

WESTFÄLISCHE WILHELMS-UNIVERSITÄT MÜNSTER
Institut für Mathematische Statistik

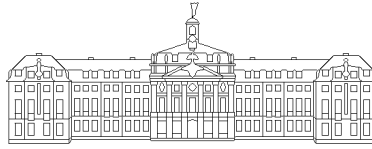
Diplomarbeit

Fixpunktsätze für Verteilungen

Themenvergabe
Prof. Dr. Alsmeyer

vorgelegt von
Adrian Hirsch

Dezember 1998



WESTFÄLISCHE WILHELMS-UNIVERSITÄT MÜNSTER
Institut für Mathematische Statistik

Diplomarbeit

Fixpunktsätze für Verteilungen

Themenvergabe
Prof. Dr. Alsmeyer

vorgelegt von
Adrian Hirsch

Dezember 1998

Inhaltsverzeichnis

1	Das allgemeine Fixpunktproblem	9
2	Fixpunktsätze für Verteilungen	15
2.1	Existenz der Abbildung S	15
2.2	Die Mallow-Metrik	18
2.2.1	Nachweis der Eigenschaften einer Metrik	19
2.2.2	Konvergenz bezüglich der Mallow-Metrik	24
2.2.3	Die Wasserstein-Metrik	28
2.3	Fixpunkte der Abbildung S	29
2.3.1	Der Spezialfall: (\mathcal{M}_2, d_2) als Definitions- und Wertebereich der Abbildung S	29
	Repräsentation des Fixpunktes durch eine Zufallsgröße	35
2.3.2	(\mathcal{M}_p, d_p) als Definitions- und Wertebereich: Die Existenz von höheren Momenten des Fixpunktes	41
2.3.3	Exponentielle Momente des Fixpunktes	47
3	Anwendungen	53
3.1	Der Galton-Watson-Verzweigungsprozeß	53
3.2	Quicksort	57
3.2.1	Einführung	57
3.2.2	Modellierung und Herleitung einer rekursiven Beziehung	58
	Berechnung des Erwartungswertes der Zufallsgröße X_n	62
3.2.3	Bestimmung der Fixpunktgleichung	64
3.2.4	Nachweis der Existenz eines Fixpunktes	67
3.2.5	Konvergenz von Y_n gegen den Fixpunkt	68
3.2.6	Abschätzungen für das Auftreten schlechter Laufzeiten	71
3.2.7	Momenterzeugende Funktionen	73
3.2.8	Eine verbesserte Abschätzung für schlechtes Laufzeitverhalten	78

A	Hilfsresultate	81
A.1	Abschätzungen und Konvergenzaussagen	81
A.2	Die verallgemeinerte Höldersche Ungleichung	83
A.3	Die Bernoullischen Zahlen	84
A.4	Die Bernoullischen Polynome und die Eulersche Summenformel	85

Einleitung

Hat man eine Abbildung $f : M \rightarrow N$ gegeben, so ist ein zugehöriger Fixpunkt ein Punkt ξ mit der Eigenschaft $f(\xi) = \xi$. Es handelt sich also um einen Punkt, der bei der Abbildung $x \mapsto f(x)$ fest bleibt. Von Bedeutung ist die Frage, unter welchen Bedingungen ein Fixpunkt einer Abbildung existiert und wann dieser eindeutig ist. Dabei spielt der Banachsche Fixpunktsatz eine bedeutende Rolle. Dieser stellt eine weitreichende Methode zum Beweis von Existenz- und Eindeutigkeitsätzen dar, sofern als Definitionsbereich und Wertebereich ein metrischer Raum gegeben ist.

Im Rahmen dieser Arbeit werden wir uns speziell mit Abbildungen S der Form

$$S(F) \stackrel{\mathcal{D}}{=} \sum_i T_i X_i + C$$

auseinandersetzen. Dabei ist F eine eindimensionale Verteilungsfunktion, die Zufallsvariablen $(C, T = (T_1, T_2, \dots))$, X_i , $i \in \mathbb{N}$, sind unabhängig, $X_i \sim F$ und \mathcal{D} bezeichne die Gleichheit in Verteilung. Definitionsbereich und Wertebereich ist der Raum der eindimensionalen Verteilungsfunktionen. Fixpunkte solcher Abbildungen treten im Zusammenhang mit verschiedenen Problemstellungen, wie etwa dem Galton-Watson-Verzweigungsprozeß oder dem Sortieralgorithmus „Quicksort“, auf.

Dementsprechend gliedert sich die Arbeit in drei Teile. Im ersten Kapitel werden allgemeine Aussagen über die Existenz von Fixpunkten von Abbildungen auf metrischen Räumen zusammengestellt und in diesem Zusammenhang das Prinzip der iterativen Approximation vorgestellt, dessen eindeutige Lösung im Falle einer kontrahierenden Abbildung durch den Banachschen Fixpunktsatz garantiert wird. Im zweiten Teil wird dann die allgemeine Theorie auf den oben eingeführten Abbildungstyp S übertragen. Dazu ist eine Metrik auf dem Raum der eindimensionalen Verteilungsfunktionen erforderlich. Als geeignet erweist sich die Mallows-Metrik, die von Mallows 1972 eingeführt wurde [↔ [18]] und die den Abstand von zwei Verteilungsfunktionen als minimalen \mathfrak{L}_p -Abstand ($1 \leq p < \infty$) von zwei Zufallsgrößen mit den entsprechenden Verteilungen definiert. Eine ausführliche Untersuchung dieser Metrik erfolgt in Abschnitt 2.2.

Desweiteren werden wir uns mit der Existenz von höheren und auch exponentiellen Momen-

ten der Fixpunktverteilung beschäftigen. Es sei angemerkt, daß Fixpunkte solcher Abbildungen aus diesem Blickwinkel zum ersten Mal von Rösler in [19] im Jahr 1992 untersucht wurden.

Im letzten Teil erfolgt schließlich die Anwendung der zuvor bewiesenen Fixpunktsätze auf konkrete Grenzverteilungsprobleme. Interessant in diesem Zusammenhang ist der Galton-Watson-Verzweigungsprozeß: Betrachtet man eine Population, die zu Beginn (0-te Generation) aus einem Mitglied, genannt Urahn, besteht und gibt M_n die Anzahl der Individuen der n -ten Generation an (d.h. $M_0 = 1$), so bezeichnet man den Prozeß $M = (M_n)_{n \geq 0}$ als Galton-Watson-Verzweigungsprozeß, wenn folgende Annahmen erfüllt sind: Jedes Individuum habe eine Lebenszeit von einer Zeiteinheit und produziere am Lebensende eine zufällige Anzahl von Nachkommen. Dabei sei die Anzahl der Nachkommen für jedes Individuum identisch verteilt gemäß der Reproduktionsverteilung $(p_k)_{k \geq 0}$. Zudem mögen die Individuen unabhängig voneinander und von der Anzahl der Mitglieder der eigenen Generation Nachkommen produzieren. Der Erwartungswert μ der Reproduktionsverteilung wird Reproduktionsmittel genannt.

Es ist bekannt, daß das Martingal M_n/μ^n fast sicher gegen eine Zufallsvariable W konvergiert [↔ [1]]. Seien $C \equiv 0$, $T = (T_1, T_2, \dots) \in \{0, 1\}^{\mathbb{N}}$ Zufallsvariablen mit $P(\sum_i T_i < \infty) = 1$, und definiere $p_k := P(\sum_i T_i = k)$. Betrachte die Abbildung

$$S(F) \stackrel{\mathcal{D}}{=} \sum_i \frac{T_i}{m} X_i.$$

Ausgehend von der Punktverteilung F_1 in 1, läßt sich leicht erkennen, daß die Verteilung der Zufallsgröße M_n/μ^n gleich der n -ten Iteration $S^n(F_1)$ ist, und die Verteilung der Grenzzufallsgröße W entspricht dem Fixpunkt der Abbildung S . Einzelheiten hierzu werden in Kapitel 3, Abschnitt 3.1 behandelt.

Ein weiteres Beispiel liefert der Sortieralgorithmus „Quicksort“. Wählt man T_1 speziell als $\mathcal{R}(0, 1)$ -verteilte Zufallsgröße und $T_2 = 1 \Leftrightarrow T_1$, $T_3 = 0 = T_4 = \dots$, $C(x) = 2x \ln x + 2(1 \Leftrightarrow x) \ln(1 \Leftrightarrow x) + 1$, so stellt sich heraus, daß die Abbildung

$$S(F) \stackrel{\mathcal{D}}{=} T_1 X_1 + (1 \Leftrightarrow T_1) X_2 + C(T_1)$$

einen bis auf Translation eindeutigen Fixpunkt besitzt. Die Anzahl der Vergleiche, die der Algorithmus Quicksort zum Sortieren einer Liste benötigt, konvergiert bei entsprechender Normierung für gegen unendlich laufende Listenlänge schwach gegen diesen Fixpunkt. Genauer liegt sogar Konvergenz in der Mallow-Metrik vor (für beliebiges $1 \leq p < \infty$), was sich als stärker erweist als die schwache Konvergenz. Man vergleiche hierzu etwa Satz 2.2.8 in Kapitel 2.2.

Das Beispiel des Quicksortalgorithmus werden wir detailliert im Abschnitt 3.2 untersuchen.

Dabei lassen wir es nicht bei dem Nachweis der Existenz des Fixpunktes und der Konvergenzaussage bewenden, sondern werden zeigen, wie man unter Ausnutzung dieses Wissens eine Abschätzung für ungünstiges Laufzeitverhalten des Algorithmus angeben kann, wenn man unterstellt, daß die Laufzeit proportional zur benötigten Anzahl von Vergleichen ist. Anhaltspunkte liefert eine Veröffentlichung zu diesem Thema von U. Rösler aus dem Jahre 1991 [↔ [20]].

Mathematische Notationen

- ⇒ (Ω, \mathcal{A}, P) allen Betrachtungen zugrundegelegter Wahrscheinlichkeitsraum
- ⇒ $\mathcal{L}_p(\Omega, \mathcal{A}, P)$ Vektorraum aller p -fach P -integrierbaren meßbaren Funktionen auf (Ω, \mathcal{A})
- ⇒ $\|\cdot\|_p$ \mathcal{L}_p -Norm auf $\mathcal{L}_p(\Omega, \mathcal{A}, P)$
- ⇒ Ist X eine reellwertige Zufallsgröße, so bezeichne $L(X)$ den Übergang zur Verteilungsfunktion.
- ⇒ $E(X)$, $\text{Var}(X)$ Erwartungswert bzw. Varianz der Zufallsgröße X . Dabei wird im vorliegenden Text auch häufiger die Schreibweise $E(L(X))$, $\text{Var}(L(X))$ benutzt.
- ⇒ $\mathcal{M} := \{\text{Menge aller eindimensionalen Verteilungsfunktionen}\}$
- ⇒ $\mathcal{M}_p := \{F \in \mathcal{M} : \int |x|^p dF(x) < \infty\}$, $1 \leq p < \infty$
- ⇒ $\mathcal{M}_p^b := \{F \in \mathcal{M}_p : E(F) = b\}$
- ⇒ $\sum_i T_i \in \mathcal{L}_p$ bedeutet: $\sum_{i=1}^n T_i \in \mathcal{L}_p$ für alle $n \in \mathbb{N}$ und $\left(\sum_{i=1}^n T_i\right)_n$ konvergiert in \mathcal{L}_p -Norm, wobei die Grenzzufallsgröße mit $\sum_i T_i$ bezeichnet sei.
- ⇒ $\|\sum_i T_i\|_p < \infty$ analog.
- ⇒ Für $F \in \mathcal{M}_1$ wird mit F^b die so verschobene Verteilungsfunktion bezeichnet, daß $E(F^b) = b$.
- ⇒ $\mathcal{H}(F, G)$ Menge aller zweidimensionalen Verteilungsfunktionen mit Randverteilungsfunktionen F bzw. G .
- ⇒ Hat man eine Folge von nichtnegativen Zufallsgrößen Y_n , und betrachtet man $\sum_i Y_i$, so ist dieser Grenzwert - sofern nichts anderes gesagt ist - als fast sicherer Limes aufzufassen.

Kapitel 1

Das allgemeine Fixpunktproblem

Bevor wir mit der speziellen Theorie beginnen, sei unser Blick allgemein auf das Fixpunktproblem gerichtet. Hat man eine Funktion $f : M \rightarrow N$ zwischen zwei Teilmengen eines metrischen Raumes (X, d) gegeben, so ist ein Fixpunkt von f eine Lösung der Gleichung $x = f(x)$. Um solch eine Lösung zu erhalten, kann man sich nun einen Startpunkt x_0 aus dem Definitionsbereich M von f wählen, und die Folge (x_n) rekursiv (iterativ) durch $x_{n+1} := f(x_n)$, $n = 0, 1, 2, \dots$ definieren. Dies ist immer dann möglich, wenn jedes x_n in M liegt (z.B. wenn $N \subset M$). Explizit ist das n -te Folgenglied x_n nichts anderes als die n -te Iteration f^n (n -fache Hintereinanderschaltung der Abbildung f) angewendet auf den Startwert x_0 . Man muß nun untersuchen, ob (x_n) gegen einen Grenzwert $\xi \in M$ konvergiert. Ist dies der Fall, konvergiert aufgrund der Gleichheit $f(x_n) = x_{n+1}$ natürlich auch die Folge der $f(x_n)$, und wir erhalten unter der Voraussetzung, daß f stetig ist, aus $x_{n+1} = f(x_n)$ durch Grenzübergang die Gleichung $\xi = f(\xi)$, die ξ als Fixpunkt ausweist. Es handelt sich hierbei um die Methode der sukzessiven Approximation [↔ [24], S.312].

Allgemein ist die Existenz des Grenzwertes der iterativen Folge x_n keinesfalls gesichert. Betrachtet man speziell den metrischen Raum $(\mathbb{R}, |\cdot|)$, wobei $|\cdot|$ die gewöhnliche Betragsmetrik kennzeichnet, so läßt sich die Iterationsvorschrift $x_{n+1} = f(x_n)$ sehr gut anhand der Kurven $y = f(x)$ und $y = x$ veranschaulichen. In Abbildung 1.1 sind vier typische Fälle skizziert, in denen monotone und alternierende Konvergenz, monotone und alternierende Divergenz vorliegt.

Es stellt sich also die Frage, unter welchen Voraussetzungen die Iterationsfolge einen Grenzwert besitzt. In diesem Zusammenhang sind kontrahierende Abbildungen von besonderer Bedeutung [↔ [3], S146]:

1.1 Definition Eine Abbildung

$$f : M \rightarrow N$$

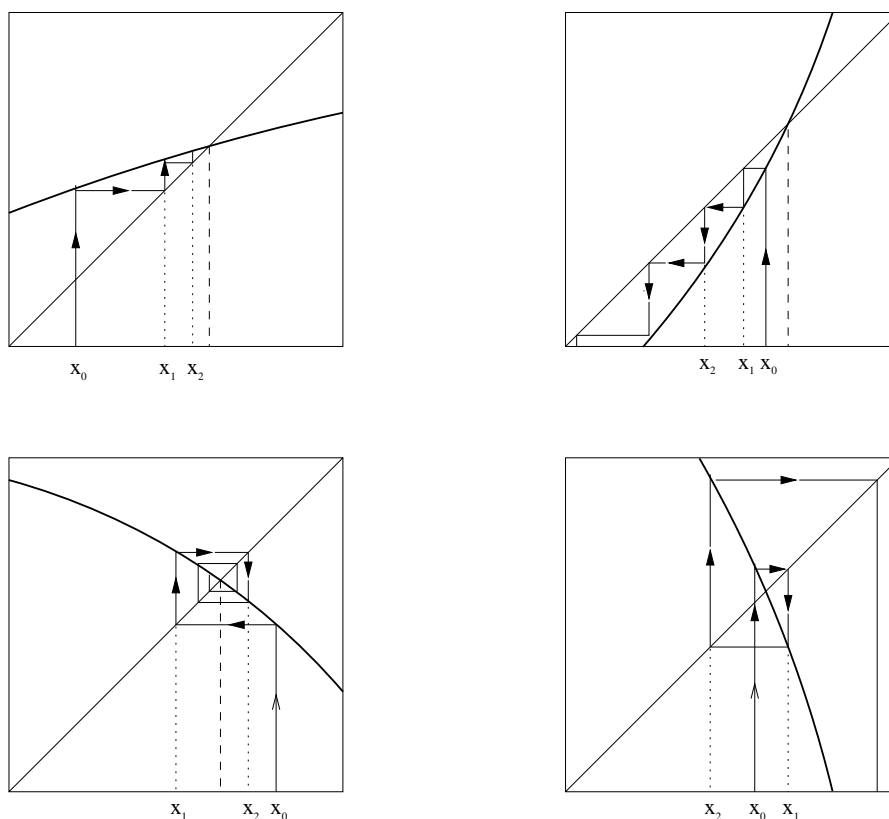


Abbildung 1.1: Veranschaulichung der Iterationsvorschrift $x_{n+1} = f(x_n)$ anhand von vier typischen Fällen: monotone Konvergenz, monotone Divergenz, alternierende Konvergenz, alternierende Divergenz (von links oben nach rechts unten).

heißt *kontrahierend*, wenn es eine reelle Zahl α mit $0 \leq \alpha < 1$ gibt, derart daß für alle $x, y \in M$ gilt:

$$d(f(x), f(y)) \leq \alpha \cdot d(x, y).$$

Auch Formulierungen wie α -Kontraktion, α -kontrahierend sind gebräuchlich. Eine solche Abbildung „kontrahiert“ den Abstand $d(x, y)$ der Urbilder. Der Abstand der Bilder ist mindestens um den Faktor α kleiner. Es ist wesentlich, daß die Konstante $\alpha < 1$ ist. Der folgende Satz [⇔ [3], S.146], der auch als Banachscher Fixpunktsatz bekannt ist, besagt, daß in einem vollständigen metrischen Raum, die kontrahierende Eigenschaft der Abbildung f hinreichend für die Existenz eines eindeutigen Fixpunktes ist:

1.2 Satz (Banachscher Fixpunktsatz) *Jede kontrahierende Abbildung $f : M \rightarrow M$ eines vollständigen metrischen Raumes (M, d) besitzt genau einen Fixpunkt, d.h. es gibt genau ein $\xi \in M$ derart, daß gilt*

$$f(\xi) = \xi.$$

BEWEIS: Wir weisen zunächst die Existenz eines Fixpunktes ξ nach. Sei dazu $x_0 \in M$ und (wie bereits weiter oben angedeutet)

$$x_{k+1} := f(x_k).$$

Um zu beweisen, daß (x_k) eine Cauchy-Folge ist, müssen wir $d(x_m, x_n)$ für $m \geq n$ abschätzen. Unter Beachtung der vorausgesetzten Kontraktionseigenschaft der Abbildung f erhält man für $j \geq 1$:

$$d(x_{j+1}, x_j) \leq \alpha \cdot d(x_j, x_{j-1})$$

und durch wiederholte Anwendung

$$(1.1) \quad d(x_{j+1}, x_j) \leq \alpha^j \cdot d(x_1, x_0).$$

Die Dreiecksungleichung liefert zudem:

$$d(x_m, x_n) \leq d(x_m, x_{m-1}) + \cdots + d(x_{n+1}, x_n)$$

und somit unter Beachtung von (1.1):

$$d(x_m, x_n) \leq (\alpha^{m-1} + \alpha^{m-2} + \cdots + \alpha^n) \cdot d(x_1, x_0) = \frac{1 \Leftrightarrow \alpha^{m-n}}{1 \Leftrightarrow \alpha} \cdot \alpha^n \cdot d(x_1, x_0).$$

Da $0 \leq \alpha < 1$ gilt, ergibt sich für beliebiges $m \geq n$ die Abschätzung:

$$d(x_m, x_n) \leq \frac{\alpha^n}{1 \Leftrightarrow \alpha} \cdot d(x_1, x_0).$$

Wegen $\lim_{k \rightarrow \infty} \alpha^k = 0$ ist (x_k) also eine Cauchy-Folge und aufgrund der vorausgesetzten Vollständigkeit des metrischen Raumes (M, d) gibt es ein $\xi \in M$, so daß $\lim_k x_k = \xi$. Dieses ξ ist ein Fixpunkt von f , denn aufgrund der Kontraktionseigenschaft ist die Abbildung f offenbar insbesondere stetig, woraus zusammen mit

$$x_{k+1} = f(x_k)$$

unmittelbar:

$$\xi = \lim_k x_{k+1} = \lim_k f(x_k) = f(\lim_k x_k) = f(\xi)$$

folgt.

Schließlich ergibt sich die Eindeutigkeit des Fixpunktes folgendermaßen: Aus $f(\xi) = \xi$ und $f(\psi) = \psi$ erhält man die Abschätzung:

$$d(\xi, \psi) = d(f(\xi), f(\psi)) \leq \alpha \cdot d(\xi, \psi).$$

Wegen $0 \leq \alpha < 1$, also der Striktheit der Kontraktion, folgt hieraus $d(\xi, \psi) = 0$, d.h. $\xi = \psi$. \square

Die Aussage des Banachschen Fixpunktsatzes kann im einfachsten Fall, nämlich der kontrahierenden Abbildung eines abgeschlossenen Intervalls in sich, gut veranschaulicht werden:

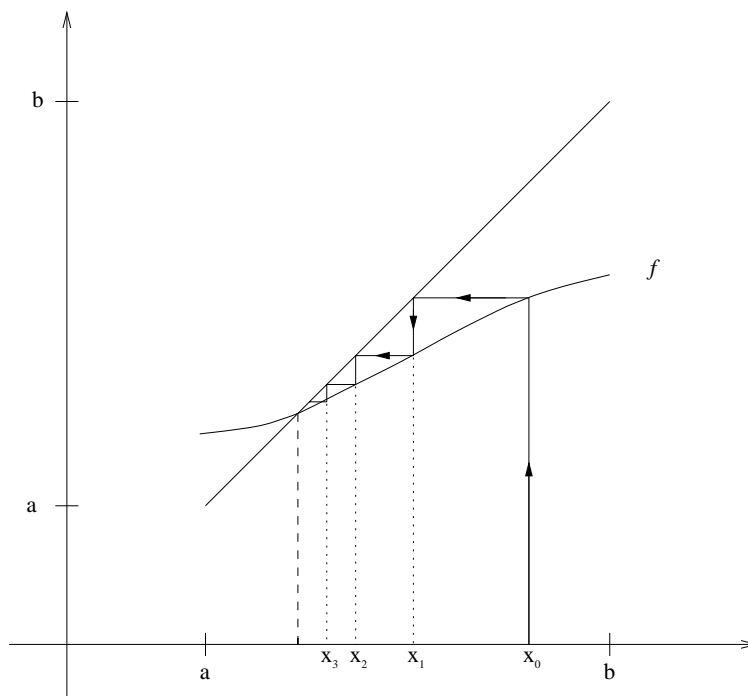


Abbildung 1.2: Veranschaulichung der Kontraktionseigenschaft im Fall einer Abbildung $f : [a, b] \rightarrow [a, b]$.

Daß f kontrahierend ist, zeigt sich darin, daß der Graph von f flacher als die Hauptdiagonale verläuft.

Der Beweis des Banachschen Fixpunktsatzes zeigt weiterhin, daß der Fehler, der nach dem n -ten Iterationsschritt noch vorhanden ist, explizit abgeschätzt werden kann. Aus der bewiesenen Ungleichung

$$d(x_m, x_n) \leq \frac{\alpha^n}{1 \Leftrightarrow \alpha} \cdot d(x_1, x_0)$$

folgt nämlich als Abschätzung für den Fehler nach n -Iterationen

$$d(\xi, x_n) \leq \frac{\alpha^n}{1 \Leftrightarrow \alpha} \cdot d(x_1, x_0).$$

Benutzt haben wir dabei nur die Stetigkeit der Abbildung $M \rightarrow \mathbb{R}$, $y \mapsto d(y, x_n)$. Diese ist eine einfache Folgerung aus der Dreiecksungleichung und sei im folgenden Satz festgehalten:

1.3 Satz *Gegeben einen metrischen Raum (M, d) , ist die Abbildung $g : M \rightarrow \mathbb{R}$, $y \mapsto d(y, x)$ für jedes $x \in M$ stetig.*

BEWEIS: Seien $x, y' \in M$ vorgegeben und (y_k) eine Folge in M mit Grenzwert y' . Aus der Dreiecksungleichung folgen die Abschätzungen

$$d(y', x) \leq d(y', y_k) + d(y_k, x) \quad \text{und} \quad d(y_k, x) \leq d(y_k, y') + d(y', x)$$

und somit für beliebiges $k \in \mathbb{N}$

$$|d(y', x) - d(y_k, x)| \leq d(y', y_k).$$

Es ergibt sich

$$\lim_{k \rightarrow \infty} |d(y', x) - d(y_k, x)| = 0.$$

□

Als Folgerung aus dem Banachschen Fixpunktsatz, dessen Beweis und den vorherigen Ausführungen können wir festhalten:

1.4 Korollar *Gegeben die Voraussetzungen des Banachschen Fixpunktsatzes gilt: Für jedes $x_0 \in \mathcal{M}$ konvergiert die Folge $x_{k+1} := f(x_k)$ exponentiell schnell gegen den eindeutigen Fixpunkt der Abbildung f .*

Der Banachsche Fixpunktsatz wird sich auch für unsere Zwecke als wichtig herausstellen, da er Aussagen für allgemeine vollständige metrische Räume bereitstellt; vgl. Satz 2.3.5. Zudem liefert der Beweis wichtige Grundideen für den Nachweis der Existenz von Fixpunkten. Unmittelbar ersichtlich ist, daß die Kontraktionseigenschaft keine notwendige Bedingung darstellt, sondern daß es ausreicht, für vorgegebenen Startpunkt x_0 die Existenz von Konstanten $\lambda < 1$ und $c \in \mathbb{R}^+$ nachzuweisen, mit der Eigenschaft

$$d(f^{n+1}(x_0), f^n(x_0)) \leq c \cdot \lambda^n.$$

Dies wird etwa in Satz 2.3.6 im anschließenden Kapitel Verwendung finden. Dann ist allerdings die Forderung der Stetigkeit der Abbildung f eine weitere Voraussetzung, und die Eindeutigkeit des Fixpunktes muß nicht mehr gegeben sein; vgl. Korollar 2.3.7.

Kapitel 2

Fixpunktsätze für Verteilungen

In diesem zweiten Kapitel werden die allgemeinen Überlegungen des vorherigen Kapitels auf Abbildungen des speziellen, bereits in der Einleitung erwähnten Abbildungstyps übertragen. Unter Verwendung der eingeführten mathematischen Notationen ist dieser gekennzeichnet durch eine Abbildungsvorschrift der Form

$$S : \mathcal{M}_p \rightarrow \mathcal{M}_p, \quad F \mapsto L\left(\sum_i T_i X_i + C\right).$$

Dabei sind $(C, T = (T_1, T_2, \dots))$, $X_i, i \in \mathbb{N}$, \mathbb{R}^d - bzw. \mathbb{R} -wertige unabhängige Zufallsvariablen mit $L(X_i) = F$. p ist eine reelle Zahl größer oder gleich Eins. Abhängigkeiten zwischen den einzelnen Koordinatenvariablen T_i bzw. zwischen C und T sind erlaubt.

Es sei daran erinnert, daß \mathcal{M}_p die Menge aller Verteilungsfunktionen mit existierendem p -ten absoluten Moment bezeichnet.

Als Einstieg möge das folgende kleine Beispiel einer Abbildung des obigen Typs dienen, bei der man einen Fixpunkt direkt ablesen kann. Betrachtet man die Zuordnung

$$F \mapsto L(2^{-\frac{1}{2}}X_1 + 2^{-\frac{1}{2}}X_2),$$

so erweist sich die Standardnormalverteilung als Fixpunkt. Für X_1, X_2 Zufallsgrößen mit $X_i \sim N(0, 1)$ gilt nämlich $2^{-1/2}X_i \sim N(0, 1/2)$ und somit $2^{-1/2}X_1 + 2^{-1/2}X_2 \sim N(0, 1/2) * N(0, 1/2) = N(0, 1)$.

2.1 Existenz der Abbildung S

Die in der Definition der Abbildung S auftretende unendliche Summe wirft die Frage nach der Existenz der Abbildung und der Bedeutung von $\sum_i T_i X_i$ auf. Im folgenden wollen wir unter $\sum_i T_i X_i$ immer den \mathfrak{L}_p -Limes der zugehörigen Folge der Partialsummen verstehen

(bei variierendem p). Die Benutzung gerade dieser Konvergenzart liegt an dem engen Zusammenhang zwischen der Mallow-Metrik und der \mathfrak{L}_p -Norm; vgl. Abschnitt 2.2.

Um die Existenz des \mathfrak{L}_p -Limes zu gewährleisten, ist es natürlich erforderlich, entsprechende konvergenzerzwingende Forderungen an die auftretenden Zufallsgrößen zu stellen. Einige hinreichende Voraussetzungen beinhaltet der folgende

2.1.1 Satz *Gegeben unabhängige Zufallsvariablen $T = (T_1, T_2, \dots)$, $X_i, i \in \mathbb{N}$ mit Werten in $\mathbb{R}^{\mathbb{N}}$ bzw. \mathbb{R} und $X_i, i \in \mathbb{N}$, identisch verteilt, gilt:*

- 1) Aus $\sum_i |T_i|, X_1 \in \mathfrak{L}_p$ für ein $1 \leq p < \infty$ folgt $\sum_i T_i X_i \in \mathfrak{L}_p$ ($1 \leq p < \infty$).
- 2) Aus $\sum_i T_i, X_i \in \mathfrak{L}_2$ und $E(\sum_i T_i^2) < \infty$ folgt $\sum_i T_i X_i \in \mathfrak{L}_2$.
- 3) Aus $E(\sum_i T_i^2) < \infty$ und $L(X_1) \in \mathcal{M}_2^0$ folgt $\sum_i T_i X_i \in \mathfrak{L}_2$.

BEWEIS: Für den Nachweis der einzelnen Aussagen reicht es aufgrund der Vollständigkeit des metrischen Raumes $(\mathfrak{L}_p, \|\cdot\|_p)$ [\Leftrightarrow [2]] zu zeigen, daß $(\sum_{i=1}^n T_i X_i)_n$ eine Cauchy-Folge bildet. Zu 1): Wir betrachten zunächst den Fall $p = 1$. Es gelte $L(X_i) = F \in \mathcal{M}_1$. Eine einfache Rechnung ergibt:

$$\left\| \sum_{i=1}^m T_i X_i \right\|_1 \leq E|X_1| \cdot E\left(\sum_{i=1}^m |T_i|\right) < \infty$$

und damit $\sum_{i=1}^m T_i X_i \in \mathfrak{L}_1$ für alle $m \in \mathbb{N}$. Analog erhält man:

$$\left\| \sum_{i=n}^m T_i X_i \right\|_1 \leq E|X_1| \cdot E\left(\sum_{i=n}^m |T_i|\right),$$

wobei aufgrund der Voraussetzung $\sum_i |T_i| \in \mathfrak{L}_1$ die rechte Seite gegen Null konvergiert ($n, m \rightarrow \infty$). Die Folge $(\sum_{i=1}^m T_i X_i)_m$ ist somit eine Cauchy-Folge in $(\mathfrak{L}_1, \|\cdot\|_1)$.

Der Fall $p \geq 2$ ist etwas komplizierter. Es gelte $L(X_i) = F \in \mathcal{M}_p$. Zur Abkürzung sei $t_N := \sum_{i=1}^N |T_i|$ gesetzt. Dann folgt

$$\begin{aligned} E\left(\left|\sum_{i=1}^N T_i X_i\right|^p\right) &\leq E\left(\left(\sum_{i=1}^N |T_i| |X_i|\right)^p\right) = E\left(E\left(\left(\sum_{i=1}^N \frac{|T_i|}{t_N} t_N |X_i|\right)^p \mid T\right)\right) \\ &= E_\omega\left(E_{\omega_1}\left(\sum_{i=1}^N \frac{|T_i(\omega)|}{t_N(\omega)} t_N(\omega) |X_i(\omega_1)|\right)^p\right) \leq E\left(E\left(\sum_{i=1}^N |T_i| t_N^{p-1} |X_i|^p \mid T\right)\right) \\ &= E(|X_1|^p) E\left(\sum_{i=1}^N |T_i|^p\right) \leq E(|X_1|^p) E\left(\sum_i |T_i|^p\right) < \infty. \end{aligned}$$

Angewendet wurde dabei die Jensensche Ungleichung auf den Ausdruck

$$\left(\sum_{i=1}^N \frac{|T_i(\omega)|}{t_N(\omega)} t_N(\omega) |X_i(\omega_1)| \right)^p,$$

dessen innerer Term nichts anderes darstellt als den Erwartungswert einer diskreten Verteilung. In Analogie dazu erhält man außerdem:

$$E\left(\left| \sum_{i=1}^N T_i X_i \right| \right)^p \leq E\left(\left| \sum_{i=M+1}^N |T_i| |X_i| \right|^p \right) = E(|X_1|^p) E\left(\left| \sum_{i=M+1}^N |T_i| \right|^p \right),$$

wobei aufgrund der Voraussetzung folgt, daß der letzte Term gegen Null konvergiert (für $N, M \rightarrow \infty$) und somit $\sum_{i=1}^n T_i X_i$, $n \in \mathbb{N}$, eine Cauchy Folge in $(\mathfrak{L}_p, \|\cdot\|_p)$ darstellt.

Zu 2): Es gelte $L(X_i) = F \in \mathcal{M}_2$. Zunächst wird wieder $\sum_{i=1}^N T_i X_i \in \mathfrak{L}_2$ für alle $N \in \mathbb{N}$ gezeigt. Dazu können wir zunächst festhalten, daß

$$\begin{aligned} E\left(\left(\sum_{i=1}^N T_i X_i \right)^2 \right) &= E\left(\sum_{i=1}^N \sum_{k=1}^N T_i X_i T_k X_k \right) = E\left(\sum_{i=1}^N \sum_{k=1}^N T_i T_k X_i X_k \right) \\ &= \sum_{i=1}^N \sum_{k=1}^N E(T_i T_k) \cdot E(X_i X_k) < \infty. \end{aligned}$$

Ausgenutzt wurde dabei, daß aus der Voraussetzung $E(\sum T_i^2) < \infty$ die Quadratintegrierbarkeit der T_i und damit die Integrierbarkeit von $T_i T_k$ folgt. Für $M \leq N$ gilt außerdem:

$$\begin{aligned} \left\| \sum_{i=1}^N T_i X_i \right\|_2^2 &= \left\| \sum_{i=M+1}^N T_i X_i \right\|_2^2 = E\left(\left(\sum_{i=M+1}^N T_i X_i \right)^2 \right) \\ &= E\left(\sum_{\substack{i,j=M+1 \\ i \neq j}}^N T_i T_j \right) \cdot (E(X_1))^2 + E\left(\sum_{i=M+1}^N T_i^2 \right) \cdot E(X_1^2) \\ &\leq E\left(\sum_{i=M+1}^N T_i^2 \right) \cdot (E(X_1))^2 + E\left(\sum_{i=M+1}^N T_i^2 \right) \cdot E(X_1^2). \end{aligned}$$

Da nun $\sum_i T_i \in \mathfrak{L}_2$ und $E(\sum_i T_i^2) < \infty$, ergibt sich die Behauptung.

Zu 3): Es gelte $L(X_i) = F \in \mathcal{M}_2^0$. Aus der Abschätzung

$$\begin{aligned} E\left(\left(\sum_{i=1}^N T_i X_i \right)^2 \right) &= E\left(\sum_{i=1}^N \sum_{k=1}^N T_i T_k X_i X_k \right) = \sum_{i=1}^N \sum_{k=1}^N E(T_i T_k) E(X_i X_k) \\ &\stackrel{E X_i = 0}{=} E(X_1^2) \cdot E\left(\sum_{i=1}^N T_i^2 \right) < \infty \end{aligned}$$

und der Gleichungskette

$$\left\| \sum_{i=1}^N T_i X_i \Leftrightarrow \sum_{i=1}^M T_i X_i \right\|_2^2 = \left\| \sum_{i=M+1}^N T_i X_i \right\|_2^2 = E(X_1^2) \cdot E\left(\sum_{i=M+1}^N T_i^2 \right)$$

resultiert wiederum die Eigenschaft der Cauchy-Folge. Der letzte Term konvergiert dabei gegen Null ($N, M \rightarrow \infty$), da aufgrund monotoner Konvergenz und der Voraussetzung $E(\sum_i T_i^2) < \infty$ die Folge $(E(\sum_{i=1}^n T_i^2))_n$ konvergiert und folglich insbesondere eine Cauchy Folge darstellt. \square

Aus der p -fachen Integrierbarkeit einer Zufallsgröße folgt insbesondere deren fast sichere Reellwertigkeit [\Leftrightarrow [2], Korollar 11.8. auf Seite 47 und Satz 10.6 auf S.42]. Zusammen mit der vorausgesetzten Reellwertigkeit der Zufallsgröße C ergibt sich die Existenz der Abbildung S , wenn gemäß der Forderung $X_i \in \mathcal{L}_p$ als Definitionsbereich \mathcal{M}_p gewählt wird. Da $\sum_i T_i X_i \in \mathcal{L}_p$, sagen wir, die Abbildung S existiert im \mathcal{L}_p -Sinn. Wir können festhalten:

- 2.1.2 Korollar**
1. $\sum_i |T_i|, C \in \mathcal{L}_p \Rightarrow S : \mathcal{M}_p \rightarrow \mathcal{M}_p$ existiert im \mathcal{L}_p -Sinn ($1 \leq p < \infty$).
 2. $\sum_i T_i, C \in \mathcal{L}_2, E(\sum_i T_i^2) < \infty \Rightarrow S : \mathcal{M}_2 \rightarrow \mathcal{M}_2$ existiert im \mathcal{L}_2 -Sinn.
 3. $E(\sum_i T_i^2) < \infty, C \in \mathcal{L}_2 \Rightarrow S : \mathcal{M}_2^0 \rightarrow \mathcal{M}_2$ existiert im \mathcal{L}_2 -Sinn.

Um die Voraussetzungen des Banachschen Fixpunktsatzes zu erfüllen, ist eine Metrik d_p auf \mathcal{M}_p erforderlich, die zum einen unter entsprechenden Voraussetzungen die Kontraktionseigenschaft der Abbildung S gewährleistet und zum anderen die Vollständigkeitsanforderung erfüllt. Als geeignet erweist sich die Mallow-Metrik, die im folgenden Abschnitt einer genauen Betrachtung unterzogen wird.

2.2 Die Mallow-Metrik

Für $F, G \in \mathcal{M}_p$ ordnet die Abbildung $d_p : \mathcal{M}_p \times \mathcal{M}_p \rightarrow \mathbb{R}, (F, G) \mapsto \inf_{\substack{X \sim F \\ Y \sim G}} \|X \Leftrightarrow Y\|_p$, wobei das Infimum über alle Zufallsgrößen X und Y auf einem beliebigen Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) genommen wird mit $L(X) = F$ und $L(Y) = G$, diesen Verteilungsfunktionen einen reellen Wert zu ($1 \leq p < \infty$). In diesem Abschnitt werden wir zeigen, daß durch diese Abbildungsvorschrift bereits eine Metrik auf dem Raum \mathcal{M}_p aller eindimensionalen Verteilungsfunktionen mit existierendem p -ten absoluten Moment bestimmt ist, und daß das Infimum angenommen wird. Außerdem werden Verbindungen zwischen der durch die Mallow-Metrik determinierten Konvergenz und anderen Konvergenzarten der Wahrscheinlichkeitstheorie aufgezeigt und ein kleiner Seitenblick auf die sogenannte Wasserstein-Metrik geworfen, um Unterschiede und Gemeinsamkeiten aufzuzeigen.

2.2.1 Nachweis der Eigenschaften einer Metrik

Wir notieren den wichtigen

2.2.1 Satz Die Abbildung $d_p : \mathcal{M}_p \times \mathcal{M}_p$, $d_p(F, G) := \inf_{\substack{X \sim F \\ Y \sim G}} \|X \Leftrightarrow Y\|_p$ mit $1 \leq p < \infty$, besitzt die für alle $F, G, H \in \mathcal{M}_p$ gültigen Eigenschaften:

- i) $d_p(F, G) = d_p(G, F)$ (Symmetrie).
- ii) $d_p(F, H) \leq d_p(F, G) + d_p(G, H)$ (Dreiecksungleichung).
- iii) $d_p(F, G) = 0 \Leftrightarrow F = G$.

Zudem gilt: $d_p(F, G) = \|F^{-1}(U) \Leftrightarrow G^{-1}(U)\|_p$ für $U \sim \mathcal{R}(0, 1)$ und F^{-1}, G^{-1} Pseudo-Inverse von F bzw. G . Das Infimum wird folglich angenommen.

Wir wollen uns zunächst mit der Frage beschäftigen, wie man zeigen kann, daß für die Zufallsgrößen $F^{-1}(U), G^{-1}(U)$ das Infimum angenommen wird.

Dafür müssen wir herausfinden, für welche Zufallsgrößen X und Y auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathcal{P})$ der Ausdruck $E(|X \Leftrightarrow Y|^p)$ „klein“ wird unter der zusätzlichen Forderung $L(X) = F$ und $L(Y) = G$. Da $E|X \Leftrightarrow Y|^p = \int_{\mathbb{R}^2} |x \Leftrightarrow y|^p dH(x, y)$, falls H die Verteilungsfunktion von (X, Y) ist, bedeutet dies nichts anderes, als die Abhängigkeit von $E(|X \Leftrightarrow Y|^p)$ von der zweidimensionalen gemeinsamen Verteilungsfunktion $H(x, y)$ bei gegebenen Randverteilungsfunktionen F und G zu studieren.

Wir führen diese Untersuchungen direkt in einem allgemeineren Rahmen durch [↔ [7]] und betrachten den Ausdruck $Ek(X, Y)$ für eine Funktion $k : \mathbb{R}^2 \rightarrow \mathbb{R}$, die die folgenden drei Eigenschaften besitzt:

- Symmetrie, d.h. $k(x, y) = k(y, x)$,
- linksseitige Stetigkeit,
- Δ -Antitonie, d.h. $\Delta_{(x, x')}^{(y, y')} k := k(x, y) + k(x', y') \Leftrightarrow k(x, y') \Leftrightarrow k(x', y) \leq 0$ für $x \leq x'$ und $y \leq y'$.

Die Δ -Antitonie von k ist gleichbedeutend mit der Δ -Monotonie von $\Leftrightarrow k$. Für den oben angesprochenen Fall definiere man k durch $(x, y) \mapsto |x \Leftrightarrow y|^p$. Diese spezielle Funktion erfüllt die drei angesprochenen Eigenschaften wie der nachfolgende Satz zeigt.

2.2.2 Satz Die Funktion $k : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto |x \Leftrightarrow y|^p$ ist symmetrisch und stetig. Zudem erfüllt sie für $p \geq 1$ die Eigenschaft der Δ -Antitonie.

BEWEIS: Zu den ersten beiden Eigenschaften ist nichts zu sagen. Es verbleibt die Behauptung, daß für $x \leq x'$ und $y \leq y'$ gilt:

$$\Delta_{(x,x')}^{(y,y')} k := k(x, y) + k(x', y') \Leftrightarrow k(x, y') \Leftrightarrow k(x', y) \leq 0.$$

Dies wird mit Hilfe einer Fallunterscheidung bewiesen.

1.Fall: $x \leq y \leq x' \leq y'$ oder $y \leq x \leq y' \leq x'$. Die Behauptung ergibt sich hier unter Anwendung von Satz A.1 aus dem Anhang.

2.Fall: $x \leq y \leq y' \leq x'$ oder $y \leq x \leq x' \leq y'$. Diese beiden Fälle sind trivial und ergeben sich unter Benutzung der Monotonie der Potenzfunktion.

3.Fall: $x \leq x' \leq y \leq y'$ oder $y \leq y' \leq x \leq x'$. Anwendung von Satz A.2 aus dem Anhang liefert die angegebene Ungleichung. \square

Um die Abhängigkeit des Erwartungswertes $Ek(X, Y)$ von der gemeinsamen Verteilung und damit von der zweidimensionalen Verteilungsfunktion H zu studieren, ist es zunächst erforderlich einen einfacheren Ausdruck als die ursprüngliche Definition $\int_{\mathbb{R}^2} k(x, y) dH(x, y)$ bereitzustellen, durch den die Abhängigkeit verdeutlicht wird. Dies wird im folgenden getan: Aufgrund der Eigenschaften ii) und iii), die man von der Funktion k fordert, wird durch k ein eindeutiges, σ -endliches Maß μ auf der Borelschen σ -Algebra \mathcal{B}^2 durch die Festlegung

$$\mu([x, x'] \times [y, y']) := \Leftrightarrow \Delta_{(x,x')}^{(y,y')} k, \quad x \leq x', y \leq y',$$

determiniert. Sind nun X und Y Zufallsgrößen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) mit Verteilungsfunktion F bzw. G und gemeinsamer Verteilungsfunktion H , so betrachte die Funktion $f : \mathbb{R}^2 \times \Omega \rightarrow \mathbb{R}$, definiert durch:

$$f(x, y, \omega) := \begin{cases} 1, & \text{falls } X(\omega) \leq x, y < Y(\omega) \text{ oder } Y(\omega) \leq x, y < X(\omega) \\ 0, & \text{sonst} \end{cases}.$$

f ist eine $(\mathbb{R}^2 \times \Omega, \mathcal{B}^2 \otimes \mathcal{A})$ -meßbare Funktion. Dies ist am einfachsten einzusehen, wenn man f mit Hilfe von Indikatorfunktionen darstellt:

$$f(x, y, \omega) = 1_{(-\infty, x]}(X(\omega)) \cdot 1_{(-\infty, y]}(X(\omega)) \cdot 1_{(x, \infty)}(Y(\omega)) \cdot 1_{(y, \infty)}(Y(\omega)) \\ + 1_{(-\infty, x]}(Y(\omega)) \cdot 1_{(-\infty, y]}(Y(\omega)) \cdot 1_{(x, \infty)}(X(\omega)) \cdot 1_{(y, \infty)}(X(\omega)).$$

Um nun die Meßbarkeit zu zeigen, reicht es offenkundig, ohne Beschränkung der Allgemeinheit f von der Form $f(x, y, \omega) = 1_{(-\infty, x]}(X(\omega))$ anzunehmen. Aufgrund der linksseitigen Stetigkeit dieser Abbildung in x für festes ω und festes y erhält man dann:

$$f(x, y, \omega) = \lim_{n \rightarrow \infty} \sum_{k \in \mathbb{N}} 1_{(-\infty, \frac{k}{2^n}]}(X(\omega)) \cdot 1_{(\frac{k}{2^n}, \frac{k+1}{2^n}]}(x).$$

Dabei ist $g(x, \omega) := 1_{(-\infty, \frac{k+1}{2^n}]}(X(\omega)) \cdot 1_{(\frac{k}{2^n}, \frac{k+1}{2^n}]}(x)$ $\mathcal{B} \otimes \mathcal{A}$ -meßbar wie man leicht einsieht und somit insbesondere die Abbildung $h(x, y, \omega) = g(x, \omega)$ $\mathcal{B}^2 \otimes \mathcal{A}$ -meßbar. Es folgt die Meßbarkeit von f . Da f außerdem nichtnegativ ist, erhält man mit dem Satz von Fubini:

$$E\left(\int_{\mathbb{R}^2} f d\mu\right) = \int_{\mathbb{R}^2} E f d\mu.$$

Offensichtlich gilt nun unter Ausnutzung der Symmetrie von f

$$\begin{aligned} \int_{\mathbb{R}^2} f d\mu &= \mu([X(\omega), Y(\omega)] \times [X(\omega), Y(\omega)]) \\ &= \Leftrightarrow(k(X, X) + k(Y, Y) \Leftrightarrow k(X, Y) \Leftrightarrow k(Y, X)) \\ &= \Leftrightarrow k(X, X) \Leftrightarrow k(Y, Y) + 2k(X, Y), \end{aligned}$$

wobei ohne Einschränkung $X(\omega) \leq Y(\omega)$ angenommen wurde. Zusammen mit

$$\begin{aligned} E f(x, y, \cdot) &= P(X \leq x \wedge y, Y > x \vee y) + P(X > x \vee y, Y \leq x \wedge y) \\ &= P(X \leq x \wedge y) \Leftrightarrow P(X \leq x \wedge y, Y \leq x \vee y) \\ &\quad + P(Y \leq x \wedge y) \Leftrightarrow P(Y \leq x \wedge y, X \leq x \vee y) \\ &= F(x \wedge y) + G(x \wedge y) \Leftrightarrow H(x \vee y, x \wedge y) \Leftrightarrow H(x \wedge y, x \vee y) \\ &=: A(x, y) \end{aligned}$$

erhält man also:

$$E(\Leftrightarrow k(X, X) \Leftrightarrow k(Y, Y) + 2k(X, Y)) = \int_{\mathbb{R}^2} A(x, y) d\mu.$$

Sind nun die Erwartungswerte $E k(X, X)$ und $E k(Y, Y)$ endlich (im oben betrachteten Spezialfall stellt dies kein Problem dar), so gilt:

$$2E k(X, Y) = E k(X, X) + E k(Y, Y) + \int_{\mathbb{R}^2} A(x, y) d\mu.$$

Man erkennt: Bei fixierten Randverteilungen hängt $E k(X, Y)$ von H nur über A ab. Sei nun $\mathcal{H}(F, G)$ die Klasse aller gemeinsamen Verteilungsfunktionen $H(x, y)$ mit festen Marginalverteilungsfunktionen F und G und $E_H k(X, Y)$ der Erwartungswert, wenn H die gemeinsame Verteilungsfunktion von X und Y ist. Um auszudrücken, wie die Abhängigkeit des Erwartungswertes $E k(X, Y)$ von H aussieht, ist es notwendig eine Ordnung auf der Menge $\mathcal{H}(F, G)$ einzuführen. Dazu können wir festhalten, daß durch

$$H_1 \leq H_2 \quad :\Leftrightarrow \quad H_1(x, y) \leq H_2(x, y) \text{ für alle } x, y \in \mathbb{R}$$

eine partielle Ordnung auf dieser Menge definiert wird. Daß dabei tatsächlich die Eigenschaften der Reflexivität, Transitivität und Antisymmetrie erfüllt sind, ergibt sich direkt

aus den entsprechenden Eigenschaften der gewöhnlichen Größerrelation auf \mathbb{R} . Wir notieren: Sind $H_1, H_2 \in \mathcal{H}(F, G)$ und gilt $H_1 \leq H_2$ in Bezug auf die gerade eingeführte Ordnung, so ist $E_{H_1}(k(X, Y)) \geq E_{H_2}(k(X, Y))$, denn wachsendes H führt zu einer Verkleinerung von A und somit zu einer Abnahme von $Ek(X, Y)$. Mit anderen Worten: $Ek(X, Y)$ ist ein antitones Funktional von H .

Betrachte nun $H_+(x, y) := \min \{F(x), G(y)\}$. Es gilt: $H_+ \in \mathcal{H}(F, G)$ und $H \leq H_+$ für jedes $H \in \mathcal{H}(F, G)$, d.h. H_+ ist größtes Element in $\mathcal{H}(F, G)$. Dabei ergibt sich der zweite Teil der Aussage direkt aus den Ungleichungen

$$F(x) \Leftrightarrow H(x, y) = P(X \leq x) \Leftrightarrow P(X \leq x, Y \leq y) = P(X \leq x, Y > y) \geq 0$$

und $G(y) \Leftrightarrow H(x, y) \geq 0$, die für jedes $H \in \mathcal{H}(F, G)$ gelten. Für unsere Zwecke ist dies ausreichend, denn man weiß nun, daß aufgrund der Antitonie des Funktionals $Ek(X, Y)$ der minimale Wert bei gemeinsamer Verteilung H_+ angenommen wird, d.h. sogar ein Minimum darstellt. Allgemeiner ist es aber sogar möglich zu zeigen

2.2.3 Satz Für die Funktionen $H_-(x, y) := \max \{F(x) + G(y) \Leftrightarrow 1, 0\}$ und $H_+(x, y) := \min \{F(x), G(y)\}$ gilt:

$$H(x, y) \in \mathcal{H}(F, G) \Leftrightarrow H_-(x, y) \leq H(x, y) \leq H_+(x, y)$$

BEWEIS: siehe [7] auf Seite 289 bzw. [10],[15]. □

Nachdem wir nun wissen für welche Verteilungsfunktion H das Minimum angenommen wird, verbleibt eine nähere Betrachtung des zugehörigen Integrals.

2.2.4 Satz Ist (X, Y) ein Zufallsvektor auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) mit Randverteilungsfunktionen F bzw. G und zweidimensionaler Verteilungsfunktion H_+ , so gilt:

$$E_{H_+}k(X, Y) = Ek(F^{-1}(U), G^{-1}(U)) = \int_0^1 k(F^{-1}(u), G^{-1}(u))du,$$

wobei U eine auf $(0, 1)$ rechteck-verteilte Zufallsgröße ist.

BEWEIS: Für eine auf $(0, 1)$ rechteck-verteilte Zufallsgröße U gilt:

$$\begin{aligned} P^{(F^{-1}(U), G^{-1}(U))}((x, x'] \times (y, y']) &= P(F^{-1}(U) \in (x, x'], G^{-1}(U) \in (y, y']) \\ &= P^U(\{u : \inf \{t : F(t) \geq u\} \in (x, x']\} \cap \{u : \inf \{s : G(s) \geq u\} \in (y, y']\}) \\ &= P^U((F(x), F(x')] \cap (G(y), G(y')]) \\ &= \min \{F(x), G(y)\} + \min \{F(x'), G(y')\} \Leftrightarrow \min \{F(x), G(y')\} \Leftrightarrow \min \{F(x'), G(y)\} \\ &= H_+(x, y) + H_+(x', y') \Leftrightarrow H_+(x, y') \Leftrightarrow H(x', y) \\ &= P(X \in (x, x'], Y \in (y, y']), \end{aligned}$$

wobei man das vierte Gleichheitszeichen leicht mittels einer Fallunterscheidung nachvollziehen kann. Unter H_+ gilt folglich: $(X, Y) \sim (F^{-1}(U), G^{-1}(U))$, woraus die Behauptung folgt. \square

Betrachten wir nun speziell wieder die Funktion $k(x, y) := |x \Leftrightarrow y|^p$, so können wir für unsere ursprüngliche Fragestellung festhalten:

$$d_p(F, G) := \inf_{\substack{X \sim F \\ Y \sim G}} \|X \Leftrightarrow Y\|_p = \|F^{-1}(U) \Leftrightarrow G^{-1}(U)\|_p = \left(\int_0^1 |F^{-1}(u) \Leftrightarrow G^{-1}(u)|^p du \right)^{1/p}.$$

Damit ist die letzte Aussage von Satz 2.2.1 gezeigt und gleichzeitig die Vorarbeit geleistet, um den gesamten Beweis zu komplettieren:

BEWEIS VON SATZ 2.2.1: Die Vorbetrachtungen haben gezeigt, daß das Infimum für die im Satz angegebenen Zufallsgrößen angenommen wird. Sind nun F, G zwei Verteilungsfunktionen aus \mathcal{M}_p mit $F = G$, so folgt offenkundig $d_p(F, G) = 0$, da für eine Zufallsgröße X mit $L(X) = F$ gilt: $0 \leq d_p(F, G) = d_p(F, F) \leq \|X \Leftrightarrow X\|_p = 0$. Ist umgekehrt $d_p(F, G) = 0$ für zwei Funktionen $F, G \in \mathcal{M}_p$, so folgt $\|F^{-1}(U) \Leftrightarrow G^{-1}(U)\|_p = d_p(F, G) = 0$. Damit ist aber $F^{-1}(U) = G^{-1}(U)$ f.s. und somit $F = L(F^{-1}(U)) = L(G^{-1}(U)) = G$. Für den Beweis der Dreiecksungleichung betrachte Verteilungsfunktionen $F, G, H \in \mathcal{M}_p$. Dann gilt für $U \sim \mathcal{R}(0, 1)$:

$$\begin{aligned} d_p(F, G) &= \|F^{-1}(U) \Leftrightarrow G^{-1}(U)\|_p \\ &\leq \|F^{-1}(U) \Leftrightarrow H^{-1}(U)\|_p + \|H^{-1}(U) \Leftrightarrow G^{-1}(U)\|_p \\ &= d_p(F, H) + d_p(H, G). \end{aligned}$$

Dabei wurde beim ersten Ungleichheitszeichen die Dreiecksungleichung für die \mathfrak{L}_p -Norm ausgenutzt. Die Symmetrie der Abbildung ergibt sich direkt aus deren Definition. \square

Nun sind wir in der Lage den Begriff der Mallow-Metrik einzuführen:

2.2.5 Definition Die Abbildung $d_p : \mathcal{M}_p \times \mathcal{M}_p \rightarrow \mathbb{R}$, $(F, G) \mapsto \inf_{\substack{X \sim F \\ Y \sim G}} \|X \Leftrightarrow Y\|_p$ heißt *Mallow-Metrik*.

Als interessant erweist sich bei der Mallow-Metrik insbesondere der Fall $p = 1$. Dann gilt

$$d_1(F, G) = \int_0^1 |F^{-1}(t) \Leftrightarrow G^{-1}(t)| dt = \int_{-\infty}^{\infty} |F(t) \Leftrightarrow G(t)| dt$$

[\Leftrightarrow [23]] und folglich ist der durch die Metrik gegebene Abstand nichts anderes als die Fläche zwischen den beiden Graphen von F und G .

2.2.6 Bemerkung Für die allgemeinen Betrachtungen, die in diesem Abschnitt für den Term $Ek(X, Y)$ durchgeführt wurden, ist nicht unbedingt die Δ -Antitonie von k und die linksseitige Stetigkeit erforderlich. Hinreichend sind auch Δ -Isotonie und rechtsseitige Stetigkeit. Dabei ist allerdings zu beachten, daß man für Δ -isotone Funktionen k ein isotones Funktional erhält und sich somit Minimum und Maximum gerade vertauschen.

Zudem sind genauere Aussagen über den Wertebereich des Funktionals möglich. Dieser ist nämlich das möglicherweise unbeschränkte Intervall $[E_{H_-}k(X, Y), E_{H_+}k(X, Y)]$, falls k Δ -isoton ist und $[E_{H_+}k(X, Y), E_{H_-}k(X, Y)]$ bei Δ -Antitonie von k . Auch die Voraussetzung der Symmetrie von k ist nicht notwendig, sondern durch entsprechende andere Annahmen zu ersetzen. Details finden sich in [7].

2.2.2 Konvergenz bezüglich der Mallow-Metrik

Durch den letzten Abschnitt sind die Voraussetzungen geschaffen, um von Konvergenz von Folgen mit Elementen aus \mathcal{M}_p zu sprechen. Dabei soll bedeuten, daß eine Folge F_n von Elementen aus (\mathcal{M}_p, d_p) gegen eine Verteilungsfunktion $G \in \mathcal{M}_p$ konvergiert, wenn $\lim_{n \rightarrow \infty} d_p(F_n, G) = 0$. Ähnlich wie bei dem metrischen Raum $(\mathcal{L}_p, \|\cdot\|_p)$ kann man sich die Frage stellen, ob ein Zusammenhang zwischen der d_p -Konvergenz und der d_q -Konvergenz besteht für $1 \leq q \leq p < \infty$. Daß das gleiche Resultat wie bei der \mathcal{L}_p -Konvergenz zu erwarten ist, ergibt sich aus dem engen Zusammenhang zwischen der d_p -Metrik und der $\|\cdot\|_p$ -Norm. Die Richtigkeit dieser Überlegung zeigt der folgende

2.2.7 Satz *Es seien $1 \leq q \leq p$ reelle Zahlen und F_n , $n \in \mathbb{N}$, eine Folge von Verteilungsfunktionen aus \mathcal{M}_p , die in der d_p -Metrik gegen eine Grenzverteilungsfunktion $F \in \mathcal{M}_p$ konvergiert. Dann liegt auch Konvergenz gegen die gleiche Funktion in d_q -Metrik vor.*

BEWEIS: Sei U eine auf $(0,1)$ rechteck-verteilte Zufallsgröße. Es gilt unter Benutzung der Jensenschen Ungleichung:

$$\begin{aligned} \lim_{n \rightarrow \infty} d_q(F_n, F) &= \lim_{n \rightarrow \infty} \|F_n^{-1}(U) \Leftrightarrow F^{-1}(U)\|_q \leq \lim_{n \rightarrow \infty} \|F_n^{-1}(U) \Leftrightarrow F^{-1}(U)\|_p \\ &= \lim_{n \rightarrow \infty} d_p(F_n, F) = 0, \end{aligned}$$

woraus die Behauptung folgt. □

Interessant bei der Einführung einer „neuen“ Konvergenzart ist, ob man einen Bezug zu bereits bekannten Konvergenzarten herstellen kann. Durch die bereits bewiesene Gleichheit $d_p(F, G) = \|F^{-1}(U) \Leftrightarrow G^{-1}(U)\|_p$, die für beliebige Verteilungsfunktionen $F, G \in \mathcal{M}_p$ gilt, erhalten wir in unserem Fall einen Anhaltspunkt für einen möglichen Zusammenhang zur schwachen Konvergenz von Verteilungen. Dazu sei zum einen daran erinnert, daß aus

der schwachen Konvergenz einer Folge von Verteilungen Q_n gegen ein Q_0 insbesondere die fast sichere Konvergenz von $F_n^{-1}(U)$ gegen $F_0^{-1}(U)$ folgt, falls $U \sim \mathcal{R}(0,1)$ und F_i^{-1} Pseudoinverse der Verteilungsfunktion von Q_i , $0 \leq i$, und andererseits aus der fast sicheren Konvergenz einer Folge von p -fach integrierbaren Zufallsgrößen in Verbindung mit der Konvergenz der p -ten absoluten Momente die \mathfrak{L}_p -Konvergenz folgt [\Leftrightarrow [2], Satz von Riesz].

2.2.8 Satz Die Konvergenz einer Folge $(F_n)_n$ aus \mathcal{M}_p in d_p -Metrik gegen eine Funktion F ist äquivalent zu jeder der beiden folgenden Aussagen:

- a) die zugehörigen Verteilungen konvergieren schwach, und es liegt Konvergenz der p -ten absoluten Momente vor, d.h. $\int |x|^p dF_n(x) \xrightarrow{n \rightarrow \infty} \int |x|^p dF(x)$,
- b) die zugehörigen Verteilungen konvergieren schwach und es liegt gleichgradige Integrierbarkeit von $|x|^p$ bezüglich der Folge F_n vor.

BEWEIS: Für den Beweis sei U eine $\mathcal{R}(0,1)$ -verteilte Zufallsgröße auf dem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) . Zunächst zur Äquivalenz der Konvergenz in d_p -Metrik mit der Aussage a): " \Rightarrow " Aus $\lim_{n \rightarrow \infty} d_p(F_n, F) = 0$ folgt $\lim_{n \rightarrow \infty} \|F_n^{-1}(U) \Leftrightarrow F^{-1}(U)\|_p = 0$ und somit die \mathfrak{L}_p -Konvergenz der Folge $F_n^{-1}(U)$ gegen $F^{-1}(U)$. Damit ergibt sich aber direkt

$$\lim_{n \rightarrow \infty} \int |x|^p dF_n = \lim_{n \rightarrow \infty} \int |F_n^{-1}(U)|^p dP = \int |F^{-1}(U)|^p dP = \int |x|^p dF,$$

d.h. die Konvergenz der p -ten absoluten Momente. Da \mathfrak{L}_p -Konvergenz zudem stärker ist als schwache Konvergenz [\Leftrightarrow Seite 94 in [8] in Kombination mit Satz 36.3 in [2]], folgt $F_n^{-1}(U) \xrightarrow{d} F^{-1}(U)$ und somit $F_n \xrightarrow{w} F$, da $F_n^{-1}(U) \sim F_n$, $F^{-1}(U) \sim F$.

" \Leftarrow " Hat man umgekehrt schwache Konvergenz der Folge F_n , $n \in \mathbb{N}$, gegen F gegeben, so folgt die fast sichere Konvergenz der Folge $F_n^{-1}(U)$, $n \in \mathbb{N}$, gegen die Zufallsgröße $F^{-1}(U)$, und somit zusammen mit der vorausgesetzten Konvergenz der p -ten absoluten Momente die \mathfrak{L}_p -Konvergenz dieser Folge. Daher ist:

$$\lim_{n \rightarrow \infty} d_p(F_n, F) = \lim_{n \rightarrow \infty} \|F_n^{-1}(U) \Leftrightarrow F^{-1}(U)\|_p = 0.$$

Für a) \Leftrightarrow b) sei auf Korollar 50.6. in [2] verwiesen □

Als direkte Konsequenz ergibt sich

2.2.9 Bemerkung Die schwache Konvergenz von Verteilungen in Verbindung mit der Konvergenz der p -ten absoluten Momente wird durch die Mallow-Metrik d_p metrisiert.

Satz 2.2.8 beinhaltet unter anderem die Aussage, daß aus der Konvergenz in der d_p -Metrik insbesondere die Konvergenz der p -ten absoluten Momente folgt. Diese Aussage kann noch verallgemeinert werden.

2.2.10 Korollar Gegeben eine Folge $F_n \in \mathcal{M}_p$ mit $d_p\text{-}\lim_{n \rightarrow \infty} F_n = F$, gilt für $1 \leq q \leq p$

$$\lim_{n \rightarrow \infty} \int |x|^q dF_n = \int |x|^q dF,$$

d.h. die zugehörigen q -ten absoluten Momente konvergieren gegeneinander, sowie im Fall $q = 1$ außerdem

$$\lim_{n \rightarrow \infty} \int x dF_n = \int x dF.$$

BEWEIS: Die erste Aussage ist trivial unter Berücksichtigung von Satz 2.2.7. Für den Fall $q = 1$ beachte man, daß wie im Beweis zu Satz 2.2.7 die \mathfrak{L}_1 -Konvergenz der Folge $F_n^{-1}(U)$ folgt und somit etwa unter Benutzung von Satz 12.2 in [2] die Konvergenz der Folge der Erwartungswerte, was der angegebenen Konvergenzaussage entspricht. \square

Es gilt folgender wichtiger

2.2.11 Satz Der metrische Raum (\mathcal{M}_p, d_p) ist vollständig.

BEWEIS: Sei F_n eine Cauchy-Folge in (\mathcal{M}_p, d_p) , d.h. $\lim_{n, m \rightarrow \infty} d_p(F_n, F_m) = 0$. Dann gilt für $U \sim \mathcal{R}(0, 1)$:

$$\lim_{n, m \rightarrow \infty} \|F_n^{-1}(U) \Leftrightarrow F_m^{-1}(U)\|_p = \lim_{n, m \rightarrow \infty} d_p(F_n, F_m) = 0,$$

was nichts anderes bedeutet, als daß $F_n^{-1}(U)$ eine Cauchy-Folge im vollständigen metrischen Raum $(\mathfrak{L}_p, \|\cdot\|_p)$ ist. Folglich existiert eine Zufallsgröße $X \in \mathfrak{L}_p$ mit $F_n^{-1}(U) \xrightarrow{\mathfrak{L}_p} X$. Sei G die Verteilungsfunktion von X . Dann ergibt sich

$$0 \leq \lim_{n \rightarrow \infty} d_p(F_n, G) \leq \lim_{n \rightarrow \infty} \|F_n^{-1}(U) \Leftrightarrow X\|_p = 0$$

und damit die Vollständigkeit von (\mathcal{M}_p, d_p) . \square

Man kann durch die Metrik d_p eine Pseudo-Metrik auf dem Raum \mathfrak{L}_p definieren, indem man $d_p(X, Y) := d_p(L(X), L(Y))$ setzt. Pseudo-Metrik bedeutet dabei, daß aus „Abstand zweier Elemente gleich Null“ nicht unbedingt die Gleichheit der Elemente selbst folgt. Dies ist hier gegeben, denn hat man zwei Zufallsgrößen deren Verteilungsfunktionen übereinstimmen, so folgt noch nicht die Gleichheit der Zufallsgrößen (vielmehr können diese

grundverschieden sein), sondern lediglich die der zugehörigen Verteilungen.

Man kann nun, ähnlich wie dies bei der Definition der schwachen Konvergenz geschieht, d_p -Konvergenz von reellwertigen Zufallsgrößen definieren, und zwar durch die Festlegung

$$X_n \xrightarrow{d_p} X \Leftrightarrow F_n \xrightarrow{d_p} F,$$

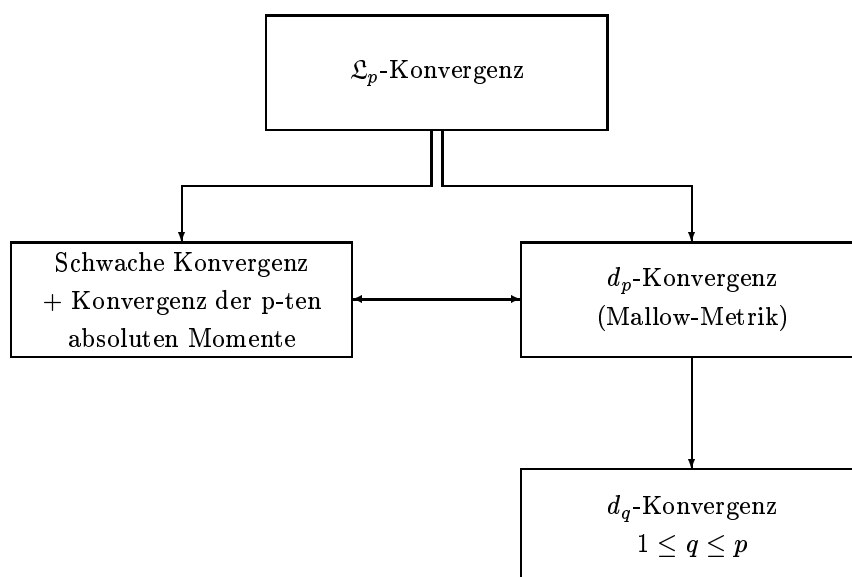
wobei F_n und F die zu X_n bzw. X gehörigen Verteilungsfunktionen sind. Dies entspricht natürlich einer Konvergenz der Zufallsgrößen in der oben eingeführten Pseudo-Metrik.

Insbesondere ist es nun auch möglich, einen Zusammenhang zwischen \mathcal{L}_p -Konvergenz und d_p -Konvergenz herzustellen.

2.2.12 Satz *Gegeben seien reellwertige Zufallsgrößen $X, X_n, n \in \mathbb{N}$. Dann folgt aus der \mathcal{L}_p -Konvergenz von X_n gegen X insbesondere die d_p -Konvergenz.*

BEWEIS: Trivial unter Berücksichtigung der Definition der d_p -Metrik. □

Die in diesem Abschnitt ausgearbeiteten Zusammenhänge zwischen den Konvergenzarten faßt das folgende Schema noch einmal zusammen:



2.2.3 Die Wasserstein-Metrik

In der Literatur tritt in Verbindung mit der Mallow-Metrik häufig der Begriff der Wasserstein-Metrik auf. Ziel dieses Abschnitts ist es, die Unterschiede und Gemeinsamkeiten dieser beiden Metriken aufzuzeigen.

Betrachten wir allgemein einen metrischen Raum (\mathfrak{X}, ρ) und die zugehörige Borelsche σ -Algebra $\mathcal{B}(\mathfrak{X})$, so kann man den Abstand R zweier Wahrscheinlichkeitsmaße P und Q auf $(\mathfrak{X}, \mathcal{B}(\mathfrak{X}))$ definieren durch

$$R(P, Q) := \inf_{\substack{X \sim P \\ Y \sim Q}} E\rho(X, Y),$$

wobei das Infimum über alle Zufallsvariablen X und Y mit Werten in $(\mathfrak{X}, \mathcal{B}(\mathfrak{X}))$ gebildet wird, so daß X die Verteilung P und Y die Verteilung Q besitzt. Zu beachten ist dabei, daß die Metrik ρ gewisse Meßbarkeitsanforderungen erfüllen muß, damit die Erwartungswertbildung möglich ist.

Der Wert $R(P, Q)$ heißt Wasserstein-Abstand [\Leftrightarrow [9]] und kann den Wert $+\infty$ annehmen. Wir notieren:

2.2.13 Satz *Seien (\mathfrak{X}, ρ) ein metrischer Raum und $\mathcal{B}(\mathfrak{X})$ die σ -Algebra der Borelschen Mengen. Sei weiterhin P_0 ein Wahrscheinlichkeitsmaß auf $(\mathfrak{X}, \mathcal{B}(\mathfrak{X}))$ und $U(P_0)$ die Menge aller Wahrscheinlichkeitsmaße auf $(\mathfrak{X}, \mathcal{B}(\mathfrak{X}))$ mit $R(P_0, P) < \infty$. Dann ist $U(P_0)$ zusammen mit der Abbildung $R(\cdot, \cdot)$ ein metrischer Raum und R wird Wasserstein-Metrik genannt. Ist weiterhin der metrische Raum \mathfrak{X} separabel, so ist $(U(P_0), R)$ vollständig.*

BEWEIS: siehe [9], Theorem 2, S.469 ff. □

Zur oben erwähnten Meßbarkeitsanforderung an die Metrik ρ können wir den folgenden Satz festhalten, der dem Buch von Gänsler und Stute über Wahrscheinlichkeitstheorie entnommen ist [\Leftrightarrow [11]].

2.2.14 Satz *Sei (\mathfrak{X}, ρ) ein separabler metrischer Raum und ξ, η Zufallsvariable auf dem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) mit Werten in $(\mathfrak{X}, \mathcal{B}(\mathfrak{X}))$. Dann ist $\rho(\xi, \eta)$ $(\Omega, \mathcal{A}) \Leftrightarrow (\mathbb{R}, \mathcal{B})$ -meßbar.*

BEWEIS: siehe [11], Satz 8.1.4 auf Seite 329. □

Betrachten wir nun speziell den metrischen Raum $(\mathbb{R}, |\cdot|)$, so erkennt man, daß in diesem Fall die Wasserstein-Metrik mit der Mallow-Metrik d_1 übereinstimmt. Für $p \geq 2$ kann man einen direkten Bezug nicht mehr herstellen, da in diesem Fall die Metrik $|\cdot|$ noch mit der Transformation $g(x) = x^p$ verknüpft und der gesamte Wert mit der Funktion $h(x) = x^{1/p}$ renormiert wird.

2.3 Fixpunkte der Abbildung S

Für gegebene Verteilungen der Zufallsvariablen $T = (T_1, T_2, \dots)$ und C entsprechend den Anforderungen aus Abschnitt 2.1 ist $S : (\mathcal{M}_p, d_p) \rightarrow (\mathcal{M}_p, d_p)$, $F \mapsto \sum_i T_i X_i + C$, eine Abbildung des vollständigen metrischen Raumes (\mathcal{M}_p, d_p) (\mathcal{M}_p zusammen mit der Mallows-Metrik) in sich selbst.

Im folgenden Abschnitt wenden wir uns nun der noch fehlenden, entscheidenden Kontraktionseigenschaft der Abbildung S zu, d.h. wir werden systematisch das kontrahierende Verhalten von Abbildungen des obigen Typs studieren. Dabei werden wir uns zunächst auf den Spezialfall $p = 2$ beschränken, da die durchzuführenden Rechnungen einfacher sind und zudem schwächere Voraussetzungen an die Zufallsvariablen ausreichen. Allerdings sei darauf hingewiesen, daß folglich auch nur der Nachweis von Fixpunkten mit existierendem 2-ten absoluten Moment erfolgt.

Im Unterabschnitt 2.3.2 gehen wir dann zur Betrachtung des allgemeinen Falles über.

2.3.1 Der Spezialfall: (\mathcal{M}_2, d_2) als Definitions- und Wertebereich der Abbildung S

Wie in Kapitel 1 bereits ausgeführt, spielt für den Nachweis der Existenz von Fixpunkten das Prinzip der iterativen Approximation und damit Folgen der Art $(S^n(F))$ mit Startverteilung $F \in \mathcal{M}_2$ eine entscheidende Rolle. Wir hatten bereits angemerkt, daß es für die Existenz eines Limes ausreicht, Konstanten $c \in \mathbb{R}^+$ und $\lambda < 1$ zu finden mit der Eigenschaft

$$d_2(S^{n+1}(F), S^n(F)) \leq c\lambda^n;$$

vgl. Seite 13. Um dann trotzdem noch Eindeutigkeitsaussagen machen zu können, ist es erforderlich, sich die zugehörige Folge der Erwartungswerte anzuschauen. Dies geschieht im folgenden Lemma. Besonders wichtig für spätere Anwendungen sind die Punkte 2) und 5).

2.3.1 Lemma *Es seien entsprechende Voraussetzungen gegeben, die die Existenz der Abbildung $S : \mathcal{M}_p \rightarrow \mathcal{M}_p$ im \mathfrak{L}_p -Sinn gewährleisten (für ein $1 \leq p < \infty$) (vgl. Abschnitt 2.1). Zudem sei $\sum_i T_i \in \mathfrak{L}_1$. Dann gelten die folgenden Aussagen:*

- 1) Ist $E(C)=0$, so gilt: $S: \mathcal{M}_p^0 \mapsto \mathcal{M}_p^0$.
- 2) Ist $|E(\sum_i T_i)| < 1$, so konvergiert $E(S^n(F))$ für n gegen unendlich exponentiell schnell gegen $E(C)(1 \Leftrightarrow E(\sum_i T_i))^{-1}$.

3) Ist $|E(\sum_i T_i)| > 1$ und $E(F) + E(C) \cdot (E(\sum_i T_i) \Leftrightarrow 1)^{-1} \neq 0$, so gilt

$$\lim_{n \rightarrow \infty} |E(S^n(F))| = \infty.$$

4) Ist $|E(\sum_i T_i)| > 1$, $E(C)$ und $E(F)$ nicht beide Null und $E(F) + E(C) \cdot (E(\sum_i T_i) \Leftrightarrow 1)^{-1} = 0$, so gilt $\lim_{n \rightarrow \infty} E(S^n(F)) = E(C) \cdot (1 \Leftrightarrow E(\sum_i T_i))^{-1}$.

5) Ist $E(\sum_i T_i) = 1$ und $E(C) = 0$, so gilt $E(F) = E(S^n(F))$ für alle n .

6) Ist $E(\sum_i T_i) = 1$ und $E(C) \neq 0$, so gilt $\lim_{n \rightarrow \infty} |E(S^n(F))| = \infty$.

7) Ist $E(\sum_i T_i) = \Leftrightarrow 1$, so gilt für alle $n \in \mathbb{N}$: $E(S^{2n}(F)) = E(F)$, $E(S^{2n+1}(F)) = E(C) \Leftrightarrow E(F)$.

BEWEIS: Die Zufallsgrößen (C, T) , X_i , $i \in \mathbb{N}$, seien unabhängig und $X_i \sim S^{n-1}(F)$. Da für eine in \mathfrak{L}_1 -Norm konvergente Folge von Zufallsgrößen auch deren erste Momente konvergieren, folgt:

$$\begin{aligned} E(S^n(F)) &= E\left(\sum_i T_i X_i + C\right) = E\left(\lim_{m \rightarrow \infty} \sum_{i=1}^m T_i X_i + C\right) = \lim_{m \rightarrow \infty} E\left(\sum_{i=1}^m T_i X_i\right) + E(C) \\ &= E(X_1) \cdot \lim_{m \rightarrow \infty} E\left(\sum_{i=1}^m T_i\right) + E(C) = E(S^{n-1}(F)) \cdot E\left(\lim_{m \rightarrow \infty} \sum_{i=1}^m T_i\right) + E(C) \\ &= E(S^{n-1}(F)) \cdot E\left(\sum_i T_i\right) + E(C). \end{aligned}$$

Dies liefert sofort:

$$E(S^n(F)) = E(F) \cdot \left(E\left(\sum_i T_i\right)\right)^n + E(C) \cdot \sum_{i=1}^{n-1} \left(E\left(\sum_i T_i\right)\right)^i.$$

Alle Behauptungen ergeben sich nun leicht aus dieser Gleichung. \square

2.3.2 Bemerkung Die zusätzliche Forderung $\sum_i T_i \in \mathfrak{L}_1$ ergibt sich in den Fällen 1) \Leftrightarrow 3) aus Satz 2.1.1 automatisch aus den die Existenz von S sichernden Voraussetzungen und ist somit unproblematisch.

Sei im folgenden für $b \in \mathbb{R}$ und $F \in \mathcal{M}_1$

$$F^b := F(\cdot + (E(F) \Leftrightarrow b)).$$

Es gilt also $E(F^b) = b$.

2.3.3 Lemma Gegeben seien $F, G \in \mathcal{M}_2$.

(a) Sind X, Y Zufallsgrößen mit $L(X) = F, L(Y) = G$ und $d_2(F, G) = \|X \Leftrightarrow Y\|_2$, so gilt:

$$d_2(F^0, G^0) = \|X \Leftrightarrow EX \Leftrightarrow (Y \Leftrightarrow EY)\|_2.$$

(b) Sind X, Y Zufallsgrößen mit $L(X) = F^0, L(Y) = G^0$ und $d_2(F^0, G^0) = \|X \Leftrightarrow Y\|_2$, so gilt:

$$d_2(F, G) = \|X + EF \Leftrightarrow (Y + EG)\|_2.$$

(c) Es gilt:

$$d_2(F^0, G^0) \leq d_2(F, G) \leq d_2(F^0, G^0) + |E(F) \Leftrightarrow E(G)|.$$

BEWEIS: Zu a): Seien X_1, Y_1 Zufallsgrößen mit $L(X_1) = F^0, L(Y_1) = G^0$ und $d_2(F^0, G^0) = \|X_1 \Leftrightarrow Y_1\|_2$. Dann ist offensichtlich $L(X_1 + EX) = F, L(Y_1 + EY) = G$ und

$$\begin{aligned} d_2^2(F^0, G^0) &= E(X_1 \Leftrightarrow Y_1)^2 = E(X_1 + EX \Leftrightarrow (Y_1 + EY) \Leftrightarrow (EX \Leftrightarrow EY))^2 \\ &= E((X_1 + EX) \Leftrightarrow (Y_1 + EY))^2 \Leftrightarrow (EX \Leftrightarrow EY)^2 \\ &\geq d_2^2(F, G) \Leftrightarrow (EX \Leftrightarrow EY)^2, \end{aligned}$$

wobei wir ausgenutzt haben, daß $EX_1 = EY_1 = 0$. Andererseits:

$$\begin{aligned} d_2^2(F^0, G^0) &\leq E(X \Leftrightarrow EX \Leftrightarrow Y + EY)^2 = E((X \Leftrightarrow Y) \Leftrightarrow (EX \Leftrightarrow EY))^2 \\ &= E(X \Leftrightarrow Y)^2 \Leftrightarrow (EX \Leftrightarrow EY)^2 \\ &= d_2^2(F, G) \Leftrightarrow (EX \Leftrightarrow EY)^2. \end{aligned}$$

Es folgen die Behauptungen (a) und (c). Der Beweis von (b) erfolgt analog. \square

Das gerade bewiesene Lemma findet nun Anwendung im folgenden

2.3.4 Lemma Seien $C \in \mathfrak{L}_2, F, G \in \mathfrak{M}_2$.

(1) Falls $E(\sum_i T_i^2) < \infty$, so existiert $S : \mathfrak{M}_2^0 \rightarrow \mathfrak{M}_2$ im \mathfrak{L}_2 -Sinn und

$$d_2^2(S(F^0), S(G^0)) \leq E\left(\sum_i T_i^2\right) \cdot d_2^2(F^0, G^0).$$

(2) Falls $\sum_i T_i \in \mathfrak{L}_2$ und $E(\sum_i T_i^2) < \infty$, so existiert $S : \mathfrak{M}_2 \rightarrow \mathfrak{M}_2$ im \mathfrak{L}_2 -Sinn und

$$d_2^2(S(F), S(G)) \leq E\left(\sum_i T_i^2\right) \cdot d_2^2(F^0, G^0) + \left\| \sum_i T_i \right\|_2^2 \cdot (E(F) \Leftrightarrow E(G))^2.$$

BEWEIS: Die erste Ungleichung ist aufgrund des günstigen Definitionsbereichs der Abbildung S leicht zu beweisen. Gegeben unabhängige Zufallsvariablen $(X_i, Y_i), i \in \mathbb{N}$, (C, T) mit $L(X_i) = F^0, L(Y_i) = G^0, \|X_i \Leftrightarrow Y_i\|_2 = d_2(F^0, G^0)$ (vgl. Satz 2.2.1) gilt nämlich

$$\begin{aligned} d_2^2(S(F^0), S(G^0)) &\leq \left\| \sum_i T_i X_i \Leftrightarrow \sum_i T_i Y_i \right\|_2^2 = \lim_{n \rightarrow \infty} E \left(\sum_{i=1}^n T_i (X_i \Leftrightarrow Y_i) \right)^2 \\ &= \lim_{n \rightarrow \infty} E(X_1 \Leftrightarrow Y_1)^2 \cdot E \left(\sum_{i=1}^n T_i^2 \right) = E \left(\sum_i T_i^2 \right) \cdot d_2^2(F^0, G^0), \end{aligned}$$

wobei insbesondere ausgenutzt wurde, daß $E(X_i \Leftrightarrow Y_i) = 0$ ist.

Für die zweite Aussage wählen wir unabhängige Zufallsvariablen $(C, T), (X_i, Y_i), i \in \mathbb{N}$, mit $L(X_i) = F, L(Y_i) = G, \|X_i \Leftrightarrow Y_i\|_2 = d_2(F, G)$. Es folgt unter Verwendung der Bezeichnungen $X'_i := X_i \Leftrightarrow EX_i, Y'_i := Y_i \Leftrightarrow EY_i$ und $c := (EX_1 \Leftrightarrow EY_1)$:

$$\begin{aligned} d_2^2(S(F), S(G)) &\leq \left\| \sum_i T_i X_i \Leftrightarrow \sum_i T_i Y_i \right\|_2^2 \\ &= E \left(\sum_i T_i (X_i \Leftrightarrow Y_i) \right)^2 = \lim_{n \rightarrow \infty} E \left(\sum_{i=1}^n T_i (X_i \Leftrightarrow Y_i) \right)^2 \\ &= \lim_{n \rightarrow \infty} E \left(\sum_{i=1}^n T_i (X'_i \Leftrightarrow Y'_i) + (EX_1 \Leftrightarrow EY_1) \cdot \sum_{i=1}^n T_i \right)^2 \\ &= \lim_{n \rightarrow \infty} E \left(\left(\sum_{i=1}^n T_i (X'_i \Leftrightarrow Y'_i) \right)^2 + 2c \cdot \left(\sum_{i=1}^n T_i (X'_i \Leftrightarrow Y'_i) \right) \sum_{i=1}^n T_i + c^2 \left(\sum_{i=1}^n T_i \right)^2 \right) \\ &= \lim_{n \rightarrow \infty} \left(E \left(\sum_{i=1}^n T_i (X'_i \Leftrightarrow Y'_i) \right)^2 + 2cE \left(\sum_{i=1}^n \sum_{k=1}^n T_i (X'_i \Leftrightarrow Y'_i) T_k \right) + c^2 E \left(\sum_{i=1}^n T_i \right)^2 \right). \end{aligned}$$

Vermöge der geforderten Unabhängigkeiten und $E(X'_i \Leftrightarrow Y'_i) = 0$, ergibt sich das Wegfallen sowohl des kompletten zweiten Terms als auch der gemischten Terme des ersten Erwartungswertes und wir können festhalten:

$$\begin{aligned} d_2^2(S(F), S(G)) &\leq \lim_{n \rightarrow \infty} E \left(\sum_{i=1}^n T_i^2 \right) \cdot E(X'_1 \Leftrightarrow Y'_1)^2 + \lim_{n \rightarrow \infty} c^2 E \left(\sum_{i=1}^n T_i \right)^2 \\ &= E \left(\sum_i T_i^2 \right) \cdot d_2^2(F^0, G^0) + \left\| \sum_i T_i \right\|_2^2 \cdot (E(F) \Leftrightarrow E(G))^2. \end{aligned}$$

Im letzten Schritt wurde dabei der Satz von der monotonen Konvergenz, die Voraussetzung $\sum_i T_i \in \mathfrak{L}_2$ und Lemma 2.3.3 benutzt. \square

Nun sind alle Vorbereitungen für die beiden folgenden Hauptresultate dieses Abschnitts getroffen.

2.3.5 Satz *Es gelte:*

$$E\left(\sum_i T_i^2\right) < 1, \quad E(C) = 0, \quad \|C\|_2 < \infty.$$

Dann existiert $S : \mathcal{M}_2^0 \rightarrow \mathcal{M}_2^0$ im \mathfrak{L}_2 -Sinn und hat einen eindeutigen Fixpunkt. Die Folge $(S^n(F))_{n \in \mathbb{N}}$ konvergiert für jedes $F \in \mathcal{M}_2^0$ in d_2 -Metrik exponentiell schnell gegen diesen Fixpunkt.

BEWEIS: Die Einhaltung des Wertebereichs \mathcal{M}_2^0 ist eine direkte Konsequenz der Voraussetzung $E(C) = 0$ und des gewählten Definitionsbereichs \mathcal{M}_2^0 ; vgl. Lemma 2.3.1, 1). Nach Lemma 2.3.4 ist die Abbildung S eine α -Kontraktion mit $\alpha := (E(\sum_i T_i^2))^{1/2}$. Die Aussage des Satzes ist nun eine direkte Konsequenz aus dem Banachschen Fixpunktsatz 1.2 und Korollar 1.4. \square

2.3.6 Satz *Es gelte:*

$$E\left(\sum_i T_i^2\right) < 1, \quad \left\| \sum_i T_i \right\|_2 < \infty, \quad |E(\sum_i T_i)| < 1, \quad \|C\|_2 < \infty.$$

Dann existiert die Abbildung $S : \mathcal{M}_2 \rightarrow \mathcal{M}_2$ im \mathfrak{L}_2 -Sinn, und für jedes $F \in \mathcal{M}_2$ konvergiert die Folge $(S^n(F))_{n \in \mathbb{N}}$ in d_2 -Metrik exponentiell schnell gegen den eindeutigen Fixpunkt der Abbildung S .

BEWEIS: Wie schon angekündigt, kommt nun die abschließende Bemerkung des ersten Kapitels zur Anwendung. Wir werden die Existenz von Konstanten c_F und λ nachweisen mit der Eigenschaft:

$$b_n := d_2^2(S^{n+1}(F), S^n(F)) \leq c_F \lambda^n.$$

Wenden wir Lemma 2.3.4, (2) und (3), an, ergibt sich die Abschätzung

$$d_2^2(S^{n+1}(F), S^n(F)) \leq c_1 \cdot d_2^2(S^n(F), S^{n-1}(F)) + c_2 \cdot (E(S^n(F)) \Leftrightarrow E(S^{n-1}(F)))^2,$$

mit $c_1 := E(\sum_i T_i^2)$ und $c_2 := \left\| \sum_i T_i \right\|_2^2$. Unter Verwendung von

$$\begin{aligned} (E(S^n(F)) \Leftrightarrow E(S^{n-1}(F)))^2 &= (E(S^{n-1}(F)) \Leftrightarrow E(S^{n-2}(F)))^2 \cdot (E(\sum_i T_i))^2 \\ &= (E(S(F)) \Leftrightarrow E(F))^2 \cdot (E(\sum_i T_i))^{2(n-1)}, \end{aligned}$$

erhalten wir insgesamt

$$(2.1) \quad b_n = d_2^2(S^{n+1}(F), S^n(F)) \leq c_1 \cdot b_{n-1} + c_3 \cdot \eta^{n-1},$$

wobei $c_3 := c_2 \cdot (E(S(F)) \Leftrightarrow E(F))^2$ und $\eta := (E(\sum_i T_i))^2$. Die Konstanten c_1 und η sind aufgrund der Voraussetzungen echt kleiner als Eins. Wähle nun $\lambda < 1$ mit

$$\eta \leq \lambda \text{ und } \lambda > c_1$$

und c_F mit

$$c_3 \leq c_F \cdot (\lambda \Leftrightarrow c_1) \text{ und } c_F \geq d_2^2(S(F), F).$$

Eine vollständige Induktion nach n liefert schließlich:

$$b_n \leq c_F \cdot \lambda^n.$$

Induktionsanfang $n=0$: $b_0 = d_2^2(S(F), F) \leq c_F \cdot \lambda^0$.

Induktionsschritt $n \rightarrow n+1$: Vermöge (2.1) folgt $b_{n+1} \leq c_1 \cdot b_n + c_3 \cdot \eta^n$. Wendet man darauf die Induktionsvoraussetzung unter Berücksichtigung der Wahl von λ und c_F an, so ergibt sich:

$$\begin{aligned} b_{n+1} &\leq c_1 \cdot c_F \cdot \lambda^n + \eta^n \cdot c_3 \leq c_1 \cdot c_F \cdot \lambda^n + \lambda^n \cdot c_3 = \lambda^n \cdot (c_1 \cdot c_F + c_3) \leq \lambda^n \cdot c_F \cdot \lambda \\ &= c_F \cdot \lambda^{n+1}. \end{aligned}$$

Analog zum Beweis des Banachschen Fixpunktsatzes erhalten wir nun, daß $(S^n(F))$ eine Cauchy-Folge in \mathcal{M}_2 ist und daher aufgrund der Vollständigkeit von (\mathcal{M}_2, d_2) gegen eine Grenzverteilungsfunktion G in \mathcal{M}_2 konvergiert; vgl. Satz 2.2.11. Die Konvergenz ist auch hier wiederum exponentiell schnell. Darüber hinaus liefert Lemma 2.3.4, (2), die Stetigkeit von S , wenn man beachtet, daß mit $F_n \xrightarrow{d_2} F$ sowohl $F_n^0 \xrightarrow{d_2} F^0$ als auch $\lim_{n \rightarrow \infty} E(F_n) = E(F)$ gilt; vgl. Lemma 2.3.3, (c) und Korollar 2.2.10. Die Grenzverteilungsfunktion G ist folglich Fixpunkt der Abbildung S .

Da S nicht wie im vorherigen Satz die Kontraktionseigenschaft besitzt, erweist sich der Beweis der Eindeutigkeit des Fixpunktes als schwieriger. Seien F, G zwei Fixpunkte der Abbildung S . Mit Lemma 2.3.1, (2), folgt die Übereinstimmung von $E(F)$ mit $E(G)$ und Lemma 2.3.4, (2), und Lemma 2.3.3, (c), liefern die Abschätzung:

$$(2.2) \quad d_2^2(F, G) = d_2^2(S(F), S(G)) \leq E\left(\sum_i T_i^2\right) \cdot d_2^2(F, G).$$

Diese kann aber aufgrund der Voraussetzung $E(\sum_i T_i^2) < 1$ nur erfüllt sein, falls $d_2^2(F, G) = 0$ und somit $F = G$. \square

Es ist möglich in Satz 2.3.6 die Voraussetzung $|E(\sum_i T_i)| < 1$ durch $E(\sum_i T_i) = 1$ und $E(C) = 0$ zu ersetzen. Statt (2.1) erhält man dann direkt $b_n \leq E(\sum_i T_i^2) \cdot b_{n-1}$ und damit wiederum die Eigenschaft einer Cauchy-Folge von $F, S(F), S^2(F), \dots$. Jedoch geht die Eindeutigkeit des Fixpunktes verloren, wobei im Fall $\sum_i T_i \equiv 1$ eine genauere Aussage über die Beziehung der einzelnen Fixpunkte zueinander gemacht werden kann:

2.3.7 Satz Ersetzt man in Satz 2.3.6 die Annahme $|E(\sum_i T_i)| < 1$ durch $E(C) = 0$ und $E(\sum_i T_i) = 1$, so erhält man für jedes $F \in \mathcal{M}_2$ exponentiell schnelle Konvergenz der Folge $S^n(F)$, $n \in \mathbb{N}$, in d_2 -Metrik gegen einen Fixpunkt G der Abbildung $S : \mathcal{M}_2 \rightarrow \mathcal{M}_2$, der denselben Erwartungswert hat wie die Startverteilung F . Insbesondere geht die Eindeutigkeit des Fixpunktes verloren und man erhält vielmehr für jedes $b \in \mathbb{R}$ einen Fixpunkt von S . Für vorgegebenen Erwartungswert b ist der Fixpunkt allerdings wieder eindeutig. Weiß man zudem $\sum_i T_i \equiv 1$, so sind die einzelnen Fixpunkte Translationen voneinander, d.h. es existiert ein $G \in \mathcal{M}_2$, so daß $\{G^b, b \in \mathbb{R}\}$ die Menge aller Fixpunkte darstellt.

BEWEIS: Die Konvergenzaussage ergibt sich aus dem Beweis des vorherigen Satzes in Verbindung mit der zuvor gemachten Anmerkung. Daß der Fixpunkt für vorgegebenen Erwartungswert b eindeutig ist, folgt aus Lemma 2.3.4, (2). Zu zeigen bleibt also nur noch die Translationseigenschaft der Fixpunkte im Spezialfall $\sum_i T_i \equiv 1$. Dazu reicht es offensichtlich nachzuweisen, daß falls G^b Fixpunkt der Abbildung S ist (mit Erwartungswert b), auch die Verteilungsfunktionen G^a , $a \in \mathbb{R}$, Fixpunkte darstellen. Seien (T, C) , X_i , $i \in \mathbb{N}$, unabhängige Zufallsgrößen mit $L(X_i) = G^b$. Aufgrund der Fixpunkteigenschaft von G^b gilt dann: $G^b = L(\sum_i T_i X_i + C)$. Betrachte weiter $Y_i := X_i \Leftrightarrow b + a$. Offensichtlich ist $L(Y_i) = G^a$ und

$$\begin{aligned} S(G^a) &= L\left(\sum_i T_i Y_i + C\right) = L\left(\sum_i T_i X_i + \sum_i T_i (a \Leftrightarrow b) + C\right) \\ &= L\left(\sum_i T_i X_i + C + (a \Leftrightarrow b)\right) = G^a. \end{aligned}$$

□

2.3.8 Bemerkung Es ist möglich, auch für Abbildungen $S : \mathcal{M}_2 \rightarrow \mathcal{M}_2$ Voraussetzungen anzugeben, so daß die Kontraktionseigenschaft erfüllt ist. Dies werden wir aber in allgemeinerem Rahmen in Abschnitt 2.3.2 untersuchen. Es sei an dieser Stelle bereits auf den entsprechenden Satz 2.3.11 verwiesen.

Repräsentation des Fixpunktes durch eine Zufallsgröße

Bisher haben wir uns nur mit Existenzaussagen beschäftigt. In diesem Unterabschnitt werden wir eine Zufallsgröße angeben, deren Verteilung dem Fixpunkt der Abbildung S (sofern die Existenz durch entsprechende Annahmen gesichert ist) entspricht. Diese Zufallsgröße wird sich dabei als fast sicherer Limes einer geeigneten Folge von Zufallsgrößen herausstellen.

Um dieses Ziel zu erreichen, ist es sinnvoll, folgende Notationen einzuführen:

$I := \emptyset \cup \bigcup_{n=1}^{\infty} \mathbb{N}^n$. Für $\sigma \in \bigcup_{n \geq i} \mathbb{N}^n$, $i = 1, 2, \dots$, sei $\sigma|i := (\sigma_1, \sigma_2, \dots, \sigma_i)$ und $\sigma|0 := \emptyset$. $|\sigma|$ bezeichne die Länge von σ , d.h. $|\sigma| = n$ für $\sigma = (\sigma_1, \dots, \sigma_n)$, $|\emptyset| = 0$, und $\sigma i := (\sigma_1, \sigma_2, \dots, \sigma_n, i)$ für $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_n)$ und $i \in \mathbb{N}$.

Desweiteren seien $(C(\sigma), T(\sigma))$, $\sigma \in I$, unabhängige, identisch verteilte Zufallsvektoren und $X(\sigma)$, $\sigma \in I$, unabhängig identisch verteilte Zufallsgrößen mit Verteilungsfunktion F , die zudem unabhängig von $(C(\sigma), T(\sigma))$, $\sigma \in I$, seien. Definiere dann Zufallsgrößen Y_n^F , $n \in \mathbb{N} \cup \{0\}$, sukzessive durch:

$$\begin{aligned} Y_0^F &:= X(\emptyset), & Y_1^F &:= C(\emptyset) + \sum_{|\sigma|=1} T_{\sigma_1}(\emptyset)X(\sigma) \\ Y_2^F &:= C(\emptyset) + \sum_{|\sigma|=1} T_{\sigma_1}(\emptyset)C(\sigma) + \sum_{|\sigma|=2} T_{\sigma_1}(\emptyset)T_{\sigma_2}(\sigma|1)X(\sigma) \end{aligned}$$

und allgemein:

$$Y_n^F := C(\emptyset) + \sum_{k=1}^{n-1} \sum_{\substack{\sigma \in I \\ |\sigma|=k}} C(\sigma) \prod_{j=1}^k T_{\sigma_j}(\sigma|(j \Leftrightarrow 1)) + \sum_{\substack{\sigma \in I \\ |\sigma|=n}} X(\sigma) \prod_{j=1}^n T_{\sigma_j}(\sigma|(j \Leftrightarrow 1)),$$

wobei natürlich $\sum_{|\sigma|=n}^{\sigma \in I} \dots = \sum_{\sigma_1=1}^{\infty} \sum_{\sigma_2=1}^{\infty} \dots \sum_{\sigma_n=1}^{\infty} \dots$ bedeuten soll. Man kann bereits an dieser Definition erkennen, daß $L(Y_n^F) = S^n(F)$. Dies und die rekursive Struktur der Folge von Verteilungsfunktionen $L(Y_n)$ wird jedoch noch offensichtlicher, wenn man Y_n^F folgendermaßen umformt:

$$\begin{aligned} Y_n^F &= C(\emptyset) + \sum_{k=1}^{n-1} \sum_{\substack{\sigma \in I \\ |\sigma|=k}} C(\sigma) \prod_{j=1}^k T_{\sigma_j}(\sigma|(j \Leftrightarrow 1)) + \sum_{\substack{\sigma \in I \\ |\sigma|=n}} X(\sigma) \prod_{j=1}^n T_{\sigma_j}(\sigma|(j \Leftrightarrow 1)) \\ &= \sum_{k=1}^{n-1} \sum_{i=1}^{\infty} \sum_{\substack{\sigma_1=i \\ |\sigma|=k}} C(\sigma) \prod_{j=1}^k T_{\sigma_j}(\sigma|(j \Leftrightarrow 1)) + \sum_{i=1}^{\infty} \sum_{\substack{\sigma_1=i \\ |\sigma|=n}} X(\sigma) \prod_{j=1}^n T_{\sigma_j}(\sigma|(j \Leftrightarrow 1)) + C(\emptyset) \\ &= \sum_{i=1}^{\infty} \left(\sum_{k=1}^{n-1} \sum_{\substack{\sigma_1=i \\ |\sigma|=k}} C(\sigma) \prod_{j=1}^k T_{\sigma_j}(\sigma|(j \Leftrightarrow 1)) + \sum_{\substack{\sigma_1=i \\ |\sigma|=n}} X(\sigma) \prod_{j=1}^n T_{\sigma_j}(\sigma|(j \Leftrightarrow 1)) \right) + C(\emptyset) \\ &= \sum_{i=1}^{\infty} \left(C(i) \cdot T_i(\emptyset) + \sum_{k=2}^{n-1} \sum_{\substack{\sigma_1=i \\ |\sigma|=k}} C(\sigma) \prod_{j=1}^k T_{\sigma_j}(\sigma|(j \Leftrightarrow 1)) \right. \\ &\quad \left. + \sum_{\substack{\sigma_1=i \\ |\sigma|=n}} X(\sigma) \prod_{j=1}^n T_{\sigma_j}(\sigma|(j \Leftrightarrow 1)) \right) + C(\emptyset) \end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^{\infty} \left(C(i) \cdot T_i(\emptyset) + \sum_{k=2}^{n-1} \sum_{\substack{\sigma_1=i \\ |\sigma|=k}} C(\sigma) T_i(\emptyset) \prod_{j=2}^k T_{\sigma_j}(\sigma|(j \Leftrightarrow 1)) \right. \\
&\quad \left. + \sum_{\substack{\sigma_1=i \\ |\sigma|=n}} X(\sigma) T_i(\emptyset) \prod_{j=2}^n T_{\sigma_j}(\sigma|(j \Leftrightarrow 1)) \right) + C(\emptyset) \\
&= \sum_{i=1}^{\infty} T_i(\emptyset) \left(C(i) + \sum_{k=2}^{n-1} \sum_{\substack{\sigma_1=i \\ |\sigma|=k}} C(\sigma) \prod_{j=2}^k T_{\sigma_j}(\sigma|(j \Leftrightarrow 1)) \right. \\
&\quad \left. + \sum_{\substack{\sigma_1=i \\ |\sigma|=n}} X(\sigma) \prod_{j=2}^n T_{\sigma_j}(\sigma|(j \Leftrightarrow 1)) \right) + C(\emptyset)
\end{aligned}$$

Hierbei sind die Zufallsgrößen $W_i^F := C(i) + \sum_{k=2}^{n-1} \sum_{\substack{\sigma_1=i \\ |\sigma|=k}} C(\sigma) \prod_{j=2}^k T_{\sigma_j}(\sigma|(j \Leftrightarrow 1)) + \sum_{\substack{\sigma_1=i \\ |\sigma|=n}} X(\sigma) \prod_{j=2}^n T_{\sigma_j}(\sigma|(j \Leftrightarrow 1))$, $i \in \mathbb{N}$, unabhängig, identisch verteilt mit $W_i^F \sim Y_{n-1}^F$. Es gilt nun folgender

2.3.9 Satz *Es gelte*

$$F \in \mathcal{M}_2, \quad E\left(\sum_i T_i^2\right) < 1, \quad \left\| \sum_i T_i \right\|_2 < \infty, \quad \left| E\left(\sum_i T_i\right) \right| < 1, \quad \|C\|_2 < \infty.$$

Dann konvergiert die Folge der Zufallsgrößen Y_n^F fast sicher und in d_2 -Metrik gegen die Zufallsgröße

$$Y = C(\emptyset) + \sum_{k=1}^{\infty} \sum_{\substack{\sigma \in I \\ |\sigma|=k}} C(\sigma) \prod_{j=1}^k T_{\sigma_j}(\sigma|(j \Leftrightarrow 1)),$$

wobei die Zufallsvariablen $(C(\sigma), T(\sigma))$, $\sigma \in I$ und (C, T) die gleiche Verteilung besitzen. Die Verteilung dieser Zufallsgröße ist der eindeutige Fixpunkt der Abbildung $S : \mathcal{M}_2 \rightarrow \mathcal{M}_2$. Der Prozeß

$$\sum_{\substack{\sigma \in I \\ |\sigma|=n}} X(\sigma) \prod_{j=1}^n T_{\sigma_j}(\sigma|(j \Leftrightarrow 1)), \quad n \in \mathbb{N},$$

konvergiert fast sicher gegen Null. Falls F die Verteilungsfunktion der Dirac-Verteilung in $b := E(C)(1 \Leftrightarrow E(\sum_i T_i))^{-1}$ ist, bildet $Y_n^F, n \in \mathbb{N}$, ein \mathcal{L}_2 -beschränktes Martingal bezüglich der Filtration $(\mathcal{F}_n)_{n \in \mathbb{N}}$, wobei \mathcal{F}_n durch die Zufallsgrößen $C(\sigma)$, $T(\sigma)$, $|\sigma| < n$ und $X(\sigma)$, $\sigma \in I$, $|\sigma| \leq n$ erzeugt wird.

BEWEIS: Wir betrachten zunächst den Spezialfall der Dirac-Verteilung in $b := E(C)(1 \Leftrightarrow E(\sum_i T_i))^{-1}$ und zeigen die behauptete Martingaleigenschaft. Dazu bezeichne F_b die Verteilungsfunktion der Dirac-Verteilung in b . Da $L(Y_n^{F_b}) = S^n(F_b) \in \mathcal{M}_2$, folgt insbesondere

die Integrierbarkeit der Zufallsgröße $Y_n^{F_b}$ für alle $n \in \mathbb{N}$. Weiterhin können wir unter Ausnutzung der Definition der Filtration $(\mathcal{F}_n)_{n \in \mathbb{N}}$ und der Unabhängigkeitseigenschaften der Zufallsvariablen $(C(\sigma), T(\sigma)), X(\sigma)$, $\sigma \in I$, festhalten:

$$(2.3) \quad \begin{aligned} E(Y_{n+1}^{F_b} | \mathcal{F}_n) &= C(\emptyset) + \sum_{k=1}^{n-1} \sum_{\substack{\sigma \in I \\ |\sigma|=k}} C(\sigma) \prod_{j=1}^k T_{\sigma_j}(\sigma | (j \Leftrightarrow 1)) \\ &+ \sum_{\substack{\sigma \in I \\ |\sigma|=n}} E(C(\sigma)) \prod_{j=1}^n T_{\sigma_j}(\sigma | (j \Leftrightarrow 1)) \\ &+ \sum_{\substack{\sigma \in I \\ |\sigma|=n+1}} E(X(\sigma)) E(T_{\sigma_{n+1}}(\sigma | n)) \prod_{j=1}^n T_{\sigma_j}(\sigma | (j \Leftrightarrow 1)). \end{aligned}$$

Eine nähere Untersuchung der letzten beiden Summanden ergibt

$$\begin{aligned} &\sum_{\substack{\sigma \in I \\ |\sigma|=n}} E(C(\sigma)) \prod_{j=1}^n T_{\sigma_j}(\sigma | (j \Leftrightarrow 1)) + \sum_{\substack{\sigma \in I \\ |\sigma|=n+1}} E(X(\sigma)) E(T_{\sigma_{n+1}}(\sigma | n)) \prod_{j=1}^n T_{\sigma_j}(\sigma | (j \Leftrightarrow 1)) \\ &= \sum_{\substack{\sigma \in I \\ |\sigma|=n}} \left(E(C(\sigma)) \prod_{j=1}^n T_{\sigma_j}(\sigma | (j \Leftrightarrow 1)) + \sum_{i=1}^{\infty} E(X(\sigma_i)) E(T_i(\sigma)) \prod_{j=1}^n T_{\sigma_j}(\sigma | (j \Leftrightarrow 1)) \right) \\ &= \sum_{\substack{\sigma \in I \\ |\sigma|=n}} \left(E(C(\sigma)) \prod_{j=1}^n T_{\sigma_j}(\sigma | (j \Leftrightarrow 1)) + E(X(\sigma)) \prod_{j=1}^n T_{\sigma_j}(\sigma | (j \Leftrightarrow 1)) \sum_{i=1}^{\infty} E(T_i(\sigma)) \right) \\ &= \sum_{\substack{\sigma \in I \\ |\sigma|=n}} \left(\left(E(C) + E(C)(1 \Leftrightarrow E(\sum_i T_i))^{-1} E(\sum_{i=1}^{\infty} T_i) \right) \prod_{j=1}^n T_{\sigma_j}(\sigma | (j \Leftrightarrow 1)) \right) \\ &= \sum_{\substack{\sigma \in I \\ |\sigma|=n}} E(C) \left(1 + \frac{E(\sum_i T_i)}{1 \Leftrightarrow E(\sum_i T_i)} \right) \prod_{j=1}^n T_{\sigma_j}(\sigma | (j \Leftrightarrow 1)) \\ &= \sum_{\substack{\sigma \in I \\ |\sigma|=n}} E(C) (1 \Leftrightarrow E(\sum_i T_i))^{-1} \prod_{j=1}^n T_{\sigma_j}(\sigma | (j \Leftrightarrow 1)) \\ &= \sum_{\substack{\sigma \in I \\ |\sigma|=n}} X(\sigma) \prod_{j=1}^n T_{\sigma_j}(\sigma | (j \Leftrightarrow 1)) \quad \text{f.s.,} \end{aligned}$$

wobei wir die identische Verteilung der verwendeten Zufallsvariablen und $E(X(\sigma)) = E(C) \cdot (1 \Leftrightarrow E(\sum_i T_i))^{-1}$ ausgenutzt haben. Dies zusammen mit (2.3) liefert schließlich die Gleichheit:

$$E(Y_{n+1}^{F_b} | \mathcal{F}_n) = Y_n^{F_b}$$

und damit die behauptete Martingaleigenschaft. Da $L(Y_n^{F_b}) = S^n(F_b)$ und $S^n(F_b)$ nach Satz 2.3.6 einen d_2 -Limes in \mathcal{M}_2 besitzt (nämlich gerade den Fixpunkt der Abbildung S), folgt weiterhin $\lim_{n \rightarrow \infty} E((Y_n^{F_b})^2) = \lim_{n \rightarrow \infty} E(S^n(F_b))^2 < \infty$ (vgl. Korollar 2.2.10) und somit die \mathfrak{L}_2 -Beschränktheit des Martingals $Y_n^{F_b}$, $n \in \mathbb{N}$. Folglich existiert eine Zufallsgröße, gegen die die Folge $Y_n^{F_b}$ fast sicher und in \mathfrak{L}_2 konvergiert [\Leftrightarrow [1]].

Es verbleibt die nähere Bestimmung dieser Zufallsgröße. Betrachte dazu die Folge

$$Z_n^F := \sum_{\substack{\sigma \in I \\ |\sigma|=n}} X(\sigma) \prod_{j=1}^n T_{\sigma_j}(\sigma|(j \Leftrightarrow 1)),$$

$n \in \mathbb{N}$, für allgemeines $F \in \mathcal{M}_2$. Die folgenden Ausführungen werden zeigen, daß diese Folge fast sicher gegen Null konvergiert. Wie schon die Folge $L(Y_n^F)$, so hat auch $L(Z_n^F)$ eine rekursive Struktur. Diese wird gegeben durch

$$L(Z_n^F) = \bar{S}^n(F), \quad \bar{S} : \mathcal{M}_2 \rightarrow \mathcal{M}_2,$$

mit $\bar{S}(F) = L(\sum_i T_i V_i)$, $L(V_i) = F$ und T, V_i , $i \in \mathbb{N}$, unabhängig. Damit ergibt sich in bekannter Argumentationsweise:

$$(2.4) \quad E(\bar{S}^n(F)) = E\left(\sum_i T_i\right) E(\bar{S}^{n-1}(F)) = \left(E\left(\sum_i T_i\right)\right)^n E(F).$$

Da $|E(\sum_i T_i)| < 1$, konvergiert dieser Term exponentiell schnell gegen Null. Eine analoge Aussage gilt für die Varianz von $\bar{S}^n(F)$. Dazu seien V_i , $i \in \mathbb{N}$, identisch verteilte Zufallsgrößen mit $L(V_i) = \bar{S}^{n-1}(F)$ und T, V_i , $i \in \mathbb{N}$, unabhängig:

$$\begin{aligned} \text{Var}(\bar{S}^n(F)) &= \text{Var}\left(\sum_i T_i V_i\right) = E\left(\sum_i T_i V_i\right)^2 \Leftrightarrow \left(E\left(\sum_i T_i V_i\right)\right)^2 \\ &= \lim_{n \rightarrow \infty} \left(E(V_1)^2 \cdot E\left(\sum_{i=1}^n T_i^2\right) + E\left(\sum_{\substack{j,k=1 \\ j \neq k}}^n T_j T_k V_j V_k\right) \right) \Leftrightarrow \left(E\left(\sum_i T_i V_i\right)\right)^2 \\ &= \lim_{n \rightarrow \infty} \left(E(V_1^2) E\left(\sum_{i=1}^n T_i^2\right) + (E(V_1))^2 E\left(\sum_{\substack{j,k=1 \\ j \neq k}}^n T_j T_k\right) \right) \Leftrightarrow \left(E\left(\sum_i T_i V_i\right)\right)^2 \\ &= E(V_1^2) E\left(\sum_i T_i^2\right) + (E(V_1))^2 \cdot \left(E\left(\sum_i T_i\right)^2 \Leftrightarrow E\left(\sum_i T_i^2\right) \right) \Leftrightarrow (E(V_1))^2 \left(E\left(\sum_i T_i\right) \right)^2 \\ &= \left(E(V_1^2) \Leftrightarrow (E(V_1))^2 \right) \cdot E\left(\sum_i T_i^2\right) + (E(V_1))^2 \cdot E\left(\sum_i T_i\right)^2 \Leftrightarrow (E(V_1))^2 \left(E\left(\sum_i T_i\right) \right)^2 \\ &= \text{Var}(\bar{S}^{n-1}(F)) \cdot E\left(\sum_i T_i^2\right) + (E(V_1))^2 \cdot \text{Var}\left(\sum_i T_i\right) \\ &= \text{Var}(\bar{S}^{n-1}(F)) \cdot E\left(\sum_i T_i^2\right) + (E(\bar{S}^{n-1}(F)))^2 \cdot \text{Var}\left(\sum_i T_i\right). \end{aligned}$$

Vermöge (2.4) folgt daraus leicht die exponentiell schnelle Konvergenz von $\text{Var}(\bar{S}^n(F))$ gegen Null.

Aufgrund der beiden letzten Konvergenzaussagen gilt nun für genügend großes n :

$$(1 \Leftrightarrow nE(Z_n^F) \Leftrightarrow n^2(\text{Var}(Z_n^F))^{1/2}) \cdot (n(1 \Leftrightarrow nE(Z_n^F)))^{-1} \geq 0$$

und somit

$$(\text{Var}(Z_n^F))^{1/2} \cdot (1/n \Leftrightarrow E(Z_n^F))^{-1} \leq \frac{1}{n}.$$

Dies zusammen mit der Tschebyscheffschen-Ungleichung ergibt:

$$P(|Z_n^F| > 1/n) \leq P(|Z_n^F \Leftrightarrow E(Z_n^F)| \geq 1/n \Leftrightarrow E(Z_n^F)) \leq \text{Var}(Z_n^F)(1/n \Leftrightarrow E(Z_n^F))^{-2} \leq 1/n^2.$$

und folglich die fast sichere Konvergenz von Z_n^F gegen Null aufgrund des Lemmas von Borell-Cantelli.

Es folgt nun offenkundig

$$(2.5) \quad Y_n^{F_b} \rightarrow Y \text{ f.s. und in } \mathcal{L}_2$$

sowie unter Beachtung der Tatsache, daß bei der Folge

$$\left(C(\emptyset) + \sum_{k=1}^{n-1} \sum_{\substack{\sigma \in I \\ |\sigma|=k}} C(\sigma) \prod_{j=1}^k T_{\sigma_j}(\sigma|(j \Leftrightarrow 1)) \right)_n$$

die spezielle Wahl von F nicht einfließt:

$$Y_n^F \xrightarrow{\text{f.s.}} Y$$

für alle $F \in \mathcal{M}_2$. Zu zeigen verbleibt also nur noch, daß diese letzte Konvergenzaussage auch in der d_2 -Metrik gültig bleibt, und daß Y die Fixpunktverteilung besitzt. Dies ist aber nun einfach einzusehen. Nach Satz 2.3.6 besitzt die Abbildung S einen eindeutigen Fixpunkt gegen den die Folge $L(Y_n^F) = S^n(F)$ für jedes $F \in \mathcal{M}_2$ in d_2 -Metrik konvergiert. Andererseits folgt aus (2.5) nach Satz 2.2.12

$$Y_n^{F_b} \xrightarrow{d_2} Y.$$

Dann muß aber bereits $L(Y)$ der Verteilungsfunktion der Fixpunktverteilung entsprechen und allgemein

$$Y_n^F \xrightarrow{d_2} Y$$

gelten. □

2.3.10 Bemerkung Die Annahme $|E(\sum_i T_i)| < 1$ kann wiederum durch $E(\sum_i T_i) = 1$ und $E(C) = 0$ ersetzt werden; vgl. auch Satz 2.3.7.

2.3.2 (\mathcal{M}_p, d_p) als Definitions- und Wertebereich: Die Existenz von höheren Momenten des Fixpunktes

Wir gehen nun zu (\mathcal{M}_p, d_p) als Definitions- und Wertebereich der Abbildung S über. Der Vorteil liegt in dem Nachweis von Fixpunkten mit endlichen höheren Momenten.

Um die Existenz der Abbildung im \mathfrak{L}_p -Sinn zu gewährleisten, sind stärkere Forderungen an die Zufallsvariable T zu stellen. Gemäß Abschnitt 2.1 ist dafür etwa $\sum_i |T_i| \in \mathfrak{L}_p$ hinreichend.

Im nachfolgenden Satz erweist sich die Abbildung S wiederum als Kontraktion, was eine Anwendung des Banachschen Fixpunktsatzes ermöglicht.

2.3.11 Satz *Es gelte für ein $1 \leq p$:*

$$\left\| \sum_i |T_i| \right\|_p < 1, \quad \|C\|_p < \infty.$$

Dann existiert die Abbildung $S : \mathcal{M}_p \rightarrow \mathcal{M}_p$ im \mathfrak{L}_p -Sinn und besitzt einen eindeutigen Fixpunkt. Die Folge $S^n(F)$, $n \in \mathbb{N}$, mit $F \in \mathcal{M}_p$, konvergiert in d_p -Metrik exponentiell schnell gegen diesen Fixpunkt.

BEWEIS: Seien $X_i, i \in \mathbb{N}$ Zufallsvariablen mit $L(X_i) = F \in \mathcal{M}_p$, die unabhängig untereinander und von dem Vektor T sind. Mit den Bezeichnungen $t := \sum_i |T_i|$, $\alpha := \left\| \sum_i |T_i| \right\|_p$ erhält man, wenn man berücksichtigt, daß mit $\left\| \sum_i |T_i| \right\|_p < 1$ insbesondere $\sum_i |T_i| < \infty$ fast sicher:

$$\begin{aligned} E\left(\sum_i |T_i X_i|\right)^p &= E\left(E\left(\left(\sum_i \frac{|T_i|}{t} t |X_i|\right)^p \mid T\right)\right) = E_\omega\left(E_{\omega_1}\left(\sum_i \frac{|T_i(\omega)|}{t} t |X_i(\omega_1)|^p\right)\right) \\ &\leq E\left(E\left(\sum_i |T_i| t^{p-1} |X_i|^p \mid T\right)\right) = E(|X_1|^p) E\left(\left(\sum_i |T_i|\right)^p\right) \\ &= \alpha^p E(|X_1|^p). \end{aligned}$$

Angewendet wurde dabei die Jensensche Ungleichung auf den Ausdruck

$$\left(\sum_i \frac{|T_i(\omega)|}{t} t |X_i(\omega_1)|\right)^p,$$

dessen innerer Term nichts anderes darstellt als den Erwartungswert einer diskreten Verteilung.

Sind nun insbesondere $(C, T), (X_i, Y_i), i \in \mathbb{N}$ unabhängige Zufallsvariable mit $L(X_i) = F$, $L(Y_i) = G$ und $\|X_i \Leftrightarrow Y_i\|_p = d_p(F, G)$ (vgl. Satz 2.2.1), so folgt mit obiger Ungleichung:

$$\begin{aligned} d_p^p(S(F), S(G)) &\leq \left\| \sum_i T_i(X_i \Leftrightarrow Y_i) \right\|_p^p \\ &\leq E\left(\left(\sum_i |T_i(X_i \Leftrightarrow Y_i)|\right)^p\right) \leq \alpha^p \cdot E(|X_1 \Leftrightarrow Y_1|^p) = \alpha^p \cdot d_p^p(F, G) \end{aligned}$$

für $F, G \in \mathcal{M}_p$, d.h. die Abbildung S ist eine α -Kontraktion mit $\alpha = \|\sum_i |T_i|\|_p < 1$. Die Behauptungen folgen nun wiederum nach dem Banachschen Fixpunktsatz 1.2 und Korollar 1.4. \square

Der vorherige Satz ist noch insofern unbefriedigend, als er die sehr restriktive Voraussetzung $\|\sum_i |T_i|\|_p < 1$ benutzt; vgl. Bemerkung am Ende dieses Abschnitts. In den beiden folgenden Sätzen sollen deshalb ähnliche Resultate unter schwächeren Voraussetzungen bewiesen werden. Dazu sind die beiden folgenden Lemmata hilfreich [\heartsuit Beweis zu Satz 14.1 in [4]].

2.3.12 Lemma Gegeben $p, q > 1$ mit $1/p + 1/q = 1$, gilt folgende Ungleichung:

$$x^{\frac{1}{p}} \cdot y^{\frac{1}{q}} \leq \frac{1}{p}x + \frac{1}{q}y \quad \text{für alle } x, y \in \mathbb{R}^+.$$

BEWEIS: Betrachte die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto (1+x)^{\frac{1}{p}}$. Anwendung des Mittelwertsatzes der Differentialrechnung auf diese Funktion ergibt für $\eta \in \mathbb{R}^+ \Leftrightarrow \{0\}$:

$$\frac{f(\eta) - f(0)}{\eta - 0} = f'(\xi) \quad \text{mit } \xi \in (0, \eta)$$

und folglich

$$\frac{(1+\eta)^{\frac{1}{p}} - 1}{\eta} = \frac{1}{p}(1+\xi)^{-\frac{1}{q}} \leq \frac{1}{p} \quad \text{für } \xi \in (0, \eta).$$

Für alle $\eta \in \mathbb{R}^+$ gilt also die Ungleichung:

$$(1+\eta)^{\frac{1}{p}} \leq \frac{\eta}{p} + 1 \quad \text{bzw.} \quad \zeta^{\frac{1}{p}} \leq \frac{\zeta}{p} + \frac{1}{q} \quad \text{für alle } \zeta \geq 1.$$

Sind nun $x > 0$ und $y > 0$ reelle Zahlen, so ist xy^{-1} oder $x^{-1}y \geq 1$, also ein derartiges ζ . Einsetzen dieses ζ in die letzte Ungleichung (im zweiten Fall nach Vertauschung von p und q) liefert

$$x^{\frac{1}{p}} \cdot y^{\frac{1}{q}} \leq \frac{1}{p}x + \frac{1}{q}y,$$

wobei diese Ungleichung offenbar auch noch für alle $x, y \in \mathbb{R}^+$ Gültigkeit besitzt. \square

2.3.13 Lemma Gegeben sei eine Folge reeller Zahlen $(a_n)_n$. Dann gilt für $1 \leq p \leq q$:

$$\left(\sum_k |a_k|^q\right)^{1/q} \leq \left(\sum_k |a_k|^p\right)^{1/p}.$$

BEWEIS: Wir zeigen die Behauptung für den Spezialfall endlich vieler Summanden. Der allgemeine Fall ergibt sich dann durch Grenzübergang. Seien p und q gegeben und $1 \leq p \leq q$. Betrachte die Funktionen

$$f_k(j) := \begin{cases} |a_k|^p & ; j = k \\ 0 & ; j \neq k \end{cases}.$$

Es gilt vermöge der Minkowskischen Ungleichung

$$\begin{aligned} \left(\sum_{k=1}^n |a_k|^q \right)^{\frac{p}{q}} &= \left(\sum_{k=1}^n (|a_k|^p)^{\frac{q}{p}} \right)^{\frac{p}{q}} = \left\| \sum_{k=1}^n f_k \right\|_{\frac{q}{p}} \leq \sum_{k=1}^n \|f_k\|_{\frac{q}{p}} \\ &= \sum_{k=1}^n \left((|a_k|^p)^{\frac{q}{p}} \right)^{\frac{p}{q}} = \sum_{k=1}^n |a_k|^p. \end{aligned}$$

$\|\cdot\|_{\frac{q}{p}}$ bezeichne dabei die $\mathfrak{L}_{\frac{q}{p}}$ -Norm bezüglich des Zählmaßes auf \mathbb{N} . Ziehen der p -ten Wurzel auf beiden Seiten liefert die Behauptung. \square

2.3.14 Satz *Es gelte für ein $2 \leq p \in \mathbb{N}$:*

$$\begin{aligned} E\left(\sum_i T_i^2\right) < 1, \quad E\left(\sum_i |T_i|^p\right) < 1, \quad \left\| \sum_i T_i^2 \right\|_{p/2} < \infty, \quad \left\| \sum_i |T_i| \right\|_p < \infty, \\ E(C) = 0, \quad \|C\|_p < \infty. \end{aligned}$$

Dann konvergiert die Folge $S^m(F)$, $m \in \mathbb{N}$, für jedes $F \in \mathcal{M}_p^0$ in d_q -Metrik, $1 \leq q \leq p$, exponentiell schnell gegen den eindeutigen Fixpunkt der Abbildung $S : \mathcal{M}_p^0 \rightarrow \mathcal{M}_p^0$.

BEWEIS: Aufgrund der Voraussetzungen gilt für $q \in [2, p]$

$$E\left(\sum_i |T_i|^q\right) \leq \max \left\{ E\left(\sum_i T_i^2\right), E\left(\sum_i |T_i|^p\right) \right\} < 1.$$

Dazu betrachte $p' := \frac{p-2}{p-q}$, $q' := \frac{p-2}{q-2}$, falls $q \in (2, p)$. Es gilt: $1/p' + 1/q' = 1$. Folglich liefert eine Anwendung von Lemma 2.3.12 für $a \geq 0$ die Ungleichung:

$$\begin{aligned} a^q &= a^{\frac{2(p-q)}{p-2}} \cdot a^{\frac{p(q-2)}{p-2}} = (a^2)^{\frac{p-q}{p-2}} \cdot (a^p)^{\frac{q-2}{p-2}} \leq a^2 \cdot \frac{p \Leftrightarrow q}{p \Leftrightarrow 2} + a^p \cdot \frac{q \Leftrightarrow 2}{p \Leftrightarrow 2} \\ &= a^2 \cdot \left(1 \Leftrightarrow \frac{q \Leftrightarrow 2}{p \Leftrightarrow 2}\right) + a^p \cdot \frac{q \Leftrightarrow 2}{p \Leftrightarrow 2}. \end{aligned}$$

Im Fall $q = 2$ bzw. $q = p$ ist diese Ungleichung trivialerweise gültig. Summieren und Bilden des Erwartungswertes ergibt schließlich:

$$\begin{aligned} E\left(\sum_i |T_i|^q\right) &\leq E\left(\sum_i T_i^2\right) \left(1 \Leftrightarrow \frac{q \Leftrightarrow 2}{p \Leftrightarrow 2}\right) + E\left(\sum_i |T_i|^p\right) \frac{q \Leftrightarrow 2}{p \Leftrightarrow 2} \\ &\leq \max \left\{ E\left(\sum_i T_i^2\right), E\left(\sum_i |T_i|^p\right) \right\} < 1. \end{aligned}$$

Es wird nun zunächst folgende Aussage bewiesen: Für $n \in \mathbb{N}$ mit $2 \leq n \leq p$ und $F \in \mathcal{M}_p^0$ existieren Konstanten $\text{const}, \lambda > 0$, so daß:

$$d_\alpha^\alpha(S^{m+1}(F), S^m(F)) \leq \text{const} \cdot e^{-\lambda m}$$

für $\alpha \in \{1, 2, \dots, n\}$. λ und const sind dabei unabhängig von m .

Der Beweis erfolgt durch Induktion nach n .

Induktionsanfang $n = 1$: Anwendung von Lemma 2.3.4, (1), liefert:

$$d_2^2(S^{m+1}(F), S^m(F)) \leq E\left(\sum_i T_i^2\right) \cdot d_2^2(S^m(F), S^{m-1}(F)) \leq \left(E\left(\sum_i T_i^2\right)\right)^m \cdot d_2^2(S(F), F).$$

und folglich

$$d_1(S^{m+1}(F), S^m(F)) \leq d_2(S^{m+1}(F), S^m(F)) \leq \text{const} \cdot e^{-\lambda m},$$

mit $\text{const} := d_2(S(F), F)$ und $\lambda := \Leftrightarrow \ln\left(E\left(\sum_i T_i^2\right)\right)^{1/2}$ (der Fall $E\left(\sum_i T_i^2\right) = 0$ ist trivial).

Induktionsschritt $n \rightarrow n + 1$: Für $n < p$ gelte die behauptete Ungleichung. Seien nun $U_i, i \in \mathbb{N}$, unabhängige $\mathcal{R}(0, 1)$ -verteilte Zufallsgrößen, die auch von T unabhängig sind. Definiere:

$$Y_{i,m} := (S^m(F))^{-1}(U_i) \Leftrightarrow (S^{m-1}(F))^{-1}(U_i), \quad m = 1, 2, \dots$$

Dann ist $d_q(S^m(F), S^{m-1}(F)) = \|Y_{i,m}\|_q$ für $1 \leq q \leq n + 1$, und für $m \geq 1$ gilt die Abschätzung:

$$\begin{aligned} & d_{n+1}^{n+1}(S^{m+1}(F), S^m(F)) \\ & \leq E\left(\left|\sum_i T_i Y_{i,m}\right|^{n+1}\right) = \lim_{k \rightarrow \infty} E\left(\left|\sum_{i=1}^k T_i Y_{i,m}\right|^{n+1}\right) \leq \lim_{k \rightarrow \infty} E\left(\sum_{i=1}^k |T_i| |Y_{i,m}|^{n+1}\right) \\ & = \lim_{k \rightarrow \infty} \left(E\left(\sum_{i=1}^k |T_i|^{n+1} |Y_{i,m}|^{n+1}\right) + E\left(\sum_{\substack{0 \leq \alpha_1, \dots, \alpha_k \leq n \\ \sum \alpha_i = n+1}} \frac{(n+1)!}{\alpha_1! \cdots \alpha_k!} |T_1|^{\alpha_1} |Y_{1,m}|^{\alpha_1} \cdots |T_k|^{\alpha_k} |Y_{k,m}|^{\alpha_k}\right) \right) \\ & = \lim_{k \rightarrow \infty} \left(E(|Y_{1,m}|^{n+1}) \cdot E\left(\sum_{i=1}^k |T_i|^{n+1}\right) \right. \\ & \quad \left. + E\left(\sum_{l=2}^{n+1} \sum_{\substack{1 \leq \alpha_1, \dots, \alpha_l \leq n \\ \sum \alpha_i = n+1}} \sum_{\substack{(i_1, \dots, i_l) \in \{1, \dots, k\}^l \\ i_1 < \dots < i_l}} \frac{(n+1)!}{\alpha_1! \cdots \alpha_l!} |T_{i_1}|^{\alpha_1} |Y_{i_1,m}|^{\alpha_1} \cdots |T_{i_l}|^{\alpha_l} |Y_{i_l,m}|^{\alpha_l}\right) \right) \\ & = \lim_{k \rightarrow \infty} \left(d_{n+1}^{n+1}(S^m(F), S^{m-1}(F)) \cdot E\left(\sum_{i=1}^k |T_i|^{n+1}\right) \right. \\ & \quad \left. + \sum_{l=2}^{n+1} \sum_{\substack{1 \leq \alpha_1, \dots, \alpha_l \leq n \\ \sum \alpha_i = n+1}} \sum_{\substack{(i_1, \dots, i_l) \in \{1, \dots, k\}^l \\ i_1 < \dots < i_l}} \frac{(n+1)!}{\alpha_1! \cdots \alpha_l!} E\left(\prod_{j=1}^l |T_{i_j}|^{\alpha_j}\right) \cdot \prod_{j=1}^l d_{\alpha_j}^{\alpha_j}(S^m(F), S^{m-1}(F)) \right) \end{aligned}$$

$$\begin{aligned}
& \stackrel{I.V.}{\leq} d_{n+1}^{n+1}(S^m(F), S^{m-1}(F)) \cdot E\left(\sum_i |T_i|^{n+1}\right) \\
& + \lim_{k \rightarrow \infty} \left(\sum_{l=2}^n \sum_{\substack{1 \leq \alpha_1, \dots, \alpha_l \leq n \\ \sum \alpha_i = n+1}} \sum_{\substack{(i_1, \dots, i_l) \in \{1, \dots, k\}^l \\ i_1 < \dots < i_l}} \frac{(n+1)!}{\alpha_1! \cdots \alpha_l!} \cdot E\left(\prod_{j=1}^l |T_{i_j}|^{\alpha_j}\right) \cdot c_1^n e^{-\lambda \cdot (m-1)} \right)
\end{aligned}$$

mit $c_1 := \max\{\text{const}, 1\}$. Vermöge der verallgemeinerten Hölderschen Ungleichung (vgl. Satz A.5 im Anhang) erhalten wir zudem:

$$\begin{aligned}
\sum_{i \in \mathbb{N}^l} E\left(\prod_{j=1}^l |T_{i_j}|^{\alpha_j}\right) &= \sum_{i_1} \cdots \sum_{i_l} E\left(\prod_{j=1}^l |T_{i_j}|^{\alpha_j}\right) = E\left(\sum_{i_1} \cdots \sum_{i_l} \prod_{j=1}^l |T_{i_j}|^{\alpha_j}\right) \\
&= E\left(\sum_{i_1} |T_{i_1}|^{\alpha_1} \cdots \sum_{i_l} |T_{i_l}|^{\alpha_l}\right) = E\left(\prod_{j=1}^l \left(\sum_{i \in \mathbb{N}} |T_i|^{\alpha_j}\right)\right) \leq \prod_{j=1}^l \left\| \sum_i |T_i|^{\alpha_j} \right\|_{(n+1)/\alpha_j}.
\end{aligned}$$

Da $\left\| \sum_i |T_i| \right\|_{n+1} < \infty$, sei somit im folgenden direkt $\alpha_j \geq 2$ für $j = 1, \dots, l$ angenommen. Dann liefert Lemma 2.3.13 unter Beachtung von $\sum_i |T_i|^2 < \infty$ fast sicher für alle $2 \leq q \leq p$:

$$\sum_{i \in \mathbb{N}^l} E\left(\prod_{j=1}^l |T_{i_j}|^{\alpha_j}\right) \leq \prod_{j=1}^l \left\| \sum_i T_i^2 \right\|_{(n+1)/2}^{\alpha_j/2} = \left\| \sum_i T_i^2 \right\|_{(n+1)/2}^{(n+1)/2} < \infty.$$

Folglich existiert eine von m unabhängige Konstante c_2 , so daß

$$d_{n+1}^{n+1}(S^{m+1}(F), S^m(F)) \leq d_{n+1}^{n+1}(S^m(F), S^{m-1}(F)) \cdot E\left(\sum_i |T_i|^{n+1}\right) + c_2 \cdot c_1^n \cdot e^{-\lambda \cdot (m-1)}.$$

Analog zum Beweis von Satz 2.3.6 ergibt sich schließlich:

$$d_{n+1}^{n+1}(S^{m+1}(F), S^m(F)) \leq c \cdot e^{-\lambda_1 \cdot m}$$

mit Konstanten c und $\lambda_1 > 0$ unabhängig von m und für $\text{const}_1 := \max\{\text{const}, c\}$ und $\lambda_2 := \min\{\lambda, \lambda_1\}$ gilt dann:

$$d_\alpha^\alpha(S^{m+1}(F), S^m(F)) \leq \text{const}_1 \cdot e^{-\lambda_2 m}$$

für alle $\alpha = 1, 2, \dots, n+1$. Damit ist die Induktion abgeschlossen und insbesondere die Gültigkeit der Aussage für $n = p$ gezeigt. Es ergibt sich in nunmehr bekannter Schlußweise die exponentiell schnelle Konvergenz der Folge $S^m(F)$ in d_p -Metrik und insbesondere auch die in d_q -Metrik für $1 \leq q \leq p$, wenn man $d_q \leq d_r$, $1 \leq q \leq r$, berücksichtigt.

Wegen $\left\| \sum_i |T_i| \right\|_p < \infty$ folgt:

$$d_p(S(F), S(G)) \leq \left\| \sum_i |T_i| \right\|_p \cdot d_p(F, G)$$

(siehe auch Beweis zu Satz 2.3.11), und damit die Stetigkeit der Abbildung S bezüglich der d_p -Metrik. Folglich ist der Limes der Folge $S^m(F)$ (bezüglich jeder d_q -Metrik $1 \leq q \leq p$) wirklich Fixpunkt von S . Sind $F, G \in \mathcal{M}_p^0$ zwei Fixpunkte der Abbildung S , so gilt:

$$d_2(F, G) = d_2(S(F), S(G)) \leq (E(\sum_i T_i^2))^{1/2} \cdot d_2(F, G)$$

mit $E(\sum_i T_i^2) < 1$ und somit $d_2(F, G) = 0$ bzw. $F = G$. □

2.3.15 Satz *Es gelte für ein $2 \leq p \in \mathbb{N}$:*

$$\begin{aligned} |E(\sum_i T_i)| < 1, \quad E(\sum_i T_i^2) < 1, \quad E(\sum_i |T_i|^p) < 1, \quad \left\| \sum_i T_i^2 \right\|_{p/2} < \infty, \\ \left\| \sum_i |T_i| \right\|_p < \infty, \quad \|C\|_p < \infty. \end{aligned}$$

Dann konvergiert die Folge $S^m(F)$, $m \in \mathbb{N}$, für jedes $F \in \mathcal{M}_p$ in d_q -Metrik, $1 \leq q \leq p$, exponentiell schnell gegen den eindeutigen Fixpunkt der Abbildung $S : \mathcal{M}_p \rightarrow \mathcal{M}_p$.

BEWEIS: Die Beweisidee ist die gleiche wie im vorherigen Satz. Der einzige Unterschied besteht im Beweis des Induktionsanfangs. Hier wird die zusätzliche Voraussetzung $|E(\sum_i T_i)| < 1$ benötigt, da allgemeiner Abbildungen $S : \mathcal{M}_p \rightarrow \mathcal{M}_p$ betrachtet werden (vergleiche dazu auch die Sätze 2.3.5 und 2.3.6). Analog zum Beweis von Satz 2.3.6 erhält man die Abschätzung:

$$d_2^2(S^{m+1}(F), S^m(F)) \leq c' \cdot \eta^m$$

mit Konstanten c' und $\eta < 1$ unabhängig von m . Daraus ergibt sich wie zuvor die entsprechende Abschätzung für $d_1(S^{m+1}(F), S^m(F))$ und somit der Nachweis des Induktionsanfangs. Der restliche Beweis erfolgt analog, wobei sich die Aussagen über die Eindeutigkeit des Fixpunktes wiederum wie in Satz 2.3.6 ergeben. □

2.3.16 Bemerkung An dieser Stelle ist es wichtig zu erwähnen, daß auch hier die Voraussetzung $|E(\sum_i T_i)| < 1$ wiederum ersetzt werden kann durch $E(\sum_i T_i) = 1$ und $E(C) = 0$. Der Beweis - mit kleinen Modifikationen innerhalb des Induktionsanfangs - vollzieht sich wie in den vorherigen Sätzen. Es sei dabei an Satz 2.3.7 aus Abschnitt 2.3.1 erinnert. Der gleiche Verweis gilt für den Sonderfall $\sum_i T_i \equiv 1$, für den sich die Fixpunkte als gegenseitige Translationen herausstellen.

2.3.17 Bemerkung Die beiden letzten Sätze beinhalten stärkere Aussagen als der erste Satz dieses Abschnitts, da sie mit schwächeren Voraussetzungen zum gleichen Ergebnis gelangen. Dies ergibt sich aus der Gültigkeit der folgenden Abschätzungen, wobei zur Abkürzung $l_p(T) := (\sum_i |T_i|^p)^{1/p}$ gesetzt sei:

- $E(\sum_i T_i^2) \leq E(\sum_i |T_i|)^2 = \|\sum_i |T_i|\|_2^2 \leq \|\sum_i |T_i|\|_p^2 < 1$
- $E(\sum_i |T_i|^p) = \|l_p(T)\|_p^p \leq \|l_1(T)\|_p^p = \|\sum_i |T_i|\|_p^p < 1,$
- $\left\| \sum_i T_i^2 \right\|_{p/2} = \left(E((\sum_i T_i^2)^{p/2}) \right)^{2/p} = \|l_2(T)\|_p^2 \leq \|l_1(T)\|_p^2 = \|\sum_i |T_i|\|_p^2 < 1,$
- $\left| E(\sum_i T_i) \right| \leq E(\sum_i |T_i|) = \|\sum_i |T_i|\|_1 \leq \|\sum_i |T_i|\|_p < 1.$

Für den Nachweis der obigen Abschätzungen sei an Lemma 2.3.13 und die Jensensche Ungleichung erinnert.

2.3.3 Exponentielle Momente des Fixpunktes

Gegenstand dieses letzten Abschnitts ist die Untersuchung von exponentiellen Momenten (momentenerzeugende Funktion) der iterativen Folge $(S^n(F))$ und der Fixpunktverteilung. Die Voraussetzung $\|\sum_i T_i\|_2 < \infty$ sichert dabei wiederum die Existenz der Abbildung S . Wir notieren das folgende Hauptresultat:

2.3.18 Satz *Es gelte*

$$\left\| \sum_i T_i \right\|_2 < \infty, \quad E\left(\sum_i T_i^2\right) < 1, \quad |T_i| \leq 1$$

für $i \in \mathbb{N}$. Außerdem sei für $\lambda \in \mathbb{R}$ in einer offenen Umgebung von 0

$$E(\exp(\lambda \sum_i T_i^2)) < \infty, \quad E(\exp(\lambda \sum_i |T_i|)) < \infty, \quad E(\exp(\lambda C)) < \infty$$

und entweder $E(\sum_i T_i) \neq 1$ oder $E(C) = 0$. Definiere b durch $E(C) = b(1 \Leftrightarrow E(\sum_i T_i))$, falls $E(\sum_i T_i) \neq 1$ und beliebig sonst. Sei $F \in \mathcal{M}_2$ eine Verteilungsfunktion mit

$$E(\exp(\lambda X)) \leq \exp(b\lambda + K\lambda^2)$$

in einer offenen Umgebung von 0, wobei K eine positive Konstante ist und $L(X)=F$. Dann existiert eine offene Umgebung von 0 und ein $K_0 \in \mathbb{R}$, so daß für alle λ in dieser Umgebung und alle Zufallsgrößen Z mit Verteilungsfunktion $S^n(F)$ für ein $n \in \mathbb{N}$ oder der Fixpunktverteilung gilt:

$$E(\exp(\lambda Z)) \leq \exp(b\lambda + K_0\lambda^2).$$

BEWEIS: Betrachte die Funktion

$$f_{K_0}(\lambda) = E(\exp(\lambda \cdot C + b\lambda \cdot (\sum_i T_i \Leftrightarrow 1) + K_0\lambda^2 \cdot (\sum_i T_i^2 \Leftrightarrow 1)))$$

für ein $K_0 > 0$. Für $|\lambda|$ klein genug erhält man vermöge des Differentiationslemmas [\Leftrightarrow [4], Lemma 16.2 auf S.102]

$$\begin{aligned} \frac{d}{d\lambda} f_{K_0}(\lambda) &= E\left(\left(C + b\left(\sum_i T_i \Leftrightarrow 1\right) + 2K_0\lambda\left(\sum_i T_i^2 \Leftrightarrow 1\right)\right)\right. \\ &\quad \left.\cdot \exp\left(\lambda C + b\lambda\left(\sum_i T_i \Leftrightarrow 1\right) + K_0\lambda^2 \cdot \left(\sum_i T_i^2 \Leftrightarrow 1\right)\right)\right) \end{aligned}$$

und zusammen mit der Definition von b

$$\frac{d}{d\lambda} f_{K_0}(0) = 0.$$

Analog ergibt sich als zweite Ableitung:

$$\begin{aligned} \frac{d^2}{d\lambda^2} f_{K_0}(\lambda) &= E\left(\left(\left(C + b\left(\sum_i T_i \Leftrightarrow 1\right) + 2\lambda K_0\left(\sum_i T_i^2 \Leftrightarrow 1\right)\right)^2\right)\right. \\ &\quad \left.\cdot \exp\left(\lambda C + b\lambda \cdot \left(\sum_i T_i \Leftrightarrow 1\right) + K_0\lambda^2 \cdot \left(\sum_i T_i^2 \Leftrightarrow 1\right)\right)\right) \end{aligned}$$

und

$$\frac{d^2}{d\lambda^2} f_{K_0}(0) = E\left(\left(C + b\left(\sum_i T_i \Leftrightarrow 1\right)\right)^2 + 2K_0\left(\sum_i T_i^2 \Leftrightarrow 1\right)\right).$$

Wähle nun K_0 so groß, daß $K_0 > K$ und $(d^2/d\lambda^2)f_{K_0}(0) < 0$. Dies ist aufgrund der Voraussetzung $E(\sum_i T_i^2) < 1$ möglich. Da $f_{K_0}(0) = 1$ erhält man folglich

$$(2.6) \quad f_{K_0}(\lambda) \leq 1$$

für alle λ in einer entsprechend gewählten Umgebung um 0.

Der eigentliche Beweis erfolgt nun durch eine Induktion nach n .

Induktionsanfang $n = 0$: Sei $Z \sim F$ und F wie im Satz angegeben. Dann gilt

$$E(\exp(\lambda Z)) \leq \exp(b\lambda + K\lambda^2) \leq \exp(b\lambda + K_0\lambda^2)$$

für alle λ aus einer geeignet gewählten Umgebung um Null. Man beachte $K_0 > K$.

Induktionsschritt $n \Leftrightarrow 1 \rightarrow n$: Sei $Z \sim S^n(F) = S(S^{n-1}(F))$. Aufgrund der Induktionsvoraussetzung ist $E(\exp(\lambda X)) \leq \exp(b\lambda + K_0\lambda^2)$ für alle λ aus der vorgegebenen Umgebung,

falls $X \sim S^{n-1}(F)$. Seien weiter (C, T) , X_i , $i \in \mathbb{N}$, unabhängig mit $L(X_i) = S^{n-1}(F)$. Da $S^{n-1}(F) \in \mathcal{M}_2$, ist $\sum_i T_i X_i \in \mathfrak{L}_2$ (vgl. Satz 2.1.1), und es existiert insbesondere eine fast sicher konvergente Teilfolge $(\sum_i^{N_n} T_i X_i)_n$. Vermöge des Lemmas von Fatou ergibt sich:

$$\begin{aligned} E(\exp(\lambda Z)) &= E(\exp(\lambda(\sum_i T_i X_i + C))) = E(\exp(\lambda(\lim_{n \rightarrow \infty} \sum_{i=1}^{N_n} T_i X_i + C))) \\ &\leq \liminf_{n \rightarrow \infty} E(\exp(\lambda(\sum_{i=1}^{N_n} T_i X_i + C))) \\ &= \liminf_{n \rightarrow \infty} E\left(\exp(\lambda C) \cdot E(\exp(\lambda \sum_{i=1}^{N_n} T_i X_i) \mid T, C)\right) \\ &= \liminf_{n \rightarrow \infty} E_\omega\left(\exp(\lambda C(\omega)) \cdot \prod_{i=1}^{N_n} E(\exp(\lambda T_i(\omega) X_i))\right) \end{aligned}$$

Im letzten Schritt wurde dabei die Unabhängigkeit der verwendeten Zufallsvariablen ausgenutzt. Da nun $|T_i| \leq 1$ und folglich mit $|\lambda| < \varepsilon$ auch $|\lambda T_i| = |\lambda| \cdot |T_i| < \varepsilon$, gilt für genügend kleine $|\lambda|$

$$\begin{aligned} E(\exp(\lambda Z)) &\leq \liminf_{n \rightarrow \infty} E_\omega\left(\exp(\lambda C(\omega)) \cdot \prod_{i=1}^{N_n} E(\exp(b\lambda T_i(\omega) + K_0 \lambda^2 T_i^2(\omega)))\right) \\ &= \liminf_{n \rightarrow \infty} E(\exp(\lambda C) \cdot \prod_{i=1}^{N_n} \exp(b\lambda T_i + K_0 \lambda^2 T_i^2)) \\ &= \liminf_{n \rightarrow \infty} E(\exp(\lambda C + b\lambda \cdot \sum_{i=1}^{N_n} T_i + K_0 \lambda^2 \cdot \sum_{i=1}^{N_n} T_i^2)) \\ &= E(\exp(\lambda C + b\lambda \cdot \sum_i T_i + K_0 \lambda^2 \cdot \sum_i T_i^2)) \\ &= \exp(b\lambda + K_0 \lambda^2) \cdot f_{K_0}(\lambda). \end{aligned}$$

Beim vorletzten Gleichheitszeichen wurde der Satz von der majorisierten Konvergenz verwendet. Dessen Anwendbarkeit begründet sich durch die Abschätzungen

$$\begin{aligned} \exp\left(\lambda C + b\lambda \sum_{i=1}^{N_n} T_i + K_0 \lambda^2 \sum_{i=1}^{N_n} T_i^2\right) &\leq \exp\left(\lambda C + |b||\lambda| \sum_{i=1}^{N_n} |T_i| + K_0 \lambda^2 \sum_{i=1}^{N_n} T_i^2\right) \\ &\leq \exp\left(\lambda C + |b||\lambda| \sum_i |T_i| + K_0 \lambda^2 \sum_i T_i^2\right) \end{aligned}$$

und

$$E\left(\exp\left(\lambda C + |b||\lambda| \cdot \sum_i |T_i| + K_0 \lambda^2 \cdot \sum_i T_i^2\right)\right)$$

$$\begin{aligned}
&\leq E\left(\exp(3\lambda C) + \exp\left(3|b|\lambda \sum_i |T_i|\right) + \exp\left(3K_0\lambda^2 \sum_i T_i^2\right)\right) \\
&= E(\exp(3\lambda C)) + E\left(\exp\left(2|b|\lambda \sum_i |T_i|\right)\right) + E\left(\exp\left(2K_0\lambda^2 \sum_i T_i^2\right)\right) < \infty,
\end{aligned}$$

wobei $|\lambda|$ klein genug zu wählen ist.

Vermöge (2.6) folgt nun

$$E(\exp(\lambda Z)) \leq \exp(b\lambda + K_0\lambda^2)$$

in einer Umgebung von $\lambda = 0$, d.h. die Behauptung, da die Wahl von λ in der obigen Rechnung nicht von n abhängt.

Zu zeigen bleibt, daß die Abschätzung auch für den Fixpunkt Gültigkeit besitzt, daß also die Abschätzung beim Übergang zum Grenzwert erhalten bleibt. Dies ergibt sich aber sofort aus einer Anwendung des Lemmas von Fatou, wenn man beachtet, daß aus der d_2 -Konvergenz die schwache Konvergenz folgt; vgl. Satz 2.2.8. \square

Eine Erweiterung des vorherigen Satzes beinhaltet

2.3.19 Satz *Zusätzlich zu den Voraussetzungen des vorherigen Satzes gelte für ein $L > 0$:*

$$E(\exp(3L|C|)) < \infty.$$

Außerdem sei $\sum_i T_i^2 \leq 1$ f.s., $P(\sum_i T_i^2 < 1) > 0$ und $E(\exp(3|bL \sum_i T_i|)) < \infty$, falls $E(\sum_i T_i) \neq 1$. Dann ist die Aussage des Satzes 2.3.18 richtig für alle $\lambda \in [\Leftrightarrow L, +L]$.

BEWEIS: Es reicht zu zeigen, daß $f_K(\lambda) \leq 1$ für alle $|\lambda| \leq L$ und K genügend groß. Dies weiß man nach Satz 2.3.18 bereits für alle $|\lambda| \leq \varepsilon$ (ε entsprechend gewählt). Für $|\lambda| \in [\varepsilon, L]$ liefert die erweiterte Hölder-Ungleichung (vgl. Satz A.5 im Anhang) folgende Abschätzung:

$$\begin{aligned}
f_K(\lambda) &\leq \|\exp(\lambda C)\|_3 \cdot \|\exp(\lambda^2 K (\sum_i T_i^2 \Leftrightarrow 1))\|_3 \cdot \|\exp(b\lambda (\sum_i T_i \Leftrightarrow 1))\|_3 \\
&\leq \|\exp(L|C|)\|_3 \cdot \|\exp(\varepsilon_0^2 K (\sum_i T_i^2 \Leftrightarrow 1))\|_3 \cdot \|\exp(|bL \sum_i T_i| + bL)\|_3.
\end{aligned}$$

Aufgrund der getroffenen Voraussetzungen konvergiert der zweite Term für K gegen unendlich gegen Null und somit $f_K(\lambda)$ ebenfalls für alle $|\lambda| \in [\varepsilon, L]$. Die Konvergenz ist dabei sogar gleichmäßig. Der Satz ist damit bewiesen. \square

2.3.20 Bemerkung In den beiden vorherigen Sätzen wurde für eine Zufallsgröße X mit Werten in \mathbb{R} die Bedingung $E(e^{\lambda X}) \leq e^{b\lambda + K\lambda^2}$ für λ aus einer Umgebung von Null gefordert.

Dies impliziert insbesondere die Existenz aller Momente von X . Es gibt nämlich offenkundig ein $\lambda > 0$, so daß

$$E(e^{\lambda|X|}) \leq E(e^{\lambda X}) + E(e^{-\lambda X}) < \infty,$$

und es folgt mit dem Satz von der monotonen Konvergenz

$$\sum_{j=1}^{\infty} \frac{E(|X|^j) \lambda^j}{j!} = E\left(\sum_{j=1}^{\infty} \frac{(\lambda \cdot |X|)^j}{j!}\right) = E(e^{\lambda|X|}) < \infty,$$

d.h. insbesondere die Endlichkeit aller absoluten Momente.

Kapitel 3

Anwendungen

In der Einleitung wurden bereits zwei Anwendungsbeispiele der zuvor entwickelten Theorie, nämlich der Galton-Watson-Verzweigungsprozesses und der Quicksort-Algorithmus, angesprochen. Im folgenden Kapitel sollen diese beiden Beispiele näher untersucht werden. Im Fall des Quicksort-Algorithmus werden wir zusätzlich demonstrieren, wie die Ergebnisse benutzt werden können, um eine Abschätzung für schlechtes Laufzeitverhalten dieses Algorithmus zu erhalten.

3.1 Der Galton-Watson-Verzweigungsprozeß

In der Einleitung wurde bereits auf die definierenden Eigenschaften eines Galton-Watson-Verzweigungsprozesses eingegangen. Man betrachte nun folgendes Modell: Es seien $I := \emptyset \cup \bigcup_{n \in \mathbb{N}} \mathbb{N}^n$ die Indexmenge und $T(\sigma) : \Omega \rightarrow \{0, 1\}^{\mathbb{N}}$, $\sigma \in I$, unabhängige identisch verteilte Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) . Definiere $l : \Omega \times I \rightarrow \{0, 1\}$ durch $l_\emptyset := 1$,

$$l_{\sigma i} := l_\sigma T_i(\sigma), \quad i \in \mathbb{N}, \sigma \in I \quad .$$

Die Notationen σi und $i\sigma$, $i \in \mathbb{N}$, $\sigma \in I$, haben dabei die offenkundigen Bedeutungen (vgl. hierzu auch Abschnitt 2.3.1). Anschaulich entspricht dies der „Beschriftung“ eines Baumes, der dadurch gekennzeichnet ist, daß von jedem Knoten (auch von der Wurzel) unendlich viele Zweige (Mächtigkeit von \mathbb{N}) ausgehen. Die Zweige des Knotens σ (\emptyset entspricht der Wurzel) erhalten dabei eine Kennzeichnung gemäß des Vektors $T(\sigma)$. Dies geschieht unabhängig voneinander und nach identischer Verteilung. Die Zufallsgrößen l_σ geben an, ob ein Ast (bei Betrachtung des Baumes bis zur $|\sigma|$ -ten Generation) nur aus Zweigen der Kennzeichnung Eins besteht ($l_\sigma=1$) oder nicht. Die Struktur und die Bezeichnungen verdeutlicht Abbildung 3.1

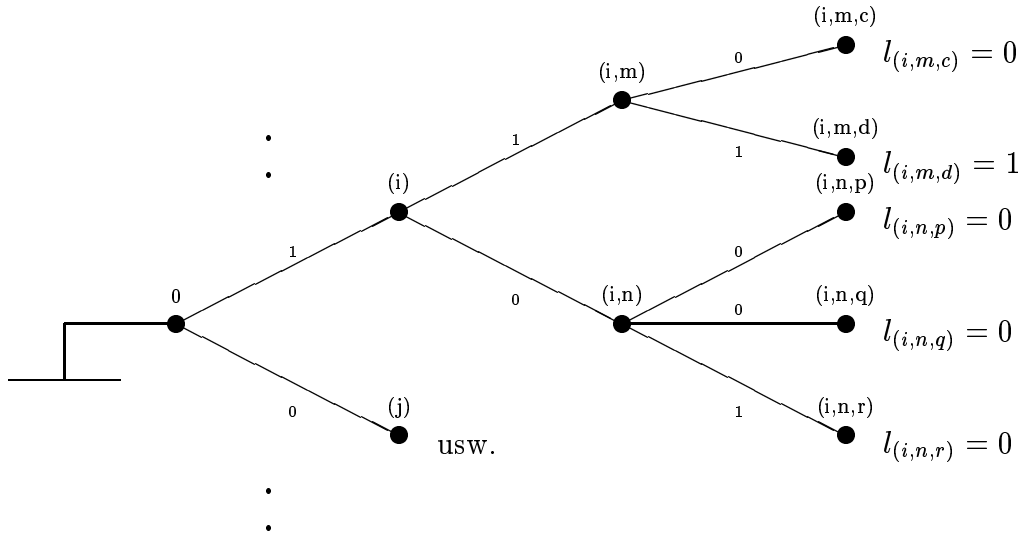


Abbildung 3.1: Das Diagramm stellt einen Auszug aus dem Baum der 3ten Generation dar. Dabei wird eine spezielle Realisierung betrachtet. Es gilt: $T_i(0)(\omega) = 1$, $T_j(0)(\omega) = 0$, $T_m(i)(\omega) = 1$, $T_n(i)(\omega) = 0$, $T_c((i, m))(\omega) = 0$, $T_d((i, m))(\omega) = 1$, $T_p, ((i, n))(\omega) = 0$ usw. Das einzige, in dem Baumausschnitt lebende Individuum der 3-ten Generation ist das zum Knoten (i, m, d) gehörige. Dies entspricht einem Wert von $l_{(i, m, d)}$ von eins, während alle anderen Zweige den Wert Null erhalten.

Nun ist es möglich, dies als Populationsmodell auszulegen, indem man folgende Interpretation wählt:

- $\sum_i T_i(\sigma)$: Anzahl der produzierten Individuen des Mitglieds σ der $|\sigma|$ -ten Generation
- $M_n := \sum_{\sigma \in \mathbb{N}^n} l_\sigma$ Gesamtanzahl der lebenden Individuen der n -ten Generation.

Aufgrund der Nichtnegativität aller vorkommenden Zufallsgrößen ist die Existenz der Zufallsgröße M_n gesichert. Es ist leicht einzusehen, daß der so definierte Prozeß $M = (M_n)_{n \geq 0}$ einen Galton-Watson-Verzweigungsprozeß mit Reproduktionsverteilung $p_k := P(\sum_i T_i = k)$ bildet (wenn man noch $P(\sum_i T_i < \infty) = 1$ fordert). Dies wird noch verdeutlicht durch folgende Umformung:

$$M_n = \sum_{\sigma \in \mathbb{N}^n} l_\sigma = \sum_{\sigma \in \mathbb{N}^{n-1}} \sum_{i \in \mathbb{N}} l_{\sigma i} = \sum_{\sigma \in \mathbb{N}^{n-1}} l_\sigma \sum_{i \in \mathbb{N}} T_i(\sigma).$$

Für M_n gilt die rekursive Beziehung:

$$\begin{aligned} L(M_{n+1}) &= L\left(\sum_{\sigma \in \mathbb{N}^{n+1}} l_\sigma\right) = L\left(\sum_{i \in \mathbb{N}} \sum_{\sigma \in \mathbb{N}^n} l_{i\sigma}\right) = L\left(\sum_{i \in \mathbb{N}} \sum_{\sigma \in \mathbb{N}^n} T_i(\emptyset) l_\sigma\right) \\ &= L\left(\sum_{i \in \mathbb{N}} T_i(\emptyset) \sum_{\sigma \in \mathbb{N}^n} l_\sigma\right) = L\left(\sum_{i \in \mathbb{N}} T_i(\emptyset) X_i\right), \end{aligned}$$

wobei X_i , $i \in \mathbb{N}$, unabhängige Zufallsgrößen bilden, die auch noch unabhängig von $T(\emptyset)$ und verteilt wie M_n sind. Beim Galton-Watson-Verzweigungsprozeß interessiert bei endlichem Reproduktionsmittel μ insbesondere das Verhalten der Zufallsgröße M_n/μ^n für n gegen unendlich und somit die Verteilung der Grenzzufallsgröße. Ziel ist es dabei, eine genauere Beschreibung des Wachstums von M_n im Überlebensfall zu erhalten. Definiere dazu eine Abbildung $S : \mathcal{M}_2 \rightarrow \mathcal{M}_2$ durch die Zuordnung:

$$S(F) = L\left(\sum_{i \in \mathbb{N}} T_i X_i\right),$$

wobei die Zufallsgrößen X_i , $i \in \mathbb{N}$, T unabhängig sind mit

$$T = T(\emptyset)/\mu = (T_1(\emptyset)/\mu, T_2(\emptyset)/\mu, \dots) \text{ und } X_i \sim F.$$

Um die Existenz der Abbildung im \mathcal{L}_2 -Sinn zu gewährleisten, ist die Bedingung

$$\left\| \sum_i T_i \right\|_2 < \infty$$

hinreichend; vgl. Korollar 2.1.2.

Bezeichnet nun F_1 die Verteilungsfunktion der Punktverteilung in 1, so gilt offenkundig:

$$S^n(F_1) = L(M_n/\mu^n)$$

und die Zufallsgröße M_n/μ^n entspricht der Zufallsgröße $Y_n^{F_1}$ aus Abschnitt 2.3.1.

Es sei weiterhin $1 < \mu < \infty$, d.h. wir betrachten den superkritischen Fall (nur dieser ist von Interesse, da für $\mu \leq 1$ die Population fast sicher ausstirbt [\Leftrightarrow [1]]). Dann gilt:

$$E\left(\sum_i T_i/\mu\right) = E\left(\sum_i T_i\right)/\mu = 1 \text{ und } E\left(\sum_i T_i^2/\mu^2\right) = E\left(\sum_i T_i\right)/\mu^2 < 1.$$

Somit sind alle Voraussetzungen für Satz 2.3.9 bzw. Bemerkung 2.3.10 erfüllt, und man erhält

$$M^n/\mu^n \xrightarrow[\text{f.s.}]{d_2\text{-Metrik}} W,$$

wobei $L(W)$ ein Fixpunkt der Abbildung S ist.

Das einzige, was sich innerhalb dieser Betrachtungen als störend erweist, ist die wenig anschauliche, die Existenz der Abbildung sichernde Forderung $\sum_i T_i \in \mathcal{L}_2$. Dazu sei folgendes angemerkt: Hat man $E((\sum_i T_i)^2) < \infty$ gegeben, d.h. die Endlichkeit der Varianz der Reproduktionsverteilung, wobei der auftretende Limes zunächst fast sicher zu verstehen ist und dessen Existenz aufgrund der Nichtnegativität der Zufallsgrößen T_i gesichert

ist, so ergibt sich automatisch $\sum_i T_i \in \mathcal{L}_2$. Aufgrund der in der Modellierung festgelegten Nichtnegativität der Zufallsgrößen T_i erhält man nämlich die für alle $M \geq N$ gültige Abschätzung:

$$E\left(\left(\sum_{i=N}^M T_i\right)^2\right) \leq E\left(\left(\sum_{i=N}^{\infty} T_i\right)^2\right).$$

Dabei konvergiert die rechte Seite für N gegen unendlich gegen Null, da nach Voraussetzung $E\left(\left(\sum_i T_i\right)^2\right) < \infty$. Aus Monotoniegründen ergibt sich zudem für alle n $E\left(\left(\sum_{i=1}^n T_i\right)^2\right) < \infty$ und die Folge $\sum_{i=1}^n T_i$ bildet somit eine Cauchy-Folge in $(\mathcal{L}_2, \|\cdot\|_2)$. Folglich ist $\sum_i T_i \in \mathcal{L}_2$.

3.2 Quicksort

3.2.1 Einführung

Quicksort ist wahrscheinlich der am häufigsten benutzte Sortieralgorithmus und wurde 1961/62 von C.A.R. Hoare erfunden. In Unix Systemen stellt er den Standard-Sortieralgorithmus dar. Eine kurze Zusammenfassung der Arbeitsweise von Quicksort möge als Basis für weitere Diskussionen dienen: Hat man eine Liste $a(j)$, $1 \leq j \leq n$, von n Elementen (Realzahlen) vorgegeben, die sortiert werden soll, so wird beim Algorithmus Quicksort zunächst ein Element x (Partitionselement) dieser Liste ausgewählt. Dann werden die Elemente so angeordnet, daß man eine Unterteilung in drei Teile erhält:

- a) ein Teilungselement x ,
- b) eine vordere Liste, die aus allen Elementen besteht, die kleiner oder gleich x sind,
- c) eine hintere Liste, deren Elemente alle größer als x sind.

Während sich das Element x nach diesem Vorgang bereits an seiner entsprechenden Position befindet und keiner weiteren Umstellung mehr bedarf, müssen die beiden verbleibenden Listen keinesfalls bereits in sich geordnet sein. Vielmehr wird nun mit diesen beiden Listen entsprechend verfahren, bis man zu entsprechenden Listen mit nur noch einem Element gelangt. Die Vorgehensweise des Algorithmus verdeutlicht die folgende Abbildung:

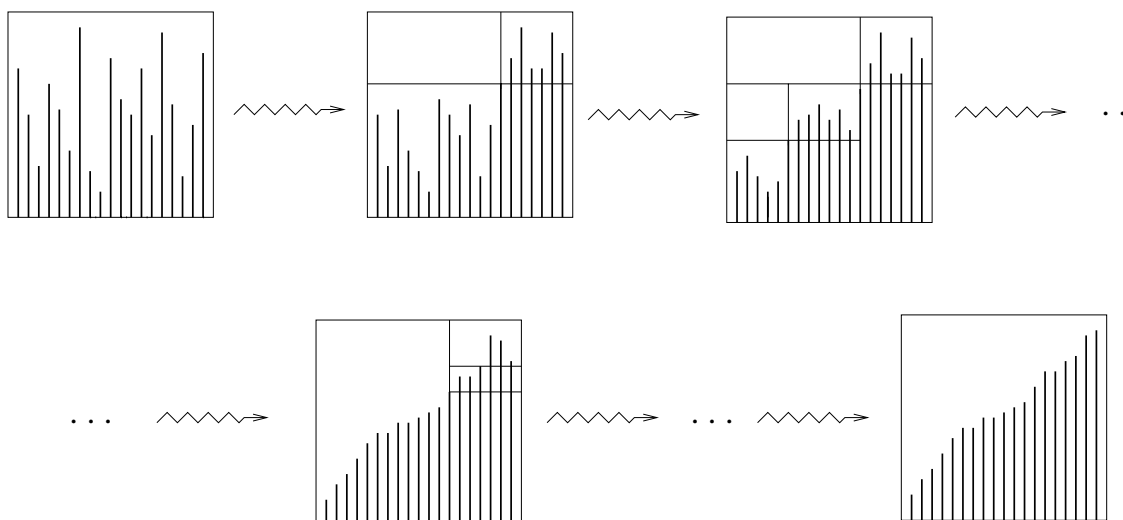


Abbildung 3.2: Graphische Darstellung der Vorgehensweise des Algorithmus Quicksort anhand einer Liste mit 19 Elementen. Die Höhe der Balken entspricht der Größe der Zahlen.

Die oben erwähnte Auswahl eines Elementes x aus der vorgegebenen Liste kann dabei nach verschiedenen Vorgaben erfolgen. Als erste Möglichkeit bietet sich ein deterministisches Auswahlkriterium an. So wäre zum Beispiel die Vorgabe, in jedem Partitionierungsschritt das erste Element der gegebenen Liste auszuwählen, solch ein Kriterium. Eine weitere Möglichkeit besteht in der Auswahl nach dem Zufallsprinzip: Jedes Element wird mit der gleichen Wahrscheinlichkeit als Partitionierungselement genommen. Welches Auswahlkriterium unter dem Gesichtspunkt der Schnelligkeit vorzuziehen ist, hängt von der Art der Listen mit denen in der Realität gearbeitet wird ab.

Das Hauptinteresse bei einem solchen Sortiervorgang gilt der Gesamtzahl aller benötigten Vergleiche zweier Elemente. Alle anderen Operationen seien im folgenden vernachlässigt. Man nimmt also eine Proportionalität von benötigter Zeit und Anzahl der Vergleichsoperationen an.

Das Ziel besteht darin, eine Aussage über die Verlässlichkeit des Algorithmus Quicksort zu machen, also etwas über die Abweichung vom Erwartungswert auszusagen. Um dabei unsere Theorie über die Fixpunktsätze von Verteilungen auszunutzen, ist es zunächst erforderlich eine rekursive Beziehung für die Verteilung der Zufallsgröße X_n (Anzahl der benötigten Vergleiche zum Sortieren einer Liste mit n Elementen) herzuleiten, um einen Hinweis auf die zu betrachtende Abbildung zu erlangen. Wir wollen uns dabei im Rahmen dieser Arbeit mit der deterministischen Version des Quicksort-Algorithmus beschäftigen. Orientierungspunkt ist ein Artikel von Uwe Rösler zu diesem Thema aus dem Jahre 1991 [↔ [20]].

3.2.2 Modellierung und Herleitung einer rekursiven Beziehung

Legt man ein deterministisches Auswahlkriterium zugrunde, so kommt der Zufall dadurch ins Spiel, daß man ein zufälliges Auftreten der zu sortierenden Listen annimmt. Zur Vereinfachung nehmen wir an, daß es keine doppelt auftretenden Elemente gibt. Ohne Einschränkung seien deshalb die Elemente mit 1 bis n gekennzeichnet. Jede Permutation der Zahlen $1, \dots, n$ trete zudem mit der gleichen Wahrscheinlichkeit auf.

Unter diesen Voraussetzungen bietet sich folgende stochastische Modellierung an, wobei von einer zu sortierenden Liste der Länge n ausgegangen wird:

Als Ergebnisraum betrachte die Menge

$$\Omega_n = \{\pi, \pi \text{ Permutation der Zahlen } 1, \dots, n\}$$

und auf dieser die Laplace-Verteilung P_n bestimmt durch die Wahrscheinlichkeitsfunktion $p : \Omega_n \rightarrow [0, 1]$, $\omega \mapsto 1/|\Omega_n| = 1/n!$

X_n sei die Abbildung von Ω_n nach \mathbb{N}_0 , die jeder Permutation π die bei Anwendung von

Quicksort benötigte Anzahl von Vergleichen zuordnet. X_n ist dann offensichtlich eine Zufallsgröße mit Werten in \mathbb{N}_0 auf dem Wahrscheinlichkeitsraum $(\Omega_n, \mathcal{P}(\Omega_n), P_n)$. Um nun eine rekursive Beziehung herzuleiten, spielen weiterhin folgende Zufallsgrößen eine wichtige Rolle:

$Z_n : \Omega_n \rightarrow \{1, \dots, n\}$	Rang des ersten ausgewählten Partitionselementes (stimmt bei Nummerierung von $1, \dots, n$ mit der Nummer des ausgewählten Elementes überein)
$L_1, L_2 : \Omega_n \rightarrow \emptyset \cup \bigcup_{k=1}^{n-1} \{\pi _k : \pi \in \Omega_n\} =: \mathcal{X}$	Die nach der ersten Partitionierung vorliegende vordere bzw. hintere Liste.
$Y_1, Y_2 : \mathcal{X} \rightarrow \mathbb{N}$	Benötigte Anzahl von Vergleichen für die verbleibende vordere bzw. hintere Liste

Unter Benutzung dieser Zufallsgrößen erhält man nach dem Algorithmus für Quicksort die Gleichheit:

$$(3.1) \quad X_n = Y_1 \circ L_1 + Y_2 \circ L_2 + n \Leftrightarrow 1.$$

$n \Leftrightarrow 1$ bezeichnet dabei die für die Partitionierung erforderliche Anzahl von Vergleichen. Weiterhin gilt:

1. Z_n ist auf $\{1, \dots, n\}$ Laplace-verteilt, was sich direkt aus Verwendung der deterministischen Auswahlregel und der Gleichverteilung auf den Permutationen ergibt.
2. L_1 bzw. L_2 sind unter $Q_i := P_n(\cdot | Z_n = i)$, $i \in \{1, \dots, n\}$, Laplace-verteilt auf den Permutationen der $i \Leftrightarrow 1$ kleineren Elemente bzw. der $n \Leftrightarrow i$ größeren Elemente.

BEWEIS: Der Beweis ist eng mit der Frage verbunden, wie viele Permutationen $\pi \in \Omega_n$ existieren, die nach der ersten Partitionierung der Liste als vordere (bzw. hintere) Liste eine vorgegebene Permutation σ liefern mit $\sigma \in \{\text{Permutationen der Elemente } 1, \dots, i \Leftrightarrow 1 \text{ (bzw. } i + 1, \dots, n)\}$. Dazu kann man folgendes anmerken (es sei hier auf L_1 eingegangen; entsprechende analoge Überlegungen gelten für die Liste der größeren Elemente):

- i) Es kann die Lage der $n \Leftrightarrow i$ größeren Elemente variieren. Dafür gibt es $\binom{n-1}{n-i}$ Möglichkeiten, da die Position des i -ten Elementes vorgegeben ist (durch die Bedingung $Z_n = i$).
- ii) Für jede vorgegebene Lage der größeren Elemente kann es nur eine Möglichkeit für die Anordnung der kleineren Elemente innerhalb der freigeblichen Plätze geben, so daß man nach der Partitionierung zur vorgegebenen Liste σ gelangt.
- iii) In welcher Reihenfolge die größeren Elemente auf den vorherbestimmten Plätzen angeordnet sind, spielt keine Rolle (bei der Partitionierung erfolgt immer nur ein Vergleich mit dem Partitionselement, d.h. bei vertauschten größeren Elementen muß bei gleicher „Lage“ und gleichem Aufspaltungsalgorithmus die gleiche Liste der kleineren Elemente herauskommen (relative Gleichwertigkeit der größeren Elemente)).

Mit diesen Betrachtungen erhält man nun leicht: Ist $x_j \in \{1, \dots, i \Leftrightarrow 1\}$ für $1 \leq j \leq i \Leftrightarrow 1$ mit $x_{j_1} \neq x_{j_2}$ für $j_1 \neq j_2$, so gilt:

$$\begin{aligned} P_n(L_1 = (x_1, \dots, x_{i-1}) \mid Z_n = i) &= \frac{P_n(L_1 = (x_1, \dots, x_{i-1}), Z_n = i)}{P_n(Z_n = i)} = \frac{\binom{n-1}{n-i} \cdot (n \Leftrightarrow i)!}{\frac{1}{n}} \\ &= \frac{(n \Leftrightarrow 1)! \cdot (n \Leftrightarrow i)!}{(n \Leftrightarrow i)! \cdot (i \Leftrightarrow 1)! \cdot (n \Leftrightarrow 1)!} = \frac{1}{(i \Leftrightarrow 1)!}, \end{aligned}$$

wie man sich mittels kombinatorischer Überlegungen leicht klar macht. Analog ergibt sich für L_2 und $y_k \in \{i+1, \dots, n\}$ für $1 \leq k \leq n \Leftrightarrow i$ mit $y_{k_1} \neq y_{k_2}$ für $k_1 \neq k_2$:

$$P_n(L_2 = (y_1, \dots, y_{n-i}) \mid Z_n = i) = \frac{\binom{n-1}{i-1} \cdot (i \Leftrightarrow 1)!}{\frac{1}{n}} = \frac{1}{(n \Leftrightarrow i)!}.$$

3. L_1 und L_2 sind unter $P_n(\cdot \mid Z_n = i)$ unabhängig.

BEWEIS: Sei σ_1 eine Permutation der Elemente $1, \dots, i \Leftrightarrow 1$ und σ_2 eine Permutation der Elemente $i+1, \dots, n$. Dann gilt unter Benutzung der vorherigen Aussage:

$$\begin{aligned} P_n(L_1 = \sigma_1, L_2 = \sigma_2 \mid Z_n = i) &= \frac{P_n(L_1 = \sigma_1, L_2 = \sigma_2 \mid Z_n = i)}{P_n(Z_n = i)} = \frac{\binom{n-1}{i-1}}{\frac{1}{n}} \\ &= \frac{1}{(i \Leftrightarrow 1)!} \cdot \frac{1}{(n \Leftrightarrow i)!} = P_n(L_1 = \sigma_1 \mid Z_n = i) \cdot P_n(L_2 = \sigma_2 \mid Z_n = i). \end{aligned}$$

Betrachtet man nun die Verteilung der Zufallsgröße X_n , so erhält man unter Benutzung von (3.1) und der vorherigen Anmerkungen, wenn man wiederum $Q_i := P_n(\cdot | Z_n = i)$ setzt für $i \in \{1, \dots, n\}$:

$$(3.2) \quad \begin{aligned} P_n^{X_n} &= \sum_{i=1}^n P_n^{X_n | Z_n=i} P_n(Z_n = i) = \sum_{i=1}^n \frac{1}{n} \cdot P_n^{Y_1 \circ L_1 + Y_2 \circ L_2 + n-1 | Z_n=i} \\ &= \sum_{i=1}^n \frac{1}{n} \cdot Q_i^{Y_1 \circ L_1 + Y_2 \circ L_2 + n-1} = \sum_{i=1}^n \frac{1}{n} \cdot Q_i^{Y_1 \circ L_1} * Q_i^{Y_2 \circ L_2} * Q_i^{n-1}. \end{aligned}$$

Dabei sind $Q_i^{L_1}$ und $Q_i^{L_2}$ Laplace-Verteilungen auf $\{\pi : \pi \text{ Permutation der Elemente } 1, \dots, i \Leftrightarrow 1\}$ bzw. $\{\pi : \pi \text{ Permutation der Elemente } i+1, \dots, n\}$. Da für den Algorithmus die Numerierung der Elemente keine Rolle spielt, gilt:

$$Q_i^{Y_1 \circ L_1} = P_{i-1}^{X_{i-1}} \quad \text{und} \quad Q_i^{Y_2 \circ L_2} = P_{n-i}^{X_{n-i}}.$$

Hat man nun Zufallsgrößen $X'_0, \dots, X'_{n-1}, X''_0, \dots, X''_{n-1}$ auf einem gemeinsamen Wahrscheinlichkeitsraum $(\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{P})$ gegeben, wobei X'_i unabhängig von X''_j ist, für alle $i, j \in \{0, \dots, n \Leftrightarrow 1\}$ und $X'_0 = 0, X'_1 = 0, X'_2 = 1, X'_i \sim X''_i$ für $0 \leq i \leq n \Leftrightarrow 1$, $\tilde{P}^{X'_i} = P_i^{X_i}$ für $0 \leq i \leq n \Leftrightarrow 1$ und eine Zufallsgröße Z'_n Laplace-verteilt auf $\{1, \dots, n\}$ und unabhängig von $(X'_1, \dots, X'_{n-1}, X''_1, \dots, X''_{n-1})$, so gilt unter Benutzung von (3.2):

$$(3.3) \quad \begin{aligned} P_n^{X_n} &= \sum_{i=1}^n \frac{1}{n} \cdot \tilde{P}^{X'_{i-1}} * \tilde{P}^{X''_{n-1}} * \delta_{n-1} \\ &= \sum_{i=1}^n \frac{1}{n} \cdot \tilde{P}^{X'_{i-1} + X''_{n-1} + n-1} = \sum_{i=1}^n \frac{1}{n} \cdot \tilde{P}^{X'_{i-1} + X''_{n-1} + n-1 | Z'_n=i} \\ &= \tilde{P}^{X'_{Z'_n-1} + X''_{n-Z'_n} + n-1}. \end{aligned}$$

Folglich gilt für die Verteilung $L(X_n)$ der Zufallsgröße X_n die rekursive Beziehung $X_0 = 0, X_1 = 0, X_2 = 1$,

$$(3.4) \quad L(X_n) = L(X'_{Z_n-1} + X''_{n-Z_n} + n \Leftrightarrow 1), \quad n \geq 2.$$

Dabei gelte $X'_i \sim X_i \sim X''_i$ und die Zufallsgrößen seien untereinander unabhängig sowie unabhängig von Z_n .

Diese rekursive Beziehung ist naheliegend, da nach der Partitionierung ein Aufruf des gleichen Algorithmus auf niedrigerer Stufe erfolgt, die sich nach dem Rang des ausgewählten Elementes richtet. Die obigen Betrachtungen sind dabei notwendig, um zu zeigen, daß sich die Gleichverteilung auf die entsprechenden Teillisten überträgt.

Um später gewisse Wahrscheinlichkeiten für Abweichungen von der erwarteten Anzahl von Vergleichen berechnen zu können, ist es zunächst erforderlich, den Erwartungswert zu kennen. Dessen Berechnung ist Gegenstand des nächsten Abschnitts.

Berechnung des Erwartungswertes der Zufallsgröße X_n

Mit Hilfe der rekursiven Gleichung (3.4) erhält man die Gleichheit:

$$(3.5) \quad E(X_n) = n \Leftrightarrow 1 + \sum_{i=1}^n P(Z_n = i)(E(X_{i-1}) + E(X_{n-i}))$$

und damit weiter:

$$\begin{aligned} \frac{E(X_n)}{n+1} &= \frac{n \Leftrightarrow 1}{n+1} + \frac{1}{n \cdot (n+1)} \cdot \sum_{i=1}^n (E(X_{i-1}) + E(X_{n-i})) \\ &= \frac{n \Leftrightarrow 1}{n+1} + \frac{2 \cdot E(X_{n-1})}{n \cdot (n+1)} + \frac{1}{n \cdot (n+1)} \cdot \sum_{i=2}^{n-1} (E(X_{i-1}) + E(X_{n-i})) \\ &= \frac{n \Leftrightarrow 1}{n+1} + \frac{2 \cdot E(X_{n-1})}{n \cdot (n+1)} + \frac{1}{n \cdot (n+1)} \cdot \sum_{i=1}^{n-1} (E(X_{i-1}) + E(X_{n-1-i})) \\ &= \frac{n \Leftrightarrow 1}{n+1} + \frac{2 \cdot E(X_{n-1})}{n \cdot (n+1)} + \frac{n \Leftrightarrow 1}{n \cdot (n+1)} \cdot (E(X_{n-1}) \Leftrightarrow (n \Leftrightarrow 2)) \\ &= \frac{E(X_{n-1})}{n} + \frac{2 \cdot (n \Leftrightarrow 1)}{n \cdot (n+1)} \end{aligned}$$

Sukzessive Anwendung liefert dann:

$$\begin{aligned} (3.6) \quad \frac{E(X_n)}{n+1} &= \frac{E(X_{n-1})}{n} + \frac{2 \cdot (n \Leftrightarrow 1)}{n \cdot (n+1)} = \sum_{k=1}^{n-1} \frac{2 \cdot k}{(k+1)(k+2)} = \sum_{k=1}^{n-1} \frac{2 \cdot ((k+2) \Leftrightarrow 2)}{(k+1)(k+2)} \\ &= \sum_{k=1}^{n-1} \frac{2}{k+1} \Leftrightarrow \sum_{k=1}^{n-1} \frac{4}{(k+1)(k+2)} = 2 \cdot \sum_{k=2}^n \frac{1}{k} \Leftrightarrow 4 \cdot \sum_{k=2}^n \frac{1}{k(k+1)} \\ &= 2 \cdot \sum_{k=1}^n \frac{1}{k} \Leftrightarrow 4 \cdot \sum_{k=1}^n \frac{1}{k(k+1)} = 2 \cdot \sum_{k=1}^{n+1} \frac{1}{k} \Leftrightarrow \frac{2}{n+1} \Leftrightarrow \frac{4n}{n+1} \\ &= 2 \cdot \sum_{k=1}^{n+1} \frac{1}{k} + \frac{2}{n+1} \Leftrightarrow 4. \end{aligned}$$

Weiterhin ist nun bekannt, daß es sich bei der Folge $\sum_{k=1}^n \frac{1}{k}$, $n \in \mathbb{N}$, um eine divergente Reihe handelt (harmonische Reihe), die in ihrem Wachstumsverhalten der Logarithmusfunktion gleichkommt, und zwar in dem Sinn, daß der Grenzwert

$$\gamma = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{2} + \cdots + \frac{1}{n} \Leftrightarrow \ln n \right)$$

existiert. Der Grenzwert $\gamma = 0,5772156\dots$ wird Eulersche Konstante genannt. Die Existenz des Grenzwertes und einen Weg zu seiner Berechnung wird in [24] in Abschnitt 12.17 aufgezeigt, und zwar unter Verwendung der Eulerschen Summenformel; vgl auch Satz A.6 im Anhang bzw. S. 349 in [24]. Wendet man die Eulersche Summenformel speziell auf die Funktion $f(x) = 1/(x+n)$, $n \in \mathbb{N}$, $n \geq 1$, an, so erhält man die Gleichheit

$$\begin{aligned} \sum_{k=0}^m \frac{1}{n+k} &= \ln \frac{n+m}{n} + \frac{1}{2} \cdot \left(\frac{1}{n} + \frac{1}{n+m} \right) \\ &+ \sum_{i=1}^p \frac{B_{2i}}{2i} \left(\frac{1}{n^{2i}} \Leftrightarrow \frac{1}{(n+m)^{2i}} \right) \Leftrightarrow \int_0^m \frac{C_{2p+1}(x)}{(x+n)^{2p+2}} dx \end{aligned}$$

für ein beliebiges $p \in \mathbb{N}$. B_j , $j \in \mathbb{N}$, bezeichnen dabei die Bernoullischen Zahlen und $C_j(x)$ ist die Funktion, die sich aus dem Bernoullischen Polynom $B_j(x)$ durch Einschränkung auf $[0, 1)$ und periodische Fortsetzung ergibt (d.h. $C_j(x) = B_j(x \Leftrightarrow [x])$). Für Details sei auf den Anhang und dort insbesondere auf die Abschnitte A.3 und A.4 verwiesen.. Addition von $1 + \dots + 1/(n \Leftrightarrow 1) \Leftrightarrow \ln(m+n)$ auf beiden Seiten, Grenzübergang $m \rightarrow \infty$ und anschließendes Umstellen der Gleichung liefert schließlich

$$(3.7) \quad H_n := \sum_{k=1}^n \frac{1}{k} = \gamma + \frac{1}{2n} + \ln n + \sum_{i=1}^p \frac{B_{2i}}{2in^{2i}} \Leftrightarrow \int_0^\infty \frac{C_{2p+1}(x)}{(x+n)^{2p+2}} dx,$$

für $n \in \mathbb{N}$. Weiterhin sieht man leicht, daß

$$(3.8) \quad g(n) := \sum_{i=1}^p \frac{B_{2i}}{2in^{2i}} \Leftrightarrow \int_0^\infty \frac{C_{2p+1}(x)}{(x+n)^{2p+2}} dx = O\left(\frac{1}{n^2}\right),$$

für $p \geq 2$. Insgesamt liefern also (3.7) und (3.8) die Beziehung:

$$(3.9) \quad H_n = \gamma + \frac{1}{2n} + \ln n + O\left(\frac{1}{n^2}\right).$$

Damit erhält man für den Erwartungswert von X_n , $n \in \mathbb{N}$:

$$\begin{aligned} (3.10) \quad E(X_n) &= (n+1) \cdot 2 \cdot \sum_{k=1}^{n+1} \frac{1}{k} + 2 \Leftrightarrow 4 \cdot (n+1) = (n+1) \cdot 2 \cdot H_n \Leftrightarrow 4n \\ &= 2\gamma n + 2\gamma + 1 + \frac{1}{n} + 2 \cdot n \cdot \ln n + 2 \cdot \ln n \Leftrightarrow 4n + 2(n+1) \cdot O\left(\frac{1}{n^2}\right) \\ &= 2n \ln n + n(2\gamma \Leftrightarrow 4) + 2 \ln n + 2\gamma + 1 + O\left(\frac{1}{n}\right). \end{aligned}$$

Im folgenden werden nun die Zufallsgrößen $Y_n := (X_n \Leftrightarrow E(X_n))/n$ betrachtet. Wir wollen zeigen, daß diese Folge in d_2 -Metrik gegen eine Zufallsgröße Y konvergiert, deren Verteilungsfunktion Fixpunkt einer Abbildung ist, die es noch zu bestimmen gilt.

3.2.3 Bestimmung der Fixpunktgleichung

Unter Ausnutzung von (3.4) für die Zufallsgröße X_n erhält man für Y_n die rekursive Beziehung $Y_0 = 0, Y_1 = 0$,

$$(3.11) \quad L(Y_n) = L\left(Y'_{Z'_n-1} \frac{Z'_n \Leftrightarrow 1}{n} + \bar{Y}'_{n-Z'_n} \frac{n \Leftrightarrow Z'_n}{n} + C_n(Z'_n)\right), \quad n \geq 2.$$

Dabei sind für jedes feste n die Zufallsvariablen $Y'_i, \bar{Y}'_j, i, j \in \{1, \dots, n \Leftrightarrow 1\}$ unabhängig, $Y'_i \sim \bar{Y}'_i$ und Z'_n ist gleichverteilt auf $\{1, \dots, n\}$ und unabhängig von $(Y'_1, \dots, Y'_{n-1}, \bar{Y}'_1, \dots, \bar{Y}'_{n-1})$. C_n ist eine Funktion auf $\{1, \dots, n\}$, definiert durch:

$$(3.12) \quad C_n(i) = \frac{n \Leftrightarrow 1}{n} + \frac{1}{n}(E(X_{i-1}) + E(X_{n-i}) \Leftrightarrow E(X_n)).$$

Für den Nachweis von (3.11) hat man nur zu beachten, daß aufgrund von (3.4) $L(Y_n) = L(((X_{Z'_n-1} + \bar{X}_{n-Z'_n} + n \Leftrightarrow 1) \Leftrightarrow E(X_n))/n)$ ist.

Betrachte nun die Funktion $C : (0, 1) \rightarrow \mathbb{R}, x \mapsto 1 + 2t \ln t + 2(1 \Leftrightarrow t) \ln(1 \Leftrightarrow t)$. Bezeichnet $[x]$ die kleinste ganze Zahl größer oder gleich x und $h_n(t) := [nt]$, so besteht folgender Bezug zwischen der Funktionenfolge $C_n \circ h_n$ und C [\Leftrightarrow [6], Lemma 6.1]:

3.2.1 Lemma Die Funktionenfolge $C_n \circ h_n = C_n([n \cdot \cdot])$ konvergiert für alle $\delta_0, \delta_1 \in (0, 1)$ mit $0 < \delta_0 < \delta_1$ gleichmäßig auf $[\delta_0, \delta_1]$ gegen die Funktion C . Desweiteren sind die C_n gleichmäßig beschränkt, d.h. es gilt $\sup \{C_n(i) : n \in \mathbb{N}, i \in \{1, \dots, n\}\} < \infty$.

BEWEIS: Führt man die Bezeichnungen

$$f : [0, \infty) \rightarrow \mathbb{R}, f(t) = \begin{cases} 0 & : t = 0 \\ t \cdot \ln t & : t > 0 \end{cases},$$

$$R : \mathbb{N}_0 \rightarrow \mathbb{R}, R(k) := \begin{cases} 0 & : k = 0 \\ E(X_k) \Leftrightarrow f(k) & : k > 0 \end{cases}$$

ein, so gilt nach (3.10) und (3.12) für alle $t \in (0, 1)$ und $n \in \mathbb{N}$:

$$\begin{aligned}
 (3.13) \quad C_n(\lceil nt \rceil) &= \frac{n \Leftrightarrow 1}{n} + \frac{1}{n} \cdot (E(X_{\lceil nt \rceil - 1}) + E(X_{n - \lceil nt \rceil}) \Leftrightarrow E(X_n)) \\
 &= \frac{n \Leftrightarrow 1}{n} + \frac{1}{n} \cdot (2 \cdot f(\lceil nt \rceil \Leftrightarrow 1) + R(\lceil nt \rceil \Leftrightarrow 1)) \\
 &\quad + 2 \cdot f(n \Leftrightarrow \lceil nt \rceil) + R(n \Leftrightarrow \lceil nt \rceil) \Leftrightarrow 2 \cdot f(n) \Leftrightarrow R(n) \\
 &= \frac{n \Leftrightarrow 1}{n} + 2 \cdot f\left(\frac{\lceil nt \rceil \Leftrightarrow 1}{n}\right) + 2 \cdot f\left(\frac{n \Leftrightarrow \lceil nt \rceil}{n}\right) \Leftrightarrow \frac{2}{n} \cdot \ln n \\
 &\quad + \frac{\lceil nt \rceil \Leftrightarrow 1}{n} \cdot \frac{R(\lceil nt \rceil \Leftrightarrow 1)}{\lceil nt \rceil \Leftrightarrow 1} + \frac{n \Leftrightarrow \lceil nt \rceil}{n} \cdot \frac{R(n \Leftrightarrow \lceil nt \rceil)}{n \Leftrightarrow \lceil nt \rceil} \Leftrightarrow \frac{R(n)}{n}.
 \end{aligned}$$

Da offenkundig $\lim_{k \rightarrow \infty} \frac{R(k)}{k}$ existiert und endlich ist, erkennt man hieran direkt, daß auf jeden Fall punktweise Konvergenz im Intervall $(0, 1)$ vorliegt. Zu zeigen bleibt, daß diese Konvergenz auf vorgegebenem Intervall $[\delta_0, \delta_1] \subset (0, 1)$ gleichmäßig ist. Offenbar ist f stetig und beschränkt auf $[0, 1]$ und damit insbesondere auch gleichmäßig stetig, d.h. $\forall \varepsilon > 0 \exists \delta_\varepsilon > 0 \forall x, y \in [0, 1]$ mit $|x \Leftrightarrow y| \leq \delta_\varepsilon : |f(x) \Leftrightarrow f(y)| \leq \varepsilon$. Sei nun ein $\varepsilon > 0$ vorgegeben. Dann existiert ein $n_0 \in \mathbb{N}$, so daß für alle $n \geq n_0$ gilt: $\frac{2}{n} \leq \delta_\varepsilon$. Also erhält man für alle $t \in (0, 1)$ und $n \geq n_0$:

$$\left| \frac{\lceil nt \rceil \Leftrightarrow 1}{n} \Leftrightarrow t \right| = \left| \frac{nt + (\lceil nt \rceil \Leftrightarrow nt) \Leftrightarrow 1}{n} \Leftrightarrow t \right| \leq \frac{\lceil nt \rceil \Leftrightarrow nt}{n} + \frac{1}{n} \leq \frac{2}{n} \leq \delta_\varepsilon$$

und damit:

$$\left| f\left(\frac{\lceil nt \rceil \Leftrightarrow 1}{n}\right) \Leftrightarrow f(t) \right| \leq \varepsilon.$$

Ein analoges Vorgehen liefert

$$\exists n_1 \in \mathbb{N} \forall n \geq n_1 \forall t \in (0, 1) : \left| f\left(\frac{n \Leftrightarrow \lceil nt \rceil}{n}\right) \Leftrightarrow f(t) \right| \leq \varepsilon.$$

Folglich konvergiert sowohl die Funktionenfolge $2 \cdot f\left(\frac{\lceil nt \rceil - 1}{n}\right)$ als auch $2 \cdot f\left(\frac{n - \lceil nt \rceil}{n}\right)$ gleichmäßig auf $(0, 1)$ für n gegen unendlich gegen $2 \cdot f(t) = 2 \cdot t \ln t$ bzw. $2 \cdot f(1 \Leftrightarrow t) = 2 \cdot (1 \Leftrightarrow t) \ln(1 \Leftrightarrow t)$. Da $\frac{n-1}{n}$ und $\Leftrightarrow 2 \cdot \frac{\ln n}{n}$ nicht von t abhängen und somit ebenfalls gleichmäßig gegen 1 bzw. 0 konvergieren, bleibt nur noch die Funktionenfolge

$$g_n(t) := \frac{\lceil nt \rceil \Leftrightarrow 1}{n} \cdot \frac{R(\lceil nt \rceil \Leftrightarrow 1)}{\lceil nt \rceil \Leftrightarrow 1} + \frac{n \Leftrightarrow \lceil nt \rceil}{n} \cdot \frac{R(n \Leftrightarrow \lceil nt \rceil)}{n \Leftrightarrow \lceil nt \rceil} \Leftrightarrow \frac{R(n)}{n}$$

zu untersuchen. Dazu sei nun $\varepsilon > 0$ und $[\delta_0, \delta_1] \subset (0, 1)$ vorgegeben, wobei ohne Einschränkung $\delta_1 = 1 \Leftrightarrow \delta_0$ gelte. Bezeichnet $a := \lim_{k \rightarrow \infty} \frac{R(k)}{k} = 2\gamma \Leftrightarrow 4$, so existiert ein $n_0 \in \mathbb{N}$, so daß für alle $n \geq n_0$ $\left| \frac{R(n)}{n} \Leftrightarrow a \right| \leq \frac{\varepsilon}{3}$ ist. Wählt man nun $n \geq \max\left\{\frac{n_0+1}{\delta_0}, \frac{3|a|}{\varepsilon}\right\}$, so gilt für $t \in [\delta_0, \delta_1]$:

$$(i) \lceil nt \rceil \Leftrightarrow 1 \geq nt \Leftrightarrow 1 \geq n\delta_0 \Leftrightarrow 1 \geq n_0,$$

$$(ii) n \Leftrightarrow \lceil nt \rceil \geq n \Leftrightarrow \lceil n\delta_1 \rceil \geq n \Leftrightarrow n\delta_1 \Leftrightarrow 1 = n(1 \Leftrightarrow \delta_1) \Leftrightarrow 1 = n\delta_0 \Leftrightarrow 1 \geq n_0,$$

$$(iii) \left| \frac{a}{n} \right| \leq \frac{\varepsilon}{3}.$$

Folglich erhält man nun:

$$\begin{aligned} |g_n(t) \Leftrightarrow 0| &= \left| \frac{\lceil nt \rceil \Leftrightarrow 1}{n} \cdot \left(\frac{R(\lceil nt \rceil \Leftrightarrow 1)}{\lceil nt \rceil \Leftrightarrow 1} \Leftrightarrow a \right) + \frac{n \Leftrightarrow \lceil nt \rceil}{n} \cdot \left(\frac{R(n \Leftrightarrow \lceil nt \rceil)}{n \Leftrightarrow \lceil nt \rceil} \Leftrightarrow a \right) \right. \\ &\quad \left. \Leftrightarrow \left(\frac{R(n)}{n} \Leftrightarrow a \right) + \left(\frac{n \Leftrightarrow 1}{n} \Leftrightarrow 1 \right) \cdot a \right| \\ &\leq \frac{\lceil nt \rceil \Leftrightarrow 1}{n} \cdot \left| \frac{R(\lceil nt \rceil \Leftrightarrow 1)}{\lceil nt \rceil \Leftrightarrow 1} \Leftrightarrow a \right| + \frac{n \Leftrightarrow \lceil nt \rceil}{n} \cdot \left| \frac{R(n \Leftrightarrow \lceil nt \rceil)}{n \Leftrightarrow \lceil nt \rceil} \Leftrightarrow a \right| \\ &\quad + \left| \frac{R(n)}{n} \Leftrightarrow a \right| + \left| \frac{a}{n} \right| \\ &\leq \frac{\lceil nt \rceil \Leftrightarrow 1}{n} \cdot \frac{\varepsilon}{3} + \frac{n \Leftrightarrow \lceil nt \rceil}{n} \cdot \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \left| \frac{a}{n} \right| \leq 2 \cdot \frac{\varepsilon}{3} + \left| \frac{a}{n} \right| \leq \varepsilon. \end{aligned}$$

Als endliche Summe von auf $[\delta_0, \delta_1]$ gleichmäßig konvergenten Funktionenfolgen ist $C_n \circ h_n$ ebenfalls auf diesem Intervall gleichmäßig konvergent, und es folgt die erste Behauptung des Lemmas.

Die zweite Behauptung ist eine einfache Konsequenz aus (3.13), wenn man dort $\lceil nt \rceil$ durch i ersetzt, und die Beschränktheit der Funktion f auf $[0, 1]$ und die Konvergenz von $R(k)/k$ für k gegen unendlich ausnutzt. \square

Um die Frage nach der Fixpunktgleichung beantworten zu können, fehlt nun nur noch eine Aussage über die Konvergenz von Z'_n/n . Dies liefert das folgende Lemma:

3.2.2 Lemma *Ist Z'_n eine auf $\{1, \dots, n\}$ Laplace-verteilte Zufallsgröße, so gilt für die Verteilung Q_n von Z'_n/n :*

$$Q_n \xrightarrow{w} \mathcal{R}(0, 1).$$

BEWEIS: Man erkennt leicht, daß für alle $t \in \mathbb{R}$ die zu Q_n gehörige Verteilungsfunktion an der Stelle t gegen den entsprechenden Wert der Verteilungsfunktion der $\mathcal{R}(0, 1)$ -Verteilung konvergiert. Daraus ergibt sich die Behauptung. \square

Beachtet man nun, daß $C_n(Z_n) = C_n(n \cdot \frac{Z_n}{n}) = C_n(\lceil n \cdot \frac{Z_n}{n} \rceil)$ für eine auf $\{1, \dots, n\}$ Laplace-verteilte Zufallsgröße Z_n gilt, und nimmt man für einen Augenblick an, daß Y_n

in Verteilung gegen eine Zufallsgröße Y konvergiert, so liegt aufgrund der beiden letzten Lemmata und der rekursiven Gleichung (3.11) die Vermutung nahe, daß

$$L(Y) = L(Y\tau + \bar{Y}(1 \Leftrightarrow \tau) + C(\tau))$$

mit τ, Y, \bar{Y} unabhängig, $L(Y) = L(\bar{Y})$ und $\tau \sim \mathcal{R}(0, 1)$.

Unsere weitere Argumentation wird nun darin bestehen zu zeigen, daß solch ein Fixpunkt existiert, und daß die Folge Y_n in d_2 -Metrik gegen diesen Fixpunkt konvergiert.

3.2.4 Nachweis der Existenz eines Fixpunktes

Definiere eine Abbildung $S : \mathcal{M}_2 \rightarrow \mathcal{M}_2$ durch:

$$S(F) := L(\tau X + (1 \Leftrightarrow \tau)\bar{X} + C(\tau))$$

mit X, \bar{X}, τ unabhängig, $L(X) = L(\bar{X}) = F$, $\tau \sim \mathcal{R}(0, 1)$, $C(t)$ wie im vorherigen Abschnitt. Definiert man nun

$$T_1 := \tau, T_2 := 1 \Leftrightarrow \tau, 0 = T_3 = T_4 = \dots, C := C(\tau),$$

so kann man für $2 \leq p < \infty$ festhalten:

- 1) $\sum_i T_i \equiv 1$;
- 2) $E(C(\tau)) = 0$, denn man rechnet leicht nach, daß $E(\tau \cdot \ln \tau) = E((1 \Leftrightarrow \tau) \cdot \ln(1 \Leftrightarrow \tau)) = \Leftrightarrow 1/4$, und erhält somit $E(C(\tau)) = 2 \cdot E(\tau \cdot \ln \tau) + 2 \cdot E((1 \Leftrightarrow \tau) \cdot \ln(1 \Leftrightarrow \tau)) + 1 = 0$;
- 3) $E(\sum_i T_i^2) < 1$;
- 4) $E(\sum_i |T_i|^p) < 1$, denn $T_i < 1$ f.s. für alle $i \in \mathbb{N}$ und $p > 1$;
- 5) $\|\sum_i T_i^2\|_{p/2} \leq \|1\|_{p/2} < \infty$;
- 6) $\|\sum_i |T_i|\|_p < \infty$;
- 7) $\|C(\tau)\|_p < \infty$, denn die Funktion $C(t)$ ist auf $(0, 1)$ beschränkt.

Es sind somit alle Voraussetzungen von Satz 2.3.7 und Satz 2.3.15 bzw. der anschließende Bemerkung erfüllt und man erhält die Existenz eines bis auf Translation eindeutigen Fixpunktes der Abbildung S . Da die Anwendung für alle $2 \leq p$ gesichert ist, existieren zudem für diesen Fixpunkt alle p -ten Momente.

3.2.5 Konvergenz von Y_n gegen den Fixpunkt

Die Konvergenz der Folge Y_n betreffend notieren wir:

3.2.3 Satz $L(Y_n)$ konvergiert in d_2 -Metrik gegen den zum Erwartungswert 0 gehörenden Fixpunkt F der Abbildung S .

BEWEIS: Für $n \in \mathbb{N}$ sei F_n die Verteilungsfunktion der Zufallsgröße $Y_n = (X_n \Leftrightarrow E(X_n))/n$. U, V, W seien unabhängige $\mathcal{R}(0, 1)$ verteilte Zufallsgrößen. G_n sei die Verteilungsfunktion einer auf $\{1/n, 2/n, \dots, 1\}$ Laplace-verteilten Zufallsgröße. Definiere:

$$\begin{aligned} Y'_n &:= F_n^{-1}(U), \quad n \in \mathbb{N}, & Y' &:= F^{-1}(U), \\ \bar{Y}'_n &:= F_n^{-1}(V), \quad n \in \mathbb{N}, & \bar{Y}' &:= F^{-1}(V), \\ Z'_n &:= n \cdot G_n^{-1}(W), \quad n \in \mathbb{N}. \end{aligned}$$

Dann ist Z'_n offenbar Laplace-verteilt auf $\{1, \dots, n\}$. G sei die Verteilungsfunktion einer $\mathcal{R}(0, 1)$ -Verteilung und $\tau := G^{-1}(W)$. Es gilt:

$$\begin{aligned} (3.14) \quad d_2(L(Y_n), F) &= d_2(F_n, F) \\ &\leq \left\| \frac{Z'_n \Leftrightarrow 1}{n} \cdot Y'_{Z'_n-1} + \frac{n \Leftrightarrow Z'_n}{n} \cdot \bar{Y}'_{n-Z'_n} + C_n(Z'_n) \Leftrightarrow \tau \cdot Y' \Leftrightarrow (1 \Leftrightarrow \tau) \cdot \bar{Y}' \Leftrightarrow C(\tau) \right\|_2 \\ &\leq \left\| \frac{Z'_n \Leftrightarrow 1}{n} \cdot (Y'_{Z'_n-1} \Leftrightarrow Y') + \frac{n \Leftrightarrow Z'_n}{n} \cdot (\bar{Y}'_{n-Z'_n} \Leftrightarrow \bar{Y}') \right\|_2 \\ &\quad + \left\| \left(\frac{Z'_n \Leftrightarrow 1}{n} \Leftrightarrow \tau \right) \cdot Y' + \left(\frac{n \Leftrightarrow Z'_n}{n} \Leftrightarrow (1 \Leftrightarrow \tau) \right) \cdot \bar{Y}' \right\|_2 \\ &\quad + \|C_n(Z'_n) \Leftrightarrow C(\tau)\|_2. \end{aligned}$$

Wenden wir uns nun den drei letzten Termen einzeln zu. Für den ersten Term erhält man:

$$\begin{aligned} (3.15) \quad &\left\| \frac{Z'_n \Leftrightarrow 1}{n} \cdot (Y'_{Z'_n-1} \Leftrightarrow Y') + \frac{n \Leftrightarrow Z'_n}{n} \cdot (\bar{Y}'_{n-Z'_n} \Leftrightarrow \bar{Y}') \right\|_2 \\ &= \left(\sum_{i=1}^n \left(P(Z'_n = i) \cdot E \left(\left(\frac{i \Leftrightarrow 1}{n} \cdot (Y'_{i-1} \Leftrightarrow Y') + \frac{n \Leftrightarrow i}{n} \cdot (\bar{Y}'_{n-i} \Leftrightarrow \bar{Y}') \right)^2 \right) \right) \right)^{1/2} \\ &= \left(\frac{1}{n} \cdot \sum_{i=1}^n \left(\left(\frac{i \Leftrightarrow 1}{n} \right)^2 \cdot E(Y'_{i-1} \Leftrightarrow Y')^2 + \left(\frac{n \Leftrightarrow i}{n} \right)^2 \cdot E(\bar{Y}'_{n-i} \Leftrightarrow \bar{Y}')^2 \right) \right)^{1/2} \\ &= \left(\frac{1}{n} \cdot \left(\sum_{i=1}^n \left(\frac{i \Leftrightarrow 1}{n} \right)^2 \cdot d_2^2(F_{i-1}, F) + \sum_{i=1}^n \left(\frac{n \Leftrightarrow i}{n} \right)^2 \cdot d_2^2(F_{n-i}, F) \right) \right)^{1/2} \\ &= \left(\frac{2}{n} \cdot \sum_{i=1}^{n-1} \left(\frac{i}{n} \right)^2 \cdot d_2^2(F_i, F) \right)^{1/2}. \end{aligned}$$

Für das zweite Gleichheitszeichen hat man zu beachten, daß die gemischten Terme wegfallen, da Y'_{i-1}, Y' unabhängig von \bar{Y}'_{n-i}, \bar{Y}' sind und jeweils den Erwartungswert Null haben. Betrachte für den zweiten Term die Abschätzung:

$$\begin{aligned} & \left\| \left(\frac{Z'_n \Leftrightarrow 1}{n} \Leftrightarrow \tau \right) \cdot Y' + \left(\frac{n \Leftrightarrow Z'_n}{n} \Leftrightarrow (1 \Leftrightarrow \tau) \right) \cdot \bar{Y}' \right\|_2 \\ & \leq \left(\left\| \frac{Z'_n \Leftrightarrow 1}{n} \Leftrightarrow \tau \right\|_2 + \left\| \frac{n \Leftrightarrow Z'_n}{n} \Leftrightarrow (1 \Leftrightarrow \tau) \right\|_2 \right) \cdot \|Y'\|_2 \\ & = \left(\left\| G_n^{-1}(W) \Leftrightarrow \frac{1}{n} \Leftrightarrow G^{-1}(W) \right\|_2 + \left\| 1 \Leftrightarrow G_n^{-1}(W) \Leftrightarrow (1 \Leftrightarrow G^{-1}(W)) \right\|_2 \right) \cdot \|Y'\|_2. \end{aligned}$$

Da aus $G_n^{-1}(W) \xrightarrow{f.s.} G^{-1}(W)$ (vgl. Lemma 3.2.2) sowohl $G_n^{-1}(W) \Leftrightarrow 1/n \xrightarrow{f.s.} G^{-1}(W)$ als auch $1 \Leftrightarrow G_n^{-1}(W) \xrightarrow{f.s.} 1 \Leftrightarrow G^{-1}(W)$ folgt, ergibt sich leicht mittels des Satzes von der majorisierten Konvergenz:

$$(3.16) \quad \lim_{n \rightarrow \infty} \left\| \left(\frac{Z'_n \Leftrightarrow 1}{n} \Leftrightarrow \tau \right) \cdot Y' + \left(\frac{n \Leftrightarrow Z'_n}{n} \Leftrightarrow (1 \Leftrightarrow \tau) \right) \cdot \bar{Y}' \right\|_2 = 0.$$

Es verbleibt die Untersuchung des dritten Terms, wozu wir zunächst festhalten können, daß $\lim_{n \rightarrow \infty} G_n^{-1}(w) = G^{-1}(w)$ für alle $w \in (0, 1)$ gilt (der Stetigkeitsbereich von G^{-1} ist gerade das offene Intervall $(0, 1)$). Außerdem ist $G^{-1}(w) > 0$ für alle $w \in (0, 1)$. Für vorgegebenes $w \in (0, 1)$ existiert deshalb ein $\varepsilon_w > 0$ mit $[G^{-1}(w) \Leftrightarrow \varepsilon_w, G^{-1}(w) + \varepsilon_w] \subset (0, 1)$ und ein $n_0 \in \mathbb{N}$, so daß $G_n^{-1}(w) \in [G^{-1}(w) \Leftrightarrow \varepsilon_w, G^{-1}(w) + \varepsilon_w]$ für alle $n \geq n_0$. Zusammen mit der gleichmäßigen Konvergenz von $C_n(\lceil n \cdot \rceil)$ auf dieser Umgebung nach Lemma 3.2.1 ergibt dies $\lim_{n \rightarrow \infty} C_n(nG_n^{-1}(w)) = C(G^{-1}(w))$ für alle $w \in (0, 1)$ und damit auch

$$C_n(nG_n^{-1}(W)) \xrightarrow{f.s.} C(G^{-1}(W)).$$

Nutzen wir nun noch die gleichmäßige Beschränktheit von $C_n(i)$ (vgl. Lemma 3.2.1) und den Satz von der majorisierten Konvergenz aus, folgt:

$$(3.17) \quad \|C_n(Z'_n) \Leftrightarrow C(\tau)\|_2 \xrightarrow{n \rightarrow \infty} 0.$$

Kombination der Ergebnisse (3.14), (3.15), (3.16) und (3.17) ergibt

$$d_2(F_n, F) \leq \left(\frac{2}{n} \cdot \sum_{i=1}^{n-1} \left(\frac{i}{n} \right)^2 \cdot d_2^2(F_i, F) \right)^{1/2} + b_n,$$

wobei b_n eine Nullfolge darstellt. Anwendung von Satz A.3 aus dem Anhang für $p = 2$ liefert die Behauptung. \square

In dem vorherigen Satz wurde speziell die Konvergenz in der d_2 -Metrik nachgewiesen. Nach Abschnitt 2.3.2 existieren für den Fixpunkt aber sogar alle p -ten Momente, und da X_n eine beschränkte Zufallsgröße für alle n darstellt, ist auch die p -fache Integrierbarkeit der Zufallsgrößen Y_n gesichert. Man kann somit die Frage aufwerfen, ob nicht (Y_n) sogar allgemein in d_p -Metrik gegen den Fixpunkt F der Abbildung S konvergiert. Wir halten fest:

3.2.4 Satz $L(Y_n)$ konvergiert in d_p -Metrik, $p \geq 2$, gegen den zum Erwartungswert 0 gehörenden Fixpunkt der Abbildung S .

BEWEIS: Die Wahl der Zufallsvariablen ist wörtlich aus dem vorherigen Beweis zu übernehmen. Wir zeigen im folgenden zuerst durch eine Induktion über p , daß die Aussage für alle $p \in \mathbb{N}$ mit $p \geq 2$ richtig ist. Den Induktionsanfang bildet Satz 3.2.3. Im Induktionsschritt betrachten wir nicht den Übergang von $p \Leftrightarrow 1$ nach p , sondern von $p \Leftrightarrow 2$ nach p . Dies hat den Vorteil, daß man p als gerade annehmen kann, wenn man beachtet, daß der Induktionsanfang bei $p = 2$ liegt. Daß die entsprechende Aussage dann auch für $p \Leftrightarrow 1$ gilt, folgt aus der Beziehung $d_{p-1} \leq d_p$. Die Induktionsvoraussetzung besagt gerade, daß Konvergenz in d_q -Metrik für alle $2 \leq q \leq p \Leftrightarrow 2$ vorliegt. Analog zum vorherigen Beweis erhält man die Abschätzung:

$$(3.18) \quad \begin{aligned} d_p(L(Y_n), F) \leq & \left\| \frac{Z'_n \Leftrightarrow 1}{n} \cdot (Y'_{Z'_n-1} \Leftrightarrow Y') + \frac{n \Leftrightarrow Z'_n}{n} \cdot (\bar{Y}'_{n-Z'_n} \Leftrightarrow \bar{Y}') \right\|_p \\ & + \left\| \left(\frac{Z'_n \Leftrightarrow 1}{n} \Leftrightarrow \tau \right) \cdot Y' + \left(\frac{n \Leftrightarrow Z'_n}{n} \Leftrightarrow (1 \Leftrightarrow \tau) \right) \cdot \bar{Y}' \right\|_p \\ & + \|C_n(Z'_n) \Leftrightarrow C(\tau)\|_p. \end{aligned}$$

Für den zweiten und dritten Term folgt durch eine analoges Vorgehen wie im Spezialfall $p = 2$ die Konvergenz gegen Null für n gegen unendlich. Allein der erste Term erweist sich als etwas schwieriger, da die gemischten Terme im Kalkül verbleiben. Es gilt:

$$\begin{aligned} & \left\| \frac{Z'_n \Leftrightarrow 1}{n} \cdot (Y'_{Z'_n-1} \Leftrightarrow Y') + \frac{n \Leftrightarrow Z'_n}{n} \cdot (\bar{Y}'_{n-Z'_n} \Leftrightarrow \bar{Y}') \right\|_p \\ &= \left(\sum_{i=1}^n P(Z'_n = i) \cdot E \left(\left(\frac{i \Leftrightarrow 1}{n} \cdot (Y'_{i-1} \Leftrightarrow Y') + \frac{n \Leftrightarrow i}{n} \cdot (\bar{Y}'_{n-i} \Leftrightarrow \bar{Y}') \right)^p \right) \right)^{1/p} \\ &\leq \left(\sum_{i=1}^n \frac{1}{n} \cdot E \left(\left(\frac{i \Leftrightarrow 1}{n} \right)^p \cdot (Y'_{i-1} \Leftrightarrow Y')^p + \left(\frac{n \Leftrightarrow i}{n} \right)^p \cdot (\bar{Y}'_{n-i} \Leftrightarrow \bar{Y}')^p \right) \right)^{1/p} \\ &\quad + \left| \sum_{k=1}^{p-1} \binom{p}{k} \cdot \frac{1}{n} \cdot \sum_{i=1}^n E \left(\left(\frac{i \Leftrightarrow 1}{n} \right)^k \cdot (Y'_{i-1} \Leftrightarrow Y')^k \cdot \left(\frac{n \Leftrightarrow i}{n} \right)^{p-k} \cdot (\bar{Y}'_{n-i} \Leftrightarrow \bar{Y}')^{p-k} \right) \right|^{1/p} \end{aligned}$$

$$\leq \left(\frac{2}{n} \cdot \sum_{i=1}^{n-1} \left(\frac{i}{n} \right)^p \cdot d_p^p(F_i, F) \right)^{1/p} \\ + \left(\sum_{k=2}^{p-2} \binom{p}{k} \cdot \frac{1}{n} \cdot \sum_{i=1}^n E(|Y'_{i-1} \Leftrightarrow Y'|^k) \cdot E(|\bar{Y}'_{n-i} \Leftrightarrow \bar{Y}'|^{p-k}) \right)^{1/p}.$$

Zu beachten hat man bei dieser Rechnung nur, daß die Betragsstriche der \mathfrak{L}_p -Norm entfallen, da p gerade ist, und daß aufgrund von $E(Y'_{i-1} \Leftrightarrow Y') = E(\bar{Y}'_{n-i} \Leftrightarrow \bar{Y}') = 0$ für $i = 1, \dots, n$ die entsprechenden Terme für $k = 1$ und $k = p \Leftrightarrow 1$ wegfallen. Nach Induktionsvoraussetzung gilt weiterhin $\lim_{j \rightarrow \infty} d_{p-k}^{p-k}(F_j, F) = 0$ für $k = 2, \dots, p \Leftrightarrow 2$, woraus die Existenz von Konstanten $\text{const}(k)$ folgt mit $d_{p-k}^{p-k}(F_j, F) \leq \text{const}(k)$ für alle j . Somit erhält man:

(3.19)

$$\| \frac{Z'_n \Leftrightarrow 1}{n} \cdot (Y'_{Z'_n-1} \Leftrightarrow Y') + \frac{n \Leftrightarrow Z'_n}{n} \cdot (\bar{Y}'_{n-Z'_n} \Leftrightarrow \bar{Y}') \|_p \\ \leq \left(\frac{2}{n} \cdot \sum_{i=1}^{n-1} \left(\frac{i}{n} \right)^p \cdot d_p^p(F_i, F) \right)^{1/p} + \left(\sum_{k=2}^{p-2} \binom{p}{k} \cdot \frac{1}{n} \cdot \sum_{i=1}^n d_k^k(F_{i-1}, F) \cdot \text{const}(k) \right)^{1/p}.$$

Nochmalige Anwendung der Induktionsvoraussetzung zusammen mit Satz A.4 aus dem Anhang zeigt, daß der zweite Summand des letzten Terms eine Nullfolge bildet (in n). Insgesamt erhält man also aus (3.18) und (3.19):

$$d_p(L(Y_n), F) \leq \left(\frac{2}{n} \cdot \sum_{i=1}^{n-1} \left(\frac{i}{n} \right)^p \cdot d_p^p(F_i, F) \right)^{1/p} + b_n,$$

mit b_n Nullfolge, $b_n \geq 0$ für alle $n \in \mathbb{N}$. Anwendung von Satz A.3 aus dem Anhang komplettiert den Induktionsschritt, womit die Aussage des Satzes für alle $p \in \mathbb{N}$ gezeigt ist. Aus der Beziehung $d_q \leq d_p$ für alle $1 \leq q \leq p < \infty$ folgt schließlich die Behauptung für alle reellen $p \geq 1$. \square

3.2.6 Abschätzungen für das Auftreten schlechter Laufzeiten

Quicksort weist im worst-case eine äußerst hohe Zahl von Vergleichen auf. Wählt man als Partitionsregel zum Beispiel die Auswahl des ersten Elementes der Liste und ist die zu sortierende Liste bereits in geordneter Form gegeben, so benötigt Quicksort die erschreckend hohe Zahl von $n^2/2$ Vergleichen. Deshalb ist eine genauere Kenntnis der Verteilung der Anzahl der Vergleiche über den Erwartungswert hinaus wünschenswert. Wir haben bereits gezeigt, daß für die Anzahl der Vergleiche bei geeigneter Normierung eine Grenzverteilung

existiert, die sich zudem als Lösung einer Fixpunktgleichung herausgestellt hat. Die normierten Verteilungen konvergieren dabei in d_p -Metrik für beliebig gewähltes $p \in \mathbb{N}$ gegen diese Grenzverteilung, von der wir wissen, daß alle Momente endlich sind. Im folgenden wollen wir nun dieses Wissen ausnutzen, um zu zeigen, daß schlechtes Verhalten von Quicksort selten auftritt. Beginnen möchten wir jedoch mit einer näherungsweise Berechnung der Standardabweichung der Zufallsgröße X_n für großes n .

Aufgrund der Fixpunktgleichung

$$L(Y) = L(\tau Y + (1 \Leftrightarrow \tau)\bar{Y} + C(\tau))$$

mit $Y \sim \bar{Y}$ erhält man:

$$E(Y^2) = E((\tau Y + (1 \Leftrightarrow \tau)\bar{Y} + C(\tau))^2) = \frac{2}{3}E(Y^2) + E((C(\tau))^2).$$

Dabei wurde lediglich die Unabhängigkeit der Zufallsgrößen τ, Y, \bar{Y} ausgenutzt. Folglich gilt:

$$\begin{aligned} \text{Var}(Y) &= E(Y^2) = 3 \cdot E((C(\tau))^2) = 3 \cdot \int_0^1 (C(x))^2 dx \\ &= 3 \cdot \left(2 \cdot \int_0^1 4x^2 (\ln x)^2 dx + 8 \cdot \int_0^1 x \ln x \cdot (1 \Leftrightarrow x) \ln(1 \Leftrightarrow x) dx + 8 \cdot \int_0^1 x (\ln x) dx + 1 \right) \\ &= 3 \cdot \left(2 \cdot \frac{8}{27} + 8 \cdot \frac{37 \Leftrightarrow 3\pi^2}{108} \Leftrightarrow 8 \cdot \frac{1}{4} + 1 \right) = 7 \Leftrightarrow \frac{2}{3}\pi^2. \end{aligned}$$

Nach Satz 3.2.3 konvergiert die Folge $Y_n = (X_n \Leftrightarrow E(X_n))/n$ in d_2 -Metrik gegen die Zufallsgröße Y . Man weiß somit insbesondere, daß die zugehörigen zweiten Momente ebenfalls gegeneinander konvergieren. Es gilt:

$$\text{Var}(X_n) = n^2 \cdot \text{Var}(Y_n) = n^2 \cdot E(Y_n^2) \approx n^2 \cdot E(Y^2) = n^2 \cdot \left(7 \Leftrightarrow \frac{2}{3}\pi^2 \right) \approx 0.42 \cdot n^2$$

und $\sigma_{X_n} \approx 0.65 \cdot n$ [vgl. zu diesem Resultat auch [21]].

Im folgenden soll nun eine Abschätzung der Wahrscheinlichkeit für schlechtes Verhalten von Quicksort in bezug auf die benötigten Vergleiche erfolgen. Um schlechtes Verhalten zu definieren, ist es sinnvoll, Quicksort etwa mit dem Sortieralgorithmus Heapsort zu vergleichen. Dieser hat den Vorteil, daß die durchschnittlich benötigte Anzahl von Vergleichen $n \log_2 n$ beträgt (der führende Term bei Quicksort hat die Größenordnung $1.39 \cdot n \log_2 n$) und immer weniger als $4n \ln n$ Vergleiche benötigt werden. Deshalb interessieren wir uns insbesondere für die Wahrscheinlichkeit dafür, daß Quicksort mehr als $2E(X_n) \approx 4n \ln n$ Vergleiche braucht. Nach (3.10) wissen wir, daß $E(X_n)/n = 2 \cdot \ln n + R(n)/n$ mit $\lim_{n \rightarrow \infty} R(n)/n \in \mathbb{R}$. Daher erhalten wir unter Ausnutzung der Markov-Ungleichung die für genügend große n

gültige Abschätzung:

$$\begin{aligned} P(X_n \geq 2E(X_n)) &= P(|X_n \ominus E(X_n)| \geq E(X_n)) = P(|Y_n| \geq E(X_n)/n) \\ &\leq \frac{E(Y_n^2)}{(E(X_n)/n)^2} \leq \frac{\text{const}}{(E(X_n)/n)^2} \approx \frac{\text{const}}{(2 \cdot \ln n)^2}. \end{aligned}$$

Zu beachten hat man dabei, daß aufgrund der d_2 -Konvergenz von Y_n gegen Y auch die zugehörigen 2-ten Momente gegeneinander konvergieren und somit aus $L(Y_n), L(Y) \in \mathcal{M}_2$ die Existenz einer Konstanten const folgt, mit $E(Y_n^2) \leq \text{const}$ für alle n .

Benutzt man hingegen den allgemeineren Satz 3.2.4 mit dem insbesondere die Endlichkeit und Konvergenz der p -ten absoluten Momente gegen eine reelle Zahl folgt, so können wir die vorherige Abschätzung dahingehend verbessern, daß man die Konvergenz des Nenners gegen unendlich beschleunigt (unter Inkaufnahme einer Vergrößerung der Konstanten im Zähler). Eine analoge Abschätzung liefert nämlich:

$$P(X_n \geq 2E(X_n)) \leq \frac{E(|Y_n|^p)}{(E(X_n)/n)^p} \leq \frac{\text{const}(p)}{(E(X_n)/n)^p} \approx \frac{E(|Y|^p)}{(2 \cdot \ln n)^p} \quad \text{für } n \text{ genügend groß.}$$

Trotzdem ist dieses Resultat noch insofern unbefriedigend, als im Nenner immer noch die Logarithmusfunktion auftritt, die bekanntermaßen ein äußerst langsames Konvergenzverhalten für n gegen unendlich aufweist. Deswegen muß es nun noch unser Ziel sein eine entsprechende Abschätzung mit besserem Konvergenzverhalten zu finden, was uns unweigerlich zur Betrachtung von exponentiellen Momenten der Zufallsgrößen Y_n und des Fixpunktes Y treibt.

3.2.7 Momenterzeugende Funktionen

In diesem Abschnitt werden wir zeigen, daß die momenterzeugenden Funktionen der Zufallsgrößen Y_n auf ganz \mathbb{R} endlich sind und daß man in jedem vorgegebenem Intervall $[\ominus L, L]$ sogar gleichmäßige Beschränktheit (über alle n) vorliegen hat. Ziel dieses Abschnitts ist somit der Beweis des folgenden Satzes:

3.2.5 Satz Für alle $L > 0$ existiert ein $K_L \in \mathbb{R}^+$, so daß für alle $n \in \mathbb{N}$ und alle $\lambda \in [\ominus L, L]$ gilt:

$$E e^{\lambda Y_n} \leq e^{\lambda^2 K_L}.$$

Der Beweis dieses Satzes ist keinesfalls einfach und bedarf einiger Mühe. Für die folgenden Betrachtungen ist es notwendig, die Funktion

$$g_n : [1, n] \rightarrow \mathbb{R}, \quad g_n(x) := \left(\frac{x \ominus 1}{n} \right)^2 + \left(\frac{n \ominus x}{n} \right)^2 \ominus 1$$

zu untersuchen. Dies leistet

3.2.6 Lemma Die Funktion g_n ($n \geq 2$) nimmt Werte im Intervall $(\Leftrightarrow 1, 0)$ an. Zudem ist sie auf $[1, \frac{n+1}{2})$ monoton fallend, auf $[\frac{n+1}{2}, n]$ monoton steigend und nimmt für $x = \frac{n+1}{2}$ ihr Minimum an. Es gilt die Symmetrieeigenschaft: $g_n(n \Leftrightarrow i + 1) = g_n(i)$ für $i = 1, \dots, n$.

BEWEIS: Die erste und die letzte Behauptung ergibt sich direkt aus der Definition von g_n . Man rechnet leicht nach, daß $g'_n(x) = \frac{4x-2-2n}{n^2}$. Somit gilt $g'_n(x) = 0 \Leftrightarrow x = \frac{n+1}{2}$ und man erkennt, daß es sich tatsächlich um ein Minimum handelt, denn $g''_n(x) = \frac{4}{n^2} > 0$. Offensichtlich ist $g'_n(x) < 0$ auf $[1, \frac{n+1}{2})$ und ≥ 0 auf $[\frac{n+1}{2}, n]$ und es folgen die behaupteten Monotonieeigenschaften. \square

Nun können wir uns daran begeben, einen weiteren Satz zu beweisen, der für den induktiven Beweis von Satz 3.2.5 von entscheidender Bedeutung ist. Dazu sei im folgenden für jedes $n \in \mathbb{N}$ Z_n eine auf $\{1, \dots, n\}$ Laplace-verteilte Zufallsgröße,

$$U_n := \left(\frac{Z_n \Leftrightarrow 1}{n} \right)^2 + \left(\frac{n \Leftrightarrow Z_n}{n} \right)^2 \Leftrightarrow 1$$

und

$$f(K, \lambda, n) := E(\exp(\lambda C_n(Z_n) + \lambda^2 K U_n)) \text{ für } \lambda, K \in \mathbb{R}.$$

3.2.7 Satz Für alle $L > 0$ existiert ein $K_L > 0$, so daß für alle $n \in \mathbb{N}$ und $\lambda \in [\Leftrightarrow L, L]$:

$$f(K_L, \lambda, n) \leq 1.$$

Der Beweis dieses Satzes unterteilt sich in mehrere Schritte. Zunächst können wir festhalten:

3.2.8 Lemma Es existieren $K_1 > 0$, $L_1 > 0$, so daß für alle $n \in \mathbb{N}$ und $\lambda \in [\Leftrightarrow L_1, L_1]$:

$$f(K_1, \lambda, n) \leq 1.$$

BEWEIS: Unter Ausnutzung des Differentiationslemmas zeigt man leicht, daß f , aufgefaßt als Funktion in λ , zweimal stetig differenzierbar ist mit den Ableitungen

$$\begin{aligned} \frac{d}{d\lambda} f(K, \lambda, n) &= E((C_n(Z_n) + 2\lambda K U_n) \cdot \exp(\lambda C_n(Z_n) + \lambda^2 K U_n)), \\ \frac{d^2}{d\lambda^2} f(K, \lambda, n) &= E((2K U_n + (C_n(Z_n) + 2\lambda K U_n)^2) \cdot \exp(\lambda C_n(Z_n) + \lambda^2 K U_n)). \end{aligned}$$

Somit erhält man unter Ausnutzung von (3.12) und (3.5)

$$\begin{aligned} \frac{d}{d\lambda} f(K, \lambda, n) \Big|_{\lambda=0} &= E(C_n(Z_n)) = 0, \\ \frac{d^2}{d\lambda^2} f(K, \lambda, n) \Big|_{\lambda=0} &= E(C_n^2(Z_n) + 2K U_n). \end{aligned}$$

Nun existiert ein $K_1 > 0$ mit $d^2/d\lambda^2 f(K_1, \lambda, n)|_{\lambda=0} < 0$ für alle $n \in \mathbb{N}$. Daß dabei K_1 unabhängig von n gewählt werden kann, ist nicht ganz einfach einzusehen und bedarf einer näheren Untersuchung. Es gilt unter Benutzung von Lemma 3.2.1 mit $c := \sup \{C_n(i) : n \in \mathbb{N}, i \in \{1, \dots, n\}\}$:

$$(3.20) \quad \begin{aligned} \frac{d^2}{d\lambda^2} f(K, \lambda, n)|_{\lambda=0} &= E(C_n^2(Z_n) + 2KU_n) = E(C_n^2(Z_n)) + 2K \cdot E(U_n) \\ &\leq c^2 + 2K \cdot E(U_n). \end{aligned}$$

Eine genauere Betrachtung von $E(U_n)$ ergibt dabei:

$$(3.21) \quad \begin{aligned} E(U_n) &= \frac{1}{n^2} \cdot (E(Z_n \Leftrightarrow 1)^2 + E(n \Leftrightarrow Z_n)^2 \Leftrightarrow n^2) \\ &= \frac{1}{n^2} \cdot \left(2 \cdot \sum_{i=1}^{n-1} \frac{i^2}{n} \Leftrightarrow n^2\right) = \frac{1}{n^2} \cdot \left(\frac{1}{n} \cdot \frac{(n \Leftrightarrow 1) \cdot n \cdot (2n \Leftrightarrow 1)}{3} \Leftrightarrow n^2\right) \\ &= \Leftrightarrow \frac{1}{3} \Leftrightarrow \frac{1}{n} + \frac{1}{3n^2} < \Leftrightarrow \frac{1}{3}. \end{aligned}$$

Diese Ungleichung gilt für alle n . Unter Beachtung von (3.20) ergibt sich die Existenz einer Konstanten $K' > 0$, so daß

$$\frac{d^2}{d\lambda^2} f(K', \lambda, n)|_{\lambda=0} < 0$$

für alle $n \in \mathbb{N}$. Wir haben somit gezeigt, daß die Funktion $f(K', \lambda, n)$ an der Stelle $\lambda = 0$ für alle $n \in \mathbb{N}$ ein Maximum hat mit Wert $f(K_1, 0, n) = 1$. Dies ist für unsere Zwecke aber noch nicht ausreichend, denn wir wollen schließlich zeigen, daß eine von n unabhängige Umgebung existiert, auf welcher die Funktion Werte kleiner oder gleich eins annimmt. Dazu überlegt man sich leicht (z.B. mit dem Satz von Rolle), daß es ausreicht, die Existenz einer Umgebung um Null zu zeigen, in der die zweite Ableitung für alle n Werte kleiner als Null annimmt (für genügend großes K). Präzise formuliert bedeutet dies: Wir müssen noch zeigen, daß $L_1 > 0$ und $K_1 > 0$ existieren, so daß für alle $\lambda \in [\Leftrightarrow L_1, L_1]$ und alle $n \in \mathbb{N}$ gilt: $d^2/d\lambda^2 f(K_1, \lambda, n) < 0$. Es bleibt uns also nichts anderes übrig, als noch einmal einen genaueren Blick auf die zweite Ableitung zu werfen. Da die Umgebung beliebig klein werden darf, können wir ohne Einschränkung annehmen, daß $|\lambda \cdot K| \leq 1$ und $|\lambda| \leq 1$. Dann erhalten wir mit $c := \sup \{C_n(i) : n \in \mathbb{N}, i \in \{1, \dots, n\}\}$ und $U_n < 0$ f.s.:

$$\begin{aligned} \frac{d^2}{d\lambda^2} f(K, \lambda, n) &= E((2KU_n + (C_n(Z_n) + 2\lambda KU_n)^2) \cdot \exp(\lambda C_n(Z_n) + \lambda^2 KU_n)) \\ &= E(2KU_n \cdot \exp(\lambda C_n(Z_n) + \lambda^2 KU_n)) \\ &\quad + E((C_n(Z_n) + 2\lambda KU_n)^2 \cdot \exp(\lambda C_n(Z_n) + \lambda^2 KU_n)) \\ &\leq E(2KU_n \cdot \exp(\Leftrightarrow |\lambda| \cdot c \Leftrightarrow \lambda^2 K)) + (c + 2)^2 \cdot \exp(|\lambda| \cdot c + |\lambda|) \\ &\leq 2KE(U_n) \cdot \exp(\Leftrightarrow c \Leftrightarrow 1) + (c + 2)^2 \cdot \exp(c + 1). \end{aligned}$$

Zusammen mit (3.21) liefert dies die Existenz von $K_1 > 0$ und $L_1 > 0$ mit den behaupteten Eigenschaften. \square

Den zweiten Schritt bildet

3.2.9 Lemma Für vorgegebenes $L > 0$ existiert ein $K_2 > 0$, so daß für alle $n \in \mathbb{N}$:

$$f(K_2, L, n) \leq 1, \quad f(K_2, \Leftrightarrow L, n) \leq 1.$$

Für den Beweis dieses Lemmas ist es wichtig, die folgende Aussage zu beachten.

3.2.10 Lemma Die Funktion $f(K, \lambda, n)$ ist monoton fallend in K für festgehaltenes λ und n .

BEWEIS: Die Aussage des Lemmas kann man direkt an der ersten Ableitung

$$\frac{d}{dK} f(K, \lambda, n) = E(\lambda^2 U_n \exp(\lambda C_n(Z_n) + \lambda^2 K U_n)) \leq 0$$

ablesen. \square

BEWEIS VON LEMMA 3.2.9: Unter Benutzung von Lemma 3.2.10 reicht es offenkundig, die Aussage

$$\lim_{K \rightarrow \infty} \sup_n f(K, \lambda, n) = 0 \text{ für } \lambda \neq 0$$

zu beweisen. Sei dazu $\lambda \neq 0$ vorgegeben und, wie bereits weiter oben verwendet, $c := \sup \{C_n(i) : n \in \mathbb{N}, i \in \{1, \dots, n\}\}$. Dann erhält man

$$(3.22) \quad \begin{aligned} f(K, \lambda, n) &\leq E(\exp(|\lambda| \cdot |C_n(Z_n)| + \lambda^2 K U_n)) \\ &\leq \exp(|\lambda| \cdot c) \cdot E(\exp(\lambda^2 K U_n)) \end{aligned}$$

und weiter für den zweiten Faktor für vorgegebenes $j \in \mathbb{N}$ und $n \geq j$:

$$\begin{aligned} E(\exp(\lambda^2 K U_n)) &= \frac{1}{n} \cdot \sum_{i=1}^n \exp\left(\lambda^2 K \left(\left(\frac{i \Leftrightarrow 1}{n}\right)^2 + \left(\frac{n \Leftrightarrow i}{n}\right)^2 \Leftrightarrow 1\right)\right) \\ &\leq \frac{1}{n} \cdot (2([n/j] \Leftrightarrow 1) + (n \Leftrightarrow 2([n/j] \Leftrightarrow 1))) \cdot \exp(\lambda^2 K g_n([n/j])). \end{aligned}$$

Da nun offensichtlich $\lim_{n \rightarrow \infty} g_n([n/j]) = (1/j)^2 + (1 \Leftrightarrow 1/j)^2 \Leftrightarrow 1 =: a_j < 0$, gibt es ein $n_0 \in \mathbb{N}$, so daß für alle $n \geq n_0$: $g_n([n/j]) \leq 1/2 \cdot a_j$. Folglich gilt für $n \geq \max\{4, j, n_0\}$:

$$\begin{aligned} E(\exp(\lambda^2 K U_n)) &\leq \frac{2}{n} \cdot ([n/j] \Leftrightarrow 1) + (1 \Leftrightarrow \frac{2}{n} \cdot ([n/j] \Leftrightarrow 1)) \cdot \exp(\lambda^2 K \frac{a_j}{2}) \\ &\leq \frac{2}{j} + (1 \Leftrightarrow \frac{2}{j} + \frac{4}{n}) \cdot \exp(\lambda^2 K \frac{a_j}{2}) \\ &\leq \frac{2}{j} + (2 \Leftrightarrow \frac{2}{j}) \cdot \exp(\lambda^2 K \frac{a_j}{2}). \end{aligned}$$

Zusammen mit der Abschätzung

$$\begin{aligned} E(\exp(\lambda^2 K U_n)) &\leq \frac{1}{n} \cdot n \cdot \exp\left(\lambda^2 K \left(\left(\frac{n \Leftrightarrow 1}{n}\right)^2 \Leftrightarrow 1\right)\right) \\ &\leq \exp\left(\lambda^2 K \underbrace{\left(\left(1 \Leftrightarrow \frac{1}{b_j}\right)^2 \Leftrightarrow 1\right)}_{<0}\right) \end{aligned}$$

für $1 \leq n < \max\{4, j, n_0\} =: b_j$ liefert dies:

$$\begin{aligned} \sup_n E(\exp(\lambda^2 K U_n)) &\leq \sup_{1 \leq n < b_j} E(\exp(\lambda^2 K U_n)) + \sup_{n \geq b_j} E(\exp(\lambda^2 K U_n)) \\ &\leq \exp\left(\lambda^2 K \cdot \left(\left(1 \Leftrightarrow \frac{1}{b_j}\right)^2 \Leftrightarrow 1\right)\right) + \frac{2}{j} + \left(2 \Leftrightarrow \frac{2}{j}\right) \cdot \exp\left(\lambda^2 K \frac{a_j}{2}\right) \end{aligned}$$

und somit unter Berücksichtigung von (3.22)

$$\begin{aligned} \lim_{K \rightarrow \infty} \sup_n f(K, \lambda, n) &\leq \exp(|\lambda| \cdot c) \cdot \lim_{K \rightarrow \infty} E(\exp(\lambda^2 K U_n)) \\ &\leq \exp(|\lambda| \cdot c) \cdot \frac{2}{j}. \end{aligned}$$

Da dabei $j \in \mathbb{N}$ beliebig ist, folgt die Behauptung. \square

Nun sind wir in der Lage, in einem dritten Schritt den endgültigen Beweis von Satz 3.2.7 zu führen.

BEWEIS VON SATZ 3.2.7: Es sei $L > 0$ vorgegeben. Definiere K_L als das Maximum von K_1 aus Lemma 3.2.8 und $L \cdot K_2/L_1$, K_2 wie in Lemma 3.2.9. Ist nun $|\lambda| \leq L_1$, so folgt $f(K_L, \lambda, n) \leq 1$ aus Lemma 3.2.8 und der Monotonie von f in K (vgl. Lemma 3.2.10). Ist hingegen $L_1 \leq |\lambda| \leq L$, dann gilt für $\alpha := L/|\lambda| \geq 1$, $\lambda > 0$ die Abschätzung:

$$\begin{aligned} f(K_L, \lambda, n) &\leq (E(\exp(\alpha \cdot (\lambda C_n(Z_n) + \lambda^2 K_L U_n))))^{1/\alpha} \\ &= (E(\exp(L \cdot C_n(Z_n) + L \lambda K_L U_n)))^{1/\alpha} \\ &\leq (E(\exp(L \cdot C_n(Z_n) + L \lambda \frac{L \cdot K_2}{L_1} U_n)))^{1/\alpha} \\ &\leq (E(\exp(L \cdot C_n(Z_n) + L^2 K_2 U_n)))^{1/\alpha} \leq 1. \end{aligned}$$

Für die erste Zeile wurde dabei die Jensensche Ungleichung, für die dritte die Abschätzungen $K_L \geq L \cdot K_2/L_1$, $U_n \leq 0$ und für die letzte Zeile Lemma 3.2.9 benutzt. Da man eine analoge Abschätzung für $\lambda < 0$ erhält, ist der Beweis abgeschlossen. \square

Damit sind alle Vorbereitungen beendet und wir befinden uns quasi auf der Zielgeraden. Das eigentliche Ziel dieses Abschnitts, der Beweis von Satz 3.2.5, ist in greifbarer Nähe.

BEWEIS VON SATZ 3.2.5: Wie bereits angekündigt bildet das Kernstück dieses Beweises eine Induktion nach n . Sei dazu $L > 0$ vorgegeben und K_L wie in Satz 3.2.7 gewählt.

Für den Induktionsanfang $n = 1$ erhält man ohne Mühe

$$E(e^{\lambda Y_1}) = 1 \leq e^{\lambda^2 K_L} \text{ für alle } \lambda \in [\Leftrightarrow L, L],$$

wenn man $Y_1 = X_1 \Leftrightarrow EX_1 = 0$ beachtet. Für den Induktionsschritt $n \Leftrightarrow 1 \rightarrow n$ sei $\lambda \in [\Leftrightarrow L, L]$ vorgegeben. Für $j \in \{1, \dots, n \Leftrightarrow 1\}$ und $\lambda \in [\Leftrightarrow L, L]$ gelte $Ee^{\lambda Y_j} \leq e^{\lambda^2 K_L}$ (Induktionsvoraussetzung). Dann gilt unter Benutzung der rekursiven Gleichung $L(Y_n) = L(Y'_{Z_n-1} \frac{Z_n-1}{n} + \bar{Y}'_{n-Z_n} \frac{n-Z_n}{n} + C_n(Z_n))$ (vgl. (3.11)):

$$\begin{aligned} E(\exp(\lambda Y_n)) &= E\left(\exp\left(\lambda \cdot \left(Y'_{Z_n-1} \frac{Z_n \Leftrightarrow 1}{n} + \bar{Y}'_{n-Z_n} \frac{n \Leftrightarrow Z_n}{n} + C_n(Z_n)\right)\right)\right) \\ &= \sum_{i=1}^n \frac{1}{n} \cdot E\left(\exp\left(\lambda \cdot \left(Y'_{i-1} \cdot \frac{(i \Leftrightarrow 1)}{n} + \bar{Y}'_{n-i} \cdot \frac{n \Leftrightarrow i}{n} + C_n(i)\right)\right)\right) \\ &= \sum_{i=1}^n \frac{1}{n} \cdot E\left(\exp\left(\frac{\lambda \cdot (i \Leftrightarrow 1)}{n} \cdot Y'_{i-1}\right)\right) \cdot E\left(\exp\left(\frac{\lambda \cdot (n \Leftrightarrow i)}{n} \cdot \bar{Y}'_{n-i}\right)\right) \cdot \exp(\lambda C_n(i)) \\ &\stackrel{I.V.}{\leq} \sum_{i=1}^n \frac{1}{n} \cdot \exp\left(\frac{\lambda^2 (i \Leftrightarrow 1)^2}{n^2} \cdot K_L\right) \cdot \exp\left(\frac{\lambda^2 (n \Leftrightarrow i)^2}{n^2} \cdot K_L\right) \cdot \exp(\lambda \cdot C_n(i)) \\ &= E\left(\exp\left(\lambda^2 \cdot \left(\frac{Z_n \Leftrightarrow 1}{n}\right)^2 \cdot K_L + \lambda^2 \cdot \left(\frac{n \Leftrightarrow Z_n}{n}\right)^2 \cdot K_L + \lambda \cdot C_n(Z_n)\right)\right) \\ &= E(\exp(\lambda^2 K_L U_n + \lambda \cdot C_n(Z_n))) \\ &\leq \exp(\lambda^2 K_L) \cdot E(\exp(\lambda^2 K_L U_n + \lambda \cdot C_n(Z_n))) \\ &= \exp(\lambda^2 K_L) \cdot f(K_L, \lambda, n) \leq \exp(\lambda^2 K_L). \end{aligned}$$

Dabei wurde für die letzte Ungleichung Satz 3.2.7 verwendet. Der Beweis ist somit abgeschlossen. \square

3.2.8 Eine verbesserte Abschätzung für schlechtes Laufzeitverhalten

Wir haben bereits im vorletzten Abschnitt die Wahrscheinlichkeit für ungünstiges Verhalten (in bezug auf die benötigte Anzahl von Vergleichen) des Quicksort-Algorithmus abgeschätzt und zugleich darauf hingewiesen, daß das dortige Ergebnis einer Verbesserung bedarf. Mit Beendigung des letzten Abschnitts stehen uns nun die Mittel zur Verfügung, um dieses Vorhaben in die Tat umzusetzen.

In Satz 3.2.5 haben wir bereits gezeigt, daß $E(\exp(\lambda Y_n))$ in n gleichmäßig beschränkt ist durch $\exp(\lambda^2 K_{|\lambda|})$ für vorgegebenes $\lambda \in \mathbb{R}$.

Wir erhalten deshalb für $\lambda > 0$, $\varepsilon > 0$:

$$\begin{aligned} P(|X_n \ominus EX_n| \geq \varepsilon \cdot EX_n) &= P(\exp(\lambda|Y_n|) \geq \exp(\lambda\varepsilon EX_n/n)) \\ &\leq \frac{E(\exp(\lambda|Y_n|))}{\exp(\lambda\varepsilon EX_n/n)} = \frac{E(\exp(\lambda|Y_n|))}{n^{2\varepsilon\lambda} \cdot \exp(\lambda\varepsilon R(n)/n)} \leq \frac{\text{const}(\lambda, \varepsilon)}{n^{2\varepsilon\lambda}}, \end{aligned}$$

wobei man wiederum $\lim_{n \rightarrow \infty} \frac{R(n)}{n} \in \mathbb{R}$ beachte (vgl. Abschnitt 3.2.6). Wählt man speziell $\varepsilon = 2$, so gelangt man zu der für jedes $\lambda > 0$ und $n \in \mathbb{N}$ gültigen Ungleichung

$$P(X_n \geq 2E(X_n)) \leq \frac{\text{const}(\lambda)}{n^{2\lambda}}.$$

Diese Abschätzung stellt eine wesentliche Verbesserung des Resultats aus Abschnitt 3.2.6 dar, da die Funktion $n \mapsto n^{2\lambda}$ bedeutend schneller gegen unendlich konvergiert als die natürliche Logarithmusfunktion. Als Fazit können wir festhalten, daß schlechtes Laufzeitverhalten des Algorithmus Quicksort bei genügend großer Listenlänge extrem selten vorkommt, und daß folglich der Algorithmus verläßlich ist. Möchte man numerische Abschätzungen der Wahrscheinlichkeit durchführen, so muß man sich noch näher mit der Berechnung der Konstanten $\text{const}(\lambda, \varepsilon)$ beschäftigen. Dann ist es auch möglich, Aussagen darüber zu treffen, welche Größenordnung eine zu sortierende Liste haben sollte, um Quicksort als verläßlich einzustufen. Im Rahmen dieser Arbeit wollen wir darauf verzichten. Es sei nur noch darauf hingewiesen, daß man zeigen kann, daß für $\lambda > 0$ gilt: $E(\exp(\lambda|Y_n|)) \rightarrow E(\exp(\lambda|Y|))$ [\Leftrightarrow [20]].

Anhang A

Hilfsresultate

In diesem Kapitel werden einige Hilfsresultate zusammengestellt.

A.1 Abschätzungen und Konvergenzaussagen

A.1 Satz Seien $a, b \in \mathbb{R}_+$ und $p \in [1, \infty)$. Dann gilt: $(a + b)^p \geq a^p + b^p$.

BEWEIS: Für $p \in \mathbb{N}$ ergibt sich die Behauptung direkt aus dem binomischen Lehrsatz. Ist nun allgemein $p \in [1, \infty)$, so existieren $k \in \mathbb{N}$, $r \in [0, 1)$ mit $p = k + r$. Folglich erhält man:

$$\begin{aligned}(a + b)^p &= (a + b)^{k+r} = (a + b)^k \cdot (a + b)^r \geq (a^k + b^k) \cdot (a + b)^r \\ &= (a + b)^r \cdot a^k + (a + b)^r \cdot b^k \geq a^r a^k + b^r b^k = a^p + b^p.\end{aligned}$$

□

Anmerkung: Man kann den Beweis auch mittels der Jensenschen Ungleichung führen.

A.2 Satz Seien $a, b, c \in \mathbb{R}_+$ mit $c \geq b$ und $p \geq 1$. Dann gilt: $(a + b)^p + c^p \leq (a + c)^p + b^p$.

BEWEIS: Es seien $b, c \in \mathbb{R}_+$ beliebig, aber fest vorgegeben. Betrachte die Funktion $f_{(b,c)} : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto (x + c)^p \Leftrightarrow (x + b)^p + b^p \Leftrightarrow c^p$. Es gilt:

$$f'_{(b,c)}(x) = p \cdot (x + c)^{p-1} \Leftrightarrow p \cdot (x + b)^{p-1} = p \cdot ((x + c)^{p-1} \Leftrightarrow (x + b)^{p-1})$$

und somit $f'_{(b,c)}(x) \geq 0$ für alle $x \in \mathbb{R}$, denn $p \Leftrightarrow 1 \geq 0$ und $c \geq b$. Daraus folgt insbesondere, daß f monoton wachsend ist auf \mathbb{R}_+ , und zusammen mit $f_{(b,c)}(0) = 0$ liefert dies die Behauptung. □

A.3 Satz Es seien a_n, b_n , $n \in \mathbb{N}$ zwei Folgen mit Elementen aus \mathbb{R}_0^+ , für die für ein $2 \leq p \in \mathbb{N}$ gilt:

i) $\lim_{n \rightarrow \infty} b_n = 0$,

ii) $a_n \leq \left(\frac{2}{n} \cdot \sum_{i=1}^{n-1} \left(\frac{i}{n}\right)^p a_i^p\right)^{1/p} + b_n$.

Dann konvergiert die Folge a_n gegen 0.

BEWEIS: Für a_n erhält man die rekursive Abschätzung:

$$(A.1) \quad a_n \leq \left(\frac{2}{3}\right)^{1/p} \cdot \left(\sup_{1 \leq i \leq n-1} a_i\right) + b_n.$$

Diese ergibt sich aus folgender Rechnung:

$$\begin{aligned} a_n &\leq \left(\frac{2}{n} \cdot \sum_{i=1}^{n-1} \frac{i^p}{n^p} a_i^p\right)^{1/p} + b_n \leq \left(\frac{2}{n} \cdot \left(\sup_{1 \leq i \leq n-1} a_i^p\right) \cdot \sum_{i=1}^{n-1} \frac{i^p}{n^p}\right)^{1/p} + b_n \\ &\leq \left(\sup_{1 \leq i \leq n-1} a_i\right) \cdot \left(\frac{2}{n} \sum_{i=1}^{n-1} \frac{i^2}{n^2}\right)^{1/p} + b_n \\ &= \left(\sup_{1 \leq i \leq n-1} a_i\right) \cdot \left(\frac{2}{n} \cdot \frac{1}{n^2} \cdot \frac{(n \Leftrightarrow 1)n(2n \Leftrightarrow 1)}{6}\right)^{1/p} + b_n \\ &= \left(\sup_{1 \leq i \leq n-1} a_i\right) \cdot \left(\frac{2}{3} \Leftrightarrow \frac{1}{n} + \frac{1}{3n^2}\right)^{1/p} + b_n \leq \left(\frac{2}{3}\right)^{1/p} \cdot \left(\sup_{1 \leq i \leq n-1} a_i\right) + b_n. \end{aligned}$$

Diese rekursive Abschätzung impliziert insbesondere die Beschränktheit der Folge a_n . Da nämlich $\lim_{n \rightarrow \infty} b_n = 0$, existiert $n_0 \in \mathbb{N}$, so daß für alle $n \geq n_0$ gilt: $b_n \leq (1 \Leftrightarrow (2/3)^{1/p}) \cdot \sup_{1 \leq i \leq n-1} a_i$ (der Fall $a_n = 0$ für alle n ist trivial). Man erhält nun leicht mittels einer Induktion und unter Verwendung von (A.1) die Abschätzung $a_n \leq \sup_{1 \leq i \leq n_0-1} a_i$ für alle $n \in \mathbb{N}$. Definiere nun: $a := \limsup a_n < \infty$. Also: Für alle $\varepsilon > 0$ existiert ein $n_0 \in \mathbb{N}$, so daß für alle $n \geq n_0$ gilt: $a_n \leq a + \varepsilon$. Es folgt unter Verwendung von Voraussetzung ii) für fest vorgegebenes ε , $n_0(\varepsilon)$ und alle $n \geq n_0(\varepsilon)$:

$$\begin{aligned} a_n &\leq \left(\frac{2}{n} \cdot \sum_{i=1}^{n_0} \frac{i^p}{n^p} a_i^p\right)^{1/p} + \left(\frac{2}{n} \cdot \sum_{i=n_0+1}^{n-1} \frac{i^p}{n^p} (a + \varepsilon)^p\right)^{1/p} + b_n \\ &\leq \left(\frac{2}{n} \cdot \sum_{i=1}^{n_0} \frac{i^p}{n^p} a_i^p\right)^{1/p} + \left(\frac{2}{3}\right)^{1/p} (a + \varepsilon) + b_n \\ &= \left(\frac{2}{3}\right)^{1/p} (a + \varepsilon) + o(1). \end{aligned}$$

Dabei hat man für das erste Ungleichheitszeichen nur zu beachten, daß sich mit Hilfe von Satz A.1 leicht die Abschätzung $(x + y)^{1/p} \leq x^{1/p} + y^{1/p}$ für $x, y \in \mathbb{R}_0^+$ ergibt.

Man erhält nun:

$$(A.2) \quad 0 \leq a = \limsup_n a_n \leq \left(\frac{2}{3}\right)^{1/p} (a + \varepsilon),$$

woraus die Behauptung folgt, da (A.2) für alle $\varepsilon > 0$ Gültigkeit besitzt. \square

A.4 Satz (Zusammenhang zwischen „normalem“ und Cesaro-Limes) Sei $(a_i)_{i \in \mathbb{N}}$ eine Folge reeller Zahlen mit $\lim_{i \rightarrow \infty} a_i = c \in \mathbb{R}$. Dann gilt:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \cdot \sum_{i=1}^n a_i = c.$$

BEWEIS: Sei $\varepsilon > 0$ vorgegeben. Da $\lim_{i \rightarrow \infty} a_i = c$ existiert ein $n_0 \in \mathbb{N}$ mit $|a_i - c| \leq \varepsilon/2$ für alle $i \geq n_0$. Folglich gilt für $n \geq n_0$:

$$\begin{aligned} \left| \frac{1}{n} \cdot \sum_{i=1}^n a_i - c \right| &\leq \frac{1}{n} \cdot \sum_{i=1}^n |a_i - c| = \frac{1}{n} \cdot \sum_{i=1}^{n_0} |a_i - c| + \frac{1}{n} \cdot \sum_{i=n_0+1}^n |a_i - c| \\ &\leq \frac{1}{n} \cdot \sum_{i=1}^{n_0} |a_i - c| + \frac{1}{n} \cdot \sum_{i=n_0+1}^n \varepsilon/2 \leq \frac{1}{n} \cdot \sum_{i=1}^{n_0} |a_i - c| + \frac{\varepsilon}{2}. \end{aligned}$$

Nun existiert aber offensichtlich ein $n_1 \in \mathbb{N}$, so daß für alle $n \geq n_1$: $\frac{1}{n} \cdot \sum_{i=1}^{n_0} |a_i - c| \leq \frac{\varepsilon}{2}$. Folglich erhält man für $n \geq \max\{n_0, n_1\}$

$$\left| \frac{1}{n} \cdot \sum_{i=1}^n a_i - c \right| \leq \varepsilon$$

und somit die Behauptung, da ε beliebig gewählt war. \square

A.2 Die verallgemeinerte Höldersche Ungleichung

A.5 Satz Für je endlich viele meßbare numerische Funktionen f_1, \dots, f_n auf einem Maßraum und zugehörige reelle Zahlen $p_1 > 1, \dots, p_n > 1$ mit $\sum_{j=1}^n p_j^{-1} = 1$ gilt die verallgemeinerte Höldersche Ungleichung:

$$\|f_1 \cdot \dots \cdot f_n\|_1 \leq \|f_1\|_{p_1} \cdot \dots \cdot \|f_n\|_{p_n}.$$

BEWEIS: Eine Induktion nach n ist hier der Schlüssel zum Erfolg. Den Induktionsanfang $n = 2$ liefert dabei die bekannte Höldersche Ungleichung. Wenden wir uns also direkt dem Induktionsschritt $n \rightarrow n + 1$ zu. Seien dazu f_1, \dots, f_{n+1} meßbare numerische Funktionen und p_1, \dots, p_{n+1} entsprechend gewählte reelle Zahlen. Nach Induktionsvoraussetzung gilt nun: Sind für $k \in \{2, \dots, n\}$ g_1, \dots, g_k meßbare numerische Funktionen und $q_1 > 1, \dots, q_k > 1$ reelle Zahlen mit $\sum_{i=1}^k (q_i)^{-1} = 1$, so ist $\|g_1 \cdot \dots \cdot g_k\|_1 \leq \|g_1\|_{q_1} \cdot \dots \cdot \|g_k\|_{q_k}$. Unter Verwendung dieser Induktionsvoraussetzung für den Fall $k = 2$ können wir deshalb abschätzen:

$$(A.3) \quad \|f_1 \cdot \dots \cdot f_{n+1}\|_1 \leq \|f_1\|_{p_1} \cdot \|f_2 \cdot \dots \cdot f_{n+1}\|_{1/\sum_{i=2}^{n+1} p_i^{-1}}.$$

Definieren wir nun

$$q_1 := p_2 \cdot \sum_{i=2}^{n+1} p_i^{-1}, \dots, q_n := p_{n+1} \cdot \sum_{i=2}^{n+1} p_i^{-1},$$

so gilt offensichtlich $\sum_{i=1}^n (q_i)^{-1} = 1$, woraus weiter unter Benutzung der Induktionsvoraussetzung folgt:

$$\begin{aligned} \text{(A.4)} \quad \|f_2 \cdot \dots \cdot f_{n+1}\|_{1/\sum_{i=2}^{n+1} p_i^{-1}} &= \left(\|f_2^{p_2/q_1} \cdot \dots \cdot f_{n+1}^{p_{n+1}/q_n}\|_1 \right)^{\sum_{i=2}^{n+1} p_i^{-1}} \\ &\leq \left(\|f_2^{p_2/q_1}\|_{q_1} \cdot \dots \cdot \|f_{n+1}^{p_{n+1}/q_n}\|_{q_n} \right)^{\sum_{i=2}^{n+1} p_i^{-1}} \\ &= \left(\|f_2\|_{p_2}^{p_2/q_1} \cdot \dots \cdot \|f_{n+1}\|_{p_{n+1}}^{p_{n+1}/q_n} \right)^{\sum_{i=2}^{n+1} p_i^{-1}} \\ &= \|f_2\|_{p_2} \cdot \dots \cdot \|f_{n+1}\|_{p_{n+1}}. \end{aligned}$$

(A.3) und (A.4) zusammen ergeben die Behauptung. □

A.3 Die Bernoullischen Zahlen

Entwickelt man die Funktion $f(x) = \frac{x}{e^x - 1}$ in eine Potenzreihe $\sum_{n=0}^{\infty} b_n x^n$ um $x = 0$ (zur Existenz vgl. 7.14 in [24]), so gelangt man zu der Darstellung

$$f(x) = \sum_{n=0}^{\infty} \frac{B_n}{n!} x^n,$$

wenn man $B_n := n!b_n$ setzt. Die B_n werden Bernoullische Zahlen genannt und sind nichts anderes als die Koeffizienten der Potenzreihenentwicklung multipliziert mit $n!$. Multiplikation der Reihen für f und $1/f = (1/x)(e^x \Leftrightarrow 1)$ und Koeffizientenvergleich führt auf

$$1 = \sum_{n=0}^{\infty} \frac{B_n}{n!} x^n \cdot \sum_{n=0}^{\infty} \frac{x^n}{(n+1)!} = \sum_{n=0}^{\infty} c_n x^n$$

mit $c_0 = 1$ und $c_1 = c_2 = \dots = 0$. Es ist

$$c_n = \sum_{i=0}^n \frac{B_i}{i!(n+1 \Leftrightarrow i)!} = \frac{1}{(n+1)!} \cdot \sum_{i=0}^n \binom{n+1}{i} B_i,$$

woraus sich der Reihe nach die Bestimmungsgleichungen

$$\begin{aligned} B_0 &= 1 \\ B_0 + 2B_1 &= 0 \\ B_0 + 3B_1 + 3B_2 &= 0 \\ &\dots \end{aligned}$$

und allgemein

$$(A.5) \quad \sum_{i=0}^n \binom{n+1}{i} B_i = 0 \quad \text{für } n = 1, 2, 3, \dots$$

ergeben. Für die ersten zehn Zahlen können wir festhalten:

$$B_0 = 1, \quad B_1 = -1/2, \quad B_2 = 1/6, \quad B_3 = 0, \quad B_4 = -1/30, \quad B_5 = 0, \\ B_6 = 1/42, \quad B_7 = 0, \quad B_8 = -1/30, \quad B_9 = 0, \quad B_{10} = 5/66.$$

Ihrer Berechnungsweise nach müssen alle Bernoullischen Zahlen rational sein. Zudem kann man zeigen (wie die obigen Beispiele bereits vermuten lassen), daß

$$B_{2n+1} = 0 \quad \text{für } n = 1, 2, \dots$$

Vergleiche dazu etwa [13].

A.4 Die Bernoullischen Polynome und die Eulersche Summenformel

Die Bernoullischen Polynome $B_n(x)$ sind für $n \in \mathbb{N}$ mit Hilfe der Bernoullischen Zahlen B_n gemäß der Formel

$$B_n(x) := \sum_{k=0}^n \binom{n}{k} B_k x^{n-k} = x^n + \binom{n}{1} B_1 x^{n-1} + \binom{n}{2} B_2 x^{n-2} + \dots + B_n$$

definiert. Wegen $(B_0, B_1, B_2, B_3) = (1, -1/2, 1/6, 0)$ ist

$$B_0(x) = 1, \quad B_1(x) = x - \frac{1}{2}, \quad B_2(x) = x^2 - x + \frac{1}{6}, \quad B_3(x) = x^3 - \frac{3}{2}x^2 + \frac{1}{2}x.$$

Betrachtet man die Bernoullischen Polynome nur im Intervall $[0, 1)$ und setzt sie periodisch mit der Periode 1 fort, so gelangt man zu Funktionen $C_n(x)$ mit $C_n(x) := B_n(x - [x])$. Da $B_n(0) = B_n(1) = B_n$ für $n \neq 1$, wie man sofort anhand der Definition der Bernoullischen Polynome und der rekursiven Beziehung (A.5) ablesen kann, ist $C_n(x)$ für $n \neq 1$ eine in \mathbb{R} stetige Funktion. Mit Hilfe dieser Funktionen gelangt man zur Eulerschen Summenformel:

A.6 Satz (Eulersche Summenformel) Für $p \in \mathbb{N}$ und $f \in C^{2p+1}[0, n]$ gilt:

$$(A.6) \quad f(0) + f(1) + \dots + f(n) = \int_0^n f(x) dx + \frac{1}{2}(f(0) + f(n)) \\ + \left[\frac{B_2}{2!} f' + \frac{B_4}{4!} f''' + \dots + \frac{B_{2p}}{(2p)!} f^{(2p-1)} \right]_0^n + R_p$$

mit

$$R_p = \frac{1}{(2p+1)!} \int_0^n f^{(2p+1)}(x) C_{2p+1}(x) dx.$$

Ein Beweis findet sich zum Beispiel in [24].

Quellenverzeichnis

Satz 1.2	⇔ [3], S. 146
Beweis von Satz 2.2.1	⇔ [7]
Satz 2.2.3	⇔ [7], S. 289
Satz 2.2.13	⇔ [9], Theorem 2, S 469
Satz 2.2.14	⇔ [11], Satz 8.14, S. 329
Lemma 2.3.1	⇔ [19], Lemma 1, S. 198
Lemma 2.3.4	⇔ [19], Lemma 2, S. 199
Satz 2.3.5	⇔ [19], Theorem 3, S. 200
Satz 2.3.6	⇔ [19], Theorem 4, S. 200
Satz 2.3.7	⇔ [19], Remark, S. 200f
Satz 2.3.9	⇔ [19], Theorem 5, S. 202
Satz 2.3.11	⇔ [19], Theorem 8, S. 205
Lemma 2.3.12	⇔ [4], Beweis zu Satz 14.1
Satz 2.3.14	⇔ [19], Theorem 9, S. 206
Satz 2.3.15	⇔ [19], Theorem 10, S 208
Bemerkung 2.3.17	⇔ [19], S. 209
Satz 2.3.18	⇔ [19], Theorem 6, S. 203
Satz 2.3.19	⇔ [19], Theorem 7, S. 205
Lemma 3.2.1	⇔ [6], Lemma 6.1

Satz 3.2.3	⇔ [20], Theorem 3.1, S. 92
Satz 3.2.5	⇔ [20], Lemma 4.1, S. 95
Satz 3.2.7	⇔ [20], (4.2), S. 96
Lemma 3.2.8	⇔ [20], Claim 1, S. 96
Lemma 3.2.9	⇔ [20], Claim 2, S. 96
Satz A.3	⇔ [20], Proposition 3.3, S. 93

Ein abschließender Blick auf den Quelltext

In der nachfolgenden Tabelle sind einige Fehler in der Arbeit „A fixed point theorem for distributions“ von U. Rösler aufgelistet:

U. Rösler, Distribution fix points [⇔ [19]]	A. Hirsch, Fixpunktsätze für Verteilungen	Art des Fehlers bzw. durchgeführte Veränderung
Lemma 1, S. 198, Unterpunkt (3)	Lemma 2.3.1, (3) und (4)	Aussage ist nicht richtig
Beweis von Theorem 4, S. 200	Beweis von Satz 2.3.6	Der Beweis in der ursprünglichen Version ist falsch.
Beweis von Theorem 5, S. 203	Beweis von Satz 2.3.9	Die Gleichheit $\text{Var}(\bar{S}^n(F)) =$ $E(\sum_i T_i^2)\text{Var}(\bar{S}^{n-1}(F))$ ist nicht richtig.
Theorem 6, S. 203	Satz 2.3.18	Voraussetzung $\lim_{N \rightarrow \infty} E(e^{\lambda \sum_{i=1}^N T_i}) < 1$ geändert in $E(\exp(\lambda \sum_i T_i)) < \infty$
Beweis von Theorem 6, S. 204	Beweis von Satz 2.3.18	die Gleichheit $\frac{d^2}{d\lambda^2} f_K(0) = \dots$ ist nicht richtig
Beweis von Theorem 8, S. 205	Beweis von Satz 2.3.11	zweiten Teil des Beweises verändert
Theorem 9, S. 206	Satz 2.3.14	Voraussetzung $\ \sum_i T_i \ _p < \infty$ eingefügt, um Existenz der Abbildung S zu gewährleisten

Beweis von Theorem 9, S. 207	Beweis von Satz 2.3.14	$d_q(S^m(F), S^m(G))$ umgewandelt in $d_q(S^m(F), S^{m-1}(F))$
Beweis von Theorem 9, S. 207	Beweis von Satz 2.3.14	$\leq E(\sum_i T_i Y_{i,m})^{n+1} \dots$ hier fehlen Betragstriche, wodurch bestimmte Terme innerhalb der nachfolgenden Rechnung nicht wegfallen
Theorem 10, S. 208	Satz 2.3.15	Voraussetzung $\ \sum_i T_i\ _p < \infty$ umgewandelt in $\ \sum_i T_i \ _p < \infty$, um Existenz der Abb. S zu gewährleisten
Beweis von Theorem 10, S. 208	Beweis von Satz 2.3.15	komplett verändert; vgl. auch Theorem 4 und geänderte Version

Literaturverzeichnis

- [1] Alsmeyer, G. *Script zur Vorlesung über Stochastische Prozesse*, WWU Münster
- [2] Alsmeyer, G. *Wahrscheinlichkeitstheorie*, Skripten zur Mathematischen Statistik Nr.30 (1998)
- [3] Barner, M., Flohr, F. *Analysis II*, de Gruyter Lehrbuch (1989), 2. Auflage
- [4] Bauer, H. *Maß- und Integrationstheorie*, de Gruyter Lehrbuch, 2. Auflage (1992).
- [5] Bickel, P.J. and Freedman, D.A. *Some asymptotic theory for the bootstrap*. Ann. Probab. **9** (1981), 1196-1217.
- [6] Bruhn, V. *Eine Methode zur asymptotischen Behandlung einer Klasse von Rekursionsgleichungen mit einer Anwendung in der stochastischen Analyse des Quicksort-Algorithmus*. Dissertation zu Erlangung des Doktorgrades, 1996.
- [7] Cambanis, S., Simon, G., Stout, W. *Inequalities for $E_k(X, Y)$ when the Marginals are Fixed*. Z. Wahrsch. Verw. Gebiete **36** (1976), 285-294.
- [8] Chow, Y.S., Teicher, H. *Probability Theory*, Springer-Verlag (1978).
- [9] Dobrushin, R.L. *Describing a system of random variables by conditional distributions*. Theory Probab. Appl. **15** (1970), 458-486.
- [10] Fréchet, M. *Sur les tableaux de corrélation dont les marges sont données*. Ann. Univ. Lyon, Sect. A. Sér. **3**,14,(1951),53-77
- [11] Gänszler, P., Stute, W. *Wahrscheinlichkeitstheorie*, Springer-Verlag (1977).
- [12] Heuser, H. *Funktionalanalysis*, B.G. Teubner Stuttgart, 3. Auflage (1986).
- [13] Heuser, H. *Lehrbuch der Analysis, Teil 1*, B.G. Teubner Stuttgart (1993), 10. Auflage
- [14] Hoare, C.A.R. *Quicksort*. The Computer Journal **5** (1962), 10-15.

- [15] Hoeffding, W. *Maßstabsvariante Korrelationstheorie*. Schr. Math. Inst. Univ. Berlin **5** (1940), 181-233
- [16] Loeser, R. *Some Performance Tests of Quicksort and Descendants*. Communications of The Association for Computing Machinery **17** (1974), 143-152.
- [17] Major, P. *On the invariance principle for sums of independent, identically distributed random variables*. Jour. Of Multivariate Anal. **8** (1978), 487-501.
- [18] Mallows, C.L. *A note on asymptotic joint normality*. Ann. Math. Statist. **43** (1972), 508-515.
- [19] Rösler, U. *A fixed point theorem for distributions*. Stochastic Processes and their Applications **42** (1992), 195-214.
- [20] Rösler, U. *A limit theorem for Quicksort*. Inform. Théor. Appl./ Theor. Inform. Appl. **25** (1991), 85-100.
- [21] Sedgewick, R. *The analysis of Quicksort programs*. Acta Inform. **7** (1977), 327-355.
- [22] Tanaka, H. *An inequality for a functional of probability distribution and its application to Kac's one dimensional model of a Maxwellian gas*, Z. Wahrsch. Verw. Gebiete **27** (1973), 47-52.
- [23] Vallender, S.S. *Calculation of the Wasserstein distance between probability distributions on the line*. Theor. Probability Appl. **18** (1973), 184-186.
- [24] Walter, W. *Analysis 1*, Springer-Verlag (1992), 3. Auflage.

Hiermit versichere ich, daß ich diese Arbeit selbständig verfaßt und keine anderen als die angegebenen Hilfsmittel benutzt habe. Alle Stellen, die den Ausführungen anderer Autoren wörtlich oder sinngemäß entnommen sind, habe ich durch Angabe der Quellen als solche kenntlich gemacht.

Münster, den 10. 12. 1998

