④ Markov Ketten

Informell:

> Ein Markov-Prozess ist ein spezieller stochastischer Prozess, der kein (oder nur ein begrenztes) Gedächtnis besitzt. Ziel ist die Vorhersage zukünftiger Wahrscheinlichkeiten, oder die Simulation komplexer W-Verteilungen.

- Wahrscheinlichkeitsraum (Ω, Σ, P)

  Ω = Grundmenge („Elementarereignisse")

  Σ = σ-Algebra von Teilmengen von Ω

  - $\Omega \in \Sigma$
  - $A \in \Sigma \Rightarrow \bar{A} = \Omega \setminus A \in \Sigma$
  - $A_1, A_2, A_3, \ldots \in \Sigma \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \Sigma$

  * falls Ω diskret (endlich ~~oder abzählbar~~) normalerweise

  $\Sigma = \mathcal{P}(\Omega)$ die Potenzmenge von Ω

  * falls $\Omega \subseteq \mathbb{R}^n \rightarrow \Sigma = \mathcal{B}(\mathbb{R}^n)$ Borel-Algebra

  P = Wahrscheinlichkeitsmaß (W-Maß):

  - $P(\emptyset) = 0$
  - $P(\Omega) = 1$
  - $A_1, A_2, A_3 \in \Sigma$ paarweise disjunkt: $P(A_1 \cup A_2 \cup \ldots) = P(A_1) + P(A_2) + \ldots$
    („σ-Additivität")

- Falls Ω endlich ~~oder abzählbar~~ (in Numerik immer)

  $\Sigma = \mathcal{P}(\Omega)$

  $p_i = P(\{\omega_i\}) \geq 0$ charakterisiert W-Maß eindeutig, $\sum_i p_i = 1$
  („Zähldichte")

- <u>Zufallsvariable</u>
  - Messbare Funktion von einem W-Raum (Ω, Σ, P) in einen Messraum (Ω', Σ')
  - <u>Reelle</u> Zufallsvariable $\Omega' \subseteq \mathbb{R}$, $\Sigma' =$ Borel-Algebra
  - Ereignisraum abzählbar: $X : \Omega \to \mathbb{R}$, so daß

    $$\forall x \in \mathbb{R} : \{\omega \in \Omega \mid X(\omega) \leq x\} \in \Sigma$$ (Verteilung messbar)
  - Ereignisraum endlich: $\Sigma = \mathcal{P}(\Omega)$, Messbarkeit automatisch ✓

    $$X : \Omega \to \mathbb{R}$$
  - Messbarkeit eine Abbdg $X : \Omega \to \Omega'$ falls $X^{-1}(\Sigma') = \{X^{-1}(B) \mid B \in \Sigma'\} \subset \Sigma$

- <u>Zufallsvektor</u>: Tupel / Sequenz von Zufallsvariablen $X = (X_1, \ldots, X_n)$

- <u>Stochastischer Prozess</u>: Folge / Fkt von Zufallsvektoren über demselben W-Raum

- <u>Zustandsraum</u> = Bildraum von $X \cong \mathbb{R}^n$

- <u>Ensemble</u> $\mathcal{E} = (S, X)$ Zufallsvektor über S mit W-Verteilung

  $$W(\omega) = X(\omega)$$

  Ensemble $\hat{=}$ Zufallsvektor über S

Markov-Eigenschaft (Zeit diskret)

$$\boxed{\begin{aligned} &P\left(X^{(\alpha+1)} = x \mid X^{(\alpha)} = y^{(\alpha)}, X^{(\alpha-1)} = y^{(\alpha-1)}, \dots, X^{(1)} = y^{(1)}\right) \\ &= P\left(X^{(\alpha+1)} = x \mid X^{(\alpha)} = y^{(\alpha)}\right) \end{aligned}}$$

- Zukunft hängt nur von Gegenwart, nicht Vergangenheit ab („Kein Gedächtnis")
- Ketten höher Ordnung hängen von begrenzter Anzahl der Vorgängerereignisse ab („begrenztes Gedächtnis") → hier nicht betrachtet.

Anwendungen

- Wirtschaftswissenschaften: Warteschlangentheorie
- Biologie: Untersuchung und Charakterisierung von Gen-Sequenzabschnitten
- Versicherungsmathematik: Berechnung von (biometrischen) Risiken
- Finanzmathematik: Modellierung von Aktienkursen und Zinsentwicklung
- Musik: Algorithmische Kompositionen
- Spieltheorie: Mehr-rundige, zufallsbasierte Entscheidungen
- Automatisierte Spam-Filter
- Statistische Physik
- Simulation komplexer W-Verteilungen
- Qualitätsmanagement: Zuverlässigkeit und Ausfallswahrscheinl. von Systemen
- Internet: Page-Rank (Google)

  Zustandsraum = alle bekannten Web-Seiten

  Page rank P_i = score (Relevanz) der Seite #i

  $$P(i \leftarrow k) = \begin{cases} (1-\alpha)/N + \frac{\alpha}{\nu_k} ; & k \text{ hat einen Link auf } i \\ (1-\alpha)/N ; & k \text{ hat keinen Link auf } i \end{cases}$$ (ν_i = #links auf i)

Notation: diskrete Markov Ketten

- Dyn. System mit n Freiheitsgraden $\boxed{\omega^a \in \mathbb{R}, \quad a = 1, \dots, n}$
- Zustand = Konfiguration $\boxed{\omega = (\omega^1, \dots, \omega^n) \in \mathbb{R}^n}$

  = eine ~~leb~~ Instanz aller dyn. Variablen

  = charakterisiert Zustand des Systems vollständig

  (klass. m-Teilchen System : Phasenraum $\mathbb{R}^{6m}$)

Bsp: Spin-Modell $\omega^a \in \{\pm 1\}$ $a = 1, \dots, n$ numeriert Spin

$$\omega = (\underset{\omega^1}{-1}, \underset{\omega^2}{+1}, \underset{\omega^3}{+1}, \dots, \underset{\omega^{n-1}}{-1}, \underset{\omega^n}{+1})$$

- Zustandsraum $S = \{\omega = (\omega^1, \dots, \omega^n)\} \overset{\text{endlich}}{=} \{\omega_i \; ; \; i = 1, \dots, N\}$

NB : $\omega^a \leftarrow$ numeriert Freiheitsgrad = eine Komponente von ω

$\omega_i \leftarrow$ numeriert Zustand in S = ein kompletter Zustand des Systems

Bsp Spin-Modell : $N = |S| = 2^n$

Bei endlichen Zustandsräumen wird oft nur der Index i statt der Konfiguration ω_i notiert (bei fester Numerierung der N Zustände)

- Zufallsvariable oder -vektor nimmt zufällige Werte aus ~~S~~ S an mit bekannter Verteilung

$$\boxed{W(\omega_i) = W_i = P(X = \omega_i)}$$

(S, W) = Ensemble

Markov Kette der Ordnung 1 (= einzig hier betrachtet) vollständig charakterisiert durch (falls zeitdiskret)

• Übergangswahrscheinlichkeit

$$\boxed{\begin{aligned} P(X^{(\alpha+1)} = \omega_i \mid X^{(\alpha)} = \omega_k) &= P^{(\alpha)}(\omega_i \leftarrow \omega_k) \\ &= P^{(\alpha)}(i \leftarrow k) = P_{ik}^{(\alpha)} \end{aligned}}$$

1.3 Hier nur zeitdiskrete Markov-Prozesse (= Markov-Ketten)

(Für kontinuierliche Zeit: Markov-Eigenschaft für alle n-tupel von Zeiten $0 < t_1 < t_2 < \ldots < t_n$; hier nicht näher betrachtet)

• Zeitliche Homogenität $\boxed{P_{ik}^{(\alpha+1)} = P_{ik}^{(\alpha)}}$

$P_{ik}^{(\alpha)}$ hängt nicht von Zeitübergangsindex α ab.

Zeitlich homogene Markov-Kette durch eine Ü-Matrix P_{ik} bestimmt

$$P_{ik} = P(i \leftarrow k) = P(\omega_i \leftarrow \omega_k)$$

• Chapman-Kolmogorov $\boxed{\overline{P}_{ik}^{(\alpha)} = \sum_{\omega_j \in S} \overline{P}_{ij}^{(\beta)} \, \overline{P}_{jk}^{(\alpha-\beta)}}$ für beliebiges $1 \leq \beta < \alpha$

Falls zeitlich homogen ist dies einfach Matrix-Multiplikation $\overline{P}^{(\alpha)} = \overline{P}^{(\beta)} \overline{P}^{(\alpha-\beta)}$, also mit $\overline{P}^{(1)} = P$ Ü-Wahrscheinlichkeit

$$\boxed{\overline{P}^{(\alpha)} = P^{\alpha}}$$

$\overline{P}_{ik}^{(\alpha)}$ = Wahrscheinl. für Übergang $\omega_k \to \omega_i$ in α Schritten.

Für zeitlich homogene Markov-Ketten ist die Zeitevolution also

$$W(X^{(\alpha)} = \omega_i) = \sum_j P(i \leftarrow j)\, W(X^{(\alpha-1)} = \omega_j)$$

$$= \sum_j \sum_\ell P(i \leftarrow j)\, P(j \leftarrow \ell)\, W(X^{(\alpha-2)} = \omega_\ell)$$

$$= \dots$$

$$= \sum_{i_1} \cdots \sum_{i_n} P(i \leftarrow i_1)\, P(i_1 \leftarrow i_2) \cdots P(i_{n-1} \leftarrow i_n)\, W(X^{(0)} = \omega_n)$$

$$= \sum_k \overline{P}^{(n)}(\omega_i \leftarrow \omega_k)\, W(X^{(0)} = \omega_k)$$

$$= \sum_k \overline{P}_{ik}^{(n)}\, W(X^0 = \omega_k)$$

$$= \sum_k [P^n]_{ik}\, W(X^0 = \omega_k)$$

Für endliche Zustandsräume $|S| = N < \infty$ ist die Ü-Wahrscheinlichkeit eine stochastische $N \times N$ Matrix. Aussagen über P^n folgen daher aus dem klassischen Matrix Theorem [falls symmetrisch, s.u.] von Froebenius-Perron ✓. Bei nicht-endlichen Zustandsräumen ist Spektraltheorie + Funktionalanalysis notwendig!

Froebenius-Perron: Für jede symmetrische stochastische Matrix P

1. Spektralradius $r = \sup\{|\lambda_i|,\ \lambda_i \text{ Eigenwert}\}$: $\min_i \sum_j P_{ij} \le r \le \max_i \sum_j P_{ij}$
2. Für alle Eigenwerte gilt $|\lambda_i| \le 1$ (folgt aus 1)
3. $\lambda_0 = 1$ ist einfach
4. Eigenvektor w zu λ_0 hat alle Komponenten mit gleichem Vorzeichen
   O.B.d.A. $w_i \ge 0$ mit $\sum w_i = 1$ normierbar.
5. $\lim_{n\to\infty} P^n = P_\infty$ ist Projektor auf W : $P_\infty = |w\rangle\langle w|$

Mögliche Eigenschaften (Attribute) von Markov-Ketten

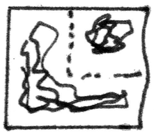
1. <u>Irreduzibilität</u> : Jeder Zustand von jedem anderen Zustand in endlicher Zahl von Schritten (mit pos. Wahrscheinl.) <u>erreichbar</u>:

Zu jedem $\omega_i, \omega_k \in S$ gibt es $m \in \mathbb{N}$, so daß

$$\bar{P}_{ik}^{(m)} = [P^m]_{ik} = P(X^{(m)} = \omega_i \mid X^{(0)} = \omega_k) > 0$$

= <u>schwach ergodisch</u>

Zustandsraum zerfällt also nicht in unverbundene Teilbereiche



zwei reduzible Unterräume

2. <u>Periodizität</u> : Ist Markov-Kette periodisch mit Periode d, so kann sie höchstens alle d Zeitschritte zum Startpunkt zurückkehren.

Genauer : Zustand $\omega_i \in S$

Mögliche Rückkehrzeiten $T_i = \{ t \mid P(X^{(t)} = i \mid X^{(0)} = i) > 0 \}$

Periode $d_i = \gcd(T_i)$ größter gemeinsamer Teiler

d_i ist also größte gemeinsame Teil-Rückkehrzeit

$T_i = \{2, 4, 6, 8, \ldots\} \rightarrow d_i = 2$

$T_i = \{4, 8, 10, 12, \ldots\} \rightarrow d_i = 2$ obwohl $t = 2 \notin T_i$

Falls MK <u>irreduzibel</u> haben alle Zustände dieselbe Periode $d = d_i$

Falls Periode d_i, kann das System höchstens zu den Zeiten $n \cdot d_i$ zurückkehren, muß aber nicht ($T_i = \{4, 8, 10, 12\}$ ~~da~~ hat $d_i = 2$ aber $2, 6, ..$ sind keine mögl. Rückkehrzeiten). Andere als die periodischen Zeitpunkte sind aber unmöglich, im Bsp. alle ungeraden t.

$d_i > 1$ ~~Sr~~ Zustand periodisch
$d_i = 1$ Zustand aperiodisch (schließt keine Zeiten aus)

- Irred. Markov Ketten sind also entweder periodisch oder aperiodisch.

3. Rekurrenz / Transienz : Langzeitverhalten!

Kehrt das System in endlicher Schrittzahl zu Zustand ω_i fast sicher zurück, so heißt ω_i rekurrent. Ein rekurrenter Zustand wird mit Wahrsch. 1 (unendlich) oft besucht.

Ist ein Zustand nicht rekurrent, so heißt er transient.

Sind alle Zustände rekurrent, so heißt die MK rekurrent (transient entsprechend).

Formal Wiederkehrzeit („hitting time") $\inf\{\alpha \geq 1 : X^{(\alpha)} = \omega_i \mid X^{(0)} = \omega_i\} = \tau_i$ Zufallsvariable

Ersteinkehrwahrscheinlichkeit $q_i^n = P(\tau_i = n)$

ω_i ist rekurrent, falls $P(\tau_i < \infty) = \sum_{n=1}^{\infty} q_i^n = 1$

ω_i ist transient, falls $P(\tau_i < \infty) < 1$, d.h. $P(\tau_i = \infty) > 0$

Falls auch Erwartungswert der Rückkehrzeit endlich

$$\langle \tau_i \rangle = \sum_{n=1}^{\infty} n \cdot P(\tau_i = n) < \infty \quad : \quad \text{positiv rekurrent}$$

4. <u>Ergodizität</u> : System kommt jedem Zustand beliebig nahe oder (bei diskretem S) besucht jeden Zustand ohne „bias"

- schwach ergodisch = irreduzibel
- quasi ergodisch = irreduzibel + (positiv) rekurrent
- ergodisch = irreduzibel + pos. rekurrent + aperiodisch

5. <u>Reversibilität</u> : Invarianz under Zeitumkehr

$$\boxed{P_{ik} = P(X^{(\alpha+1)} = \omega_i \mid X^{(\alpha)} = \omega_k) \stackrel{!}{=} P(X^{(\alpha)} = \omega_i \mid X^{(\alpha+1)} = \omega_k) = \tilde{P}_{ik}}$$

Sichert die Existenz einen stationären Zustandes, also einer stat. Endverteilung ~~$\lim_{n\to\infty} P^n =$~~, s. <u>stationäre Ensemble</u> unten.

Reversibilität impliziert <u>detailed balance</u> (s.u.)

Stationäre Ensemble

- Voraussetzung: MK homogen mit Ü-Matrix $P = P_{ik} = P(\omega_i \leftarrow \omega_k)$
  (endlich oder abzählbar dimensional)

  Zustandsraum abzählbar

- <u>Stationäre Verteilung</u> Zufallsvariable X auf S mit Verteilung
  $W_i = P(X = \omega_i)$, so daß

$$\boxed{W_i = \sum_k P_{ik} W_k \qquad \text{oder } W = PW}$$

(Eigenvektor zu P zum Eigenwert 1)

Theorem 1: W existiert genau dann, wenn die Markov-Kette quasi-ergodisch ist (irreduzibel + pos. rekurrent)
In diesem Fall ist W eindeutig und $W_i = 1/\tau_i$
(τ_i = mittlere Rückkehrzeit)

Theorem 2: Ist die Markov-Kette ergodisch (irred, pos. rekurrent, aperiodisch)

1. $W_i = \lim_{n \to \infty} (P^n)_{ij} \quad \forall j$ (unabh. von Startzustand)
2. $W_i > 0$ für alle i
3. $\sum_i W_i = 1$

Theorem 3: Voraussetzung wie Theorem 2, zusätzlich

$$\tau_i^{(2)} = \sum_{n=1}^{\infty} n^2 \, P(\tau_i = n) < \infty \quad \text{für alle } \omega_i \in S$$

Dann Scharmittel = Zeitmittel, also

$$\bar{A} = \cancel{\frac{1}{N}\sum_{i=\alpha}^{\alpha+N} A} = \cancel{\frac{1}{m}\left[A(i) + P(j,i)\,A(i) + P^2(\right.}$$

$$= \frac{1}{M}\sum_{\alpha=1}^{M} P_{ki}^{(\alpha)} A(k) \qquad m \to \infty$$

$$\langle A \rangle = \sum_{i=1}^{N} W_i \, A(i)$$

↑ stat. Verteilung

für Observable $A: S \to \mathbb{R}$ mit $A(i) = A(\omega_i)$

<u>Bem.</u> Für endliche Zustandsräume Voraussetzungen automatisch erfüllt, Beweis durch Froebenius-Perron.

<u>Froebenius-Perron</u> : Für jede symmetrische stochastische Matrix P gilt:

1. Spektralradius $r \equiv \sup\{|\lambda_i| ; \lambda_i \text{ Eigenwert}\}$

   $$\min_i \sum_j P_{ij} \leq r \leq \max_i \sum_j P_{ij}$$, bei stochastischer Matrix also $r = 1$

2. Für alle Eigenwerte λ_i gilt daher $|\lambda_i| \leq 1$

3. $\lambda_0 = 1$ ist immer einfacher Eigenwert

4. Eigenvektor W zu λ_0 hat positive Komponenten $W_i \geq 0$ mit $\sum_i W_i = 1$

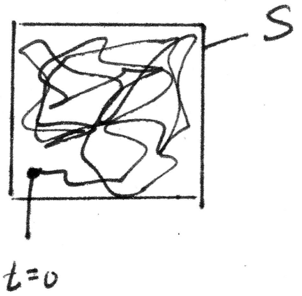
5. $\lim_{N \to \infty} P^N \equiv P_\infty$ ist stochastische Matrix mit Rang 1 (also Projektor) auf W

   $$P_\infty = |W\rangle\langle W|$$

Anwendung in QM $\quad P_{ij} = Z^{-1} e^{-T\hat{H}} = \sum_n e^{-TE_n} |n\rangle\langle n| \Big/ \sum_n e^{-TE_n}$

Markov-Ketten und statistische Physik

- System mit vielen dynamischen Freiheitsgraden bewegt sich unter der tatsächlichen Zeitentwicklung chaotisch durch den Zustandsraum

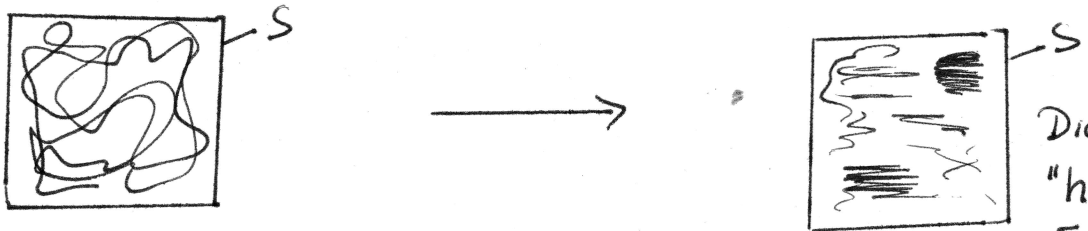


- <u>Observable</u> $A: S \to \mathbb{R}$

  Zeitskala T für Messung $\gg$ mikroskopische Zeitskala der Dynamik, also der Zeitdauer von Zustandsänderungen.
  System durchläuft während Messung große Zahl von Zuständen

$$\boxed{\text{Zeitmittel} \qquad \bar{A} = \frac{1}{T}\int_t^{t+T} d\tau\, A(\omega(\tau))}$$

- Ersetze unter geeigneten Voraussetzungen (Zeitentwicklung ergodisch) die Zeitmittelung durch eine Wahrscheinlichkeitsmittelung



Dichte = "heat map"
Ensemble W

<u>Zeitmittel</u>

$$\bar{A} = \frac{1}{T}\int_t^{t+T} d\tau\, A(\omega(\tau))$$

<u>Scharmittel</u>

$$\langle A \rangle = \int_S d\omega\, W(\omega)\, A(\omega)$$

- Die "Dichte" $W(\omega)$ des Ensembles entspricht der Häufigkeit, in der der Zustand ω während der Beobachtungszeit unter der tatsächlichen Zeitentwicklung besucht wird. („heat map")

- W-Mittel ist theoretisch einfacher und oft auch numerisch einfacher zu implementieren. Unter recht allgemeinen Voraussetzungen sind die Klass. Ensemble unter bestimmten Randbedingungen (und großer Zahl von Freiheitsgraden) bekannt

① Isoliertes System → mikrokanonisches Ensemble

$$\boxed{W(\omega) \sim \delta(H(\omega) - E)}$$

② Kontakt mit Wärmebad → Kanonisches Ensemble

$$\boxed{W(\omega) \sim \exp\left[-\beta H(\omega)\right]}$$

Boltzmann-Verteilung

$\beta = \frac{1}{k_B T}$ inverse Temperatur

③ Kontakt mit Wärme- und Teilchenbad → großkanonisches Ensemble

$$\boxed{W(\omega) \sim \exp\left[-\beta\left\{H(\omega) + \mu N(\omega)\right\}\right]}$$

$N(\omega)$ = Teilchenzahl (# Freiheitsgrade) im Zustand ω

μ = chem. Potential

- Statistische Gleichgewichtsphysik ist zeitunabhängig (eher Thermostatik statt Thermodynamik)

- In Numerik # Freiheitsgrade normalerweise Konstant
  → <u>Kanonisches Ensemble</u> (Boltzmann-Verteilung)

- Erzeugung der Boltzmann-Verteilung kompliziert (viele gekoppelte Freiheitsgrade, $\omega = (\omega^1, \omega^2, .., \omega^n)$ hochdimensional)
  ⟹ <u>Erzeugung durch Markov-Kette iterativ</u>

$$X^0(S, W^0) \xrightarrow{P} X^1(S, W^1) \xrightarrow{P} X^2(S, W^2) \xrightarrow{P} \dots \xrightarrow{P} X^\infty(S, W^\infty)$$

Steuere Übergangswahrscheinlichkeit P so, daß die stationäre Verteilung W^∞ Boltzmann ist.

- Die Schritte der MK können reale Zeitschritte sein (wenn P die reale Dynamik beschreibt), i.a. können aber bel. (und einfachere) Wahrscheinlichkeiten P gewählt werden, sofern sie nur zur richtigen stationären Verteilung führen.

  Man nennt die MK-Schritte daher <u>MC-Zeit</u> (Monte-Carlo-Zeit) im Gegensatz zur realen Zeit.

- Mögliche Ü-Wahrscheinlichkeiten P die ~~zur~~ richtigen stationären Boltzmann-Verteilung führen, heißen <u>thermodyn. Algorithmen</u>

Detailed balance

- Wie kann man verifizieren, daß Algorithmus korrekt ist, also die Ü-Wahrscheinlichkeit P in der Markov-Kette zum richtigen stationären Ensemble führt?

Hinreichendes (Haupt-) Kriterium:

<u>Detailed balance</u>:

Sei Markov-Kette quasi-ergodisch und es gebe Zahlen $r_i > 0$ (Funktionen $S \to \mathbb{R}_+$ mit $r_i := r(\omega_i)$) mit

$$\boxed{P_{ki}\, r_i = P_{ik}\, r_k} \quad \forall\, i, k \quad (\text{Keine Summe!})$$

Dann existiert das Gleichgewichtsensemble $X^{(\infty)} = (S, W^{(\infty)})$ und

$$W_i^{(\infty)} = \frac{r_i}{\sum_j r_j}$$

Normalerweise ist die Zielverteilung Boltzmann, d.h.

$$r_i = e^{-\beta H_i} = e^{-\beta H(\omega_i)} > 0$$

und detailed balance bedeutet

$$\boxed{P_{ki}\, e^{-\beta H_i} = P_{ik}\, e^{-\beta H_k} \quad \forall\, i, k}$$

Beweis:

1) $W^{(\infty)}$ existiert (Theorem 1)

2) $W_i^{(\infty)} > 0$ und $\sum_i w_i^{(\infty)} = 1$ trivial

3)
$$(PW^{(\infty)})_i = \sum_k P_{ik} W_k^{(\infty)} = \sum_k P_{ik} \frac{r_k}{Z}$$
$$\overset{(*)}{=} \sum_k P_{ki} \frac{r_i}{Z} = \frac{r_i}{Z} \sum_k P_{ki} = \frac{r_i}{Z} = W_i^{(\infty)}$$

wobei $Z = \sum_e r_e > 0$ und d.b. wurde bei (*) verwendet ■

Bemerkung: Detailed balance entspricht lokaler Reversibilität, also Zeitumkehr invarianz:

vorwärts: $P_{ik} = P(X^{(\alpha+1)} = \omega_i \mid X^{(\alpha)} = \omega_k)$

rückwärts: $\tilde{P}_{ik} = P(X^{(\alpha)} = \omega_i \mid X^{(\alpha+1)} = \omega_k)$

Bayes: $P(A|B)\, P(B) = P(B|A)\, P(A)$

Hier:
$$P(X^{(\alpha)} = \omega_i \mid X^{(\alpha+1)} = \omega_k)\, P(X^{(\alpha+1)} = \omega_k)$$
$$= P(X^{(\alpha+1)} = \omega_k \mid X^{(\alpha)} = \omega_i)\, P(X^{(\alpha)} = \omega_i)$$

$$\Rightarrow \quad \underline{\tilde{P}_{ik}\, W_k^{(\alpha+1)} = P_{ki}\, W_i^{(\alpha)}} \quad \text{für alle } i, k$$

Prozess reversibel $P = \tilde{P}$ und $W^{(\alpha+1)} = W^{(\alpha)}$ $\Rightarrow$ d.b.

Umgekehrt d.b. im stationären Zustand $W^{(\alpha)} = W^{(\alpha+1)} = W^{(\infty)}$ und daher $\tilde{P} = P$ $\Rightarrow$ lokale Reversibilität ■

Satz (Creutz 1983)

Falls eine Markov-Kette detailed balance erfüllt (und quasi-ergodisch ist), so ist die Abweichung vom Gleichgewicht nach n Schritten

$$\sigma_n(i) \equiv \sum_k \left| (P^n)_{ki} - W_k^{(\infty)} \right|$$

(in der L_1-Norm über alle Zustände k) ~~streng~~ monoton fallend

$$\sigma_{n+1}(i) \leq \sigma_n(i) \qquad \text{für jeden Anfangszustand } i$$

Beweis:

$$\sigma_{n+1}(i) = \sum_k \left| (P^{n+1})_{ki} - W_k^{(\infty)} \right| = \sum_k \left| \sum_l P_{kl} P_{li}^n - W_k^{(\infty)} \right|$$

$$= \sum_k \left| \sum_l \{ P_{kl} P_{li}^n - P_{lk} W_k^{(\infty)} \} \right| \qquad \text{da } \sum_l P_{lk} = 1$$

$$\overset{d.b.}{=} \sum_k \left| \sum_l \{ P_{kl} P_{li}^n - P_{kl} W_l^{(\infty)} \} \right|$$

$$= \sum_k \left| \sum_l P_{kl} (P_{li}^n - W_l^{(\infty)}) \right|$$

$$\leq \sum_k \sum_l P_{kl} \left| P_{li}^n - W_l^{(\infty)} \right| \qquad \Delta\text{-Ungleichung}$$

$$= \sum_l \left| P_{li}^n - W_l^{(\infty)} \right| \underbrace{\sum_k P_{kl}}_{1}$$

$$= \sum_l \left| P_{li}^n - W_l^{(\infty)} \right|$$

$$= \sigma_n(i) \qquad \blacksquare$$

Beweist nicht $\sigma_n(i) = 0$ im Limes $n \to \infty$:

$$\left.\begin{array}{l} \text{Aus } \sigma_n > 0 \\ \sigma_n \text{ streng monoton fallend} \end{array}\right\} \quad \sigma \equiv \lim_{n\to\infty} \sigma_n \geq 0 \text{ existiert, aber } \sigma = 0 \text{ folgt nicht } \blacksquare$$