geburts- und Todesprozess mit Zustandsraum  $\mathcal{S} = \mathbb{N}_0$  und Intensitätsmatrix

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & \dots \\ \mu & -(\mu + \lambda) & \lambda & 0 & \dots \\ 0 & \mu & -(\mu + \lambda) & \lambda & \dots \\ 0 & 0 & \mu & -(\mu + \lambda) & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

mit  $\lambda, \mu > 0$ , wobei  $\lambda$  die Ankunftsrate und  $\mu$  die Bedienrate darstellt.

Zu behandeln sind folgende Fragen:

- Gibt es eine stationäre Verteilung bzw. einen Gleichgewichtszustand, wie sieht  $\lim_{t \rightarrow \infty} P(X_t = j)$  aus?
- Wie ist die Antwortzeit (also die Gesamtverweilzeit im System) verteilt?

Benutze Proposition 3.24 (stationäre Verteilung von Geburts- und Todesprozessen):

$$\vartheta_0 = 1, \quad \vartheta_j = \frac{\lambda_0 \cdot \dots \cdot \lambda_{j-1}}{\mu_1 \cdot \dots \cdot \mu_j} = \left(\frac{\lambda}{\mu}\right)^j, \quad j \in \mathbb{N}$$

Es existiert eine stationäre Verteilung  $\mathbf{p}^* = (p_i^*)_{i \in \mathcal{S}}$  genau dann, wenn

$$\sum_{j=0}^{\infty} \vartheta_j = \sum_{j=0}^{\infty} \left(\frac{\lambda}{\mu}\right)^j < \infty$$

gilt. Dies ist eine geometrische Reihe. Sie konvergiert genau dann, wenn  $\rho = \frac{\lambda}{\mu} < 1$  ist.  $\rho = \frac{\lambda}{\mu}$  heißt Auslastung [traffic intensity, utilization].

Insgesamt: Es existiert eine stationäre Verteilung  $\mathbf{p}^* = (p_j^*)_{j \in \mathcal{S}}$  genau dann, wenn  $\rho < 1$ . In diesem Fall gilt

$$p_0^* = \frac{1}{\sum_{l=0}^{\infty} \rho^l} = \frac{1}{\frac{1}{1-\rho}} = 1 - \rho, \quad p_j^* = \rho^j (1 - \rho), \quad j \in \mathbb{N}.$$

Ferner gilt dann

$$\lim_{t \rightarrow \infty} P(X_t = j) = p_j^* = \rho^j (1 - \rho), \quad j \in \mathbb{N}_0,$$

d.h. die stationäre Verteilung ist identisch mit der Grenzverteilung. Sie ist eine geometrische Verteilung mit dem Parameter  $1 - \rho$ .

$X^* \sim (p_0^*, p_1^*, \dots)$  bezeichnet eine Zufallsvariable, die gemäß  $(p_0^*, p_1^*, \dots)$  verteilt ist ( $X^* \sim \text{Geo}(1 - \rho)$ ). Das System im Gleichgewicht (equilibrium) wird beschrieben durch die Zufallsvariable  $X^*$ . Es folgt für das M/M/1-System im Gleichgewicht, d.h.  $X_t \sim X^*$  für alle  $t \geq 0$  mit  $\rho < 1$ :

a) Die Verteilung der Anzahl der Anforderungen im System weist folgende Momente auf:

$$\begin{aligned} E(X^*) &= \frac{\rho}{1 - \rho} = \frac{\frac{\lambda}{\mu}}{1 - \frac{\lambda}{\mu}} = \frac{\lambda}{\mu - \lambda} \xrightarrow{\rho \rightarrow 1} \infty \\ V(X^*) &= \frac{\rho}{(1 - \rho)^2} \xrightarrow{\rho \rightarrow 1} \infty \end{aligned}$$

Dies folgt direkt aus  $X^* \sim \text{Geo}(1 - \rho)$ .

b) Die Wahrscheinlichkeit, dass eine neu ankommende Anforderung sofort bedient wird, ist  $P(X^* = 0) = 1 - \rho$ .

c) Die Wahrscheinlichkeit, dass der Server beschäftigt ist, bzw. der Beschäftigungsanteil beträgt  $P(X^* > 0) = \rho$ .

d) Wartezeitverteilung in der Warteschlange  $W_Q^*$ : Eine Anforderung kommt zum Zeitpunkt  $t$  an. Die Warteschlangendisziplin sei FIFO. Welche Verteilung hat die Wartezeit in der Warteschlange  $W_Q^*$ ?

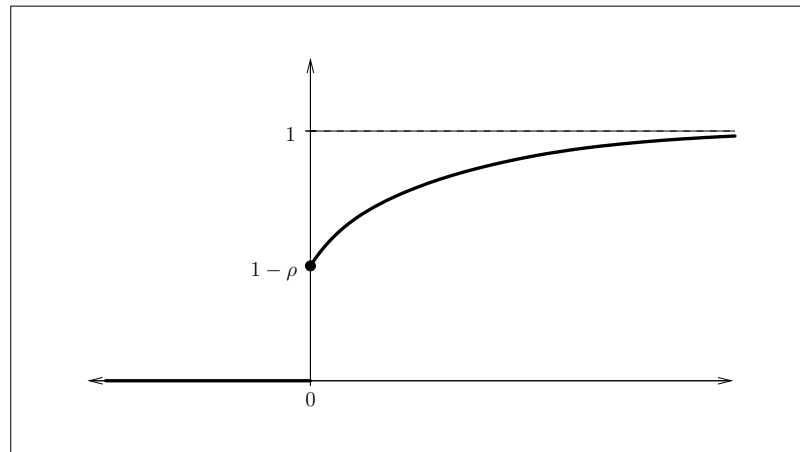
$$P(W_Q^* \leq z) = \begin{cases} 0, & z < 0 \\ P(X^* = 0), & z = 0 \\ P(X^* = 0) + \sum_{n=1}^{\infty} P(Y_1 + \dots + Y_n \leq z) P(X^* = n), & z > 0 \end{cases}$$

$Y_1, \dots, Y_n \sim \text{Exp}(\mu)$  i.i.d., d.h.  $Y_1, \dots, Y_n$  sind stochastisch unabhängig  $\text{Exp}(\mu)$ -verteilt, sie sind die Wartezeit bzw. Restwartezeit der Anforderungen vor der angekommenen Anforderung.

Beachte:  $Y_1 + \dots + Y_n \sim \text{Erl}(n, \mu)$  (Erl( $n, \mu$ )-Verteilung),  $X^* \sim \text{Geo}(1 - \rho)$  (Geo( $1 - \rho$ )-Verteilung). Es folgt für  $z > 0$ :

$$\begin{aligned}
 P(W_Q^* \leq z) &= 1 - \rho + \sum_{n=1}^{\infty} \left[ \int_0^z \frac{\mu^n}{(n-1)!} y^{n-1} e^{-\mu y} dy \right] (1 - \rho) \underbrace{\rho^n}_{=\frac{\lambda^n}{\mu^n}} \\
 &= 1 - \rho + \int_0^z e^{-\mu y} (1 - \rho) \lambda \underbrace{\sum_{n=1}^{\infty} \frac{y^{n-1} \lambda^{n-1}}{(n-1)!}}_{=e^{\lambda y}} dy \\
 &= 1 - \rho + \frac{\mu - \lambda}{\mu} \lambda \int_0^z e^{-(\mu - \lambda)y} dy \\
 &= 1 - \rho + \rho(1 - e^{-(\mu - \lambda)z}) \\
 &= 1 - \rho e^{-(\mu - \lambda)z}
 \end{aligned}$$

Skizze der Verteilungsfunktion von  $W_Q^*$ :



Die Verteilungsfunktion besitzt an der Stelle 0 ein sogenanntes *Atom* [atom], einen Punkt, der positive Wahrscheinlichkeit trägt. Die zugehörige Verteilung ist weder diskret noch absolut-stetig.

Es gilt:

$$\begin{aligned}
 E(W_Q^*) &= \frac{\rho}{\mu(1 - \rho)} \\
 V(W_Q^*) &= \frac{2\rho - \rho^2}{\mu^2(1 - \rho)^2}
 \end{aligned}$$

e) Für die Antwortzeit (Gesamtverweilzeit im System)  $W^*$  gilt

$$\begin{aligned}
 P(W^* \leq z) &= P(W_Q^* + Y \leq z) \\
 &= \int_0^z P(W_Q^* \leq z - y) \mu e^{-\mu y} dy \\
 &= \dots = 1 - e^{-(\mu - \lambda)z}, \quad z \geq 0,
 \end{aligned}$$

wobei  $Y$  die von  $W_Q^*$  stochastisch unabhängige Bedienzeit der ausgewählten Anforderung ist.  $W^*$  ist also  $\text{Exp}(\mu - \lambda)$ -verteilt, es gilt:

$$E(W^*) = \frac{1}{\mu - \lambda}$$

$$V(W^*) = \frac{1}{(\mu - \lambda)^2}$$

Beachte den folgenden Zusammenhang:

$$E(X^*) = \frac{\lambda}{\mu - \lambda} = \lambda E(W^*)$$

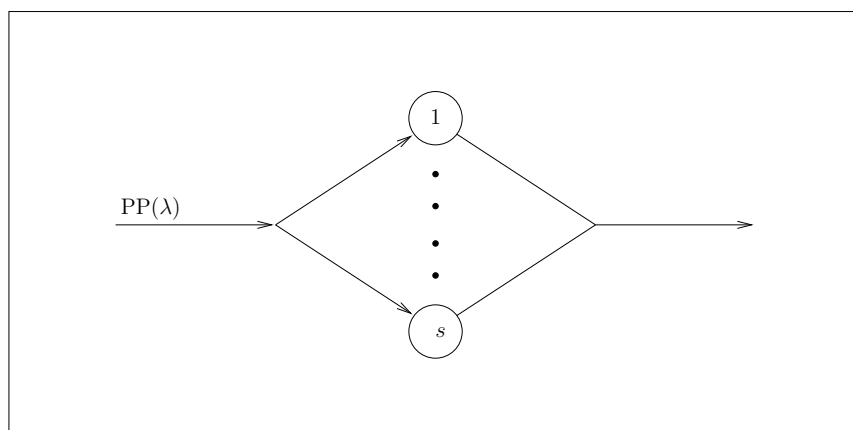
Dies heißt *Little's Gesetz* [Little'S law]. Die erwartete Anzahl von Anforderungen im System ist gleich der Ankunftsintensität multipliziert mit der erwarteten Verweilzeit im System. Dies ist intuitiv einsehbar. Dieser Zusammenhang gilt wesentlich allgemeiner auch für  $G/G/s/K$ -Systeme, die ein Gleichgewicht besitzen. Dieses muss allerdings geeignet definiert werden. Der Wert  $\lambda$  bedeutet dann

$$\lim_{T \rightarrow \infty} \frac{1}{T} (\text{Anzahl der Ankünfte bis zum Zeitpunkt } T).$$

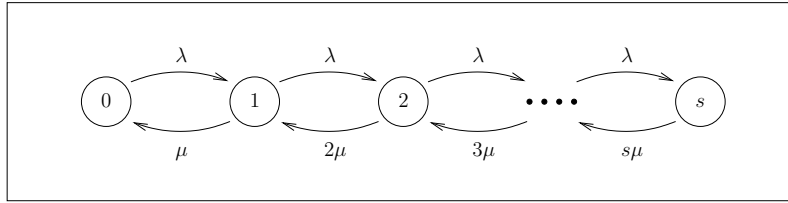
## 4.2 M/M/s/0

$s$  Server arbeiten parallel mit stochastisch unabhängigen  $\text{Exp}(\mu)$ -verteilten Bedienzeiten. Es gibt keine Wartekapazität (Puffer), neue Anforderungen werden abgewiesen oder blockiert, wenn alle Server belegt sind. Es handelt sich um ein *Verlust-* oder *Blockiersystem*.

Das Modell ist geeignet zur Beschreibung des Verkehrs in GSM-Zellen.  $s$  bedeutet die Anzahl verfügbarer Kanäle. Gespräche werden abgewiesen, wenn alle Kanäle belegt sind.



Betrachte  $X(t)$ , die Anzahl der Anforderungen im System zum Zeitpunkt  $t \geq 0$ . Das Modell für  $X(t)$  ist ein Geburts- und Todesprozess mit Intensitätsgraph



und Intensitätsmatrix

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & \dots & 0 & 0 \\ \mu & -(\mu + \lambda) & \lambda & \dots & 0 & 0 \\ 0 & 2\mu & -(2\mu + \lambda) & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -((s-1)\mu + \lambda) & \lambda \\ 0 & 0 & 0 & \dots & s\mu & -s\mu \end{pmatrix}.$$

Die Bestimmung der stationären Verteilung erfolgt mit Proposition 3.24:

$$\vartheta_0 = 1, \quad \vartheta_j = \frac{\lambda_0 \cdot \dots \cdot \lambda_{j-1}}{\mu_1 \cdot \dots \cdot \mu_j} = \frac{\lambda^j}{j! \mu^j}, \quad j = 1, \dots, s$$

Mit  $\rho = \frac{\lambda}{\mu}$  folgt

$$p_0^* = \frac{1}{\sum_{j=0}^s \vartheta_j} = \frac{1}{\sum_{j=0}^s \frac{\rho^j}{j!}}.$$

Insgesamt lautet die stationäre Verteilung

$$p_k^* = \frac{\frac{\rho^k}{k!}}{\sum_{j=0}^s \frac{\rho^j}{j!}}, \quad k = 0, 1, \dots, s \quad (\text{abgebrochene Poissonverteilung}).$$

Wie oben bezeichne die Zufallsvariable  $X^*$  mit der Verteilung  $\mathbf{p}^* = (p_0^*, \dots, p_s^*)$  den Systemzustand im Gleichgewicht. Die Wahrscheinlichkeit, dass eine ankommende Anforderung abgewiesen wird, beträgt dann

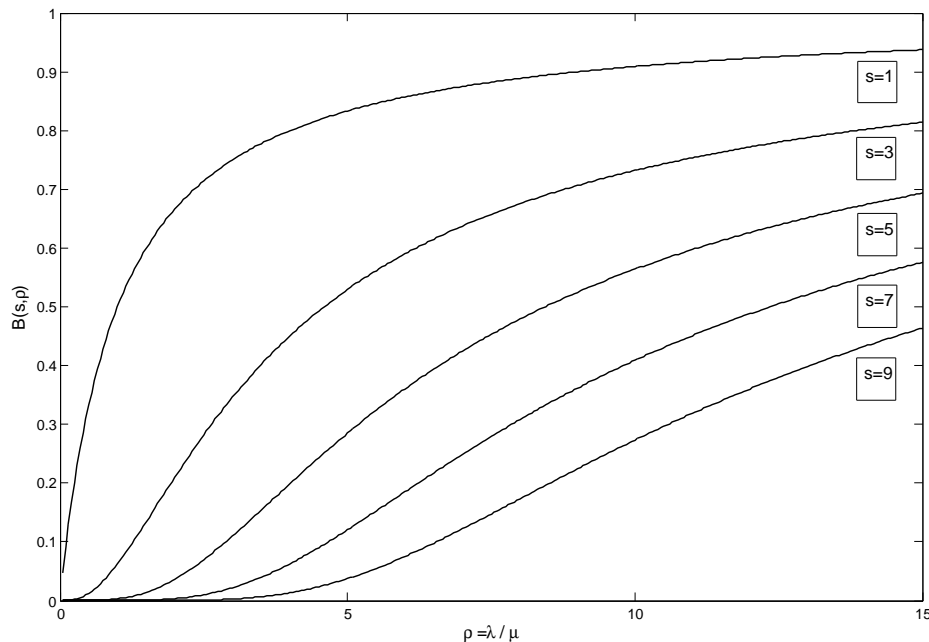
$$P(X^* = s) = \frac{\frac{\rho^s}{s!}}{\sum_{j=0}^s \frac{\rho^j}{j!}}.$$

Diese wichtige Formel heißt *Erste Erlangfunktion* oder *Erlang-Blockierwahrscheinlichkeit* (Erlang's B-formula, Erlang 1917). Bezeichnung:

$$E_1(s, \rho) = B(s, \rho) = \frac{\frac{\rho^s}{s!}}{\sum_{j=0}^s \frac{\rho^j}{j!}}$$

Der Parameter  $\rho$  wird mit der fiktiven Maßeinheit „Erlang“ bezeichnet. Die folgende Graphik zeigt einige Werte dieser Blockierwahrscheinlichkeit für verschiedene Werte von  $s$  und  $\rho$ :





Die Erlang-B-Formel gilt auch für  $M/G/s/0$ -Systeme mit beliebiger Bedienzeitverteilung mit endlichem Erwartungswert im „Gleichgewicht“. Genauer; ist  $X(t)$  die Anzahl der Anforderungen im  $M/G/s/0$ -System, so gilt

$$\lim_{t \rightarrow \infty} P(X(t) = k) = p_k^*, \quad k = 0, 1, \dots, s.$$

Beachte, dass man nicht von „Gleichgewicht“ im Sinn einer stationären Verteilung sprechen kann, da kein Markov-Prozess (Geburts- und Todesprozess) vorliegt.

**Beispiel 4.1.** Betrachtet wird eine Zelle in einem GSM-Mobilfunknetz. Die Gesprächsdauern seien stochastisch unabhängige, identisch verteilte Zufallsvariablen, Exp-verteilt mit Erwartungswert 2 [min]. Die Zwischenankunftszeiten seien Exp-verteilt, die erwartete Anzahl von Anforderungen pro Minute sei 8. Wieviele Kanäle werden benötigt, damit die Blockierwahrscheinlichkeit höchstens 1 % beträgt?

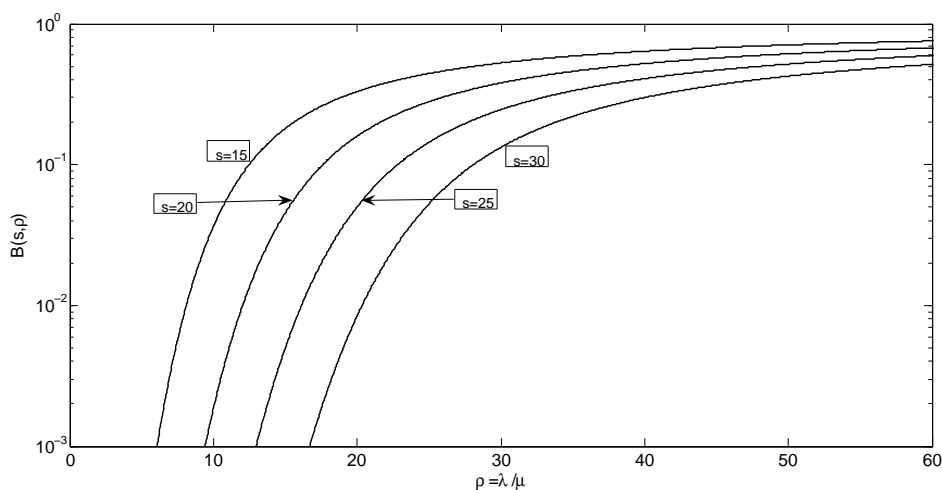
Zur Einbettung in das  $M/M/s/0$ -Modell entsprechen die Server den Kanälen, die Anforderungen den Gesprächen und die Bedienzeiten den Gesprächsdauern, die Zeiteinheiten sind Minuten. Da die erwartete Gesprächsdauer 2 beträgt, sind die Bedienzeiten  $\text{Exp}(\mu = \frac{1}{2})$ -verteilt. Die erwartete Anzahl von Anforderungen pro Minute ist  $\lambda = 8$ , also folgt mit Proposition 3.21, dass der Ankunftsstrom ein Poissonprozess mit Intensität  $\lambda = 8$  ist. Es folgt

$$\rho = \frac{\lambda}{\mu} = \frac{8}{\frac{1}{2}} = 16.$$

Zu lösen ist nun die Ungleichung

$$E_1(s, 16) \leq 0.01.$$

Das Ergebnis lautet  $s = 25$  (aus Kurven wie der nachfolgend abgebildeten, Programmen oder z.B. <http://www.cas.mcmaster.ca/~qiao/publications/erlang/newerlang.html>).



### 4.3 $M/M/\infty$

Es gibt unendliche viele Server, jede Anforderung geht sofort in Bedienung. Dieses System wird als Grenzfall von  $M/M/s$  mit  $s \rightarrow \infty$  behandelt. Es gilt mit Abschnitt 4.2

$$p_k^* = \lim_{s \rightarrow \infty} \frac{\frac{\rho^k}{k!}}{\sum_{j=0}^s \frac{\rho^j}{j!}} = e^{-\rho} \frac{\rho^k}{k!}, \quad k \in \mathbb{N}_0,$$

es liegt eine Poisson-Verteilung mit Parameter  $\rho$  vor, also  $X^* \sim \text{Poi}(\rho)$ . Jeder Kunde wird sofort bedient, so dass für die Antwortzeit  $W^* \sim \text{Exp}(\mu)$  gilt. Es folgt

$$E(X^*) = \rho, \quad E(W^*) = \frac{1}{\mu},$$

also

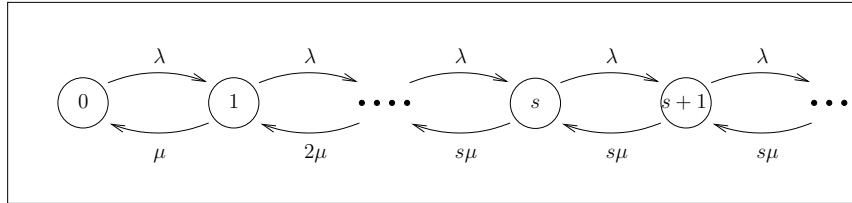
$$E(X^*) = \lambda E(W^*) \quad (\text{Little's Gesetz}).$$

## 4.4 M/M/s/∞

Modell:  $X(t)$  ist ein Geburts- und Todesprozess mit

$$\lambda_k = \lambda > 0, \quad k \in \mathbb{N}_0,$$

$$\mu_k = \begin{cases} k\mu, & k = 1, \dots, s \\ s\mu, & k > s \end{cases}.$$



Zur Bestimmung der stationären Verteilung ( $\sim X^*$ ) setze zunächst

$$\vartheta_0 = 1, \quad \vartheta_j = \frac{\lambda_0 \cdot \dots \cdot \lambda_{j-1}}{\mu_1 \cdot \dots \cdot \mu_j} = \begin{cases} \frac{1}{j!} \left(\frac{\lambda}{\mu}\right)^j, & j = 1, \dots, s-1 \\ \frac{1}{s!} \left(\frac{\lambda}{\mu}\right)^s \left(\frac{\lambda}{s\mu}\right)^{j-s}, & j \geq s \end{cases}.$$

Dann folgt mit  $\rho = \frac{\lambda}{s\mu}$

$$(p_0^*)^{-1} = \sum_{j=0}^{\infty} \vartheta_j = \sum_{j=0}^{s-1} \frac{1}{j!} \left(\frac{\lambda}{\mu}\right)^j + \underbrace{\sum_{j=s}^{\infty} \frac{1}{s!} \left(\frac{\lambda}{\mu}\right)^s \left(\frac{\lambda}{s\mu}\right)^{j-s}}_{\text{geometrische Reihe}} = \sum_{j=0}^{s-1} \frac{(s\rho)^j}{j!} + \frac{(s\rho)^s}{s!} \frac{1}{1-\rho},$$

falls  $\rho = \frac{\lambda}{s\mu} < 1$ . Es folgt

$$p_j^* = \begin{cases} \frac{(s\rho)^j}{j!} p_0^*, & j = 1, \dots, s-1 \\ \frac{(s\rho)^s}{s!} \rho^{j-s} p_0^* = \frac{s^s}{s!} \rho^j p_0^*, & j \geq s \end{cases}.$$

Weiterhin gilt im Gleichgewichtszustand:

- a) Die Wahrscheinlichkeit, dass ein ankommender Kunde in der Warteschlange warten muss, berechnet sich als

$$P(X^* \geq s) = p_0^* \frac{s^s}{s!} \sum_{j=s}^{\infty} \rho^j = p_0^* \frac{s^s}{s!} \rho^s \sum_{j=0}^{\infty} \rho^j = p_0^* \frac{s^s}{s!} \rho^s \frac{1}{1-\rho} =$$

$$E_2(s, \rho) = \frac{\frac{(s\rho)^s}{s!} \frac{1}{1-\rho}}{\sum_{j=0}^{s-1} \frac{(s\rho)^j}{j!} + \frac{(s\rho)^s}{s!} \frac{1}{1-\rho}}, \quad \rho = \frac{\lambda}{s\mu} < 1,$$

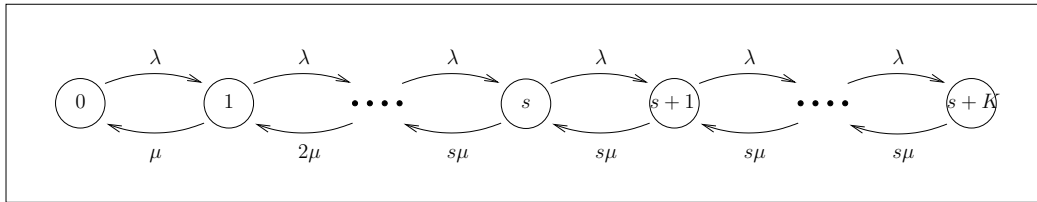
die sogenannte *Zweite Erlangfunktion*, auch als  $C(s, \rho)$  bezeichnet (Erlang's C-formula).

b) Die erwartete Warteschlangenlänge ergibt sich als

$$L_Q = \sum_{j=s+1}^{\infty} (j-s)P(X^* = j) = p_0^* \sum_{j=0}^{\infty} j \frac{s^s}{s!} \rho^{j+s} = p_0^* \frac{s^s}{s!} \rho^s \underbrace{\sum_{j=0}^{\infty} j \rho^j}_{= \frac{\rho}{(1-\rho)^2}} = p_0^* \frac{(s\rho)^s}{s!} \frac{\rho}{(1-\rho)^2}.$$

## 4.5 M/M/s/K

Modell:  $X(t)$  ist ein Geburts- und Todesprozess mit dem folgenden Intensitätsgraph:



Zur Berechnung der stationären Verteilung ( $\sim X^*$ ) setze wiederum

$$\vartheta_0 = 1, \quad \vartheta_j = \begin{cases} \frac{1}{j!} \left(\frac{\lambda}{\mu}\right)^j, & j = 1, \dots, s-1 \\ \frac{1}{s!} \left(\frac{\lambda}{\mu}\right)^s \left(\frac{\lambda}{s\mu}\right)^{j-s}, & j = s, \dots, s+K \end{cases}.$$

Mit  $\rho = \frac{\lambda}{s\mu}$  folgt

$$\begin{aligned} (p_0^*)^{-1} &= \sum_{j=0}^{s-1} \frac{1}{j!} \left(\frac{\lambda}{\mu}\right)^j + \sum_{j=s}^{s+K} \frac{1}{s!} \left(\frac{\lambda}{\mu}\right)^s \left(\frac{\lambda}{s\mu}\right)^{j-s} \\ &= \sum_{j=0}^{s-1} \frac{(s\rho)^j}{j!} + \frac{s^s}{s!} \rho^s \sum_{j=0}^K \rho^j = \sum_{j=0}^{s-1} \frac{(s\rho)^j}{j!} + \begin{cases} \frac{(s\rho)^s}{s!} \frac{1-\rho^{K+1}}{1-\rho}, & \rho \neq 1 \\ \frac{(s\rho)^s}{s!} (K+1), & \rho = 1 \end{cases}. \end{aligned}$$

Es folgt

$$p_j^* = \begin{cases} \frac{(s\rho)^j}{j!} p_0^*, & j = 1, \dots, s-1 \\ \frac{s^s}{s!} \rho^j p_0^*, & j = s, \dots, s+K \end{cases}.$$

Insbesondere gilt bei einem M/M/1/K-System

$$(p_0^*)^{-1} = \begin{cases} 1 + \rho \frac{1-\rho^{K+1}}{1-\rho} = \frac{1-\rho^{K+2}}{1-\rho}, & \rho = \frac{\lambda}{\mu} \neq 1 \\ 1 + \rho(K+1) = K+2, & \rho = \frac{\lambda}{\mu} = 1 \end{cases}.$$

Daraus folgt

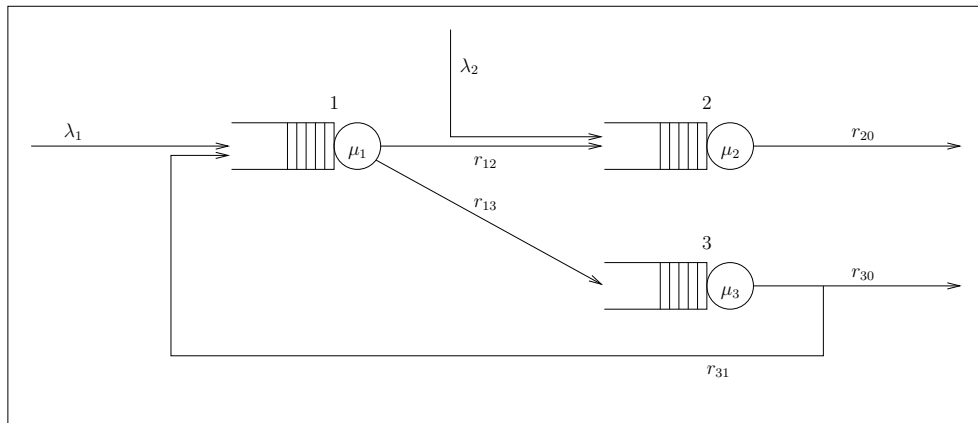
$$p_j^* = \begin{cases} \frac{1-\rho}{1-\rho^{K+2}} \rho^j, & \rho \neq 1 \\ \frac{1}{K+2}, & \rho = 1 \end{cases} \quad j = 0, \dots, K+1.$$

Im Fall  $\rho = 1$  ist die stationäre Verteilung eine Gleichverteilung.

# 5 Warteschlangennetze

Mit Markov-Prozessen lassen sich auch Warteschlangennetze modellieren und analysieren. Der Zustandsraum ist hierbei mehrdimensional, der Systemzustand ist die Anzahl von Anforderungen in jedem Knoten des Netzes. Mögliche Zustandswechsel sind: Ankünfte von Anforderungen an Knoten, Wechsel von Anforderungen zwischen den Knoten und Abgänge von Anforderungen bei Knoten.

**Beispiel 5.1.** Ein Warteschlangennetz mit  $J = 3$  Bedienstationen.



Zustandsraum:

$$\mathcal{S} = \{\mathbf{n} = (n_1, n_2, n_3) \mid n_i \in \mathbb{N}_0, i = 1, 2, 3\} = \mathbb{N}_0^3$$

Die Bedienknoten sind 1, 2, 3, „0“ bedeutet die „Außenwelt“.

Benötigt werden *Routingwahrscheinlichkeiten* [routing probabilities]

$r_{ij}$ : Wahrscheinlichkeit, dass eine Anforderung von Knoten  $i$  zu Knoten  $j$  wechselt,  $i = 1, \dots, 3, j = 0, \dots, 3$ .

Ankunftsströme werden charakterisiert durch Parameter  $\lambda_1, \dots, \lambda_3$ , Bedienintensitäten durch Parameter  $\mu_1, \dots, \mu_3$ . Der Systemzustand wird beschrieben durch einen Markov-Prozess  $\mathbf{X}(t) = (X_1(t), X_2(t), X_3(t))$  mit Zustandsraum  $\mathcal{S} = \mathbb{N}_0^3$ . Um geschlossene analytische Ergebnisse zu erhalten, sind bestimmte Einschränkungen an die Intensitäten nötig.

Begriffsbildung: Ein Netzwerk mit  $J$  Stationen heißt

- *offen* [open], wenn Zugänge und Abgänge bei gewissen Stationen möglich sind. Zustandsraum

$$\mathcal{S} = \mathcal{S}_O = \{\mathbf{n} = (n_1, \dots, n_J) \mid n_i \in \mathbb{N}_0, i = 1, \dots, J\} = \mathbb{N}_0^J.$$

- *geschlossen* [closed], wenn keine Zu- und Abgänge möglich sind und die Anzahl der Anforderungen im Netz konstant  $M$  ist. Zustandsraum

$$\mathcal{S} = \mathcal{S}_C = \{\mathbf{n} = (n_1, \dots, n_J) \mid n_i \in \mathbb{N}_0, \sum_{i=1}^J n_i = M\}.$$

Die Modellierung des Netzwerks erfolgt durch einen  $J$ -dimensionalen Markov-Prozess

$$\mathbf{X}(t) = (X_1(t), \dots, X_J(t)), \quad t \geq 0,$$

mit Intensitäten

$$q(\mathbf{n}, \mathbf{m}), \quad \mathbf{n}, \mathbf{m} \in \mathcal{S},$$

und stationärer Verteilung (falls diese existiert)

$$p^*(\mathbf{n}) = P(\mathbf{X}^* = \mathbf{n}) = P(X_1^* = n_1, \dots, X_J^* = n_J), \quad \mathbf{n} \in \mathcal{S},$$

wobei  $\mathbf{X}^* = (X_1^*, \dots, X_J^*)$  eine Zufallsvariable mit Verteilung  $p^*$  ist, die das Netz im Gleichgewicht beschreibt. Es bezeichne

$$p_i^*(n_i) = P(X_i^* = n_i) = \sum_{n_1, \dots, n_{i-1}, n_{i+1}, \dots, n_J} p^*(n_1, \dots, n_{i-1}, n_i, n_{i+1}, \dots, n_J)$$

die Wahrscheinlichkeit für  $n_i$  Anforderungen im Knoten  $i$  im Gleichgewicht.

Alle möglichen Zustandswechsel sind nun modellierbar, z.B. gleichzeitige Ankünfte oder Wechsel von mehreren Kunden. Eine explizite Form der stationären Verteilung, eine sogenannte Produktform, wird erzielt, wenn man sich auf Zustandswechsel mit nur einer Anforderung konzentriert.

Bezeichne  $\mathbf{e}_i = \underbrace{(0, \dots, 0, 1, 0, \dots, 0)}_{i\text{-te Komp. hat Wert 1}}$  den  $i$ -ten Einheitsvektor in  $\mathbb{R}^J$ . Die formale Beschreibung der Zustandswechsel erfolgt durch folgende Operatoren:

$$S_{ij}\mathbf{n} = \mathbf{n} - \mathbf{e}_i + \mathbf{e}_j, \quad \mathbf{n} \in \mathcal{S}, \quad n_i > 0, \quad 1 \leq i, j \leq J, \quad i \neq j$$

Shift von Station  $i$  nach  $j$ , eine Anforderung wechselt von  $i$  nach  $j$ .

$$A_i\mathbf{n} = \mathbf{n} + \mathbf{e}_i, \quad \mathbf{n} \in \mathcal{S}, \quad 1 \leq i \leq J$$

Ankunft einer Anforderung bei Station  $i$  von außen.

$$D_i\mathbf{n} = \mathbf{n} - \mathbf{e}_i, \quad \mathbf{n} \in \mathcal{S}, \quad n_i > 0, \quad 1 \leq i \leq J$$

Eine Anforderung verläßt das System bei Station  $i$ . Die zugehörigen Intensitäten werden bezeichnet mit

$$\begin{aligned} q(\mathbf{n}, S_{ij}\mathbf{n}) &= \mu_{ij}(\mathbf{n}) && \text{(Wechsel von } i \text{ nach } j) \\ q(\mathbf{n}, A_i\mathbf{n}) &= \lambda_i(\mathbf{n}) && \text{(Ankunft bei } i) \\ q(\mathbf{n}, D_i\mathbf{n}) &= \mu_{i0}(\mathbf{n}) && \text{(Wechsel von } i \text{ in die Außenwelt)} \end{aligned}$$

---

Ferner bezeichnet

$$\mu_i(\mathbf{n}) = \sum_{j=0}^J \mu_{ij}(\mathbf{n}), \quad i = 1, \dots, J,$$

die Bedienintensität bei Station  $i$  bei Belegung  $\mathbf{n}$  des Netzes.

**Definition 5.2.** Ein Netzwerk heißt *Jackson-Netz* (JN) [Jackson network], wenn für den beschreibenden Markov-Prozess  $\mathbf{X}(t) = (X_1(t), \dots, X_J(t))$ ,  $t \geq 0$ , folgendes gilt:

- (i)  $\lambda_i(\mathbf{n}) = \lambda_i$ ,  $i = 1, \dots, J$ .
- (ii)  $\mu_i(\mathbf{n}) = \mu_i(n_i)$ ,  $i = 1, \dots, J$ ,  $\mu_i(n_i) > 0$ , falls  $n_i > 0$ .
- (iii)  $\mu_{ij}(\mathbf{n}) = \mu_i(n_i)r_{ij}$ ,  $i = 1, \dots, J$ ,  $j = 0, \dots, J$ ,  $\mathbf{n} \in \mathcal{S}$ , wobei  $r_{ij} \geq 0$ ,  $\sum_{j=0}^J r_{ij} = 1$  für alle  $i$ .

Ein Jackson-Netz heißt *geschlossen* [closed], wenn  $\mathcal{S} = \mathcal{S}_C$ ,  $\lambda_i = 0$  für alle  $i = 1, \dots, J$ ,

- (iv)  $r_{i0} = 0$  für alle  $i = 1, \dots, J$  und  $\mathbf{R} = (r_{ij})_{1 \leq i, j \leq J}$  irreduzibel ist.

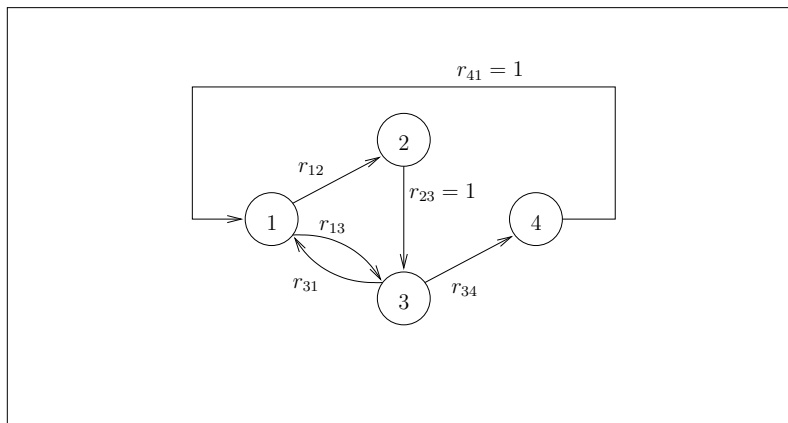
Es heißt *offen* [open], wenn  $\mathcal{S} = \mathcal{S}_O$ ,

- (v)  $\exists i : \lambda_i > 0$ ,  $\exists j : r_{j0} > 0$  und  $\mathbf{R} = (r_{ij})_{0 \leq i, j \leq J}$  irreduzibel ist, wobei  $r_{0k} = \frac{\lambda_k}{\sum_{l=1}^J \lambda_l}$ ,  $k = 1, \dots, J$ , und  $r_{00} = 0$ .

Interpretation der Bedingungen.

- (i): Die Ankunftsrate bei Station  $i$  ist unabhängig vom Zustand des Netzes.
- (ii): Die Bedienintensität bei Station  $i$  hängt nur ab von der Anzahl der Anforderungen in Station  $i$ .
- (iii):  $r_{ij}$  ist die Wahrscheinlichkeit, dass ein  $i$  verlassender Kunde nach  $j$  geht. Die Intensität für Zustandswechsel wird proportional aufgeteilt.
- (iv): Bei geschlossenen Jackson-Netzen gibt es keine Zugänge aus und Abgänge in die Außenwelt.
- (v): Die Wahrscheinlichkeiten, das Netz von außen bei Station  $k$  zu betreten, werden proportional zu den Ankunftsintensitäten gesetzt.

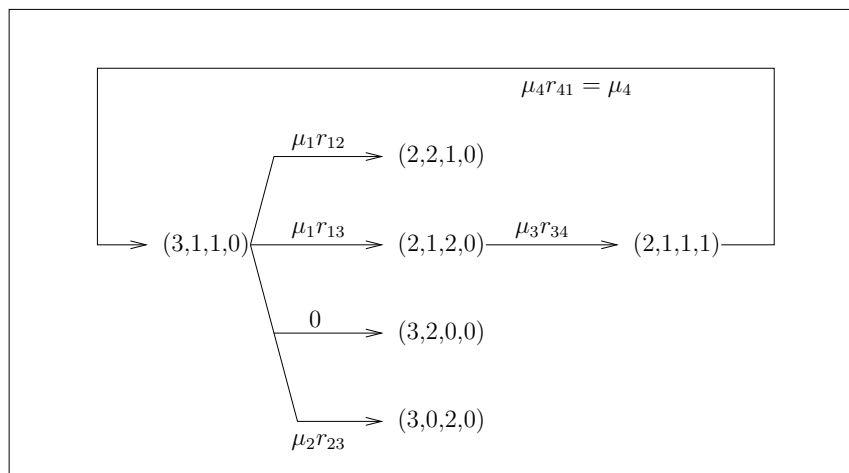
**Beispiel 5.3.** Ein geschlossenes Jackson-Netz aus  $J = 4$  Stationen mit  $M = 5$  Anforderungen.



$$\mathcal{S} = \{(5, 0, 0, 0), (4, 1, 0, 0), (4, 0, 1, 0), \dots\}$$

$$|\mathcal{S}| = \binom{M + J - 1}{J - 1} = \binom{5 + 4 - 1}{4 - 1} = 56$$

Ausschnitt aus dem Intensitätsgraphen:



## 5.1 Offene Jackson-Netze

Ansatz: Im Gleichgewicht sollte für jede Station  $i \in \{1, \dots, J\}$  gelten:

$$\Lambda_i = \lambda_i + \sum_{j=1}^J \Lambda_j r_{ji}, \quad i = 1, \dots, J, \quad (5.1)$$

wobei die Variablen  $\Lambda_i$  als Fluss aus dem Knoten  $i$  interpretiert werden können. Die rechte Seite repräsentiert den Gesamtfluss in den Knoten  $i$ . Beide Seiten sollten im Gleichgewicht übereinstimmen.



Fakt: Das Gleichungssystem 5.1 ist für offene Jackson-Netze eindeutig lösbar. Die eindeutige Lösung wird im Folgenden bezeichnet mit

$$\mathbf{\Lambda}^* = (\Lambda_1^*, \dots, \Lambda_J^*).$$

In Matrixform lautet das System 5.1 mit  $\mathbf{\Lambda} = (\Lambda_1, \dots, \Lambda_J)$ ,  $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_J)$ :

$$\begin{aligned} \mathbf{\Lambda} &= \boldsymbol{\lambda} + \mathbf{\Lambda} \hat{\mathbf{R}}, & \text{wobei } \hat{\mathbf{R}} &= (r_{ij})_{1 \leq i, j \leq J}, \\ \Leftrightarrow \mathbf{\Lambda} (\mathbf{I} - \hat{\mathbf{R}}) &= \boldsymbol{\lambda}. \end{aligned}$$

Unter den Voraussetzungen eines offenen Jackson-Netzes kann man nachweisen, dass  $\mathbf{I} - \hat{\mathbf{R}}$  regulär ist, also  $(\mathbf{I} - \hat{\mathbf{R}})^{-1}$  existiert.

Mit der eindeutigen Lösung  $\mathbf{\Lambda}^* = (\Lambda_1^*, \dots, \Lambda_J^*)$  setze

$$k_i = \left[ \sum_{l=0}^{\infty} \frac{(\Lambda_i^*)^l}{\mu_i(1) \cdot \dots \cdot \mu_i(l)} \right]^{-1}, \quad i = 1, \dots, J.$$

Das für  $l = 0$  im Nenner auftretende leere Produkt ist dabei als 1 definiert.

J.R. Jackson hat 1957 gezeigt, dass die stationäre Verteilung von offenen Jackson-Netzen sogenannte *Produktform* [product form] besitzt, vorausgesetzt sie existiert.

**Theorem 5.4.** Ein offenes Jackson-Netz besitzt eine eindeutige stationäre Verteilung  $\mathbf{p}^*$  genau dann, wenn  $\frac{1}{k_i} < \infty$  für alle  $i = 1, \dots, J$ . In diesem Fall gilt

$$p^*(\mathbf{n}) = p_1^*(n_1) \cdot \dots \cdot p_J^*(n_J), \quad \mathbf{n} = (n_1, \dots, n_J) \in \mathcal{S}_O = \mathbb{N}_0^J$$

mit

$$p_i^*(n_i) = k_i \cdot \frac{(\Lambda_i^*)^{n_i}}{\mu_i(1) \cdot \dots \cdot \mu_i(n_i)}, \quad i = 1, \dots, J.$$

Im Fall  $n_i = 0$  tritt im Nenner ein leeres Produkt auf, dieses ist als 1 definiert.

*Beweis.* Betrachte die Gleichgewichtsgleichungen des Markov-Prozesses. Allgemein:

$$\mathbf{p}^* \mathbf{Q} = \mathbf{0} \quad \Leftrightarrow \quad \underbrace{p_j^* \sum_{i \neq j} q_{ji}}_{\text{ausgehende Raten}} = \underbrace{\sum_{i \neq j} p_i^* q_{ij}}_{\text{eingehende Raten}} \quad \forall j \in \mathcal{S}.$$

Sei  $\mathbf{n} = (n_1, \dots, n_J) \in \mathcal{S}_O = \mathbb{N}_0^J$ . Dann lautet die obige Gleichung (mit den ausgehenden Raten in der ersten und den eingehenden Raten in der zweiten Zeile):

$$\begin{aligned} p^*(\mathbf{n}) & \left[ \sum_{j=1}^J q(\mathbf{n}, A_j \mathbf{n}) + \sum_{i=1}^J q(\mathbf{n}, D_i \mathbf{n}) + \sum_{i,j=1}^J q(\mathbf{n}, S_{ij} \mathbf{n}) \right] \\ &= \sum_{j=1}^J p^*(A_j \mathbf{n}) q(A_j \mathbf{n}, \mathbf{n}) + \sum_{i=1}^J p^*(D_i \mathbf{n}) q(D_i \mathbf{n}, \mathbf{n}) + \sum_{i,j=1}^J p^*(S_{ij} \mathbf{n}) q(S_{ij} \mathbf{n}, \mathbf{n}) \quad \forall \mathbf{n} \in \mathcal{S}_O. \end{aligned}$$

$\mathbf{p}^*$  ist Lösung dieses Gleichungssystems, wenn  $\mathbf{p}^*$  Lösung der *partiellen* Gleichgewichtsgleichungen ist, und zwar

$$p^*(\mathbf{n}) \left[ q(\mathbf{n}, D_i \mathbf{n}) + \sum_j q(\mathbf{n}, S_{ij} \mathbf{n}) \right] = p^*(D_i \mathbf{n}) q(D_i \mathbf{n}, \mathbf{n}) + \sum_j p^*(S_{ij} \mathbf{n}) q(S_{ij} \mathbf{n}, \mathbf{n}) \quad (5.2)$$

für alle  $i = 1, \dots, J$  und  $\mathbf{n} \in \mathcal{S}_O$  (partielle Gleichgewichtsgleichungen für Abgänge) und

$$p^*(\mathbf{n}) \sum_j q(\mathbf{n}, A_j \mathbf{n}) = \sum_j p^*(A_j \mathbf{n}) q(A_j \mathbf{n}, \mathbf{n}) \quad (5.3)$$

für alle  $\mathbf{n} \in \mathcal{S}_O$  (partielle Gleichgewichtsgleichungen für Zugänge).  
Beachte hierbei:

$$\begin{aligned} q(\mathbf{n}, D_i \mathbf{n}) &:= 0, \quad p^*(D_i \mathbf{n}) := 0, \quad \text{falls } D_i \mathbf{n} \notin \mathcal{S}_O \\ q(\mathbf{n}, S_{ij} \mathbf{n}) &:= 0, \quad p^*(S_{ij} \mathbf{n}) := 0, \quad \text{falls } S_{ij} \mathbf{n} \notin \mathcal{S}_O \end{aligned}$$

Gelte  $k_i^{-1} < \infty$  für alle  $i = 1, \dots, J$ . Das angegebene  $p^*(\mathbf{n})$  ist Lösung der partiellen Gleichgewichtsgleichungen, wie im Folgenden gezeigt wird. Zu Gleichung 5.2:

$$\begin{aligned} p^*(\mathbf{n}) \left[ \underbrace{q(\mathbf{n}, D_i \mathbf{n})}_{\substack{= \mu_{i0}(\mathbf{n}) \\ = \mu_i(n_i) r_{i0}}} + \sum_j \underbrace{q(\mathbf{n}, S_{ij} \mathbf{n})}_{\substack{= \mu_{ij}(\mathbf{n}) \\ = \mu_i(n_i) r_{ij}}} \right] &= p^*(D_i \mathbf{n}) \underbrace{q(D_i \mathbf{n}, \mathbf{n})}_{= \lambda_i} + \sum_j p^*(S_{ij} \mathbf{n}) \underbrace{q(S_{ij} \mathbf{n}, \mathbf{n})}_{= \mu_j(n_j+1) r_{ji}} \\ \Leftrightarrow \underbrace{p^*(\mathbf{n}) \mu_i(n_i)}_{= p^*(D_i \mathbf{n}) \Lambda_i^*} &= p^*(D_i \mathbf{n}) \lambda_i + \sum_j p^*(S_{ij} \mathbf{n}) \mu_j(n_j + 1) r_{ji} \\ \Leftrightarrow \Lambda_i^* &= \lambda_i + \sum_j \Lambda_j^* r_{ji} \quad \forall i = 1, \dots, J. \end{aligned}$$

Benutzt wird dabei im zweiten Schritt der Zusammenhang

$$\frac{p^*(S_{ij} \mathbf{n})}{p^*(D_i \mathbf{n})} \mu_j(n_j + 1) = \frac{p_i^*(n_i - 1) p_j^*(n_j + 1)}{p_i^*(n_i - 1) p_j^*(n_j)} \mu_j(n_j + 1) = \Lambda_j^*.$$

Die Beziehung  $p^*(\mathbf{n}) \mu_i(n_i) = p^*(D_i \mathbf{n}) \Lambda_i^*$  gilt wegen  $p_i^*(n_i) \mu_i = p_i^*(n_i - 1) \Lambda_i^*$  (Definition von  $p_i^*$ ).

Zu zeigen ist jetzt noch Gleichung 5.3. Durch Summation ergibt sich

$$\sum_i \Lambda_i^* = \sum_i \lambda_i + \sum_j \Lambda_j^* \sum_i r_{ji} = \sum_i \lambda_i + \sum_j \Lambda_j^* (1 - r_{j0}),$$

was äquivalent ist zu

$$\sum_i \lambda_i = \sum_i (\Lambda_i^* - \Lambda_i^* (1 - r_{i0})) = \sum_i \Lambda_i^* r_{i0}. \quad (5.4)$$

Es folgt

$$\begin{aligned}
 p^*(\mathbf{n}) \sum_j \underbrace{q(\mathbf{n}, A_j \mathbf{n})}_{=\lambda_j} &= \sum_j p^*(A_j \mathbf{n}) \underbrace{q(A_j \mathbf{n}, \mathbf{n})}_{=\mu_j(n_j+1)r_{j0}} \\
 \Leftrightarrow p^*(\mathbf{n}) \sum_j \lambda_j &= \sum_j p^*(A_j \mathbf{n}) \mu_j(n_j+1)r_{j0} \\
 \Leftrightarrow \sum_j \lambda_j &= \sum_j \underbrace{\frac{p^*(A_j \mathbf{n})}{p^*(\mathbf{n})} \mu_j(n_j+1)r_{j0}}_{=\Lambda_j^*} = \sum_j \Lambda_j^* r_{j0}.
 \end{aligned}$$

Letzteres gilt wegen Gleichung 5.4.

Da  $\{\mathbf{X}(t)\}_{t \geq 0}$  irreduzibel ist, gilt insgesamt:  $(p^*(\mathbf{n}))_{\mathbf{n} \in \mathcal{S}}$  ist ein stochastischer Vektor, der die Gleichgewichtsgleichungen erfüllt, also ist  $\mathbf{p}^*$  die eindeutige stationäre Verteilung.

Umgekehrt gilt:

$$r(\mathbf{n}) = r_1(n_1) \cdot \dots \cdot r_J(n_J), \quad \mathbf{n} = (n_1, \dots, n_J) \in \mathcal{S}_0$$

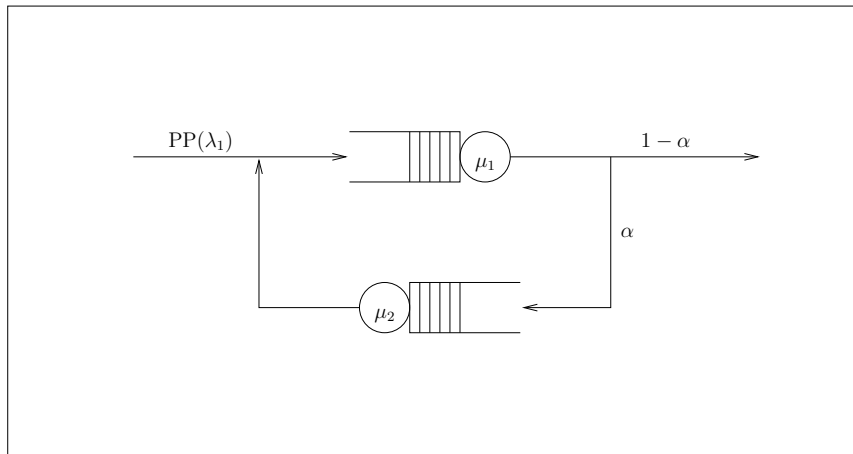
mit

$$r_i(n_i) = \frac{(\Lambda_i^*)^{n_i}}{\mu_i(1) \cdot \dots \cdot \mu_i(n_i)}, \quad n_i \in \mathbb{N}_0$$

löst die Gleichgewichtsgleichungen, wobei  $\Lambda_i^* > 0$  für alle  $i = 1, \dots, J$ . Also gilt  $r(\mathbf{n}) > 0$  für alle  $\mathbf{n} \in \mathcal{S}$  und  $(r(\mathbf{n}))_{\mathbf{n} \in \mathcal{S}}$  löst die Gleichgewichtsgleichungen. Da eine stationäre Verteilung  $\mathbf{p}^*$  existiert, folgt  $\sum_{l=0}^{\infty} r_i(l) < \infty$  für alle  $i$ , also  $k_i^{-1} < \infty$ .  $\square$

Mit Theorem 5.4 folgt: Im Gleichgewicht sind die Anzahlen von Kunden bei den einzelnen Stationen  $X_1(t), \dots, X_J(t)$  zu jedem festen  $t \in \mathbb{R}$  stochastisch unabhängig. Die Prozesse  $\{X_1(t)\}_{t \geq 0}, \dots, \{X_J(t)\}_{t \geq 0}$  sind nicht stochastisch unabhängig.

**Beispiel 5.5.**



Modellierung als offenes Jackson-Netz mit  $J = 2$  Stationen mit Routingmatrix

$$\mathbf{R} = \left( \begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ 1 - \alpha & 0 & \alpha & 0 \\ \hline 0 & 1 & 1 & 0 \end{array} \right),$$

wobei der rechte untere Teil die Matrix  $\hat{\mathbf{R}}$  darstellt, und Ankunftsraten  $\boldsymbol{\lambda} = (\lambda_1, \lambda_2) = (\lambda_1, 0)$ .

Zur Bestimmung der stationären Verteilung löse das Gleichungssystem

$$(\Lambda_1, \Lambda_2) = (\lambda_1, \lambda_2) + (\Lambda_1, \Lambda_2)\hat{\mathbf{R}}.$$

Es ergibt sich die Lösung

$$(\Lambda_1^*, \Lambda_2^*) = (\lambda_1, \lambda_2) (\mathbf{I} - \hat{\mathbf{R}})^{-1} = (\lambda_1, \lambda_2) \begin{pmatrix} 1 & -\alpha \\ -1 & 1 \end{pmatrix}^{-1} = (\lambda_1, 0) \frac{1}{1-\alpha} \begin{pmatrix} 1 & \alpha \\ 1 & 1 \end{pmatrix},$$

also

$$\Lambda_1^* = \frac{1}{1-\alpha} \lambda_1, \quad \Lambda_2^* = \frac{\alpha}{1-\alpha} \lambda_1.$$

Mit Theorem 5.4 folgt:

$$p_i^*(n_i) = k_i \cdot \frac{(\Lambda_i^*)^{n_i}}{\mu_i(1) \cdot \dots \cdot \mu_i(n_i)} = k_i \left( \frac{\Lambda_i^*}{\mu_i} \right)^{n_i}.$$

Also im Einzelnen:

$$p_1^*(n_1) = k_1 \left( \frac{\lambda_1}{(1-\alpha)\mu_1} \right)^{n_1} = (1-\rho_1)\rho_1^{n_1}, \quad n_1 \in \mathbb{N}_0, \quad \text{wobei } \rho_1 = \frac{\lambda_1}{(1-\alpha)\mu_1} \stackrel{!}{<} 1,$$

$$p_2^*(n_2) = k_2 \left( \frac{\alpha\lambda_1}{(1-\alpha)\mu_2} \right)^{n_2} = (1-\rho_2)\rho_2^{n_2}, \quad n_2 \in \mathbb{N}_0, \quad \text{wobei } \rho_2 = \frac{\alpha\lambda_1}{(1-\alpha)\mu_2} \stackrel{!}{<} 1.$$

Die stationäre Verteilung des gesamten Netzes hat die Produktform

$$p^*(n_1, n_2) = p_1^*(n_1) \cdot p_2^*(n_2) \sim \mathbf{X}^* = (X_1^*, X_2^*), \quad n_1, n_2 \in \mathbb{N}_0.$$

Beachte: Die Bedienrate  $\mu_2$  in Server 2 spielt keine Rolle für die stationäre Verteilung in Server 1. Mit  $\mu_2 \rightarrow \infty$  ergibt sich ein  $M/M/1$ -System mit Feedback.

## 5.2 Geschlossene Jackson-Netze

Geschlossene Jackson-Netze werden in der Literatur auch mit Gordon-Newell-Netzen bezeichnet. Sie werden beschrieben durch einen Markov-Prozess

$$\mathbf{X}(t) = (X_1(t), \dots, X_J(t)), \quad t \geq 0,$$

mit Zustandsraum

$$\mathcal{S} = \mathcal{S}_M = \left\{ \mathbf{n} \in \mathbb{N}_0^J \mid \sum_{i=1}^J n_i = M \right\}.$$

Im Gleichgewicht sollte für jede Station  $i$  gelten

$$\Lambda_i = \sum_{j=1}^J \Lambda_j r_{ji}, \quad i = 1, \dots, J, \tag{5.5}$$

wobei  $\Lambda_i$  als Fluss aus Knoten  $i$  interpretiert werden kann. Nach Voraussetzung ist  $\mathbf{R} = (r_{ij})_{1 \leq i, j \leq J}$  irreduzibel. Setze  $\mathbf{\Lambda} = (\Lambda_1, \dots, \Lambda_J)$ . Dann lautet das Gleichungssystem 5.5 in Matrixform

$$\mathbf{\Lambda} = \mathbf{\Lambda R}.$$

$\mathbf{R}$  ist eine stochastische Matrix, obige Gleichung ist die Bedingung für eine stationäre Verteilung  $\mathbf{\Lambda}$ . Aus den Ergebnissen zu endlichen Markov-Ketten folgt: Das System 5.5 besitzt eine eindeutige Lösung  $\mathbf{\Lambda}^* = (\Lambda_1^*, \dots, \Lambda_J^*)$  in der Menge der stochastischen Vektoren. Mit dieser Lösung wird die stationäre Verteilung des geschlossenen Jackson-Netzes konstruiert.

**Theorem 5.6.** (Gordon-Newell, 1967)

Ein geschlossenes Jackson-Netz besitzt die eindeutige stationäre Verteilung

$$p^*(\mathbf{n}) = K_M \prod_{i=1}^J \frac{(\Lambda_i^*)^{n_i}}{\mu_i(1) \cdot \dots \cdot \mu_i(n_i)}, \quad \mathbf{n} = (n_1, \dots, n_J) \in \mathcal{S}_M,$$

wobei  $K_M$  die normierende Konstante und  $\Lambda_1^*, \dots, \Lambda_J^* > 0$  eine Lösung des Systems 5.5 ist.

Der Beweis der Aussage geht analog zu dem von Theorem 5.4. Es werden die partiellen Gleichgewichtsgleichungen überprüft.

Man beachte:

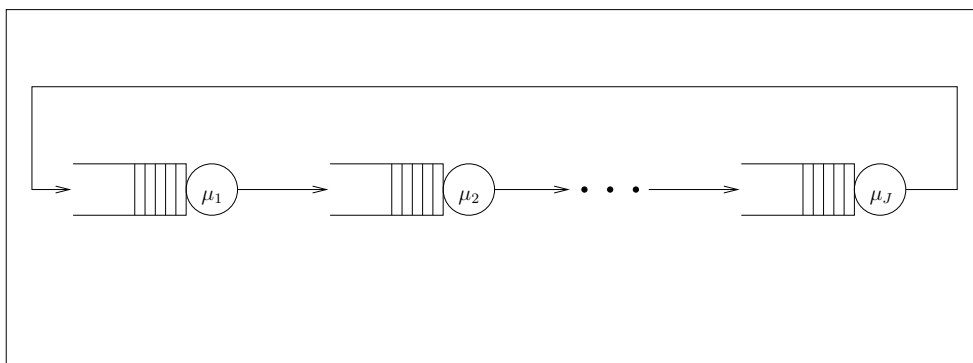
- a) Die stationäre Verteilung der Anzahl von Anforderungen in den Knoten  $1, \dots, J$  besitzt *Produktform* [product form]. Wie üblich bezeichne

$$\mathbf{X}^* = (X_1^*, \dots, X_J^*) \sim (p^*(\mathbf{n}))_{\mathbf{n} \in \mathcal{S}_M}.$$

Die Zufallsvariablen  $X_1^*, \dots, X_J^*$  sind nicht stochastisch unabhängig, da der Träger  $\mathcal{S}_M$  kein kartesisches Produkt ist, im Gegensatz zu dieser Eigenschaft bei den offenen Jackson-Netzen.

- b) Jede andere Lösung  $\Lambda'_1, \dots, \Lambda'_J$  des Systems 5.5 mit positiven Komponenten führt zu derselben stationären Verteilung  $p^*(\mathbf{n})$ , da sich die Konstante  $K_M$  entsprechend ändert.

**Beispiel 5.7.** (Geschlossene Tandemnetze)



Es gibt  $M$  Anforderungen im System. Die Bedienzeiten sind stochastisch unabhängig  $\text{Exp}(\mu_i)$ -verteilt ( $i = 1, \dots, J$ ).

Modell: Geschlossenes Jackson-Netz mit  $J$  Stationen und  $M$  Anforderungen.

$$r_{12} = r_{23} = \dots = r_{J1} = 1, \\ \mu_{ij}(\mathbf{n}) = \mu_i(n_i)r_{ij} = \mu_i, \quad j = i + 1, i = 1, \dots, J - 1 \text{ oder } j = 1, i = J.$$

Berechne  $\Lambda_i^*$  aus  $\Lambda_{i+1} = \Lambda_i$ ,  $i = 1, \dots, J - 1$ ,  $\Lambda_1 = \Lambda_J$ . Es ergibt sich die Lösung  $\Lambda_i^* = \frac{1}{j}$ ,  $i = 1, \dots, J$  (oder  $\Lambda_i' = 1$ ,  $i = 1, \dots, J$ ). Mit Theorem 5.6 folgt:

$$p^*(\mathbf{n}) = K_M \prod_{i=1}^J \frac{1}{\mu_i(1) \cdot \dots \cdot \mu_i(n_i)} = K_M \prod_{i=1}^J \left( \frac{1}{\mu_i} \right)^{n_i}, \quad n_i \geq 0, \sum_{i=1}^J n_i = M,$$

mit

$$K_M = \left[ \sum_{\substack{n_i \geq 0 \\ n_1 + \dots + n_J = M}} \frac{1}{\mu_1^{n_1} \cdot \dots \cdot \mu_J^{n_J}} \right]^{-1}.$$

Das Beispiel zeigt folgendes Problem bei der Berechnung der Konstanten  $K_M$  in geschlossenen Jackson-Netzen:

$$K_M^{-1} = \sum_{\mathbf{n} \in \mathcal{S}_M} p^*(\mathbf{n}) \quad \text{mit} \quad |\mathcal{S}_M| = \binom{J + M - 1}{M}.$$

Für großes  $J$  und  $M$  liegt eine große Anzahl kleiner Summanden vor, durch Rundungsfehler entstehen unbrauchbare Ergebnisse. Hierzu:

### Rekursive Berechnung von $K_M$ , falls $\mu_i(n_i) = \mu_i$ , $i = 1, \dots, J$

Es gilt

$$p^*(\mathbf{n}) = K_M \prod_{i=1}^J \left( \underbrace{\frac{\Lambda_i^*}{\mu_i}}_{=\gamma_i} \right)^{n_i} = K_M \prod_{i=1}^J \gamma_i^{n_i}, \quad K_M = \left[ \sum_{\mathbf{n} \in \mathcal{S}_M} \prod_{i=1}^J \gamma_i^{n_i} \right]^{-1}.$$

Setze, falls  $|z| < \frac{1}{\gamma_i}$ ,  $i = 1, \dots, J$ ,

$$G_J(z) = \prod_{i=1}^J \frac{1}{1 - \gamma_i z} = \prod_{i=1}^J \sum_{l=0}^{\infty} (\gamma_i z)^l \stackrel{\text{„Cauchy-Summe“}}{=} \sum_{l=0}^{\infty} \underbrace{\left( \sum_{\mathbf{n} \in \mathcal{S}_l} \gamma_1^{n_1} \cdot \dots \cdot \gamma_J^{n_J} \right)}_{=C_l(J) = \frac{1}{K_l}} z^l = \sum_{l=0}^{\infty} C_l(J) z^l \tag{5.6}$$

$C_l(J) = \frac{1}{K_l}$  sind die Koeffizienten der Potenzreihe. Gesucht ist  $C_M(J)$ .

Für  $G_j(z) = \prod_{i=1}^j \frac{1}{1 - \gamma_i z}$ ,  $j = 1, \dots, J$ , gilt folgende Rekursion:

$$G_1(z) = \frac{1}{1 - \gamma_1 z} = \sum_{l=0}^{\infty} \gamma_1^l z^l = \sum_{l=0}^{\infty} C_l(1) z^l, \tag{5.7}$$

$$G_j(z) = G_{j-1}(z) \frac{1}{1 - \gamma_j z}, \quad j = 2, \dots, J.$$

Mit Gleichung 5.6 ergibt sich

$$\sum_{l=0}^{\infty} C_l(j) z^l (1 - \gamma_j z) = \sum_{l=0}^{\infty} C_l(j-1) z^l$$

und hieraus

$$\sum_{l=0}^{\infty} C_l(j) z^l = \sum_{l=0}^{\infty} C_l(j-1) z^l + \gamma_j \sum_{l=0}^{\infty} C_l(j) z^{l+1}.$$

Durch Koeffizientenvergleich folgt die Rekursion

$$C_l(j) = C_l(j-1) + \gamma_j C_{l-1}(j), \quad l = 1, \dots, M, \quad j = 2, \dots, J$$

mit den Anfangsbedingungen

$$C_l(1) = \gamma_1^l, \quad l = 1, \dots, M \quad (\text{aus Gleichung 5.7})$$

$$C_0(j) = 1, \quad j = 2, \dots, J \quad (\text{aus Gleichung 5.6})$$

Folgende Tabelle beschreibt das Rekursionsschema für  $C_l(j)$ :

$l \setminus j$	1	2	...	$J$
0	1	1	...	1
1	$\gamma_1$			$\vdots$
2	$\gamma_1^2$			$\vdots$
$\vdots$	$\vdots$	$C_l(j) = C_l(j-1) + \gamma_j C_{l-1}(j)$		$\vdots$
$M$	$\gamma_1^M$	...	...	$C_M(J)$

Beachte: Nur die letzte aktuelle Spalte muss gespeichert werden.

Insgesamt: Mit Aufwand proportional zu  $J \cdot M$  stehen zur Verfügung  $C_l(J)$ ,  $l = 0, \dots, M$ , wobei  $C_M(J) = \frac{1}{K_M}$ . Hiermit können weitere Größen berechnet werden. Modelliere hierzu  $(X_1^*, \dots, X_J^*) \sim p^*$  das geschlossene Jackson-Netz im Gleichgewicht. Es gilt:

$$\rho_i = u_i(M) = P(X_i^* > 0) = \gamma_i \frac{C_{M-1}(J)}{C_M(J)}$$

(Auslastung des  $i$ -ten Knotens, utilization, traffic intensity)

$$P(X_i^* \geq k) = \gamma_i^k \frac{C_{M-k}(J)}{C_M(J)}, \quad k = 0, \dots, M$$

(Komplementäre Verteilungsfunktion)

$$E(X_i^*) = \frac{1}{C_M(J)} \sum_{l=1}^M \gamma_i^l C_{M-l}(J)$$

(Erwartungswerte)

$$\lambda_i = \mu_i P(X_i^* > 0) = \Lambda_i^* \frac{C_{M-1}(J)}{C_M(J)}$$

(Durchsatz bei Knoten  $i$ )





# 6 Allgemeine Bedienzeitverteilung, $M/G/1$

Modellkomponenten sind der Ankunftsprozess vom Typ  $PP(\lambda)$  (Poissonprozess mit Intensität  $\lambda$ ) und die Bedienzeiten  $\{Y_n\}_{n \in \mathbb{N}}$  i.i.d. (stochastisch unabhängig identisch verteilt), diese Komponenten sind stochastisch unabhängig.

Bezeichnungen:

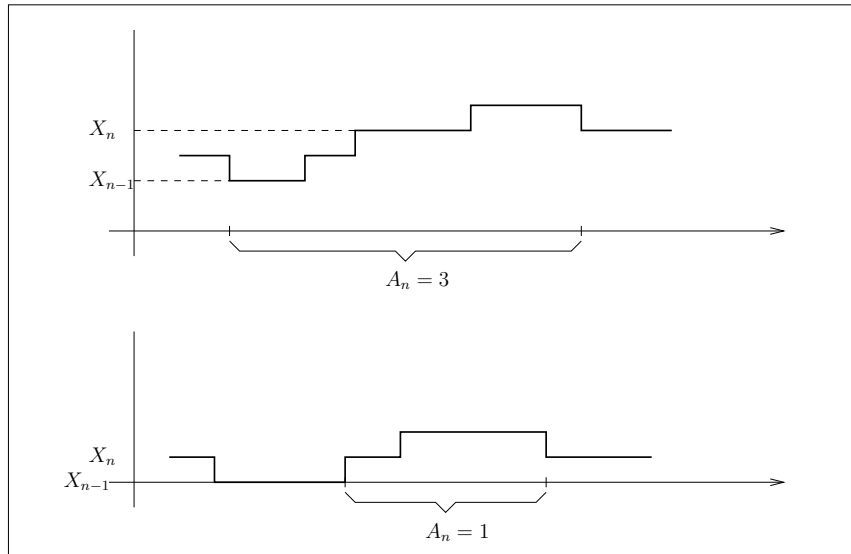
- $Y_n \sim G$  (Verteilungsfunktion, nicht notwendigerweise eine Exponentialverteilung).
- $E(Y_n) = \nu$ ,  $\mu = \frac{1}{\nu}$  heißt Bedienrate (in Anlehnung an  $M/M/1$ ).
- $X(t)$ ,  $t \geq 0$ : Anzahl Anforderungen im System zum Zeitpunkt  $t$  (im allgemeinen *kein* Markov-Prozess).
- $D_n$ : Abgangszeitpunkt der  $n$ -ten Anforderung.
- $X_n = X(D_n+)$ : Anzahl Anforderungen im System sofort nach dem Abgang der  $n$ -ten Anforderung.
- $A_n$ : Anzahl der Anforderungen, die während der Bedienzeit der  $n$ -ten Anforderung ankommen. Die  $A_n$  sind stochastisch unabhängig, da der Ankunftsprozess ein Poisson-Prozess ist.

Es gilt

$$\begin{aligned} X_n &= \begin{cases} X_{n-1} + A_n - 1, & \text{falls } X_{n-1} > 0 \\ A_n, & \text{falls } X_{n-1} = 0 \end{cases} \\ &= (X_{n-1} - 1)^+ + A_n, \quad n \in \mathbb{N}, \end{aligned}$$

wobei  $x^+ = \max\{x, 0\}$  den Positivteil von  $x \in \mathbb{R}$  bezeichnet.

Skizze:



Setze  $X_0 = 0$  oder eine beliebige von  $\{A_n\}_{n \in \mathbb{N}}$  stochastisch unabhängige Zufallsvariable. Die so definierte Folge  $\{X_n\}_{n \in \mathbb{N}_0}$  bildet eine Markov-Kette, da die Wahrscheinlichkeit für das Auftreten des aktuellen Zustands nur abhängt vom Zustand zum direkten Vorgängerzeitpunkt und von stochastisch unabhängigen Zufallsvariablen  $A_n$ .

## 6.1 $M/G/1$ – Systemzustand im Gleichgewicht

**Theorem 6.1.** Die oben definierte Folge  $\{X_n\}_{n \in \mathbb{N}_0}$  bildet eine homogene Markov-Kette mit Zustandsraum  $\mathcal{S} = \mathbb{N}_0$  und Übergangswahrscheinlichkeiten

$$p_{ij} = P(X_n = j \mid X_{n-1} = i) = \begin{cases} \gamma_{j-i+1}, & i \geq 1, j \geq i - 1 \\ \gamma_j, & i = 0, j \geq 0 \\ 0, & \text{sonst} \end{cases}, \quad i, j \in \mathbb{N}_0,$$

wobei

$$\gamma_k = P(A_n = k) = \int_0^\infty \frac{(\lambda y)^k}{k!} e^{-\lambda y} dG(y), \quad k \in \mathbb{N}_0.$$

Beachte:  $\int_0^\infty f(y) dG(y)$  bezeichnet das Stieltjes-Integral. Ist  $G$  absolut-stetig mit Dichte  $g$ , so gilt

$$\int_0^\infty f(y) dG(y) = \int_0^\infty f(y) g(y) dy.$$

In Theorem 6.1 ist nur noch zu klären, dass die Übergangswahrscheinlichkeiten die angegebene Form haben. Dies sieht man aus

$$\begin{aligned} P(X_n = i + k \mid X_{n-1} = i) &= P(A_n = k + 1), \quad \text{falls } i \geq 1, k \geq -1, \\ P(X_n = k \mid X_{n-1} = 0) &= P(A_n = k), \quad \text{falls } k \geq 0. \end{aligned}$$

Ferner gilt

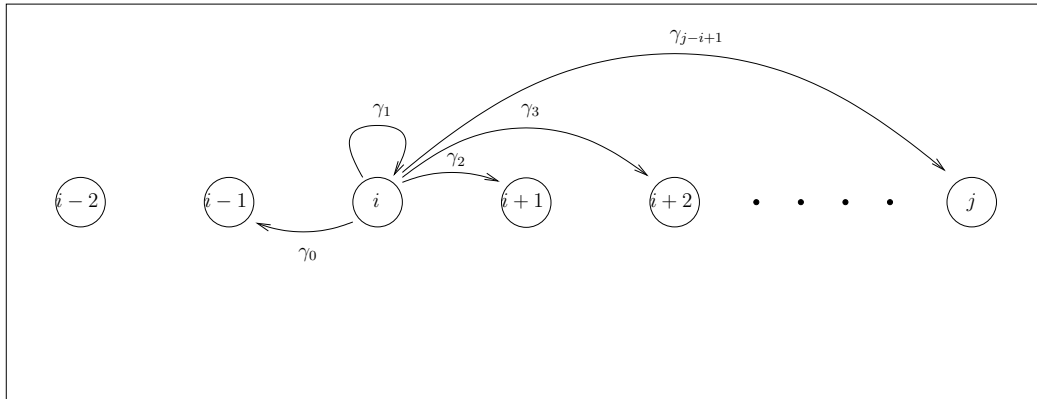
$$P(A_n = k) = \int_0^\infty P(A_n = k \mid Y_n = y) dG(y) = \int_0^\infty \frac{(\lambda y)^k}{k!} e^{-\lambda y} dG(y),$$

da der Ankunftsprozess ein Poissonprozess ist und die Anzahl der Ankünfte in einem Intervall fester Länge  $y$  daher  $\text{Poi}(\lambda y)$ -verteilt ist.

Die Übergangsmatrix von  $\{X_n\}_{n \in \mathbb{N}_0}$  hat also die Gestalt

$$\mathbf{\Pi} = \begin{pmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \gamma_3 & \dots \\ \gamma_0 & \gamma_1 & \gamma_2 & \gamma_3 & \dots \\ 0 & \gamma_0 & \gamma_1 & \gamma_2 & \dots \\ 0 & 0 & \gamma_0 & \gamma_1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad \text{mit } \gamma_j = \mathbb{E} \left( \frac{(\lambda Y)^j}{j!} e^{-\lambda Y} \right).$$

Übergangsgraph (Ausschnitt):



$\mathbf{\Pi}$  ist irreduzibel und aperiodisch, da  $\gamma_j > 0$  und  $p_{jj} > 0$ .  $\mathbf{\Pi}$  ist positiv-rekurrent, falls für die *Auslastung* [utilization]  $\rho$  gilt

$$\rho = \lambda \nu = \frac{\lambda}{\mu} < 1.$$

Also besitzt  $\{X_n\}_{n \in \mathbb{N}_0}$  eine eindeutige stationäre Verteilung  $\mathbf{p}^* = (p_i^*)_{i \in \mathbb{N}_0}$  mit

$$p_i^* = \lim_{n \rightarrow \infty} P(X_n = i) = \lim_{n \rightarrow \infty} p_{ji}^{(n)} \quad \forall j \in \mathbb{N}_0.$$

Man kann  $\mathbf{p}^*$  aus dem Gleichungssystem  $\mathbf{p}\mathbf{\Pi} = \mathbf{p}$  bestimmen, was allerdings nicht ganz einfach ist.

Die Verteilung  $\mathbf{p}^*$  repräsentiert das Gleichgewicht einer speziellen eingebetteten Markov-Kette, den Anzahlen von Anforderungen kurz nach Abgangzeitpunkten. Gleichzeitig ist  $\mathbf{p}^*$  die Grenzverteilung des ursprünglichen Prozesses  $X(t)$ ,  $t \geq 0$ .

**Theorem 6.2.** Im M/G/1-System mit  $\rho = \lambda \nu < 1$  besitzt  $X(t)$ ,  $t \geq 0$ , die Grenzverteilung  $\mathbf{p}^* = (p_i^*)_{i \in \mathbb{N}_0}$ , d.h.

$$\lim_{t \rightarrow \infty} P(X(t) = i) = p_i^* \quad \forall i \in \mathbb{N}_0.$$

Die Aussage von Theorem 6.2 gilt nur für  $M/G/1$ -Systeme, sie kann nicht ohne weiteres auf allgemeine Systeme übertragen werden. Das Problem ist die explizite Bestimmung der stationären Verteilung  $\mathbf{p}^*$ . Eine geschlossene Form erhält man für die erzeugende Funktion. Berechne hierzu

$$A(z) = \sum_{k=0}^{\infty} \gamma_k z^k, \quad |z| \leq 1, \quad \gamma_k = \mathbb{E} \left[ \frac{(\lambda Y)^k}{k!} e^{-\lambda Y} \right],$$

die erzeugende Funktion von  $(\gamma_0, \gamma_1, \gamma_2, \dots)$ , und

$$L_Y(s) = \mathbb{E}(e^{-sY}) = \int e^{-sy} dG(y),$$

die Laplace-Transformierte der Bedienzeit  $Y$  ( $\sim Y_n$ ).

Man kann zeigen, dass

$$A(z) = L_Y(\lambda(1-z)), \quad |z| \leq 1.$$

Ferner gilt der *Satz von Pollaczek-Khintchine* (*P-K transform equation*):

**Proposition 6.3.** Im  $M/G/1$ -System mit  $\rho = \lambda\nu < 1$  gilt für die erzeugende Funktion der stationären Verteilung  $\mathbf{p}^* = (p_i^*)_{i \in \mathbb{N}_0}$

$$H(z) = (1-\rho)(1-z) \frac{L_Y(\lambda(1-z))}{L_Y(\lambda(1-z)) - z}, \quad |z| \leq 1.$$

Insbesondere gilt  $p_0^* = 1 - \rho$ .

Aus der Rekursion  $X_n = (X_{n-1} - 1)^+ + A_n$  kann die *Pollaczek-Khintchine-Formel* (*P-K mean-value formula*) direkt abgeleitet werden:

**Proposition 6.4.** Im  $M/G/1$ -System mit  $\rho = \lambda\nu < 1$  gilt

$$\mathbb{E}(X^*) = \rho + \rho^2 \frac{1 + c_Y^2}{2(1-\rho)},$$

wobei  $c_Y = \frac{\sigma}{\nu} = \frac{\sqrt{\text{V}(Y)}}{\mathbb{E}(Y)}$  den Variationskoeffizienten der Bedienzeitverteilung bezeichnet.

Mit Littles Gesetz folgt

$$\mathbb{E}(W^*) = \frac{1}{\lambda} \mathbb{E}(X^*) = \nu + \frac{\lambda(\sigma^2 + \nu^2)}{2(1-\rho)}.$$

Interessant ist:  $\mathbb{E}(W^*)$  und  $\mathbb{E}(X^*)$  sind beide monoton steigend in  $\sigma^2$ . Beide sind minimal, wenn  $\sigma^2 = 0$  gilt, also  $Y = \text{const}$  (deterministische Bedienzeiten).

## 6.2 Wartezeitverteilungen bei $M/G/1$

Die Verteilung der Wartezeit hängt von der Bedienstrategie ab. Exemplarisch werden hier die Ergebnisse für FCFS und LCFS angegeben. Bei LCFS bleibt die aktuelle Anforderung in Bedienung bis sie vollständig bearbeitet ist.

**M/G/1 – FCFS**

**Proposition 6.5.** Im M/G/1-FCFS-System mit  $\rho = \lambda\nu < 1$  gilt im Gleichgewicht

$$L_W(s) = \frac{(1-\rho)sL_Y(s)}{s-\lambda+\lambda L_Y(s)}, \quad s \geq 0,$$

$$L_{W_Q}(s) = \frac{(1-\rho)s}{s-\lambda+\lambda L_Y(s)}, \quad s \geq 0,$$

wobei  $L_W$  die Laplace-Transformierte der Antwortzeit und  $L_{W_Q}$  die der Wartezeit in der Schlange bezeichnet.

Aus Proposition 6.5 folgt

$$L_W(\lambda(1-z)) = \frac{(1-\rho)\lambda(1-z)L_Y(\lambda(1-z))}{\lambda(1-z)-\lambda+\lambda L_Y(\lambda(1-z))} = H(z), \quad |z| \leq 1.$$

Es gilt für die linke Seite

$$\begin{aligned} L_W(\lambda(1-z)) &= \int_0^\infty e^{-\lambda(1-z)y} dF_W(y) \\ &= \int_0^\infty e^{-\lambda y} e^{\lambda z y} dF_W(y) \\ &= \int_0^\infty e^{-\lambda y} \sum_{k=0}^\infty \frac{(\lambda y)^k z^k}{k!} dF_W(y) \\ &= \sum_{k=0}^\infty z^k \int_0^\infty e^{-\lambda y} \frac{(\lambda y)^k}{k!} dF_W(y). \end{aligned}$$

Die Vertauschung von Integration und Summation ist hier möglich.

Nach Definition gilt für die rechte Seite (erzeugende Funktion der stationären Verteilung)

$$H(z) = \sum_{k=0}^\infty z^k p_k^*.$$

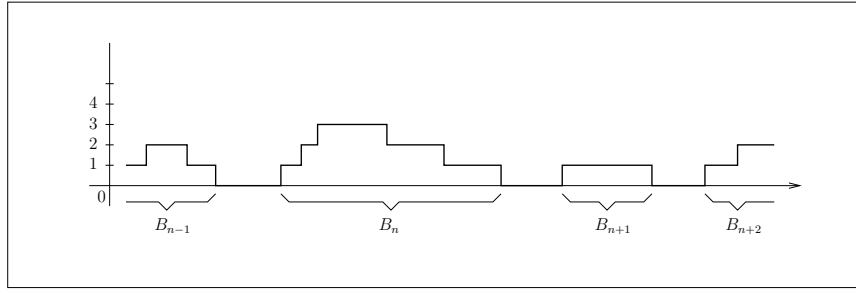
Durch Koeffizientenvergleich der beiden Gleichungen folgt

$$p_k^* = \int_0^\infty e^{-\lambda y} \frac{(\lambda y)^k}{k!} dF_W(y).$$

Die stationäre Verteilung kann also aus der Wartezeitverteilung bestimmt werden.

**M/G/1 – LCFS**

Zur Bestimmung der Wartezeitverteilung wird die Laplace-Transformierte der Beschäftigungsperioden im Gleichgewicht benötigt. Die Verteilung der Beschäftigungsperioden ist unabhängig von der Bedienstrategie. Graphisch:



Die Beschäftigungsperioden  $B_n$  besitzen dieselbe Verteilung, also  $B_n \sim B$ . Ihre Laplace-Transformierte erfüllt die Funktionalgleichung

$$L_B(s) = L_Y(s + \lambda - \lambda L_B(s)), \quad s \geq 0.$$

Diese hat eine eindeutige Lösung  $L_B(s) \leq 1$ , gehörend zu einer nicht degenerierten Verteilung, falls  $\rho = \lambda\nu < 1$ .

**Proposition 6.6.** Im  $M/G/1$ -System mit  $\rho < 1$  und nicht arithmetischer Bedienzeitverteilung  $G_Y$  gilt im Gleichgewicht

$$L_W(s) = L_Y(s) \left( 1 - \rho + \frac{\lambda(1 - L_B(s))}{s + \lambda(1 - L_B(s))} \right), \quad s \geq 0,$$

$$L_{W_Q}(s) = 1 - \rho + \frac{\lambda(1 - L_B(s))}{s + \lambda(1 - L_B(s))}, \quad s \geq 0.$$

Die Verteilung einer Zufallsvariable  $X$  heißt *arithmetisch*, wenn ein  $\alpha > 0$  existiert mit

$$P(X \in \{\alpha k \mid k \in \mathbb{Z}\}) = 1.$$

Für die Wartezeiten bei  $M/M/1$ -Systemen (Ankunftsprozess  $PP(\lambda)$ , Bedienzeiten  $\text{Exp}(\mu)$ -verteilt,  $\lambda < \mu$ ) gilt

$$L_Y(s) = \frac{\mu}{\mu + s}, \quad s \geq 0,$$

$$L_B(s) = \frac{\lambda + \mu + s - \sqrt{(\lambda + \mu + s)^2 - 4\lambda\mu}}{2\lambda}, \quad s \geq 0.$$

Daraus ergibt sich für FCFS

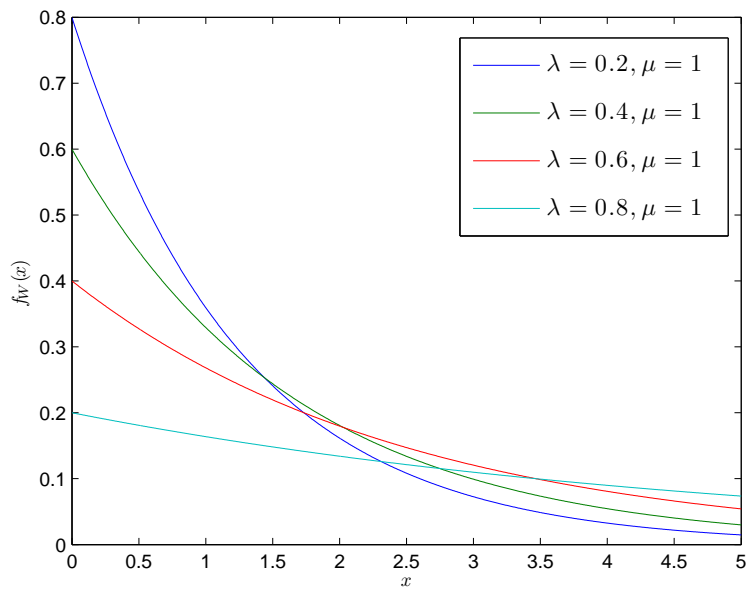
$$L_W(s) = \frac{\mu - \lambda}{s + \mu - \lambda}, \quad s \geq 0, \quad (\rightarrow \text{Exp}(\mu - \lambda))$$

und für LCFS

$$L_W(s) = \frac{s - \lambda + 3\mu - \sqrt{s^2 + 2s(\lambda + \mu) + (\mu - \lambda)^2}}{2(\mu + s)}, \quad s \geq 0.$$

Die asymptotische erwartete Wartezeit ist identisch (siehe Littles Gesetz), höhere Momente nicht.

Die Verteilung der Antwortzeit bei  $M/M/1$ -FCFS als Graphik:



Die Verteilung der Antwortzeit bei  $M/M/1$ -LCFS als Graphik (Dichten nach numerischer Inversion der Laplace-Transformierten):

