

# Algorithmische Mathematik II

Prof. Dr. Andreas Eberle

31. Juli 2009

# Inhaltsverzeichnis

<b>1</b>	<b>Diskrete Zufallsvariablen</b>	<b>5</b>
1.1	Ereignisse und ihre Wahrscheinlichkeit . . . . .	7
	Ereignisse als Mengen . . . . .	7
	Wahrscheinlichkeitsverteilungen . . . . .	9
	Diskrete Wahrscheinlichkeitsverteilungen . . . . .	11
	Spezielle Wahrscheinlichkeitsverteilungen . . . . .	14
	Gleichverteilungen / Laplace-Modelle . . . . .	14
	Empirische Verteilungen . . . . .	15
1.2	Zufallsvariablen und ihre Verteilung . . . . .	19
	Binomialverteilung . . . . .	21
	Poissonverteilung . . . . .	23
	Hypergeometrische Verteilung . . . . .	24
1.3	Simulation von Gleichverteilungen . . . . .	25
	Lineare Kongruenzgeneratoren (LCG) . . . . .	26
	Shift-Register-Generatoren . . . . .	31
	Kombination von Zufallszahlengeneratoren . . . . .	31
	Zufallszahlen aus $[0,1)$ . . . . .	31
	Zufallspermutationen . . . . .	32
1.4	Erwartungswert . . . . .	33
<b>2</b>	<b>Bedingte Wahrscheinlichkeiten und Unabhängigkeit</b>	<b>39</b>
2.1	Bedingte Wahrscheinlichkeiten . . . . .	39
	Berechnung von Wahrscheinlichkeiten durch Fallunterscheidung . . . . .	40
	Bayessche Regel . . . . .	42
2.2	Mehrstufige diskrete Modelle . . . . .	43
	Produktmodelle . . . . .	46

---

Markov-Ketten . . . . .	47
Berechnung von Wahrscheinlichkeiten . . . . .	50
2.3 Unabhängigkeit von Ereignissen . . . . .	51
Verteilungen für unabhängige Ereignisse . . . . .	53
2.4 Unabhängige Zufallsvariablen und Random Walk . . . . .	59
Der Random Walk auf $\mathbb{Z}$ . . . . .	60
Symmetrischer Random Walk . . . . .	63
2.5 Simulationsverfahren . . . . .	66
Das direkte Verfahren . . . . .	67
Acceptance-Rejection-Verfahren . . . . .	68
<b>3 Konvergenzsätze und Monte Carlo-Methoden</b>	<b>71</b>
3.1 Varianz und Kovarianz . . . . .	72
3.2 Gesetze der großen Zahlen . . . . .	76
3.3 Monte Carlo-Verfahren . . . . .	78
Varianzreduktion durch Importance Sampling . . . . .	81
3.4 Gleichgewichte von Markov-Ketten . . . . .	83
Konstruktion von Markov-Ketten mit vorgegebenen Gleichgewichtsverteilungen .	87
3.5 Konvergenz ins Gleichgewicht . . . . .	92
<b>4 Iterationsverfahren für lineare und nichtlineare Gleichungssysteme</b>	<b>97</b>
4.1 Der Banachsche Fixpunktsatz . . . . .	98
4.2 Klassische Iterationsverfahren für lineare Gleichungssysteme . . . . .	102
Klassische Iterationsverfahren . . . . .	106
Keine Präkonditionierung . . . . .	106
Jacobi-Verfahren / Gesamtschrittverfahren . . . . .	106
Gauß-Seidel-Verfahren/Einzelschrittverfahren . . . . .	107
Relaxationsverfahren . . . . .	112
Relaxiertes Jacobi-Verfahren (Jacobi over-relaxation – »JOR«) . . . . .	112
Sukzessives Relaxationsverfahren (Successive over-relaxation – »SOR«) .	113
4.3 Abstiegsverfahren . . . . .	114
Gradientenverfahren . . . . .	115
Verfahren der konjugierten Gradienten (CG) . . . . .	118
4.4 Das Newton-Verfahren für nichtlineare Gleichungssysteme . . . . .	121
Andere Iterationsverfahren im eindimensionalen Fall . . . . .	125

---

<b>5</b>	<b>Interpolation und Quadratur</b>	<b>128</b>
5.1	Stückweise lineare Interpolation . . . . .	128
5.2	Polynominterpolation . . . . .	132
	Berechnung des Interpolationspolynoms . . . . .	135
	Lagrange-Darstellung . . . . .	135
	Newton-Darstellung, dividierte Differenzen . . . . .	136
5.3	Numerische Integration . . . . .	138
	Newton-Cotes-Formeln . . . . .	139
	Gauß-Quadratur . . . . .	142
5.4	Kubische Splines . . . . .	148
	B-Splines («Bell splines») . . . . .	149
	Konstruktion einer Basis von $S_m(\Delta)$ . . . . .	150
	Interpolationsproblem für kubische Splines ( $m = 3$ ) . . . . .	151
	Berechnung in B-Spline-Basis . . . . .	153
	<b>Stichwortverzeichnis</b>	<b>154</b>

# Kapitel 1

## Diskrete Zufallsvariablen

Unser Ziel in diesem Kapitel ist die mathematische Modellierung von *Zufallsvorgängen*. Einfache Beispiele für Zufallsvorgänge sind das Werfen eines Würfels oder Münzwürfe. Anhand dieser Beispiele wollen wir zunächst einige grundlegende Begriffe der Wahrscheinlichkeitstheorie veranschaulichen.

NOTATIONEN:  $|A|$  bezeichnet die Anzahl der Elemente einer Menge  $A$ ,  $A^C$  bezeichnet das Komplement der Menge  $A$  innerhalb einer bestimmten Menge  $B$ , die  $A$  enthält.

**Beispiel** (Werfen eines Würfels).

- *Mögliche Fälle* sind 1, 2, 3, 4, 5, 6. Mit  $\Omega = \{1, 2, 3, 4, 5, 6\}$  wird die Menge aller möglichen Fälle bezeichnet. Ein *Elementarereignis* ist ein möglicher Fall, also ein Element  $\omega \in \Omega$ .
- *Ereignisse* sind die Objekte, denen man eine Wahrscheinlichkeit zuordnen kann, zum Beispiel:

»Augenzahl ist 3«	$\{3\}$
»Augenzahl ist gerade«	$\{2, 4, 6\}$
»Augenzahl ist <i>nicht</i> gerade«	$\{1, 3, 5\} = \{2, 4, 6\}^C$
»Augenzahl ist größer als 3«	$\{4, 5, 6\}$
»Augenzahl ist gerade <i>und</i> größer als 3«	$\{4, 6\} = \{2, 4, 6\} \cap \{4, 5, 6\}$
»Augenzahl gerade <i>oder</i> größer als 3«	$\{2, 4, 5, 6\} = \{2, 4, 6\} \cup \{4, 5, 6\}$

Jedes *Ereignis* kann durch eine *Teilmenge*  $A$  von  $\Omega$  dargestellt werden!

- *Wahrscheinlichkeiten* werden mit  $P$  (für »probability«) bezeichnet. Zum Beispiel sollte für einen »fairen« Würfel gelten:

$$P[\text{»3«}] = \frac{1}{6},$$

$$P[\text{»Augenzahl gerade«}] = \frac{\text{Anzahl günstige Fälle}}{\text{Anzahl mögliche Fälle}} = \frac{|\{2, 4, 6\}|}{|\{1, 2, 3, 4, 5, 6\}|} = \frac{3}{6} = \frac{1}{2},$$

$$P[\text{»Augenzahl gerade oder größer als 3«}] = \frac{4}{6} = \frac{2}{3}.$$

- *Zufallsvariablen* sind Abbildungen  $X : \Omega \rightarrow S$ , wobei  $S$  eine beliebige Menge ist, zum Beispiel:

$$X(\omega) = \omega, \quad \text{»Augenzahl des Wurfs«, oder}$$

$$X(\omega) = \begin{cases} 1 & \text{falls } \omega \in \{1, 2, 3, 4, 5\}, \\ -5 & \text{falls } \omega \in 6, \end{cases} \quad \text{»Gewinn bei einem fairen Spiel«.}$$

### Beispiel (Münzwürfe).

#### a) EIN MÜNZWURF:

Die Menge der möglichen Fälle ist  $\Omega = \{0, 1\}$ , wobei 0 für »Kopf« und 1 für »Zahl« steht. Die Wahrscheinlichkeiten sind

$$P[\{1\}] = p \quad \text{und} \quad P[\{0\}] = 1 - p \quad \text{mit } 0 \leq p \leq 1.$$

Für  $p = \frac{1}{2}$  ist der Münzwurf fair.

#### b) ENDLICH VIELE FAIRE MÜNZWÜRFE:

Die Menge der möglichen Fälle lautet

$$\Omega = \{\omega = (x_1, \dots, x_n) \mid x_i \in \{0, 1\}\} =: \{0, 1\}^n.$$

Alle Ausgänge sind genau dann gleich wahrscheinlich, wenn  $P[\{\omega\}] = 2^{-n}$  für alle  $\omega \in \Omega$  gilt. Dies wird im folgenden angenommen. Zufallsvariablen von Interesse sind beispielsweise:

- $X_i(\omega) := x_i$ , das Ergebnis des  $i$ -ten Wurfs. Das Ereignis » $i$ -ter Wurf ist Kopf« wird durch die Menge  $A_i = \{\omega \in \Omega \mid X_i(\omega) = 0\} =: \{X_i = 0\}$  beschrieben, und hat die Wahrscheinlichkeit  $P[A_i] = \frac{1}{2}$ .

- $S_n(\omega) := \sum_{i=1}^n X_i(\omega)$ , die Anzahl der Einsen in  $n$  Münzwürfen. Das Ereignis »genau  $k$ -mal Zahl« wird durch die Menge  $A = \{\omega \in \Omega \mid S_n(\omega) = k\} =: \{S_n = k\}$  beschrieben und hat die Wahrscheinlichkeit  $P[A] = \binom{n}{k} 2^{-n}$ .

c) UNENDLICH VIELE MÜNZWÜRFE:

Die Menge der möglichen Fälle ist nun

$$\Omega = \{\omega = (x_1, x_2, \dots) \mid x_i \in \{0, 1\}\} = \{0, 1\}^{\mathbb{N}}.$$

Diese Menge ist überabzählbar, da die Abbildung

$$\begin{aligned} \Omega &\rightarrow [0, 1] \\ (x_1, x_2, \dots) &\mapsto 0.x_1x_2\dots \end{aligned}$$

surjektiv ist, (wobei das Einheitsintervall binär dargestellt wird). Die Definition von Ereignissen und Wahrscheinlichkeiten ist daher in diesem Fall aufwändiger. Wahrscheinlichkeitsverteilungen auf überabzählbaren Mengen werden systematisch in der Vorlesung »Einführung in die Wahrscheinlichkeitstheorie« betrachtet.

In dieser Vorlesung ist die Menge der möglichen Fälle  $\Omega$  abzählbar. Solche Zufallsvorgänge werden *diskret* genannt.

## 1.1 Ereignisse und ihre Wahrscheinlichkeit

### Ereignisse als Mengen

Sei  $\Omega$  die Menge der möglichen Fälle und  $A \subseteq \Omega$  ein Ereignis. Als Notationen für die Menge  $A$  werden wir auch verwenden:

$$A = \{\omega \in \Omega \mid \omega \in A\} = \{\omega \in A\} = \{\text{»}A \text{ tritt ein«}\}.$$

Wir wollen nun *Kombinationen von Ereignissen* betrachten.

Seien  $A, B, A_i, i \in I$ , Ereignisse. Was bedeuten Ereignisse wie  $A^C, A \cup B, \bigcap_{i \in I} A_i$  anschaulich? Um dies herauszufinden, betrachten wir einen möglichen Fall  $\omega$  und untersuchen, wann dieser eintritt:

- $A \cup B$ :

$$\begin{aligned} \omega \in A \cup B &\Leftrightarrow \omega \in A \text{ oder } \omega \in B, \\ \text{»}A \cup B \text{ tritt ein«} &\Leftrightarrow \text{»}A \text{ tritt ein oder } B \text{ tritt ein«}. \end{aligned}$$

- $\bigcup_{i \in I} A_i$ :

$$\omega \in \bigcup_{i \in I} A_i \quad \Leftrightarrow \quad \text{es gibt ein } i \in I \text{ mit } \omega \in A_i.$$

$$\text{»}\bigcup_{i \in I} A_i \text{ tritt ein«} \quad \Leftrightarrow \quad \text{»mindestens eines der Ereignisse } A_i \text{ tritt ein«.}$$

- WEITERE BEISPIELE:

$$A \cap B \quad \Leftrightarrow \quad \text{»}A \text{ und } B \text{ treten ein«,}$$

$$\bigcap_{i \in I} A_i \quad \Leftrightarrow \quad \text{»jedes der } A_i \text{ tritt ein«,}$$

$$A^C = \Omega \quad \Leftrightarrow \quad \text{»}A \text{ tritt nicht ein«,}$$

$$A = \emptyset \quad \Leftrightarrow \quad \text{»unmögliches Ereignis« (tritt nie ein),}$$

$$A = \Omega \quad \Leftrightarrow \quad \text{»sicheres Ereignis« (tritt immer ein),}$$

$$A = \{\omega\} \quad \Leftrightarrow \quad \text{»Elementarereignis« (tritt nur im Fall } \omega \text{ ein).}$$

Sei  $\mathcal{A}$  die Kollektion aller im Modell zugelassenen bzw. in Betracht gezogenen Ereignisse.  $\mathcal{A}$  besteht aus Teilmengen von  $\Omega$ , d.h.

$$\mathcal{A} \subseteq \mathcal{P}(\Omega), \quad \text{wobei}$$

$$\mathcal{P}(\Omega) := \{A \mid A \subseteq \Omega\}$$

die Potenzmenge von  $\Omega$ , d.h. die Menge aller Teilmengen von  $\Omega$  bezeichnet. Die Kollektion  $\mathcal{A}$  sollte unter den obigen Mengenoperationen, also abzählbaren Vereinigungen, Durchschnitten und Komplementbildung abgeschlossen sein. Wir fordern daher:

**Axiom.**  $\mathcal{A} \subseteq \mathcal{P}(\Omega)$  ist eine  $\sigma$ -Algebra, d.h.

(i)  $\Omega \in \mathcal{A}$ ,

(ii) Für alle  $A \in \mathcal{A}$  gilt:  $A^C \in \mathcal{A}$ ,

(iii) Für  $A_1, A_2, \dots \in \mathcal{A}$  gilt:  $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$ .

**Bemerkung.** Für  $\sigma$ -Algebren gilt auch:

a) Nach (i) und (ii) ist  $\emptyset = \Omega^C \in \mathcal{A}$ .

b) Sind  $A, B \in \mathcal{A}$ , so gilt nach (iii) und a):  $A \cup B = A \cup B \cup \emptyset \cup \emptyset \cup \dots \in \mathcal{A}$ .

c) Sind  $A_1, A_2, \dots \in \mathcal{A}$ , so ist nach (ii) und (iii):  $\bigcap_{i=1}^{\infty} A_i = (\bigcup_{i=1}^{\infty} A_i^C)^C \in \mathcal{A}$ .

**Beispiel.** Die Potenzmenge  $\mathcal{A} = \mathcal{P}(\Omega)$  ist eine  $\sigma$ -Algebra.

Üblicherweise verwendet man  $\mathcal{A} = \mathcal{P}(\Omega)$  bei *diskreten* Modellen, d.h. für abzählbare  $\Omega$ . Bei nichtdiskreten Modellen kann man *nicht* jede Wahrscheinlichkeitsverteilung  $P$  auf einer  $\sigma$ -Algebra  $\mathcal{A} \subset \mathcal{P}(\Omega)$  zu einer Wahrscheinlichkeitsverteilung auf  $\mathcal{P}(\Omega)$  erweitern (siehe »Einführung in die Wahrscheinlichkeitstheorie«).



## Wahrscheinlichkeitsverteilungen

Sei  $\Omega$  eine nichtleere Menge und  $\mathcal{A} \subseteq \mathcal{P}(\Omega)$  eine  $\sigma$ -Algebra. Wir wollen nun Ereignissen  $A \in \mathcal{A}$  eine Wahrscheinlichkeit  $P[A]$  zuordnen. Für Ereignisse  $A, B \in \mathcal{A}$  gilt:

$$A \cup B \text{ tritt ein} \Leftrightarrow A \text{ oder } B \text{ tritt ein.}$$

Angenommen,  $A$  und  $B$  treten nicht gleichzeitig ein, d.h.

$$A \cap B = \emptyset, \quad (A \text{ und } B \text{ sind »disjunkt«.})$$

Dann sollte »endliche Additivität« gelten:

$$P[A \cup B] = P[A] + P[B].$$

**Axiom.** Eine Abbildung

$$\begin{aligned} P: \mathcal{A} &\rightarrow [0, \infty] \\ A &\mapsto P[A] \end{aligned}$$

ist eine **Wahrscheinlichkeitsverteilung** auf  $(\Omega, \mathcal{A})$ , wenn gilt:

(i)  $P$  ist » $\sigma$ -additiv«, d.h. für Ereignisse  $A_1, A_2, \dots \in \mathcal{A}$  mit  $A_i \cap A_j = \emptyset$  für  $i \neq j$  gilt:

$$P\left[\bigcup_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} P[A_i].$$

(ii)  $P$  ist »normiert«, d.h.

$$P[\Omega] = 1.$$

Ein **Wahrscheinlichkeitsraum**  $(\Omega, \mathcal{A}, P)$  besteht aus einer Menge  $\Omega$ , einer  $\sigma$ -Algebra  $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ , und einer Wahrscheinlichkeitsverteilung  $P$  auf  $(\Omega, \mathcal{A})$ .

**Satz 1.1** (Elementare Rechenregeln). Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum.

i) Es gilt  $P[\emptyset] = 0$ ,

ii) Für  $A, B \in \mathcal{A}$  mit  $A \cap B = \emptyset$  gilt endliche Additivität:

$$P[A \cup B] = P[A] + P[B].$$

iii) Für  $A, B \in \mathcal{A}$  mit  $A \subseteq B$  gilt:

$$P[B] = P[A] + P[B \setminus A].$$

Insbesondere gilt:

$$\begin{aligned} P[A] &\leq P[B], && \text{»Monotonie«,} \\ P[A^C] &= 1 - P[A], && \text{»Gegenereignis«,} \\ P[A] &\leq 1. \end{aligned}$$

iv) Für  $A, B \in \mathcal{A}$  gilt:

$$P[A \cup B] = P[A] + P[B] - P[A \cap B] \leq P[A] + P[B].$$

*Beweis.* i) Wegen der  $\sigma$ -Additivität von  $P$  gilt

$$1 = P[\Omega] = P[\Omega \cup \emptyset \cup \emptyset \cup \dots] = \underbrace{P[\Omega]}_{=1} + \underbrace{P[\emptyset]}_{\geq 0} + \underbrace{P[\emptyset]}_{\geq 0} + \dots,$$

und damit

$$P[\emptyset] = 0.$$

ii) Für disjunkte Ereignisse  $A, B$  folgt aus der  $\sigma$ -Additivität und mit i):

$$\begin{aligned} P[A \cup B] &= P[A \cup B \cup \emptyset \cup \emptyset \cup \dots] \\ &= P[A] + P[B] + P[\emptyset] + \dots \\ &= P[A] + P[B]. \end{aligned}$$

iii) Falls  $A \subseteq B$ , ist  $B = A \cup (B \setminus A)$ . Da diese Vereinigung disjunkt ist, folgt mit ii):

$$P[B] = P[A] + P[B \setminus A] \geq P[A].$$

Insbesondere ist  $1 = P[\Omega] = P[A] + P[A^C]$  und somit  $P[A] \leq 1$ .

iv) Nach iii) gilt:

$$\begin{aligned} P[A \cup B] &= P[A] + P[(A \cup B) \setminus A] \\ &= P[A] + P[B \setminus (A \cap B)] \\ &= P[A] + P[B] - P[A \cap B]. \end{aligned}$$

□

Aussage iv) des Satzes lässt sich für endlich viele Ereignisse verallgemeinern. Nach iv) gilt für die Vereinigung von drei Ereignissen:

$$\begin{aligned} P[A \cup B \cup C] &= P[A \cup B] + P[C] - P[(A \cup B) \cap C] \\ &= P[A] + P[B] + P[C] - P[A \cap B] - P[A \cap C] - P[B \cap C] + P[A \cap B \cap C]. \end{aligned}$$

Mit vollständiger Induktion folgt:

**Korollar** (Einschluss-/Ausschlussprinzip). Für  $n \in \mathbb{N}$  mit Ereignissen  $A_1, \dots, A_n \in \mathcal{A}$  gilt:

$$P[\underbrace{A_1 \cup A_2 \cup \dots \cup A_n}_{\text{»eines der } A_i \text{ tritt ein«}}] = \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} P[\underbrace{A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}}_{\text{»}A_{i_1}, A_{i_2}, \dots \text{ und } A_{i_k} \text{ treten ein«}}].$$

Das Einschluss-/Ausschlussprinzip werden wir auf eine elegantere Weise am Ende dieses Kapitels beweisen (siehe Satz 1.9).

## Diskrete Wahrscheinlichkeitsverteilungen

Als Beispiel für eine diskrete Wahrscheinlichkeitsverteilung haben wir den Münzwurf betrachtet:

$$\begin{aligned} \Omega &= \{0, 1\}, & \mathcal{A} &= \{\{\emptyset\}, \{0\}, \{1\}, \{0, 1\}\}, \\ P[\{1\}] &= p, & P[\emptyset] &= 0, \\ P[\{0\}] &= 1 - p, & P[\Omega] &= 1. \end{aligned}$$

ALLGEMEIN: Ist die Menge der möglichen Fälle  $\Omega$  endlich oder abzählbar unendlich, dann setzen wir als zugehörige  $\sigma$ -Algebra  $\mathcal{A} = \mathcal{P}[\Omega]$ .

**Satz 1.2.** i) Sei  $0 \leq p(\omega) \leq 1$ ,  $\sum_{\omega \in \Omega} p(\omega) = 1$  eine Gewichtung der möglichen Fälle. Dann ist durch

$$P[A] := \sum_{\omega \in A} p(\omega), \quad (A \subseteq \Omega),$$

eine Wahrscheinlichkeitsverteilung auf  $(\Omega, \mathcal{A})$  definiert.

ii) Umgekehrt ist jede Wahrscheinlichkeitsverteilung  $P$  auf  $(\Omega, \mathcal{A})$  von dieser Form mit

$$p(\omega) = P[\{\omega\}] \quad (\omega \in \Omega).$$

$p: \Omega \rightarrow [0, 1]$  heißt **Massenfunktion** (»probability mass function«) der diskreten Wahrscheinlichkeitsverteilung  $P$ .

Für den Beweis des Satzes brauchen wir einige Vorbereitungen.

**Bemerkung** (Vorbemerkung zu Summen mit positiven Summanden). Sei  $A$  eine abzählbare Menge,  $p(\omega) \geq 0$  für alle  $\omega \in A$ . Dann definieren wir

$$\sum_{\omega \in A} p(\omega) := \sum_{i=1}^{\infty} p(\omega_i),$$

wobei  $\omega_1, \omega_2, \dots$  eine beliebige Abzählung von  $A$  ist.

**Lemma 1.3.** *i)  $\sum_{\omega \in A} p(\omega) \in [0, \infty]$  und ist wohldefiniert (d.h. unabhängig von der Abzählung). Es gilt:*

$$\sum_{\omega \in A} p(\omega) = \sup_{\substack{F \subseteq A \\ |F| < \infty}} \sum_{\omega \in F} p(\omega). \quad (1.1.1)$$

*Insbesondere gilt Monotonie:*

$$\sum_{\omega \in A} p(\omega) \leq \sum_{\omega \in B} p(\omega), \quad (A \subseteq B). \quad (1.1.2)$$

*ii) Ist  $A = \bigcup_{i=1}^{\infty} A_i$  eine disjunkte Zerlegung, dann gilt:*

$$\sum_{\omega \in A} p(\omega) = \sum_{i=1}^{\infty} \sum_{\omega \in A_i} p(\omega).$$

*Beweis.* i) Sei  $\omega_1, \omega_2, \dots$  eine beliebige Abzählung von  $A$ . Aus  $p(\omega_i) \geq 0$  für alle  $i \in \mathbb{N}$  folgt, dass die Partialsummen  $\sum_{i=1}^n p(\omega_i)$  monoton wachsend sind. Daraus folgt:

$$\sum_{i=1}^{\infty} p(\omega_i) = \sup_{n \in \mathbb{N}} \sum_{i=1}^n p(\omega_i).$$

Falls die Menge der Partialsummen von oben beschränkt ist, existiert dieses Supremum in  $[0, \infty)$ . Andernfalls divergiert die Folge der Partialsummen bestimmt gegen  $+\infty$ . Zu zeigen bleibt:

$$\sup_{n \in \mathbb{N}} \sum_{i=1}^n p(\omega_i) = \sup_{\substack{F \subseteq A \\ |F| < \infty}} \sum_{\omega \in F} p(\omega) \quad \text{ist unabhängig von der Abzählung von } A.$$

»≤« Für alle  $n \in \mathbb{N}$  gilt:

$$\sum_{i=1}^n p(\omega_i) \leq \sup_{\substack{F \subseteq A \\ |F| < \infty}} \sum_{\omega \in F} p(\omega),$$

da das Supremum auch über  $F = \{\omega_1, \dots, \omega_n\}$  gebildet wird. Damit folgt »≤«.

» $\geq$ « Da  $F \subseteq A$  endlich ist, gibt es ein  $n \in \mathbb{N}$ , so dass  $F \subseteq \{\omega_1, \dots, \omega_n\}$ . Also gilt:

$$\sum_{\omega \in F} p(\omega) \leq \sum_{i=1}^n p(\omega_i) \leq \sum_{i=1}^{\infty} p(\omega_i).$$

Damit folgt » $\geq$ «.

- ii) • Falls  $A$  endlich ist, gilt  $A_i \neq \emptyset$  nur für endlich viele  $i \in \mathbb{N}$  und alle  $A_i$  sind endlich. Die Behauptung folgt dann aus dem Kommutativ- und dem Assoziativgesetz.
- Sei andernfalls  $A$  abzählbar unendlich.

» $\leq$ «: Da  $F \subseteq A$  endlich, ist  $F = \bigcup_i^{\infty} F \cap A_i$ . Da diese Vereinigung disjunkt ist, folgt mit  $\sigma$ -Additivität und Gleichung (1.1.2):

$$P[F] = \sum_{i \in \mathbb{N}} P[F \cap A_i] \leq \sum_{i \in \mathbb{N}} P[A_i].$$

Mit (i)) gilt auch:

$$P[A] = \sup_{\substack{F \subseteq A \\ |F| < \infty}} P[F] \leq \sum_{i \in \mathbb{N}} P[A_i].$$

Damit folgt » $\leq$ «.

» $\geq$ «: Seien  $F_i \subseteq A_i$  endlich. Da die  $F_i$  disjunkt sind, folgt mit  $\sigma$ -Additivität und Gleichung (1.1.2) für alle  $n \in \mathbb{N}$ :

$$\sum_{i=1}^n P[F_i] = P\left[\bigcup_{i=1}^n F_i\right] \leq P\left[\bigcup_{i=1}^{\infty} A_i\right] = P[A].$$

Mit (1.1.1) folgt

$$\sum_{i=1}^n P[A_i] \leq P[A],$$

und für  $n \rightarrow \infty$  schließlich

$$\sum_{i=1}^{\infty} P[A_i] \leq P[A].$$

Damit folgt » $\geq$ «.

□

*Beweis von Satz 1.2.* i) Es ist  $P[\Omega] = \sum_{\omega \in \Omega} p(\omega) = 1$  nach Voraussetzung.

Seien  $A_i$ , ( $i \in \mathbb{N}$ ) disjunkt und  $A := \bigcup_{i=1}^{\infty} A_i$ . Die  $\sigma$ -Additivität von  $P$  folgt aus Lemma 1.3.ii):

$$P\left[\bigcup_{i=1}^{\infty} A_i\right] = P[A] = \sum_{\omega \in A} p(\omega) = \sum_{i=1}^{\infty} \sum_{\omega \in A_i} p(\omega) = \sum_{i=1}^{\infty} P[A_i]$$

ii) Aus der  $\sigma$ -Additivität von  $P$  folgt:

$$P[A] = P\left[\underbrace{\bigcup_{\omega \in A} \{\omega\}}_{\text{disjunkt}}\right] = \sum_{\omega \in A} P[\{\omega\}].$$

□

## Spezielle Wahrscheinlichkeitsverteilungen

### Gleichverteilungen / Laplace-Modelle

Sei  $\Omega$  endlich und nichtleer,  $\mathcal{A} = \mathcal{P}(\Omega)$  und  $p(\omega) = \frac{1}{|\Omega|}$  für alle  $\omega \in \Omega$ . Dann ist

$$P[A] = \frac{|A|}{|\Omega|} = \frac{\text{Anzahl »günstiger« Fälle}}{\text{Anzahl aller Fälle}}, \quad (A \subseteq \Omega),$$

die Wahrscheinlichkeitsverteilung zu  $p$  und wird **Gleichverteilung** genannt.

**Beispiele.** a)  $n$  FAIRE MÜNZWÜRFE:

Sei  $\Omega = \{0, 1\}^n$  und  $P$  die Gleichverteilung. Dann ist

$$p(\omega) = \frac{1}{2^n}.$$

b) ZUFÄLLIGE PERMUTATIONEN:

Sei  $\Omega = \mathcal{S}_n = \{\omega : \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\} \text{ bijektive Abbildungen}\}$  und  $P$  die Gleichverteilung. Dann ist

$$P[A] = \frac{|A|}{n!}.$$

Beispiele für zufällige Permutationen sind das Mischen eines Kartenspiels, Vertauschen von Hüten oder Umzug in die LWK, wobei  $n$  Schlüssel zufällig vertauscht werden. Es gilt:

$$P[\text{»der } k\text{-te Schlüssel passt auf Schloß } i\ll] = P[\{\omega \in \mathcal{S}_n \mid \omega(i) = k\}] = \frac{(n-1)!}{n!} = \frac{1}{n}.$$

Wie groß ist die Wahrscheinlichkeit, daß einer der Schlüssel sofort passt?

Das Ereignis »Schlüssel  $i$  passt« ist  $A_i = \{\omega \mid \omega(i) = i\} = \{\text{»}i \text{ ist Fixpunkt«}\}$ . Die Wahrscheinlichkeit für das Ereignis »ein Schlüssel passt« ist nach dem Einschluss-/Ausschlussprinzip (Satz 1.9):

$$\begin{aligned} P[\text{»es gibt mindestens einen Fixpunkt«}] &= P[A_1 \cup A_2 \cup \dots \cup A_n] \\ &= \sum_{k=1}^n (-1)^{k+1} \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} P[A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}] \\ &= \sum_{k=1}^n (-1)^{k+1} \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \frac{(n-k)!}{n!}, \end{aligned}$$

wobei die innere Summe über alle  $k$ -elementigen Teilmengen läuft. Es folgt:

$$\begin{aligned} &= \sum_{k=1}^n (-1)^{k+1} \binom{n}{k} \frac{(n-k)!}{n!} \\ &= - \sum_{k=1}^n \frac{(-1)^k}{k!} \end{aligned}$$

Für das Gegenereignis erhalten wir:

$$\begin{aligned} P[\text{»kein Schlüssel passt«}] &= P[\text{»kein Fixpunkt«}] - P[\text{»mindestens ein Fixpunkt«}] \\ &= 1 + \sum_{k=1}^n \frac{(-1)^k}{k!} \\ &= \sum_{k=0}^n \frac{(-1)^k}{k!}. \end{aligned}$$

Die letzte Summe konvergiert für  $n \rightarrow \infty$  gegen  $e^{-1}$ . Der Grenzwert existiert also und ist weder 0 noch 1! Die Wahrscheinlichkeit hängt für große  $n$  nur wenig von  $n$  ab.

### Empirische Verteilungen

Seien  $x_1, x_2, \dots, x_n \in \Omega$  Beobachtungsdaten oder Merkmalsausprägungen, zum Beispiel das Alter aller Einwohner von Bonn. Sei

$$\begin{aligned} N[A] &:= |\{i \in \{1, \dots, n\} \mid x_i \in A\}|, && \text{die Anzahl bzw. Häufigkeit der Werte in } A, \text{ und} \\ P[A] &:= \frac{N[A]}{n}, && \text{die relative Häufigkeit der Werte in } A. \end{aligned}$$

Dann ist  $P$  eine Wahrscheinlichkeitsverteilung auf  $(\Omega, \mathcal{P}(\Omega))$  mit Massenfunktion

$$p(\omega) = \frac{N[\{\omega\}]}{n}, \quad \text{der relativen Häufigkeit der Merkmalausprägungen.}$$

**Beispiele.** a) ABZÄHLUNG ALLER MÖGLICHEN FÄLLE:

Sei  $x_1, \dots, x_n$  eine Abzählung der Elemente in  $\Omega$ . Dann stimmt die empirische Verteilung mit der Gleichverteilung überein.

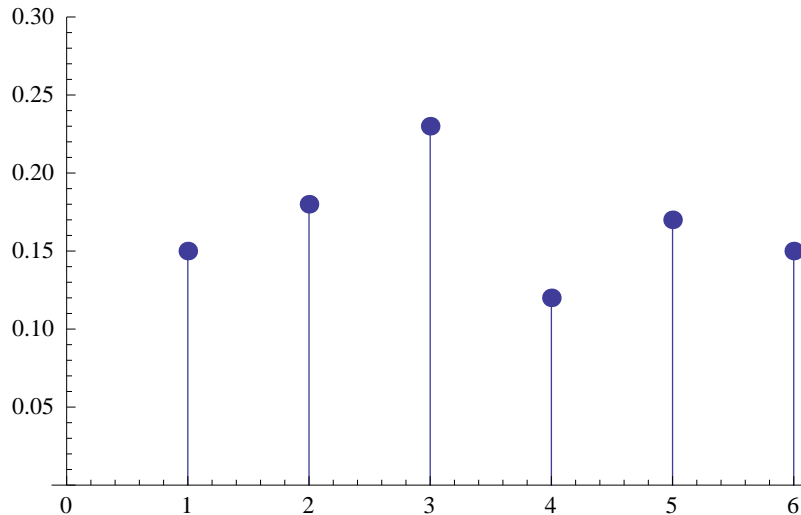
b) EMPIRISCHE VERTEILUNG VON  $n$  ZUFALLSZAHLN AUS  $\{1, 2, 3, 4, 5, 6\}$ :

Seien  $x_1, \dots, x_n$  Pseudozufallszahlen auf  $\{1, 2, 3, 4, 5, 6\}$ .

- $n = 100$ :

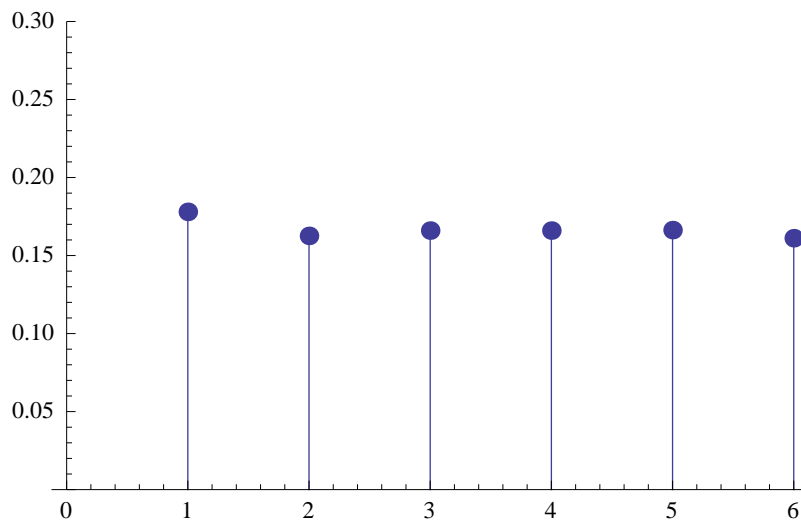
ListPlot[BinCounts[x[[1;;100]], {1, 7, 1}]/100, Filling  $\rightarrow$  Axis,

PlotRange  $\rightarrow \{0, 0.3\}$ , PlotStyle  $\rightarrow$  PointSize[Large]]



- $n = 10000$ :

ListPlot[BinCounts[x[[1;;10000]], {1, 7, 1}]/10000, Filling  $\rightarrow$  Axis,  
PlotRange  $\rightarrow \{0, 0.3\}$ , PlotStyle  $\rightarrow$  PointSize[Large]]

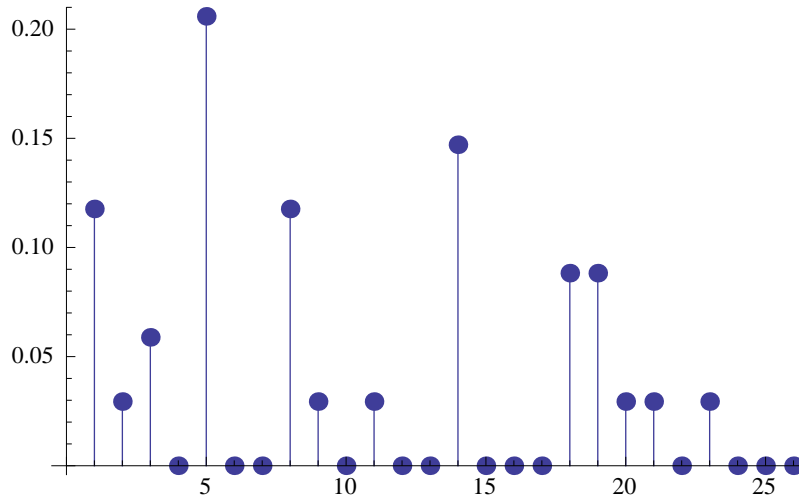


### c) EMPIRISCHE VERTEILUNG DER BUCHSTABEN »A« BIS »Z«:

- in dem Wort »Eisenbahnschrankenwaerterhaeuschen«:  
 freq = StringCount[»eisenbahnschrankenwaerterhaeuschen«, #]  
 &z/@CharacterRange[»a«, »z«];  
 relfreq = freq/Total[freq];



```
ListPlot[relfreq, Filling → Axis, PlotStyle → PointSize[Large]]
```

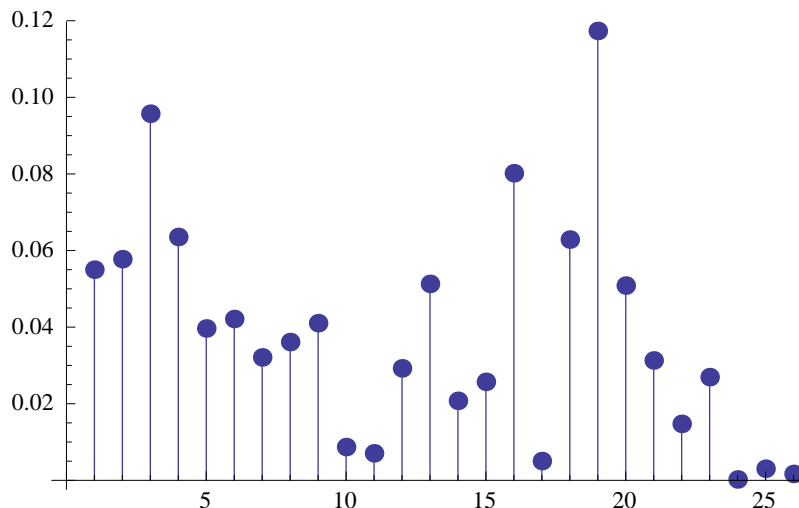


- in einem englischen Wörterbuch:

```
freq = Length[DictionaryLookup[# ~ ~ ____]]&/@CharacterRange[»a«, »z«];
```

```
relfreq = freq/Total[freq];
```

```
ListPlot[relfreq, Filling → Axis, PlotStyle → PointSize[Large]]
```



#### d) BENFORDSCHES GESETZ:

Das Benfordsche Gesetz, auch Newcomb-Benford's Law (NBL) beschreibt eine Gesetzmäßigkeit in der Verteilung der Ziffernstrukturen von Zahlen in empirischen Datensätzen, zum Beispiel ihrer ersten Ziffern. Es lässt sich etwa in Datensätzen über Einwohnerzahlen von Städten, Geldbeträge in der Buchhaltung, Naturkonstanten etc. beobachten. Kurz gefasst besagt es:

*»Je niedriger der zahlenmäßige Wert einer Ziffernsequenz bestimmter Länge an einer bestimmten Stelle einer Zahl ist, umso wahrscheinlicher ist ihr Auftreten. Für die Anfangsziffern in Zahlen des Zehnersystems gilt zum Beispiel: Zahlen mit der Anfangsziffer 1 treten etwa 6,5-mal so häufig auf wie solche mit der Anfangsziffer 9.«*

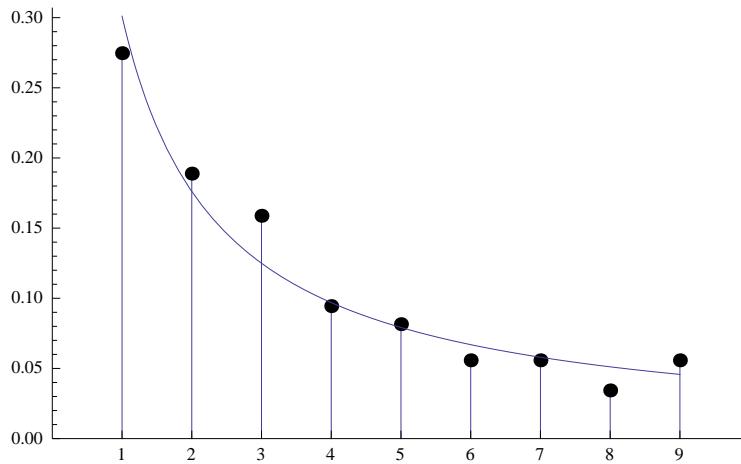
1881 wurde diese Gesetzmäßigkeit von dem Mathematiker Simon Newcomb entdeckt und im „American Journal of Mathematics“ publiziert. Er soll bemerkt haben, dass in den benutzten Büchern mit Logarithmentafeln, die Seiten mit Tabellen mit Eins als erster Ziffer deutlich schmutziger waren als die anderen Seiten, weil sie offenbar öfter benutzt worden seien. Die Abhandlung Newcombs blieb unbeachtet und war schon in Vergessenheit geraten, als der Physiker Frank Benford (1883–1948) diese Gesetzmäßigkeit wiederentdeckte und darüber 1938 neu publizierte. Seither war diese Gesetzmäßigkeit nach ihm benannt, in neuerer Zeit wird aber durch die Bezeichnung »Newcomb-Benford’s Law« (NBL) dem eigentlichen Urheber wieder Rechnung getragen. Bis vor wenigen Jahren war diese Gesetzmäßigkeit nicht einmal allen Statistikern bekannt. Erst seit der US-amerikanische Mathematiker Theodore Hill versucht hat, die Benford-Verteilung zur Lösung praktischer Probleme nutzbar zu machen, ist ihr Bekanntheitsgrad gewachsen.

(Quelle: Wikipedia) HÄUFIGKEITSVERTEILUNG DER ANFANGSZIFFERN VON ZAHLEN: Ist  $d$  die erste Ziffer einer Dezimalzahl, so tritt sie nach dem Benfordschen Gesetz in empirischen Datensätzen näherungsweise mit folgenden relativen Häufigkeiten  $p(d)$  auf:

$$p(d) = \log_{10} 1 + \frac{1}{d} = \log_{10} d + 1 - \log_{10} d.$$

In der Grafik unten (Quelle: Wolfram Demonstrations Project) werden die relativen Häufigkeiten der Anfangsziffern 1 bis 9 in den Anzahlen der Telefonanschlüsse in allen Ländern der Erde mit den nach dem Benfordschen Gesetz prognostizierten relativen Häufigkeiten

verglichen.



## 1.2 Zufallsvariablen und ihre Verteilung

Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum.

**Definition.** i) Eine *diskrete Zufallsvariable (ZV)* ist eine Abbildung

$$X: \Omega \rightarrow S, \quad S \text{ abzählbar,}$$

so daß für alle  $a \in S$  gilt:

$$X^{-1}(a) := \{\omega \in \Omega \mid X(\omega) = a\} \in \mathcal{A}. \quad (1.2.1)$$

Für  $X^{-1}(a)$  (das Urbild von  $a$  unter  $X$ ) schreiben wir im folgenden  $\{X = a\}$ .

ii) Die *Verteilung* von  $X$  ist die Wahrscheinlichkeitsverteilung  $\mu_X$  auf  $S$  mit Gewichten

$$p_X(a) := P[\{X = a\}], \quad (a \in S).$$

Für  $P[\{X = a\}]$  schreiben wir im folgenden  $P[X = a]$ .

**Bemerkung.** a) In der Tat ist  $p_X$  Massenfunktion einer Wahrscheinlichkeitsverteilung (siehe Satz 1.2):

i) Für alle  $a \in S$  gilt:  $p_X(a) \geq 0$

ii) Da die Ereignisse  $\{X = a\}$  disjunkt sind, folgt:

$$\sum_{a \in S} p_X(a) = \sum_{a \in S} P[X = a] = P\left[\bigcup_{a \in S} \{X = a\}\right] = P[\Omega] = 1.$$

b) Für  $B \subseteq S$  gilt:

$$\{X \in B\} := \underbrace{\{\omega \in \Omega \mid X(\omega) \in B\}}_{X^{-1}(B)} = \bigcup_{a \in B} \underbrace{\{X = a\}}_{\in \mathcal{A}}, \quad \text{sowie}$$

$$P[X \in B] = \sum_{a \in B} P[X = a] = \sum_{a \in B} p_X(a) = \mu_X(B).$$

Die Verteilung  $\mu_X$  gibt also an, mit welchen Wahrscheinlichkeiten die Zufallsvariable  $X$  Werte in bestimmten Mengen annimmt.

**Beispiel** (Zweimal würfeln). Sei  $\Omega = \{\omega = (\omega_1, \omega_2) \mid \omega_i \in \{1, \dots, 6\}\}$  und sei  $P$  die Gleichverteilung.

a)

$$\text{Sei } X_i: \Omega \rightarrow S := \{1, 2, 3, 4, 5, 6\},$$

$$X(\omega) := \omega_i, \quad \text{die Augenzahl des } i\text{-ten Wurfs.}$$

$X_i$  ist eine diskrete Zufallsvariable mit Verteilung  $\mu_{X_i}$ . Die Gewichte von  $\mu_{X_i}$  sind

$$p_{X_i}(a) = P[X_i = a] = \frac{6}{36} = \frac{1}{6} \quad \text{für alle } a \in S,$$

d.h.  $X_i$  ist gleichverteilt.

b)

$$\text{Sei } Y: \Omega \rightarrow \tilde{S} := \{2, 3, \dots, 12\}$$

$$Y(\omega) := X_1(\omega) + X_2(\omega), \quad \text{die Summe der Augenzahlen.}$$

Die Gewichte der Verteilung von  $Y$  sind

$$p_Y(a) = P[Y = a] = \begin{cases} \frac{1}{36} & \text{falls } a \in \{2, 12\}, \\ \frac{2}{36} & \text{falls } a \in \{3, 11\}, \\ \dots & \end{cases}$$

d.h.  $Y$  ist nicht mehr gleichverteilt!

Allgemeiner:

**Beispiel.** Sei  $\Omega = \{\omega_1, \dots, \omega_n\}$  endlich,  $P$  die Gleichverteilung,  $X: \Omega \rightarrow S$  eine Zufallsvariable und  $x_i := X(\omega_i)$ . Dann ist

$$P[X = a] = \frac{|\{\omega \in \Omega \mid X(\omega) = a\}|}{|\Omega|} = \frac{|\{1 \leq i \leq n \mid x_i = a\}|}{n},$$

also ist  $\mu_x$  die empirische Verteilung von  $x_1, \dots, x_n$ .

## Binomialverteilung

**Beispiel** (»Ziehen mit Zurücklegen«). Wir betrachten eine endliche Grundgesamtheit (Population, Zustandsraum)  $S$ , zum Beispiel Kugeln in einer Urne, Vögel im Wald, Einwohner in NRW. Wir wollen nun die zufällige Entnahme von  $n$  Einzelstichproben mit Zurücklegen aus  $S$  beschreiben und setzen daher

$$\Omega = S^n = \{\omega = (x_1, \dots, x_n) \mid x_i \in S\}.$$

Wir nehmen an, daß alle kombinierten Stichproben gleich wahrscheinlich sind, d.h.  $P$  sei die Gleichverteilung auf  $\Omega$ .

RELEVANTE ZUFALLSVARIABLEN UND EREIGNISSE:

- $i$ -ter Stichprobenwert:

$$X_i(\omega) = x_i,$$

$$P[X_i = a] = \frac{|S|^{n-1}}{|\Omega|} = \frac{|S|^{n-1}}{|S|^n} = \frac{1}{|S|}, \quad \text{für alle } a \in S,$$

d.h.  $X_i$  ist gleichverteilt auf  $S$ .

Sei  $E \subseteq S$  eine bestimmte Merkmalsausprägung der Stichprobe, die wir im folgenden als »Erfolg« bezeichnen (zum Beispiel schwarze Kugel, Beobachtung einer Amsel). Dann können wir die Ereignisse

$$\{X_i \in E\}, \text{ »Erfolg bei } i\text{-ter Stichprobe«},$$

betrachten. Es gilt:

$$P[X_i \in E] = \mu_{X_i}(E) = \frac{|E|}{|S|}.$$

Wir setzen

$$q := \frac{|E|}{|S|}, \quad \text{»Erfolgswahrscheinlichkeit«}$$

- Häufigkeit von  $E$  / »Anzahl der Erfolge«:

Sei nun

$$N: \Omega \rightarrow \{0, 1, 2, \dots, n\},$$

$$N(\omega) := |\{1 \leq i \leq n \mid X_i(\omega) \in E\}|$$

die Anzahl der Einzelstichproben mit Merkmalsausprägung  $E$ .

**Lemma 1.4.** Für  $k \in \{0, 1, \dots, n\}$  gilt:

$$P[N = k] = \binom{n}{k} p^k (1 - p)^{n-k}.$$

*Beweis.* Es gilt

$$|\{\omega \in \Omega \mid N(\omega) = k\}| = \binom{n}{k} |E|^k |S \setminus E|^{n-k},$$

wobei

$$\binom{n}{k} = \text{Anzahl der Möglichkeiten } k \text{ Indizes aus } \{1, \dots, n\} \text{ auszuwählen,}$$

für die ein Erfolg eintritt,

$$|E|^k = \text{Anzahl der Möglichkeiten für jeden Erfolg,}$$

$$|S \setminus E|^{n-k} = \text{Anzahl der Möglichkeiten für jeden Mißerfolg.}$$

Also gilt:

$$\begin{aligned} P[N = k] &= \frac{\binom{n}{k} |E|^k |S \setminus E|^{n-k}}{|S|^n} = \binom{n}{k} \left(\frac{|E|}{|S|}\right)^k \left(\frac{|S \setminus E|}{|S|}\right)^{n-k} \\ &= \binom{n}{k} p^k (1 - p)^{n-k}. \end{aligned}$$

□

**Definition.** Sei  $n \in \mathbb{N}$  und  $p \in [0, 1]$ . Die Wahrscheinlichkeitsverteilung auf  $\{0, 1, \dots, n\}$  mit Massenfunktion

$$p_{n,p}(k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

heißt **Binomialverteilung mit Parametern  $n$  und  $p$**  (kurz:  $\text{Bin}(n, p)$ ).

**Bemerkung.** Daß  $p_{n,p}$  eine Massenfunktion einer Wahrscheinlichkeitsverteilung ist, folgt aus Lemma 1.3!

**Bemerkung.** Ereignisse  $E_1, \dots, E_n$  heißen **unabhängig**, falls

$$P[E_{i_1} \cap E_{i_2} \cap \dots \cap E_{i_k}] = P[E_{i_1}] \cdot P[E_{i_2}] \cdots P[E_{i_k}]$$

für alle  $k \leq n$  und  $1 \leq i_1 < i_2 < \dots < i_k \leq n$  gilt.

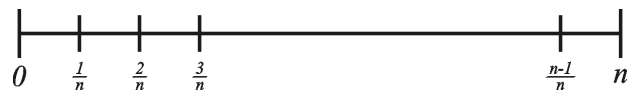
Sind  $E_1, \dots, E_n$  unabhängig und  $P[E_i] = p$ , dann ist

$$P[\text{»genau } k \text{ der } E_i \text{ treten ein«}] = \binom{n}{k} p^k (1 - p)^{n-k},$$

d.h. die Anzahl der Ereignisse, die eintreten, ist binomialverteilt. Der Beweis folgt weiter unten.

## Poissonverteilung

**Beispiel** (Warteschlange). Angenommen, die Kunden in einer Warteschlange kommen *unabhängig voneinander* zu *zufälligen* (gleichverteilten) Zeitpunkten. Wieviele Kunden kommen in einer Zeitspanne der Länge  $t_0$  an? Sei  $N$  die Anzahl dieser Kunden und  $t_0 = 1$ . Wir unterteilen das Intervall  $[0, 1]$ :



Wir machen die folgende Annahme (die natürlich in zu modellierenden Anwendungsproblemen zu überprüfen ist):

»Wenn  $n$  sehr groß ist, dann kommt in einer Zeitspanne der Länge  $\frac{1}{n}$  fast immer höchstens ein Kunde«.

$E_i$  stehe für das Ereignis, daß ein Kunde im Zeitintervall  $[\frac{i-1}{n}, \frac{i}{n}]$  ankommt ( $1 \leq i \leq n$ ). Wir nehmen außerdem an, daß die Wahrscheinlichkeit unabhängig von  $i$  und näherungsweise proportional zu  $\frac{1}{n}$  ist, also:

$$P[E_i] \approx \frac{\lambda}{n}, \quad \lambda \in (0, \infty).$$

Für das Ereignis, daß genau  $k$  Kunden im Zeitintervall  $[0, 1]$  ankommen, sollte dann gelten, dass

$$P[N = k] \approx P[\text{»genau } k \text{ der } E_i \text{ treten ein«}] \approx p_{n, \frac{\lambda}{n}}(k),$$

wobei  $p_{n, \frac{\lambda}{n}}(k)$  das Gewicht von  $k$  unter der Binomialverteilung mit Parametern  $n$  und  $\frac{\lambda}{n}$  ist. Diese Näherung sollte »für große  $n$  immer genauer werden«.

**Satz 1.5** (Poissonapproximation der Binomialverteilung). Sei  $\lambda \in (0, \infty)$ . Dann gilt:

$$\lim_{n \rightarrow \infty} p_{n, \frac{\lambda}{n}}(k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots$$

*Beweis.* Es gilt:

$$\begin{aligned} p_{n, \frac{\lambda}{n}}(k) &= \frac{n!}{k!(n-k)!} \cdot \left(\frac{\lambda}{n}\right)^k \cdot \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \frac{\lambda^k}{k!} \cdot \underbrace{\frac{n \cdot (n-1) \cdot \dots \cdot (n-k+1)}{n^k}}_{\rightarrow 1} \cdot \underbrace{\left(1 - \frac{\lambda}{n}\right)^n}_{\rightarrow e^{-\lambda}} \cdot \underbrace{\left(1 - \frac{\lambda}{n}\right)^{-k}}_{\rightarrow 1} \\ &\rightarrow \frac{\lambda^k}{k!} e^{-\lambda} \quad \text{für } n \rightarrow \infty. \end{aligned}$$

□

**Definition.** Die Wahrscheinlichkeitsverteilung auf  $\{0, 1, 2, \dots\}$  mit Massenfunktion

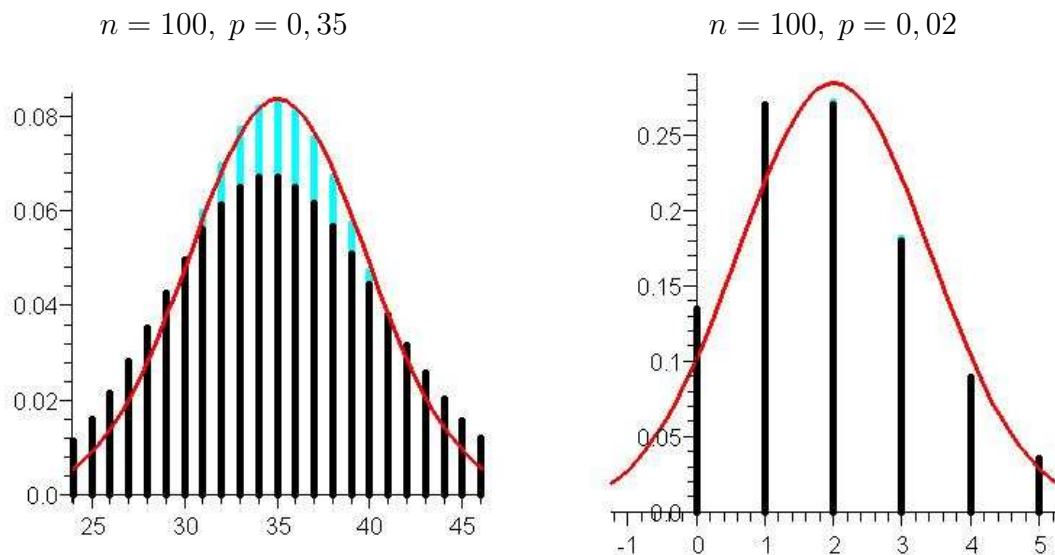
$$p(k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots,$$

heißt **Poissonverteilung mit Parameter  $\lambda$** .

Aufgrund des Satzes verwendet man die Poissonverteilung zur näherungsweisen Modellierung der Häufigkeit seltener Ereignisse (zum Beispiel Tippfehler in einem Buch, Schadenfälle bei Versicherung, Zusammenbrüche des T-Mobile-Netzes, ...) und damit zur »Approximation« von Binomialverteilungen mit kleiner Erfolgswahrscheinlichkeit  $p$ .

Für häufigere Ereignisse (zum Beispiel wenn Erfolgswahrscheinlichkeit  $p$  unabhängig von  $n$  ist) verwendet man hingegen besser eine Normalverteilung zur näherungsweisen Modellierung der (geeignet reskalierten) relativen Häufigkeit  $\frac{k}{n}$  des Ereignisses für große  $n$ . Definition und Eigenschaften von Normalverteilungen werden wir später kennenlernen.

Die folgenden (mit »Maple« erstellten) Graphiken zeigen die Poisson- und Normalapproximation (Poisson schwarz, Normalverteilung rot) der Binomialverteilung (blau) für unterschiedliche Parameterwerte:



## Hypergeometrische Verteilung

**Beispiel** (Ziehen ohne Zurücklegen). Wir betrachten  $m$  Kugeln in einer Urne (Wähler, Fische im See, ...), davon  $r$  rote und  $m - r$  schwarze. Gezogen wird eine zufällige Stichprobe von  $n$



Kugeln,  $n \leq \min(r, m - r)$ . Sind alle Stichproben gleich wahrscheinlich, dann ist ein geeignetes Modell gegeben durch:

$$\begin{aligned}\Omega &= \mathcal{P}(\{1, \dots, m\}) = \text{alle Teilmengen von } \{1, \dots, m\} \text{ der Kardinalität } n, \\ P &= \text{Gleichverteilung auf } \Omega.\end{aligned}$$

Wir definieren eine Zufallsvariable  $N: \Omega \rightarrow \{1, \dots, m\}$  durch

$$N(\omega) := \text{Anzahl der roten Kugeln in } \omega.$$

Für das Ereignis, daß genau  $k$  rote Kugeln in der Stichprobe sind, gilt:

$$P[N = k] = \frac{|\{\omega \in \Omega \mid N(\omega) = k\}|}{|\Omega|} = \frac{\binom{r}{k} \cdot \binom{m-r}{n-k}}{\binom{m}{n}}, \quad (k = 0, 1, \dots, n).$$

Diese Wahrscheinlichkeitsverteilung wird **hypergeometrische Verteilung mit Parametern  $m$ ,  $r$  und  $n$**  genannt.

**Bemerkung.** Untersucht man die Asymptotik der hypergeometrischen Verteilung für  $m \rightarrow \infty$ ,  $r \rightarrow \infty$ ,  $p = \frac{r}{m}$  fest und  $n$  fest, so gilt:

$$P[N = k] \longrightarrow \binom{n}{k} p^k (1 - p)^{n-k},$$

d.h. die hypergeometrische Verteilung nähert sich der Binomialverteilung an. Eine anschauliche Erklärung dafür ist:

Befinden sich sehr viele Kugeln in der Urne, dann ist der Unterschied zwischen Ziehen mit und ohne Zurücklegen vernachlässigbar, da nur sehr selten dieselbe Kugel zweimal gezogen wird.

### 1.3 Simulation von Gleichverteilungen

Ein (*Pseudo-*) *Zufallszahlengenerator* ist ein Algorithmus, der eine deterministische Folge von ganzen Zahlen  $x_1, x_2, x_3, \dots$  mit Werten zwischen 0 und einem Maximalwert  $m - 1$  erzeugt, welche durch eine vorgegebene Klasse statistischer Tests nicht von einer Folge von Stichproben unabhängiger, auf  $\{0, 1, 2, \dots, m - 1\}$  gleichverteilter Zufallsgrößen unterscheidbar ist. Ein Zufallszahlengenerator erzeugt also nicht wirklich zufällige Zahlen. Die von »guten« Zufallszahlengeneratoren erzeugten Zahlen haben aber statistische Eigenschaften, die denen von echten Zufallszahlen in vielerlei (aber nicht in jeder) Hinsicht sehr ähnlich sind.

Konkret werden die Pseudozufallszahlen üblicherweise über eine deterministische Rekurrenzrelation vom Typ

$$x_{n+1} = f(x_{n-k+1}, x_{n-k+2}, \dots, x_n), \quad n = k, k+1, k+2, \dots,$$

aus *Saatwerten*  $x_1, x_2, \dots, x_k$  erzeugt. In vielen Fällen hängt die Funktion  $f$  nur von der letzten erzeugten Zufallszahl  $x_n$  ab. Wir betrachten einige Beispiele:

### Lineare Kongruenzgeneratoren (LCG)

Bei linearen Kongruenzgeneratoren ist die Rekurrenzrelation vom Typ

$$x_{n+1} = (ax_n + c) \bmod m, \quad n = 0, 1, 2, \dots$$

Hierbei sind  $a$ ,  $c$  und  $m$  geeignet zu wählende positive ganze Zahlen, zum Beispiel:

$$\text{ZX81-Generator :} \quad m = 2^{16} + 1, \quad a = 75, \quad c = 0.$$

$$\text{RANDU, IBM 360/370 :} \quad m = 2^{31}, \quad a = 65539, \quad c = 0.$$

$$\text{Marsaglia-Generator :} \quad m = 2^{32}, \quad a = 69069, \quad c = 1.$$

$$\text{Langlands-Generator :} \quad m = 2^{48}, \quad a = 142412240584757, \quad c = 11.$$

Um einen ersten Eindruck zu erhalten, wie die Qualität der erzeugten Pseudozufallszahlen von  $a$ ,  $c$  und  $m$  abhängt, implementieren wir die Generatoren mit »Mathematica«:

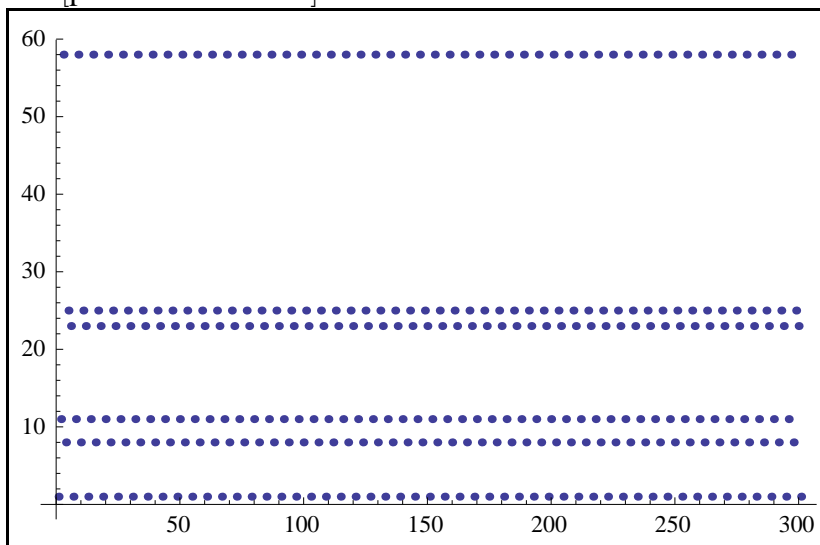
$$f[x\_]:=Mod[ax + c, m]$$

**Beispiel.** Wir beginnen zur Demonstration mit dem Beispiel eines ganz schlechten LCG:

$$a = 11; c = 0; m = 63;$$

$$\text{pseudorandomdata} = \text{NestList}[f, 1, 300];$$

$$\text{ListPlot}[\text{pseudorandomdata}]$$



Die Folge von Zufallszahlen ist in diesem Fall periodisch mit einer Periode, die viel kleiner ist als die maximal mögliche (63). Dies rechnet man auch leicht nach.

Periodizität mit Periode kleiner als  $m$  kann man leicht ausschließen. Es gilt nämlich:

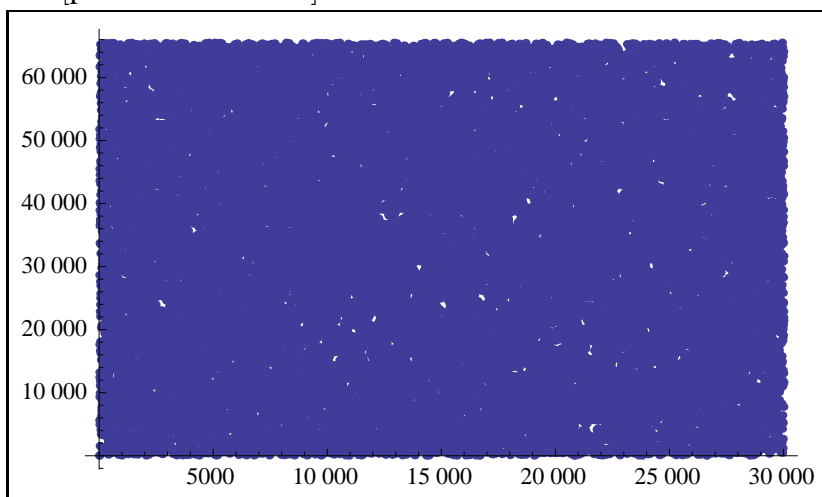
**Satz (Knuth).** *Die Periode eines LCG ist gleich  $m$  genau dann, wenn*

- i)  $c$  und  $m$  teilerfremd sind,*
- ii) jeder Primfaktor von  $m$  ein Teiler von  $a - 1$  ist, und*
- iii) falls 4 ein Teiler von  $m$  ist, dann auch von  $a - 1$ .*

*Beweis.* siehe D. Knuth: »The art of computer programming, Vol. 2.« □

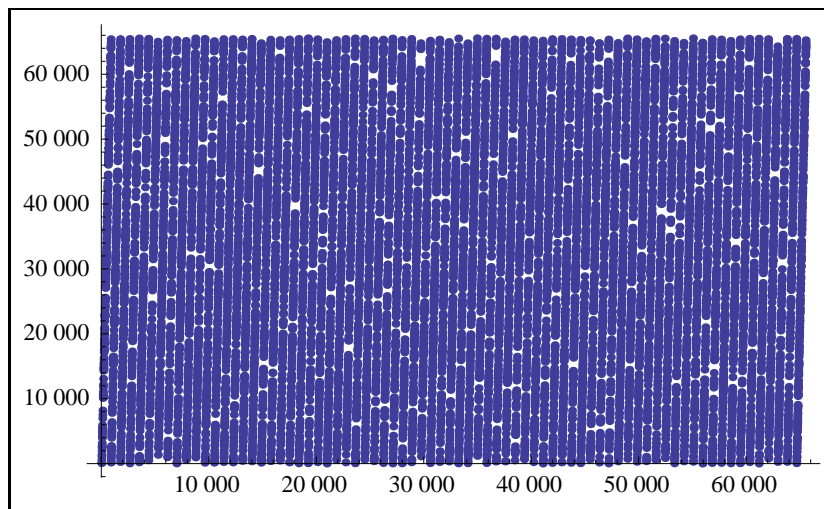
**Beispiel (ZX 81-Generator).** Hier ergibt sich ein besseres Bild, solange wir nur die Verteilung der einzelnen Zufallszahlen betrachten:

```
a = 75; c = 0; m = 2^16 + 1;
pseudorandomdata = NestList[f, 1, 30000];
ListPlot[pseudorandomdata]
```



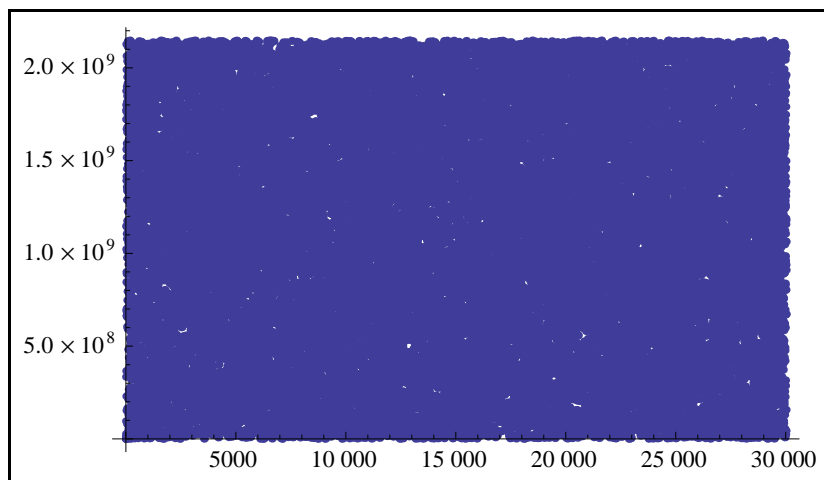
Fassen wir jedoch Paare  $(x_i, x_{i+1})$  von aufeinanderfolgenden Pseudozufallszahlen als Koordinaten eines zweidimensionalen Pseudozufallsvektors auf, und betrachten die empirische Verteilung dieser Vektoren, so ergibt sich keine besonders gute Approximation einer zweidimensionalen Gleichverteilung:

```
blocks = Partition[pseudorandomdata, 2];
ListPlot[blocks]
```

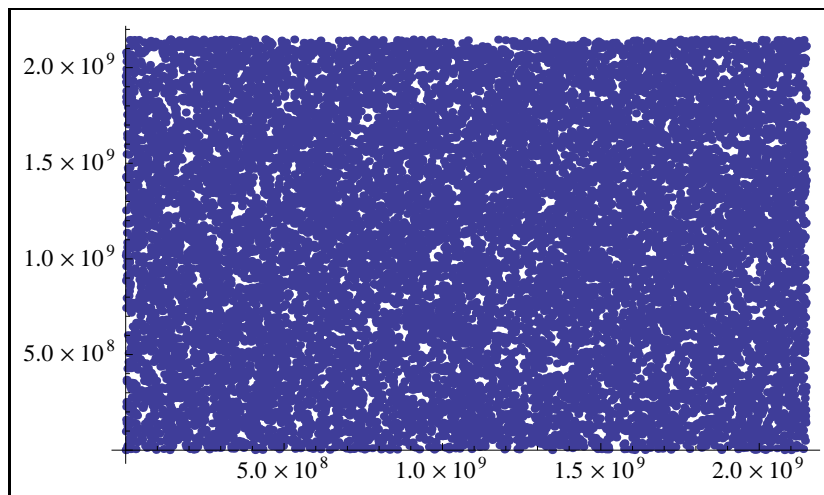


**Beispiel (RANDU).** Hier scheinen sowohl die einzelnen Pseudozufallszahlen  $x_i$  als auch die Vektoren  $(x_i, x_{i+1})$  näherungsweise gleichverteilt zu sein:

```
a = 65539; c = 0; m = 2^31;
pseudorandomdata = NestList[f, 1, 30000];
ListPlot[pseudorandomdata]
```



```
blocks = Partition[pseudorandomdata, 2];
ListPlot[blocks]
```

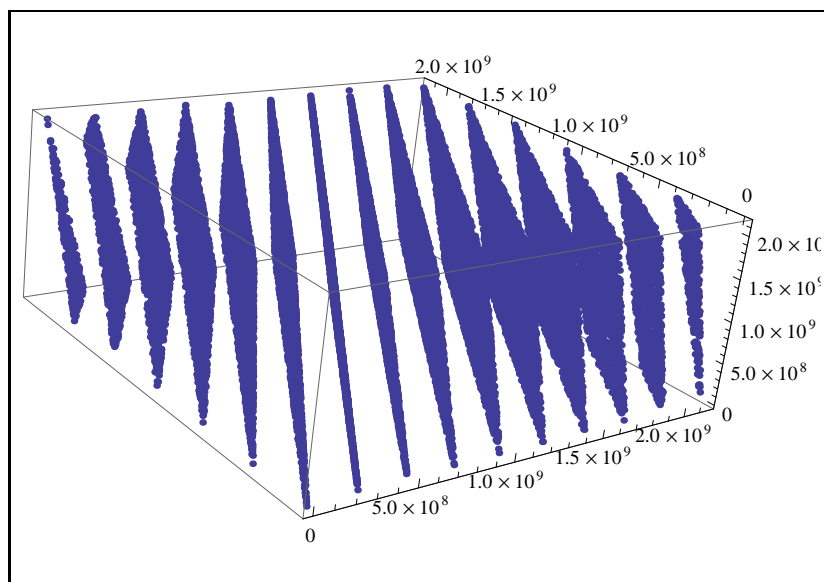


Fassen wir aber jeweils drei aufeinanderfolgende Pseudozufallszahlen als Koordinaten eines Vektors

$(x_i, x_{i+1}, x_{i+2})$  im  $\mathbb{Z}^3$  auf, dann ist die empirische Verteilung dieser Pseudozufallsvektoren keine Gleichverteilung mehr, sondern konzentriert sich auf nur 15 zweidimensionalen Hyperebenen:

```
blocks3 = Partition[pseudorandomdata, 3];
```

```
ListPointPlot3D[blocks3]
```

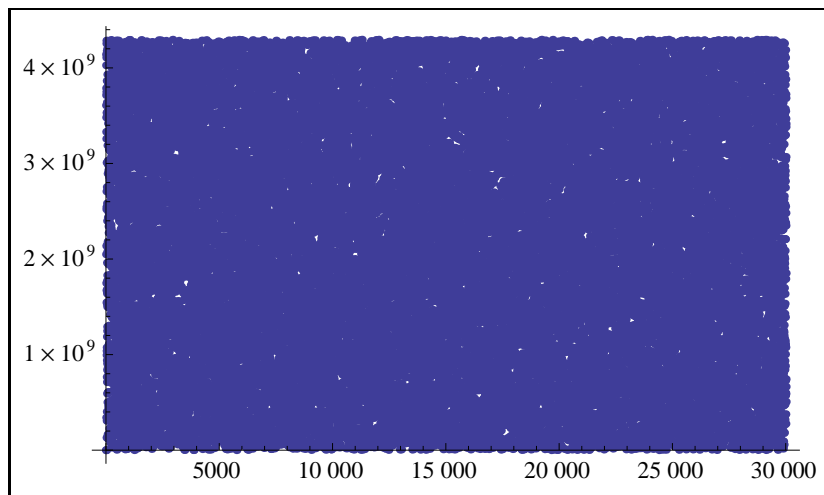


**Beispiel** (Marsaglia-Generator). Der von Marsaglia 1972 vorgeschlagene LCG besteht dagegen alle obigen Tests (und einige weitere):

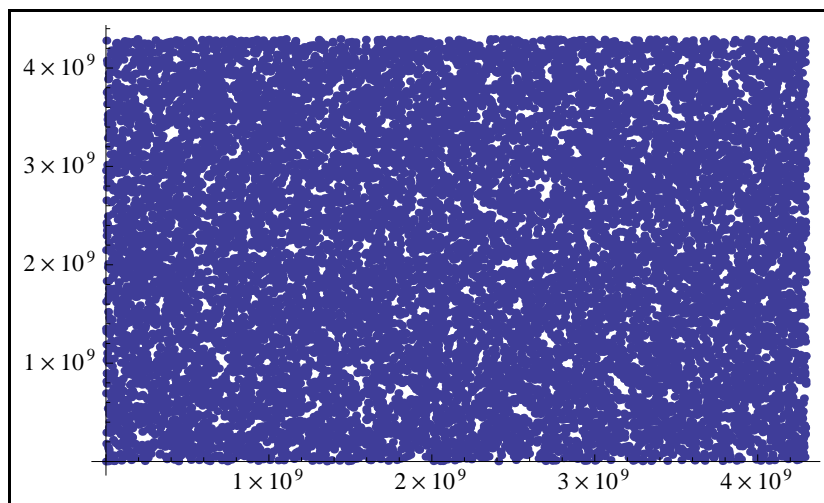
```
a = 60069; c = 1; m = 2^32;
```

```
pseudorandomdata = NestList[f, 1, 30000];
```

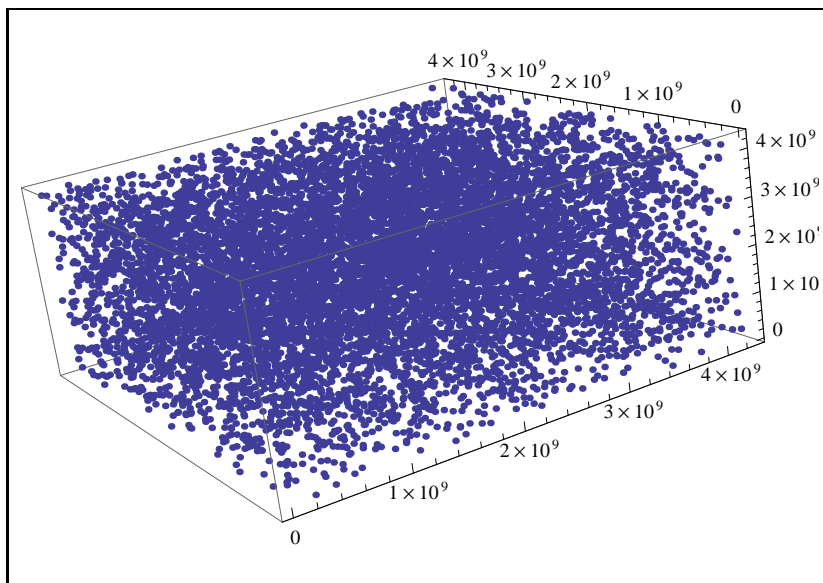
```
ListPlot[pseudorandomdata]
```



```
blocks = Partition[pseudorandomdata, 2];  
ListPlot[blocks]
```



```
blocks3 = Partition[pseudorandomdata, 3];  
ListPointPlot3D[blocks3]
```



Dies bedeutet natürlich nicht, daß die vom Marsaglia-Generator erzeugte Folge eine für *alle* Zwecke akzeptable Approximation einer Folge von unabhängigen Stichproben von der Gleichverteilung ist. Da die Folge in Wirklichkeit deterministisch ist, kann man einen Test konstruieren, der sie von einer echten Zufallsfolge unterscheidet.

### Shift-Register-Generatoren

Bei Shift-Register-Generatoren interpretiert man eine Zahl  $x_n \in \{0, 1, \dots, 2^k - 1\}$  zunächst als Binärzahl bzw. als Vektor aus  $\{0, 1\}^k$ , und wendet dann eine gegebene Matrix  $T$  darauf an, um  $x_{n+1}$  zu erhalten:

$$x_{n+1} = T x_n, \quad n = 0, 1, 2, \dots$$

### Kombination von Zufallszahlengeneratoren

Zufallszahlengeneratoren lassen sich kombinieren, zum Beispiel indem man die von mehreren Zufallszahlengeneratoren erzeugten Folgen von Pseudozufallszahlen aus  $\{0, 1, \dots, m - 1\}$  modulo  $m$  addiert. Auf diese Weise erhält man sehr leistungsfähige Zufallszahlengeneratoren, zum Beispiel den Kiss-Generator von Marsaglia, der einen LCG und zwei Shift-Register-Generatoren kombiniert, Periode  $2^{95}$  hat, und umfangreiche statistische Tests besteht.

### Zufallszahlen aus $[0, 1)$

Ein Zufallszahlengenerator kann natürlich nicht wirklich reelle Pseudozufallszahlen erzeugen, die die Gleichverteilung auf dem Intervall  $[0, 1)$  simulieren, denn dazu würden unendlich vie-

le »zufällige« Nachkommastellen benötigt. Stattdessen werden üblicherweise (pseudo-)zufällige Dezimalzahlen vom Typ

$$u_n = \frac{x_n}{m}, \quad x_n \in \{0, 1, \dots, m-1\},$$

erzeugt, wobei  $m$  vorgegeben ist (zum Beispiel Darstellungsgenauigkeit des Computers), und  $x_n$  eine Folge ganzzahliger Pseudozufallszahlen aus  $\{0, 1, \dots, m-1\}$  ist. In »Mathematica« kann man mit

`RandomReal` [*spec*, WorkingPrecision  $\rightarrow k$ ]

pseudozufällige Dezimalzahlen mit einer beliebigen vorgegebenen Anzahl  $k$  von Nachkommastellen erzeugen.

## Zufallspermutationen

Der folgende Algorithmus erzeugt eine (pseudo-)zufällige Permutation aus  $S_n$  :

**Algorithmus 1.6** (RPERM). `rperm[n_]` :=

```
Module[{x = Range[n], k, a},
  (* Beginn mit Liste {1, 2, ..., n} *)
  Do[
    k = RandomInteger[{i, n}];
    a = x[[i]];
    x[[i]] = x[[k]];
    x[[k]] = a;
    (* Vertausche x[[i]] und x[[k]] *)
    , {i, n - 1}];
  (* Schleife, i läuft von 1 bis n-1 *)
  x (* Ausgabe von x *)]
```

`rperm[17]` {12, 5, 13, 8, 17, 9, 10, 6, 1, 7, 16, 15, 14, 4, 2, 3, 11}

ÜBUNG:

Sei  $\Omega_n = \{1, 2, \dots, n\} \times \{2, 3, \dots, n\} \times \dots \times \{n-1, n\}$ .

- Zeigen Sie, daß die Abbildung  $X(\omega) = \tau_{n-1, \omega_{n-1}} \circ \dots \circ \tau_{2, \omega_2} \circ \tau_{1, \omega_1}$  eine Bijektion von  $\Omega_n$  nach  $S_n$  ist ( $\tau_{i,j}$  bezeichnet die Transposition von  $i$  und  $j$ ).
- Folgern Sie, daß der Algorithmus oben tatsächlich eine Stichprobe einer gleichverteilten Zufallspermutation aus  $S_n$  simuliert.



## 1.4 Erwartungswert

Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum,  $S \subseteq \mathbb{R}$  abzählbar und  $X: \Omega \rightarrow S$  eine Zufallsvariable auf  $(\Omega, \mathcal{A}, P)$ .

**Definition.** Der *Erwartungswert* von  $X$  bzgl.  $P$  ist definiert als

$$E[X] := \sum_{a \in S} a \cdot P[X = a] = \sum_{a \in S} a \cdot p_X(a),$$

sofern die Summe auf der rechten Seite wohldefiniert ist (d.h. unabhängig von der Abzählung von  $S$ ).

**Bemerkung.** a) Falls  $X(\omega) \geq 0$  für alle  $\omega \in \Omega$  gilt, sind alle Summanden der Reihe nichtnegativ und der Erwartungswert  $E[X] \in [0, \infty]$  wohldefiniert.

b) Falls die Reihe absolut konvergiert, d.h. falls  $\sum_{a \in S} |a| \cdot P[X = a]$  endlich ist, ist der Erwartungswert  $E[X] \in \mathbb{R}$  wohldefiniert.

$E[X]$  kann als der *Prognosewert* oder (*gewichteter*) *Mittelwert* für  $X(\omega)$  interpretiert werden.

**Beispiel** (Indikatorfunktion eines Ereignisses  $A \in \mathcal{A}$ ). Sei

$$X(\omega) = I_A(\omega) := \begin{cases} 1 & \text{falls } \omega \in A, \\ 0 & \text{falls } \omega \in A^C. \end{cases}$$

Dann ist der Erwartungswert

$$E[X] = 1 \cdot P[X = 1] + 0 \cdot P[X = 0] = P[A].$$

Ein Beispiel dafür ist ein elementarer Versicherungsvertrag mit Leistung

$$Y = \begin{cases} c & \text{falls } \omega \in A, \quad \text{»Schadensfall«,} \\ 0 & \text{sonst.} \end{cases}$$

Dann gilt:

$$Y = c \cdot I_A \quad \text{und} \quad E[Y] = c \cdot P[A].$$

**Beispiel** (Poissonverteilung). Sei  $X$  Poisson-verteilt mit Parameter  $\lambda$ . Dann ist der Erwartungswert

$$E[X] = \sum_{k=0}^{\infty} k \cdot P[X = k] = \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k}{k!} e^{-\lambda} = \lambda \cdot \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} = \lambda \cdot \sum_{k=0}^{\infty} \frac{\lambda^k}{(k)!} e^{-\lambda} = \lambda.$$

Wir können daher den Parameter  $\lambda$  als Erwartungswert oder die mittlere Häufigkeit des Ereignisses interpretieren.

Sei nun  $S$  eine beliebige abzählbare Menge,  $g: S \rightarrow \mathbb{R}$  eine Funktion und  $X: \Omega \rightarrow S$  eine Zufallsvariable. Wir definieren

$$\begin{aligned} g(X): \Omega &\rightarrow \mathbb{R}, \\ \omega &\mapsto g(X(\omega)). \end{aligned}$$

$g(X)$  ist eine *reellwertige Zufallsvariable*.

**Satz 1.7** (Transformationssatz). *Es gilt*

$$E[g(X)] = \sum_{a \in S} g(a) \cdot P[X = a],$$

*falls die Summe wohldefiniert ist (zum Beispiel falls  $g$  nichtnegativ ist oder die Summe absolut konvergiert).*

*Beweis.* Es gilt mit Verwendung der  $\sigma$ -Additivität

$$\begin{aligned} E[g(X)] &= \sum_{b \in g(S)} b \cdot P[g(X) = b] = \sum_{b \in g(S)} b \cdot P\left[\bigcup_{a \in g^{-1}(b)} \{X = a\}\right] \\ &= \sum_{b \in g(S)} b \cdot \sum_{a \in g^{-1}(b)} P[X = a] \\ &= \sum_{b \in g(S)} \sum_{a \in g^{-1}(b)} g(a) \cdot P[X = a] \\ &= \sum_{a \in S} g(a) \cdot P[X = a]. \end{aligned}$$

□

**Bemerkung.** a) Insbesondere gilt:

$$E[|X|] = \sum_{a \in S} |a| \cdot P[X = a].$$

Ist  $E[|X|]$  endlich, dann konvergiert  $E[X] = \sum a \cdot P[X = a]$  absolut.

b) Ist  $\Omega$  abzählbar, dann folgt für  $X: \Omega \rightarrow \mathbb{R}$ :

$$E[X] = E[X \circ id_{\Omega}] = \sum_{\omega \in \Omega} X(\omega) \cdot P[\{\omega\}] = \sum_{\omega \in \Omega} X(\omega) p(\omega),$$

wobei  $id_{\Omega}$  die identische Abbildung auf  $\Omega$  bezeichnet. Der Erwartungswert ist das *gewichtete Mittel*. Ist  $P$  die Gleichverteilung auf  $\Omega$ , folgt weiter:

$$E[X] = \frac{1}{|\Omega|} \sum_{\omega \in \Omega} X(\omega).$$

Der Erwartungswert ist in diesem Spezialfall das *arithmetische Mittel*.

**Beispiel (Sankt-Petersburg-Paradoxon).** Wir betrachten ein Glücksspiel mit fairen Münzwürfen  $X_1, X_2, \dots$ , wobei sich der Gewinn in jeder Runde verdoppelt bis zum ersten Mal »Kopf« fällt, dann ist das Spiel beendet. Wie hoch wäre eine faire Teilnahmegebühr für dieses Spiel?

Der Gewinn ist

$$G(\omega) = 2^{T(\omega)}, \quad \text{mit}$$

$$T(\omega) := \min\{n \in \mathbb{N} \mid X_n(\omega) = 1\}, \quad \text{der Wartezeit auf »Kopf«.}$$

Für den erwarteten Gewinn ergibt sich

$$\begin{aligned} E[G] &= \sum_{k=1}^{\infty} 2^k \cdot P[T = k] = \sum_{k=1}^{\infty} 2^k \cdot P[X_1 = \dots = X_{k-1} = 1, X_k = 0] = \sum_{k=1}^{\infty} 2^k 2^{-k} \\ &= \infty. \end{aligned}$$

Das Spiel sollte also auf den ersten Blick bei beliebig hoher Teilnahmegebühr attraktiv sein – dennoch wäre wohl kaum jemand bereit, einen sehr hohen Einsatz zu zahlen.

Eine angemessenere Beschreibung – vom Blickwinkel des Spielers aus betrachtet – erhält man, wenn man eine (üblicherweise als monoton wachsend und konkav vorausgesetzte) Nutzenfunktion  $u(x)$  einführt, die den Nutzen beschreibt, den der Spieler vom Kapital  $x$  hat. Für kleine  $x$  könnte etwa  $u(x) = x$  gelten, aber für große  $x$  wäre plausibler  $u(x) < x$ . Dann ist  $c$  ein fairer Einsatz aus Sicht des Spielers, wenn  $u(c) = E[u(G)]$  gilt.

**Satz 1.8** (Linearität des Erwartungswerts). *Seien  $X : \Omega \rightarrow S_X \subseteq \mathbb{R}$  und  $Y : \Omega \rightarrow S_Y \subseteq \mathbb{R}$  diskrete reellwertige Zufallsvariablen auf  $(\Omega, \mathcal{A}, P)$ , für die  $E[|X|]$  und  $E[|Y|]$  endlich sind, dann gilt:*

$$E[\lambda X + \mu Y] = \lambda E[X] + \mu E[Y] \quad \text{für alle } \lambda, \mu \in \mathbb{R}.$$

*Beweis.* Wir definieren  $g : S_X \times S_Y \rightarrow \mathbb{R}$ ,  $(x, y) \mapsto \lambda x + \mu y$ . Dann ist  $g(X, Y) = \lambda X + \mu Y$  eine Zufallsvariable mit Werten in  $S_X \times S_Y$ . Mit dem Transformationssatz folgt:

$$\begin{aligned} E[\lambda X + \mu Y] &= E[g(X, Y)] \\ &= \sum_{a \in S_X} \sum_{b \in S_Y} g(a, b) P[X = a, Y = b] \\ &= \sum_{a \in S_X} \sum_{b \in S_Y} (\lambda a + \mu b) P[X = a, Y = b] \\ &= \lambda \sum_{a \in S_X} a \sum_{b \in S_Y} P[X = a, Y = b] + \mu \sum_{b \in S_Y} b \sum_{a \in S_X} P[X = a, Y = b] \\ &= \lambda \sum_{a \in S_X} a P[X = a] + \mu \sum_{b \in S_Y} b P[Y = b] \\ &= \lambda E[X] + \mu E[Y]. \end{aligned} \tag{1.4.1}$$

Hierbei konvergiert die Reihe in (1.4.1) absolut, da

$$\begin{aligned} \sum_{a \in S_X} \sum_{b \in S_Y} |\lambda a + \mu b| P[X = a, Y = b] &\leq |\lambda| \sum_{a \in S_X} |a| P[X = a] + |\mu| \sum_{b \in S_Y} |b| P[Y = b] \\ &= |\lambda| E[|X|] + |\mu| E[|Y|] \end{aligned}$$

nach Voraussetzung endlich ist. □

**Korollar** (Monotonie des Erwartungswerts). *Seien die Voraussetzungen von Satz 1.8 erfüllt. Sei zusätzlich  $X(\omega) \leq Y(\omega)$  für alle  $\omega \in \Omega$ , dann gilt:*

$$E[X] \leq E[Y].$$

*Beweis.* Nach Voraussetzung gilt  $(Y - X)(\omega) \geq 0$  für alle  $\omega \in \Omega$ , weshalb der Erwartungswert  $E[Y - X]$  nichtnegativ ist. Aufgrund der Linearität des Erwartungswerts folgt:

$$0 \leq E[Y - X] = E[Y] - E[X].$$

□

**Beispiel** (Unabhängige 0-1-Experimente). Seien  $A_1, A_2, \dots, A_n \in \mathcal{A}$  unabhängige Ereignisse mit Wahrscheinlichkeit  $p$ , und sei

$$X_i = I_{A_i}, \quad \text{die Indikatorfunktion des Ereignisses } A_i, i = 0, \dots, n.$$

a) Die Zufallsvariablen  $X_i$  sind *Bernoulli-verteilt mit Parameter  $p$* , d.h.

$$X_i = \begin{cases} 1 & \text{mit Wahrscheinlichkeit } p, \\ 0 & \text{mit Wahrscheinlichkeit } 1 - p. \end{cases}$$

Also gilt:

$$E[X_i] = E[I_{A_i}] = P[A_i] = p,$$

analog zu Beispiel 1.4.

b) Die Anzahl

$$S_n = X_1 + X_2 + \dots + X_n$$

der Ereignisse, die eintreten, ist binomialverteilt mit Parametern  $n$  und  $p$  (siehe Übung), d.h.

$$P[S_n = k] = \binom{n}{k} p^k (1 - p)^{n-k}.$$

Den Erwartungswert kann man daher wie folgt berechnen:

$$\begin{aligned} E[S_n] &= \sum_{k=0}^n k \cdot P[S_n = k] = \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k} \\ &= \dots = np. \end{aligned}$$

Einfacher benutzt man aber die Linearität des Erwartungswerts, und erhält

$$E[S_n] = E\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n E[X_i] = np,$$

sogar *ohne Verwendung der Unabhängigkeit!*

**Beispiel** (Abhängige 0-1-Experimente). Wir betrachten eine Population aus  $m$  Objekten, davon  $r$  rote, aus der eine Zufallsstichprobe aus  $n$  Objekten ohne Zurücklegen entnommen wird,  $n \leq \min(r, m-r)$ . Sei  $A_i$  das Ereignis, daß das  $i$ -te Objekt in der Stichprobe rot ist, und  $X_i = I_{A_i}$ . Die Anzahl

$$S_n = X_1 + \dots + X_n$$

der roten Objekte in der Zufallsstichprobe ist dann hypergeometrisch verteilt mit Parametern  $m$ ,  $r$  und  $n$ . Als Erwartungswert dieser Verteilung erhalten wir analog zum letzten Beispiel:

$$E[S_n] = \sum_{i=1}^n E[X_i] = \sum_{i=1}^n P[A_i] = n \frac{r}{m}.$$

**Beispiel** (Inversionen von Zufallspermutationen). Seien  $\Omega = S_n$  die Menge aller Permutationen  $\omega: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ ,  $P$  die Gleichverteilung auf  $\Omega$ , und

$$N(\omega) = |\{(i, j) \mid i < j \text{ und } \omega(i) > \omega(j)\}|,$$

die Anzahl der Inversionen einer Permutation  $\omega \in \Omega$ . Dann gilt

$$\begin{aligned} N &= \sum_{1 \leq i < j \leq n} I_{A_{i,j}}, \quad \text{wobei} \\ A_{i,j} &= \{\omega \in S_n \mid \omega(i) > \omega(j)\} \end{aligned}$$

das Ereignis ist, daß eine Inversion von  $i$  und  $j$  auftritt. Es folgt:

$$E[N] = \sum_{i < j} E[I_{A_{i,j}}] = \sum_{i < j} P[\{\omega \in S_n \mid \omega(i) > \omega(j)\}] = \sum_{i < j} \frac{1}{2} = \frac{1}{2} \binom{n}{2} = \frac{n(n-1)}{4}.$$

**ANWENDUNG:** Beim Sortieralgorithmus »Insertion Sort« wird der Wert  $\omega(i)$  einer Liste  $\{\omega(1), \omega(2), \dots, \omega(n)\}$  beim Einfügen von  $\omega(j)$  genau dann verschoben, wenn  $\omega(j) < \omega(i)$  gilt. Ist die Anfangsanordnung eine Zufallspermutation der korrekten Anordnung, dann ist die mittlere Anzahl der Verschiebungen, die der Algorithmus vornimmt, also gleich  $\frac{n(n-1)}{4}$ .

**Satz 1.9** (Einschluss-/Ausschlussprinzip). Für  $n \in \mathbb{N}$  und Ereignisse  $A_1, \dots, A_n \in \mathcal{A}$  gilt:

$$P[A_1 \cup A_2 \cup \dots \cup A_n] = \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} P[A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}].$$

*Beweis.* Wir betrachten zunächst das Gegenereignis, und drücken die Wahrscheinlichkeiten als Erwartungswerte von Indikatorfunktionen aus. Unter Ausnutzung der Linearität des Erwartungswerts erhalten wir:

$$\begin{aligned} P[(A_1 \cup \dots \cup A_n)^C] &= P[A_1^C \cap \dots \cap A_n^C] = E[I_{A_1^C \cap \dots \cap A_n^C}] \\ &= E\left[\prod_{i=1}^n I_{A_i^C}\right] = E\left[\prod_{i=1}^n (1 - I_{A_i})\right] \\ &= \sum_{k=0}^n (-1)^k \sum_{1 \leq i_1 < \dots < i_k \leq n} E[I_{A_{i_1}} \cdot \dots \cdot I_{A_{i_k}}] \\ &= \sum_{k=0}^n (-1)^k \sum_{1 \leq i_1 < \dots < i_k \leq n} E[I_{A_{i_1} \cap \dots \cap A_{i_k}}] \\ &= \sum_{k=0}^n (-1)^k \sum_{1 \leq i_1 < \dots < i_k \leq n} P[A_{i_1} \cap \dots \cap A_{i_k}]. \end{aligned}$$

Es folgt:

$$\begin{aligned} P[A_1 \cup \dots \cup A_n] &= 1 - P[(A_1 \cup \dots \cup A_n)^C] \\ &= \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} P[A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}]. \end{aligned}$$

□

# Kapitel 2

## Bedingte Wahrscheinlichkeiten und Unabhängigkeit

### 2.1 Bedingte Wahrscheinlichkeiten

Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum und  $A, B \in \mathcal{A}$  Ereignisse. Was ist die Wahrscheinlichkeit dafür, daß  $A$  eintritt, wenn wir schon wissen, daß  $B$  eintritt?

Relevante Fälle:  $\omega \in B$

Davon günstige Fälle:  $\omega \in A \cap B$

**Definition.** Sei  $P[B] \neq 0$ . Dann heißt

$$P[A|B] := \frac{P[A \cap B]}{P[B]}$$

die *bedingte Wahrscheinlichkeit von  $A$  gegeben  $B$* .

**Bemerkung.** a)  $P[\bullet|B]: \mathcal{A} \rightarrow [0, 1]$  ist eine Wahrscheinlichkeitsverteilung auf  $(\Omega, \mathcal{A})$ , die *bedingte Verteilung gegeben  $B$* . Der Erwartungswert

$$E[X|B] = \sum_{a \in S} a \cdot P[X = a|B]$$

einer diskreten Zufallsvariable  $X: \Omega \rightarrow S$  bzgl. der bedingten Verteilung heißt *bedingte Erwartung von  $X$  gegeben  $B$* .

b) Ist  $P$  die Gleichverteilung auf einer endlichen Menge  $\Omega$ , dann gilt:

$$P[A|B] = \frac{|A \cap B|/|\Omega|}{|B|/|\Omega|} = \frac{|A \cap B|}{|B|} \quad \text{für alle } A, B \subseteq \Omega.$$

**Beispiel.** a) Wir betrachten eine Familie mit 2 Kindern, und stellen die Frage nach dem Geschlecht der Kinder. Sei daher

$$\Omega = \{JJ, JM, MJ, MM\}.$$

Angenommen, alle Fälle wären gleichwahrscheinlich. Dann gilt:

$$P[\text{»beide Mädchen«} \mid \text{»eines Mädchen«}] = \frac{|\{MM\}|}{|\{MM, JM, MJ\}|} = \frac{1}{3},$$

$$P[\text{»beide Mädchen«} \mid \text{»das erste ist Mädchen«}] = \frac{|\{MM\}|}{|\{MM, MJ\}|} = \frac{1}{2}.$$

In Wirklichkeit sind die Kombinationen  $JJ$  und  $MM$  wahrscheinlicher.

b) Bei 20 fairen Münzwürfen fällt 15-mal »Zahl«. Wie groß ist die Wahrscheinlichkeit, daß die ersten 5 Würfe »Zahl« ergeben haben? Sei

$$\Omega = \{\omega = (x_1, \dots, x_{20}) \mid x_i \in \{0, 1\}\}, \quad \text{und} \\ X_i(\omega) = x_i, \quad \text{der Ausgang des } i\text{-ten Wurfs.}$$

Es gilt:

$$P[X_1 = \dots = X_5 = 1 \mid \sum_{i=1}^{20} X_i = 15] = \frac{P[X_1 = \dots = X_5 = 1 \text{ und } \sum_{i=6}^{20} X_i = 10]}{P[\sum_{i=1}^{20} X_i = 15]} \\ = \frac{2^{-5} \cdot 2^{-15} \binom{15}{10}}{2^{-20} \binom{20}{15}} = \frac{15 \cdot 14 \cdot \dots \cdot 11}{20 \cdot 19 \cdot \dots \cdot 16} \approx \frac{1}{5}.$$

Dagegen ist  $P[X_1 = \dots = X_5 = 1] = \frac{1}{32}$ .

## Berechnung von Wahrscheinlichkeiten durch Fallunterscheidung

Sei  $\Omega = \bigcup H_i$  eine disjunkte Zerlegung von  $\Omega$  in abzählbar viele Fälle (»Hypothesen«)  $H_i$ ,  $i \in I$ .

**Satz 2.1.** Für alle  $A \in \mathcal{A}$  gilt:

$$P[A] = \sum_{\substack{i \in I \\ P[H_i] \neq 0}} P[A|H_i] \cdot P[H_i]$$



*Beweis.* Es ist  $A = A \cap \left(\bigcup_{i \in I} H_i\right) = \bigcup_{i \in I} (A \cap H_i)$  eine disjunkte Vereinigung, also gilt nach  $\sigma$ -Additivität:

$$P[A] = \sum_{i \in I} P[A \cap H_i] = \sum_{i \in I} \underbrace{P[A \cap H_i]}_{=0, \text{ falls } P[H_i]=0} = \sum_{\substack{i \in I, \\ P[H_i] \neq 0}} P[A|H_i] \cdot P[H_i].$$

□

**Beispiel.** Urne 1 enthalte 2 rote und 3 schwarze Kugeln, Urne 2 enthalte 3 rote und 4 schwarze Kugeln. Wir legen eine Kugel  $K_1$  von Urne 1 in Urne 2 und ziehen eine Kugel  $K_2$  aus Urne 2. Mit welcher Wahrscheinlichkeit ist  $K_2$  rot?

$$\begin{aligned} P[K_2 \text{ rot}] &= P[K_2 \text{ rot} | K_1 \text{ rot}] \cdot P[K_1 \text{ rot}] + P[K_2 \text{ rot} | K_1 \text{ schwarz}] \cdot P[K_1 \text{ schwarz}] \\ &= \frac{4}{8} \cdot \frac{2}{5} + \frac{3}{8} \cdot \frac{3}{5} = \frac{17}{40}. \end{aligned}$$

**Beispiel (Simpson-Paradoxon).** Bewerbungen in Berkeley:

BEWERBUNGEN IN BERKELEY				
Statistik 1:	Männer	angenommen ( $A$ )	Frauen	angenommen ( $A$ )
	2083	996	1067	349
Empirische Verteilung:	$P[A M] \approx 0,48$		$P[A F] \approx 0,33$	

GENAUERE ANALYSE DURCH UNTERTEILUNG IN 4 FACHBEREICHE

Statistik 2:	Männer	angenommen ( $A$ )	Frauen	angenommen ( $A$ )
Bereich 1	825	511	108	89
Bereich 2	560	353	25	17
Bereich 3	325	110	593	219
Bereich 4	373	22	341	24

Sei  $P_M[A] := P[A|M]$  die empirische Verteilung unter Männern und  $P_F[A] := P[A|F]$  die empirische Verteilung unter Frauen, angenommen zu werden. Die Aufgliederung nach Fachbereichen ergibt folgende Zerlegung in Hypothesen:

$$P_M[A] = \sum_{i=1}^4 P_M[A|H_i] P_M[H_i], \quad P_F[A] = \sum_{i=1}^4 P_F[A|H_i] P_F[H_i].$$

Im Beispiel ist  $P_F[A|H_i] > P_M[A|H_i]$  für alle  $i$ , aber dennoch  $P_F[A] < P_M[A]$ . Die erste Statistik vermischt verschiedene Populationen und legt deshalb eventuell eine falsche Schlussfolgerung nahe.

## Bayessche Regel

Wie wahrscheinlich sind die Hypothesen  $H_i$ ? Ohne zusätzliche Information ist  $P[H_i]$  die Wahrscheinlichkeit von  $H_i$ . In der Bayesschen Statistik interpretiert man  $P[H_i]$  als unsere subjektive Einschätzung (aufgrund von vorhandenem oder nicht vorhandenem Vorwissen) über die vorliegende Situation (»a priori degree of belief«).

Angenommen, wir wissen nun zusätzlich, daß ein Ereignis  $A \in \mathcal{A}$  mit  $P[A] \neq 0$  eintritt, und wir kennen die bedingte Wahrscheinlichkeit (»likelihood«)  $P[A|H_i]$  für das Eintreten von  $A$  unter der Hypothese  $H_i$  für jedes  $i \in I$  mit  $P[H_i] \neq 0$ . Wie sieht dann unsere neue Einschätzung der Wahrscheinlichkeiten der  $H_i$  (»a posteriori degree of belief«) aus?

**Korollar** (Bayessche Regel). Für  $A \in \mathcal{A}$  mit  $P[A] \neq 0$  gilt:

$$P[H_i|A] = \frac{P[A|H_i] \cdot P[H_i]}{\sum_{\substack{j \in I \\ P[H_j] \neq 0}} P[A|H_j] \cdot P[H_j]} \quad \text{für alle } i \in I \text{ mit } P[H_i] \neq 0, \text{ d.h.}$$

$$P[H_i|A] = c \cdot P[H_i] \cdot P[A|H_i],$$

wobei  $c$  eine von  $i$  unabhängige Konstante ist.

*Beweis.* Es gilt:

$$P[H_i|A] = \frac{P[A \cap H_i]}{P[A]} = \frac{P[A|H_i] \cdot P[H_i]}{\sum_{\substack{j \in I \\ P[H_j] \neq 0}} P[A|H_j] \cdot P[H_j]}.$$

□

**Beispiel.** Von 10.000 Personen eines Alters habe einer die Krankheit  $K$ . Ein Test sei positiv (+) bei 96% der Kranken und 0,1% der Gesunden.

$$\text{A priori: } P[K] = \frac{1}{10000}, \quad P[K^C] = \frac{9999}{10000}.$$

$$\text{Likelihood: } P[+|K] = 0,96, \quad P[+|K^C] = 0,001.$$

A posteriori:

$$\begin{aligned} P[K|+] &= \frac{P[+|K] \cdot P[K]}{P[+|K] \cdot P[K] + P[+|K^C] \cdot P[K^C]} \\ &= \frac{0,96 \cdot 10^{-4}}{0,96 \cdot 10^{-4} + 10^{-3} \cdot 0,9999} \approx \frac{1}{11}. \end{aligned}$$

Daraus folgt insbesondere:  $P[K^C|+] \approx \frac{10}{11}$ , d.h. ohne zusätzliche Informationen muss man davon ausgehen, daß  $\frac{10}{11}$  der positiv getesteten Personen in Wirklichkeit gesund sind!

## 2.2 Mehrstufige diskrete Modelle

Wir betrachten ein  $n$ -stufiges Zufallsexperiment. Sind  $\Omega_1, \dots, \Omega_n$  abzählbare Stichprobenräume der Telexperimente, dann können wir

$$\Omega = \Omega_1 \times \dots \times \Omega_n = \{(\omega_1, \dots, \omega_n) \mid \omega_i \in \Omega_i\}$$

als Stichprobenraum des Gesamtexperiments auffassen und setzen  $\mathcal{A} = \mathcal{P}(\Omega)$ . Für  $\omega \in \Omega$  und  $k = 1, \dots, n$  sei

$$X_k(\omega) = \omega_k, \quad \text{der Ausgang des } k\text{-ten Telexperiments.}$$

Angenommen, wir kennen

$$P[X_1 = x_1] = p_1(x_1), \quad \text{für alle } x_1 \in \Omega_1, \quad (2.2.1)$$

die Verteilung (Massenfunktion) von  $X_1$ , sowie

$$P[X_k = x_k \mid X_1 = x_1, \dots, X_{k-1} = x_{k-1}] = p_k(x_k \mid x_1, \dots, x_{k-1}), \quad (2.2.2)$$

die bedingte Verteilung von  $X_k$  gegeben  $X_1, \dots, X_{k-1}$  für  $k = 2, \dots, n$ ,  $x_i \in \Omega_i$  mit  $P[X_1 = x_1, \dots, X_{k-1} = x_{k-1}] \neq 0$ .

Wie sieht die gesamte Wahrscheinlichkeitsverteilung  $P$  auf  $\Omega$  aus?

**Satz 2.2.** Seien  $p_1$  und  $p_k(\bullet \mid x_1, \dots, x_{k-1})$  für jedes  $k = 2, \dots, n$  und  $x_1 \in \Omega_1, \dots, x_{k-1} \in \Omega_{k-1}$  die Massenfunktion einer Wahrscheinlichkeitsverteilung auf  $\Omega_k$ . Dann existiert genau eine Wahrscheinlichkeitsverteilung  $P$  auf  $(\Omega, \mathcal{A})$  mit (2.2.1) und (2.2.2). Diese ist bestimmt durch die Massenfunktion

$$p(x_1, \dots, x_n) = p_1(x_1) p_2(x_2 \mid x_1) p_3(x_3 \mid x_1, x_2) \cdots p_n(x_n \mid x_1, \dots, x_{n-1}).$$

*Beweis.*

- EINDEUTIGKEIT:

Wir behaupten, dass für eine Verteilung  $P$  mit (2.2.1) und (2.2.2) gilt:

$$P[X_1 = x_1, \dots, X_k = x_k] = p_1(x_1) \cdot p_2(x_2 \mid x_1) \cdots p_k(x_k \mid x_1, \dots, x_{k-1}), \quad k = 1, \dots, n.$$

Der Induktionsanfang folgt aus Bedingung (2.2.1). Sei die Induktionsbehauptung für  $k - 1$  wahr, dann folgt nach Induktionsannahme und (2.2.2):

$$\begin{aligned} P[X_1 = x_1, \dots, X_k = x_k] &= P[X_1 = x_1, \dots, X_{k-1} = x_{k-1}] \\ &\quad \cdot P[X_1 = x_1, \dots, X_k = x_k \mid X_1 = x_1, \dots, X_{k-1} = x_{k-1}] \\ &= p_1(x_1) \cdot p_2(x_2 \mid x_1) \cdots p_{k-1}(x_{k-1} \mid x_1, \dots, x_{k-2}) \\ &\quad \cdot p_k(x_k \mid x_1, \dots, x_{k-1}), \end{aligned}$$

falls  $P[X_1 = x_1, \dots, X_{k-1} = x_{k-1}] \neq 0$ . Andernfalls verschwinden beide Seiten und die Behauptung folgt. Für  $k = n$  erhalten wir als Massenfunktion von  $P$ :

$$p(x_1, \dots, x_n) = P[X_1 = x_1, \dots, X_n = x_n] = p_1(x_1) \cdots p_n(x_n \mid x_1, \dots, x_{n-1}).$$

- EXISTENZ:

$p$  ist Massenfunktion einer Wahrscheinlichkeitsverteilung  $P$  auf  $\Omega_1 \times \cdots \times \Omega_n$ , denn:

$$\begin{aligned} \sum_{x_1 \in \Omega_1} \cdots \sum_{x_n \in \Omega_n} p(x_1, \dots, x_n) &= \sum_{x_1 \in \Omega_1} p_1(x_1) \sum_{x_2 \in \Omega_2} p_2(x_2 \mid x_1) \cdots \underbrace{\sum_{x_n \in \Omega_n} p_n(x_n \mid x_1, \dots, x_{n-1})}_{=1} \\ &= 1. \end{aligned}$$

Für  $P$  gilt:

$$\begin{aligned} P[X_1 = x_1, \dots, X_k = x_k] &= \sum_{x_{k+1} \in \Omega_{k+1}} \cdots \sum_{x_n \in \Omega_n} p(x_1, \dots, x_n) \\ &= p_1(x_1) p_2(x_2 \mid x_1) \cdots p_k(x_k \mid x_1, \dots, x_{k-1}), \quad k = 1, \dots, n. \end{aligned}$$

Damit folgen (2.2.1) und (2.2.2). □

**Beispiel.** Wie groß ist die Wahrscheinlichkeit, daß beim Skat jeder Spieler genau einen der vier Buben erhält? Sei

$$\begin{aligned} \Omega &= \{(\omega_1, \omega_2, \omega_3) \mid \omega_i \in \{0, 1, 2, 3, 4\}\}, \\ X_i(\omega) &= \omega_i = \text{Anzahl der Buben von Spieler } i. \end{aligned}$$

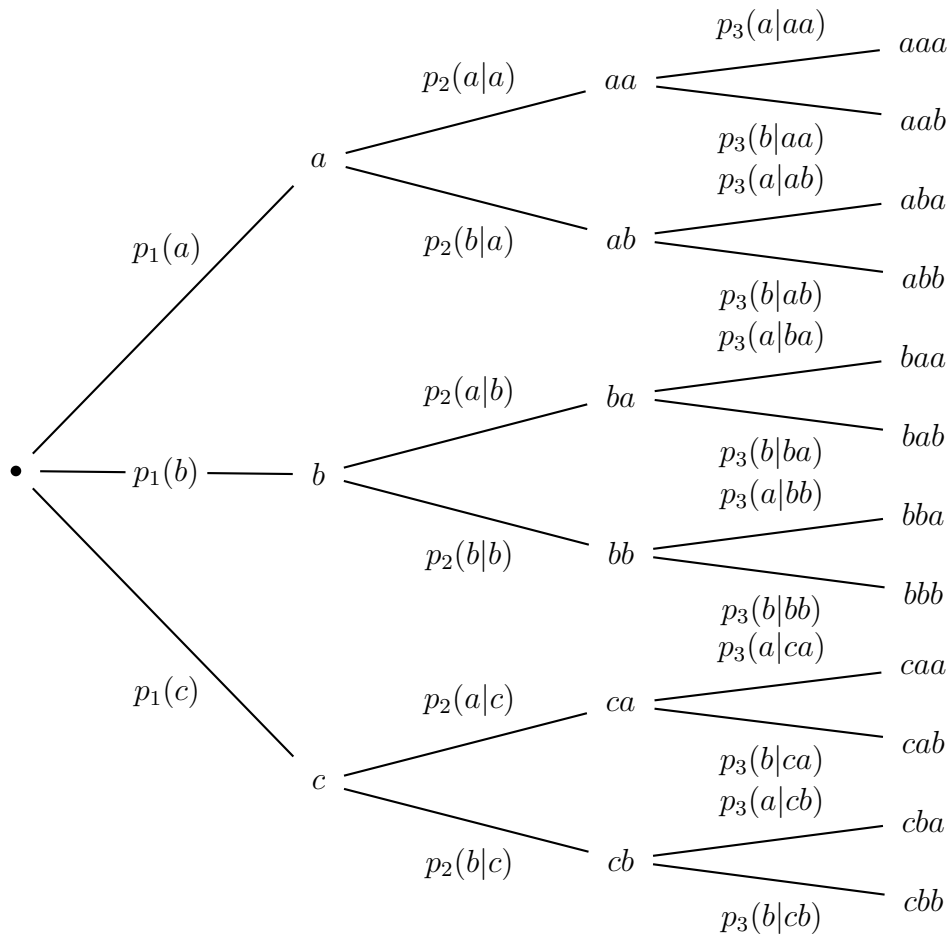


Abbildung 2.1: Baumdarstellung der Fallunterscheidungen

Es gilt:

$$p_1(x_1) = \frac{\binom{4}{x_1} \binom{28}{10-x_1}}{\binom{32}{10}}, \quad \text{hypergeometrische Verteilung,}$$

$$p_2(x_2 | x_1) = \frac{\binom{4-x_1}{x_2} \binom{18+x_1}{10-x_2}}{\binom{22}{10}}$$

$$p_3(x_3 | x_1, x_2) = \begin{cases} \frac{\binom{4-x_1-x_2}{x_3} \binom{18+x_1+x_2}{10-x_3}}{\binom{12}{10}} & \text{falls } 2 \leq x_1 + x_2 + x_3 \leq 4, \\ 0 & \text{sonst.} \end{cases}$$

Damit folgt:

$$p(1, 1, 1) = p_1(1) p_2(1 | 1) p_3(1 | 1, 1) \approx 5,56\%.$$

Im folgenden betrachten wir zwei fundamentale Klassen von mehrstufigen Modellen, Produktmodelle und Markov-Ketten.

## Produktmodelle

Angenommen, der Ausgang des  $i$ -ten Experiments hängt nicht von  $x_1, \dots, x_{i-1}$  ab. Dann sollte gelten:

$$p_i(x_i \mid x_1, \dots, x_{i-1}) = p_i(x_i)$$

mit einer von  $x_1, \dots, x_{i-1}$  unabhängigen Massenfunktion  $p_i$  einer Wahrscheinlichkeitsverteilung  $P_i$  auf  $\Omega_i$ . Die Wahrscheinlichkeitsverteilung  $P$  auf  $\Omega$  hat dann die Massenfunktion

$$p(x_1, \dots, x_n) = \prod_{i=1}^n p_i(x_i), \quad x \in \Omega. \quad (2.2.3)$$

**Definition.** Die Wahrscheinlichkeitsverteilung  $P$  auf  $\Omega = \Omega_1 \times \dots \times \Omega_n$  mit Massenfunktion (2.2.3) heißt **Produkt** von  $P_1, \dots, P_n$  und wird mit  $P_1 \otimes \dots \otimes P_n$  notiert.

**Beispiel** ( $n$ -dimensionale Bernoulliverteilung). Wir betrachten  $n$  unabhängige 0-1-Experimente mit Erfolgswahrscheinlichkeit  $p$ :

$$\Omega_1 = \dots = \Omega_n = \{0, 1\}, \quad p_i(1) = p, \quad p_i(0) = 1 - p, \quad i = 1, \dots, n.$$

Sei  $k = \sum_{i=1}^n x_i$  die Anzahl der Einsen. Dann ist

$$p(x_1, \dots, x_n) = \prod_{i=1}^n p_i(x_i) = p^k (1 - p)^{n-k}$$

die  **$n$ -dimensionale Bernoulliverteilung**.

**Bemerkung.** Sind die Mengen  $\Omega_i$ ,  $i = 1, \dots, n$  endlich, und ist  $P_i$  die Gleichverteilung auf  $\Omega_i$ , dann ist  $P_1 \otimes \dots \otimes P_n$  die Gleichverteilung auf  $\Omega_1 \times \dots \times \Omega_n$ .

Die Multiplikativität im Produktmodell gilt nicht nur für die Massenfunktion, sondern allgemeiner für die Wahrscheinlichkeiten, daß in den Telexperimenten bestimmte Ereignisse  $A_1, \dots, A_n$  eintreten:

**Satz 2.3.** Im Produktmodell gilt für beliebige Ereignisse  $A_i \subseteq \Omega_i$ ,  $i = 1, \dots, n$ :

$$\begin{aligned} P[X_1 \in A_1, \dots, X_n \in A_n] &= \prod_{i=1}^n P[X_i \in A_i] & (2.2.4) \\ &\parallel & \\ &\parallel & \\ P[A_1 \times \dots \times A_n] &= \prod_{i=1}^n P_i[A_i] \end{aligned}$$

(d.h.  $X_1, \dots, X_n$  sind unabhängige Zufallsvariablen, siehe nächsten Abschnitt).

*Beweis.* Es gilt:

$$\begin{aligned} P[X_1 \in A_1, \dots, X_n \in A_n] &= P[(X_1, \dots, X_n) \in A_1 \times \dots \times A_n] = P[A_1 \times \dots \times A_n] \\ &= \sum_{x \in A_1 \times \dots \times A_n} p(x) = \sum_{x_1 \in A_1} \dots \sum_{x_n \in A_n} \prod_{i=1}^n p_i(x_i) \\ &= \prod_{i=1}^n \sum_{x_i \in A_i} p_i(x_i) = \prod_{i=1}^n P_i[A_i]. \end{aligned}$$

Insbesondere gilt:

$$P[X_i \in A_i] = P[X_1 \in \Omega, \dots, X_{i-1} \in \Omega, X_i \in A_i, X_{i+1} \in \Omega, \dots, X_n \in \Omega] = P_i[A_i].$$

□

## Markov-Ketten

Zur Modellierung einer zufälligen zeitlichen Entwicklung mit abzählbarem Zustandsraum  $S$  betrachten wir den Stichprobenraum

$$\Omega = S^{n+1} = \{(x_0, x_1, \dots, x_n) \mid x_i \in S\}.$$

Oft ist es naheliegend anzunehmen, daß die Weiterentwicklung des Systems nur vom gegenwärtigen Zustand, aber nicht vom vorherigen Verlauf abhängt (»kein Gedächtnis«), d.h. es sollte gelten:

$$p_k(x_k \mid x_0, \dots, x_{k-1}) = \underbrace{p_k(x_{k-1}, x_k)}_{\text{»Bewegungsgesetz«}}, \quad (2.2.5)$$

wobei  $p_k : S \times S \rightarrow [0, 1]$  folgende Bedingungen erfüllt:

- i)  $p_k(x, y) \geq 0$  für alle  $x, y \in S$
- ii)  $\sum_{y \in S} p_k(x, y) = 1$  für alle  $x \in S$

d.h.  $p_k(x, \bullet)$  ist für jedes  $x \in S$  die Massenfunktion einer Wahrscheinlichkeitsverteilung auf  $S$ .

**Definition.** Eine Matrix  $p_k(x, y)$  ( $x, y \in S$ ) mit i) und ii) heißt **stochastische Matrix** (oder **stochastischer Kern**) auf  $S$ .

Im Mehrstufenmodell folgt aus Gleichung (2.2.5):

$$p(x_0, x_1, \dots, x_n) = \underbrace{p_0(x_0)}_{\text{»Startverteilung«}} p_1(x_0, x_1) p_2(x_1, x_2) \cdots p_n(x_{n-1}, x_n), \quad \text{für } x_0, \dots, x_n \in S.$$

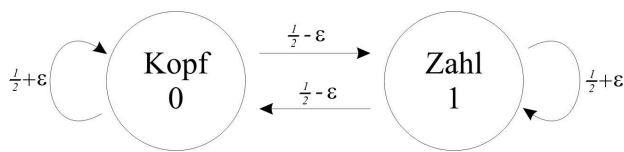
Den Fall, in dem der Übergangsmechanismus  $p_k(x, y) = p(x, y)$  unabhängig von  $k$  ist, nennt man **zeitlich homogen**.

**Beispiele.** a) PRODUKTMODELL (siehe oben):

$$p_k(x, y) = p_k(y) \text{ ist unabhängig von } x.$$

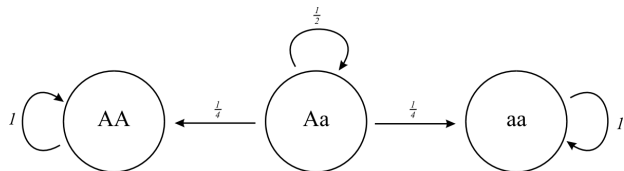
b) ABHÄNGIGE MÜNZWÜRFE:

$$S = \{0, 1\}, \quad \varepsilon \in \left[-\frac{1}{2}, \frac{1}{2}\right].$$



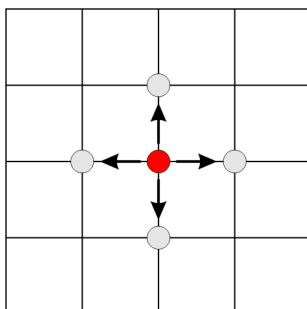
$$p = \begin{pmatrix} \frac{1}{2} + \varepsilon & \frac{1}{2} - \varepsilon \\ \frac{1}{2} - \varepsilon & \frac{1}{2} + \varepsilon \end{pmatrix}.$$

c) SELBSTBEFRUCHTUNG VON PFLANZEN:



$$p = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & 0 & 1 \end{pmatrix}$$

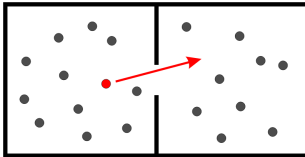
d) RANDOM WALK AUF  $S = \mathbb{Z}^d$ , ( $d \in \mathbb{N}$ ):





$$p(x, y) = \begin{cases} \frac{1}{2d} & \text{falls } |x - y| = 1, \\ 0 & \text{sonst.} \end{cases}$$

- e) URNENMODELL VON P. UND T. EHRENFEST (Austausch von Gasmolekülen in zwei Behältern):



Es seien  $N$  Kugeln auf zwei Urnen verteilt. Zu jedem Zeitpunkt  $t \in \mathbb{N}$  wechselt eine zufällig ausgewählte Kugel die Urne.

MAKROSKOPISCHES MODELL:

$$S = \{0, 1, 2, \dots, n\}.$$

$x \in S$  beschreibt die Anzahl Kugeln in der ersten Urne.

$$p(x, y) = \begin{cases} \frac{x}{n} & \text{falls } y = x - 1, \\ \frac{n-x}{n} & \text{falls } y = x + 1, \\ 0 & \text{sonst.} \end{cases}$$

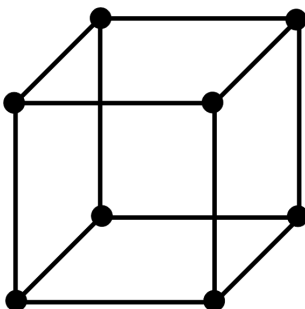
MIKROSKOPISCHES MODELL:

$$S = \{0, 1\}^n = \{(\sigma_1, \dots, \sigma_n) \mid \sigma_i \in \{0, 1\}\}.$$

Es ist  $\sigma_i = 1$  genau dann, wenn sich die  $i$ -te Kugel in Urne 1 befindet.

$$p(\sigma, \tilde{\sigma}) = \begin{cases} \frac{1}{N} & \text{falls } \sum_{i=1}^n |\sigma_i - \tilde{\sigma}_i| = 1, \\ 0 & \text{sonst.} \end{cases}$$

Die resultierende Markov-Kette ist ein Random Walk auf dem Hyperwürfel  $\{0, 1\}^n$ , d.h. sie springt in jedem Schritt von einer Ecke des Hyperwürfels zu einer zufällig ausgewählten benachbarten Ecke.



**Berechnung von Wahrscheinlichkeiten**

**Satz 2.4** (Markov-Eigenschaft). Für alle  $0 \leq k < l \leq n$  und  $x_0, \dots, x_l \in S$  mit  $P[X_0 = x_0, \dots, X_k = x_k] \neq 0$  gilt:

$$\begin{aligned} P[X_l = x_l \mid X_0 = x_0, \dots, X_k = x_k] &= P[X_l = x_l \mid X_k = x_k] \\ &= (p_{k+1} p_{k+2} \cdots p_l)(x_k, x_l), \end{aligned}$$

wobei

$$(pq)(x, y) := \sum_{z \in S} p(x, z) q(z, y)$$

das Produkt der Matrizen  $p$  und  $q$  ist.

**Bemerkung.** a) MARKOV-EIGENSCHAFT:

Die Weiterentwicklung hängt jeweils nur vom gegenwärtigen Zustand  $x_k$  ab, und nicht vom vorherigen Verlauf  $x_0, x_1, \dots, x_{k-1}$ .

b)  $n$ -SCHRITT-ÜBERGANGSWAHRSCHEINLICHKEITEN:

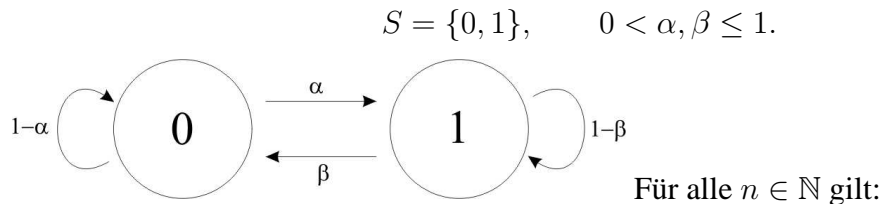
$$\begin{aligned} P[X_n = y \mid X_0 = x] &= (p_1 p_2 \cdots p_n)(x, y) \\ &= p^n(x, y) \quad \text{falls zeitlich homogen, d.h. } p_i \equiv p. \end{aligned}$$

*Beweis.*

$$\begin{aligned} P[X_l = x_l \mid X_0 = x_0, \dots, X_k = x_k] &= \frac{P[X_0 = x_0, \dots, X_k = x_k, X_l = x_l]}{P[X_0 = x_0, \dots, X_k = x_k]} \\ &= \frac{\sum_{x_{k+1}, \dots, x_{l-1}} p_0(x_0) p_1(x_0, x_1) \cdots p_l(x_{l-1}, x_l)}{p_0(x_0) p_1(x_0, x_1) \cdots p_k(x_{k-1}, x_k)} \\ &= \sum_{x_{k+1}} \cdots \sum_{x_{l-1}} p_{k+1}(x_k, x_{k+1}) p_{k+2}(x_{k+1}, x_{k+2}) \cdots p_l(x_{l-1}, x_l) \\ &= (p_{k+1} p_{k+2} \cdots p_l)(x_k, x_l). \\ P[X_l = x_l \mid X_k = x_k] &= \frac{P[X_k = x_k, X_l = x_l]}{P[X_k = x_k]} \\ &= \frac{\sum_{x_1, \dots, x_{k-1}} \sum_{x_{k+1}, \dots, x_{l-1}} p_0(x_0) p_1(x_0, x_1) \cdots p_l(x_{l-1}, x_l)}{\sum_{x_1, \dots, x_{k-1}} p_0(x_0) p_1(x_0, x_1) \cdots p_k(x_{k-1}, x_k)} \\ &= (p_{k+1} p_{k+2} \cdots p_l)(x_k, x_l). \end{aligned}$$

□

**Beispiel.**



$$\begin{aligned} p^n(0, 0) &= p^{n-1}(0, 0) \cdot p(0, 0) + p^{n-1}(0, 1) \cdot p(1, 0) \\ &= p^{n-1}(0, 0) \cdot (1 - \alpha) + (1 - p^{n-1}(0, 0)) \cdot \beta \\ &= (1 - \alpha - \beta) \cdot p^{n-1}(0, 0) + \beta. \end{aligned}$$

Daraus folgt mit Induktion:

$$\begin{aligned} p^n(0, 0) &= \frac{\beta}{\alpha + \beta} + \frac{\alpha}{\alpha + \beta} (1 - \alpha - \beta)^n, \quad \text{und} \\ p^n(0, 1) &= 1 - p^n(0, 0). \end{aligned}$$

Analoge Formeln erhält man für  $p^n(1, 0)$  und  $p^n(1, 1)$  durch Vertauschung von  $\alpha$  und  $\beta$ . Für die  $n$ -Schritt-Übergangsmatrix ergibt sich:

$$p^n = \underbrace{\begin{pmatrix} \frac{\beta}{\alpha + \beta} & \frac{\alpha}{\alpha + \beta} \\ \frac{\beta}{\alpha + \beta} & \frac{\alpha}{\alpha + \beta} \end{pmatrix}}_{\text{Gleiche Zeilen}} + \underbrace{(1 - \alpha - \beta)^n \begin{pmatrix} \frac{\alpha}{\alpha + \beta} & \frac{-\alpha}{\alpha + \beta} \\ \frac{-\beta}{\alpha + \beta} & \frac{\beta}{\alpha + \beta} \end{pmatrix}}_{\rightarrow 0 \text{ exponentiell schnell, falls } \alpha < 1 \text{ oder } \beta < 1}.$$

Insbesondere gilt  $p^n(0, \cdot) \approx p^n(1, \cdot)$  für große  $n \in \mathbb{N}$ . Die Kette »vergisst« also ihren Startwert exponentiell schnell (»Exponentieller Gedächtnisverlust«)!

## 2.3 Unabhängigkeit von Ereignissen

Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum. Hängen zwei Ereignisse  $A, B \in \mathcal{A}$  *nicht voneinander ab*, dann sollte gelten:

$$P[A|B] = P[A], \quad \text{falls } P[B] \neq 0,$$

sowie

$$\underbrace{P[B|A]}_{\frac{P[B \cap A]}{P[A]}} = P[B], \quad \text{falls } P[A] \neq 0,$$

also insgesamt

$$P[A \cap B] = P[A] \cdot P[B]. \tag{2.3.1}$$

**Definition.** i) Zwei Ereignisse  $A, B \in \mathcal{A}$  heißen **unabhängig** (bzgl.  $P$ ), falls (2.3.1) gilt.

ii) Eine beliebige Kollektion  $A_i, i \in I$ , von Ereignissen heißt **unabhängig** (bzgl.  $P$ ), falls

$$P[A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_n}] = \prod_{k=1}^n P[A_{i_k}]$$

für alle  $n \in \mathbb{N}$  und alle paarweise verschiedenen  $i_1, \dots, i_n \in I$  gilt.

**Beispiele.** a) Falls  $P[A] \in \{0, 1\}$  gilt, ist  $A$  unabhängig von  $B$  für alle  $B \in \mathcal{A}$ .

b) Wir betrachten das Modell für ZWEI FAIRE MÜNZWÜRFE, also  $\Omega = \{0, 1\}^2$  und  $P$  sei die Gleichverteilung. Die Ereignisse

$$A_1 = \{(1, 0), (1, 1)\}, \quad \text{»erster Wurf Zahl«},$$

$$A_2 = \{(0, 1), (1, 1)\}, \quad \text{»zweiter Wurf Zahl«},$$

$$A_3 = \{(0, 0), (1, 1)\}, \quad \text{»beide Würfe gleich«},$$

sind *paarweise unabhängig*, denn es gilt:

$$P[A_i \cap A_j] = \frac{1}{4} = P[A_i] \cdot P[A_j] \quad \text{für alle } i \neq j.$$

Allerdings ist die Kollektion  $A_1, A_2, A_3$  *nicht unabhängig*, denn es gilt

$$P[A_1 \cap A_2 \cap A_3] = \frac{1}{4} \neq P[A_1] \cdot P[A_2] \cdot P[A_3].$$

**Lemma 2.5.** Seien die Ereignisse  $A_1, \dots, A_n \in \mathcal{A}$  unabhängig,  $B_j = A_j$  oder  $B_j = A_j^C$  für alle  $j = 1, \dots, n$ . Dann sind die Ereignisse  $B_1, \dots, B_n$  unabhängig.

*Beweis.* Sei ohne Beschränkung der Allgemeinheit:

$$B_1 = A_1, \quad \dots, \quad B_k = A_k, \quad B_{k+1} = A_{k+1}^C, \quad \dots, \quad B_n = A_n^C$$

. Dann gilt unter Verwendung der Linearität des Erwartungswerts und der Unabhängigkeit von  $A_1, \dots, A_n$ :

$$\begin{aligned} P[B_1 \cap \dots \cap B_n] &= P[A_1 \cap \dots \cap A_k \cap A_{k+1}^C \cap \dots \cap A_n^C] \\ &= E[I_{A_1} \cdots I_{A_k} \cdot (1 - I_{A_{k+1}}) \cdots (1 - I_{A_n})] \\ &= E[I_{A_1} \cdots I_{A_k} \cdot \sum_{J \subseteq \{k+1, \dots, n\}} (-1)^{|J|} \prod_{j \in J} I_{A_j}] \\ &= \sum_{J \subseteq \{k+1, \dots, n\}} (-1)^{|J|} P[A_1 \cap \dots \cap A_k \cap \bigcap_{j \in J} A_j] \\ &= \sum_{J \subseteq \{k+1, \dots, n\}} (-1)^{|J|} P[A_1] \cdots P[A_k] \cdot \prod_{j \in J} P[A_j] \\ &= P[A_1] \cdots P[A_k] \cdot (1 - P[A_{k+1}]) \cdots (1 - P[A_n]) = P[B_1] \cdots P[B_n]. \end{aligned}$$

□

### Verteilungen für unabhängige Ereignisse

Seien  $A_1, A_2, \dots \in \mathcal{A}$  unabhängige Ereignisse (bzgl.  $P$ ) mit  $P[A_i] = p \in [0, 1]$ . Die Existenz von unendlich vielen unabhängigen Ereignissen auf einem geeigneten Wahrscheinlichkeitsraum setzen wir hier voraus – ein Beweis wird erst in der Vorlesung »Einführung in die Wahrscheinlichkeitstheorie« gegeben.

Die Wartezeit auf das erste Eintreten eines der Ereignisse ist

$$T(\omega) = \inf\{n \in \mathbb{N} \mid \omega \in A_n\}, \quad \text{wobei } \min \emptyset := \infty.$$

Mit Lemma 2.5 folgt:

$$\begin{aligned} P[T = n] &= P[A_1^C \cap A_2^C \cap \dots \cap A_{n-1}^C \cap A_n] \\ &= P[A_n] \cdot \prod_{i=1}^{n-1} P[A_i^C] \\ &= p \cdot (1 - p)^{n-1}. \end{aligned}$$

**Definition.** Sei  $p \in [0, 1]$ . Die Wahrscheinlichkeitsverteilung auf  $\mathbb{N}$  mit Massenfunktion

$$p(n) = p \cdot (1 - p)^{n-1}$$

heißt *geometrische Verteilung zum Parameter  $p$* .

**Bemerkung.** a) Für  $p \neq 0$  gilt:

$$\sum_{n=1}^{\infty} p \cdot (1 - p)^{n-1} = 1,$$

d.h. die geometrische Verteilung ist eine Wahrscheinlichkeitsverteilung auf den natürlichen Zahlen, und

$$P[T = \infty] = 0.$$

b) Allgemein gilt:

$$P[T > n] = P[A_1^C \cap \dots \cap A_n^C] = (1 - p)^n.$$

c) Es gilt:

$$E[T] = \sum_{n=0}^{\infty} P[T > n] = \frac{1}{1 - (1 - p)} = \frac{1}{p},$$

(siehe Übung).

Die Anzahl der Ereignisse unter  $A_1, \dots, A_n$ , die eintreten, ist

$$S_n(\omega) = |\{1 \leq i \leq n \mid \omega \in A_i\}| = \sum_{i=1}^n I_{A_i}(\omega).$$

Es gilt:

$$\begin{aligned} P[S_n = k] &= \sum_{\substack{I \subseteq \{1, \dots, n\} \\ |I|=k}} P\left[\bigcap_{i \in I} A_i \cap \bigcap_{i \in \{1, \dots, n\} \setminus I} A_i^C\right] \\ &= \sum_{\substack{I \subseteq \{1, \dots, n\} \\ |I|=k}} \prod_{i \in I} P[A_i] \cdot \prod_{i \in I^C} P[A_i^C] \\ &= \sum_{\substack{I \subseteq \{1, \dots, n\} \\ |I|=k}} \prod_{i \in I} p \cdot \prod_{i \in I^C} (1-p) \\ &= \sum_{\substack{I \subseteq \{1, \dots, n\} \\ |I|=k}} p^{|I|} \cdot (1-p)^{|I^C|} \\ &= \binom{n}{k} p^k (1-p)^{n-k}, \end{aligned}$$

d.h.  $S_n$  ist die *Binomialverteilung* mit Parametern  $n$  und  $p$ .

**Satz 2.6** (»Law of Averages«, Bernstein-Ungleichung). Für alle  $\varepsilon > 0$  und  $n \in \mathbb{N}$  gilt:

$$P\left[\frac{S_n}{n} \geq p + \varepsilon\right] \leq e^{-2\varepsilon^2 n}, \quad \text{und}$$

$$P\left[\frac{S_n}{n} \leq p - \varepsilon\right] \leq e^{-2\varepsilon^2 n}.$$

Insbesondere gilt:

$$P\left[\left|\frac{S_n}{n} - p\right| > \varepsilon\right] \leq 2e^{-2\varepsilon^2 n},$$

d.h. die Wahrscheinlichkeit für eine Abweichung des Mittelwerts  $\frac{S_n}{n}$  vom Erwartungswert  $p$  um mehr als  $\varepsilon$  fällt exponentiell in  $n$ .

**Bemerkung.** a) Satz 2.6 ist eine erste Version des »Gesetzes der großen Zahlen«.

b) Der Satz liefert eine nachträgliche Rechtfertigung der frequentistischen Interpretation der Wahrscheinlichkeit als asymptotische relative Häufigkeit.

c) Anwendung auf Schätzen von  $p$ :

$$p \approx \frac{S_n}{n} = \text{relative Häufigkeit des Ereignisses bei } n \text{ unabhängigen Stichproben.}$$

d) Anwendung auf näherungsweise Monte Carlo-Berechnung von  $p$ :

Simuliere  $n$  unabhängige Stichproben,  $p \sim$  relative Häufigkeit.

*Beweis.* Sei  $q := 1 - p$ ,  $S_n \sim \text{Bin}(n, p)$ . Dann gilt für  $\lambda > 0$ :

$$\begin{aligned} P[S_n \geq n(p + \varepsilon)] &= \sum_{k \geq np + n\varepsilon} \binom{n}{k} p^k q^{n-k} \\ &\leq \sum_{k \geq np + n\varepsilon} \binom{n}{k} e^{\lambda k} p^k q^{n-k} e^{-\lambda(np + n\varepsilon)} \\ &\leq \sum_{k=0}^n \binom{n}{k} (pe^\lambda)^k q^{n-k} e^{-\lambda np} e^{-\lambda n\varepsilon} \\ &= (pe^\lambda + q)^n e^{-\lambda np} e^{-\lambda n\varepsilon} \leq (pe^{\lambda q} + qe^{-\lambda p})^n e^{-\lambda n\varepsilon}. \end{aligned}$$

Wir behaupten:

$$pe^{\lambda q} + qe^{-\lambda p} \leq e^{\frac{\lambda^2}{8}}.$$

Damit folgt:

$$P[S_n \geq n(p + \varepsilon)] \leq e^{n(\frac{\lambda^2}{8} - \lambda\varepsilon)}.$$

Der Exponent ist minimal für  $\lambda = 4\varepsilon$ . Für diese Wahl von  $\lambda$  folgt schließlich

$$P[S_n \geq n(p + \varepsilon)] \leq e^{-2n\varepsilon^2}.$$

*Beweis der Behauptung:*

$$f(\lambda) := \log(pe^{\lambda q} + qe^{-\lambda p}) = \log(e^{-\lambda p}(pe^\lambda + q)) = -\lambda p + \log(pe^\lambda + q).$$

Zu zeigen ist nun

$$f(\lambda) \leq \frac{\lambda^2}{8} \quad \text{für alle } \lambda \geq 0.$$

Es gilt:

$$\begin{aligned} f(0) &= 0, \\ f'(\lambda) &= -p + \frac{pe^\lambda}{pe^\lambda + q} = -p + \frac{p}{p + qe^{-\lambda}}, \quad f'(0) = 0, \\ f''(\lambda) &= \frac{pqe^{-\lambda}}{(p + qe^{-\lambda})^2} \leq \frac{1}{4}. \end{aligned}$$

Die letzte Ungleichung folgt aus::

$$(a + b)^2 = a^2 + b^2 + 2ab \geq 4ab$$

Damit folgt

$$\begin{aligned} f(\lambda) &= f(0) + \int_0^\lambda f'(x) dx \\ &= \int_0^\lambda \int_0^x f''(y) dy dx \leq \int_0^\lambda \frac{x}{4} dx \leq \frac{\lambda^2}{8} \quad \text{für alle } \lambda \geq 0. \end{aligned}$$

□

**Beispiel.** Im letzten Satz wurde gezeigt:

$$S_n = \sum_{i=1}^n I_{A_i}, \quad A_i \text{ unabhängig mit } P[A_i] = p \implies P\left[\left|\frac{S_n}{n} - p\right| \geq \varepsilon\right] \longrightarrow 0 \quad \text{für } n \rightarrow \infty.$$

Zur Demonstration simulieren wir den Verlauf von  $S_n$  und  $\frac{S_n}{n}$  mehrfach ( $m$ -mal):

VERLAUF VON  $S_n$

$m = 30; n_{\max} = 1000; p = 0.7;$

(\* Wir erzeugen  $m \times n_{\max}$  Bernoulli-Stichproben mit Wahrscheinlichkeit  $p$ \*)

$x = \text{RandomChoice}[\{1 - p, p\} \rightarrow \{0, 1\}, \{n_{\max}, m\}]; s = \text{Accumulate}[x];$

(\*Das Feld  $s$  enthält  $m$  Verläufe von  $s_n = x_1 + \dots + x_n, n = 1, \dots, n_{\max}$ \*)

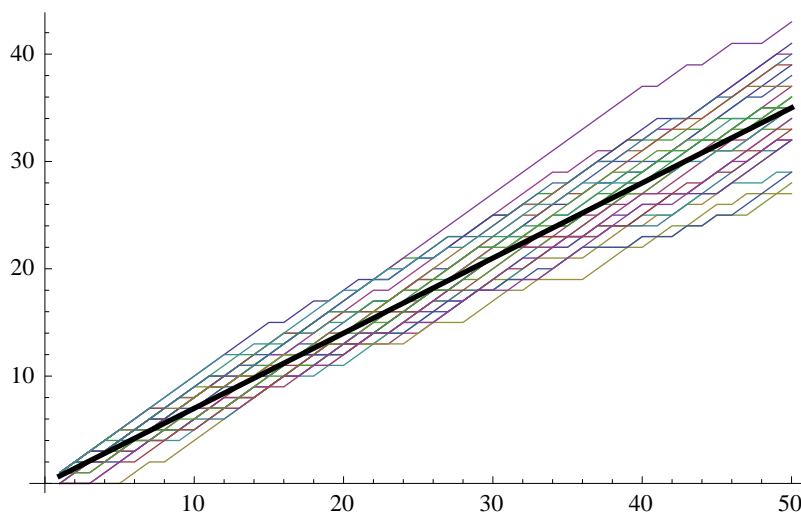
Manipulate[Show[

ListLinePlot[Transpose[s[[1;;n]]],

ListLinePlot[p \* Range[n], PlotStyle  $\rightarrow$  {Black, Thick}]

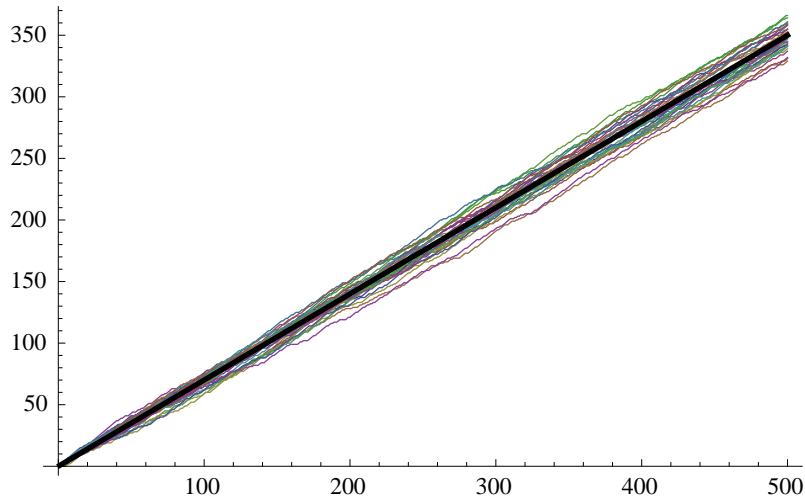
, {{n, 50}, 1, n\_{\max}, 1}]

•  $n = 50:$





- $n = 500$ :



VERLAUF VON  $\frac{S_n}{n}$

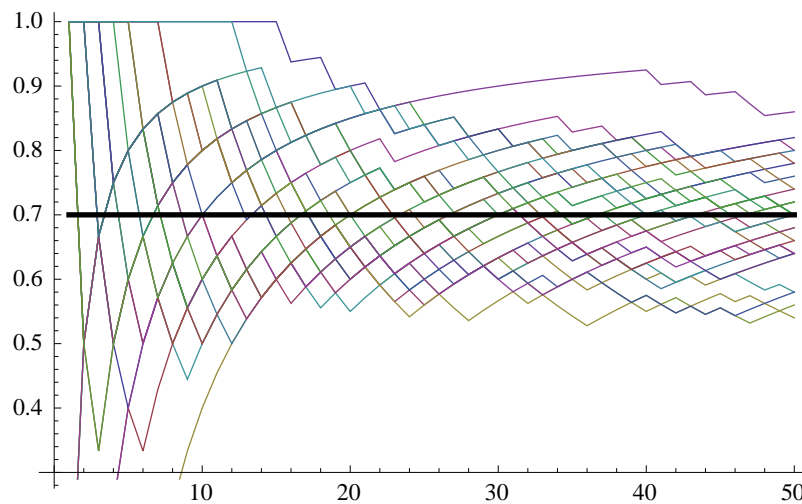
```
mean = s/Range[nmax];
```

(\*Das Feld mean enthält  $m$  Verläufe der Werte von  $\frac{S_n}{n}$ \*)

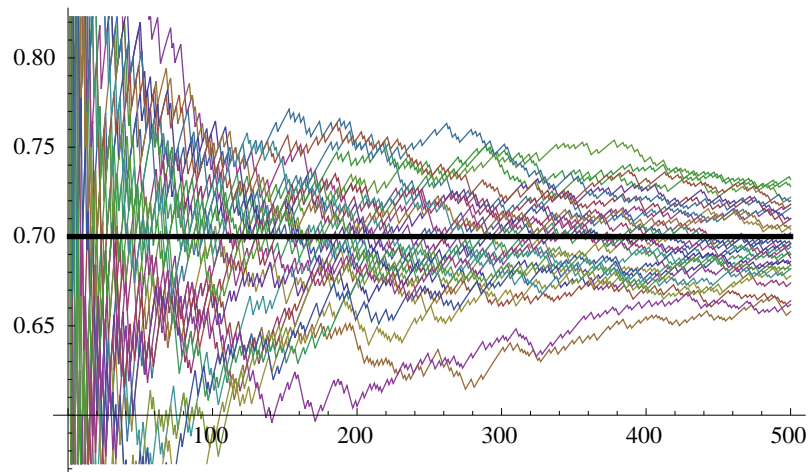
```
Manipulate[Show[
```

```
ListLinePlot[Transpose[mean[[1;;n]]],
```

```
ListLinePlot[ConstantArray[p, n], PlotStyle -> {Black, Thick}], {{n, 50}, 1, nmax, 1}]
```



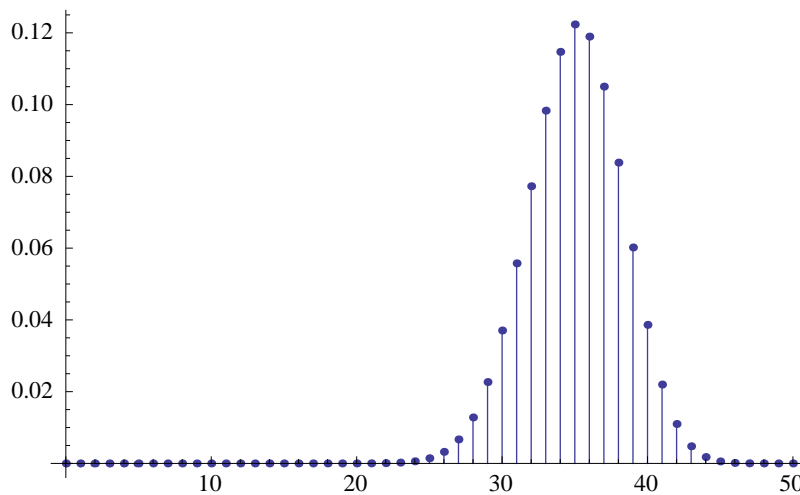
- $n = 50$ :



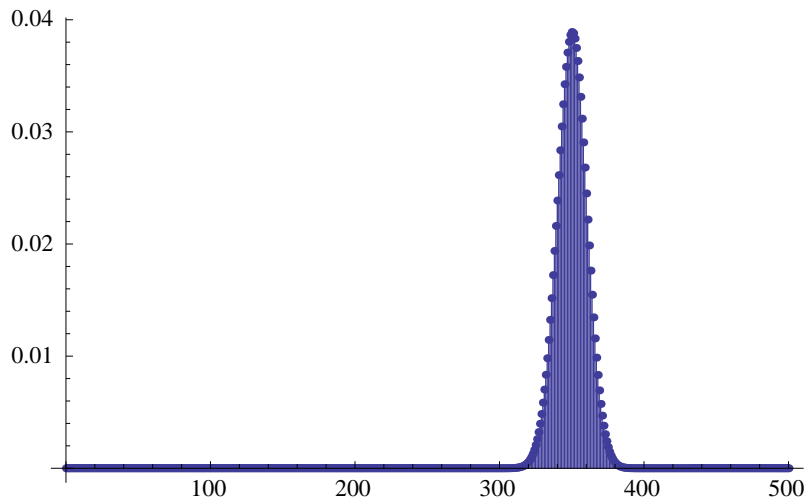
VERTEILUNG VON  $S_n$

```
Manipulate[ListPlot[Table[{k, PDF[BinomialDistribution[n, p], k]}, {k, 0, n}],
PlotRange -> All, Filling->Axis], {{n, 50}, 1, nmax, 1}]
```

•  $n = 50$ :



- $n = 500$ :



## 2.4 Unabhängige Zufallsvariablen und Random Walk

Seien  $X_i : \Omega \rightarrow S_i$ ,  $i = 1, \dots, n$ , diskrete Zufallsvariablen auf dem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, P)$ . Dann ist  $(X_1, \dots, X_n)$  eine Zufallsvariable mit Werten im Produktraum  $S_1 \times \dots \times S_n$ .

**Definition.** Die Verteilung  $\mu_{X_1, \dots, X_n}$  des Zufallsvektors  $(X_1, \dots, X_n)$  heißt **gemeinsame Verteilung** der Zufallsvariablen  $X_1, \dots, X_n$ . Die Massenfunktion der gemeinsamen Verteilung lautet

$$p_{X_1, \dots, X_n}(a_1, \dots, a_n) = P[X_1 = a_1, \dots, X_n = a_n].$$

**Definition.** Die diskreten Zufallsvariablen  $X_1, \dots, X_n$  heißen **unabhängig**, falls gilt:

$$P[X_1 = a_1, \dots, X_n = a_n] = \prod_{i=1}^n P[X_i = a_i] \quad \text{für alle } a_i \in S_i, \quad i = 1, \dots, n.$$

Die gemeinsame Verteilung enthält Informationen über den Zusammenhang zwischen den Zufallsgrößen  $X_i$ .

**Satz 2.7.** Die folgenden Aussagen sind äquivalent:

- (i)  $X_1, \dots, X_n$  sind unabhängig.
- (ii)  $p_{X_1, \dots, X_n}(a_1, \dots, a_n) = \prod_{i=1}^n p_{X_i}(a_i)$ .
- (iii)  $\mu_{X_1, \dots, X_n} = \otimes_{i=1}^n \mu_{X_i}$ .
- (iv) Die Ereignisse  $\{X_1 \in A_1\}, \dots, \{X_n \in A_n\}$  sind unabhängig für alle  $A_i \subseteq S_i$ ,  $i = 1, \dots, n$ .

(v) Die Ereignisse  $\{X_1 = a_1\}, \dots, \{X_n = a_n\}$  sind unabhängig für alle  $a_i \in S_i, i = 1, \dots, n$ .

*Beweis.*

• (i)  $\Leftrightarrow$  (ii) nach Definition von  $p_{X_1, \dots, X_n}$ .

• (ii)  $\Leftrightarrow$  (iii) nach Definition von  $\bigotimes_{i=1}^n \mu_{X_i}$ .

• (iii)  $\Rightarrow$  (iv):

Seien  $1 \leq i_1 < i_2 < \dots < i_m \leq n$  und  $A_{i_k} \subseteq S_{i_k}, (k = 1, \dots, m)$ . Wir setzen  $A_i := \Omega$  für  $i \notin \{i_1, \dots, i_m\}$ . Mit (iii) folgt dann nach Satz 2.2:

$$\begin{aligned} P[X_{i_1} \in A_{i_1}, \dots, X_{i_m} \in A_{i_m}] &= P[X_1 \in A_1, \dots, X_n \in A_n] \\ &= P[(X_1, \dots, X_n) \in A_1 \times \dots \times A_n] \\ &= \mu_{X_1, \dots, X_n}(A_1 \times \dots \times A_n) \\ &= \prod_{i=1}^n \mu_{X_i}(A_i) = \prod_{i=1}^n P[X_i \in A_i] \\ &= \prod_{i=1}^m P[X_{i_k} \in A_{i_k}]. \end{aligned}$$

• (iv)  $\Rightarrow$  (v)  $\Rightarrow$  (i) ist klar.

□

**Definition.** Eine beliebige Kollektion  $X_i: \Omega \rightarrow S_i, i \in I$ , von diskreten Zufallsvariablen heißt **unabhängig**, falls die Ereignisse  $\{X_i = a_i\}, i \in I$ , für alle  $a_i \in S_i$  unabhängig sind.

## Der Random Walk auf $\mathbb{Z}$

Seien  $X_1, X_2, \dots$  unabhängige identisch verteilte (»i.i.d.« – independent and identically distributed) Zufallsvariablen auf dem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, P)$  mit

$$P[X_i = +1] = p, \quad P[X_i = -1] = 1 - p, \quad p \in (0, 1).$$

Die Existenz von unendlich vielen unabhängigen identisch verteilten Zufallsvariablen auf einem geeigneten Wahrscheinlichkeitsraum (unendliches Produktmodell) wird in der Vorlesung »Einführung in die Wahrscheinlichkeitstheorie« gezeigt. Sei  $a \in \mathbb{Z}$  ein fester Startwert. Wir betrachten die durch

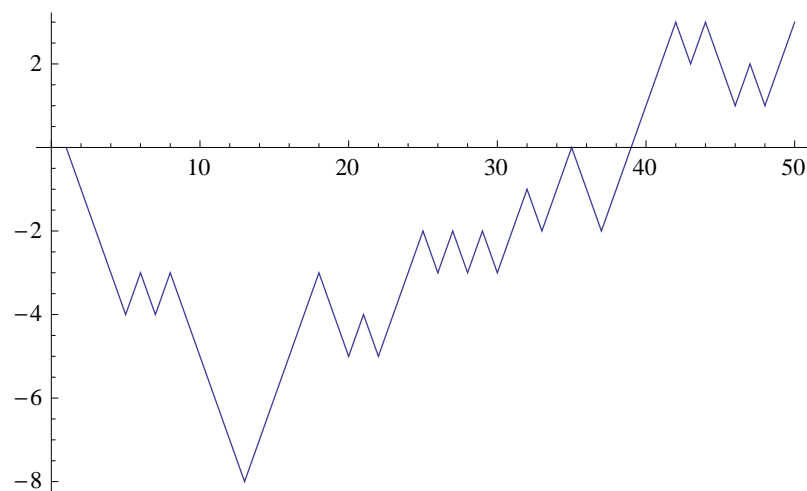
$$\begin{aligned} S_0 &= a, \\ S_{n+1} &= S_n + X_{n+1}, \end{aligned}$$

definierte zufällige Bewegung (»Irrfahrt« oder »Random Walk«) auf  $\mathbb{Z}$ . Als Position zur Zeit  $n$  ergibt sich:

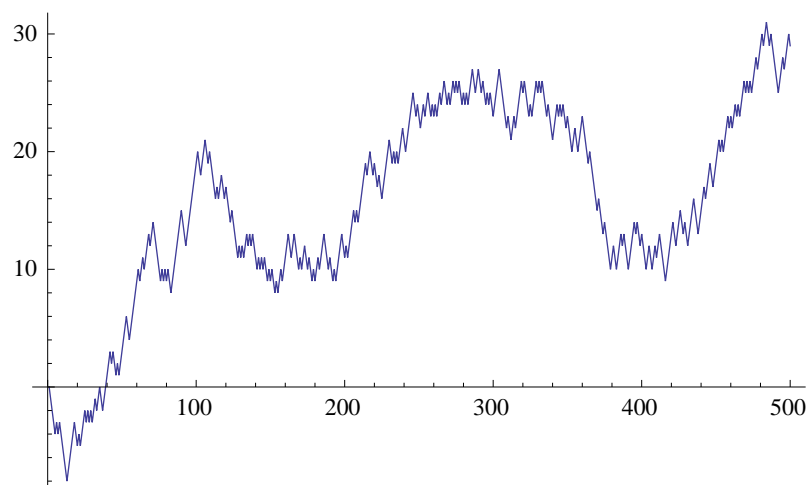
$$S_n = a + X_1 + X_2 + \cdots + X_n.$$

Irrfahrten werden unter anderem in primitiven Modellen für die Kapitalentwicklung beim Glücksspiel oder an der Börse (Aktienkurs), sowie die Brownsche Molekularbewegung (im Skalierungslimes Schrittweite  $\rightarrow 0$ ) eingesetzt.

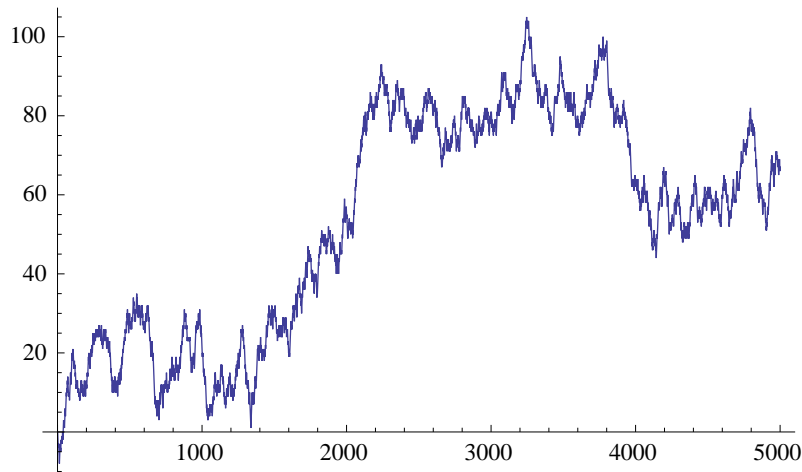
**Beispiel** (Symmetrischer Random Walk,  $p = \frac{1}{2}$ ). `zufall = RandomChoice[{-1, 1}, 10000];`  
`randomwalk = FoldList[Plus, 0, zufall];`  
`Manipulate[ListLinePlot[randomwalk[[1;;nmax]]], {nmax, 10, 10000, 10}]`



•  $nmax = 50$ :



•  $nmax = 500$ :



- $nmax = 5000$ :

**Lemma 2.8** (Verteilung von  $S_n$ ). Für  $k \in \mathbb{Z}$  gilt

$$P[S_n = a + k] = \begin{cases} 0 & \text{falls } n + k \text{ ungerade oder } |k| > n, \\ \binom{n+k}{\frac{n+k}{2}} p^{\frac{n+k}{2}} (1-p)^{\frac{n-k}{2}} & \text{sonst.} \end{cases}$$

*Beweis.* Es gilt:

$$S_n = a + k \Leftrightarrow X_1 + \dots + X_n = k \Leftrightarrow \begin{cases} X_i = 1 & \text{genau } \frac{n+k}{2}\text{-mal,} \\ X_i = -1 & \text{genau } \frac{n-k}{2}\text{-mal.} \end{cases}$$

□

**Beispiel** (Rückkehrwahrscheinlichkeit zum Startpunkt). Mithilfe der *Stirlingschen Formel*

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \quad \text{für } n \rightarrow \infty.$$

folgt:

$$\begin{aligned} P[S_{2n+1} = a] &= 0, \\ P[S_{2n} = a] &= \binom{2n}{n} p^n (1-p)^n = \frac{(2n)!}{(n!)^2} p^n (1-p)^n \\ &\sim \frac{\sqrt{4\pi n} \left(\frac{2n}{e}\right)^{2n}}{2\pi n \left(\frac{n}{e}\right)^{2n}} p^n (1-p)^n = \frac{1}{\sqrt{\pi n}} (4p(1-p))^n \quad \text{für } n \rightarrow \infty, \end{aligned}$$

wobei zwei Folgen  $a_n$  und  $b_n$  **asymptotisch äquivalent** heißen ( $a_n \sim b_n$ ), falls  $\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = 1$  gilt.

- Falls  $p \neq \frac{1}{2}$  gilt  $4p(1-p) < 1$  und  $P[S_{2n} = a]$  konvergiert exponentiell schnell gegen 0.
- Falls  $p = \frac{1}{2}$  konvergiert  $P[S_{2n} = a] \sim \frac{1}{\sqrt{\pi n}}$  nur langsam gegen 0.

### Symmetrischer Random Walk

Ab jetzt betrachten wir den *symmetrischen* Random Walk, d.h.  $p = \frac{1}{2}$ .

Sei  $\lambda \in \mathbb{Z}$ . Wir wollen die Verteilung der Zufallsvariable

$$T_\lambda(\omega) := \inf\{n \in \mathbb{N} \mid S_n(\omega) = \lambda\}, \quad (\min \emptyset := \infty),$$

bestimmen. Für  $\lambda \neq a$  ist  $T_\lambda$  die erste *Trefferzeit* von  $\lambda$ , für  $\lambda = a$  ist es die erste *Rückkehrzeit nach*  $a$ . Beschreibt der Random Walk beispielsweise die Kapitalentwicklung in einem Glücksspiel, dann kann man  $T_0$  als Ruinzeitpunkt interpretieren.

Sei  $n \in \mathbb{N}$ . Wir wollen nun die Wahrscheinlichkeit

$$P[T_\lambda \leq n] = P\left[\bigcup_{i=1}^n \{S_i = \lambda\}\right]$$

berechnen. Da das Ereignis  $\{T_\lambda \leq n\}$  von *mehreren* Positionen des Random Walks abhängt  $(S_1, S_2, \dots, S_n)$ , benötigen wir die *gemeinsame* Verteilung dieser Zufallsvariablen. Sei also

$$S(\omega) := (S_0(\omega), S_1(\omega), \dots, S_n(\omega))$$

der *Bewegungsverlauf bis zur Zeit*  $n$ . Dann ist  $S$  eine Zufallsvariable mit Werten im Raum

$$\widehat{\Omega}_a^{(n)} := \{(s_0, s_1, \dots, s_n) \mid s_0 = a, s_i \in \mathbb{Z}, \text{ so dass: } |s_i - s_{i-1}| = 1 \text{ für alle } i \in \{1, \dots, n\}\}$$

der möglichen Pfade des Random Walk. Sei  $\mu_a$  die gemeinsame Verteilung von  $S$  unter  $P$ .

**Lemma 2.9.**  $\mu_a$  ist die Gleichverteilung auf dem Pfadraum  $\widehat{\Omega}_a^{(n)}$ .

*Beweis.* Es gilt

$$\begin{aligned} \mu_a(\{(s_0, \dots, s_n)\}) &= P[S_0 = s_0, \dots, S_n = s_n] \\ &= P[S_0 = s_0, X_1 = s_1 - s_0, \dots, X_n = s_n - s_{n-1}] \\ &= \begin{cases} 0 & \text{falls } s_0 \neq a \text{ oder } |s_i - s_{i-1}| \neq 1 \text{ für ein } i \in \{1, \dots, n\}, \\ & \text{(d.h. } (s_0, \dots, s_n) \notin \widehat{\Omega}_a^{(n)}), \\ 2^{-n} & \text{sonst, d.h. falls } (s_0, \dots, s_n) \in \widehat{\Omega}_a^{(n)}. \end{cases} \end{aligned}$$

□

**Satz 2.10** (Reflektionsprinzip). Seien  $\lambda, b \in \mathbb{Z}$ . Es gelte entweder  $(a < \lambda \text{ und } b \leq \lambda)$ , oder  $(a > \lambda \text{ und } b \geq \lambda)$ . Dann gilt:

$$P[T_\lambda \leq n, S_n = b] = P[S_n = b^*],$$

wobei  $b^* := \lambda + (\lambda - b) = 2\lambda - b$  die Spiegelung von  $b$  an  $\lambda$  ist.

*Beweis.* Es gilt:

$$P[T_\lambda \leq n, S_n = b] = \mu_a[\overbrace{\{(s_0, \dots, s_n) \mid s_n = b, s_i = \lambda \text{ für ein } i \in \{1, \dots, n\}\}}^{=:A}],$$

$$P[S_n = b^*] = \mu_a[\underbrace{\{(s_0, \dots, s_n) \mid s_n = b^*\}}_{=:B}].$$

Die im Bild dargestellte Transformation (Reflexion des Pfades nach Treffen von  $\lambda$ ) definiert eine Bijektion von  $A$  nach  $B$ . Also gilt  $|A| = |B|$ . Da  $\mu_a$  die Gleichverteilung auf  $\widehat{\Omega}_a^{(n)}$  ist, folgt:

$$\mu_a(A) = \frac{|A|}{|\widehat{\Omega}_a^{(n)}|} = \frac{|B|}{|\widehat{\Omega}_a^{(n)}|} = \mu_a(B).$$

□

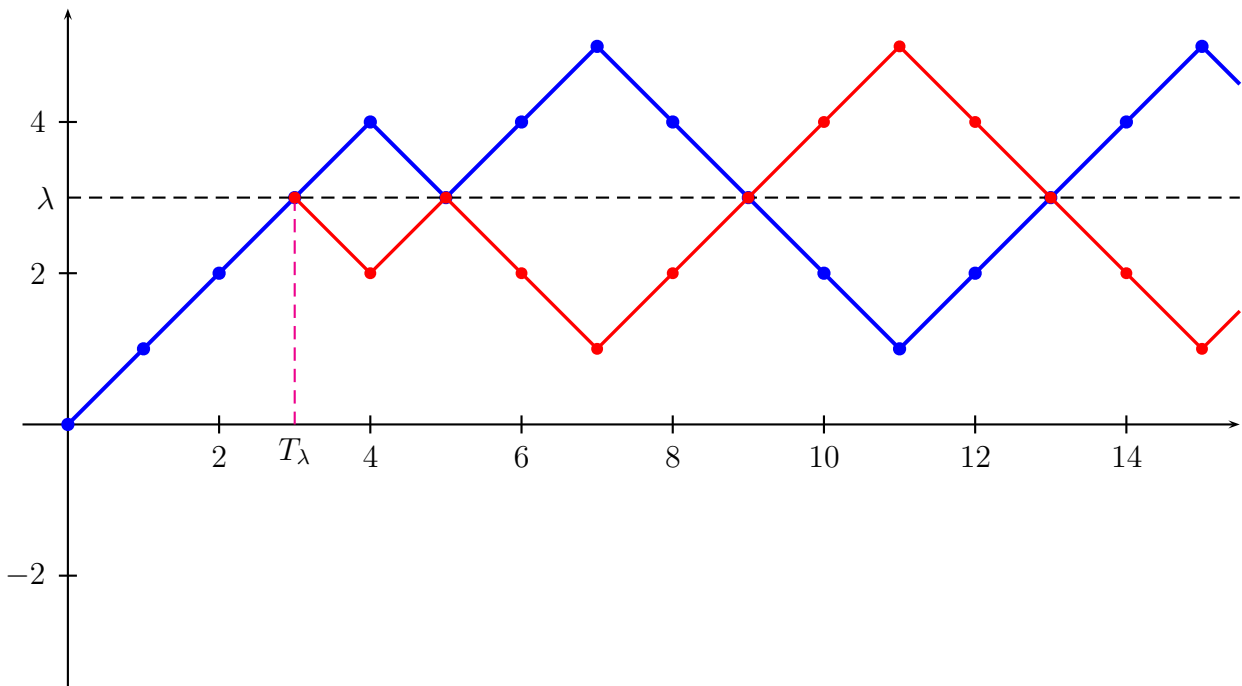


Abbildung 2.2: Spiegelung des Random Walks an  $\lambda = 3$

**Korollar** (Verteilung der Trefferzeiten). *Es gilt:*

i)

$$P[T_\lambda \leq n] = \begin{cases} P[S_n \geq \lambda] + P[S_n > \lambda], & \text{falls } \lambda > a, \\ P[S_n \leq \lambda] + P[S_n < \lambda], & \text{falls } \lambda < a. \end{cases}$$



ii)

$$\begin{aligned}
P[T_\lambda = n] &= \begin{cases} \frac{1}{2}P[S_{n-1} = \lambda - 1] - \frac{1}{2}P[S_{n-1} = \lambda + 1], & \text{falls } \lambda > a, \\ \frac{1}{2}P[S_{n-1} = \lambda + 1] - \frac{1}{2}P[S_{n-1} = \lambda - 1], & \text{falls } \lambda < a. \end{cases} \\
&= \begin{cases} \frac{\lambda-a}{n} \binom{n}{\frac{n+\lambda-a}{2}} 2^{-n} & \text{falls } \lambda > a, \\ \frac{a-\lambda}{n} \binom{n}{\frac{n+\lambda-a}{2}} 2^{-n} & \text{falls } \lambda < a. \end{cases}
\end{aligned}$$

*Beweis.* Wir beweisen die Aussagen für  $\lambda > a$ , der andere Fall wird jeweils analog gezeigt.

i)

$$\begin{aligned}
P[T_\lambda \leq n] &= \sum_{b \in \mathbb{Z}} \underbrace{P[T_\lambda \leq n, S_n = b]} \\
&= \begin{cases} P[S_n = b] & \text{falls } b \geq \lambda, \\ P[S_n = b^*] & \text{falls } b < \lambda. \end{cases} \\
&= \sum_{b \geq \lambda} P[S_n = b] + \underbrace{\sum_{b < \lambda} P[S_n = b^*]}_{= \sum_{b > \lambda} P[S_n = b]} = P[S_n \geq \lambda] + P[S_n > \lambda].
\end{aligned}$$

ii)

$$P[T_\lambda = n] = P[T_\lambda \leq n] - P[T_\lambda \leq n - 1]$$

Mit i) folgt

$$\begin{aligned}
&= \underbrace{P[S_n \geq \lambda]}_{=:A} - \underbrace{P[S_{n-1} \geq \lambda]}_{=:B} + \underbrace{P[S_n \geq \lambda + 1]}_{=:I} - \underbrace{P[S_{n-1} \geq \lambda + 1]}_{=:II}
\end{aligned}$$

Wegen

$$P[A] - P[B] = P[A \setminus B] + P[A \cap B] - P[B \setminus A] - P[B \cap A] = P[A \setminus B] - P[B \setminus A]$$

erhalten wir für den ersten Term:

$$\begin{aligned}
\mathbf{I} &= P[S_n \geq \lambda, S_{n-1} < \lambda] - P[S_{n-1} \geq \lambda, S_n < \lambda] \\
&= P[S_{n-1} = \lambda - 1, S_n = \lambda] - P[S_{n-1} = \lambda, S_n = \lambda - 1] \\
&= \frac{1}{2}P[S_{n-1} = \lambda - 1] - \frac{1}{2}P[S_{n-1} = \lambda].
\end{aligned}$$

Hierbei haben wir benutzt, dass

$$\begin{aligned}
 & |\{(s_0, \dots, s_n) \in \widehat{\Omega}_a^{(n)} \mid s_{n-1} = \lambda - 1\}| \\
 &= |\{(s_0, \dots, s_n) \mid s_{n-1} = \lambda - 1 \text{ und } s_n = \lambda\}| \\
 &\quad + |\{(s_0, \dots, s_n) \mid s_{n-1} = \lambda - 1 \text{ und } s_n = \lambda - 2\}| \\
 &= 2 \cdot |\{(s_0, \dots, s_n) \mid s_{n-1} = \lambda - 1, s_n = \lambda\}|
 \end{aligned}$$

gilt. Mit einer analogen Berechnung für den zweiten Term erhalten wir insgesamt:

$$\begin{aligned}
 P[T_\lambda = n] &= \mathbf{I} + \mathbf{II} \\
 &= \frac{1}{2} (P[S_{n-1} = \lambda - 1] - P[S_{n-1} = \lambda] \\
 &\quad + P[S_{n-1} = (\lambda + 1) - 1] - P[S_{n-1} = \lambda + 1]) \\
 &= \frac{1}{2} (P[S_{n-1} = \lambda - 1] - P[S_{n-1} = \lambda + 1]).
 \end{aligned}$$

□

Sei  $M_n := \max(S_0, S_1, \dots, S_n)$ .

**Korollar** (Verteilung des Maximums). *Für  $\lambda > a$  gilt:*

$$P[M_n \geq \lambda] = P[T_\lambda \leq n] = P[S_n \geq \lambda] + P[S_n > \lambda].$$

## 2.5 Simulationsverfahren

Die Simulation von Stichproben verschiedener Wahrscheinlichkeitsverteilungen geht von auf  $[0, 1]$  gleichverteilten Pseudo-Zufallszahlen aus. In Wirklichkeit simuliert ein Zufallszahlengenerator natürlich nur auf  $\{k m^{-1} \mid k = 0, 1, \dots, m - 1\}$  gleichverteilte Zufallszahlen, wobei  $m^{-1}$  die Darstellungsgenauigkeit des Computers ist. Dieser Aspekt wird im folgenden ignoriert. Um Simulationsverfahren zu analysieren, benötigen wir noch den Begriff einer auf  $[0, 1]$  gleichverteilten reellwertigen Zufallsvariablen. Die Existenz solcher Zufallsvariablen auf einem geeigneten Wahrscheinlichkeitsraum wird hier vorausgesetzt, und kann erst in der Vorlesung »Analysis III« bzw. in der »Einführung in die Wahrscheinlichkeitstheorie« gezeigt werden.

Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum, und  $U: \Omega \rightarrow [0, 1]$  eine Abbildung.

**Definition.** 1.  $U$  ist eine *reellwertige Zufallsvariable*, falls gilt:

$$\{\omega \in \Omega \mid U(\omega) \leq y\} \in \mathcal{A} \quad \text{für alle } y \in \mathbb{R}.$$

2. Eine reellwertige Zufallsvariable  $U : \Omega \rightarrow [0, 1]$  ist **gleichverteilt auf**  $[0, 1]$ , falls

$$P[U \leq y] = y \quad \text{für alle } y \in [0, 1].$$

Wir notieren dies im folgenden als  $(U \sim \text{Unif}[0, 1])$ .

3. Reellwertige Zufallsvariablen  $U_i : \Omega \rightarrow \mathbb{R}$ ,  $i \in I$ , heißen **unabhängig**, falls die Ereignisse  $\{U_i \leq y_i\}$ ,  $i \in I$ , für alle  $y_i \in \mathbb{R}$  unabhängig sind.

Ein *Zufallszahlengenerator* simuliert Stichproben  $u_1 = U_1(\omega)$ ,  $u_2 = U_2(\omega)$ , ... von auf  $[0, 1]$  gleichverteilten unabhängigen Zufallsvariablen. Wie erzeugt man daraus Stichproben von diskreten Verteilungen?

## Das direkte Verfahren

Sei  $S = \{a_1, a_2, \dots\}$  endlich oder abzählbar unendlich, und  $\mu$  eine Wahrscheinlichkeitsverteilung auf  $S$  mit Gewichten  $p_i = p(a_i)$ . Wir setzen

$$s_n := \sum_{i=1}^n p_i, \quad n \in \mathbb{N}, \quad \text{»kumulative Verteilungsfunktion«.}$$

Sei  $U : \Omega \rightarrow [0, 1)$  eine gleichverteilte Zufallsvariable. Wir setzen

$$X(\omega) := a_i, \quad \text{falls } s_{i-1} < U(\omega) \leq s_i, \quad i \in \mathbb{N}.$$

**Lemma 2.11.** Falls  $U \sim \text{Unif}[0, 1)$ , gilt  $X \sim \mu$ .

*Beweis.* Für alle  $i \in \mathbb{N}$  gilt:

$$P[X = a_i] = P[s_{i-1} < U \leq s_i] = P[U \leq s_i] - P[U \leq s_{i-1}] = s_i - s_{i-1} = p_i.$$

□

**Algorithmus 2.12** (Direkte Simulation einer diskreten Verteilung).

INPUT: Gewichte  $p_1, p_2, \dots$ ,

OUTPUT: Pseudozufallsstichprobe  $x$  von  $\mu$ .

$n := 1$

$s := p_1$

erzeuge Zufallszahl  $u \sim \text{Unif}[0, 1)$

**while**  $u > s$  **do**

$n := n + 1$

```

    s := s + p_n
end while
return x := a_n

```

**Bemerkung.** a) Die mittlere Anzahl von Schritten des Algorithmus ist

$$\sum_{n=1}^{\infty} n p_n = \text{Erwartungswert von Wahrscheinlichkeitsverteilung } (p_n) \text{ auf } \mathbb{N}.$$

b) Für große Zustandsräume  $S$  ist das direkte Verfahren oft nicht praktikabel, siehe Übung.

## Acceptance-Rejection-Verfahren

Sei  $S$  eine endliche oder abzählbare Menge,  $\mu$  eine Wahrscheinlichkeitsverteilung auf  $S$  mit Massenfunktion  $p(x)$ , und  $\nu$  eine Wahrscheinlichkeitsverteilung auf  $S$  mit Massenfunktion  $q(x)$ . Angenommen, wir können unabhängige Stichproben von  $\nu$  erzeugen. Wie können wir daraus Stichproben von  $\mu$  erhalten? IDEE: Erzeuge Stichprobe  $x$  von  $\nu$ , akzeptiere diese mit Wahrscheinlichkeit proportional zu  $\frac{p(x)}{q(x)}$ , sonst verwerfe die Stichprobe und wiederhole.

ANNAHME:

$$\text{es gibt ein } c \in [1, \infty) : \quad p(x) \leq c q(x) \quad \text{für alle } x \in S.$$

Aus der Annahme folgt:

$$\frac{p(x)}{c q(x)} \leq 1 \quad \text{für alle } x \in S,$$

d.h. wir können  $\frac{p(x)}{c q(x)}$  als *Akzeptanzwahrscheinlichkeit* wählen.

**Algorithmus 2.13** (Acceptance-Rejection-Verfahren).

INPUT: Gewichte  $p(y)$ ,  $q(y)$ ,  $c$  ( $y \in S$ ),

OUTPUT: Stichprobe  $x$  von  $\mu$ .

**repeat**

    erzeuge Stichprobe  $x \sim \nu$

    erzeuge Stichprobe  $u \sim \text{Unif}[0, 1]$

**until**  $\frac{p(x)}{c q(x)} \geq u$  {akzeptiere mit Wahrscheinlichkeit  $\frac{p(x)}{c q(x)}$ }

**return**  $x$

ANALYSE DES ALGORITHMUS

Für die verwendeten Zufallsvariablen gilt:

$$X_1, X_2, \dots \sim \nu, \quad (\text{Vorschläge}),$$

$$U_1, U_2, \dots \sim \text{Unif}[0, 1].$$

Es gilt Unabhängigkeit, d.h.

$$P[X_1 = a_1, \dots, X_n = a_n, U_1 \leq y_1, \dots, U_n \leq y_n] = \prod_{i=1}^n P[X_i = a_i] \cdot \prod_{i=1}^n P[U_i \leq y_i]$$

für alle  $n \in \mathbb{N}$ ,  $a_i \in S$  und  $y_i \in \mathbb{R}$ .

Seien

$$\begin{aligned} T &= \min \left\{ n \in \mathbb{N} \mid \frac{p(X_n)}{c q(X_n)} \geq U_n \right\} && \text{die »Akzeptanzzeit«, und} \\ X_T(\omega) &= X_{T(\omega)}(\omega) && \text{die ausgegebene Stichprobe.} \end{aligned}$$

des Acceptance-Rejection-Verfahrens. Wir erhalten:

**Satz 2.14.** i)  $T$  ist geometrisch verteilt mit Parameter  $1/c$ ,

ii)  $X_T \sim \mu$ .

**Bemerkung.** Insbesondere ist die mittlere Anzahl von Schritten bis Akzeptanz:

$$E[T] = c.$$

*Beweis.* i) Sei

$$A_n := \left\{ \frac{p(X_n)}{c q(X_n)} \geq U_n \right\}.$$

Aus der Unabhängigkeit der Zufallsvariablen  $X_1, U_1, X_2, U_2, \dots$  folgt, daß auch die Ereignisse  $A_1, A_2, \dots$  unabhängig sind. Dies wird in der Vorlesung »Einführung in die Wahrscheinlichkeitstheorie« bewiesen. Zudem gilt wegen der Unabhängigkeit von  $X_n$  und  $U_n$ :

$$\begin{aligned} P[A_n] &= \sum_{a \in S} P \left[ \left\{ U_n \leq \frac{p(a)}{c q(a)} \right\} \cap \{X_n = a\} \right] \\ &= \sum_{a \in S} P \left[ \left\{ U_n \leq \frac{p(a)}{c q(a)} \right\} \right] \cdot P[X_n = a] \\ &= \sum_{a \in S} \frac{p(a)}{c q(a)} \cdot q(a) = \frac{1}{c}. \end{aligned}$$

Also ist

$$T(\omega) = \min\{n \in \mathbb{N} \mid \omega \in A_n\}$$

geometrisch verteilt mit Parameter  $1/c$ .

ii)

$$\begin{aligned} P[X_T = a] &= \sum_{n=1}^{\infty} P[\{X_T = a\} \cap \{T = n\}] \\ &= \sum_{n=1}^{\infty} P[\{X_n = a\} \cap A_n \cap A_1^C \cap \dots \cap A_{n-1}^C] \\ &= \sum_{n=1}^{\infty} P[\{X_n = a\} \cap \left\{ \frac{p(a)}{c q(a)} \geq U_n \right\} \cap A_1^C \cap \dots \cap A_{n-1}^C] \\ &= \sum_{n=1}^{\infty} q(a) \frac{p(a)}{c q(a)} \left(1 - \frac{1}{c}\right)^{n-1} \\ &= \frac{p(a)}{c} \sum_{n=1}^{\infty} \left(1 - \frac{1}{c}\right)^{n-1} \\ &= \frac{p(a)}{c} \frac{1}{1 - (1 - \frac{1}{c})} = p(a). \end{aligned}$$

□

# Kapitel 3

## Konvergenzsätze und Monte Carlo-Methoden

Sei  $\mu$  eine Wahrscheinlichkeitsverteilung auf einer abzählbaren Menge  $S$ , und  $f : S \rightarrow \mathbb{R}$  eine reellwertige Zufallsvariable. Angenommen, wir wollen den Erwartungswert

$$\theta := E_\mu[f] = \sum_{x \in S} f(x) \mu(x)$$

berechnen, aber die Menge  $S$  ist zu groß, um die Summe direkt auszuführen. In einem *Monte Carlo-Verfahren* simuliert man eine große Anzahl unabhängiger Stichproben  $X_1(\omega), \dots, X_n(\omega)$  von  $\mu$ , und approximiert den Erwartungswert  $\theta$  durch den *Monte Carlo-Schätzer*

$$\hat{\theta}_n(\omega) := \frac{1}{n} \sum_{i=1}^n f(X_i(\omega)).$$

Wir wollen nun Methoden entwickeln, mit denen der Approximationsfehler  $|\hat{\theta}_n - \theta|$  abgeschätzt werden kann, und die Asymptotik des Approximationsfehlers für  $n \rightarrow \infty$  untersuchen. Nach dem Transformationssatz (1.7) und der Linearität des Erwartungswerts (1.8) gilt:

$$E[\hat{\theta}_n] = \frac{1}{n} \sum_{i=1}^n E[f(X_i)] = \frac{1}{n} \sum_{i=1}^n \sum_{x \in S} f(x) \mu(\{x\}) = E_\mu[f] = \theta,$$

d.h.  $\hat{\theta}_n$  ist ein *erwartungstreuer* Schätzer für  $\theta$ . Der mittlere quadratische Fehler (»MSE« – mean squared error) des Schätzers ist daher:

$$\text{MSE} = E[|\hat{\theta}_n - \theta|^2] = E[|\hat{\theta}_n - E[\hat{\theta}_n]|^2].$$

### 3.1 Varianz und Kovarianz

Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum und  $X : \Omega \rightarrow S$  eine Zufallsvariable auf  $(\Omega, \mathcal{A}, P)$ , so dass  $E[|X|]$  endlich ist.

**Definition.**

$$\text{Var}(X) := E[(X - E[X])^2]$$

heißt **Varianz** von  $X$  und liegt in  $[0, \infty]$ .

$$\sigma(X) := \text{Var}(X)^{\frac{1}{2}}$$

heißt **Standardabweichung** von  $X$ .

Die Varianz bzw. Standardabweichung kann als Kennzahl für die Größe der Fluktuationen (Streuung) der Zufallsvariablen  $X$  um den Erwartungswert  $E[X]$  und damit als Maß für das Risiko bei Prognose des Ausgangs  $X(\omega)$  durch  $E[X]$  interpretiert werden.

**Bemerkung.** (a) Die Varianz hängt nur von der Verteilung von  $X$  ab:

$$\text{Var}(X) = \sum_{a \in S} (a - m)^2 p_X(a), \quad \text{wobei} \quad m = E[X] = \sum_{a \in S} a p_X(a).$$

(b) Es gilt

$$\text{Var}(X) = 0 \quad \text{genau dann, wenn} \quad P[X = E[X]] = 1.$$

**Bemerkung** (Rechenregeln). i)

$$\text{Var}(X) = E[X^2] - E[X]^2.$$

Insbesondere ist die Varianz von  $X$  genau dann endlich, wenn  $E[X^2]$  endlich ist.

ii)

$$\text{Var}(aX + b) = \text{Var}(aX) = a^2 \text{Var}(X) \quad \text{für alle } a, b \in \mathbb{R}.$$

**Beweis.** i) Nach der Linearität des Erwartungswerts gilt

$$\text{Var}(X) = E[X^2 - 2X \cdot E[X] + E[X]^2] = E[X^2] - E[X]^2.$$

ii) Wiederholte Anwendung der Linearität des Erwartungswerts liefert

$$\text{Var}(aX + b) = E[(aX + b - E[aX + b])^2] = E[(aX - E[aX])^2] = a^2 \text{Var}(X).$$

□



**Beispiel.** 1. Sei  $X = 1$  mit Wahrscheinlichkeit  $p$  und  $X = 0$  mit Wahrscheinlichkeit  $1 - p$ . Dann ist der Erwartungswert von  $X$ :

$$E[X^2] = E[X] = p,$$

und die Varianz von  $X$ :

$$\text{Var}(X) = p - p^2 = p(1 - p).$$

2. Sei  $T$  geometrisch verteilt ( $T \sim \text{Geom}(p)$ ) mit Parameter  $p \in (0, 1]$ . Der Erwartungswert von  $T$  beträgt:

$$E[T] = \sum_{k=1}^{\infty} k(1-p)^{k-1}p = -p \frac{d}{dp} \sum_{k=0}^{\infty} (1-p)^k = -p \frac{d}{dp} \frac{1}{p} = \frac{1}{p}.$$

Außerdem gilt:

$$\begin{aligned} E[T(T+1)] &= \sum_{k=1}^{\infty} k(k+1)(1-p)^{k-1}p \\ &= \sum_{k=1}^{\infty} k(k-1)(1-p)^{k-2}p = p \frac{d^2}{dp^2} \sum_{k=0}^{\infty} (1-p)^k = \frac{2}{p^2}. \end{aligned}$$

Die Varianz von  $T$  ist somit:

$$\text{Var}(T) = E[T^2] - E[T]^2 = \frac{2}{p^2} - \frac{1}{p} - \frac{1}{p^2} = \frac{1-p}{p^2}.$$

**Definition.**

$$\mathcal{L}^2(\Omega, \mathcal{A}, P) := \{X: \Omega \rightarrow \mathbb{R} \mid X \text{ ist diskrete Zufallsvariable mit } E[X^2] < \infty\}$$

**Lemma 3.1.** 1. Für Zufallsvariablen  $X, Y \in \mathcal{L}^2$  gilt:

$$E[|XY|] \leq E[X^2]^{\frac{1}{2}} E[Y^2]^{\frac{1}{2}} < \infty.$$

2.  $\mathcal{L}^2$  ist ein Vektorraum, und

$$(X, Y)_{\mathcal{L}^2} := E[XY]$$

ist eine positiv semidefinite symmetrische Bilinearform (»Skalarprodukt«) auf  $\mathcal{L}^2$ .

**Bemerkung.** i) Insbesondere gilt die *Cauchy-Schwarz-Ungleichung*:

$$E[XY]^2 \leq E[|XY|] \leq E[X^2] E[Y^2] \quad \text{für alle } X, Y \in \mathcal{L}^2.$$

ii) Für eine Zufallsvariable  $X \in \mathcal{L}^2$  gilt

$$E[|X|] \leq \sqrt{E[X^2]} E[1^2]^{\frac{1}{2}} < \infty.$$

*Beweis.* i) Nach der Cauchy-Schwarz-Ungleichung gilt:

$$\begin{aligned} E[|XY|] &= \sum_{\substack{a \in X(\Omega) \\ b \in Y(\Omega)}} |ab| P[X = a, Y = b] \\ &= \sum_{\substack{a \in X(\Omega) \\ b \in Y(\Omega)}} |a| P[X = a, Y = b]^{\frac{1}{2}} |b| P[X = a, Y = b]^{\frac{1}{2}} \\ &\leq \left( \sum_{a,b} a^2 P[X = a, Y = b] \right)^{\frac{1}{2}} \left( \sum_{a,b} b^2 P[X = a, Y = b] \right)^{\frac{1}{2}} \\ &= \left( \sum_a a^2 P[X = a] \right)^{\frac{1}{2}} \left( \sum_b b^2 P[Y = b] \right)^{\frac{1}{2}} \\ &= E[X^2]^{\frac{1}{2}} E[Y^2]^{\frac{1}{2}}. \end{aligned}$$

ii) Seien  $X, Y \in \mathcal{L}^2$ ,  $a \in \mathbb{R}$ . Dann ist  $aX + Y$  eine diskrete Zufallsvariable, für die nach Monotonie und der Linearität des Erwartungswerts gilt:

$$E[(aX + Y)^2] = E[a^2X^2 + 2aXY + Y^2] \leq 2a^2E[X^2] + 2E[Y^2] < \infty.$$

$(X, Y)_{\mathcal{L}^2} = E[XY]$  ist bilinear, da  $E[\bullet]$  linear und symmetrisch ist, und positiv semidefinit, aufgrund von:

$$(X, X)_{\mathcal{L}^2} = E[X^2] \geq 0 \quad \text{für alle } X \in \mathcal{L}^2.$$

□

**Definition.** Seien  $X, Y \in \mathcal{L}^2$ .

1.

$$\text{Cov}(X, Y) := E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$$

heißt **Kovarianz** von  $X$  und  $Y$ .

2. Gilt  $\sigma(X), \sigma(Y) \neq 0$ , so heißt

$$\varrho(X, Y) := \frac{\text{Cov}(X, Y)}{\sigma(X)\sigma(Y)}$$

**Korrelationskoeffizient** von  $X$  und  $Y$ .

3.  $X$  und  $Y$  heißen **unkorreliert**, falls  $\text{Cov}(X, Y) = 0$ , d.h.

$$E[XY] = E[X] \cdot E[Y].$$

**Bemerkung.**  $\text{Cov} : \mathcal{L}^2 \times \mathcal{L}^2 \rightarrow \mathbb{R}$  ist eine symmetrische Bilinearform mit:

$$\text{Cov}(X, X) = \text{Var}(X) \geq 0 \quad \text{für alle } X \in \mathcal{L}^2.$$

**Satz 3.2** (Zusammenhang von Unabhängigkeit und Unkorreliertheit). *Seien  $X : \Omega \rightarrow S$  und  $Y : \Omega \rightarrow T$  diskrete Zufallsvariablen auf  $(\Omega, \mathcal{A}, P)$ . Dann sind äquivalent:*

(i)  $X$  und  $Y$  sind unabhängig, d.h.

$$P[X \in A, Y \in B] = P[X \in A] P[Y \in B] \quad \text{für alle } A, B \in \mathcal{A}.$$

(ii)  $f(X)$  und  $g(Y)$  sind unkorreliert für alle Funktionen  $f : S \rightarrow \mathbb{R}$  und  $g : T \rightarrow \mathbb{R}$  mit  $f(X), g(Y) \in \mathcal{L}^2$ .

*Beweis.* • (i)  $\Rightarrow$  (ii): Seien  $X$  und  $Y$  unabhängig, dann gilt:

$$\begin{aligned} E[f(X)g(Y)] &= \sum_{a \in S} \sum_{b \in T} f(a) g(b) P[X = a, Y = b] \\ &= \sum_{a \in S} f(a) P[X = a] \sum_{b \in T} g(b) P[Y = b] = E[f(X)] E[g(Y)] \end{aligned}$$

Somit folgt:

$$\text{Cov}(f(X), g(Y)) = 0.$$

• (ii)  $\Rightarrow$  (i): Aus (ii) folgt für alle  $a \in S, b \in T$ :

$$\begin{aligned} P[X = a, Y = b] &= E[I_{\{a\}}(X) I_{\{b\}}(Y)] \\ &= E[I_{\{a\}}(X)] E[I_{\{b\}}(Y)] = P[X = a] P[Y = b]. \end{aligned}$$

□

**Beispiel.** Sei  $X = +1, 0, -1$  jeweils mit Wahrscheinlichkeit  $\frac{1}{3}$ , und  $Y = X^2$ . Dann sind  $X$  und  $Y$  nicht unabhängig, aber unkorreliert:

$$E[XY] = 0 = E[X] E[Y].$$

**Satz 3.3** (Varianz von Summen). Für  $X_1, \dots, X_n \in \mathcal{L}^2$  gilt:

$$\text{Var}(X_1 + \dots + X_n) = \sum_{i=1}^n \text{Var}(X_i) + 2 \sum_{\substack{i,j=1 \\ i < j}}^n \text{Cov}(X_i, X_j).$$

Falls  $X_1, \dots, X_n$  unkorreliert sind, folgt insbesondere:

$$\text{Var}(X_1 + \dots + X_n) = \sum_{i=1}^n \text{Var}(X_i).$$

*Beweis.* Nach Bilinearität der Kovarianz gilt:

$$\begin{aligned} \text{Var}(X_1 + \dots + X_n) &= \text{Cov}\left(\sum_{i=1}^n X_i, \sum_{j=1}^n X_j\right) \\ &= \sum_{i,j=1}^n \text{Cov}(X_i, X_j) = \sum_{i=1}^n \text{Var}(X_i) + 2 \sum_{\substack{i,j=1 \\ i < j}}^n \text{Cov}(X_i, X_j). \end{aligned}$$

□

**Beispiel** (Varianz der Binomialverteilung). Sei

$$S_n = \sum_{i=1}^n X_i, \quad X_i = \begin{cases} 1 & \text{mit Wahrscheinlichkeit } p, \\ 0 & \text{mit Wahrscheinlichkeit } 1 - p, \end{cases}$$

mit unabhängigen Zufallsvariablen  $X_i$ . Mit Satz 3.2 folgt:

$$\text{Var}(S_n) = \sum_{i=1}^n \text{Var}(X_i) = n p (1 - p).$$

Analog gilt für den Random Walk:

$$\sigma(S_n) = O(\sqrt{n}).$$

## 3.2 Gesetze der großen Zahlen

Seien  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum,  $X_1, X_2, \dots \in \mathcal{L}^2$  diskrete Zufallsvariablen,

$$S_n = X_1 + \dots + X_n, \quad \frac{S_n}{n} \quad \text{das arithmetische Mittel der } X_i.$$

ZIEL: Zufall mittelt sich weg, d.h.:

$$\frac{S_n(\omega)}{n} \approx \frac{E[S_n]}{n} \quad \text{für große } n.$$

Falls  $E[X_i] = m$  für alle  $i \in \mathbb{N}$  gilt, folgt insbesondere:

$$\frac{S_n(\omega)}{n} \longrightarrow m, \quad \text{falls } E[X_i] = m \text{ für alle } i \in \mathbb{N}.$$

Im allgemeinen ist die obige Aussage falsch, zum Beispiel mittelt sich für  $X_1 = \dots = X_n$  der Zufall nicht weg:

$$\frac{S_n}{n} = X_1.$$

Wir benötigen daher zusätzliche Annahmen.

ANNAHMEN:

(i) Die Zufallsvariablen sind unkorreliert:

$$\text{Cov}(X_i, X_j) = 0 \quad \text{für alle } i \neq j.$$

(ii) Die Varianzen sind beschränkt:

$$v := \sup_{i \in \mathbb{N}} \text{Var}(X_i) < \infty.$$

Es wird *keine Unabhängigkeit vorausgesetzt!*

**Satz 3.4** (Schwaches Gesetz der großen Zahlen). *Unter den Voraussetzungen (i) und (ii) gilt für alle  $\varepsilon > 0$ :*

$$P \left[ \left| \frac{S_n}{n} - \frac{E[S_n]}{n} \right| \geq \varepsilon \right] \leq \frac{v}{\varepsilon^2 n} \longrightarrow 0 \quad \text{für } n \rightarrow \infty.$$

*Gilt außerdem  $E[X_i] = m$  für alle  $i \in \mathbb{N}$ , folgt  $\frac{E[S_n]}{n} = m$  und  $\frac{S_n}{n}$  konvergiert stochastisch gegen  $m$ .*

Zum Beweis benötigen wir:

**Lemma 3.5** (Čebyšev-Ungleichung). *Für  $X \in \mathcal{L}^2$  und  $c > 0$  gilt:*

$$P[|X - E[X]| \geq c] \leq \frac{1}{c^2} \text{Var}(X).$$

*Beweis.* Es gilt

$$I_{\{|X - E[X]| \geq c\}} \leq \frac{1}{c^2} (X - E[X])^2$$

$\frac{1}{c^2} (X - E[X])^2$  ist überall nichtnegativ und  $\geq 1$  auf  $\{|X - E[X]| \geq c\}$ . Durch Bilden des Erwartungswerts folgt:

$$P[|X - E[X]| \geq c] = E[I_{\{|X - E[X]| \geq c\}}] \leq E\left[\frac{1}{c^2} (X - E[X])^2\right] = \frac{1}{c^2} E[(X - E[X])^2]$$

□

*Beweis von Satz 3.4.* Nach der Čebyšev-Ungleichung und den Annahmen (i) und (ii) gilt für  $\varepsilon > 0$ :

$$P \left[ \left| \frac{S_n}{n} - \frac{E[S_n]}{n} \right| \geq \varepsilon \right] \leq \frac{1}{\varepsilon^2} \operatorname{Var} \left( \frac{S_n}{n} \right) = \frac{1}{n^2 \varepsilon^2} \operatorname{Var} \left( \sum_{i=1}^n X_i \right) = \frac{1}{n^2 \varepsilon^2} \sum_{i=1}^n \operatorname{Var}(X_i) \leq \frac{v}{n \varepsilon^2}.$$

□

**Bemerkung** (Starkes Gesetz der großen Zahlen).

$$\frac{S_n(\omega)}{n} \longrightarrow m \quad \text{mit Wahrscheinlichkeit 1.}$$

Dies wird in der Vorlesung »Einführung in die Wahrscheinlichkeitstheorie« bewiesen.

### 3.3 Monte Carlo-Verfahren

Sei  $S$  eine abzählbare Menge und  $\mu$  eine Wahrscheinlichkeitsverteilung auf  $S$ . Wir bezeichnen im folgenden die Massenfunktion ebenfalls mit  $\mu$ , d.h.

$$\mu(x) := \mu(\{x\}).$$

Sei  $f: S \rightarrow \mathbb{R}$  eine reellwertige Funktion mit:

$$E_\mu[f^2] = \sum_{x \in S} f(x)^2 \mu(x) < \infty.$$

Wir wollen den Erwartungswert

$$\theta := E_\mu[f] = \sum_{x \in S} f(x) \mu(x)$$

näherungsweise berechnen bzw. schätzen. Dazu approximieren wir  $\theta$  durch die *Monte Carlo-Schätzer*

$$\hat{\theta}_n := \frac{1}{n} \sum_{i=1}^n f(X_i), \quad n \in \mathbb{N},$$

wobei  $X_1, X_2, \dots$  unabhängige Zufallsvariablen auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, P)$  mit Verteilung  $\mu$  sind. Nach der Abschätzung aus dem Gesetz der großen Zahlen ergibt sich:

**Korollar.**

$$P[|\hat{\theta}_n - \theta| \geq \varepsilon] \leq \frac{1}{n \varepsilon^2} \operatorname{Var}_\mu[f] \longrightarrow 0 \quad \text{für } n \rightarrow \infty,$$

d.h.  $\hat{\theta}_n$  ist eine **konsistente Schätzfolge** für  $\theta$ .

*Beweis.* Da die Zufallsvariablen  $X_i$  unabhängig sind, sind  $f(X_i)$ ,  $i \in \mathbb{N}$ , unkorreliert. Zudem gilt

$$E[f(X_i)] = \sum_{x \in S} f(x) \mu(x) = E_\mu[f] = \theta, \quad \text{und}$$

$$\text{Var}[f(X_i)] = \sum_{x \in S} (f(x) - \theta)^2 \mu(x) = \text{Var}_\mu[f] < \infty$$

nach Voraussetzung. Die Behauptung folgt nun aus Satz 3.4.  $\square$

**Bemerkung.** a)  $\hat{\theta}_n$  ist ein *erwartungstreuer Schätzer* für  $\theta$ :

$$E[\hat{\theta}_n] = \frac{1}{n} \sum_{i=1}^n E[f(X_i)] = E_\mu[f] = \theta.$$

b) Für den mittleren quadratischen Fehler des Schätzers ergibt sich nach a):

$$E[|\hat{\theta}_n - \theta|^2] = \text{Var}(\hat{\theta}_n) = \frac{1}{n} \text{Var}_\mu[f].$$

Insbesondere gilt:

$$\|\hat{\theta}_n - \theta\|_{\mathcal{L}^2} = E[|\hat{\theta}_n - \theta|^2]^{\frac{1}{2}} = O(1/\sqrt{n}).$$

**Beispiel.** a) MONTE CARLO-SCHÄTZUNG VON  $\theta = \int_{[0,1]^d} f(x) dx$ :

Das mehrdimensionale Integral ist folgendermaßen definiert:

$$\int_{[0,1]^d} f(x) dx := \int_0^1 \dots \int_0^1 f(x_1, \dots, x_d) dx_1 \dots dx_d.$$

Der Wert von  $\theta$  kann mit dem folgenden Algorithmus geschätzt werden.

erzeuge Pseudozufallszahlen  $u_1, u_2, \dots, u_{nd} \in (0, 1)$

$$x^{(1)} := (u_1, \dots, u_d)$$

$$x^{(2)} := (u_{d+1}, \dots, u_{2d})$$

...

$$x^{(n)} := (u_{(n-1)d+1}, \dots, u_{nd})$$

$$\hat{\theta}_n = \frac{1}{n} \sum_{i=1}^n f(x^{(i)}) \text{ ist Schätzwert für } \theta.$$

b) MONTE CARLO-SCHÄTZUNG VON WAHRSCHEINLICHKEITEN:

Sei  $S$  abzählbar,  $B \subseteq S$ . Wir suchen:

$$p = \mu(B) = E_\mu[I_B]$$

Ein Monte Carlo-Schätzer ist

$$\hat{p}_n = \frac{1}{n} \sum_{i=1}^n I_B(X_i), \quad X_i \text{ unabhängig mit Verteilung } \mu.$$

FEHLERKONTROLLE:

- Mithilfe der Čebyšev-Ungleichung (Lemma 3.5) ergibt sich:

$$P[|\hat{p}_n - p| \geq \varepsilon] \leq \frac{1}{\varepsilon^2} \operatorname{Var}(\hat{p}_n) = \frac{1}{n\varepsilon^2} \operatorname{Var}_\mu(I_B) = \frac{p(1-p)}{n\varepsilon^2} \leq \frac{1}{4n\varepsilon^2}.$$

Gilt beispielsweise  $n \geq \frac{5}{\varepsilon^2}$ , dann erhalten wir:

$$P[p \notin (\hat{p}_n - \varepsilon, \hat{p}_n + \varepsilon)] \leq 5\%, \quad \text{unabhängig von } p,$$

d.h. das zufällige Intervall  $(\hat{p}_n - \varepsilon, \hat{p}_n + \varepsilon)$  ist ein *95%-Konfidenzintervall* für den gesuchten Wert  $p$ .

- Mithilfe der Bernstein-Ungleichung (Chernoff-Abschätzung) erhalten wir für  $\delta > 0$  und  $S_n := \sum_{i=1}^n I_B(X_i)$ :

$$P[p \notin (\hat{p}_n - \varepsilon, \hat{p}_n + \varepsilon)] = P\left[\left|\frac{1}{n}S_n - p\right| \geq \varepsilon\right] \leq 2e^{-2n\varepsilon^2} \leq \delta, \quad \text{falls } n \geq \frac{\log(2/\delta)}{2\varepsilon^2}.$$

Für kleine  $\delta$  ist die erhaltene Bedingung an  $n$  wesentlich schwächer als eine entsprechende Bedingung, die man durch Anwenden der Čebyšev-Ungleichung erhält. Für den *relativen Schätzfehler*  $(\hat{p}_n - p)/p$  ergibt sich:

$$P[|\hat{p}_n - p| \geq \varepsilon p] \leq 2e^{-2n\varepsilon^2 p^2} \leq \delta, \quad \text{falls } n \geq \frac{\log(2/\delta)}{2\varepsilon^2 p^2}.$$

Die benötigte Anzahl von Stichproben für eine  $(\varepsilon, \delta)$ -Approximation von  $p$  ist also polynomiell in  $\varepsilon$ ,  $\log(1/\delta)$  und  $1/p$ . Mit einer etwas modifizierten Abschätzung kann man statt der Ordnung  $O(\frac{1}{p^2})$  sogar  $O(\frac{1}{p})$  erhalten, siehe [Mitzenmacher und Upfal: »Probability and Computing«](#).

**Beispiel.** Wie viele Stichproben sind nötig, damit der *relative Fehler* mit 95% Wahrscheinlichkeit unterhalb von 10% liegt? Mithilfe der Čebyšev-Ungleichung (Lemma 3.5) ergibt sich:

$$P[|\hat{p}_n - p| \geq 0,1p] \leq \frac{p(1-p)}{n(0,1p)^2} = \frac{100(1-p)}{np} \leq 0,05, \quad \text{falls } n \geq \frac{2000(1-p)}{p}.$$

So sind zum Beispiel für  $p = 10^{-5}$  ungefähr  $n \approx 2 \cdot 10^8$  Stichproben ausreichend. Dies ist nur eine obere Schranke, aber man kann zeigen, daß die tatsächlich benötigte Stichprobenzahl immer noch sehr groß ist. Für solch kleine Wahrscheinlichkeiten ist das einfache Monte Carlo-Verfahren ineffektiv, da die meisten Summanden von  $\hat{\theta}_n$  dann gleich 0 sind. Wir brauchen daher ein alternatives Schätzverfahren mit geringerer Varianz.



### Varianzreduktion durch Importance Sampling

Sei  $\nu$  eine weitere Wahrscheinlichkeitsverteilung auf  $S$  mit Massenfunktion  $\nu(x) = \nu(\{x\})$ . Es gelte  $\nu(x) > 0$  für alle  $x \in S$ . Dann können wir den gesuchten Wert  $\theta$  auch als Erwartungswert bzgl.  $\nu$  ausdrücken:

$$\theta = E_{\mu}[f] = \sum_{x \in S} f(x) \mu(x) = \sum_{x \in S} f(x) \frac{\mu(x)}{\nu(x)} \nu(x) = E_{\nu}[f \varrho],$$

wobei

$$\varrho(x) = \frac{\mu(x)}{\nu(x)}$$

der Quotient der beiden Massenfunktionen ist. Ein alternativer Monte Carlo-Schätzer für  $\theta$  ist daher

$$\tilde{\theta}_n = \frac{1}{n} \sum_{i=1}^n f(Y_i) \varrho(Y_i),$$

wobei die  $Y_i$  unabhängige Zufallsvariablen mit Verteilung  $\nu$  sind. Auch  $\tilde{\theta}_n$  ist erwartungstreu:

$$E_{\nu}[\tilde{\theta}_n] = E_{\nu}[f \varrho] = \theta.$$

Für die Varianz erhalten wir:

$$\text{Var}_{\nu}(\tilde{\theta}_n) = \frac{1}{n} \text{Var}_{\nu}(f \varrho) = \frac{1}{n} \left( \sum_{x \in S} f(x)^2 \varrho(x)^2 \nu(x) - \theta^2 \right).$$

Bei geeigneter Wahl von  $\nu$  kann die Varianz von  $\tilde{\theta}_n$  deutlich kleiner sein als die des Schätzers  $\hat{\theta}_n$ . Faustregel für eine gute Wahl von  $\nu$ :  $\nu(x)$  sollte groß sein, wenn  $|f(x)|$  groß ist.

»Importance Sampling«: Mehr Gewicht für die wichtigen  $x$ !

**Beispiel** (Zuverlässigkeit von Netzwerken; Perkolation). Gegeben sei ein endlicher Graph  $(V, E)$ , wobei  $V$  die Menge der Knoten und  $E$  die Menge der Kanten bezeichnet. Wir nehmen an, dass die Kanten unabhängig voneinander mit Wahrscheinlichkeit  $\varepsilon \ll 1$  ausfallen. Seien  $v, w \in E$  vorgegebene Knoten. Wir wollen die Wahrscheinlichkeit

$$p = P[\text{»}v \text{ nicht verbunden mit } w \text{ durch intakte Kanten«}]$$

approximativ berechnen. Sei

$$S = \{0, 1\}^E = \{(x_e)_{e \in E} \mid x_e \in \{0, 1\}\}$$

die Menge der Konfigurationen von intakten ( $x_l = 0$ ) bzw. defekten ( $x_l = 1$ ) Kanten und  $\mu$  die Wahrscheinlichkeitsverteilung auf  $S$  mit Massenfunktion

$$\mu(x) = \varepsilon^{k(x)} (1 - \varepsilon)^{|E| - k(x)}, \quad k(x) = \sum_{e \in E} x_e.$$

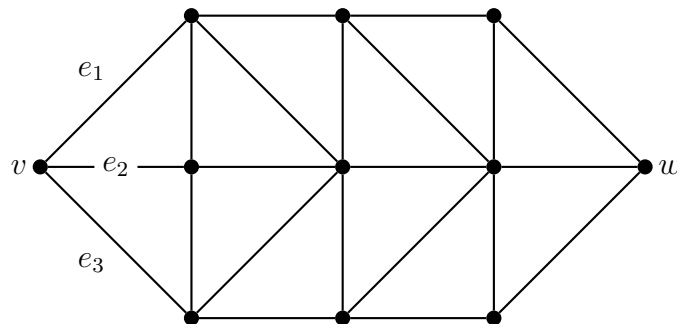


Abbildung 3.1: kleiner Beispielgraph für Perkolation

Sei

$$A = \{x \in S \mid v, w \text{ nicht verbunden durch Kanten } e \text{ mit } x_e = 0\}.$$

Dann ist

$$p = \mu(A) = E_\mu[I_A].$$

Der »klassische Monte Carlo-Schätzer« für  $p$  ist

$$\hat{p}_n = \frac{1}{n} \sum_{i=1}^n I_A(X_i), \quad X_i \text{ unabhängig mit Verteilung } \mu.$$

Fordern wir nun zum Beispiel

$$\sigma(\hat{p}_n) = \sqrt{\frac{p(1-p)}{n}} \stackrel{!}{\leq} \frac{p}{10},$$

dann benötigen wir eine Stichprobenanzahl

$$n \geq \frac{100(1-p)}{p},$$

um diese Bedingung zu erfüllen. Die Größenordnung von  $p$  für das in der obigen Graphik dargestellte Netzwerk mit  $\varepsilon = 1\%$  lässt sich wie folgt abschätzen:

$$\begin{aligned} 10^{-6} = \mu(\text{»}e_1, e_2, e_3 \text{ versagen«}) &\leq p \leq \mu(\text{»mindestens 3 Kanten versagen«}) \\ &= \binom{22}{3} \cdot 10^{-6} \approx 1,5 \cdot 10^{-3}. \end{aligned}$$

Es sind also eventuell mehrere Millionen Stichproben nötig!

Um die benötigte Stichprobenanzahl zu reduzieren, wenden wir ein Importance Sampling-Verfahren an. Sei

$$\nu(x) = t^{-k(x)} (1-t)^{|E|-k(x)}, \quad k(x) = \sum_{e \in E} x_e,$$

die Verteilung bei Ausfallwahrscheinlichkeit  $t := \frac{3}{22}$ . Da unter  $\nu$  im Schnitt 3 Kanten defekt sind, ist der Ausfall der Verbindung bzgl.  $\nu$  nicht mehr selten. Für den Schätzer

$$\tilde{p}_n = \frac{1}{n} \sum_{i=1}^n I_A(Y_i) \frac{\mu(Y_i)}{\nu(Y_i)}, \quad Y_i \text{ unabhängig mit Verteilung } \nu,$$

erhalten wir im Beispiel von oben:

$$\begin{aligned} \text{Var}(\tilde{p}_n) &= \frac{1}{n} \left( \sum_{x \in S} I_A(x)^2 \frac{\mu(x)^2}{\nu(x)} - p^2 \right) \\ &\leq \frac{1}{n} \sum_{k=3}^{22} \binom{|E|}{k} \left( \frac{\varepsilon^2}{t} \right)^k \left( \frac{(1-\varepsilon)^2}{1-t} \right)^{|E|-k} \leq 0,0053 \frac{p}{n}. \end{aligned}$$

Diese Abschätzung ist etwa um den Faktor 200 besser als die für den einfachen Monte Carlo-Schätzer erhaltene Abschätzung der Varianz.

### 3.4 Gleichgewichte von Markov-Ketten

Sei  $S$  eine abzählbare Menge,  $\nu$  eine Wahrscheinlichkeitsverteilung auf  $S$ , und  $p(x, y)$ , ( $x, y \in S$ ), eine *stochastische Matrix* bzw. *Übergangsmatrix*, d.h.  $p(x, y)$  erfüllt die folgenden Bedingungen:

- (i)  $p(x, y) \geq 0$  für alle  $x, y \in S$ ,
- (ii)  $\sum_{y \in S} p(x, y) = 1$  für alle  $x \in S$ .

Hier und im folgenden bezeichnen wir diskrete Wahrscheinlichkeitsverteilungen und die entsprechenden Massenfunktionen mit demselben Buchstaben, d.h.  $\nu(x) := \nu(\{x\})$ .

**Definition.** Eine Folge  $X_0, X_1, \dots : \Omega \rightarrow S$  von Zufallsvariablen auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, P)$  heißt **zeitlich homogene Markov-Kette** mit Startverteilung  $\nu$  und Übergangsmatrix  $p$ , falls die folgenden Bedingungen erfüllt sind:

- (i) Für alle  $x_0 \in S$  gilt:

$$P[X_0 = x_0] = \nu(x_0)$$

- (ii) Für alle  $n \in \mathbb{N}$  und  $x_0, \dots, x_{n+1} \in S$  mit  $P[X_0 = x_0, \dots, X_n = x_n] \neq 0$  gilt:

$$P[X_{n+1} = x_{n+1} \mid X_0 = x_0, \dots, X_n = x_n] = p(x_n, x_{n+1}).$$

**Bemerkung.** Die Bedingungen (i) und (ii) sind äquivalent zu:

$$P[X_0 = x_0, \dots, X_n = x_n] = \nu(x_0) p(x_0, x_1) \cdots p(x_{n-1}, x_n) \quad \text{für alle } n \in \mathbb{N}, x_i \in S.$$

Für eine Wahrscheinlichkeitsverteilung  $\mu$  mit Massenfunktion  $\mu(x) = \mu(\{x\})$  und eine stochastische Matrix  $p$  auf  $S$  setzen wir

$$(\mu p)(y) := \sum_{x \in S} \mu(x) p(x, y), \quad (y \in S),$$

d.h.  $\mu p$  ist der Zeilenvektor, den wir erhalten, wenn wir den Zeilenvektor  $(\mu(x))_{x \in S}$  von links an die Matrix  $p$  multiplizieren.

**Lemma 3.6.** i) Die Verteilung zur Zeit  $n$  einer Markov-Kette mit Startverteilung  $\nu$  und Übergangsmatrix  $p$  ist  $\nu p^n$ .

ii) Gilt  $\nu p = \nu$ , dann folgt  $X_n \sim \nu$  für alle  $n \in \mathbb{N}$ . (»Stationarität«)

*Beweis.* i) Wie im Beweis von Satz 2.4 erhalten wir

$$P[X_n = y \mid X_0 = x] = p^n(x, y)$$

für alle  $n \in \mathbb{N}$  und  $x, y \in S$  mit  $P[X_0 = x] \neq 0$ , und damit:

$$\begin{aligned} P[X_n = y] &= \sum_{\substack{x \in S \\ P[X_0 = x] \neq 0}} P[X_n = y \mid X_0 = x] P[X_0 = x] \\ &= \sum_{\substack{x \in S \\ \nu(x) \neq 0}} p^n(x, y) \nu(x) = (\nu p^n)(y). \end{aligned}$$

ii) Aus  $\nu p = \nu$  folgt  $\nu p^n = \nu$  für alle  $n \in \mathbb{N}$ . □

**Definition.** i) Eine Wahrscheinlichkeitsverteilung  $\mu$  auf  $S$  heißt **Gleichgewichtsverteilung** (oder **stationäre Verteilung**) der Übergangsmatrix  $p$ , falls  $\mu p = \mu$ , d.h. falls:

$$\sum_{x \in S} \mu(x) p(x, y) = \mu(y) \quad \text{für alle } y \in S.$$

ii)  $\mu$  erfüllt die **Detailed Balance-Bedingung** bzgl. der Übergangsmatrix  $p$ , falls gilt:

$$\mu(x) p(x, y) = \mu(y) p(y, x) \quad \text{für alle } x, y \in S \quad (3.4.1)$$

**Satz 3.7.** Erfüllt  $\mu$  die Detailed Balance-Bedingung (3.4.1), dann ist  $\mu$  eine Gleichgewichtsverteilung von  $p$ .

*Beweis.* Aus der Detailed Balance-Bedingung folgt:

$$\sum_{x \in S} \mu(x) p(x, y) = \sum_{x \in S} \mu(y) p(y, x) = \mu(y).$$

□

**Bemerkung.** Bei Startverteilung  $\mu$  gilt

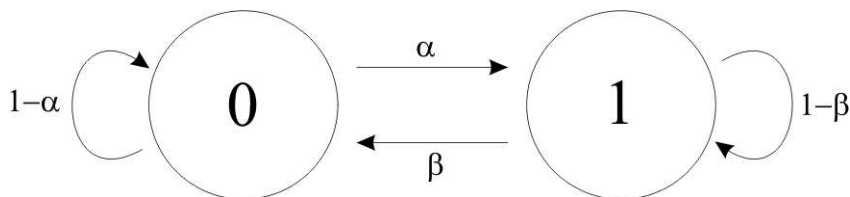
$$\mu(x) p(x, y) = P[X_0 = x, X_1 = y] \quad \text{»Fluss von } x \text{ nach } y\ll$$

DETAILED BALANCE:	$\mu(x) p(x, y)$	=	$\mu(y) p(y, x)$
	»Fluss von $x$ nach $y\ll$	=	»Fluss von $y$ nach $x\ll$
GLEICHGEWICHT:	$\sum_{x \in S} \mu(x) p(x, y)$	=	$\sum_{x \in S} \mu(y) p(y, x)$
	»Gesamter Fluss nach $y\ll$	=	»Gesamter Fluss von $y\ll$

**Beispiele.** a) MARKOV-KETTE AUF  $S = \{0, 1\}$ :

Seien  $\alpha, \beta \in [0, 1]$  und

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



Dann ist die Gleichgewichtsbedingung  $\mu p = \mu$  äquivalent zu den folgenden Gleichungen:

$$\begin{aligned} \mu(0) &= \mu(0) (1 - \alpha) + \mu(1) \beta, \\ \mu(1) &= \mu(0) \alpha + \mu(1) (1 - \beta). \end{aligned}$$

Da  $\mu$  eine Wahrscheinlichkeitsverteilung ist, sind beide Gleichungen äquivalent zu

$$\beta (1 - \mu(0)) = \alpha \mu(0).$$

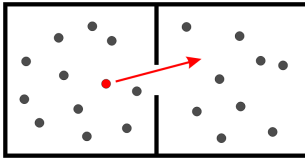
Die letzte Gleichung ist äquivalent zur Detailed Balance-Bedingung (3.4.1). Falls  $\alpha + \beta > 0$  gilt, ist  $\mu = \left( \frac{\beta}{\alpha + \beta}, \frac{\alpha}{\alpha + \beta} \right)$  die eindeutige Gleichgewichtsverteilung und erfüllt die Detailed Balance-Bedingung. Falls  $\alpha = \beta = 0$  gilt, ist jede Wahrscheinlichkeitsverteilung  $\mu$  eine Gleichgewichtsverteilung mit Detailed Balance-Bedingung.

b) ZYKLISCHER RANDOM WALK: Sei  $S = \mathbb{Z}/n\mathbb{Z}$  ein diskreter Kreis, und

$$p(k, k+1) = p, \quad p(k, k-1) = 1-p.$$

Die Gleichverteilung  $\mu(x) = \frac{1}{n}$  ist ein Gleichgewicht. Die Detailed Balance-Bedingung ist dagegen nur für  $p = \frac{1}{2}$ , d.h. im symmetrischen Fall, erfüllt.

c) EHRENFEST-MODELL:



Sei  $S = \{0, 1, \dots, n\}$ ,

$$p(k, k-1) = \frac{k}{n}, \quad p(k, k+1) = \frac{n-k}{n}.$$

Man kann erwarten, daß sich im Gleichgewicht jede Kugel mit Wahrscheinlichkeit  $\frac{1}{2}$  in jeder der beiden Urnen befindet. Tatsächlich erfüllt die Binomialverteilung  $\mu(k) = \binom{n}{k} 2^{-n}$  mit Parameter  $p = \frac{1}{2}$  die Detailed Balance-Bedingung:

$$\mu(k-1) p(k-1, k) = \mu(k) p(k, k-1) \quad k = 1, \dots, n$$

ist äquivalent zu

$$2^{-n} \frac{n!}{(k-1)!(n-(k-1))!} \frac{n-(k-1)}{n} = 2^{-n} \frac{n!}{k!(n-k)!} \frac{k}{n} \quad k = 1, \dots, n$$

d) RANDOM WALKS AUF GRAPHEN:

Sei  $(V, E)$  ein endlicher Graph,  $S = V$  die Menge der Knoten.

- Sei

$$p(x, y) = \begin{cases} \frac{1}{\deg(x)} & \text{falls } \{x, y\} \in E, \\ 0 & \text{sonst.} \end{cases}$$

Die Detailed Balance-Bedingung lautet in diesem Fall:

$$\mu(x) p(x, y) = \mu(y) p(y, x).$$

Sie ist erfüllt, falls

$$\mu(x) = c \deg(x)$$

gilt, wobei  $c$  eine Konstante ist. Damit  $\mu$  eine Wahrscheinlichkeitsverteilung ist, muss  $c$  so gewählt werden, dass gilt:

$$\sum_{x \in B} \deg(x) = 2 |E|.$$

Somit ist die Gleichgewichtsverteilung:

$$\mu(x) = \frac{\deg(x)}{2|E|}.$$

- Sei  $\Delta := \max_{x \in V} \deg(x)$ ,

$$p(x, y) = \begin{cases} \frac{1}{\Delta} & \text{falls } \{x, y\} \in E, \\ 1 - \frac{\deg(x)}{\Delta} & \text{sonst.} \end{cases}$$

Es gilt  $p(x, y) = p(y, x)$  und somit ist die Gleichverteilung auf  $V$  die stationäre Verteilung.

Ist  $\deg(x)$  konstant, dann stimmen die Random Walks in beiden Beispielen überein, und die Gleichverteilung ist ein Gleichgewicht.

Im nächsten Abschnitt zeigen wir:

**Satz** (Konvergenzsatz für Markov-Ketten). *Ist  $S$  endlich, und  $p$  eine irreduzible und aperiodische stochastische Matrix mit Gleichgewicht  $\mu$ , dann gilt für alle Wahrscheinlichkeitsverteilungen  $\nu$  auf  $S$ :*

$$\lim_{n \rightarrow \infty} (\nu p^n)(x) = \mu(x) \quad \text{für alle } x \in S.$$

Aufgrund des Konvergenzsatzes können wir Stichproben von einer Wahrscheinlichkeitsverteilung  $\mu$  näherungsweise erzeugen, indem wir eine Markov-Kette  $X_n$  mit Gleichgewicht  $\mu$  simulieren, und für großes  $n$  auswerten. Wie findet man eine Markov-Kette mit einer vorgegebenen stationären Verteilung?

## Konstruktion von Markov-Ketten mit vorgegebenen Gleichgewichtsverteilungen

### 1. DIE METROPOLIS-KETTE:

Sei  $q(x, y)$  eine symmetrische stochastische Matrix, d.h.  $q(x, y) = q(y, x)$  für alle  $x, y \in S$ .

Dann erfüllt die Gleichverteilung die Detailed Balance-Bedingung (3.4.1). Sei nun  $\mu$  eine

beliebige Wahrscheinlichkeitsverteilung auf  $S$  mit  $\mu(x) > 0$  für alle  $x \in S$ . Wie können wir die Übergangsmatrix  $q$  so modifizieren, dass die Detailed Balance-Bedingung bzgl.  $\mu$  erfüllt ist?

**Algorithmus 3.8** (Metropolis-Algorithmus (Update  $x \rightarrow y$ )). *schlage Übergang  $x \rightarrow y$  mit Wahrscheinlichkeit  $q(x, y)$  vor  
akzeptiere Übergang mit Wahrscheinlichkeit  $\alpha(x, y) \in [0, 1]$   
sonst verwerfe Vorschlag und bleibe bei  $x$*

ÜBERGANGSMATRIX:

$$p(x, y) := \begin{cases} \alpha(x, y) q(x, y) & \text{für } y \neq x, \\ 1 - \sum_{y \neq x} \alpha(x, y) q(x, y) & \text{für } y = x. \end{cases}$$

Die Detailed Balance-Bedingung lautet:

$$\mu(x) \alpha(x, y) = \mu(y) \alpha(y, x) \quad \text{für alle } x, y \in S.$$

Sie ist äquivalent dazu, dass

$$b(x, y) := \mu(x) \alpha(x, y)$$

symmetrisch in  $x$  und  $y$  ist. Was ist die größtmögliche Wahl von  $b(x, y)$ ?

Aus  $\alpha(x, y) \leq 1$  folgen

$$\begin{aligned} b(x, y) &\leq \mu(x), \\ b(x, y) &= b(y, x) \leq \mu(y), \end{aligned}$$

und somit

$$b(x, y) \leq \min(\mu(x), \mu(y)).$$

Der größtmögliche Wert  $b(x, y) = \min(\mu(x), \mu(y))$  entspricht gerade

$$\alpha(x, y) = \min\left(1, \frac{\mu(y)}{\mu(x)}\right) = \begin{cases} 1 & \text{falls } \mu(y) \geq \mu(x), \\ \frac{\mu(y)}{\mu(x)} & \text{falls } \mu(x) > \mu(y). \end{cases}$$

**Definition.** Die Markov-Kette mit Übergangsmatrix

$$p(x, y) = \min\left(1, \frac{\mu(y)}{\mu(x)}\right) \cdot q(x, y) \quad \text{für } y \neq x$$

heißt **Metropolis-Kette** mit Vorschlagsverteilung  $q(x, y)$  und Gleichgewicht  $\mu$ .



**Satz 3.9.**  $\mu$  erfüllt die Detailed Balance-Bedingung bzgl.  $p$ .

*Beweis.* siehe oben. □

2. DER GIBBS-SAMPLER Sei  $S = S_1 \times \dots \times S_d$  ein endlicher Produktraum,  $\mu(x_1, \dots, x_d)$  eine Wahrscheinlichkeitsverteilung auf  $S$  und

$$\mu_i(x_i \mid x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d) := \frac{\mu(x_1, \dots, x_d)}{\sum_{z \in S_i} \mu(x_1, \dots, x_{i-1}, z, x_{i+1}, \dots, x_d)}$$

die bedingte Verteilung der  $i$ -ten Komponente gegeben die übrigen Komponenten.

**Algorithmus 3.10** (Gibbs-Sampler (Update  $x \rightarrow y$ )).  $y := x$

**for**  $i := 1, \dots, d$  **do**

*update*  $y_i \sim \mu_i(\bullet \mid y_1, \dots, y_{i-1}, y_{i+1}, \dots, y_d)$

**end for**

**return**  $y$

ÜBERGANGSMATRIX:

$$p = p_d p_{d-1} \cdots p_1,$$

wobei

$$p_i(x, y) = \begin{cases} \mu_i(y_i \mid y_1, \dots, y_{i-1}, y_{i+1}, \dots, y_d) & \text{falls } y_k = x_k \text{ für alle } k \neq i, \\ 0 & \text{sonst.} \end{cases}$$

**Satz 3.11.** i)  $\mu$  erfüllt die Detailed Balance-Bedingung bzgl.  $p_i$  für alle  $i = 1, \dots, d$ .

ii)  $\mu$  ist ein Gleichgewicht von  $p$ .

*Beweis.* i) Der Beweis der ersten Aussage ist eine Übungsaufgabe.

ii) Nach der ersten Aussage ist  $\mu$  ein Gleichgewicht von  $p_i$  für alle  $i$ . Also gilt auch

$$\mu p = \mu p_d p_{d-1} \cdots p_1 = \mu.$$

□

**Bemerkung.** Zur Simulation von  $Y_n$ ,  $n \geq 0$ , genügt es, die Massenfunktion  $\mu(x)$  bis auf eine multiplikative Konstante zu kennen:

aus  $\mu(x) = C f(X)$  folgt

$$\alpha(x, y) = \min \left( 1, \frac{f(y)}{f(x)} \right) \quad \text{unabhängig von } C.$$

**Beispiel (Rucksackproblem).** Gegeben:

$$\begin{aligned}\omega_1, \dots, \omega_d &\in \mathbb{R}, \quad \text{»Gewichte«}, \\ v_1, \dots, v_d &\in \mathbb{R}, \quad \text{»Werte«}.\end{aligned}$$

Rucksack mit maximalem Gewicht  $b > 0$ , packe soviel Wert wie möglich ein.

$$\begin{aligned}S &= \{0, 1\}^d, && \text{alle Konfigurationen,} \\ S_b &= \{(z_1, \dots, z_d) \in S : \sum_{i=1}^d z_i \omega_i \leq b\}, && \text{zulässige Konfigurationen,} \\ & z_i = 1 : i\text{-ter Gegenstand im Rucksack.}\end{aligned}$$

RUCKSACKPROBLEM:

$$\text{maximiere } V(z) = \sum_{i=1}^d z_i v_i \text{ unter Nebenbedingung } z \in S_b.$$

Das Rucksackproblem ist *NP-vollständig*, insbesondere ist keine Lösung in  $O(d^k)$  Schritten für ein  $k \in \mathbb{N}$  bekannt.

STOCHASTISCHER ZUGANG: SIMULATED ANNEALING

Für  $\beta > 0$  betrachten wir die Wahrscheinlichkeitsverteilung

$$\mu_\beta(z) = \begin{cases} \frac{1}{Z_\beta} e^{\beta V(z)} & \text{für } z \in S_b, \\ 0 & \text{für } z \in S \setminus S_b, \end{cases}$$

auf  $S$ , wobei  $Z_\beta = \sum_{z \in S_b} e^{\beta V(z)}$  eine Konstante ist, die  $\mu$  zu einer Wahrscheinlichkeitsverteilung normiert. Für  $\beta = 0$  ist  $\mu_\beta$  die Gleichverteilung auf  $S_b$ . Für  $\beta \rightarrow \infty$  konvergiert  $\mu_\beta$  gegen die Gleichverteilung auf der Menge der globalen Maxima von  $V$ , denn:

$$\mu_\beta(z) = \frac{e^{\beta V(z)}}{Z_\beta} = \frac{1}{\sum_{y \in S_b} e^{\beta(V(y)-V(z))}} \rightarrow \begin{cases} 0 & \text{falls } V(z) \neq \max V, \\ \frac{1}{|\{y \mid V(y) = \max V\}|} & \text{falls } V(z) = \max V. \end{cases}$$

IDEE: Simuliere Stichprobe  $z$  von  $\mu_\beta$  für  $\beta$  groß ( $\beta \rightarrow \infty$ ). Dann ist  $V(z)$  *wahrscheinlich* nahe dem Maximalwert.

METROPOLIS-ALGORITHMUS: Sei  $x^+ := \max(x, 0)$  der Positivteil von  $x$ . Wir wählen als Vorschlagsmatrix die Übergangsmatrix

$$q(z, w) := \begin{cases} \frac{1}{d} & \text{falls } z_i \neq w_i \text{ für genau ein } i \in \{1, \dots, d\}, \\ 0 & \text{sonst,} \end{cases}$$

des Random Walks auf  $\{0, 1\}^d$ . Für die Akzeptanzwahrscheinlichkeit ergibt sich

$$\alpha_\beta(z, w) = \min \left( 1, \frac{\mu_\beta(w)}{\mu_\beta(z)} \right) = \begin{cases} e^{-\beta(V(z)-V(w))} & \text{für } z, w \in S_b, \\ 0 & \text{für } z \in S_b, w \notin S_b. \end{cases}$$

Der Vorschlag  $w$  wird also mit Wahrscheinlichkeit 1 akzeptiert, wenn  $V(w) \geq V(z)$  gilt – andernfalls wird der Vorschlag nur mit Wahrscheinlichkeit  $\exp -\beta (V(z) - V(w))$  akzeptiert.

**Algorithmus 3.12** (Simulation einer Markov-Kette mit Gleichgewicht  $\mu_\beta$ ). *initialisiere*  $z^{(0)} \in S_b$

```

for  $n = 1, 2, \dots$  do
   $z^{(n)} := w := z^{(n-1)}$ 
  erzeuge  $i \sim \text{Unif}\{1, \dots, d\}$ 
   $w_i := 1 - w_i$ 
  if  $w \in S_b$  then
    erzeuge  $u \sim \text{Unif}(0, 1)$ 
    if  $u \leq \alpha_\beta(z, w)$  then
       $z^{(n)} := w$ 
    end if
  end if
end for

```

**Algorithmus 3.13** (Simulated Annealing). *Wie Algorithmus 3.12 aber mit*  $\beta = \beta(n) \rightarrow \infty$  *für*  $n \rightarrow \infty$ .

**Bemerkung.** a) PHYSIKALISCHE INTERPRETATIONEN:

$\mu_\beta$  ist die Verteilung im thermodynamischen Gleichgewicht für die Energiefunktion  $H(z) = -V(z)$  bei der Temperatur  $T = 1/\beta$ . Der Grenzwert  $\beta \rightarrow \infty$  entspricht  $T \rightarrow 0$  (»simuliertes Abkühlen«).

b) Die beim Simulated Annealing-Verfahren simulierte zeitlich inhomogene Markov-Kette findet im allgemeinen nicht das globale Maximum von  $V$ , sondern kann in lokalen Maxima »steckenbleiben«. Man kann zeigen, daß die Verteilung der Markov-Kette zur Zeit  $n$  gegen die Gleichverteilung auf den Maximalstellen konvergiert, falls  $\beta(n)$  nur sehr langsam (logarithmisch) gegen  $+\infty$  geht. In praktischen Anwendungen wird der Algorithmus aber in der Regel mit einem schnelleren »Cooling schedule«  $\beta(n)$  verwendet. Das Auffinden eines globalen Maximums ist dann nicht garantiert – trotzdem erhält man ein oft nützliches *heuristisches* Verfahren.

### 3.5 Konvergenz ins Gleichgewicht

Sei  $S = \{x_1, \dots, x_m\}$  eine endliche Menge, und

$$\text{WV}(S) := \{\mu = (\mu(x_1), \dots, \mu(x_m)) \mid \mu(x_i) \geq 0, \sum_{i=1}^m \mu(x_i) = 1\} \subseteq \mathbb{R}^m$$

die Menge aller Wahrscheinlichkeitsverteilungen auf  $S$ . Geometrisch ist  $\text{WV}(S)$  ein Simplex im  $\mathbb{R}^m$ . Wir führen nun einen Abstands begriff auf  $\text{WV}(S)$  ein:

**Definition.** Die *Variationsdistanz* zweier Wahrscheinlichkeitsverteilungen  $\mu, \nu$  auf  $S$  ist:

$$d_{TV}(\mu, \nu) := \frac{1}{2} \|\mu - \nu\|_1 = \frac{1}{2} \sum_{x \in S} |\mu(x) - \nu(x)|.$$

**Bemerkung.** a) Für alle  $\mu, \nu \in \text{WV}(S)$  gilt:

$$d_{TV}(\mu, \nu) \leq \frac{1}{2} \sum_{x \in S} (\mu(x) + \nu(x)) = 1.$$

b) Seien  $\mu, \nu$  Wahrscheinlichkeitsverteilungen und  $A := \{x \in S \mid \mu(x) \geq \nu(x)\}$ . Dann gilt:

$$d_{TV}(\mu, \nu) = \sum_{x \in A} (\mu(x) - \nu(x)) = \max_{A \subseteq S} |\mu(A) - \nu(A)|.$$

Der Beweis dieser Aussage ist eine Übungsaufgabe.

Wir betrachten im folgenden eine stochastische Matrix  $p(x, y)$ , ( $x, y \in S$ ), mit Gleichgewicht  $\mu$ . Die Verteilung einer Markov-Kette mit Startverteilung  $\nu$  und Übergangsmatrix  $p$  zur Zeit  $n$  ist  $\nu p^n$ . Um Konvergenz ins Gleichgewicht zu zeigen, verwenden wir die folgende Annahme: **MINORISIERUNGSBEDINGUNG:** Es gibt ein  $\delta \in (0, 1]$  und ein  $r \in \mathbb{N}$ , so dass für alle  $x, y \in S$  gilt:

$$p^r(x, y) \geq \delta \cdot \mu(y). \quad (3.5.1)$$

**Satz 3.14.** Gilt die Minorisierungsbedingung (3.5.1), dann konvergiert  $\nu p^n$  für jede Startverteilung  $\nu$  exponentiell schnell gegen  $\mu$ . Genauer gilt für alle  $n \in \mathbb{N}$  und  $\nu \in \text{WV}(S)$ :

$$d_{TV}(\nu p^n, \mu) \leq (1 - \delta)^{\lfloor n/r \rfloor}.$$

**Bemerkung.** Insbesondere ist  $\mu$  das *eindeutige* Gleichgewicht: Betrachte eine beliebige Wahrscheinlichkeitsverteilung  $\nu$  mit  $\nu p = \nu$ . Dann folgt für  $n \rightarrow \infty$ :

$$d_{TV}(\nu, \mu) = d_{TV}(\nu p^n, \mu) \longrightarrow 0,$$

also  $d_{TV}(\mu, \nu) = 0$ , und somit  $\mu = \nu$ .

*Beweis.* 1. Durch die Zerlegung

$$p^r(x, y) = \delta \mu(y) + (1 - \delta) q(x, y)$$

der  $r$ -Schritt-Übergangswahrscheinlichkeiten wird eine *stochastische* Matrix  $q$  definiert, denn

(i) Aus der Minorisierungsbedingung (3.5.1) folgt  $q(x, y) \geq 0$  für alle  $x, y \in S$ .

(ii) Aus  $\sum_{y \in S} p^r(x, y) = 1$ ,  $\sum_{y \in S} \mu(y) = 1$  folgt  $\sum_{y \in S} q(x, y) = 1$  für alle  $x \in S$ .

Wir setzen im folgenden  $\lambda := 1 - \delta$ . Dann gilt für alle  $\nu \in \text{WV}(S)$ :

$$\nu p^r = (1 - \lambda) \mu + \lambda \nu q. \quad (3.5.2)$$

2. Wir wollen mit vollständiger Induktion zeigen:

$$\nu p^{kr} = (1 - \lambda^k) \mu + \lambda^k \nu q^k \quad \text{für alle } k \geq 0, \quad \nu \in \text{WV}(S). \quad (3.5.3)$$

Für  $k = 0$  ist die Aussage offensichtlich wahr. Gilt (3.5.3) für ein  $k \geq 0$ , dann erhalten wir durch Anwenden von Gleichung (3.5.2) auf  $\tilde{\nu} p^r$  mit  $\tilde{\nu} = \nu q^k$ :

$$\begin{aligned} \nu p^{(k+1)r} &= \nu p^{kr} p^r \\ &= ((1 - \lambda^k) \mu + \lambda^k \underbrace{\nu q^k}_{=\tilde{\nu}}) p^r \\ &= (1 - \lambda^k) \underbrace{\mu p^r}_{=\mu} + (1 - \lambda) \lambda^k \mu + \lambda^{k+1} \nu q^k q \\ &= (1 - \lambda^{k+1}) \mu + \lambda^{k+1} \nu q^{k+1}. \end{aligned}$$

3. Für  $n \in \mathbb{N}$ ,  $n = kr + i$ , ( $k \in \mathbb{N}$ ,  $0 \leq i < r$ ), folgt:

$$\nu p^n = \nu p^{kr} p^i = (1 - \lambda^k) \underbrace{\mu p^i}_{=\mu} + \lambda^k \nu q^k p^i,$$

also

$$\nu p^n - \mu = \lambda^k (\nu q^k p^i - \mu) \quad \text{für alle } \nu \in \text{WV}(S),$$

und damit

$$d_{TV}(\nu p^n, \mu) = \frac{1}{2} \|\nu p^n - \mu\|_1 = \lambda^k d_{TV}(\nu q^k p^i, \mu) \leq \lambda^k$$

nach der letzten Bemerkung.

□

Welche Übergangsmatrizen erfüllen die Minorisierungsbedingung?

**Definition.** i) Die stochastische Matrix  $p$  heißt **irreduzibel**, falls es für alle  $x, y \in S$  ein  $n \in \mathbb{N}$  gibt, so dass  $p^n(x, y) > 0$  gilt.

ii) Die **Periode** von  $x \in S$  ist definiert als

$$\text{Periode}(x) := \text{ggT}(\underbrace{\{n \in \mathbb{N} \mid p^n(x, x) > 0\}}_{=:R(x)}).$$

**Lemma 3.15.** i) Falls  $p$  irreduzibel ist, gilt  $\text{Periode}(x) = \text{Periode}(y)$  für alle  $x, y \in S$ .

ii) Falls  $p$  irreduzibel und aperiodisch (d.h.  $\text{Periode}(x) = 1$  für alle  $x \in S$ ) ist, gibt es ein  $r > 0$ , so dass  $p^r(x, y) > 0$  für alle  $x, y \in S$  gilt.

*Beweis.* Seien  $x, y \in S$ .

i) Sei  $p$  irreduzibel. Dann gibt es ein  $s$  und ein  $t \in \mathbb{N}$ , so dass gilt:

$$p^s(x, y) > 0 \quad \text{und} \quad p^t(y, x) > 0.$$

Für  $a := s + t$  folgt:

- $p^a(x, x) \geq p^s(x, y) p^t(y, x) > 0$ , also  $a \in R(x)$ .
- $p^{n+a}(x, x) \geq p^s(x, y) p^n(y, y) p^t(y, x) > 0$  für alle  $n \in R(y)$ , also  $n + a \in R(x)$  für alle  $n \in R(y)$ .

$\text{Periode}(x)$  ist ein gemeinsamer Teiler von  $R(x)$ , somit Teiler von  $a$  und  $n + a$ , also auch von  $n$  für alle  $n \in R(y)$ . Daher ist  $\text{Periode}(x)$  ein gemeinsamer Teiler von  $R(y)$  und somit gilt:

$$\text{Periode}(x) \leq \text{Periode}(y).$$

» $\geq$ « wird analog gezeigt. Es folgt:

$$\text{Periode}(x) = \text{Periode}(y).$$

ii)  $R(x)$  ist abgeschlossen unter Addition, denn falls  $s, t \in R(x)$  ist, gilt:

$$p^{s+t}(x, x) \geq p^s(x, x) p^t(x, x) > 0,$$

und somit  $s + t \in R(x)$ . Da  $p$  aperiodisch ist, folgt  $\text{ggT}(R(x)) = 1$  für alle  $x \in S$ . Nach einem Satz der Zahlentheorie gilt:

Da  $R(x)$  additiv abgeschlossen, gibt es für alle  $x$  ein  $r(x) \in \mathbb{N}$  mit  $n \in R(x)$  für alle  $n \geq r(x)$ .

$n \in R(x)$  impliziert  $p^n(x, x) > 0$ . Da  $p$  irreduzibel ist, folgt, dass es für alle  $x, y$  ein  $r(x, y) \in \mathbb{N}$  gibt, so dass gilt:

$$p^n(x, y) > 0 \quad \text{für alle } n \geq r(x, y).$$

Für  $r \geq \max_{x, y \in S} r(x, y)$  folgt dann  $p^r(x, y) > 0$  für alle  $x, y \in S$ .

□

**Satz 3.16** (Konvergenzsatz für *endliche* Markov-Ketten). *Ist  $p$  irreduzibel und aperiodisch mit Gleichgewicht  $\mu$ , dann gilt:*

$$\lim_{n \rightarrow \infty} d_{TV}(\nu p^n, \mu) = 0 \quad \text{für alle } \nu \in \text{WV}(S).$$

*Beweis.* Da  $p$  irreduzibel und aperiodisch ist, gibt es ein  $r \in \mathbb{N}$  mit:

$$p^r(x, y) > 0 \quad \text{für alle } x, y \in S.$$

Daher gibt es ein  $r \in \mathbb{N}$  und ein  $\delta > 0$ , so dass gilt:

$$p^r(x, y) > \delta \mu(y) \quad \text{für alle } x, y \in S,$$

(z.B.  $\delta := \min_{x, y \in S} p^r(x, y)$ ). Mit Satz 3.14 folgt die Behauptung. □

**Beispiel** (Metropolis-Kette). Sei  $S$  endlich,  $\mu(x) > 0$  für alle  $x \in S$ , nicht konstant, und  $q(x, y)$  irreduzibel. Dann ist  $p(x, y)$  irreduzibel und aperiodisch. Somit folgt die Konvergenz ins Gleichgewicht nach Satz 3.16, allerdings evtl. sehr langsam!

#### ANWENDUNG: MARKOV-CHAIN-MONTE CARLO-VERFAHREN

Sei  $\mu \in \text{WV}(S)$ ,  $f: S \rightarrow \mathbb{R}$ .

GESUCHT:

$$\theta = E_\mu[f],$$

MARKOV-CHAIN-MONTE CARLO-SCHÄTZER:

$$\hat{\theta}_{n,b} = \frac{1}{n} \sum_{k=b+1}^{b+n} f(X_k),$$

wobei  $b \in \mathbb{N}$  eine feste Konstante (»burn-in-Zeit«) und  $(X_k)_{k \in \mathbb{N}}$  irreduzible Markov-Ketten mit Gleichgewicht  $\mu$  sind.

**Satz** (Ergodensatz / Gesetz der großen Zahlen für Markov-Ketten). : Für alle  $b \in \mathbb{N}$  gilt:

$$\lim_{n \rightarrow \infty} \hat{\theta}_{n,b} = \theta \quad \text{mit Wahrscheinlichkeit 1,}$$

*Beweis.* siehe Vorlesung »Stochastische Prozesse«. □

Die Analyse des Schätzfehler ist im Allgemeinen diffizil!



# Kapitel 4

## Iterationsverfahren für lineare und nichtlineare Gleichungssysteme

LITERATUR ZU NUMERISCHER MATHEMATIK:

- Hanke-Bourgeois: »Grundlagen der numerischen Mathematik«
- Deuflhard/Hohmann: »Numerische Mathematik«
- Hämmerlin/Hoffmann: »Numerische Mathematik«
- Quarteroni/Sacco/Saleri: »Numerische Mathematik 1«

**Beispiel. GESUCHT:** numerische Lösung  $x^* \in (0, \frac{\pi}{2})$  von  $\tan(x) - 2x = 0$ .

1. ANSATZ:

$$\tan(x) - 2x = 0 \quad \Leftrightarrow \quad x = \frac{1}{2} \tan(x) =: \phi(x) \quad \text{»Fixpunktgleichung«.}$$

ITERATIONSVERFAHREN:

$$x^{(n+1)} = \phi(x^{(n)}).$$

Konvergiert das Iterationsverfahren?

- falls  $x^{(0)} < x^*$  konvergiert  $x^{(n)}$  gegen 0.
- falls  $x^{(0)} > x^*$  ist  $x^{(n)}$  monoton wachsend bis  $x^{(n)} > \frac{\pi}{2}$ , danach konvergiert  $x^{(n)}$  gegen 0.

2. ANSATZ:

$$\tan(x) - 2x = 0 \quad \Leftrightarrow \quad x = \arctan(2x) =: \tilde{\phi}(x)$$

$x^{(n)}$  konvergiert gegen  $x^*$  für alle  $x^0 \in \mathbb{R}^+$ .

3. ANSATZ: NEWTON-VERFAHREN

$$f(y) := \tan(y) - 2y \stackrel{!}{=} 0.$$

IDEE: Lineare Approximation von  $f$  (durch Tangente):

$$f(y) \approx f(x) + (y - x) f'(x).$$

Die rechte Seite verschwindet für

$$y := x - \frac{f(x)}{f'(x)} =: \phi(x).$$

Es gilt:

$$f(x^*) = 0 \quad \Leftrightarrow \quad \phi(x^*) = x^*.$$

Dies motiviert die *Fixpunktiteration* :

$$x^{(n+1)} = \phi(x^{(n)}).$$

- falls  $x^{(0)} \in (x^*, \frac{\pi}{2})$  oder  $x^{(0)} \in (\phi^{-1}(\frac{\pi}{2}), x^*)$  konvergiert  $x^{(n)}$  sehr schnell gegen  $x^*$ ,
- falls  $x^{(0)} \leq \phi^{-1}(\frac{\pi}{2})$  konvergiert  $x^{(n)}$  nicht gegen  $x^*$ .

Das Beispiel zeigt, dass nicht jede Fixpunktiteration konvergiert. Die Konvergenz kann (abhängig vom Startwert) gegen verschiedene Fixpunkte erfolgen. Konvergenz und Konvergenzgeschwindigkeit hängen wesentlich von der Art der Darstellung einer Gleichung als Fixpunktgleichung ab.

## 4.1 Der Banachsche Fixpunktsatz

ALLGEMEINER RAHMEN:  $(X, d)$  ein metrischer Raum,  $\phi: X \rightarrow X$ .

FIXPUNKTITERATION:

$$x^{(n+1)} = \phi(x^{(n)}).$$

**Satz 4.1** (Banachscher Fixpunktsatz). *Ist  $(X, d)$  vollständig und  $\phi$  eine Kontraktion, d.h. es existiert eine Konstante  $L \in (0, 1)$  mit*

$$d(\phi(x), \phi(y)) \leq L \cdot d(x, y) \quad \text{für alle } x, y \in X, \quad (4.1.1)$$

dann gilt:

1.  $\phi$  hat genau einen Fixpunkt  $x^* \in X$ .

2. Die Fixpunktiteration  $x^{(n+1)} = \phi(x^{(n)})$  konvergiert für jeden Startwert  $x^{(0)} \in X$  gegen  $x^*$ , und für alle  $n \in \mathbb{N}$  gilt:

$$d(x^{(n)}, x^*) \leq L d(x^{(n-1)}, x^*) \quad \text{»Monotonie«} \quad (4.1.2)$$

$$d(x^{(n)}, x^*) \leq \frac{L^n}{1-L} d(x^{(1)}, x^{(0)}) \quad \text{»A priori-Schranke«} \quad (4.1.3)$$

$$d(x^{(n)}, x^*) \leq \frac{L}{1-L} d(x^{(n)}, x^{(n-1)}) \quad \text{»A posteriori-Schranke«} \quad (4.1.4)$$

*Beweis.* i) Aus (4.1.1) folgt durch Induktion:

$$d(x^{(n+1)}, x^{(n)}) \leq L^n d(x^{(1)}, x^{(0)}) \quad \text{für alle } n \geq 0, \quad (4.1.5)$$

also nach Dreiecksungleichung

$$d(x^{(m)}, x^{(n)}) \leq \sum_{i=n}^{m-1} d(x^{(i+1)}, x^{(i)}) \leq \sum_{i=n}^{m-1} L^i d(x^{(1)}, x^{(0)}) \leq \frac{L^n}{1-L} d(x^{(1)}, x^{(0)}) \quad (4.1.6)$$

für alle  $m \geq n \geq 0$ . Wegen  $L < 1$  ist  $x^{(n)}$  eine Cauchy-Folge, also existiert wegen der vorausgesetzten Vollständigkeit

$$x^* := \lim_{n \rightarrow \infty} x^{(n)}.$$

Zudem gilt wegen der Stetigkeit von  $\phi$  und Gleichung (4.1.5):

$$d(x^*, \phi(x^*)) = \lim_{n \rightarrow \infty} d(x^{(n)}, \phi(x^{(n)})) = \lim_{n \rightarrow \infty} d(x^{(n)}, x^{(n+1)}) = 0,$$

d.h.  $x^*$  ist Fixpunkt. Ist  $y^*$  ein weiterer Fixpunkt, dann gilt:

$$d(x^*, y^*) = d(\phi(x^*), \phi(y^*)) \leq L d(x^*, y^*),$$

also  $d(x^*, y^*) = 0$ , d.h.  $x^* = y^*$  und  $x^*$  ist eindeutiger Fixpunkt.

ii) 1. Es gilt:

$$d(x^{(n)}, x^*) = d(\phi(x^{(n-1)}), \phi(x^*)) \leq L d(x^{(n-1)}, x^*).$$

2.

$$d(x^{(n)}, x^*) = \lim_{m \rightarrow \infty} d(x^{(n)}, x^{(m)}) \leq \frac{L^n}{1-L} d(x^{(1)}, x^{(0)}).$$

3. Analog zu (4.1.6) ergibt sich:

$$d(x^{(m)}, x^{(n)}) \leq \frac{L}{1-L} d(x^{(n)}, x^{(n-1)}) \quad \text{für alle } m \geq n \geq 0.$$

Die Behauptung folgt für  $m \rightarrow \infty$ .

□

NOTATION:  $d\phi(x)$  bezeichnet die Jacobi-Matrix.

**Korollar.** Sei  $C \subseteq \mathbb{R}^n$  abgeschlossen und konvex und  $\phi: C \rightarrow C$  einmal stetig differenzierbar.  
Gilt

$$L := \sup_{x \in C} \underbrace{\|d\phi(x)\|_2}_{l^2\text{-Matrixnorm}} = \sup_{x \in C} \sup_{v \in \mathbb{R}^n \setminus \{0\}} \frac{\|d\phi(x)(v)\|_2}{\|v\|_2} < 1,$$

dann hat  $\phi$  genau einen Fixpunkt  $x^*$  und  $x^{(n)}$  konvergiert gegen  $x^*$ .

*Beweis.* Für  $x, y \in C$  gilt  $[x, y] \subseteq C$ , also mit  $\gamma(t) := (1-t)x + ty$ :

$$\begin{aligned} \|\phi(y) - \phi(x)\|_2 &= \|\phi(\gamma(1)) - \phi(\gamma(0))\|_2 \\ &= \left\| \int_0^1 \frac{d}{dt} \phi(\gamma(t)) dt \right\|_2 \\ &\leq \int_0^1 \|d\phi(\gamma(t))(\gamma'(t))\|_2 dt \\ &\leq \int_0^1 L \|\gamma'(t)\|_2 dt = L \|y - x\|_2, \end{aligned}$$

d.h.  $\phi$  ist eine Kontraktion bzgl.  $d(x, y) = \|y - x\|_2$ . □

**Bemerkung** (Lokale Konvergenz). Ist  $D \subseteq \mathbb{R}^n$  offen,  $\phi \in C^1(D)$  mit Fixpunkt  $x^*$  und gilt  $\|d\phi(x^*)\|_2 < 1$ , dann existiert ein  $\varepsilon > 0$  mit:

$$\|d\phi(x)\|_2 \leq L < 1 \quad \text{für alle } x \in U_\varepsilon(x^*).$$

Nach dem Korollar konvergiert also  $x^{(n)}$  gegen  $x^*$  für alle  $x^{(0)}$  mit  $\|x^{(0)} - x^*\|_2 \leq \varepsilon$  (»lokale Konvergenz«).

**Beispiele.** a) GLEICHUNG VON OBEN:

$$\tan(x) - 2x = 0.$$

1.  $\phi(x) = \frac{1}{2} \tan(x)$ . Dann ist

$$\phi'(x) = \frac{1}{2} (1 + (\tan(x))^2) = \begin{cases} < 1 & \text{für } x < \frac{\pi}{4}, \\ > 1 & \text{für } \frac{\pi}{4} < x < \frac{\pi}{2}. \end{cases}$$

- für  $t < \frac{\pi}{4}$  ist  $\phi$  eine Kontraktion auf  $[0, t]$  und  $x^{(n)}$  konvergiert gegen 0.
- $\phi$  ist keine Kontraktion auf  $(\frac{\pi}{4}, \frac{\pi}{2})$ .

2.  $\tilde{\phi}(x) = \arctan(2x)$ . Dann gilt:

$$\tilde{\phi}'(x) = \frac{2}{1+x^2} < 1 \quad \Leftrightarrow \quad x > 1.$$

- für alle  $\varepsilon > 0$  ist  $\tilde{\phi}$  eine Kontraktion auf  $[1 + \varepsilon, \infty)$  und  $x^{(n)}$  konvergiert gegen  $x^* \in [1, \infty)$ .

b) NICHTLINEARES GLEICHUNGSSYSTEM: Sei  $c \in \mathbb{R}$  eine feste Konstante.

$$\begin{aligned} x_1 &= c \cos(x_1) - c \sin(x_2), \\ x_2 &= c \cos(x_1) - 2c \sin(x_2), \\ \phi(x) &= c \begin{pmatrix} \cos(x_1) - \sin(x_2) \\ \cos(x_1) - 2 \sin(x_2) \end{pmatrix}, \\ d\phi(x) &= -c \begin{pmatrix} \sin(x_1) & \cos(x_2) \\ \sin(x_1) & 2 \cos(x_2) \end{pmatrix}. \end{aligned}$$

Es gilt:

$$\|d\phi(x)\|_F = |c| \sqrt{2 \sin^2(x_1) + 5 \cos^2(x_2)} \leq \sqrt{7} |c|.$$

Damit folgt:

$$\|d\phi(x)\|_2 = \sup_{v \neq 0} \frac{\|d\phi(x)(v)\|_2}{\|v\|_2} \leq \sup_{v \neq 0} \frac{\|d\phi(x)\|(v)\|_F \cdot \|v\|_2}{\|v\|_2} \leq \|d\phi(x)\|_F \leq \sqrt{7} |c|.$$

Somit konvergiert die Fixpunktiteration für  $|c| < 1/\sqrt{7}$  mit  $L = \sqrt{7} |c|$ . Zum Beispiel für  $c = 0,4$  mit Startwert  $x^{(0)} = (0, 0)$ .

ZIEL: Für den Approximationsfehler soll gelten:

$$\varepsilon_n = \|x^{(n)} - x^*\|_2 \stackrel{!}{\leq} 5 \cdot 10^{-5} =: \delta.$$

- A PRIORI-ABSCHÄTZUNG:

$$\varepsilon_n \leq \frac{L^n}{1-L} \|x^{(1)} - x^{(0)}\|_2 \stackrel{!}{\leq} \delta.$$

Die Ungleichung ist erfüllt, falls gilt:

$$n \geq \log_{L^{-1}} \frac{\|x^{(1)} - x^{(0)}\|_2}{(1-L)\delta} \approx 153,6.$$

- A POSTERIORI-ABSCHÄTZUNG:

$$\frac{L}{1-L} \|x^{(n)} - x^{(n-1)}\|_2 \stackrel{!}{\leq} \delta.$$

Die Ungleichung ist schon für  $n = 39$  erfüllt!

- c) GLEICHGEWICHTE VON MARKOV-KETTEN:

Sei  $S$  endlich, und  $p(x, y)$ , ( $x, y \in S$ ) eine stochastische Matrix. Ein Gleichgewicht  $\mu$  von  $p$  ist ein Fixpunkt der Abbildung

$$\begin{aligned} \phi: \text{WV}(S) &\rightarrow \text{WV}(S), \\ \phi(\mu) &:= \mu p. \end{aligned}$$

Die Menge  $\text{WV}(S)$  ist ein Simplex im  $\mathbb{R}^{|S|}$ , also kompakt. Existiert ein  $L \in (0, 1)$  mit

$$d_{TV}(\mu p, \nu p) \leq L \cdot d_{TV}(\mu, \nu) \quad \text{für alle } \mu, \nu \in \text{WV}(S), \quad (4.1.7)$$

dann existiert nach dem Banachschen Fixpunktsatz 4.1 genau ein Gleichgewicht  $\mu$  von  $p$ , und es gilt:

$$d_{TV}(\mu p^n, \nu) \leq \frac{L^n}{1-L} \cdot d_{TV}(\mu, \nu) \leq \frac{L^n}{1-L} \quad \text{für alle } \mu, \nu \in \text{WV}(S).$$

Im allgemeinen gilt Gleichung (4.1.7) jedoch nur mit  $L = 1$ !

## 4.2 Klassische Iterationsverfahren für lineare Gleichungssysteme

Gegeben ist ein lineares Gleichungssystem

$$Ax = b, \quad \text{mit } b \in \mathbb{R}^d \text{ und invertierbarer Matrix } A \in \mathbb{R}^{d \times d}.$$

Gesucht ist die Lösung  $x^* \in \mathbb{R}^d$ .

Typisch in Anwendungen ist außerdem, dass  $d$  sehr groß ist und  $A$  »dünn besetzt« ist, d.h. dass nur wenige Einträge einer Zeile von  $A$  ungleich 0 sind.

## 4.2. KLASSISCHE ITERATIONSVERFAHREN FÜR LINEARE GLEICHUNGSSYSTEME 3

**Beispiel** (Poissongleichung). Gesucht ist  $u \in C(\overline{D}) \cap C^2(D)$  mit:

$$\Delta u(x) := \sum_{i=1}^n \frac{\partial^2 u}{\partial x_i^2}(x) = g(x) \quad \text{für alle } x \in D,$$

$$u(x) = 0 \quad \text{für alle } x \in \overline{D} \setminus D.$$

DISKRETE VERSION:

$$D \subseteq \mathbb{Z}^n \text{ endlich,}$$

$$\overline{D} := \{x \in \mathbb{Z}^n \mid \text{es gibt ein } y \in D : \|x - y\|_2 \leq 1\},$$

$$g = (g_x)_{x \in D} \in \mathbb{R}^{|\overline{D}|}.$$

Gesucht ist in diesem Fall  $u = (u_x)_{x \in \overline{D}}$  mit:

$$\sum_{i=1}^n [(u_{x+e_i} - u_x) - (u_x - u_{x-e_i})] = g_x \quad \text{für alle } x \in D,$$

$$u_x = 0 \quad \text{für alle } x \in \overline{D} \setminus D.$$

Wir erhalten ein lineares Gleichungssystem

$$A u = b$$

mit hoher Dimension  $d = |\overline{D}|$  und dünn besetzter Matrix:

$$A_{x,y} \neq 0 \quad \text{nur falls} \quad \|x - y\|_2 \leq 1.$$

In Anwendungen wie dieser sind die klassischen Eliminationsverfahren meist zu aufwändig ( $O(d^3)$ ), stattdessen verwendet man Fixpunktiterationen.

ALLGEMEINES ITERATIONSVERFAHREN:

$$A = P - N, \quad P, N \in \mathbb{R}^{d \times d},$$

wobei  $P$  eine invertierbare Matrix, deren Inverses »leicht« berechenbar ist.  $P$  wird **Präkonditionierer** genannt. Es gilt:

$$A x = b \quad \Leftrightarrow \quad P x = N x + b \quad \Leftrightarrow \quad x = T x + f =: \phi(x), \quad \text{wobei}$$

$$T := P^{-1} N = P^{-1} (P - A) = I - P^{-1} A, \quad f := P^{-1} b.$$

LINEARE FIXPUNKTITERATION:

$$x^{(k+1)} = T x^{(k)} + f.$$

RESIDUENDARSTELLUNG:

$$x^{(k+1)} = x^{(k)} - P^{-1} (A x^{(k)} - b).$$

Der Ausdruck  $(A x^{(k)} - b)$  wird als **Residuum von  $x^{(k)}$**  bezeichnet.

**Bemerkung.** Die Berechnung von  $x^{(k+1)}$  erfordert Inversion von  $P$ !

Im folgenden wollen wir untersuchen, wann die Fixpunktiteration konvergiert.

**Definition.** Der *Spektralradius* einer Matrix  $T \in \mathbb{R}^{d \times d}$  ist

$$\varrho(T) := \max\{|\lambda| \mid \lambda \in \mathbb{C} \text{ Eigenwert von } T\}.$$

**Bemerkung.** a) Sei  $\|\cdot\|_V$  eine Norm auf  $\mathbb{R}^d$ ,  $\|\cdot\|_M$  eine verträgliche Matrixnorm auf  $\mathbb{R}^{d \times d}$ .

Dann gilt:

$$\varrho(T) \leq \|T\|_M.$$

*Beweis.* Sei  $\lambda$  ein Eigenwert von  $T$  und  $v$  ein normierter Eigenvektor, dann folgt:

$$|\lambda| = \|\lambda v\|_V = \|T v\|_V \leq \|T\|_M \|v\|_V = \|T\|_M.$$

□

b)

$$\|T\|_2 = \sqrt{\varrho(T^* T)}.$$

Falls  $T$  symmetrisch ist, gilt:

$$\|T\|_2 = \varrho(T).$$

**Satz 4.2.** 1. Die folgenden Aussagen sind äquivalent:

(i) Die Iterationsfolge  $x^{(k)}$  konvergiert für jeden Startwert  $x^{(0)} \in \mathbb{R}^d$  gegen die Lösung  $x^*$  von  $A x = b$ .

(ii)  $T^k$  konvergiert gegen 0 (d.h.  $(T^k)_{i,j}$  konvergiert gegen 0 für alle  $i, j$ ).

(iii)  $\varrho(T) < 1$ .

2. Hinreichende Bedingung:

Ist  $\|\cdot\|_V$  eine Norm auf  $\mathbb{R}^d$  und  $\|\cdot\|_M$  eine verträgliche Matrixnorm mit  $\|T\|_M < 1$ , dann gilt:

$$\|x^{(k+1)} - x^*\|_V \leq \|T\|_M \|x^{(k)} - x^*\|_V \quad \text{für alle } k \in \mathbb{N},$$

d.h.  $x^{(k)}$  konvergiert bzgl.  $\|\cdot\|_V$  monoton gegen  $x^*$ .



*Beweis.* i) • (i)  $\Leftrightarrow$  (ii): Wegen

$$Ax^* = b \quad \Leftrightarrow \quad x^* = \phi(x^*)$$

gilt

$$x^{(k+1)} - x^* = \phi(x^{(k)}) - \phi(x^*) = T(x^{(k)} - x^*),$$

für alle  $k \in \mathbb{N}$ , und somit:

$$x^{(k)} - x^* = T^k(x^{(0)} - x^*).$$

$x^{(k)}$  konvergiert also genau dann gegen  $x^*$ , wenn  $T^k$  gegen 0 konvergiert.

- (ii)  $\Rightarrow$  (iii): Angenommen,  $\varrho(T) \geq 1$ . Dann gibt es einen Eigenwert  $\lambda \in \mathbb{C}$  mit  $|\lambda| \geq 1$  mit Eigenvektor  $v \in \mathbb{C}^d \setminus \{0\}$  von  $T$ :

$$Tv = \lambda v.$$

Daher konvergiert  $T^k v = \lambda^k v$  nicht gegen 0. Somit konvergiert  $T^k$  nicht gegen 0 im Widerspruch zu (ii).

- (iii)  $\Rightarrow$  (ii): Wir beweisen diese Richtung für den symmetrischen Fall ( $T = T^t$ ). Da  $T$  symmetrisch ist, gibt es eine orthogonale Matrix  $O \in \mathbb{R}^{d \times d}$ , so dass gilt:

$$T = O D O^T, \quad \text{wobei} \quad D = \text{diag}(\lambda_1, \dots, \lambda_d) = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_d \end{pmatrix}$$

die Diagonalmatrix aus Eigenwerten von  $T$  ist. Wegen  $O^T = O^{-1}$  gilt:

$$T^k = (O D O^{-1})^k = O D^k O^T.$$

Aus  $\varrho(T) < 1$  folgt nun für alle Eigenwerte  $|\lambda_i| < 1$ . Also konvergiert  $D^k$  und somit auch  $T^k$  gegen 0.

Den allgemeinen Fall zeigt man via Jordanscher Normalform, siehe z.B. »Hämmerlin/Hoffmann: *Numerische Mathematik*, Satz 8.1«.

ii) Es gilt:

$$\|\phi(x) - \phi(y)\|_V = \|Tx - Ty\|_V \leq \|T\|_M \|x - y\|_V.$$

Wegen  $\|T\|_M < 1$  ist  $\phi$  eine Kontraktion bzgl.  $\|\cdot\|_V$ . Mit dem Banachschen Fixpunktsatz folgt die Behauptung.

□

## Klassische Iterationsverfahren

### Keine Prädiktionierung

$$P = I.$$

Dann lautet die Residuendarstellung:

$$x^{(k+1)} = x^{(k)} - (Ax^{(k)} - b), \quad \text{und} \quad T = I - A.$$

Bezüglich Konvergenz gilt:

*$\lambda$  ist genau dann Eigenwert von  $A$ , wenn  $1 - \lambda$  Eigenwert von  $T$  ist.*

Also ist der Spektralradius

$$\rho(T) = \max\{|1 - \lambda| \mid \lambda \text{ Eigenwert von } A\}.$$

**Beispiel.** Ist  $A$  symmetrisch und positiv definit ( $\gg$ s.p.d. $\ll$ ), dann sind die Eigenwerte von  $A$  reell und positiv. Also gilt

$$\rho(T) = \max\{|1 - \lambda| \mid \lambda \text{ Eigenwert von } A\} < 1 \quad \Leftrightarrow \quad \lambda_{\max}(A) < 2.$$

Die Bedingung an die Eigenwerte ist bei diesem Verfahren zu restriktiv!

### Jacobi-Verfahren / Gesamtschrittverfahren

Es gilt:

$$\begin{aligned} Ax = b &\Leftrightarrow \sum_{j=1}^d a_{ij} x_j = b_i \quad i = 1, \dots, n \\ &\Leftrightarrow a_{ii} x_i = b_i - \sum_{j \neq i} a_{ij} x_j. \end{aligned}$$

#### Algorithmus 4.3.

INPUT: Matrix  $A \in \mathbb{R}^{d \times d}$  mit  $a_{ii} \neq 0$  für alle  $i \leq d$ ,  $b \in \mathbb{R}^d$ , Startwert  $x^{(0)} \in \mathbb{R}^d$ ,

OUTPUT: Iterationsfolge  $x^{(k)} \in \mathbb{R}^d$ .

**for**  $k := 0, 1, \dots$  **do**

**for**  $i := 1, \dots, d$  **do**

$$x_i^{(k+1)} := \frac{1}{a_{ii}} \left( b_i - \sum_{j \neq i} a_{ij} x_j^{(k)} \right) \tag{4.2.1}$$

*end for*  
*end for*

EINORDNUNG IN ALLGEMEINEN RAHMEN:

Die Matrix  $A$  wird in drei Matrizen  $L$ ,  $D$  und  $R$  zerlegt:

$$A = -L + D - R, \quad \text{mit}$$

$$L = - \begin{pmatrix} 0 & \cdots & \cdots & 0 \\ a_{21} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ a_{d1} & \cdots & a_{d,d-1} & 0 \end{pmatrix}, \quad D = \begin{pmatrix} a_{11} & \cdots & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & a_{dd} \end{pmatrix}, \quad R = - \begin{pmatrix} 0 & a_{12} & \cdots & a_{1d} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & a_{d-1,d} \\ 0 & \cdots & \cdots & 0 \end{pmatrix}.$$

Für das lineare Gleichungssystem gilt mit dieser Zerlegung:

$$Ax = b \quad \Leftrightarrow \quad Dx = (L + R)x + b \quad \Leftrightarrow \quad x = D^{-1}(L + R)x + D^{-1}b.$$

Als Iteration setzen wir:

$$x^{(k+1)} = D^{-1}(L + R)x^{(k)} + D^{-1}b.$$

Also gilt:

$$P = D \quad \text{und} \quad T = D^{-1}(L + R).$$

### Gauß-Seidel-Verfahren/Einzelschrittverfahren

IDEE: Verwende bereits berechnete Komponenten von  $x^{(k+1)}$  sofort.

**Algorithmus 4.4.** Wie Algorithmus 4.3, aber statt Gleichung (4.2.1) mit der Iteration

$$x_i^{(k+1)} := \frac{1}{a_{ii}} \left( b_i - \sum_{j < i} a_{ij} x_j^{(k+1)} - \sum_{j > i} a_{ij} x_j^{(k)} \right) \quad (4.2.2)$$

In Matrixform:

$$Dx^{(k+1)} = b + Lx^{(k+1)} + Rx^{(k)} \quad \Leftrightarrow \quad (D - L)x^{(k+1)} = Rx^{(k)} + b.$$

Einordnung in allgemeinen Rahmen:

$$P = D - L, \quad T = (D - L)^{-1}R.$$

**Satz 4.5.** Ist  $A$  strikt diagonaldominant . d.h.

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|, \quad i = 1, \dots, d,$$

dann konvergieren das Jacobi- und das Gauß-Seidel-Verfahren monoton bzgl. der Maximumsnorm  $\|\cdot\|_\infty$  auf  $\mathbb{R}^d$ .

*Beweis.* Wir zeigen:

$$\|T\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |t_{ij}| < 1.$$

Da  $\|\cdot\|_\infty$  die von der Maximumnorm auf Vektoren induzierte Matrixnorm ist, folgt die Behauptung dann aus Satz 4.2.

1. JACOBI-VERFAHREN:

$$T = D^{-1}(L + R), \quad t_{ij} = \begin{cases} -\frac{a_{ij}}{a_{ii}} & \text{für } j \neq i, \\ 0 & \text{für } j = i. \end{cases}$$

Damit folgt:

$$\|T\|_\infty \leq \max_{1 \leq i \leq n} \frac{1}{|a_{ii}|} \sum_{j \neq i} |a_{ij}| =: c$$

Weil  $A$  strikt diagonaldominant ist, folgt  $c < 1$ .

2. GAUSS-SEIDEL-VERFAHREN:

$$T = (D - L)^{-1} R.$$

Sei  $c$  wie oben definiert, wir behaupten:

$$\|T\|_\infty \leq c < 1.$$

Dazu ist zu zeigen:

$$\|T x\|_\infty \leq c \cdot \|x\|_\infty \quad \text{für alle } x \in \mathbb{R}^n.$$

Sei  $y := T x$ . Nach Algorithmus 4.4 ist

$$y_i = \frac{1}{a_{ii}} \left( - \sum_{j < i} a_{ij} y_j - \sum_{j > i} a_{ij} x_j \right).$$

Wir zeigen durch vollständige Induktion nach  $i$  die Aussage:

$$|y_i| \leq c \cdot \|x\|_\infty.$$

Für  $i = 1$  gilt:

$$|y_1| \leq \frac{1}{|a_{11}|} \sum_{j \neq 1} |a_{1j}| \cdot |x_j| \leq c \cdot \|x\|_\infty.$$

Mit der Induktionsvoraussetzung folgt:

$$\begin{aligned}
 |y_i| &\leq \frac{1}{|a_{ii}|} \left( - \sum_{j=1}^{i-1} |a_{ij}| |y_j| + \sum_{j=i+1}^d |a_{ij}| |x_j| \right) \\
 &\leq \frac{1}{|a_{ii}|} \left( \sum_{j \neq i} |a_{ij}| c \|x\|_\infty + \sum_{j=i+1}^d |a_{ij}| \|x\|_\infty \right) \\
 &\leq \frac{1}{|a_{ii}|} \sum_{j \neq i} |a_{ij}| \|x\|_\infty \leq c \|x\|_\infty \quad \text{für alle } i = 1, \dots, d.
 \end{aligned}$$

Somit gilt:

$$\|T x\|_\infty = \|y\|_\infty \leq c \|x\|_\infty.$$

□

Man kann nicht allgemein sagen, welches Verfahren »besser« ist. Für bestimmte Klassen von Matrizen kann man zeigen, daß das Gauß-Seidel-Verfahren schneller konvergiert. Wir werden nun zeigen, daß das Gauß-Seidel-Verfahren für symmetrische positiv definite Matrizen  $A$  stets konvergiert.

Sei  $\langle v, w \rangle = \sum_{i=1}^d v_i w_i$  das euklidische Skalarprodukt im  $\mathbb{R}^d$ . Die von einer symmetrisch positiv definiten Matrix induzierte quadratische Form

$$(x, y)_A := \langle x, A y \rangle = \langle A x, y \rangle, \quad (x, y \in \mathbb{R}^n)$$

nennt man auch **Energieskalarprodukt**, die entsprechende Norm

$$\|x\|_A = \sqrt{(x, x)_A}$$

heißt auch **Energienorm**.

**Beispiel** (Diskreter Laplace-Operator). Sei  $D \subseteq \mathbb{Z}^n$  endlich,

$$u = (u_x)_{x \in \mathbb{Z}^d} \text{ mit } u_x = 0 \quad \text{für alle } x \in D^C,$$

$$(A u)_x = -(\Delta_{\mathbb{Z}^n} u)_x = - \sum_{i=1}^n [(u_{x+e_i} - u_x) - (u_x - u_{x-e_i})].$$

Dann ist die Energienorm von  $u$ :

$$\begin{aligned} \|u\|_A^2 &= \langle u, Au \rangle = \sum_{x \in D} u_x (Au)_x = \sum_{x \in \mathbb{Z}^n} u_x (Au)_x \\ &= \sum_{i=1}^n \left( \sum_x u_x (u_x - u_{x-e_i}) - \sum_x u_x (u_{x+e_i} - u_x) \right) \\ &= \sum_{i=1}^n \left( \sum_x u_{x+e_i} (u_{x+e_i} - u_x) - \sum_x u_x (u_{x+e_i} - u_x) \right) \\ &= \sum_{x \in \mathbb{Z}^n} \sum_{i=1}^n \underbrace{\left( \underbrace{u_{x+e_i} - u_x}_{\text{diskrete Richtungs-}} \right)^2}_{\text{ableitung } \partial_i u(x)} = \sum_x \underbrace{\|\nabla_{\mathbb{Z}^n} u(x)\|^2}_{\text{Energie von } u}, \end{aligned}$$

wobei

$$\nabla_{\mathbb{Z}^n} u(x) = (u(x + e_1) - u(x), \dots, u(x + e_n) - u(x)).$$

Es folgt, daß die lineare Abbildung  $A = -\Delta_{\mathbb{Z}^n}$  auf

$$V = \{u = (u_x)_{x \in \mathbb{Z}^n} \mid u_x = 0 \text{ für alle } x \in D^C\} \cong \mathbb{R}^{|D|}$$

symmetrisch und positiv definit ist, mit Energienorm

$$\|u\|_A^2 = \sum_x \|\nabla_{\mathbb{Z}^n} u(x)\|^2.$$

KONTINUIERLICHES ANALOGON:

$$-\int_D u \Delta u = \int_D \|\nabla u\|^2 \quad \text{falls } u = 0 \text{ auf } \partial D \quad (\text{siehe »Analysis III«}).$$

**Satz 4.6.** *Ist  $A$  symmetrisch und positiv definit, dann hat  $Ax = b$  eine eindeutige Lösung  $x^*$  und das Gauß-Seidel-Verfahren konvergiert monoton bzgl.  $\|\cdot\|_A$  gegen  $x^*$ , d.h. es gibt ein  $L < 1$  mit:*

$$\|x^{(n+1)} - x^*\|_A \leq L \|x^{(n)} - x^*\|_A.$$

**Bemerkung.** Sei  $A$  symmetrisch positiv definit, dann gilt

$$a_{ii} = (e_i, A e_i) > 0 \quad \text{für alle } i.$$

*Beweis.* Zu zeigen ist:

$$\|T\|_A := \max_{\|x\|_A=1} \|Tx\|_A < 1. \quad (4.2.3)$$

Die Behauptung folgt dann aus Satz 4.2

## 4.2. KLASSISCHE ITERATIONSVERFAHREN FÜR LINEARE GLEICHUNGSSYSTEME 1

1. Es genügt zu zeigen:

$$\|T x\|_A < 1 \quad \text{für alle } x \in \mathbb{R}^n \text{ mit } \|x\|_A = 1, \quad (4.2.4)$$

da  $\{x \in \mathbb{R}^n \mid \|x\|_A = 1\}$  kompakt und die Abbildung  $x \mapsto \|x\|_A$  stetig ist und somit in Gleichung (4.2.3) das Maximum angenommen wird.

2. Sei  $T = I - P^{-1} A$ , und  $P$  der Präkonditionierer, dann gilt:

$$\begin{aligned} (T x, T x)_A &= \langle T x, A T x \rangle = \langle x, A T x \rangle - \langle P^{-1} A x, A T x \rangle \\ &= (x, T x)_A - \langle x, A^t (P^{-1})^t A T x \rangle \\ &= (x, T x)_A - (x, (P^{-1})^t A T x)_A \\ &= (x, x)_A - (x, Z x)_A \stackrel{!}{<} (x, x)_A, \end{aligned}$$

wobei

$$\begin{aligned} Z &= P^{-1} A + (P^t)^{-1} A - (P^t)^{-1} A P^{-1} A \\ &= (I + (P^t)^{-1} P - (P^t)^{-1} A) P^{-1} A \\ &= (P^t)^{-1} (P^t + P - A) P^{-1} A. \end{aligned}$$

3. Zu zeigen bleibt:

$$(x, Z x)_A > 0 \quad \text{für alle } x \in \mathbb{R}^n \setminus \{0\}.$$

Es gilt:

$$\begin{aligned} (x, Z x)_A &= \langle x, A (P^t)^{-1} (P^t + P - A) P^{-1} A x \rangle \\ &= \langle P^{-1} A x, (P^t + P - A) P^{-1} A x \rangle. \end{aligned}$$

Im Gauß-Seidel-Verfahren gilt:

$$P = D - L, \quad P^t = D - L^t = D - R,$$

da  $A$  symmetrisch ist. Es folgt:

$$P^t + P - A = (D - L) + (D - R) - (D - L - R) = D \quad \text{ist positiv definit.}$$

Also folgt

$$(x, Z x)_A > 0 \quad \text{für alle } x \in \mathbb{R}^n \setminus \{0\}.$$

□

**Bemerkung.** Im Beweis haben wir benutzt, daß  $A$  symmetrisch positiv definit und  $P + P^t - A$  positiv definit ist!

## Relaxationsverfahren

ALLGEMEINES ITERATIONSVERFAHREN:

$$x^{(k+1)} = T x^{(k)} + f, \quad T = I - P^{-1} A, \quad f = P^{-1} b.$$

IDEE: erzwingen Konvergenz durch Dämpfungsparameter  $\omega \in (0, 1]$ :

RELAXATIONSVERFAHREN:

$$x^{(k+1)} = \omega (T x^{(k)} + f) + (1 - \omega) x^{(k)}.$$

**Bemerkung.** a)  $x$  ist genau dann Fixpunkt, wenn  $x = \omega (T x + f) + (1 - \omega) x$  gilt. Falls  $\omega \neq 0$  ist dies äquivalent zu  $x = T x + f$  bzw.  $A x = b$ .

b) Die Iterationsmatrix ist

$$T_\omega = \omega T + (1 - \omega) I = I - \omega P^{-1} A,$$

d.h. das Relaxationsverfahren ist ein Iterationsverfahren mit Prädiktionierer

$$P_\omega = \frac{1}{\omega} P.$$

ZIEL: Wähle  $\omega$  so, daß  $\rho(T_\omega)$  möglichst klein ist.

## Relaxiertes Jacobi-Verfahren (Jacobi over-relaxation – »JOR«)

JACOBI-VERFAHREN:

$$x^{(k+1)} = D^{-1} (b + (L + R) x^{(k)})$$

JOR-VERFAHREN:

$$x^{(k+1)} = \omega D^{-1} (b + (L + R) x^{(k)}) + (1 - \omega) x^{(k)}, \quad \text{d.h.}$$

$$x_i^{(k+1)} = \frac{\omega}{a_{ii}} (b_i - \sum_{j \neq i} a_{ij} x_j^{(k)}) + (1 - \omega) x_i^{(k)}.$$

PRÄKONDITIONIERER:

$$P_\omega = \frac{1}{\omega} P.$$

**Satz 4.7.** Ist  $A$  eine symmetrisch positiv definite Matrix, dann konvergiert das JOR-Verfahren für alle  $\omega \in (0, 2/\rho(D^{-1} A))$ .



## 4.2. KLASSISCHE ITERATIONSVERFAHREN FÜR LINEARE GLEICHUNGSSYSTEME 3

*Beweis.*

$$T_\omega = I - \omega D^{-1} A,$$

also gilt:

$$\lambda \text{ ist genau dann Eigenwert von } D^{-1} A, \text{ wenn } 1 - \omega \lambda \text{ Eigenwert von } T_\omega \text{ ist.} \quad (4.2.5)$$

Ist  $A$  symmetrisch positiv definit, so auch  $D^{-\frac{1}{2}} A D^{-\frac{1}{2}}$ , d.h. alle Eigenwerte von  $D^{-\frac{1}{2}} A D^{-\frac{1}{2}}$  sind reell und positiv. Da  $D^{-1} A$  ähnlich zu  $D^{-\frac{1}{2}} A D^{-\frac{1}{2}}$  ist, hat auch  $D^{-1} A$  nur reelle und positive Eigenwerte. Es folgt, dass das JOR-Verfahren genau dann konvergiert, wenn für den Spektralradius von  $T_\omega$  gilt:

$$\rho(T_\omega) = \max\{|1 - \omega \lambda| \mid \lambda \text{ Eigenwert von } D^{-1} A\} < 1.$$

Dies ist äquivalent dazu, dass  $\omega \lambda < 2$  für alle Eigenwerte von  $D^{-1} A$  gilt, und da  $D^{-1} A$  nur reelle und positive Eigenwerte hat, ist dies äquivalent zu

$$\omega \rho(D^{-1} A) < 2.$$

□

### Sukzessives Relaxationsverfahren (Successive over-relaxation – »SOR«)

Modifiziere JOR-Verfahren à la Gauß-Seidel-Verfahren:

$$x_i^{(k+1)} = \frac{\omega}{a_{ii}} \left( b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} \right) + (1 - \omega) x_i^{(k)}.$$

In Matrixform:

$$x^{(k+1)} = \omega D^{-1} (b + L x^{(k+1)} + R x^{(k)}) + (1 - \omega) x^{(k)}.$$

Dies ist äquivalent zu:

$$\underbrace{\left( \frac{1}{\omega} D - L \right)}_{=P_\omega \text{ Präkonditionierer}} x^{(k+1)} = \underbrace{\left( R + \frac{1-\omega}{\omega} D \right)}_{N_\omega} x^{(k)} + b.$$

**Satz 4.8.** *Ist  $A$  symmetrisch positiv definit, dann konvergiert das SOR-Verfahren für alle  $\omega \in (0, 2)$  monoton bzgl.  $\|\cdot\|_A$ .*

*Beweis.* Verfahre wie bei Satz 4.6:

$$P_\omega = \frac{1}{\omega} D - L, \quad P_\omega^t = \frac{1}{\omega} D - R, \quad A = D - L - R.$$

Dann ist

$$P_\omega + P_\omega^t - A = \left( \frac{2}{\omega} - 1 \right) D$$

positiv definit für alle  $\omega < 2$ . Damit folgt die Behauptung. □

### 4.3 Abstiegsverfahren

Sei  $A \in \mathbb{R}^{d \times d}$  symmetrisch und positiv definit,  $b \in \mathbb{R}^d$ . Wir betrachten wieder das lineare Gleichungssystem

$$Ax = b. \tag{4.3.1}$$

Wie zuvor wollen wir die eindeutige Lösung  $x^* = A^{-1}b$  numerisch berechnen. Dieses Problem lässt sich in ein *Minimierungsproblem* der folgenden Funktion  $\phi$  umformulieren:

$$\phi(x) := \frac{1}{2} \langle x, Ax \rangle - \langle x, b \rangle.$$

**Lemma 4.9.** Die Lösung  $x^* = A^{-1}b$  des linearen Gleichungssystems ist das eindeutige Minimum von  $\phi$ , und für alle  $x \in \mathbb{R}^d$  gilt:

$$\phi(x) - \phi(x^*) = \frac{1}{2} \|x - x^*\|_A^2.$$

*Beweis.* Mit quadratischer Ergänzung sieht man, dass die Funktion

$$\phi(x) = \frac{1}{2} (x, x)_A - (x, A^{-1}b)_A = \frac{1}{2} \|x - x^*\|_A^2 - \frac{1}{2} \|x^*\|_A^2$$

minimal ist für  $x = x^*$ . Das Minimum ist dann

$$\phi(x^*) = -\frac{1}{2} \|x^*\|_A^2.$$

□

ALLGEMEINES ITERATIONSVERFAHREN ZUR MINIMIERUNG VON  $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$  (LINE SEARCH):

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}.$$

Hierbei bezeichnen

$$\begin{aligned} p^{(0)}, p^{(1)}, \dots \in \mathbb{R}^d, & \quad \text{die »Suchrichtungen«,} \\ \alpha_0, \alpha_1, \dots \in \mathbb{R}, & \quad \text{die »Schrittweiten«,} \end{aligned}$$

Für die eben betrachtete Funktion  $\phi$  wählt man die Schrittweiten  $\alpha_k$  so, dass

$$\phi(x^{(k+1)}) = \phi(x^*) + \frac{1}{2} \|x^{(k+1)} - x^*\|_A^2$$

minimal unter allen  $x^{(k)} + \alpha p^{(k)}$ ,  $\alpha \in \mathbb{R}$ , ist. Dies ist der Fall für

$$\begin{aligned} x^{(k+1)} &= A\text{-orthogonale Projektion von } x^* \text{ auf Gerade } x^{(k)} + \text{span}(p^{(k)}) \\ &= x^{(k)} + \left( x^* - x^{(k)}, \frac{p_k}{\|p_k\|_A} \right)_A \frac{p_k}{\|p_k\|_A}, \end{aligned}$$

d.h.

$$\alpha_k^{\text{opt}} = \frac{\langle x^* - x^{(k)}, p^{(k)} \rangle}{\langle p^{(k)}, p^{(k)} \rangle_A} = \frac{\langle Ax^* - Ax^{(k)}, p^{(k)} \rangle}{\langle p^{(k)}, Ap^{(k)} \rangle}.$$

Es gilt also:

$$\alpha_k^{\text{opt}} = \frac{\langle r^{(k)}, p^{(k)} \rangle}{\langle p^{(k)}, Ap^{(k)} \rangle} \quad \text{mit Residuum} \quad r^{(k)} = b - Ax^{(k)}.$$

Wir betrachten nun zwei spezielle Abstiegsverfahren.

## Gradientenverfahren

Sei

$$p^{(k)} = -\nabla\phi(x^{(k)}), \quad \text{die Richtung des steilsten Abstiegs.}$$

In unserem Fall gilt:

$$\begin{aligned} \phi(x) &= \frac{1}{2} \langle x, Ax \rangle - \langle x, b \rangle, & A \text{ symmetrisch,} \\ p^{(k)} &= -\nabla\phi(x^{(k)}) = b - Ax^{(k)} = r^{(k)}, & \text{»Residuum von } x^{(k)}\text{«,} \\ \alpha_k &= \frac{\langle r^{(k)}, r^{(k)} \rangle}{\langle r^{(k)}, Ar^{(k)} \rangle} = \frac{\|r^{(k)}\|^2}{\|r^{(k)}\|_A^2}. \end{aligned}$$

Das Residuum lässt sich rekursiv berechnen:

$$\begin{aligned} r^{(k+1)} &= b - Ax^{(k+1)} = b - Ax^{(k)} - \alpha_k Ar^{(k)} \\ &= r^{(k)} - \alpha_k Ar^{(k)}. \end{aligned}$$

Damit erhalten wir folgenden Algorithmus:

### Algorithmus 4.10 (Steepest Descent).

INPUT:  $A \in \mathbb{R}^{d \times d}$  symmetrisch positiv definit,  $b \in \mathbb{R}^d$ ,  $x^{(0)} \in \mathbb{R}^d$ ,

OUTPUT: Iterationsfolgen  $x^{(k)} \in \mathbb{R}^d$ ,  $r^{(k)} \in \mathbb{R}^d$ .

$r^{(0)} := b - Ax^{(0)}$

**for**  $k := 0, 1, \dots$  **do**

$$\begin{aligned} \alpha_k &:= \frac{\langle r^{(k)}, r^{(k)} \rangle}{\langle r^{(k)}, Ar^{(k)} \rangle} \\ x^{(k+1)} &:= x^{(k)} + \alpha_k r^{(k)} \\ r^{(k+1)} &:= r^{(k)} - \alpha_k Ar^{(k)} \end{aligned}$$

*end for*

**Bemerkung.** Die Iterationsmatrix im  $k$ -ten Schritt des Gradientenverfahrens ist

$$T^{(k)} = I - \alpha_k A, \quad (\text{keine Prakonditionierung!}).$$

Das Gradientenverfahren ist also ein **nicht-stationares** (d.h. die Iterationsmatrix hangt von  $k$  ab), *nicht prakonditioniertes Verfahren mit Relaxationsparameter  $\alpha_k$ .*

**Satz 4.11.** *Ist  $A$  symmetrisch positiv definit, dann konvergiert das Gradientenverfahren fur jeden Startwert  $x^{(0)}$  monoton bzgl. der Energienorm  $\|\cdot\|_A$ , und es gilt:*

$$\|x^{(k+1)} - x^*\|_A \leq \frac{K_2(A) - 1}{K_2(A) + 1} \cdot \|x^{(k)} - x^*\|_A,$$

wobei  $K_2(A) = \|A\|_2 \cdot \|A^{-1}\|_2$  die  $l^2$ -Kondition der Matrix  $A$  ist.

Wir zeigen zunachst:

**Lemma 4.12.** *Ist  $A$  symmetrisch positiv definit mit Eigenwerten  $0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_d$ , dann gilt:*

$$\|A\|_2 = \max_{\|x\|=1} \|Ax\| = \lambda_d = \varrho(A),$$

$A^{-1}$  ist symmetrisch positiv definit mit Eigenwerten  $\frac{1}{\lambda_1} \geq \frac{1}{\lambda_2} \geq \dots \geq \frac{1}{\lambda_d} > 0$  und es gilt:

$$K_2(A) = \|A\|_2 \cdot \|A^{-1}\|_2 = \frac{\lambda_d}{\lambda_1}.$$

*Beweis.* Da  $A$  symmetrisch positiv definit ist, gibt es eine Orthonormalbasis von Eigenvektoren  $e_i$  von  $A$ . D.h. es gilt  $\langle e_i, e_j \rangle = \delta_{ij}$  und jeder Vektor  $x$  lasst sich darstellen als  $x = \sum_{i=1}^d c_i e_i$ . Damit folgt:

$$Ax = \sum_{i=1}^d c_i \lambda_i e_i, \quad \|x\|^2 = \sum_{i=1}^d c_i^2, \quad \|Ax\|^2 = \sum_{i=1}^d c_i^2 \lambda_i^2.$$

Weiterhin gilt

$$\|A\|_2 = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \max_{i=1 \dots d} |\lambda_i| = \lambda_d.$$

Analog kann man

$$\|A^{-1}\|_2 = \max_{i=1 \dots d} \left| \frac{1}{\lambda_i} \right| = \frac{1}{\lambda_1}$$

zeigen. □

*Beweis von Satz 4.11.* Sei  $T_\alpha := I - \alpha A$  die Iterationsmatrix des Verfahrens mit Relaxationsparameter  $\alpha$ .

i) VERGLEICH MIT  $T_\alpha$ -VERFAHREN:

Wir behaupten, dass der Fehler des Gradientenverfahrens durch den Fehler des Verfahrens mit  $T_\alpha$  für alle  $\alpha \in \mathbb{R}$  beschränkt ist:

$$\|x^{(k+1)} - x^*\|_A \leq \|T_\alpha (x^{(k+1)} - x^*)\|_A.$$

Nach Definition ist  $x^{(k+1)}$  die  $A$ -orthogonale Projektion von  $x^*$  auf die Gerade  $x^{(k)} + \alpha r^{(k)}$ ,  $\alpha \in \mathbb{R}$ . Also gilt für alle  $\alpha \in \mathbb{R}$ :

$$\begin{aligned} \|x^{(k+1)} - x^*\|_A &\leq \|x^{(k)} + \alpha r^{(k)} - x^*\|_A \\ &= \|T_\alpha x^{(k)} + \alpha b - T_\alpha x^* - \alpha b\|_A = \|T_\alpha (x^{(k+1)} - x^*)\|_A. \end{aligned}$$

ii) KONTRAKTIVITÄT VON  $T_\alpha$ :

Seien  $0 < \lambda < \dots < \lambda_d$  die Eigenwerte von  $A$ . Wir behaupten:

$$\|T_\alpha y\|_A \leq \max_{i=1..d} |1 - \alpha \lambda_i| \cdot \|y\|_A \quad \text{für alle } y \in \mathbb{R}^d.$$

Seien  $e_i$  die Eigenvektoren zu  $\lambda_i$  und  $\langle e_i, e_j \rangle = \delta_{ij}$ . Dann gilt:

$$\begin{aligned} T_\alpha e_i &= (1 - \alpha \lambda_i) e_i, & \text{und} \\ A T_\alpha e_i &= \lambda_i (1 - \alpha \lambda_i) e_i. \end{aligned}$$

Für  $y = \sum_{i=1}^n c_i e_i$  folgt:

$$\begin{aligned} \|T_\alpha y\|_A^2 &= (T_\alpha y, A T_\alpha y) \\ &= \left( \sum_{i=1}^d c_i (1 - \alpha \lambda_i) e_i, \sum_{i=1}^d c_i \lambda_i (1 - \alpha \lambda_i) e_i \right) = \sum_{i=1}^d c_i^2 \lambda_i (1 - \alpha \lambda_i)^2. \end{aligned}$$

Mit

$$\|y\|_A^2 = (y, A y) = \sum_{i=1}^d c_i^2 \lambda_i$$

folgt dann

$$\|T_\alpha y\|_A^2 \leq \max_{i=1..d} (1 - \alpha \lambda_i)^2 \|y\|_A^2.$$

iii) OPTIMIEREN DER ABSCHÄTZUNG:

Nach i) und ii) gilt für alle  $\alpha \geq 0$ :

$$\|x^{(k+1)} - x^*\|_A \leq \|T_\alpha(x^{(k)} - x^*)\|_A \leq c(\alpha) \cdot \|x^{(k)} - x^*\|_A,$$

wobei

$$c(\alpha) := \max_{i=1 \dots d} |1 - \alpha \lambda_i| = \max(1 - \alpha \lambda_1, \alpha \lambda_d - 1)$$

wegen  $1 - \alpha \lambda_1 \geq 1 - \alpha \lambda_2 \geq \dots \geq 1 - \alpha \lambda_d$ . Um eine möglichst gute Abschätzung zu erhalten, wählen wir  $\alpha$  nun so, dass  $c(\alpha)$  minimal wird. Dies ist genau dann der Fall, wenn:

$$\alpha \lambda_d - 1 = 1 - \alpha \lambda_1, \quad \text{d.h.} \quad \alpha = \frac{2}{\lambda_d + \lambda_1},$$

und für diese Wahl von  $\alpha$  gilt:

$$c(\alpha) = 1 - \alpha \lambda_1 = \frac{\lambda_d - \lambda_1}{\lambda_d + \lambda_1} = \frac{\frac{\lambda_d}{\lambda_1} - 1}{\frac{\lambda_d}{\lambda_1} + 1} = \frac{K_2(A) - 1}{K_2(A) + 1}.$$

□

**Bemerkung.** a) Wegen

$$\frac{K_2(A) - 1}{K_2(A) + 1} = 1 - \frac{2}{K_2(A) + 1}$$

kann die Konvergenzgeschwindigkeit des Gradientenverfahrens bei Matrizen mit großer Kondition relativ langsam sein.

b) Das Gradientenverfahren lässt sich auch zur Minimierung nicht-quadratischer Funktionen  $\phi$  verwenden – aber es ist dann im Allgemeinen nicht klar, wie die Schrittweiten  $\alpha_k$  geeignet zu wählen sind.

## Verfahren der konjugierten Gradienten (CG)

**Beispiel** (Abstiegsverfahren im  $\mathbb{R}^2$ ). Gradientenverfahren für  $\phi(x) = \text{const}$ :

Grafik fehlt noch (!!)

Im allgemeinen gilt  $x^{k+2} \neq x^*$ , d.h. die Iteration bringt nicht ab!

Kann man im zweidimensionalen Fall erzwingen, dass  $x^*$  im nächsten Schritt getroffen wird (also das Verfahren abbricht)? Eine hinreichende Bedingung ist:

$$p^{(k+1)} \parallel x^* - x^{(k+1)}, \quad \text{d.h. die } k+1\text{-te Suchrichtung ist parallel zu } x^* - x^{(k+1)}. \quad (4.3.2)$$

Nach Konstruktion ist  $x^{(k+1)}$  die  $A$ -orthogonale Projektion von  $x^*$  auf  $x^{(k)} + \text{span}(p^{(k)})$ , also gilt:

$$x^* - x^{(k+1)} \perp_A p^{(k)}, \quad \text{d.h. die } k\text{-te Suchrichtung ist } A\text{-orthogonal zu } x^* - x^{(k+1)}.$$

Im zweidimensionalen Fall folgt aus  $p^{(k+1)} \perp_A p^{(k)}$  die Gültigkeit von (4.3.2). Dies motiviert folgende *Modifikation des Gradientenverfahrens*:

$$p^{(k+1)} = r^{(k+1)} - \frac{(r^{(k+1)}, p^{(k)})_A}{(p^{(k)}, p^{(k)})_A} p^{(k)}.$$

Es gilt nun:

$$p^{(k+1)} \perp_A p^{(k)},$$

d.h.  $p^{(k+1)}$  ist  $A$ -konjugiert zu  $p^{(k)}$  (»konjugierter Gradient«).  $r^{(k+1)}$  bezeichnet das Residuum des Gradientenverfahrens des vorigen Abschnitts.

**Algorithmus 4.13** (CG-Verfahren).

INPUT:  $A \in \mathbb{R}^{d \times d}$  symmetrisch positiv definit,  $b \in \mathbb{R}^d$ ,  $x^{(0)} \in \mathbb{R}^d$ ,

OUTPUT: Iterationsfolgen  $x^{(k)} \in \mathbb{R}^d$ ,  $r^{(k)} \in \mathbb{R}^d$ ,  $p^{(k)} \in \mathbb{R}^d$ .

$$p^{(0)} := r^{(0)} := b - Ax^{(0)} \quad (4.3.3)$$

**for**  $k := 0, 1, \dots$  **do**

$$\alpha_k := \frac{\langle r^{(k)}, p^{(k)} \rangle}{(p^{(k)}, p^{(k)})_A} \quad (4.3.4)$$

$$x^{(k+1)} := x^{(k)} + \alpha_k p^{(k)} \quad (4.3.5)$$

$$r^{(k+1)} := r^{(k)} - \alpha_k A p^{(k)} \quad (4.3.6)$$

$$\beta_k := \frac{(r^{(k+1)}, p^{(k)})_A}{(p^{(k)}, p^{(k)})_A} \quad (4.3.7)$$

$$p^{(k+1)} := r^{(k+1)} - \beta_k p^{(k)} \quad (4.3.8)$$

**end for**

Wir wollen nun zeigen, dass das CG-Verfahren im  $\mathbb{R}^d$  nach höchstens  $d$  Schritten abbricht, und die exakte Lösung  $x^*$  liefert. Sei dazu

$$V_k := \text{span}\{p^{(0)}, \dots, p^{(k)}\}.$$

Durch Induktion folgt sofort aus (4.3.3), (4.3.6) und (4.3.8):

$$V_k = \text{span}\{r^{(0)}, \dots, r^{(k)}\} = \text{span}\{r^{(0)}, A r^{(0)}, \dots, A^k r^{(0)}\}.$$

$V_k$  heißt auch **Krylov-Raum**.

**Satz 4.14.** Für alle  $k \geq 0$  gilt:

- i)  $x^{(k)} - x^* \perp_A V_k$ ,
- ii)  $r^{(k+1)} \perp_A V_k$ , d.h.  $r^{(0)}, r^{(1)}, \dots, r^{(k+1)}$  sind orthogonal,
- iii)  $p^{(k+1)} \perp_A V_k$ , d.h.  $p^{(0)}, p^{(1)}, \dots, p^{(k+1)}$  sind  $A$ -orthogonal.

Insbesondere folgt  $r^{(k)} = 0$ , also  $x^{(k)} = x^*$ , für alle  $k \geq d$ , d.h. das CG-Verfahren liefert nach maximal  $d$  Schritten die exakte Lösung.

**Korollar** (Berechnung der Koeffizienten und Residuen).

$$\alpha_k = \frac{\langle r^{(k)}, r^{(k)} \rangle}{\langle p^{(k)}, A p^{(k)} \rangle}, \quad \beta_k = -\frac{\langle r^{(k+1)}, r^{(k+1)} \rangle}{\langle r^{(k)}, r^{(k)} \rangle}.$$

*Beweis.* Der Beweis verbleibt als Übungsaufgabe. □

Diese Art der Berechnung ist effizienter und numerisch stabiler als die Formel von oben, und sollte daher in Implementationen des CG-Verfahrens verwendet werden.

*Beweis von Satz 4.14.* Die Aussagen werden durch vollständige Induktion nach  $k$  bewiesen. Der Induktionsanfang ist eine Übungsaufgabe. Seien die drei Aussagen für  $k$  wahr.

- i) Nach Konstruktion ist  $x^{(k+1)}$  die  $A$ -orthogonale Projektion auf die Gerade  $x^{(k)} + \text{span}(p^{(k)})$ , also gilt

$$(x^{(k+1)} - x^*, p^{(k)})_A = 0.$$

Zudem gilt nach ii):

$$x^{(k+1)} - x^* = x^{(k)} - x^* + \alpha p^{(k)}.$$

Der letzte Ausdruck ist nach Induktionsvoraussetzung  $A$ -orthogonal zu  $V_{k-1}$ , also folgt:

$$x^{(k+1)} - x^* \perp_A V_k.$$



ii) Für alle  $z \in V^{(k)}$  gilt nach i):

$$0 = (x^{(k+1)} - x^*, z)_A = \langle Ax^{(k+1)} - Ax^*, z \rangle = \langle -r^{(k+1)}, z \rangle.$$

iii) Nach Konstruktion ist  $p^{(k+1)}$   $A$ -orthogonal zu  $p^{(k)}$ . Zudem gilt für  $z \in V_{k-1}$  nach Induktionsvoraussetzung:

$$(p^{(k+1)}, z)_A = (r^{(k+1)}, z)_A - \beta_k (p^{(k)}, z)_A = \langle r^{(k+1)}, Az \rangle = 0.$$

Also gilt  $p^{(k+1)} \perp_A V_{k-1}$ , und damit  $p^{(k+1)} \perp_A V_k$ .

□

In der Praxis wird das CG-Verfahren häufig schon nach weniger als  $d$  Schritten beendet. Hier gilt folgende Fehlerabschätzung:

**Satz 4.15.** *Ist  $A$  symmetrisch positiv definit, dann gilt*

$$\|x^{(k)} - x^*\|_A \leq \frac{2c^k}{1+c^{2k}} \|x^{(0)} - x^*\|_A,$$

wobei

$$c = \frac{\sqrt{K_2(A)} - 1}{\sqrt{K_2(A)} + 1} = 1 - \frac{2}{\sqrt{K_2(A)} + 1}.$$

*Beweis.* siehe **Quarteroni/Sacco/Saleri**: »Numerische Mathematik 1«. □

**Bemerkung** (Aufwand des CG-Verfahrens.). a) Maximaler Aufwand:

- pro Schritt:  $O(d^2)$  (Multiplikation Matrix  $\times$  Vektor),
- insgesamt:  $O(d^3)$  (exakte Lösung nach  $d$  Schritten).

b) Matrix dünn besetzt (vgl. Programmieraufgabe):

- pro Schritt nur  $O(d)$ .
- insgesamt maximal  $O(d^2)$ .

c) Approximative Lösung: Schrittzahl  $\ll d$ . Daher ist weitere Reduktion des Aufwands möglich.

## 4.4 Das Newton-Verfahren für nichtlineare Gleichungssysteme

Sei  $D \subseteq \mathbb{R}^d$  offen,  $f : D \rightarrow \mathbb{R}^d$  einmal stetig differenzierbar.

Gesucht ist  $x^* \in D$  mit:

$$f(x^*) = 0.$$

Letztes Gleichungssystem hat  $d$  Gleichungen mit  $d$  Unbekannten und ist im Allgemeinen nicht linear.

ANSATZ: LINEARISIERUNG

$$f(x) \approx \underbrace{f(x^{(0)}) + (df)(x^{(0)})(x - x^{(0)})}_{:=l(x)} + o(\|x - x^{(0)}\|).$$

- löse  $l(x) = 0$ ,
- Lösung  $x^{(1)}$  liefert Näherung für Nullstelle von  $f$ ,
- Iteration.

Es gilt:

$$l(x) = 0 \quad \Leftrightarrow \quad df(x^{(0)})(x - x^{(0)}) = -f(x^{(0)}), \quad \text{lineares Gleichungssystem.}$$

Falls  $df(x^{(0)})$  invertierbar ist, gilt ferner:

$$l(x) = 0 \quad \Leftrightarrow \quad x = x^{(0)} - df(x^{(0)})^{-1} f(x^{(0)}).$$

Dies motiviert das Iterationsverfahren:

$$x^{(k+1)} = x^{(k)} - df(x^{(k)})^{-1} f(x^{(k)}).$$

EINDIMENSIONALER FALL:

$$x^{(k+1)} = x^{(k)} - \frac{f(x^{(k)})}{f'(x^{(k)})}.$$

$x^{(k+1)}$  ist in diesem Fall der Schnittpunkt der Tangente an  $f$  in  $x^{(k)}$  mit der  $x$ -Achse. Falls  $f$  konvex und streng monoton wachsend ist, liegt die Tangente unter dem Graph von  $f$ . Dann konvergiert die Newton-Iteration monoton für  $x^{(0)} > x^*$ .

**Beispiel.**

$$f(x) = x^n - a, \quad n \in \mathbb{N} \setminus \{1\}, \quad a \in \mathbb{R}^+.$$

Die Nullstellen von  $f$  entsprechen den  $n$ -ten Wurzeln von  $a$ .

NEWTON-VERFAHREN:

$$f'(x) = n x^{n-1}.$$

$$x^{(k+1)} = x^{(k)} - \frac{(x^{(k)})^n - a}{n (x^{(k)})^{n-1}} = \frac{n-1}{n} x^{(k)} + \frac{a}{n} (x^{(k)})^{1-n}.$$

Wir haben somit ein Iterationsverfahren zur Berechnung der  $n$ -ten Wurzel erhalten.

**Algorithmus 4.16** (Newton-Verfahren).

INPUT:  $f : D \rightarrow \mathbb{R}^d$ , Anfangsnäherung  $x^{(0)} \in D$ , Fehlertoleranz  $\varepsilon > 0$ .

$k := 0$

**repeat**

berechne Lösung  $h \in \mathbb{R}^n$  des linearen Gleichungssystems  $df(x^{(k)})h = -f(x^{(k)})$

$x^{(k+1)} := x^{(k)} + h$

**if**  $x^{(k+1)} \notin D$  **then**

error (»overflow«)

**end if**

$k := k + 1$

**until**  $\|f(x^{(k)})\| \leq \varepsilon$

Im folgenden wollen wir die Konvergenz des Newton-Verfahrens untersuchen.

**Definition** (Konvergenzordnung). Eine Nullfolge  $\varepsilon_k$  nicht-negativer reeller Zahlen konvergiert

(i) **linear**, falls es ein  $L \in (0, 1)$  und ein  $n \in \mathbb{N}$  gibt, so dass gilt:

$$\varepsilon_{k+1} \leq L \cdot \varepsilon_k \quad \text{für alle } k \geq n.$$

(ii) **superlinear**, falls es eine Nullfolge  $L_k$  gibt, so dass gilt:

$$\varepsilon_{k+1} \leq L_k \cdot \varepsilon_k \quad \text{für alle } k \in \mathbb{N}.$$

(iii) **superlinear mit Ordnung**  $p > 1$ , falls es ein  $C \in (0, \infty)$ ,  $n \in \mathbb{N}$  gibt, so dass gilt:

$$\varepsilon_{k+1} \leq C \cdot \varepsilon_k^p \quad \text{für alle } k \geq n.$$

**Beispiel.** (i)  $\varepsilon_k = e^{-\alpha k}$ ,  $\alpha > 0$ . Aufgrund von

$$\varepsilon_{k+1} = e^{-\alpha} \varepsilon_k$$

konvergiert  $\varepsilon_k$  linear.

(ii)  $\varepsilon_k = e^{-k^2}$ . Aufgrund von

$$\varepsilon_{k+1} = e^{-(k+1)^2} = e^{-2k-1} \varepsilon_k$$

konvergiert  $\varepsilon_k$  superlinear, aber nicht mit Ordnung  $p > 1$ .

(iii)  $\varepsilon_k = e^{-p^k}$ . Aufgrund von

$$\varepsilon_{k+1} = e^{-p \cdot p^k} = (\varepsilon_k)^p$$

konvergiert  $\varepsilon_k$  superlinear mit Ordnung  $p$ .

Sei  $\|\cdot\|_V$  eine Norm auf  $\mathbb{R}^d$ ,  $\|\cdot\|_M$  eine verträgliche Matrixnorm.

**Satz 4.17** (Lokal quadratische Konvergenz des Newton-Verfahrens). Sei  $D \subseteq \mathbb{R}^d$  und  $f : D \rightarrow \mathbb{R}^d$  einmal stetig differenzierbar, sei  $x^* \in D$  mit  $f(x^*) = 0$ . Existiert eine Umgebung  $U = \{x \in \mathbb{R}^d \mid \|x - x^*\|_V < \varepsilon\}$  von  $x^*$  mit  $U \subseteq D$ , und existiert ein  $L \in (0, \infty)$ , mit den Eigenschaften

$$\begin{aligned} df(x) \text{ ist invertierbar für alle } x \in U, \quad \text{und} \\ \|df(x)^{-1}(df(y) - df(x))\|_M \leq L \|x - y\|_V \quad \text{für alle } x, y \in U, \end{aligned} \tag{4.4.1}$$

dann gilt

$$\|x^{(k+1)} - x^*\|_V \leq \frac{L}{2} \|x^{(k)} - x^*\|_V^2$$

für alle  $k \geq 0$  mit  $x^{(k)} \in U$ . Insbesondere konvergiert  $\|x^{(k)} - x^*\|_V$  für alle  $x^{(0)}$  mit  $\|x^{(0)} - x^*\|_V < \min(2/L, \varepsilon)$  quadratisch gegen 0.

**Bemerkung.** a) Global gilt im allgemeinen keine Konvergenz (siehe Beispiel unten).

b) Bedingung (4.4.1) ist invariant unter linearen Transformationen

$$f(x) \mapsto \tilde{f}(x) := A f(x), \quad A \in \mathbb{R}^{d \times d} \text{ nicht-singulär.}$$

c) Falls  $df(x^*)$  invertierbar und  $df$  Lipschitz-stetig in einer Umgebung von  $x^*$  ist, gibt es eine geeignete Umgebung  $U$  und Konstante  $L$ , so dass Bedingung (4.4.1) erfüllt ist.

*Beweis.* Für  $k \geq 0$  mit  $x^{(k)} \in U$  gilt:

$$\begin{aligned} x^{(k+1)} - x^* &= x^{(k)} - x^* - df(x^{(k)})^{-1}(f(x^{(k)}) - f(x^*)) \\ &= df(x^{(k)})^{-1}(f(x^*) - f(x^{(k)}) - df(x^{(k)})(x^* - x^{(k)})). \end{aligned} \quad (4.4.2)$$

Sei nun

$$\gamma(t) := (1-t)x^{(k)} + tx^* \in U, \quad 0 \leq t \leq 1,$$

die Verbindungsstrecke von  $x^{(k)}$  nach  $x^*$ . Es gilt:

$$\gamma'(t) = x^* - x^{(k)}.$$

Damit folgt:

$$f(\gamma(1)) - f(\gamma(0)) = f(x^*) - f(x^{(k)}) = \int_0^1 \frac{d}{dt} f(\gamma(t)) dt = \int_0^1 df(\gamma(t))(x^* - x^{(k)}) dt.$$

Außerdem gilt:

$$df(x^{(k)})(x^* - x^{(k)}) = \int_0^1 df(x^{(k)})(x^* - x^{(k)}) dt.$$

Mit Bedingung (4.4.1) folgt:

$$\begin{aligned} \|x^{(k+1)} - x^*\|_V &= \left\| \int_0^1 df(x^{(k)})^{-1} [df(\gamma(t)) - df(x^{(k)})] (x^* - x^{(k)}) dt \right\|_V \\ &\leq \int_0^1 \|df(x^{(k)})^{-1} [df(\gamma(t)) - df(x^{(k)})]\|_M \|x^* - x^{(k)}\|_V dt \\ &= L \int_0^1 \|\gamma(t) - x^{(k)}\|_V dt \|x^* - x^{(k)}\|_V \\ &= L \int_0^1 t \|x^* - x^{(k)}\|_V dt \|x^* - x^{(k)}\|_V \\ &\leq \frac{L}{2} \|x^* - x^{(k)}\|_V. \end{aligned}$$

Angenommen,  $x^{(k)} \in U$ ,  $\|x^{(k)} - x^*\|_V \leq \delta$  für ein  $\delta \in (0, \frac{2}{L})$ . Dann gilt:

$$\|x^{(k+1)} - x^*\|_V \leq \frac{L}{2} \|x^{(k)} - x^*\|_V^2 \leq \frac{\delta L}{2} \|x^{(k)} - x^*\|_V < \|x^{(k)} - x^*\|_V.$$

Somit ist  $x^{(k+1)} \in U$  und  $\|x^{(k+1)} - x^*\|_V \leq \delta$ . Durch Induktion folgt die Behauptung für alle  $k \geq 0$  falls  $x^{(0)} \in U$  mit  $\delta := \|x^{(0)} - x^*\|_V \leq \frac{2}{L}$ .  $\square$

**Beispiel.**  $d = 1$ ,  $f(x) = \arctan(x)$ .  $x^* = 0$  löst  $f(x) = 0$ .

NEWTON-VERFAHREN:

$$x^{(k+1)} = x^{(k)} - (1 + x^{(k)})^2 \arctan x^{(k)},$$

$$\arctan(|x^{(0)}|) \geq \frac{2|x^{(0)}|}{1 + |x^{(0)}|^2}.$$

Dann divergiert  $x^{(k)}$  bestimmt gegen  $\infty$ . Das Newton-Verfahren konvergiert also nicht global.

### Andere Iterationsverfahren im eindimensionalen Fall

Sei  $f : [a, b] \rightarrow \mathbb{R}$  stetig. Gilt  $f(a) \cdot f(b) \leq 0$ , dann existiert nach dem Zwischenwertsatz ein  $x^* \in [a, b]$  mit  $f(x^*) = 0$ . Das Newton-Verfahren konvergiert für  $x^{(0)}$  nahe  $x^*$  quadratisch gegen  $x^*$ , konvergiert aber im allgemeinen nicht global. Außerdem benötigt es die Existenz und Berechnung der Ableitung  $f'$ . Daher werden häufig auch andere Verfahren mit einer kleineren Konvergenzordnung verwendet:

a) BISEKTIONSVERFAHREN:

Sei  $f \in C([a, b])$  mit  $f(a) \cdot f(b) \leq 0$  (Vorzeichenwechsel im Intervall  $[a, b]$ ).

**Algorithmus 4.18** (Bisektionsverfahren).

$$a^{(0)} := a$$

$$b^{(0)} := b$$

**for**  $k = 0, 1, \dots$  **do**

$$x^{(k)} := \frac{a^{(k)} + b^{(k)}}{2}$$

**if**  $f(a^{(k)}) \cdot f(x^{(k)}) \leq 0$  **then**

$$a^{(k+1)} := a^{(k)}$$

$$b^{(k+1)} := x^{(k)}$$

**else**

$$a^{(k+1)} := x^{(k)}$$

$$b^{(k+1)} := b^{(k)}$$

**end if**

**end for**

Das Bisektionsverfahren liefert Intervalle  $[a^{(k)}, b^{(k)}]$  mit  $|b^{(k)} - a^{(k)}| = 2^{-k} |b - a|$ , deren Durchschnitt eine Nullstelle  $x^*$  von  $f$  enthält.

FEHLERABSCHÄTZUNG:

$$|x^{(k)} - x^*| \leq \frac{1}{2} |b^{(k)} - a^{(k)}| \leq 2^{-k-1} |b - a|.$$

Das Bisektionsverfahren konvergiert global, die Konvergenz ist im Allgemeinen nicht monoton und die Konvergenzgeschwindigkeit entspricht linearer Konvergenzordnung.

b) SEKANTENVERFAHREN:

Das Sekantenverfahren funktioniert wie das Newton-Verfahren, wobei man statt der Tangente die Sekante durch die letzten zwei Iterationswerte nimmt:

$$\begin{aligned} x^{(0)} &= a, & x^{(1)} &= b, \\ x^{(k+1)} &= x^{(k)} - \frac{x^{(k)} - x^{(k-1)}}{f(x^{(k)}) - f(x^{(k-1)})} f(x^{(k)}) && \text{für alle } k \geq 1. \end{aligned}$$

Der letzte Bruch ist der Kehrwert der Sekantensteigung. Das Sekantenverfahren benötigt keine Berechnung der Ableitung, konvergiert jedoch im allgemeinen nicht global. Im Falle lokaler Konvergenz ist die Ordnung  $(1 + \sqrt{5})/2$  (siehe z.B. **Hanke-Bourgeois**: »Grundlagen der numerischen Mathematik«).

c) REGULA FALSI-VERFAHREN:

ANNAHME:

$$f(a) \cdot f(b) \leq 0.$$

Das Regula Falsi-Verfahren funktioniert wie das Sekantenverfahren, aber anstelle von  $x^{(k-1)}$  wird  $x^{(l)}$  mit

$$l = \max\{i < k \mid f(x^{(i)}) \cdot f(x^{(k)}) < 0\}$$

verwendet. Das Regula Falsi-Verfahren hat lineare Konvergenzordnung. Es konvergiert global: die Nullstelle befindet sich zwischen  $x^{(l)}$  und  $x^{(k)}$ .

d) KOMBINIERTES VERFAHREN:

Zum Beispiel Bisektionsverfahren für Anfangsnäherung, anschließend Newton-Verfahren für schnelle genaue Approximation von  $x^*$ .

# Kapitel 5

## Interpolation und Quadratur

### 5.1 Stückweise lineare Interpolation

Sei  $[a, b] \subseteq \mathbb{R}$  ein Intervall,  $-\infty < a < b < \infty$ ,

$$\begin{aligned} a = x_0 < x_1 < \dots < x_n = b & \quad \text{eine Partition des Intervalls} \\ \Delta = \{x_0, x_1, \dots, x_n\} & \quad \text{ein eindimensionales Gitter,} \quad \text{und} \\ y_0, y_1, \dots, y_n \in \mathbb{R}. & \end{aligned}$$

**Definition.**

$$s(x) := \frac{x_{i+1} - x}{x_{i+1} - x_i} \cdot y_i + \frac{x - x_i}{x_{i+1} - x_i} \cdot y_{i+1} \quad \text{für } x \in [x_i, x_{i+1}], \quad i = 0, \dots, n-1,$$

heißt *stückweise lineare Interpolation* der Punkte  $(x_i, y_i)$ .

**Bemerkung.**

$$S_1(\Delta) = \{s : [a, b] \rightarrow \mathbb{R} \mid s \text{ stetig, } s \text{ linear auf } [x_i, x_{i+1}] \text{ für alle } i = 0, \dots, n-1\}$$

ist ein Vektorraum mit Dimension  $n + 1$ , **der Raum der linearen Splines**.

**Satz 5.1** (Charakterisierung der stückweise linearen Interpolation als Energieminimierer). Sei  $s \in S_1(\Delta)$ , und  $f \in C([a, b])$  stetig differenzierbar auf jedem der Teilintervalle  $[x_i, x_{i+1}]$ ,  $i = 0, \dots, n$ , mit

$$f(x_i) = s(x_i) \quad i = 0, \dots, n.$$

Dann gilt:

$$\int_a^b s'(x)^2 dx \leq \underbrace{\int_a^b f'(x)^2 dx}_{\text{»Energie von } f\text{«}}$$



*Beweis.* Wir definieren das »Energieskalarprodukt«

$$(f, g) := \int_a^b f'(x)g'(x)dx$$

zweier stückweise stetig differenzierbarer Funktionen  $f, g \in C([a, b])$ . Da  $s'(x)$  konstant auf  $[x_{i-1}, x_i]$  ist, also  $s'(x) =: c_i$  und  $f = g$  auf  $\Delta$  gilt, folgt

$$(s, f - s) = \sum_{i=1}^n \int_{x_{i-1}}^{x_i} s'(x) (f'(x) - g'(x))dx = \sum_{i=1}^n c_i (f - s)|_{x_{i-1}}^{x_i} = 0,$$

d.h.  $s$  ist orthogonal zu  $f - s$ , und damit:

$$\begin{aligned} (f, f) &= (s + (f - s), s + (f - s)) \\ &= (s, s) + (f - s, f - s) + 2(s, f - s) \geq (s, s). \end{aligned}$$

□

Zu einer beliebigen stetigen Funktion  $f \in C([a, b])$  betrachten wir nun die stückweise lineare Interpolation bzgl. des Gitters  $\Delta$ , d.h. die Funktion

$$\begin{aligned} L_1 f &:= s \in S_1(\Delta), \quad \text{mit} \\ s(x_i) &= f(x_i) \quad \text{für } i = 0, 1, \dots, n. \end{aligned}$$

Die Abbildung

$$L_1 : C([a, b]) \longrightarrow S_1(\Delta) \subset C([a, b])$$

ist eine lineare Projektion. Wir wollen nun eine Abschätzung des Approximationsfehlers  $f - L_1 f$  der stückweise linearen Interpolation herleiten. Dazu betrachten wir die vom  $L^2$ -Skalarprodukt

$$(f, g)_{L^2(a,b)} := \int_a^b f(x)g(x)dx$$

induzierte Norm

$$\|f\|_{L^2(a,b)} = \sqrt{(f, f)_{L^2(a,b)}} = \sqrt{\int_a^b f(x)^2 dx}.$$

Sei  $C^2([a, b])$  der Raum aller auf  $[a, b]$  zweimal stetig differenzierbaren Funktionen  $f$ , wobei die Ableitungen in den Randpunkten einseitig sind.

**Satz 5.2.** Für  $f \in C^2([a, b])$  gilt:

$$\begin{aligned} \|f - L_1 f\|_{L^2(a,b)} &\leq \frac{h^2}{2} \|f''\|_{L^2(a,b)}, \quad \text{und} \\ \|f' - (L_1 f)'\|_{L^2(a,b)} &\leq \frac{h^2}{\sqrt{2}} \|f''\|_{L^2(a,b)}, \end{aligned}$$

wobei  $h := \max_{1 \leq i \leq n} |x_i - x_{i-1}|$  die (maximale) Gitterweite ist.

Zum Beweis dieses Satzes benötigen wir:

**Lemma 5.3** (Poincaré-Ungleichung). Für  $g \in C^2([\alpha, \beta])$ ,  $-\infty < \alpha < \beta < \infty$ , mit  $g(\alpha) = g(\beta) = 0$  gilt:

$$\int_{\alpha}^{\beta} g^2 dx \leq \frac{(\beta - \alpha)^2}{2} \int_{\alpha}^{\beta} (g')^2 dx \leq \frac{(\beta - \alpha)^2}{2} \sqrt{\int_{\alpha}^{\beta} g^2 dx} \sqrt{\int_{\alpha}^{\beta} (g'')^2 dx}.$$

*Beweis.* Für  $x \in [\alpha, \beta]$  gilt nach dem Hauptsatz der Differential- und Integralrechnung und der Cauchy-Schwarz-Ungleichung:

$$\begin{aligned} g(x)^2 &= \left( g(\alpha) + \int_{\alpha}^x g'(y) dy \right)^2 = \left( 0 + (g', 1)_{L^2(\alpha, x)} \right)^2 \\ &\leq \|g'\|_{L^2(\alpha, x)}^2 \|1\|_{L^2(\alpha, x)}^2 = \|g'\|_{L^2(\alpha, x)}^2 \int_{\alpha}^x 1^2 dy \\ &\leq (x - \alpha) \|g'\|_{L^2(\alpha, \beta)}^2, \end{aligned}$$

also

$$\int_{\alpha}^{\beta} g(x)^2 dx \leq \int_{\alpha}^{\beta} (x - \alpha) dx \int_{\alpha}^{\beta} (g')^2 dx = \frac{(\beta - \alpha)^2}{2} \int_{\alpha}^{\beta} (g')^2 dx.$$

Somit ist die erste Ungleichung bewiesen. Die zweite Ungleichung folgt durch partielle Integration und die Cauchy-Schwarz-Ungleichung:

$$\begin{aligned} \int_{\alpha}^{\beta} (g')^2 dx &= g g' \Big|_{\alpha}^{\beta} - \int_{\alpha}^{\beta} g g'' dx \\ &= 0 - (g, g'')_{L^2(\alpha, \beta)} \leq \|g\|_{L^2(\alpha, \beta)} \|g''\|_{L^2(\alpha, \beta)}. \end{aligned}$$

□

*Beweis von Satz 5.2.* Sei  $g := f - L_1 f$ . Dann gilt:

$$g(x_i) = 0 \quad \text{für } i = 0, \dots, n.$$

Aus Lemma 5.3 folgt wegen  $|x_{i+1} - x_i| \leq h$ :

$$\begin{aligned} \int_{x_i}^{x_{i+1}} g^2 dx &\leq \left( \frac{h^2}{2} \right)^2 \int_{x_i}^{x_{i+1}} (g'')^2 dx, & \text{und (damit)} \\ \int_{x_i}^{x_{i+1}} (g')^2 dx &\leq \frac{h^2}{2} \int_{x_i}^{x_{i+1}} (g'')^2 dx & \text{für } i = 0, \dots, n-1. \end{aligned}$$

Wegen  $g'' = f''$  folgt die Behauptung durch Summation über  $i$ . □

ANWENDUNG AUF NUMERISCHE INTEGRATION (»QUADRATUR«): Sei  $f \in C([a, b])$ .

Gesucht ist

$$I[f] = \int_a^b f(x) dx.$$

ANSATZ:

Stützstellen  $x_i = a + ih$ , ( $i = 0, 1, \dots, n$ ), mit Maschenweite  $h = \frac{b-a}{n}$ .

$$\int_a^b f dx \approx \int_a^b s_n dx,$$

wobei  $s_n$  stückweise lineare Interpolation von  $f$  bzgl.  $\{x_0, \dots, x_n\}$ .

**Definition.**

$$\begin{aligned} T_n[f] &:= \int_a^b s_n dx = \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} s_n dx \\ &= h \cdot \sum_{i=0}^{n-1} \frac{f(x_{i+1}) + f(x_i)}{2} \\ &= h \cdot \left( \frac{1}{2} f(x_0) + f(x_1) + f(x_2) + \dots + f(x_{n-1}) + \frac{1}{2} f(x_n) \right) \end{aligned}$$

heißt *n-te (zusammengesetzte) Trapezformel* zur Approximation von  $I[f]$ .

FEHLERABSCHÄTZUNG:

Sei

$$\varepsilon_n := I[f] - T_n[f].$$

**Satz 5.4.** Für  $f \in C^2([a, b])$  existiert ein  $\xi \in (a, b)$  mit

$$\varepsilon_n = -\frac{b-a}{12} h^2 f''(\xi).$$

*Beweis.*

$$\varepsilon_n = \int_a^b (f - s_n) dx = \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} (f - s_n) dx \quad (5.1.1)$$

Durch zweimalige partielle Integration erhalten wir

$$\begin{aligned} \int_{x_i}^{x_{i+1}} (f - s_n) dx &= \int_{x_i}^{x_{i+1}} f(x) dx - \frac{x_{i+1} - x_i}{2} \cdot (f(x_{i+1}) - f(x_i)) \\ &= - \int_{x_i}^{x_{i+1}} f'(x) \left( x - \frac{x_{i+1} - x_i}{2} \right) dx \\ &= - \int_{x_i}^{x_{i+1}} f''(x) \underbrace{\frac{1}{2} (x - x_i)(x_{i+1} - x)}_{\geq 0} dx \\ &\geq - \max_{x \in [a, b]} f''(x) \int_{x_i}^{x_{i+1}} \frac{1}{2} (x - x_i)(x_{i+1} - x) dx = -\frac{h^3}{12} \max f'' \end{aligned}$$

für  $i = 0, 1, \dots, n-1$ , und damit gilt nach Gleichung (5.1.1) auch:

$$\varepsilon_n \geq -n \cdot \frac{h^3}{12} \max f'' = -\frac{b-a}{12} h^2 \max f''.$$

Analog folgt:

$$\varepsilon_n \geq -\frac{b-a}{12} h^2 \min f''.$$

Die Behauptung folgt nun aus dem Zwischenwertsatz, da  $f''$  stetig ist.  $\square$

Um eine Approximation höherer Ordnung zu erhalten, werden wir unten stückweise lineare Funktionen durch stückweise Polynome ersetzen. Als ersten Schritt überlegen wir uns deshalb, wie wir gegebene Funktionswerte auf einem Teilintervall durch ein Polynom interpolieren können.

## 5.2 Polynominterpolation

Seien

$$\begin{aligned} \Pi_n &= \left\{ x \mapsto \sum_{i=0}^n a_i x^i \mid a_0, \dots, a_n \in \mathbb{R} \right\} && \text{Polynome vom Grad } \leq n, && \text{und} \\ -\infty < a \leq x_0 < x_1 < \dots < x_n \leq b < \infty && \text{Stützstellen bzw. Knoten} \end{aligned}$$

innerhalb des Intervalls.

LAGRANGE-INTERPOLATIONSPROBLEM:

Gesucht ist  $p \in \Pi_n$  mit

$$p(x_i) = y_i \quad \text{für } i = 0, \dots, n, \quad (5.2.1)$$

wobei  $y_0, \dots, y_n \in \mathbb{R}$  gegeben sind.

NOTATION:  $p^{(j)}$  bezeichnet im folgenden die  $j$ -te Ableitung von  $p$ . VERALLGEMEINERTES /

HERMITE-INTERPOLATIONSPROBLEM:

Seien  $k_0, \dots, k_n \in \mathbb{N}_0$ . Gesucht ist ein  $p \in \Pi_{N-1}$ , wobei  $N = \sum_{i=0}^n (k_i + 1)$  mit

$$p^{(j)}(x_i) = y_{i,j} \quad \text{für alle } i, j \text{ mit } 0 \leq i \leq n, 0 \leq j \leq k_i, \quad (5.2.2)$$

und  $y_{i,j} \in \mathbb{R}$  gegeben sind. Für  $k_0 = \dots = k_n = 0$  ergibt sich das Lagrange-Interpolationsproblem als Spezialfall.

**Satz 5.5.** *Das verallgemeinerte Interpolationsproblem hat für alle  $y_{i,j} \in \mathbb{R}$ , ( $0 \leq i \leq n, 0 \leq j \leq k_i$ ), eine eindeutige Lösung  $p \in \Pi_{N-1}$ .*

*Beweis.* Wir betrachten die lineare Abbildung

$$\begin{aligned} l : \Pi_{N-1} &\rightarrow \mathbb{R}^N, \\ p &\mapsto (p^{(j)}(x_i))_{0 \leq i \leq n, 0 \leq j \leq k_i}. \end{aligned}$$

Die Abbildung  $l$  ist injektiv, denn

$$p^{(j)}(x_i) = 0 \quad \text{für alle } i, j \text{ mit } 0 \leq i \leq n, 0 \leq j \leq k_i.$$

Also hat  $p$  eine  $(k_i + 1)$ -fache Nullstelle bei  $x_i$  für alle  $i = 0, 1, \dots, n$ . Somit hat  $p$  insgesamt  $N = \sum_{i=0}^n (k_i + 1)$  Nullstellen (gezählt mit Vielfachheit). Da  $p \in \Pi_{N-1}$  ein Polynom vom Grad  $\leq N - 1$  ist, folgt  $p \equiv 0$ . Wegen  $\dim(\Pi_{N-1}) = N = \dim(\mathbb{R}^N)$  folgt, dass  $l$  bijektiv ist, d.h. zu jedem  $y = (y_{i,j}) \in \mathbb{R}^N$  existiert genau ein  $p \in \Pi_{N-1}$  mit  $l(p) = y$ .  $\square$

ANWENDUNG: POLYNOMINTERPOLATION DER FUNKTION  $f$

Sei  $f \in C^{\max k_i}([a, b])$ . Dann gibt es genau ein Polynom  $p \in \Pi_n$  mit der Eigenschaft

$$p^{(j)}(x_i) = f^{(j)}(x_i) \quad \text{für alle } i, j \text{ mit } 0 \leq i \leq n, 0 \leq j \leq k_i.$$

Im folgenden Satz wird der Approximationsfehler des Lagrange-Interpolationspolynoms abgeschätzt.

**Satz 5.6.** *Sei  $f \in C^{n+1}([a, b])$ , und  $p \in \Pi_n$  mit*

$$p(x_i) = f(x_i) \quad \text{für alle } i = 0, 1, \dots, n.$$

*Dann gibt es zu jedem  $x \in [a, b]$  ein  $\xi \in (x_0, x_n)$  mit*

$$f(x) - p(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)(x - x_1) \cdots (x - x_n). \quad (5.2.3)$$

**Bemerkung.** a)  $\omega(x) := (x - x_0) \cdots (x - x_n)$  heißt **Knotenpolynom** zu den Stützstellen / Knoten  $x_0, \dots, x_n$ . Beachte:  $\omega$  ist ein Polynom vom Grad  $n + 1$ .

b) Für den maximalen Interpolationsfehler auf dem Intervall  $[a, b]$  ergibt sich folgende Abschätzung:

$$\|f - p\| := \max_{x_0 \leq x \leq x_n} |f(x) - p(x)| \leq \frac{1}{(n+1)!} \|f^{(n+1)}\| \|\omega\|.$$

Man kann zeigen, dass  $\|\omega\|$  für  $[a, b] = [-1, 1]$  genau dann minimal ist, wenn gilt:

$$x_i = -\cos \frac{(2i+1)\pi}{2(n+1)} \quad \text{für alle } i = 0, \dots, n.$$

Diese Knoten  $x_0, \dots, x_n$  heißen **Čebyšev-Knoten**, da sie die Nullstellen des  $n$ -ten **Čebyšev-Polynoms**

$$T_n(x) = \cos(n \arccos(x))$$

sind.

- c) Für das verallgemeinerte Interpolationsproblem gelten entsprechende Aussagen, siehe z.B. Deuffhard/Hohmann: »Numerische Mathematik«. Das Knotenpolynom ist in diesem Fall zu ersetzen durch:

$$\omega(x) = \prod_{i=0}^n (x - x_i)^{k_i+1} \in \Pi_N.$$

*Beweis von Satz 5.6.* Für  $x \in \{x_0, \dots, x_n\}$  sind beide Seiten von (5.2.3) gleich 0. Sei also  $x \notin \{x_0, \dots, x_n\}$ . Für  $\lambda \in \mathbb{R}$  betrachten wir:

$$h(t) := f(t) - p(t) - \lambda \omega(t).$$

Es gilt  $h(x_i) = 0$  für  $i = 0, 1, \dots, n$ . Zudem gilt  $h(x) = 0$ , wenn wir  $\lambda$  geeignet wählen. In diesem Fall hat  $h$   $n+2$  Nullstellen, also hat nach dem Satz von Rolle  $h'$   $n+1$  Nullstellen,  $h''$   $n$  Nullstellen und schließlich  $h^{(n+1)}$  mindestens eine Nullstelle  $\xi \in (a, b)$ . Mit  $p \in \Pi_n$  folgt:

$$\begin{aligned} 0 &= h^{(n+1)}(\xi) = f^{(n+1)}(\xi) - p^{(n+1)}(\xi) - \lambda \omega^{(n+1)}(\xi) \\ &= f^{(n+1)}(\xi) - \lambda (n+1)!, \end{aligned}$$

also

$$\lambda = \frac{f^{(n+1)}(\xi)}{(n+1)!},$$

und damit:

$$0 = h(x) = f(x) - p(x) - \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega(x).$$

□

**Beispiel (Runge).** Sei

$$f(x) = \frac{1}{1+x^2}, \quad x \in [-5, 5],$$

und  $p_n(x)$  das Interpolationspolynom zu den äquidistanten Stützstellen  $x_i = -5 + 10 \frac{i}{n}$ , ( $i = 0, \dots, n$ ). Dann divergiert  $|f(x) - p_n(x)|$  für  $|x| \geq 3.64$  (ohne Beweis).

Das Interpolationspolynom ist im allgemeinen zur Approximation ungeeignet – insbesondere bei hohem Grad und äquidistanten Stützstellen! Die Ursache sind große Oszillationen des Interpolationspolynoms.

## Berechnung des Interpolationspolynoms

### Lagrange-Darstellung

LAGRANGE-GRUNDPOLYNOME:

$$L_i \in \Pi_n \quad \text{mit} \quad L_i(x_j) = \delta_{ij}, \quad i = 0, 1, \dots, n.$$

$\{L_0, L_1, \dots, L_n\}$  ist eine Basis von  $\Pi_n$ . Offenbar gilt:

$$L_i(x_j) = 0 \quad \text{für alle } j \neq i.$$

Daher gilt:

$$\prod_{j \neq i} (x - x_j) \mid L_i(x).$$

Wegen  $L_i(x_i) = 1$  erhalten wir somit die explizite Formel für die Lagrange-Grundpolynome:

$$L_i(x) = \prod_{j \neq i} \frac{x - x_j}{x_i - x_j}.$$

Die Lösung des Lagrange-Interpolationsproblems (5.2.1) lässt sich durch die Lagrange-Grundpolynome darstellen.

### Korollar.

$$p(x) = \sum_{i=0}^n y_i L_i(x)$$

ist eindeutige Lösung des Lagrange-Interpolationsproblems (5.2.1).

*Beweis.* Es gilt  $p \in \Pi_n$  mit  $p(x_j) = y_j$ . Die Eindeutigkeit folgt aus Satz 5.5. □

Die direkte Berechnung von  $p(x)$  via Lagrange-Darstellung ist instabil und aufwändig.

Alternativ *rekursive* Auswertung:

Seien  $p_{i,k} \in \Pi_{k-i}$  mit  $p_{i,k}(x_j) = y_j$  für alle  $i, j$  mit  $i \leq j \leq k$ .

**Lemma 5.7** (Aitken-Neville-Formel). *i) Für alle  $x \in \mathbb{R}$  und  $i$  mit  $0 \leq i \leq n$  gilt:*

$$p_{i,i}(x) = y_i$$

ii) Für alle  $x \in \mathbb{R}$  und  $i, k$  mit  $0 \leq i < k \leq n$  gilt:

$$p_{i,k}(x) = \frac{(x - x_i)p_{i+1,k}(x) - (x - x_k)p_{i,k-1}(x)}{x_k - x_i}.$$

*Beweis.* Der Beweis des Lemmas verbleibt als Übungsaufgabe. □

KONSEQUENZ: NEVILLE-SCHEMA ZUR AUSWERTUNG VON  $p(x)$ :

Baumdarstellung fehlt noch (!!)

Das Neville-Schema ist geeignet für Auswertung an einer Stelle ( $\frac{3}{2}n^2 + O(n)$  Multiplikationen bzw. Divisionen). Es ist jedoch unpraktisch zur Berechnung des gesamten Polynoms.

### Newton-Darstellung, dividierte Differenzen

NEWTON-BASIS VON  $\Pi_n$ :

$$1, \quad x - x_0, \quad (x - x_0)(x - x_1), \quad \dots, \quad (x - x_0)(x - x_1) \cdots (x - x_{n-1}).$$

$$p(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + \dots + a_n(x - x_0) \cdots (x - x_{n-1}).$$

**Bemerkung.** Effiziente Auswertung von  $p(x)$  mit HORNER-SCHEMA:

$$p(x) = (\cdots((a_n \cdot (x - x_{n-1}) + a_{n-1}) \cdot (x - x_{n-2}) + a_{n-2}) \cdots + a_1) \cdot (x - x_0) + a_0$$

benötigt nur  $n$  Multiplikationen!

Berechnung der Koeffizienten  $a_i$  des Interpolationspolynoms in der Newton-Basis:

**Satz 5.8** (Newton-Interpolationsformel). *Das eindeutige Interpolationspolynom  $p \in \Pi_n$  mit  $p(x_i) = y_i$  für alle  $i$  mit  $0 \leq i \leq n$  hat die Darstellung*

$$p(x) = f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) + \cdots + f[x_0, \dots, x_n](x - x_0)(x - x_1) \cdots (x - x_{n-1}), \quad \text{wobei}$$

$$f[x_i] := y_i,$$

$$f[x_i, x_{i+1}, \dots, x_k] := \frac{f[x_{i+1}, x_{i+2}, \dots, x_k] - f[x_i, x_{i+1}, \dots, x_{k-1}]}{x_k - x_i}, \quad 0 \leq i < k \leq n.$$

**Bemerkung.** a) Die Koeffizienten  $f[x_i, \dots, x_k]$  heißen *dividierte Differenzen*.



b) Ist  $y_i = f(x_i)$ , dann gilt:

$$\begin{aligned} f[x_i] &= f(x_i) && \text{Funktionswert} \\ f[x_i, x_{i+1}] &= \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} && \text{Differenzenquotient} \\ f[x_i, x_{i+1}, x_{i+2}] &= \frac{f[x_{i+1}, x_{i+2}] - f[x_i, x_{i+1}]}{x_{i+2} - x_i} && \text{etc.} \end{aligned}$$

c) Rekursive Berechnung aller dividierten Differenzen mit  $\frac{n(n+1)}{2}$  Divisionen:

Baumdarstellung fehlt noch (!!)

*Beweis von Satz 5.8.* Wir zeigen die Behauptung durch Induktion nach  $n$ . Wie oben sei  $p_{i,k} \in \Pi_{i,k}$  das Interpolationspolynom zu den Stützstellen  $x_i, \dots, x_k$ . Für  $n = 0$  gilt:

$$p(x) \equiv y_0 = f[x_0].$$

Den Induktionsschritt zeigen wir in zwei Schritten.

i) **BERECHNUNG DES FÜHRENDEN KOEFFIZIENTEN VON  $p(x)$  IN DER NEWTON-BASIS:**  
Nach der Rekursionsformel von Aitken und Neville (Lemma 5.7) und der Induktionsvoraussetzung gilt:

$$\begin{aligned} p(x) = p_{0,\dots,n}(x) &= \frac{(x - x_0) p_{1,\dots,n}(x) - (x - x_n) p_{0,\dots,n-1}(x)}{x_n - x_0} \\ &= \frac{f[x_1, \dots, x_n] (x - x_0) \cdots (x - x_{n-1})}{x_n - x_0} \\ &\quad - \frac{f[x_0, \dots, x_{n-1}] (x - x_0) \cdots (x - x_{n-2}) (x - x_n)}{x_n - x_0} + r(x) \\ &= f[x_0, \dots, x_n] x^n + \tilde{r}(x) \\ &= f[x_0, \dots, x_n] (x - x_0) \cdots (x - x_{n-1}) + q(x), \end{aligned}$$

wobei  $r, \tilde{r}$  und  $q$  Polynome vom Grad  $\leq n - 1$  sind.

ii) **IDENTIFIKATION VON  $q$ :**

Es gilt

$$y_i = p(x_i) = q(x_i) \quad \text{für } i = 0, \dots, n - 1,$$

also wegen der Eindeutigkeit des Interpolationspolynoms und nach Induktionsvoraussetzung:

$$q(x) = p_{0,n-1}(x) = f[x_0] + \cdots + f[x_0, \dots, x_{n-1}] (x - x_0) \cdots (x - x_{n-2}).$$

□

### 5.3 Numerische Integration

Sei  $-\infty < a < b < \infty$ ,  $w : (a, b) \rightarrow \mathbb{R}$  stetig, integrierbar und positiv («Gewichtsfunktion»).

GESUCHT:

$$I_w[f] := \int_a^b f(x) w(x) dx.$$

NOTATION: Zur Approximation von  $I_w[f]$  auf kleinen Intervallen  $[a, b]$  betrachten wir **Quadraturformeln** vom Typ

$$Q_{[a,b]}[f] = \sum_{i=0}^m w_i f(x_i)$$

mit **Knoten**  $x_i \in [a, b]$  und **Gewichten**  $w_i \in \mathbb{R}$ . Zu einer **Quadraturformel** erhalten wir ein (zusammengesetztes) **Quadraturverfahren**  $Q_n[f]$ , ( $n \in \mathbb{N}$ ), indem wir das Intervall  $[a, b]$  in  $n$  Teilintervalle der Länge  $h = \frac{b-a}{n}$  unterteilen, und die Quadraturformel auf jedem Teilintervall anwenden, d.h.:

$$Q_n[f] = \sum_{k=0}^{n-1} Q_{[a+kh, a+(k+1)h]}[f].$$

**Beispiel** (Mittelpunktsformel). Sei  $w \equiv 1$ .

$$Q_{[a,b]}^{MP}[f] = (b-a) f\left(\frac{a+b}{2}\right)$$

$$Q_n^{MP}[f] = \sum_{k=0}^{n-1} h f\left(a + \left(k + \frac{1}{2}\right)h\right) \quad h = \frac{b-a}{n}.$$

**Definition.** i) Die Quadraturformel  $Q_{[a,b]}$  hat **Exaktheitsgrad**  $r \in \mathbb{N} \cup \{0\}$ , falls

$$Q_{[a,b]}[p] = I_w[p]$$

für alle  $p \in \Pi_r$  gilt.

ii) Das Quadraturverfahren  $Q_n$  hat **Konsistenzordnung**  $s \geq 0$ , falls

$$|Q_n[f] - I_w[f]| = O(n^{-s}) = O(h^s)$$

für alle hinreichend glatten Funktionen  $f : [a, b] \rightarrow \mathbb{R}$  gilt.

**Beispiel** (Mittelpunkts- und Trapezformel). Ist  $f : [a, b] \rightarrow \mathbb{R}$  affin, dann gilt:

$$\int_a^b f(x) dx = (b-a) f\left(\frac{a+b}{2}\right) = (b-a) \frac{f(a) + f(b)}{2}.$$

Also haben die Mittelpunkts- und die Trapezregel beide Exaktheitsgrad 1. Mit einer Taylorentwicklung bei  $\frac{a+b}{2}$  ergibt sich zudem für  $f \in C^2$  und  $x \in [a, b]$ :

$$f(x) - f\left(\frac{a+b}{2}\right) = f'\left(\frac{a+b}{2}\right) \left(x - \frac{a+b}{2}\right) + f''(\xi(x)) \frac{1}{2} \left(x - \frac{a+b}{2}\right)^2$$

mit  $\xi(x) \in (a, b)$ . Integrieren dieser Gleichung liefert:

$$\begin{aligned} \left| \int_a^b f(x) dx - Q_{[a,b]}^{MP}[f] \right| &= \left| \int_a^b f'\left(\frac{a+b}{2}\right) \left(x - \frac{a+b}{2}\right) + f''(\xi(x)) \frac{1}{2} \left(x - \frac{a+b}{2}\right)^2 dx \right| \\ &\leq \sup_{x \in [a,b]} |f''(x)| \frac{1}{2} \int_a^b \left(x - \frac{a+b}{2}\right)^2 dx \\ &= \frac{(b-a)^3}{24} \|f''\|. \end{aligned}$$

Für das zusammengesetzte Quadraturverfahren

$$Q_n^{MP}[f] = \sum_{i=0}^{n-1} Q_{[a+kh, a+(k+1)h]}^{MP}[f], \quad h = \frac{b-a}{n},$$

erhalten wir damit:

$$\begin{aligned} \left| \int_a^b f(x) dx - Q_n^{MP}[f] \right| &= \left| \sum_{k=0}^{n-1} \int_{a+kh}^{a+(k+1)h} f(x) dx - Q_{[a+kh, a+(k+1)h]}^{MP}[f] \right| \\ &\leq \frac{n h^3}{24} \|f''\| = \frac{b-a}{24} h^2 \|f''\| \end{aligned}$$

für alle  $f \in C^2$ . Das Mittelpunktsverfahren hat also – ebenso wie das Trapezverfahren (siehe oben) – Konsistenzordnung 2.

## Newton-Cotes-Formeln

Wie erhält man Quadraturformeln mit höherem Exaktheitsgrad? Wir betrachten eine Quadraturformel

$$Q_{[a,b]}[f] = \sum_{i=0}^m w_i f(x_i)$$

mit Stützstellen  $x_0, \dots, x_m$  zur Approximation von  $\int_a^b f(x) w(x) dx$ .

**Satz 5.9.** Die Quadraturformel hat genau dann Exaktheitsgrad  $\geq m$ , wenn

$$w_i = \int_a^b L_i(x) w(x) dx, \quad i = 0, 1, \dots, m, \quad (5.3.1)$$

gilt, wobei  $L_i$  das  $i$ -te Lagrange-Grundpolynom ist.

*Beweis.* Ist der Exaktheitsgrad  $\geq m$ , dann gilt:

$$\int_a^b L_i(x) w(x) dx = Q_{[a,b]}[L_i] = \sum_{j=0}^m w_j L_i(x_j) = \sum_{j=0}^m w_j \delta_{ij} = w_i, \quad i = 0, 1, \dots, m.$$

Umgekehrt folgt aus (5.3.1) für  $p = L_i$ , ( $i = 0, 1, \dots, m$ ):

$$Q_{[a,b]}[p] = \int_a^b p(x) w(x) dx,$$

also für alle  $p \in \Pi_m = \text{span}\{L_i \mid i = 0, \dots, m\}$ . □

**Definition.** Die *Quadraturformel mit äquidistanten Knoten*

$$x_i = a + i \frac{b-a}{m}, \quad i = 0, 1, \dots, m,$$

bzw.

$$x_i = a + \left(i + \frac{1}{2}\right) \frac{b-a}{m+1}, \quad i = 0, 1, \dots, m,$$

und Gewichten

$$w_i = \int_a^b L_i(x) w(x) dx$$

heißt **geschlossene** bzw. **offene Newton-Cotes-Formel**.

**Beispiel.** Für  $w \equiv 1$ ,  $I[f] = \int_a^b f(x) dx$ , ergeben sich unter anderem folgende Newton-Cotes-Formeln:

- $m = 0$ , *offen*:

$$x_0 = \frac{a+b}{2}, \quad w_0 = \int_a^b 1 dx = b-a,$$

$$Q[f] = (b-a) f\left(\frac{a+b}{2}\right), \quad \text{»Mittelpunktsformel«.}$$

- $m = 1$ , *abgeschlossen*:

$$x_0 = a, \quad x_1 = b, \quad w_0 = \int_a^b \frac{x-b}{a-b} dx = \frac{b-a}{2} = w_1,$$

$$Q[f] = \frac{b-a}{2} (f(a) + f(b)), \quad \text{»Trapezformel«.}$$

- $m = 1$ , *offen*:

$$Q[f] = \frac{b-a}{2} \left( f\left(\frac{3}{4}a + \frac{1}{4}b\right) + f\left(\frac{1}{4}a + \frac{3}{4}b\right) \right).$$

- $m = 2$ , *abgeschlossen*:

$$x_0 = a, \quad x_1 = \frac{a+b}{2}, \quad x_2 = b,$$

$$Q[f] = \frac{b-a}{6} \left( f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right), \quad \text{»Simpson-Formel«.}$$

- $m = 2$ , *offen*:

$$Q[f] = \frac{b-a}{3} \left( \frac{3}{8}f\left(\frac{5}{6}a + \frac{1}{6}b\right) + \frac{2}{8}f\left(\frac{a+b}{2}\right) + \frac{3}{8}f\left(\frac{1}{6}a + \frac{5}{6}b\right) \right).$$

Bei den Formeln höheren Grades erhält man zum Teil negative Gewichte, was zu numerischer Instabilität führen kann. Der Exaktheitsgrad der Newton-Cotes-Formeln ist  $\geq m$  nach Satz 5.9. Die Mittelpunktsformel hat sogar Exaktheitsgrad 1 (statt 0). Allgemeiner gilt:

**Satz 5.10.** Ist  $Q[f] = \sum_{i=0}^m w_i f(x_i)$  eine zu  $\frac{a+b}{2}$  symmetrische Quadraturformel für  $\int_a^b f(x)dx$ , d.h.

$$x_{m-i} = b - x_i \quad \text{und} \quad w_{m-i} = w_i \quad \text{für } i = 0, 1, \dots, m,$$

mit Exaktheitsgrad  $\geq 2q$ , dann hat  $Q$  mindestens Exaktheitsgrad  $2q + 1$ .

*Beweis.* Sei  $p(x) := \left(x - \frac{a+b}{2}\right)^{2q+1}$ . Aufgrund der Symmetrie gilt:

$$Q[p] = 0 = \int_a^b p(x)dx. \quad (5.3.2)$$

Da  $Q$  Exaktheitsgrad  $\geq 2q$  hat, gilt (5.3.2) auch für alle  $p \in \Pi_{2q}$ , also für alle  $p \in \Pi_{2q+1} = \text{span}(\Pi_{2q}, (\cdot - \frac{a+b}{2})^{2q+1})$ .  $\square$

**Beispiel.** Die Simpson-Formel hat Exaktheitsgrad  $\geq 2$ , also hat  $Q$  Exaktheitsgrad  $\geq 3$ . Wegen  $Q[x^4] \neq \int_a^b x^4 dx$  ist der Exaktheitsgrad genau gleich 3.

Das zusammengesetzte Simpson-Verfahren ist:

$$\begin{aligned} Q_n^{\text{Simpson}}[f] &= \sum_{k=0}^{n-1} Q_{[a+kh, a+(k+1)h]}^{\text{Simpson}}[f], \quad h = \frac{b-a}{n}, \\ &= \frac{h}{6} \left[ f(a) + 4f\left(a + \frac{h}{2}\right) + 2f(a+h) + 4f\left(a + \frac{3}{2}h\right) + 2f(a+2h) \right. \\ &\quad \left. + \dots + 4f\left(b - \frac{h}{2}\right) + f(b) \right]. \end{aligned}$$

**Satz 5.11** (Konsistenzordnung des Simpson-Verfahrens). Für  $f \in C^4([a, b])$  und  $n \in \mathbb{N}$  gilt:

$$\left| Q_n^{\text{Simpson}}[f] - \int_a^b f(x) dx \right| \leq \frac{b-a}{180} \|f^{(4)}\|_{[a,b]} \left(\frac{h}{2}\right)^4.$$

*Beweis.* Wir zeigen:

$$\left| Q_{[c, c+h]}^{\text{Simpson}}[f] - \int_c^{c+h} f(x) dx \right| \leq \frac{h}{180} \|f^{(4)}\|_{[a,b]} \left(\frac{h}{2}\right)^4 \quad \text{für alle } c \in [a, b-h]. \quad (5.3.3)$$

Die Behauptung folgt dann durch Summation dieser Gleichungen für  $c = a+kh$ , ( $0 \leq k \leq n-1$ ).

Sei  $p \in \Pi_3$  das Hermite-Interpolationspolynom mit:

$$p(c) = f(c), \quad p\left(c + \frac{h}{2}\right) = f\left(c + \frac{h}{2}\right), \quad p'\left(c + \frac{h}{2}\right) = f'\left(c + \frac{h}{2}\right), \quad p(c+h) = f(c+h).$$

Da die Simpson-Formel Exaktheitsgrad 3 hat, gilt:

$$Q_{[c, c+h]}^{\text{Simpson}}[f] = Q_{[c, c+h]}^{\text{Simpson}}[p] = \int_c^{c+h} p(x) dx,$$

also gilt mit Bemerkung c) zu Satz 5.6:

$$\begin{aligned} \left| Q_{[c, c+h]}^{\text{Simpson}}[f] - \int_c^{c+h} f(x) dx \right| &= \left| \int_c^{c+h} (f(x) - p(x)) dx \right| \\ &= \left| \int_c^{c+h} \frac{f^{(4)}(\xi)}{4!} (x-c) \underbrace{\left(x - \left(c + \frac{h}{2}\right)\right)^2 (x - (c+h))}_{\leq 0 \text{ auf } [c, c+h]} dx \right| \\ &\leq \frac{\|f^{(4)}\|_{[a,b]}}{4!} \int_c^{c+h} (x-c) \left(x - \left(c + \frac{h}{2}\right)\right)^2 (c+h-x) dx \\ &= \frac{h}{180} \|f^{(4)}\|_{[a,b]} \left(\frac{h}{2}\right)^4. \end{aligned}$$

Der Wert des Integrals ergibt sich beispielsweise mit der Substitution  $t = \frac{x-c-h/2}{h/2}$ , die das Intervall  $[c, c+h]$  auf  $[-1, 1]$  abbildet.  $\square$

## Gauß-Quadratur

Seien  $a \leq x_0 < x_1 < \dots < x_m \leq b$  Knoten,

$$w_i = \int_a^b L_i(x) w(x) dx, \quad i = 0, 1, \dots, m.$$

Dann hat

$$Q_{[a,b]}[f] = \sum_{i=0}^m w_i f(x_i)$$

Exaktheitsgrad  $\geq m$ .

Bisher haben wir die Knoten fest vorgegeben (äquidistant). Andererseits haben wir schon bei der Polynominterpolation gesehen, daß sich für äquidistante Knoten im Allgemeinen keine optimalen Approximationseigenschaften ergeben. Kann man durch geeignete Wahl der Knoten einen höheren Exaktheitsgrad als  $m$  erreichen? Was ist der maximale Exaktheitsgrad? Um diese Fragen zu beantworten, betrachten wir das *Knotenpolynom*

$$\omega(x) := \prod_{j=0}^m (x - x_j).$$

Es gilt  $\omega \in \Pi_{m+1}$ .

**Bemerkung.** Es gilt:

$$L_i(x) = \prod_{j \neq i} \frac{x - x_j}{x_i - x_j} = \frac{\omega(x)}{(x - x_i) \omega'(x_i)}.$$

**Satz 5.12.** Sei  $k \in \mathbb{N}$ . Die Quadraturformel  $Q_{[a,b]}$  hat genau dann Exaktheitsgrad  $m + k$ , wenn  $w$  orthogonal zu  $\Pi_{k-1}$  bzgl. des gewichteten  $L^2$ -Skalarprodukts

$$(f, g)_w := \int_a^b f(x)g(x)w(x)dx$$

ist.

*Beweis.* Hat  $Q$  Exaktheitsgrad  $m + k$ , dann gilt für  $p \in \Pi_{k-1}$ :

$$(\omega, p)_w = \int_a^b \underbrace{\omega(x)p(x)}_{\in \Pi_{m+k}} w(x)dx = Q_{[a,b]}[\omega p] = 0,$$

da  $\omega(x_i) = 0$  für alle  $i$  mit  $0 \leq i \leq m$  ist. Umgekehrt: Gilt  $(\omega, p)_w = 0$  für alle  $p \in \Pi_{k-1}$ , dann ist

$$\int_a^b \omega(x)p(x)w(x)dx = 0 = Q_{[a,b]}[\omega p]$$

für alle  $p \in \Pi_{k-1}$ , also auch

$$\int_a^b (\omega(x)p(x) + q(x))w(x)dx = Q_{[a,b]}[\omega p + q]$$

für alle  $p \in \Pi_{k-1}$  und  $q \in \Pi_m$ . Da sich jedes Polynom aus  $\Pi_{m+k}$  in der Form  $\omega p + q$  mit  $q \in \Pi_{k-1}$  und  $q \in \Pi_m$  darstellen lässt, folgt, daß  $Q_{[a,b]}$  Exaktheitsgrad  $m + k$  hat.  $\square$

**Korollar.** i) Der Exaktheitsgrad ist stets  $\leq 2m + 1$ .

ii) Der Exaktheitsgrad ist genau dann  $2m + 1$ , wenn  $\omega$  orthogonal zu  $\Pi_m$  bzgl.  $(\cdot, \cdot)_w$  ist.

*Beweis.* i) Angenommen, der Exaktheitsgrad ist  $\geq 2m + 2$ . Dann ist  $\omega$  orthogonal zu  $\Pi_{m+1}$  und es folgt der Widerspruch

$$(\omega, \omega)_w = \int_a^b \omega(x)^2 w(x) dx = 0.$$

ii) Die zweite Behauptung folgt sofort aus dem Satz. □

Zu klären bleibt, ob und wie wir Knoten  $x_0, \dots, x_m$  finden können, für die das Knotenpolynom  $\omega$  orthogonal zu  $\Pi_m$  ist. Ein normiertes Polynom  $\omega_{m+1} \in \Pi_{m+1}$  mit  $\omega_{m+1} \perp \Pi_m$  kann man mithilfe des Schmidtschen Orthogonalisierungsverfahrens konstruieren:

$$\begin{aligned} \omega_0(x) &:= 1 && \text{für alle } x \in [a, b], \\ \omega_{m+1}(x) &:= x^{m+1} - \sum_{i=0}^m (x^{m+1}, \omega_i)_w \omega_i(x) && \text{für alle } m \geq 0. \end{aligned}$$

**Satz 5.13** (Orthogonalpolynome). i)  $\omega_{m+1}$  ist das eindeutige Polynom in  $\Pi_{m+1}$  mit  $\omega_{m+1} \perp \Pi_m$ .

ii) Die Orthogonalpolynome  $\omega_K$  erfüllen die Dreitermrekursion

$$\begin{aligned} \omega_{-1}(x) &= 0, \\ \omega_0(x) &= 1, \\ \omega_{m+1}(x) &= (x - \alpha_m)\omega_m(x) - \beta_m \omega_{m-1}(x), \quad (m \geq 0), \end{aligned}$$

mit

$$\alpha_m = \frac{(x\omega_m, \omega_m)_w}{(\omega_m, \omega_m)_w}, \quad \beta_{m+1} = \frac{(\omega_{m+1}, \omega_{m+1})_w}{(\omega_m, \omega_m)_w}, \quad \beta_0 \text{ beliebig.}$$

iii)  $\omega_m$  hat  $m + 1$  einfache reelle Nullstellen  $x_0 < x_1 < \dots < x_m$  im offenen Intervall  $(a, b)$ .

*Beweis.* i) EINDEUTIGKEIT: Sei  $\tilde{\omega}_{m+1} \in \Pi_{m+1}$  normiert mit  $\tilde{\omega}_{m+1} \perp \Pi_m$ . Dann ist  $\omega_{m+1} - \tilde{\omega}_{m+1} \in \Pi_m$  und orthogonal zu  $\Pi_m$ . Damit folgt  $\omega_{m+1} - \tilde{\omega}_{m+1} = 0$  und  $\tilde{\omega}_{m+1} = \omega_{m+1}$ .

ii) Mit Induktion rechnet man nach, daß die durch die Dreitermrekursion definierten normierten Polynome  $\omega_{m+1} \in \Pi_{m+1}$  orthogonal sind. Die Behauptung folgt dann aus der Eindeutigkeit der Orthogonalpolynome.



iii) Seien  $x_0, x_1, \dots, x_k$ , ( $k \leq m$ ), die Nullstellen von  $\omega_{m+1}$  mit ungerader Vielfachheit, die im offenen Intervall  $(a, b)$  liegen. Zu zeigen ist  $k = m$ . Daraus folgt die Existenz von  $m + 1$  einfachen Nullstellen. Angenommen, es gilt  $k < m$ . Dann ist das Polynom

$$q(x) := \underbrace{\prod_{i=0}^k (x - x_i)}_{\text{Grad } k+1}$$

orthogonal zu  $\omega_{m+1}$ . Da  $\omega_{m+1}(x)$  und  $q(x)$  bei  $x_0, \dots, x_k$  jeweils das Vorzeichen wechseln, hat das Produkt  $\omega_{m+1}(x) q(x)$  konstantes Vorzeichen. Somit folgt der Widerspruch

$$\int_a^b \omega_{m+1}(x) q(x) dx = 0.$$

□

Damit können wir nun im Prinzip Quadraturformeln mit maximalem Exaktheitsgrad  $2m + 1$  angeben:

**Definition.** Seien  $x_0, \dots, x_m$  die Nullstellen des Orthogonalpolynoms  $\omega_{m+1}$ . Die Quadraturformel

$$Q_{[a,b]}^m[f] = \sum_{i=0}^m w_i f(x_i),$$

mit Gewichten

$$w_i = \int_a^b L_i(x) w(x) dx = \int_a^b \frac{\omega_{m+1}(x)}{(x - x_i) \omega'_{m+1}(x)} w(x) dx$$

heißt **Gauß-Formel vom Grad  $m + 1$** .

**Bemerkung.** a) Die Gauß-Formel hat nach Konstruktion Exaktheitsgrad  $2m + 1$ .

b) Die Berechnung der Stützstellen ist im Allgemeinen nichttrivial – aber diese müssen zu einer gegebenen Gewichtsfunktion nur einmal berechnet werden, und können dann immer wieder verwendet werden.

**Beispiel.** Sei ohne Beschränkung der Allgemeinheit  $[a, b] = [-1, 1]$  und

$$w(x) = \frac{1}{\sqrt{1 - x^2}}$$

Dann ist

$$(f, g)_w = \int_{-1}^1 f(x) g(x) \frac{1}{\sqrt{1 - x^2}} dx$$

Mit der Substitution  $x = \cos \theta$ ,  $\sqrt{1-x^2} = \sin \theta$  und  $dx = -\sin \theta d\theta$  folgt

$$(f, g)_w = \int_0^\pi \underbrace{f(\cos \theta)}_{=: \tilde{f}(\theta)} \underbrace{g(\cos \theta)}_{=: \tilde{g}(\theta)} d\theta = (\tilde{f}, \tilde{g})_{L^2(0, \pi)}.$$

Die Funktionen

$$\tilde{T}_k(\theta) = \cos(k\theta), \quad k = 0, 1, 2, \dots,$$

sind orthogonal bzgl.  $(\cdot, \cdot)_{L^2(0, \pi)}$ . Also sind die Funktionen

$$T_k(x) = \cos(k \cdot \arccos(x)), \quad k = 0, 1, 2, \dots,$$

orthogonal bzgl.  $(\cdot, \cdot)_w$ . Aus den Additionstheoremen für die Kosinusfunktion folgt die Dreitermrekursionsformel

$$\begin{aligned} T_0(x) &= 1, \\ T_1(x) &= x, \\ T_{k+1}(x) &= 2x T_k(x) - T_{k-1}(x). \end{aligned}$$

Insbesondere ist  $T_k$  ein Polynom  $k$ -ten Grades, das  $k$ -te *Čebyšev-Polynom*. Durch Normierung erhalten wir die eindeutigen normierten Orthogonalpolynome

$$\omega_{k+1}(x) = 2^{-k} T_{k+1}(x).$$

Die Nullstellen sind die *Čebyšev-Knoten*

$$x_i = -\cos \frac{(2i+1)\pi}{2(k+1)}, \quad 0 \leq i \leq k,$$

für die Gewichte ergibt sich

$$w_i = \frac{\pi}{k+1} \quad \text{für alle } i.$$

Andere wichtige Gewichtsfunktionen sind zum Beispiel  $w(x) = 1$  und  $w(x) = e^{-x^2}$ .

	Gauß-Legendre	Gauß-Čebyšev	Gauß-Hermite
Intervall	$[-1, 1]$	$[-1, 1]$	$(-\infty, \infty)$
$w(x)$	1	$\frac{1}{\sqrt{1-x^2}}$	$e^{-x^2}$
Orthogonalpolynome	$L_k(x) = \frac{1}{2^k k!} \frac{d^k}{dx^k} (x^2 - 1)^k$	$T_k(x) = \cos(k \arccos(x))$	$H_k(x) = (-1)^k e^{x^2} \frac{d^k}{dx^k} e^{-x^2}$
	Legendre-Polynome	Čebyšev-Polynome	Hermite-Polynome
Dreitermrekursion	$L_{k+1}(x) = \frac{2k+1}{k+1} x L_k(x) - \frac{k}{k+1} L_{k-1}(x)$	$T_{k+1}(x) = 2x T_k(x) - T_{k-1}(x)$	$H_{k+1}(x) = 2x H_k(x) - 2k H_{k-1}(x)$
Normierung	$\omega_{k+1}(x) = \frac{2^{k+1}}{(2 \binom{k+1}{k+1})} L_k(x)$	$\omega_{k+1}(x) = 2^{-k} T_{k+1}(x)$	$\omega_{k+1}(x) = 2^{-(k+1)} H_{k+1}(x)$
Nullstellen	z.B. n=4: $x_3 = -x_0 = \sqrt{\frac{3}{7} + \frac{2}{7} \sqrt{\frac{6}{5}}}$ $x_2 = -x_1 = \sqrt{\frac{3}{7} - \frac{2}{7} \sqrt{\frac{6}{5}}}$	$x_i = -\cos \frac{(2i+1)\pi}{2 \cdot (k+1)}$	
Gewichte	$\omega_3 = \omega_0 = \frac{18 - \sqrt{30}}{36}$ $\omega_2 = \omega_1 = \frac{18 + \sqrt{30}}{36}$	$\omega_i = \frac{\pi}{k+1}$	

**Satz 5.14.** Für  $s \in \mathbb{N}$  existieren konstante  $C_L, C_C > 0$  mit

$$\left| Q_{Leg}^m[f] - \int_a^b f(x) dx \right| \leq C_L \cdot m^{-s} \|f\|_S$$

bzw.

$$\left| Q_{Ceb}^m[f] - \int_a^b f(x) \frac{1}{\sqrt{1-x^2}} dx \right| \leq C_C \cdot m^{-s} \|f\|_{S,w}$$

für alle  $f \in C^S([-1, 1])$ , wobei

$$\|f\|_{S,w} := \sqrt{\sum_{k=0}^s \|f^{(k)}\|_w^2}, \quad \|f\|_S := \|f\|_{S,1}.$$

Ohne Beweis.

*Konsequenz:*

Gauß-Verfahren erreicht beliebig hohe Ordnung, wenn  $f$  hinreichend regulär ist.

## 5.4 Kubische Splines

Sei  $[a, b] \subset \mathbb{R}$  ein Intervall,  $a = x < x_1 < \dots < x_n = b$  Stützstellen,  $\Delta = \{x_0, x_1, \dots, x_n\}$  ein Gitter und  $m \in \mathbb{N}$ .

**Definition.**

$$S_m(\Delta) := \{s \in C^{m-1}([a, b]) \mid s|_{[x_{i-1}, x_i]} \text{ ist ein Polynom vom Grad } m \text{ für alle } i = 1, \dots, n\}$$

heißt **Raum der Splines vom Grad  $m$  bzgl. des Gitters  $\Delta$** .

$m = 1$ : stückweise lineare stetige Funktionen,

$m = 3$ : kubische Splines.

Falls  $s \in C^2$  ist, ist die Krümmung der Kurve  $(x, s(x))$  stetig und die Kurve wird als glatt empfunden (Anwendungen bei CAD, Schienenverlegung etc.).

**Beispiel.** Sei  $\varrho(x) = ((x - x_i)^+)^m$ ,  $a^+ := \max(a, 0)$ . Dann ist  $s \in S_m(\Delta)$ .

**Lemma 5.15.**  $S_m(\{x_0, \dots, x_n\})$  ist ein  $(m + n)$ -dimensionaler Vektorraum mit Basis

$$\underbrace{1, x - x_0, (x - x_0)^2, \dots, (x - x_0)^m}_{m + 1 \text{ Funktionen}}, \underbrace{((x - x_1)^+)^m, \dots, ((x - x_{n-1})^+)^m}_{n - 1 \text{ Funktionen}}.$$

*Beweis.* Wir beweisen das Lemma durch vollständige Induktion nach  $n$ .

Für  $n = 1$  ist  $S_m(\{x_0, x_1\}) = \Pi_m$  mit Basis  $1, x - x_0, \dots, (x - x_0)^m$ .

Sei die Behauptung für  $n$  wahr. Wegen  $((x - x_n)^+)^m \equiv 0$  auf  $[x_0, x_n]$  folgt die lineare Unabhängigkeit der Funktionen  $1, \dots, ((x - x_n)^+)^m$  aus der linearen Unabhängigkeit der Funktionen  $1, \dots, ((x - x_{n-1})^+)^m$ , die nach Induktionsvoraussetzung erfüllt ist. Zu zeigen bleibt, daß sich jeder Spline  $m$ -ter Ordnung als Linearkombination der angegebenen Basisfunktionen darstellen lässt, dass also aus  $s \in S_m(\{x_0, \dots, x_{n+1}\})$  folgt:

$$s|_{[x_0, x_n]} \in S_m(\{x_0, \dots, x_n\}).$$

Nach Induktionsvoraussetzung gibt es  $a_k, b_i \in \mathbb{R}$ , so dass gilt:

$$s(x) = \underbrace{\sum_{k=0}^m a_k (x - x_0)^k + \sum_{i=1}^{n-1} b_i ((x - x_i)^+)^m}_{=: \tilde{s}(x)} \quad \text{für alle } x \in [x_0, x_n].$$

Wegen  $s \in C^{m-1}$  gilt:

$$s^{(k)}(x_n) - \tilde{s}^{(k)}(x_n) = 0, \quad 0 \leq k \leq m-1.$$

Da  $s - \tilde{s}$  auf  $[x_n, x_{n+1}]$  ein Polynom vom Grad  $m$  ist, gibt es ein  $b_n \in \mathbb{R}$  mit:

$$s(x) - \tilde{s}(x) = b_n (x - x_n)^m \quad \text{für alle } x \in [x_n, x_{n+1}],$$

also

$$s(x) = \tilde{s}(x) + b_n ((x - x_n)^+)^m \quad \text{für alle } x \in [a, b].$$

□

### B-Splines (»Bell splines«)

Die in Lemma 5.15 angegebene Basis ist zur numerischen Berechnung eines Splines zu gegebenen Interpolationspunkten  $(x_i, y_i)$  wenig geeignet – wir betrachten daher nun eine andere Basis. Dazu nehmen wir an, daß die Stützstellen äquidistant sind:

$$x_i = a + i h, \quad h = \frac{b - a}{n}.$$

Wir setzen

$$B_0(x) := \begin{cases} 1 & \text{für } |x| \leq \frac{1}{2}, \\ 0 & \text{für } |x| > \frac{1}{2}, \end{cases}$$

$$B_{m+1}(x) := \int_{x-\frac{1}{2}}^{x+\frac{1}{2}} B_m(t) dt.$$

Zum Beispiel ist für  $m = 3$ :

$$B_3(x) = \frac{1}{6} \cdot \begin{cases} (2 - |x|)^3 - 4(1 - |x|^3) & \text{für } |x| \leq 1, \\ (2 - |x|)^3 & \text{für } 1 \leq |x| \leq 2, \\ 0 & \text{für } |x| \geq 2. \end{cases}$$

**Lemma 5.16.** Für alle  $m \in \mathbb{N}$  gilt:

i)  $B_m \in C^{m-1}(\mathbb{R})$  mit  $B_m(x) = 0$  für alle  $|x| \geq \frac{m+1}{2}$ .

ii)  $B_m \in S_m(\Delta_m)$ , wobei

$$\Delta_m = \left\{ -\frac{m+1}{2}, -\frac{m-1}{2}, \dots, \frac{m-1}{2}, \frac{m+1}{2} \right\}.$$

*Beweis.* Der Beweis erfolgt durch vollständige Induktion nach  $m$ . Der Fall  $m = 1$  ist klar. Sei die Induktionsbehauptung für  $m$  wahr. Wegen  $B_m \in C^{m-1}$  folgt  $B_{m+1} \in C^m$ , und

$$B'_{m+1}(x) = B_m\left(x + \frac{1}{2}\right) - B_m\left(x - \frac{1}{2}\right).$$

Der Minuend ist ein Polynom vom Grad  $\leq m$  auf dem Intervall  $I$  falls  $(I + \frac{1}{2})$  und  $\Delta_m$  disjunkt sind. Der Subtrahend ist ein Polynom vom Grad  $m$  auf dem Intervall  $I$  falls  $(I - \frac{1}{2})$  und  $\Delta_m$  disjunkt sind. Daher ist  $B'_{m+1}$  ein Polynom vom Grad  $\leq m$  auf offenen Teilintervallen von  $\Delta_{m+1}$ . Damit ist  $B_{m+1}$  ein Polynom vom Grad  $\leq m + 1$  auf Teilintervallen von  $\Delta_{m+1}$ .  $\square$

### Konstruktion einer Basis von $S_m(\Delta)$

*Annahme:* äquidistantes Gitter

$$\Delta = \{x_0, x_1, \dots, x_n\}, \quad x_i = a + ih, \quad h = \frac{b-a}{n}.$$

Wir definieren

$$B_{m,i}(x) := B_m\left(\frac{x - (a + ih)}{h}\right).$$

**Satz 5.17.** Die  $m + n$  B-Splines  $B_{m,i}$ ,  $i = -\frac{m-1}{2}, -\frac{m-1}{2} + 1, \dots, n + \frac{m-1}{2}$ , bilden eine Basis von  $S_m(\Delta)$ .

*Beweis.* Die lineare Unabhängigkeit von  $B_{m,i} \in S_m(\Delta)$  folgt durch Induktion. Wegen  $\dim(S_m(\Delta)) = m + n$  bilden die  $B_{m,i}$  eine Basis.  $\square$

**Beispiel.**  $m = 1$ :

$$s(x) = \sum_{i=0}^n y_i B_{1,i}(x)$$

ist stückweise lineare Interpolation der Punkte  $(x_i, y_i)$ .

**Interpolationsproblem für kubische Splines ( $m = 3$ )***Gegeben:*

$$y_0, \dots, y_n \in \mathbb{R}, y_{0,1}, y_{n,1} \in \mathbb{R}.$$

*Gesucht:*

$$s \in S_3(\Delta) \text{ mit } s(x_i) = y_i \quad \text{für alle } i = 0, 1, \dots, n, \quad (5.4.1)$$

sowie einer der folgenden Randbedingungen:

(i) *Hermite:*

$$s'(a) = y_{0,1}, \quad s'(b) = y_{n,1},$$

(ii) *periodisch:*

$$s'(a) = s'(b), \quad s''(a) = s''(b),$$

(iii) *natürlich:*

$$s''(a) = s''(b) = 0.$$

**Bemerkung.** Es gilt  $\dim(S_3(\Delta)) = n+3$ . Da in (5.4.1) nur  $n+1$  Interpolationswerte  $(y_0, \dots, y_n)$  vorgegeben sind, werden zwei zusätzliche Randbedingungen benötigt, um ein eindeutig lösbares Interpolationsproblem zu erhalten. Im folgenden betrachten wir nur Hermite-Randbedingungen – für die anderen Arten von Randbedingungen gelten aber ähnliche Aussagen.

Ähnlich wie die linearen Splines erfüllen auch die kubischen Splines eine Energieminimierungseigenschaft:

**Satz 5.18.** Sei  $s \in S_3(\Delta)$  und  $f \in C^4([a, b])$  mit  $f(x_i) = s(x_i)$  für alle  $i$  mit  $0 \leq i \leq n$ ,  $f'(a) = s'(a)$  und  $f'(b) = s'(b)$ . Dann gilt:

$$i) (f'' - s'', l) = 0 \text{ für alle } l \in S_1(\Delta),$$

$$ii) \|s''\| \leq \|f''\|,$$

$$iii) \|f'' - s''\| \leq \frac{h^2}{2} \cdot \|f^{(4)}\|,$$

$$iv) \|f - s\| \leq \frac{h^4}{4} \cdot \|f^{(4)}\|,$$

wobei  $(f, g) = \int_a^b f(x)g(x) dx$  das  $L^2(a, b)$ -Skalarprodukt und  $\|\cdot\|$  die zugehörige Norm bezeichnet.

**Bemerkung.**

$$E[f] = \|f''\|^2 = \int_a^b f''(x)^2 dx$$

ist die Biegeenergie eines Stabes oder einer Holzlatte (englisch »spline«) mit Parametrisierung  $(x, f(x))$  (in erster Näherung für kleine Auslenkungen). Der kubische Spline minimiert  $E$  unter allen Kurven durch die Punkte  $(x_i, y_i)$  mit denselben Hermite-Randbedingungen.

*Beweis.* i) Da  $l'$  auf  $[x_i, x_{i+1}]$  konstant ist, gilt:

$$\int_{x_i}^{x_{i+1}} (f'' - s'') l = (f' - s') l|_{x_i}^{x_{i+1}} - \int_{x_i}^{x_{i+1}} (f' - s') l' = (f' - s') l|_{x_i}^{x_{i+1}} - c_i (f - s)|_{x_i}^{x_{i+1}} = 0.$$

Aufsummieren über  $i$  liefert aufgrund der Randbedingungen:

$$\int_a^b (f'' - s'') l = (f' - s') l|_a^b = 0.$$

ii) Wegen  $f = s + (f - s)$ ,  $s'' \in S_1(\Delta)$  und i) gilt:

$$(f'', f'') = (s'', s'') + \underbrace{(f'' - s'', f'' - s'')}_{\geq 0} + 2 \underbrace{(s'', f'' - s'')}_{=0} \geq (s'', s'').$$

iii) Sei  $l \in S_1(\Delta)$  die stückweise lineare Interpolation von  $f''$  bzgl.  $\Delta$ . Nach Satz 5.2 gilt:

$$\|f'' - l\| \leq \frac{h^2}{2} \|f^{(4)}\|.$$

Da  $f'' - s''$  und  $s'' - l$  nach i) orthogonal sind, gilt:

$$\|f'' - l\|^2 = \|f'' - s'' + s'' - l\|^2 = \|f'' - s''\|^2 + \|s'' - l\|^2 \geq \|f'' - s''\|^2.$$

Hieraus folgt die Behauptung.

iv) Die stückweise lineare Interpolation von  $f - s$  bzgl.  $\Delta$  ist die Nullfunktion, also gilt nach Satz 5.2 und iii):

$$\|f - s\| \leq \frac{h^2}{2} \|f'' - s''\| \leq \frac{h^4}{4} \|f^{(4)}\|.$$

□

**Korollar.** Das Interpolationsproblem (5.4.1) mit Hermite-Randbedingungen hat eine eindeutige Lösung  $s \in S_3(\Delta)$ .



*Beweis.* Wir betrachten die lineare Abbildung

$$\begin{aligned}\phi: S_3(\Delta) &\rightarrow \mathbb{R}^{n+3} \\ s &\mapsto (s(x_0), s(x_1), \dots, s(x_n), s'(a), s'(b))\end{aligned}$$

Wir zeigen zunächst die *Injektivität* von  $\phi$ . Gilt  $\phi(s) = 0$ , dann löst  $s$  das Interpolationsproblem (5.4.1) mit  $y_i = 0$  für alle  $i$  und  $s'(a) = s'(b) = 0$ . Da auch die Nullfunktion dasselbe Interpolationsproblem löst, folgt nach Satz 5.17:

$$\int_a^b (s'')^2 \leq \int_a^b 0^2 = 0.$$

Da  $s''$  stetig ist, folgt  $s'' = 0$  auf  $[a, b]$ . Somit ist  $s$  affin und mit  $s(a) = s(b) = 0$  folgt  $s = 0$ . Wegen  $\dim(S_3(\Delta)) = n + 3 = \dim(\mathbb{R}^{n+3})$  ist  $\phi$  sogar *bijektiv*, also existiert für alle  $y_0, \dots, y_k$  eine eindeutige Lösung des Interpolationsproblems mit Hermite-Randbedingungen.  $\square$

**Berechnung in B-Spline-Basis** Sei  $s \in S_3(\Delta)$ . Dann ist

$$s(x) = \sum_{i=-1}^{n+1} c_i B_3\left(\frac{x-x_i}{h}\right) \quad \text{für alle } x \in [a, b],$$

wobei  $x_i = a + ih$ . Berechnung der Koeffizienten  $c_i$ :

$$\begin{aligned}B_3(0) &= \frac{2}{3}, & B_3(\pm 1) &= \frac{1}{6}, & B_3(k) &= 0 \text{ für } |k| \geq 2, \\ B_3'(0) &= 0, & B_3'(\pm 1) &= \mp \frac{1}{2}, & B_3'(k) &= 0 \text{ für } |k| \geq 2.\end{aligned}$$

$$\begin{aligned} \begin{matrix} s(x_i) = y_i \\ s'(x_0) = y_{0,1} \\ s'(x_n) = y_{n,1} \end{matrix} &\Leftrightarrow \begin{pmatrix} -\frac{1}{2} & 0 & \frac{1}{2} & & & 0 \\ \frac{1}{6} & \frac{2}{3} & \frac{1}{6} & & & \\ & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} & & \\ & & \ddots & \ddots & \ddots & \\ & & & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \\ 0 & & & -\frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} c_{-1} \\ c_0 \\ c_1 \\ \vdots \\ c_n \\ c_{n+1} \end{pmatrix} = \begin{pmatrix} h y_0 \\ y_0 \\ y_1 \\ \vdots \\ y_n \\ h y_n \end{pmatrix}. \end{aligned}$$

Dies ist ein lineares Gleichungssystem mit dünn besetzter Matrix.

# Stichwortverzeichnis

- n*-stufiges Zufallsexperiment, 43
- Übergangsmatrix, 83
- Čebyšev-Knoten, 134
- Čebyšev-Polynoms, 134
- 0-1-Experimente
  - abhängige, 37
  - unabhängige, 36, 46
- a posteriori degree of belief, 42
- a priori degree of belief, 42
- abhängige 0-1-Experimente, 37
- Acceptance-Rejection-Verfahren, 68
- Additivität
  - einer Wahrscheinlichkeitsverteilung, 9
- Akzeptanzwahrscheinlichkeit, 68
- Akzeptanzzeit, 69
- allgemeines Iterationsverfahren, 103
- asymptotische Äquivalenz von Folgen, 62
- Austausch von Gasmolekülen in zwei Behältern, 49
- Banachscher Fixpunktsatz, 98
- Bayessche Regel, 42
- Bayessche Statistik, 42
- bedingte Erwartung, 39
- bedingte Verteilung, 39
- bedingte Wahrscheinlichkeit, **39**, 39
- Beispiel von Runge, 134
- Benfordsches Gesetz, 17
- Bernoulli-Verteilung, 36
- Bernoulliverteilung
  - Bernoulliverteilung
    - n*-dimensionale, 46
- Bernstein-Ungleichung, 54
- Bewegungsverlauf eines Random Walks, 63
- Binomialverteilung, 22
  - Poissonapproximation, 23
  - Varianz, 76
- Bisektionsverfahren, 125
- Cauchy-Schwarz-Ungleichung in  $\mathcal{L}^2$ , 73
- Čebyšev-Ungleichung, 77
- CG-Verfahren, 118
- dünn besetzte Matrix, 102
- degree of belief
  - a posteriori, 42
  - a priori, 42
- Detailed Balance-Bedingung, 84
- diskrete Zufallsvariable
  - gemeinsame Verteilung, 59
  - Unabhängigkeit, 59
  - Verteilung, Massenfunktion, 19
- diskreter Laplace-Operator, 109
- diskretes Modell, 8
  - mehrstufiges, 43
- dividierte Differenzen, 136
- Ehrenfest-Modell, 49, 86
- Einschluss-/Ausschlussprinzip, 11

- Einzelschrittverfahren, 107  
 Elementarereignis, 5  
 Empirische Verteilung, 15  
 Energie einer stetig differenzierbaren Funktion, 128  
 Energienorm, 109  
 Energieskalarprodukt, 109, 129  
 Ereignis, 5  
     disjunkte Ereignisse, 9  
     Indikatorfunktion, 33  
     Kombinationen von Ereignissen, 7  
 Erfolgswahrscheinlichkeit, 21  
 Ergodensatz, 96  
 Erwartung  
     bedingte *siehe* bedingte Erwartung 39  
 Erwartungswert, 33  
     der Poissonverteilung, 33  
     Linearität, 35  
     Monotonie, 36  
 Fixpunktiteration, 98  
     lineare, 103  
 Fluss, 85  
 Gauß-Formel, 145  
 Gauß-Seidel-Verfahren, 107  
 gemeinsame Verteilung diskreter Zufallsvariablen, 59  
 geometrische Verteilung, 53  
 Gesamtschrittverfahren, 106  
 Gesetz der großen Zahlen, 54  
     schwaches, 77  
     starkes, 78  
 Gesetz der großen Zahlen für Markov-Ketten, 96  
 gewichtetes Mittel, 34  
 Gewichtung der möglichen Fälle, 11  
     Monotonie, 12  
 Gibbs-Sampler, 89  
 Gleichgewichtsverteilung, 84  
     Konvergenz, 92  
 Gleichverteilung, 14  
     reellwertiger Zufallsvariablen, 67  
 Gradientenverfahren, 115  
     konjugiertes, 118  
     nicht-stationäres, 116  
 Häufigkeitsverteilung der Anfangsziffern von Zahlen, 18  
 Hermite-Interpolationsproblem, 132  
 Horner-Schema, 136  
 hypergeometrische Verteilung, 25, 45  
 Hypothesen, 40  
 Importance Sampling, 81  
 Indikatorfunktion einer Ereignisse, 33  
 Interpolation  
     stückweise lineare, 128  
 Interpolationsproblem  
     von Lagrange, 132  
     von Hermite, 132  
 Inversionen von Zufallspermutationen, 37  
 irreduzible stochastische Matrix, 94  
 Iterationsverfahren  
     allgemeines, 103  
     Iterationsverfahren  
         Berechnung der  $n$ -ten Wurzeln, 122  
 Jacobi-Verfahren, 106  
 Knotenpolynom, 133  
 konsistente Schätzfolge, 78  
 Konsistenzordnung

- Simpson-Verfahren, 142
- Kontraktion, 98
- Konvergenz  
lokale, 100  
stochastische, 77
- Konvergenz des Newton-Verfahrens, 124
- Konvergenz ins Gleichgewicht, 92
- Konvergenzordnung, 123
- Konvergenzsatz für endliche Markov-Ketten, 95
- Korrelationskoeffizient, 74
- Kovarianz, 74
- Krylov-Raum, 119
- kumulative Verteilungsfunktion, 67
- $l^2$ -Kondition einer invertierbaren Matrix, 116
- $L^2$ -Raum von diskreten Zufallsvariablen, 73
- $L^2$ -Skalarprodukt, 129
- Lagrange-Interpolationsproblem, 132
- Laplace-Modell, 14
- Laplace-Operator  
diskreter, 109
- likelihood, 42
- linearer Kongruenzgenerator, 26
- lokal quadratische Konvergenz des Newton-Verfahrens,  $l^2$ -Kondition, 116  
124
- lokale Konvergenz, 100
- Mögliche Fälle  
Gewichtung, 11
- Möglicher Fall,  
see Menge aller möglichen Fälle 5
- Münzwürfe, 6  
abhängige, 48  
zwei faire, 52
- Münzwurf  
endlich viele faire Münzwürfe, 14
- Markov-Kette, 85
- Markov-Kette, 47  
bei einem Münzwurf, 85  
Bewegungsgesetz, 47  
Gesetz der großen Zahlen, 96  
Konstruktion mit vorgegebenen Gleichgewichtsverteilungen, 87  
Konvergenzsatz für endliche Markov-Ketten, 95  
Metropolis-Kette, 88  
Monte Carlo-Verfahren, 95  
Simulation mit vorgegebenem Gleichgewicht, 91  
zeitlich homogene, 83
- Markovkette  
Stationarität, 84
- Massenfunktion, 11  
einer diskreten Zufallsvariable, 19  
eines mehrstufigen diskreten Modells, 43
- Matrix  
dünn besetzte, 102  
Matrix  
mehrstufiges diskretes Modell, 43  
Markov-Kette, *siehe* Markov-Kette  
Produktmodell, 46  
Wahrscheinlichkeitsverteilung, 43
- Menge aller möglichen Fälle, 5
- Metropolis-Algorithmus, 90
- Metropolis-Kette, 88  
Konvergenz, 95
- Minimierungsproblem zu einem linearen Gleichungssystem, 114
- Minorisierungsbedingung, 92

- Mittel
- arithmetisches, 34
  - gewichtetes, 34
- Mittelpunktsformel, 140
- Beispiel, 138
- Monotonie
- von Gewichtungen der möglichen Fälle, 12
- Monte Carlo-Schätzer, 71, 78
- Approximationsfehler, 71
  - eines mehrdimensionalen Integrals, 79
  - erwartungstreuer, 71
  - für Wahrscheinlichkeiten, 79
  - mittlere quadratische Fehler, 71
- Monte Carlo-Verfahren, 71
- für Markov-Ketten, 95
- $n$ -dimensionale Bernoulliverteilung, 46
- Newton-Cotes-Formel, 140
- Newton-Verfahren, 122
- Konvergenzsatz, 124
- numerische Integration durch Interpolation, 131
- Paradoxon
- von Simpson, 41
- Paradoxon
- von Sankt Petersburg, 35
- Periode, 94
- Periode eines Zustands, 94
- Perkolation von Netzwerken, 81
- Permutationen
- zufällige, 14
- Poincaré-Ungleichung, 130
- Poissonapproximation der Binomialverteilung, 23
- Poissongleichung, 103
- Poissonverteilung, 24
- Erwartungswert, 33
- Polynominterpolation, 133
- Präkonditionierer, 103
- Produkt von Wahrscheinlichkeitsverteilungen, 46
- Produktmodell, 46
- Pseudo-Zufallszahlengenerator, 25
- Quadratur, 131
- Quadraturformel, 138
- Rückkehrzeit, 63
- Random Walk, 63
- auf den ganzen Zahlen, 60
  - auf einem Gitter, 48
  - auf Graphen, 86
  - Bewegungsverlauf, 63
  - symmetrischer, 63
  - Trefferzeit, 63
  - Verteilung der Positionen zur Zeit  $n$ , 62
  - zyklischer, 86
- Raum der linearen Splines, 128
- reellwertige Zufallsvariable, **66**
- gleichverteilt, 67
  - Unabhängigkeit, 67
- Reflektionsprinzip, 63
- Regula Falsi-Verfahren, 127
- Relaxationsverfahren, 112
- Residuendarstellung, 104
- Residuum, 104
- Rucksackproblem, 90
- Runges Beispiel, 134
- Sankt-Petersburg-Paradoxon, 35
- Schätzfolge
- konsistente, 78
- Schwaches Gesetz der großen Zahlen, 77

- Sekantenverfahren, 126
- Selbstbefruchtung von Pflanzen, 48
- $\sigma$ -Additivität von Wahrscheinlichkeitsverteilungen, 9
- $\sigma$ -Algebra, 8
- Simpson-Formel, 141
- Simpson-Paradoxon, 41
- Simpson-Verfahren  
Konsistenzordnung, 142
- Simulated Annealing, 90  
Algorithmus, 91
- Simulation einer diskreten Verteilung  
Acceptance-Rejection-Verfahren, 68  
direkt, 67
- Simulation einer Markov-Kette mit vorgegebenem Gleichgewicht, 91
- Skalarprodukt  
Energie-, 129
- Skalarprodukt auf  $\mathcal{L}^2$ , 73
- Spektralradius, 104
- Splines  
Raum der linearen Splines, 128  
Splines vom Grad  $m$ , 148  
stückweise lineare Interpolation, 128
- Standardabweichung, 72
- starkes Gesetz der großen Zahlen, 78
- Stationarität von Markov-Ketten, 84
- Steepest Descent-Algorithmus, 115
- Stirlingsche Formel, 62
- stochastische Konvergenz, 77
- stochastische Matrix, 47, 83  
irreduzibel, 94
- stochastischer Kern, 47
- strikt diagonaldominant, 107
- symmetrischer Random Walk, 63
- Transformationssatz, 34
- Trapezformel, 131, 140  
Beispiel, 138
- Trefferzeit, 63  
Verteilung, 64
- unabhängige 0-1-Experimente, 36, 46
- Unabhängigkeit, 39  
reellwertiger Zufallsvariablen, 67
- Unabhängigkeit von diskreten Zufallsvariablen, 59, 60
- Unabhängigkeit von Ereignissen, 22, 52
- Unkorreliertheit, 75
- Varianz, 72  
der Binomialverteilung, 76  
Reduktion durch Importance Sampling, 81  
von Summen, 76
- Variationsdistanz von Wahrscheinlichkeitsverteilungen, 92
- verallgemeinertes Interpolationsproblem, 132
- Verfahren der konjugierten Gradienten, 118
- Verteilung  
bedingte,  
seebedingte Verteilung<sup>39</sup>  
direkte Simulation einer diskreten Verteilung, 67
- Verteilung der Trefferzeiten, 64
- Verteilung des Maximums, 66
- Verteilungsfunktion  
kumulative, 67
- Wahrscheinlichkeitsverteilung  
Gleichverteilung oder Laplace-Modell, 14
- Wahrscheinlichkeit, 5

- 
- Akzeptanz-, 68
  - bedingte, *siehe* bedingte Wahrscheinlichkeit
  - Erfolgs-, 21
  - Wahrscheinlichkeitsraum, 9
  - Wahrscheinlichkeitsverteilung, 9, 11
    - einer diskreten Zufallsvariable, 19
    - Detailed Balance-Bedingung, 84
    - diskret
      - Massenfunktion, 11
    - diskrete, 11
    - eines mehrstufigen diskreten Modells, 43
    - Elementare Rechenregeln, 9
    - endliche Additivität, 9
    - geometrische, 53
    - Gleichgewichtsverteilung, 84
    - Monotonie, 10
    - Normiertheit, 9
    - Produkt, 46
    - Variationsdistanz, 92
  - Warteschlange, 23
  - Wartezeit auf das erste Eintreten eines von unendlich vielen Ereignissen, 53
  - Werfen von zwei Würfeln, 20
  - Ziehen mit Zurücklegen, *siehe* Binomialverteilung
  - Ziehen ohne Zurücklegen, *siehe* Hypergeometrische Verteilung
  - Zufallspermutation
    - Inversionen, 37
  - Zufallsvariable, 6, 19
    - diskrete, *siehe* 19reellwertige, 34, *siehe* reellwertige Zufallsvariable
    - Standardabweichung, 72
    - Varianz, 72
  - Zufallsvorgang, 5
    - diskreter, 7
  - Zufallszahlengenerator, 25, 67
  - Zuverlässigkeit von Netzwerken, 81
  - zyklischer Random Walk, 86