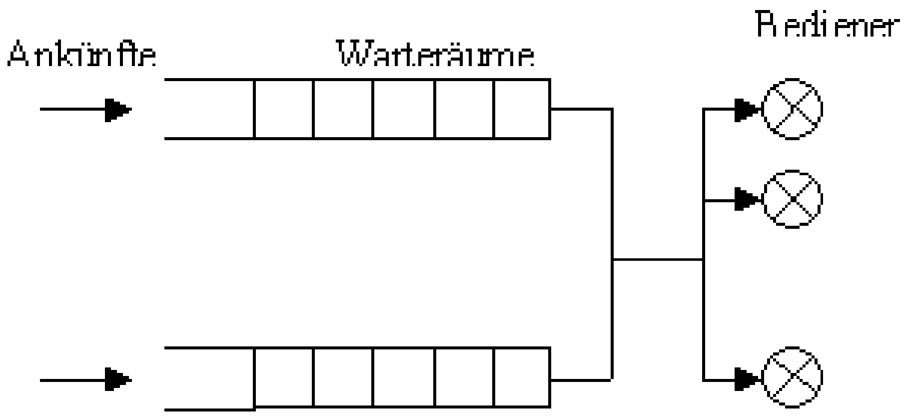
Kapitel 3
Isolierte Warteschlangensysteme im Gleichgewicht.
§ 3.1 Grundlegende Systeme

Ein isoliertes Warteschlangensystem ('Queue') hat folgende Aufbau:
• es gibt ein oder mehrere Warteräume
• es gibt sukzessive Ankünfte von Aufträgen
• daraus entstehen Warteschlangen
• es gibt ein oder mehrere Bediener
Kennzeichnend sind dabei folgende Komponenten:
• $(A_t)_{t\geq 0}$ Ankunftsprozesse
• $(B_t)_{t\geq 0}$ Bedienprozesse
• Bediendisziplin, Zahl der Bediener
• Kapazität der Warteräume
• Größe der Benutzerpopulation



Im folgenden betrachten wir Warteschlangen, die zu KHMK $(N_t)_{t\geq 0}$führen.
Ein Zustand $i \in I = \mathbb{N}_0$ der KHMK steht hier für die Information, daß sich i Benutzer im System (wartend oder in Bedienung) befinden.
Hauptinteresse der Warteschlangentheorie ist es, den Prozess $(N_t)_{t\geq 0}$ vollständig und genau zu beschreiben. Ausreichend sind dabei meist die Zustandswahrscheinlichkeiten $\mathbb{P}[N_t = i] = p_i(t)$ für $i \in I = \mathbb{N}_0$ und $t \in \mathbb{R}_0^+$
Zu Angabe von Warteschlangen wird oft die Kendall-Notation (D.G. Kendall, 1951) verwendet. Sie charakterisiert ein Warteschlangensystem durch ein 6-Tupel A/B/C/K/P/S:
• A:Ankunftsprozeß
• B: Bedienprozeß
• C: Anzahl der Bediener
• K: Maximale Anzahl der Jobs im System
• P: Größe der Benutzerpopulation
• S: Scheduling-Disziplin (Warteschlangendisziplin)

Beispiele für die Einträge bei A und B sind
• M: Markov-Prozess (bzw. Exponential-Verteilung)
• D: deterministische Verteilung (konstanter Wert!)
• E_k : k – stufige Erlangverteilung (Faltung von k Exponentialverteilungen)
• H_k : k – phasige Hyperexponentialverteilung
• h_k : k – phasige Hypoexponentialverteilung
• G: beliebige Verteilung
• GI: unabhängige Bedienzeiten beliebige Verteilung
P=∞ und K=∞ werden meist nicht angegeben, ebenso S=FIFO/FCFS, also wird etwa M/M/1/∞/∞/FIFO kurz als M/M/1-System bezeichnet.
Zunächst betrachten wir M/M/m-Systeme:
• exponentiell verteilte Zwischenankunftszeiten, mit Ankunftsrate λ
• exponentiell verteilte Bedienzeiten, mit Bedienraten μ
• *m* Bediener
• Warteraum nicht beschränkt
• Benutzerpopulation nicht beschränkt
• Warteschlangen-Disziplin FCFS (First-Come-First-Served)

Berechnungsmethoden für die Gleichgewichtsverteilung am Beispiel des M/M/1 Modells

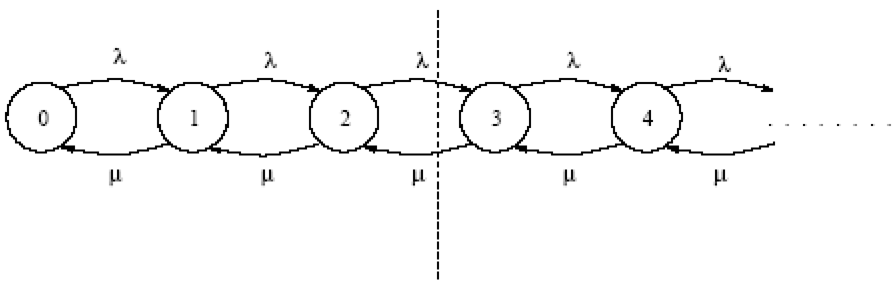
Hier sind die Zwischenankunftzeiten (Parameter λ) und die Bedienungszeiten (Parameter μ) exponentialverteilt.
Bei m=1 ist die Zeit bis zur nächsten Ankunft bzw. dem nächsten Abgang exponentiell verteilt, damit ist (Beweis analog zur überlagerung zweier Poisson-Prozesse):
$\mathbb{P}$[zwei Ereignisse bis Zeit Δt]=o(Δt), Δt→0
Also ist $\lambda_{i\,j} = 0$für j∉{i-1,i+1}, d.h. es liegt ein ***Geburts-Todes-Prozess*** (**GTP**) vor.
Für i=0 ist offensichtlich

$$\lambda_i = \lambda_{i\,i+1} = \lim_{\Delta t \to 0} \frac{\mathbb{P}[\text{eine Ankunft, kein Abgang bis zur Zeit } \Delta t]}{\Delta t}$$
$$= \lim_{\Delta t \to 0} \frac{\left(1-e^{-\lambda \Delta t}\right) e^{-\mu \Delta t}}{\Delta t} = \lim_{\Delta t \to 0} \frac{1-e^{-\lambda \Delta t}}{\Delta t} =$$
$$= \lim_{\Delta t \to 0} \frac{\lambda e^{-\lambda t}}{1} = \lambda$$

und analog

$$\mu_i = \lambda_{i\,i-1} = \mu$$

Bei M/M/1 haben wir folgende Zustandsübergänge:



Die Gleichgewichtsverteilung (**pΛ**=**0**) wird nun durch Lösung von

$-\lambda\, p_0 + \mu p_1 = 0$
$\lambda\, p_0 - (\lambda + \mu)\, p_1 + \mu\, p_2 = 0$
… … … … … … … … … … … …
$\lambda\, p_{n-2} + (\lambda + \mu)\, p_{n-1} + \mu\, p_n = 0$
… … … … … … … … … … … …
bestimmt.

1. Iterative Methode

Diese haben wir bereits beim allgemeinen Geburts- und Todesprozess kennengelernt. Wir erhalten für alle $n \in \mathbb{N}_0$

$$p_n = \left(\frac{\lambda}{\mu}\right)^n p_0.$$

Wie im allgemeinen Fall ergibt sich mit $\rho=\frac{\lambda}{\mu}$

$$p_0 = \left(1 + \sum_{n=1}^{\infty} \prod_{i=1}^{n} \frac{\lambda}{\mu}\right)^{-1}$$
$$=(1 + \sum_{n=1}^{\infty} \rho^n)^{-1}$$

Im Fall ρ gibt es genau eine Gleichgewichtsverteilung, nähmlich

$$p_n = \rho^n (1 - \rho).$$

- Der Quotient ρ heißt ***Nutzungsfaktor*** oder ***Verkehrsintensität***, er gibt den durchschnittlichen Zeitanteil an, in dem der Bediener belegt ist.
- Die Wahrscheinlichkeit dafür, daß der Bediener belegt ist, heißt ***Utilization*** (Beanspruchung/Auslastung) U des Bedieners. Bei M/M/1 gilt also speziell $U = 1 - p_0 = \rho$
- Die Anzahl der Kunden im System ist bei M/M/1 also ***geometrisch verteilt*** mit Parameter ρ.

2. Erzeugende Funktionen

Das Gleichungssystem, bei dem es sich ja um Differenzengleichungen handelt, kann auch durch erzeugende Funktionen gelöst werden. Dazu wird die (Wahrscheinlichkeits-) erzeugende Funktion (z-Transformierte)

$$P(z):=\sum_{n=0}^{\infty} p_n\, z^n$$

mit zϵ $\mathbb{C}$ und $|z|\leq 1$ eingeführt. Für diese wird aus dem gleichungssystem eine Formel gefunden, aus der dann durch Reihenentwicklung die Koeffizienten p_n gewonnen werden. Die Methode soll nun am M/M/1 Modell vorgeführt werden. Das Gleichungssystem umgeschrieben für ρ lautet

$$p_{n+1} = (\rho + 1)\, p_n - \rho\, p_{n-1},\ n\geq 1$$
$$p_1 = \rho\, p_0.$$

Daraus folgt durch Multiplikation mit z^n und Summation über n

$$z^{-1}\, p_{n+1}\, z^{n+1} = (\rho + 1)\, p_n\, z^n - z\, \rho\, p_{n-1}\, z^{n-1}$$
$$z^{-1} \sum_{n=1}^{\infty} p_{n+1}\, z^{n+1} = (\rho + 1) \sum_{n=1}^{\infty} p_n\, z^n - z\, \rho \sum_{n=1}^{\infty} p_{n-1}\, z^{n-1}$$
$$z^{-1}(P(z) - p_1\, z - p_0) = (\rho + 1)\, (P(z) - p_0) - z\, \rho\, P(z)$$

und daraus mit Hilfe der zweiten Gleichung $p_1 = \rho\, p_0$und Auflösen nach P(z)

$$P(z)=\frac{p_0}{1-z\,\rho}.$$

Da wegen der Definition von P(z) die Beziehung P(1)=1 gilt, ist $\rho\neq 1$ nötig, und es folgt $p_0 = 1 - \rho$ und damit $\rho<1$ und wir erhalten

$$P(z)=\frac{1-\rho}{1-z\,\rho},\ \ \rho < 1,\ \ |\, z\,| \leq 1$$

Indem man nun P(z) in eine Potenzreihe entwickelt, erhält man in diesem Fall sehr schnell $p_n = \rho^n (1 - \rho)$.

Häufig muss man sich damit begnügen, interessanten Größen aus der erzeugenden Funktion direkt herzuleiten, z.B. ***die erwartete Anzahl L=𝔼[N] von Kunden im System***

$$L = \sum_{n=1}^{\infty} n\, p_n = \sum_{n=1}^{\infty} n\, p_n\, z^{n-1}\, |_{z=1} = \frac{d}{d\,z} \sum_{n=1}^{\infty} p_n\, z^n\, |_{z=1} = P'(1)$$

In unserem Fall ergibt sich

$$L=\frac{\rho}{1-\rho}$$

3. Operatoren

Die Gleichgewichtsverteilung kann auch mit Hilfe von Operatoren gefunden werden, die auf Folgen angewendet werden. Ist etwa $(a_0, a_1, a_2, \ldots)$ eine Folge, so wird durch $D\,a_n = a_{n+1} (n \in \mathbb{N})$ der Schift-Operator definiert, der die Folge $(a_0, a_1, a_2, \ldots)$ auf die Folge $(a_1, a_2, a_3, \ldots)$ abbildet (einfache Shift). Man beachte, dass $D^2 = D \circ D$ eine Folge $(a_0, a_1, a_2, \ldots)$ auf die Folge $(a_2, a_3, \ldots)$ abbildet (doppelte Shift). Eine lineare Differenzengleichung hat nun die Form

$$c_n\, a_n + c_{n+1}\, a_{n+1} + \ldots + c_{n+k}\, a_{n+k} = 0$$

was sich mittels Shift-Operator als

$$\left(\sum_{i=0}^{k} c_{n+i}\, D^i\right) a_n = 0$$

schreibt, weil $D^m\, a_n = a_{n+m}$ für alle n, m gilt. Ist nun r eine Wurzel (Nullstelle) des Operator-Polynoms $\sum_{i=0}^{k} c_{n+i}\, D^i$, d.h.

$$\sum_{i=0}^{k} c_{n+i}\, D^i = (D - r) \sum_{i=0}^{k-1} \tilde{c}_{n+i}\, D^i,$$

so gilt

$$\begin{aligned}
\left(\sum_{i=0}^{k} c_{n+i}\, D^i\right) r^n &= (D - r) \left(\sum_{i=0}^{k-1} \tilde{c}_{n+i}\, D^i\right) r^n \\
&= (D - r) \sum_{i=0}^{k-1} \tilde{c}_{n+i}\, r^{n+i} \\
&= D \sum_{i=0}^{k-1} \tilde{c}_{n+i}\, r^{n+i} - r \sum_{i=0}^{k-1} \tilde{c}_{n+i}\, r^{n+i} \\
&= \sum_{i=0}^{k-1} \tilde{c}_{n+i}\, r^{n+i+1} - \sum_{i=0}^{k-1} \tilde{c}_{n+i}\, r^{n+i+1} \\
&= 0
\end{aligned}$$

d.h. $a_n = r^n$ ist eine Lösung der Differenzengleichung. Alle Lösungen der Differenzengleichung sind Linearkombinationen der Gestalt $a_n = \sum_i \alpha_n\, r_i^n$, wobei die r_iWurzeln des Operator-Polinoms sind.

Wir wenden nun die Methode auf unsere Rekursionsgleichung

$$p_{n+2} - (\rho + 1)\, p_{n+1} + \rho\, p_n = 0 \qquad n \geq 0$$

mit den Nebenbedingungen $p_1 = \rho\, p_0$ und $\sum_{n=0}^{\infty} p_n = 1$ an.

Mittels Shift-Operator schreibt sich die Rekursionsgleichung als

$$\begin{aligned}
0 &= \left(D^2 - (\rho + 1)\, D + \rho\right) p_n \\
&= (D - 1)\, (D - \rho)\, p_n
\end{aligned}$$

was auf $p_n = \alpha + \beta \rho^n$ führt. Aus der Nebenbedingung $\sum_{n=0}^{\infty} p_n = 1$ folgt α=0 (die Nebenbedingung $p_1 = \rho\, p_0$ist damit automatisch erfüllt) und ebenso ρ<1 und $p_n = \rho^n (1 - \rho)$.

Sind die p_n berechnet, so können alle interessierenden Kenngrößen des Bedienungssystems im Gleichgewicht berechnet werden, z.B. ***die erwartete Anzahl von Kunden im System***

$$\begin{aligned}
L = \mathbb{E}[N] &= \sum_{n=1}^{\infty} n\, p_n = (1 - \rho) \sum_{n=1}^{\infty} n\, \rho^n \\
&= (1 - \rho)\, \rho \frac{d}{d\,\rho} \sum_{n=0}^{\infty} \rho^n \\
&= (1 - \rho)\, \rho \frac{d}{d\,\rho} \frac{1}{1-\rho} \\
&= \frac{\rho}{1-\rho} = \frac{\lambda}{\mu - \lambda}
\end{aligned}$$

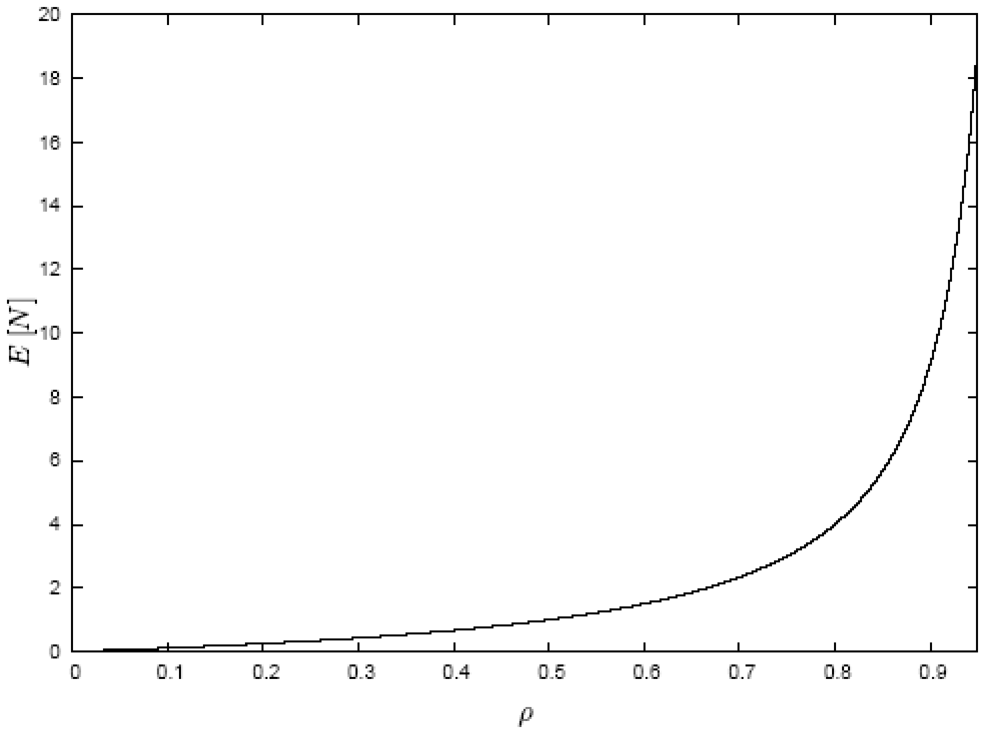


Figure Mean Number of Customers vs. Utilization

Ebenso ist auch ***die erwartete Anzahl von wartenden Kunden*** interessant,

$$L_q = \mathbb{E}[N_q] = \sum_{n=1}^{\infty} (n-1)\, p_n = \frac{\rho}{1-\rho} - (1-p_0) = \frac{\rho^2}{1-\rho}$$

Interessant ist auch die erwartete Anzahl von Kunden in der Warteschlange, wenn diese nicht leer ist,

$$\tilde{L}_q = \mathbb{E}[N_q \mid N > 1] = \frac{1}{1-\rho}$$

Aus dem ***Satz von der totalen Wahrscheinlichkeit*** erhält man also

$\mathbb{E}[N_q] = \mathbb{E}[N_q \mid N = 0]\, \mathbb{P}[N = 0] + \mathbb{E}[N_q \mid N = 1]\, \mathbb{P}[N = 1] + \mathbb{E}[N_q \mid N > 1]\, \mathbb{P}[N > 1]$

wobei $\mathbb{E}[N_q \mid N = 0] = \mathbb{E}[N_q \mid N = 1] = 0$, so gilt

$\mathbb{E}[N_q] = \mathbb{E}[N_q \mid N > 1]\, \mathbb{P}[N > 1]$ oder

$$\begin{aligned} \mathbb{E}[N_q \mid N > 1] &= \frac{\mathbb{E}[N_q]}{\mathbb{P}[N>1]} \\ &= \frac{\mathbb{E}[N_q]}{1-p_0-p_1} \\ &= \frac{\mathbb{E}[N_q]}{\rho-\rho(1-\rho)} \\ &= \frac{\mathbb{E}[N_q]}{\rho^2} \\ &= \frac{1}{1-\rho} \end{aligned}$$

Satz (Symmetrie von Geburts/Todes-Prozessen im Gleichgewicht): $(N_t)_{t\geq 0}$ sei ein GTP mit Gleichgewichtsverteilung. Definiere

r_k= Wahrscheinlichkeit, daß der Prozess im Zustand k ist unmittelbar vor einem Übergang der Form $i \to i+1$ (d.h. Ankunft)

d_k= Wahrscheinlichkeit, daß der Prozess im Zustand k ist unmittelbar nach einem Übergang der Form $i+1 \to i$ (d.h. Abgang)

Dann gilt $r_k = d_k$ für alle $k \geq 0$

Bemerkung: Bei M/M/1 ist

r_k= $\mathbb{P}$[Ankömmling findet k Benutzer im System] und

d_k= $\mathbb{P}$[Abgänger hinterläßt k Benutzer im System].

Beweis. Bezeichnen wir mit x_n die Zeitpunkten unmittelbar vor einem Übergang der Form $i \to i+1$, und mit y_n - unmittelbar nach einem Übergang der Form $i+1 \to i$.
Bezeichnen wir nun N_{x_n} mit α_n und N_{y_n} mit β_n.
1) Zeigen wir zuerst, falls $\beta_{n+i} \le k$, dann $\alpha_{n+k+1} \le k$.
Nehmen wir an, daß $\beta_{n+i} \le k$ (für beliebige $k \ge 0$, $n>0$). D.h. der Abgänger hinterlässt höchstens k Benuter.
Sei
a=die Anzahl der Ankünfte zur Zeit y_{n+i}=die Anzahl der Ankünfte unmittelbar vor ($n+i$)-ten Abgang.
Es gilt

$$a + i - (n + i) = \beta_{n+i} \le k$$
$$a \le n + k$$

und dann ergibt es sich

$$y_{n+i} \le x_{n+k+1}$$

So treten mindestens $n+i$ Abgänge vor ($n+k+1$)-ten Ankunft ein, d.h.

$$\alpha_{n+k+1} \le n + k + 1 - (n + i) = k.$$

2) Zeigen wir jetzt, falls $\alpha_{n+k+1} \le k$, dann $\beta_{n+i} \le k$.
Nehmen wir an, daß $\alpha_{n+k+1} \le k$ (für beliebige $k \ge 0$, $n>0$). D.h. der Ankömmling findet höchstens k Benuter.
Sei
b=die Anzahl der Abgänge zur Zeit x_{n+k+1}=die Anzahl der Abgänge unmittelbar vor ($n+k+1$)-ten Ankunft.
Es gilt

$$n + k + i - d = \alpha_{n+k+1} \le k$$
$$n + i \le d$$

und dann ergibt es sich

$$y_{n+i} \le x_{n+k+1}$$

So treten höchstens $n+k$ Ankünfte vor ($n+i$)-ten Abgang ein, d.h.

$$\beta_{n+i} \le n + k + i - (n + i) = k.$$

3) Aus 1) und 2) für k≥0 erhält man

$$\mathbb{P}[\beta_{n+i} \le k] = \mathbb{P}[\alpha_{n+k+1} \le k]$$

Im Gleichgewicht wenn n gross ist

$$\lim_{n\to\infty} \mathbb{P}[\beta_n \le k] = \lim_{n\to\infty} \mathbb{P}[\beta_{n+i} \le k] = \lim_{n\to\infty} \mathbb{P}[\alpha_{n+k+1} \le k] = \lim_{n\to\infty} \mathbb{P}[\alpha_n \le k]$$

oder

$$d_k = r_k$$

Satz (Random-Observer-Property bzw PASTA-Eigenschaft): $(N_t)_{t\ge 0}$ sei ein stochastischer Prozeß, der ein Warteschlangensystem mit einem Puasson-Ankunftsstrom beschreibt. $(N_t)_{t\ge 0}$ besitze die Gleichgewichtsverteilung p_k=Wahrscheinlichkeit, daß der Prozess im Zustand k ist (wobei k nicht notwendigerweise die Zahl der Jobs im System darstellen muss).

r_k=sei wieder die Wahrscheinlichkeit dafür, daß ein Ankömmling das System im Zustand k vorfindet.

Dann gilt $r_k = d_k = p_k$

Beweis.
Wir brauchen zu zeigen, daß
$r_k = \mathbb{P}[N_t = k \mid N_{t+\Delta t} - N_t = 1] = \mathbb{P}[N_t = k] = p_k$

$\mathbb{P}[N_{t+\Delta t} = k+1 \mid N_t = k] = \mathbb{P}[N_{t+\Delta t} - N_t = 1 \mid N_t - N_0 = k] \quad =$

$= \frac{\mathbb{P}[N_{t+\Delta t}-N_t=1, N_t-N_0=k]}{\mathbb{P}[N_t-N_0=k]}$

$= \frac{\mathbb{P}[N_{t+\Delta t}-N_t=1]\,\mathbb{P}[N_t-N_0=k]}{\mathbb{P}[N_t-N_0=k]}$

$= \mathbb{P}[N_{t+\Delta t} - N_t = 1] = \lambda\Delta t + o(\Delta t)$

gilt wegen ***der Unabhängigkeit der Zuwächse*** für den Poisson-Prozess $(N_t)_{t\geq 0}$.
Es folgt von hier, daß zwei Ereignisse $N_{t+\Delta t} - N_t = 1$ und $N_t = k$ unabhängig voneinander sind, d.h.
$r_k = \mathbb{P}[N_t = k \mid N_{t+\Delta t} - N_t = 1] = \mathbb{P}[N_t = k] = p_k = d_k$

Wartezeiten beim M/M/1 Modell

1. Aus der Formel von Little können die erwartete Systemzeit W und die erwartete Wartezeit W_q in der Warteschlange gewonnen werden, nämlich

$$W=\mathbb{E}[Q]=\frac{L}{\lambda}=\frac{1}{\mu-\lambda} \text{ und } W_q = \mathbb{E}[Q_q] = \frac{L_q}{\lambda}=\frac{\lambda}{\mu(\mu-\lambda)}.$$

2. Diese Formeln kann auch direkt wie folgt gewonnen werden.
Wir brauchen die folgende Notation zu verwenden:
B - die Zufallsvariable der Bedienzeit einer Kunde (unabhängig, exponentiellverteilt mit Parameter μ)
$F_B(t) = \mathbb{P}[B < t] = 1 - e^{-\mu t}$
Q_q - die Zufallsvariable der Wartezeit im Puffer bis zum Bedienungsanfang
N - die Zufallsvariable der Anzahl der Kunden im System
N_q - die Zufallsvariable von wartenden Kunden (wenn N=0, dann $N_q = 0$; wenn N>0, dann $N = N_q + 1$)
R - die Zufallsvariable der Restbedienzeit (R=0, wenn N=0; R>0 sonst)

Offenbar gilt es $Q_q = R + N_q B$

Aus dem Satz von der totalen Wahrscheinlichkeit erhält man
$\mathbb{E}[Q_q]=\sum_{n=0}^{\infty}\mathbb{E}[Q_q \mid N = n]\,\mathbb{P}[N = n] = \sum_{n=0}^{\infty}\mathbb{E}[R + N_q B \mid N = n]\,\mathbb{P}[N = n]$
$= \sum_{n=0}^{\infty}\mathbb{E}[R \mid N = n]\,\mathbb{P}[N = n] + \sum_{n=0}^{\infty}\mathbb{E}[N_q B \mid N = n]\,\mathbb{P}[N = n]$
$= \sum_{n=1}^{\infty}\mathbb{E}[R]\,\mathbb{P}[N = n] + \sum_{n=1}^{\infty}\mathbb{E}[N_q B \mid N_q + 1 = n]\,\mathbb{P}[N_q = n-1]$
$= \mathbb{E}[R]\sum_{n=1}^{\infty}\mathbb{P}[N_q = n-1] + \mathbb{E}[B]\sum_{n=0}^{\infty}(n-1)\,\mathbb{P}[N_q = n-1]$
$= \mathbb{E}[R]\,(1-p_0) + \mathbb{E}[B]\mathbb{E}[N_q]$
$= \mathbb{E}[R]\,\rho + \mathbb{E}[B]\mathbb{E}[N_q]$
weil R und N_q, B und N_q unabhängig voneinander sind (***PASTA-Eigenschaft***).
Satz von Little: $\mathbb{E}[N_q] = \lambda\mathbb{E}[Q_q]$
Dann erhält man in diesem Fall

$$\mathbb{E}[Q_q]= \mathbb{E}[\mathrm{R}]\rho+\lambda\mathbb{E}[B]\,\mathbb{E}[Q_q]$$
$$\mathbb{E}[Q_q]=\frac{\mathbb{E}[R]\,\rho}{1-\lambda\mathbb{E}[B]}.$$

Für die Exponentialverteilung ist bekannt, daß $\mathbb{E}[\mathrm{B}]=\frac{1}{\mu}$.

Nehmen wir an, daß $\mathbb{E}[\mathrm{R}]=\frac{\mathbb{E}[B^2]}{2\,\mathbb{E}[B]}$, wobei für die Exponentialverteilung die zweite Moment $\mathbb{E}[B^2]=\frac{2}{\mu^2}$ bekannt ist. Dann erhält man die Formel

$$\mathbb{E}[Q_q]=\frac{\mathbb{E}[R]\,\rho}{1-\lambda\mathbb{E}[B]}=\frac{\lambda}{\mu(\mu-\lambda)} \text{ und } \mathbb{E}[Q]=\mathbb{E}[Q_q]+\mathbb{E}[\mathrm{B}]=\frac{\lambda}{\mu(\mu-\lambda)}+\frac{1}{\mu}=\frac{1}{\mu-\lambda}$$

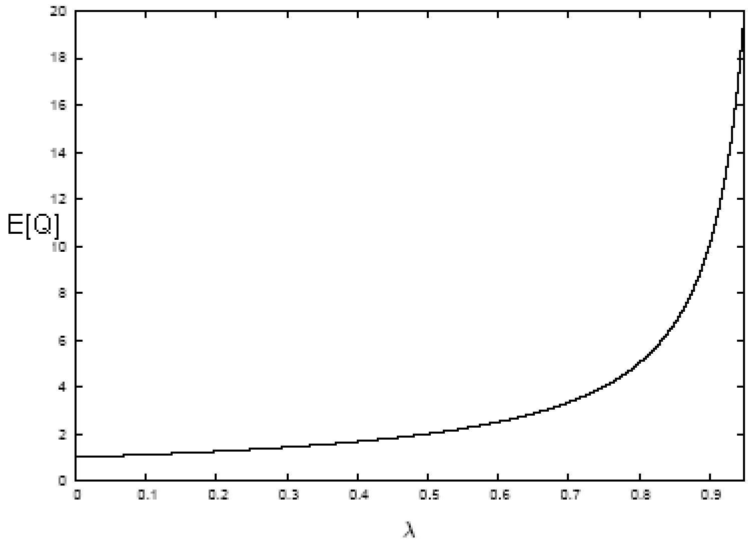
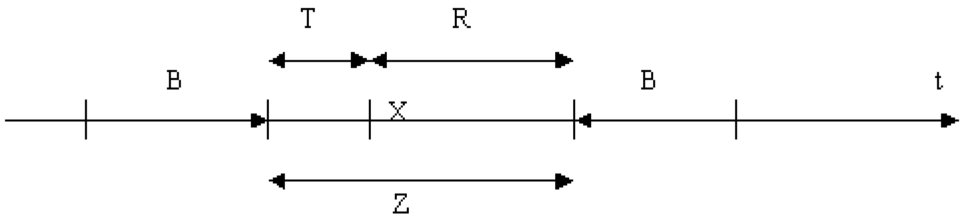


Figure Mean Delay vs. Utilization

Es bleibt nur den Wert $\mathbb{E}[R]$ auszurechnen.

Paradoxon der Restlebenszeit.



Z - beobachtete Zeitspanne mit Dichte $\mathbb{s}(t)$
Frage: Welche Restlebensdauer R hat das dabei gesehene Intervall Z.
Es sei R - Restlebenszeit mit dichte $\mathbb{r}(t)$, B - Intervall mit Dichte $\mathbb{f}(t)$.
Einfache (falsche!!!) Lösung: Ist $\mathbb{E}[B]$ die mittlere Länge der Intervalle ($\mathbb{E}[B]=\mathbb{E}[Z]$), so sollte der zufällige Beobachter einen Rest von $\mathbb{E}[R]=\frac{\mathbb{E}[B]}{2}$.
Gegenbeispiel: Sind die Intervall-Längen z.B. exponentiell verteilt, so ist die Wahrscheinlichkeit

$$\mathbb{P}[R<t]=\mathbb{P}[B<t]$$

wegen der Gedächtnislosigkeit der Exponentialverteilung, d.h. $\mathbb{E}[R]=\mathbb{E}[B]$.
Grund des Paradoxon: Wahrscheinlichkeit, daß beobachtete Periode Länge z hat, ist ungleich der Wahrscheinlichkeit, das beliebige Periode Länge z hat, $\mathbb{P}[B\leq z]\neq\mathbb{P}[Z\leq z]$: Lange Intervalle werden häufiger beobachtet als kurze Intervalle!
Z sei die Zufallsvariable für die beobachtete Periode. Dann ergibt sich:
(a) Mittlere Dauer von k aufeinanderfolgenden Bedienungsintervalle: $k\cdot\mathbb{E}[B]$
(b) Die Wahrscheinlichkeit dafür, daß genau i von k Intervallen die Länge z haben, kann mit Hilfe der Binomialverteilung gewinnen werden:

$$\binom{k}{i}\mathbb{P}[B\in[z,\,z+d z)]^i\,\mathbb{P}[B\notin[z,\,z+d z)]^{k-i}\,.$$

Die Mittlere Anzahl von Intervallen mit Längen zwischen z und z+dz unter k Intervallen:

$$k\cdot\mathbb{P}[B\in[z,\,z+d z)].$$

Mittlere Dauer von Bedienungsintervallen mit Länge zwischen z und z+dz unter k Intervallen:

$$\approx k\cdot\mathbb{P}[B\in[z,\,z+d z)]\cdot z=k\cdot\mathbb{f}(z)\cdot z\,d z$$

Daraus erhalten wir für die Dichte $\mathbb{s}(z)$ von Z:

$\mathbb{P}$[Z∈[z,z+d z)]=(Mittlere Dauer von Bedienungsintervallen mit Länge zwischen z und $z + dz$ unter k Intervallen)/
(Mittlere Dauer alle k Intervallen)

$$=\frac{k\cdot \mathbb{f}(z)\cdot z dz}{k\cdot \mathbb{E}[B]}$$

$$\mathbb{s}(z)dz=\frac{k\cdot \mathbb{f}(z)\cdot z dz}{k\cdot \mathbb{E}[B]}$$

$$\mathbb{s}(z)=\frac{\mathbb{f}(z)\cdot z}{\mathbb{E}[B]}\ ,\ z>0$$

Hat das beobachtete Intervall Z die Länge z, so ist der Beobachtungsteitpunkt X auf dem Intervall gleichverteilt, mit Dichte $\frac{1}{z}$.

Als Dichte von R ergibt dann sich mit dem Satz von der totalen Wahrscheinlichkeit (bedingte nach beobachtetem Intervall Z), beachte daß bei Z<t keine Restlebenszeit der Größe t existieren kann):

$$\mathbb{r}(t)=\int_t^\infty \frac{1}{z}\,\mathbb{s}(z)\,dz= \int_t^\infty \frac{1}{z}\,\frac{\mathbb{f}(z)\cdot z}{\mathbb{E}[B]}\,dz = \frac{1}{\mathbb{E}[B]}\int_t^\infty \mathbb{f}(z)\,dz=\frac{1-\mathbb{F}(t)}{\mathbb{E}[B]}$$

Als mittlere Restlebensdauer ergibt sich:

$$\mathbb{E}[R]=\int_0^\infty t\,\mathbb{r}(t)\,dt = \int_0^\infty t\,\frac{1-\mathbb{F}(t)}{\mathbb{E}[B]}\,dt= \frac{1}{\mathbb{E}[B]}\int_0^\infty t\,(1-\mathbb{F}(t))\,dt=$$

$$= \frac{1}{\mathbb{E}[B]}\int_0^\infty t\int_t^\infty \mathbb{f}(z)\,dz\,dt = \frac{1}{\mathbb{E}[B]}\int_0^\infty \int_t^\infty t\,\mathbb{f}(z)\,dz\,dt$$

$$= \frac{1}{\mathbb{E}[B]}\int_0^\infty \int_0^z t\,\mathbb{f}(z)\,dt\,dz= \frac{1}{\mathbb{E}[B]}\int_0^\infty \frac{z^2}{2}\,\mathbb{f}(z)\,dz =\frac{\mathbb{E}[B^2]}{2\cdot \mathbb{E}[B]}$$

Ist das System im Gleichgewicht, so kann ***die Verteilung der (Zufallsvariable) Wartezeit*** Q_q eines (virtuellen) Kunden, der zum Zeitpunkt 0 eintrifft, durch folgende Überlegungen gewonnen werden.

Satz: Die Verteilung der Wartezeit eines Kunden besitzt die folgende Verteilungsfunktion

$$\mathbb{F}_{Q_q}(t) = \mathbb{P}\left[Q_q \le t\right] = 1 - \rho\, e^{-(\mu-\lambda)\,t}$$

Beweis: Wegen der Gedächtnislosigkeit der Exponentialverteilung gilt für $n \in \mathbb{N}_0$ bei einer FCFS Strategie

$$\mathbb{P}\left[Q_q \le t \mid N = n\right] = 1 - \sum_{k=0}^{n-1} \frac{e^{-\mu t}(\mu t)^k}{k!}\ .$$

Daraus folgt mit dem Satz von der totalen Wahrscheinlichkeit

$$\mathbb{P}\left[Q_q \le t\right]=\sum_{n=0}^{\infty} \mathbb{P}\left[Q_q \le t \mid N = n\right] p_n$$

$$=1-\sum_{n=1}^{\infty}\left(\sum_{k=0}^{n-1} \frac{e^{-\mu t}(\mu t)^k}{k!}\right) p_n$$

$$=1-(1-\rho)\sum_{n=1}^{\infty}\left(\sum_{k=0}^{n-1} \frac{e^{-\mu t}(\mu t)^k}{k!}\right) \rho^n$$

$$=1-(1-\rho)e^{-\mu t}\sum_{k=0}^{\infty} \frac{(\mu t)^k}{k!}\left(\sum_{n=k+1}^{\infty} \rho^n\right)$$

$$=1-(1-\rho)e^{-\mu t}\sum_{k=0}^{\infty} \frac{(\mu t)^k}{k!}\ \frac{\rho^{k+1}}{1-\rho}$$

$$=1-\rho e^{-\mu t}\sum_{k=0}^{\infty} \frac{(\rho\mu t)^k}{k!}$$

$$=1-\rho e^{-(\mu-\lambda)\,t}$$

Ein Kunde muss also mit Wahrscheinlichkeit 1-ρ gar nicht warten und, wenn er warten muss, eine exponentialverteilte Zeit mit Parameter μ-λ. Man beachte, dass die Annahme von exponentialverteilten Bedienungszeiten wesentlich in die Berechnung eingegangen ist. Da die Zwischenankunftszeiten exponentialverteilt sind, gilt diese Verteilung für die Wartezeit eines jeden Kunden, der in ein Bedienungssystem im Gleichgewicht eintritt. Daraus könnte natürlich wieder die mittlere Wartezeit gewonnen werden, die wir ja bereits mit der Formel von Little berechnet haben.

$$\rho = \frac{\lambda}{\mu};$$
$$F_{Q_q} = 1 - \rho\, \mathrm{Exp}[-(\mu - \lambda)\, t];$$
$$f_{Q_q} = D\left[F_{Q_q},\, t\right];$$
$$W_q = \mathrm{Integrate}\left[t\, f_{Q_q},\, \{t, 0, \infty\}\right]$$
$$W = \mathrm{Simplify}\left[\frac{\rho}{\mu - \lambda} + \frac{1}{\mu}\right]$$

$$\mathrm{If}\left[\mathrm{Re}[\lambda] < \mathrm{Re}[\mu],\, \frac{\lambda}{\mu\,(-\lambda + \mu)},\right.$$
$$\left.\mathrm{Integrate}\left[e^{t\,(\lambda-\mu)}\, t\, \lambda - \frac{e^{t\,(\lambda-\mu)}\, t\, \lambda^2}{\mu},\, \{t, 0, \infty\},\, \mathrm{Assumptions} \to \mathrm{Re}[\lambda] \geq \mathrm{Re}[\mu]\right]\right]$$

$$\frac{1}{-\lambda + \mu}$$

$W_q = \frac{\lambda}{\mu(\mu-\lambda)}$

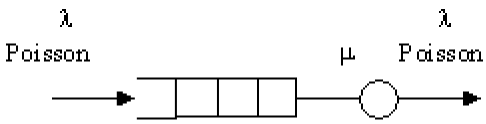
$W = W_q + \mathbb{E}[B]=\frac{\lambda}{\mu(\mu-\lambda)} + \frac{1}{\mu} = \frac{1}{\mu-\lambda}$

Man kann aber auch die Wahrscheinlichkeit angeben, dass ein Kunde länger als einen vorgegebenen Wert warten muss - es könnte ja sein, dass dann zusätzliche Kosten anfallen.

$\mathbb{P}\left[Q_q > t \mid Q_q > 0\right] = \frac{\mathbb{P}[Q_q>t, Q_q>0]}{\mathbb{P}[Q_q>0]} = \frac{\mathbb{P}[Q_q>t]}{\mathbb{P}[Q_q>0]} = e^{-(\mu-\lambda)\, t}$

Satz (Satz von Burke für M/M/1-Systeme): Für jedes M/M/1-System im Gleichgewicht stimmen Ankunfts- und Abgangsprozeß stochastisch übereit, d.h. der Abgangsprozeß ist ebenfalls Poissonprozess mit Rate λ und mittlere Durchsatz durch das System

$$S = (1 - p_0) \cdot \mu = \rho \cdot \mu = \lambda$$



Zusammenfassung (Leistungsmerkmale bei M/M/1): Betrachte ein M/M/1-System mit Ankunftsrate λ, Bedienrate μ. Es sei $\rho=\frac{\lambda}{\mu}$. Dann gilt

- mittlere Anzahl von Kunden im System:

$$L=\mathbb{E}[N]=\frac{\rho}{1-\rho}$$

- mittlere Anzahl von Kunden in der Warteschlange:

$$L_q = \mathbb{E}\left[N_q\right] = \frac{\rho^2}{1-\rho}$$

- mittlere bediengte Anzahl von Kunden in der Warteschlange:

$$\tilde{L}_q = \mathbb{E}\left[N_q \mid N > 1\right] = \frac{1}{1-\rho}$$

- mittlere Systemzeit

$$W=\mathbb{E}[Q]=\frac{1}{\mu(1-\rho)}$$

- mittlere Wartezeit

$$W_q = \mathbb{E}\left[Q_q\right] = W - \mathbb{E}[B] = \frac{\rho}{\mu(1-\rho)}$$

- mittlere Durchsatz durch das System

$$S = (1 - p_0)\cdot\mu=\rho\cdot\mu=\lambda$$

- die Verteilungsfunktion der Wartezeit eines Kunden

$$\mathbb{F}_{Q_q}(t) = \mathbb{P}\left[Q_q \leq t\right] = 1 - \rho\, e^{-(\mu-\lambda)\,t}$$

- die bedingte Verteilungsfunktion der Wartezeit eines Kunden

$$\mathbb{F}_{Q_q|Q_q>0}(t) = \mathbb{P}\left[Q_q \leq t \mid Q_q > 0\right] = 1 - e^{-(\mu-\lambda)\,t}$$

Beispiel 12. Betrachten Sie eine einfache Warteschlange M/M/1 mit exponential verteilten Zwischenankunftszeiten (mit fester Rate λ=1) und einem Bediener mit ebenfalls exponential verteilten Bedienzeiten (mit Bedienrate μ), wobei diese Bedienrate frei gewählt werden kann. Zu jedem solchen μ ergibt sich eine mittlere Wartezeit $W_q(\mu)$ in der Warteschlange für die Jobs im System. Ziel ist die Untersuchung des folgenden Optimierungsproblems: Durch den Betrieb des Servers entstehen Kosten (die von der Leistungsfähigkeit des Servers abhängen und es wird angenommen, daß $c(\mu) = 2\,\mu$), es entstehen aber ebenso Kosten durch die Belegung von Plätzen in der Warteschlange vor dem server, d.h. $c\left(W_q\right)$=1.5 $W_q(\mu)$. Die Gesamtkosten des Systems ergeben sich als C=$c(\mu)$+$c\left(W_q\right)$. Bestimmen Sie nun das μ mit den geringsten Gesamtkosten.

Lösung:

$W_q(\mu) = \frac{1}{\mu(\mu-1)}, \frac{1}{\mu} < 1$

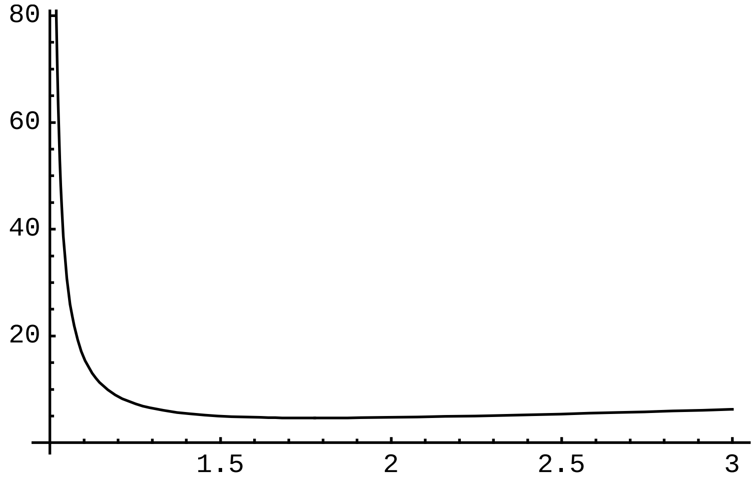
c=2 μ+1.5 $\frac{1}{\mu(\mu-1)}$

$\frac{\mathrm{dC}}{\mathrm{d}\mu} = 0 \Rightarrow \mu\approx 1.7786$

```
c = 2 mu + 1.5 / (mu (mu - 1))
D[f, mu]
Plot[f, {mu, 1, 3}]
Solve[D[f, mu] == 0, mu] // N
```

$$\frac{1.5}{(-1+\mathrm{mu})\,\mathrm{mu}}+2\,\mathrm{mu}$$

$$2-\frac{1.5}{(-1+\mathrm{mu})\,\mathrm{mu}^2}-\frac{1.5}{(-1+\mathrm{mu})^2\,\mathrm{mu}}$$




```
- Graphics -
{{mu → -0.159869 - 0.868171 i},
 {mu → -0.159869 + 0.868171 i}, {mu → 0.541105}, {mu → 1.77863}}
```

Beispiel 13. Eine besondere Klasse von Markov-Kette bilden sogenannten Geburts- und Todes-Markov-Ketten. Dies sind Markov-Ketten, bei denen ein Zustandsübergang nur zwischen benachbarten Zuständen möglich ist. So ist z.B. die folgende Markov-Kette eine solche Gburts-Todes-Markov-Kette mit der Besonderheit, daß alle Geburtsraten gleich λ sind, alle Todesraten gleich μ sind und es unendlich viele Zustände gibt (M/M/1-Warteschlangensystem).

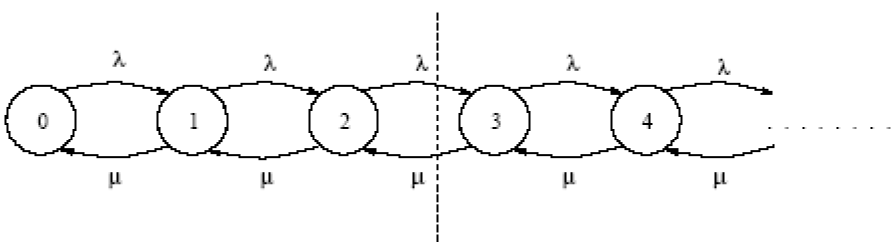
Aufgaben:

(a) Bestimmen Sie die Zustandsübergangssratenmatrix Λ.

(b) Bestimmen Sie die Zustandsübergangswahrscheinlichkeitsmatrix P

(c) Eine solche Markov-Kette besitzt einen Gleichgewichtswahrscheinlichkeitsvektor π genau dann,wenn $\lambda<\mu$ gilt. Bestimmen Sie $p=(p_n)_{n\in\mathbb{N}_0}$ für λ=2 und μ=4.

Lösung:



$$\Lambda=\begin{pmatrix} -\lambda & \lambda & 0 & \ldots \\ \mu & -(\lambda+\mu) & \lambda & \ldots \\ 0 & \mu & -(\lambda+\mu) & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

$$P=\begin{pmatrix} -1 & 1 & 0 & \ldots \\ \frac{\mu}{\lambda+\mu} & -1 & \frac{\lambda}{\lambda+\mu} & \ldots \\ 0 & \frac{\mu}{\lambda+\mu} & -1 & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$ (für die eingebettete Markov-Kette)

$p_n = \rho^n(1-\rho)$, wobei $\rho=\frac{\lambda}{\mu} = 0.5 \Rightarrow$

$p_n = 0.5^{n+1}$

Beispiel 14. Berechnen Sie für das System M/M/1 mit λ=1000 Pakete/s und μ=4000 Pakete/s die Wahrscheinlichkeit, dass kein Paket auf seinen Versand wartet bzw. gerade versandet wird. Vergleichen Sie das mit der Wahrscheinlichkeit, dass 10 Pakete im System sind.

Lösung:

Zur Vereinfachung kann man λ und μ normieren:
λ=1 und μ=4.
Die Wahrscheinlichkeit, dass sich kein Paket im System befindet ist
$p_0 = 1-\rho$
$p_0 = 0.75$
Und für 10 Pakete:
$p_{10} = p_0\, \rho^{10} = 0.75 \cdot 0.25^{10} \approx 7.153 \cdot 10^{-7} \Rightarrow \frac{p_{10}}{p_0} \approx 9.54 \cdot 10^{-7}$

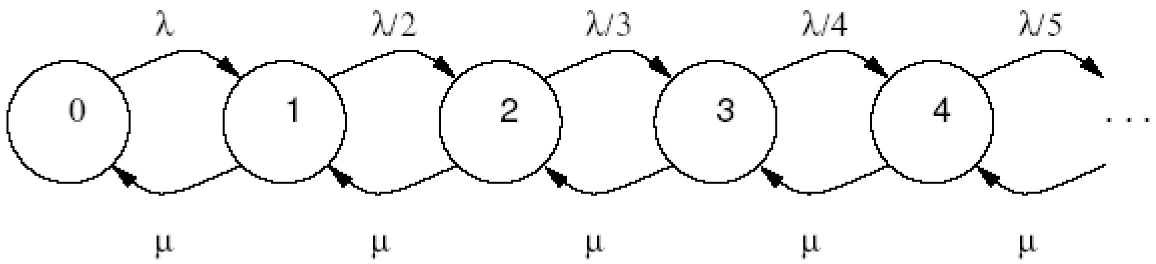
Beispiel 15. Nehmen Sie an, dass die Ankunftsrate bei steigender Anzahl k von zu versendenden Paketen folgendermaßen sinkt: $\lambda'(k) = \frac{\lambda}{k+1}$.Die Bedienrate bleibt aber konstant, d.h. μ(k)=μ für alle k.

(a) Zeichnen und beschriften Sie das Zustandsübergangsdiagramm.

(b) Drücken Sie die Wahrscheinlichkeit p_k, daß sich k Pakete im System aufhalten, als Funktion von λ, μ und p_0 aus.

(c) Finden Sie eine geschlossene Lösung für p_0. Hinweis: Nutzen Sie die Reihenentwicklung der Exponentialfunktion e^x.

Lösung:
(a)



Zustandsübergangsdiagramm

(b) Beim stationären System kann man folgende Gleichungen aufstellen:

$\lambda\ p_0 = \mu\ p_1 \ \Rightarrow\ p_1 = \frac{\lambda}{\mu}\ p_0$

$\frac{\lambda}{2}\ p_1 = \mu\ p_2 \ \Rightarrow\ p_2 = \frac{\lambda^2}{1 \cdot 2 \cdot \mu^2}\ p_0$

$\frac{\lambda}{3}\ p_2 = \mu\ p_3 \ \Rightarrow\ p_3 = \frac{\lambda^3}{1 \cdot 2 \cdot 3 \cdot \mu^3}\ p_0$

....

Daraus läßt sich

$$p_k = \frac{\lambda^k}{k!\ \mu^k}\ p_0 = \frac{\rho^k}{k!}\ p_0 \quad (\rho = \frac{\lambda}{\mu})$$

induktiv für k>0 ableiten.

(c) Wegen

$$\sum_{k=0}^{\infty} p_k = p_0 + \sum_{k=1}^{\infty} \frac{\rho^k}{k!}\ p_0 = 1$$

d.h. mit Vorzihen des konstanten Faktors p_0 und dem Ausdruck von p_0 als Indexverschiebung

$$1 = p_0 \sum_{k=0}^{\infty} \frac{\rho^k}{k!}$$

Mit der Reihenentwicklung der Exponentialfunktion

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

folgt

$$1 = p_0\ e^{\rho}$$

und somit

$$p_0 = e^{-\rho}$$