Iterierte Funktionensysteme und Fraktale

Konrad Kiener

Author address:

(K. Kiener) Institut für Mathematik, Johannes Kepler Universität, A–4040 Linz, Österreich

Inhalt

KAPITEL 1

Metrische Räume

1. Notation und Beispiele

(X, d) bezeiche einen metrischen Raum mit zugrundeliegender Menge X und mit der Metrik $d : X \times X \to [0, \infty)$.

Beispiel 1.1.
Der m–dimensionale euklidische Raum $(\mathbf{R}^m, |\,.\,|)$. Die Metrik ist durch den euklidischen Abstand für $x = (\xi_1, \dots, \xi_m)$, $y = (\eta_1, \dots, \eta_m) \in \mathbf{R}^m$ definiert durch

$$|x - y| := d(x, y) := \left(\sum_{i=1}^{m} |\xi_i - \eta_i|^2 \right)^{\frac{1}{2}}.$$

Beispiel 1.2.
Der Adreß–Raum (Σ, d) über $N \geq 2$ Symbolen, etwa $\{0, 1, \dots, N-1\}$ ist definiert als die Menge $\Sigma = \prod_{i=1}^{\infty} \{0, 1, \dots, N-1\}$ der Folgen $\sigma = \sigma_1 \sigma_2 \dots$ mit $\sigma_i \in \{0, 1, \dots, N-1\}$ und mit der Metrik

$$d(\sigma, \omega) := \sum_{i=1}^{\infty} \frac{|\sigma_i - \omega_i|}{(N+1)^i}.$$

Definition 1.1. Zwei Metriken d_1, d_2 auf X heißen äquivalent, wenn das Streckungsverhältnis von Abständen, gemessen in beiden Metriken global von 0 und ∞ weg beschränkt ist:

$$\bigwedge_{x \neq y \in X} \qquad 0 < c_1 \leq \frac{d_2(x, y)}{d_1(x, y)} \leq c_2 < \infty.$$

Definition 1.2. Zwei metrische Räume $(X_1, d_1), (X_2, d_2)$ heißen äquivalent, falls eine bijektive Abbildung $h : X_1 \to X_2$ existiert, für die durch $d_2(h(x), h(y))$ eine zu d_1 äquivalente Metrik auf X_1 definiert ist.

Bemerkung 1.1. Auf $\mathbf{R}^m$ sind alle, von Normen induzierten Metriken äquivalent. Definieren wir auf $X = (0, 1]$ die Metriken $d_1(x, y) = |x - y|$ und $d_2(x, y) = |1/x - 1/y|$, dann sind d_1 und d_2 nicht äquivalent. Für $X_1 = [1, 2]$, $X_2 = [0, 1]$, $d_1 = |x - y|$, $d_2 = 2|x - y|$ sind (X_1, d_1), (X_2, d_2) metrisch äquivalente Räume.

Bemerkung 1.2. d_1, d_2 sind äquivalente Metriken auf X genau dann, wenn die identische Abbildung $1_X : X \to X$ die metrische Äquivalenz zwischen den Räumen (X, d_1), (X, d_2) vermittelt. Jede metrische Äquivalenz ist ein Homöomorphismus. Daher sind alle "topologischen" Eigenschaften (z.B. offen, abgeschlossen, kompakt, zusammenhängend usw.) invariant unter metrischer Äquivalenz. Wir werden später sehen, daß ebenso die fraktale Dimension zu den "metrischen Invarianten" gehört. Auch der Begriff der Vollständigkeit ist hier zu erwähnen. Es seien kurz die relevanten Definitionen wiederholt.

Definition 1.3. Eine Folge $\{x_n\}_1^\infty$ im metrischen Raumes (X,d) heißt Cauchy–Folge, falls

$$\bigwedge_{\epsilon>0}\bigvee_{N>0}\bigwedge_{n,m>N} d(x_m,x_n)<\epsilon.$$

(X,d) heißt vollständig , falls jede Cauchy–Folge $\{x_n\}_1^\infty$ in (X,d) gegen ein $x\in X$ konvergiert, in Zeichen:

$$\bigvee_{x\in X} x=\lim_{n\to\infty} x_n.$$

Satz 1.1. *Der Adreß–Raum (Σ,d) über N Symbolen ist vollständig.*

Wir beweisen dazu und für spätere Zwecke folgendes

Lemma 1.2. *Für beliebige $\sigma,\omega\in\Sigma$ gilt*

$$d(\sigma,\omega)<\frac{1}{(N+1)^n}\iff\bigwedge_{1\le i\le n}\sigma_i=\omega_i.$$

Beweis. Die Bedingung ist hinreichend:

$$\begin{aligned}d(\sigma,\omega)&=\sum_{i=1}^\infty\frac{|\sigma_i-\omega_i|}{(N+1)^i}\\&=\sum_{i=n+1}^\infty\frac{|\sigma_i-\omega_i|}{(N+1)^i}\le\frac{N-1}{(N+1)^{n+1}}\left(1+\frac{1}{N+1}+\frac{1}{(N+1)^2}+\dots\right)\\&=\frac{N-1}{(N+1)^{n+1}}\,\frac{1}{1-1/(N+1)}<\frac{1}{(N+1)^n}.\end{aligned}$$

Die Bedingung ist notwendig: Ist nämlich für $1\le i_0\le n \quad \sigma_{i_0}\ne\omega_{i_0}$, dann folgt

$$\begin{aligned}d(\sigma,\omega)&\ge\frac{|\sigma_{i_0}-\omega_{i_0}|}{(N+1)^{i_0}}+\sum_{i=n+1}^\infty\frac{|\sigma_i-\omega_i|}{(N+1)^i}\\&\ge\frac{1}{(N+1)^{i_0}}+\sum_{i=n+1}^\infty\frac{|\sigma_i-\omega_i|}{(N+1)^i}\ge\frac{1}{(N+1)^n}.\end{aligned}$$

□

Auf Σ haben wir zwei Topologien zur Verfügung; einerseits die Produkttopologie τ_1 der diskreten Räume $\{0,1,\dots,N-1\}$, andererseits die metrische Topologie τ_2. Beide Topologien stimmen jedoch überein: Ist nämlich U eine offene τ_1–Umgebung von $\omega=\omega_1\omega_2\dots$, dann umfaßt sie eine Umgebung von ω aus der Basis der Produkttopologie, also eine Menge der Form $\prod_{n=1}^\infty\Omega_n$, wobei nur für endlich viele n_i, etwa $n_1,\dots,n_k \quad \Omega_{n_i}=\{\omega_{n_i}\}$ gilt, für alle anderen n jedoch $\Omega_n=\{0,1,\dots,N-1\}$. Sei $n_0=Max\{n_1,\dots,n_k\}$. Dann ist $U':=\{\sigma:\sigma_i=\omega_i, 1\le i\le n_0, \sigma_i \text{ beliebig, sonst}\}\subset U$ und $U'=\{\sigma:d(\sigma,\omega)<1/(N+1)^{i_0}\}$ τ_2–Umgebung von ω. Umgekehrt gilt für $1/(N+1)^{i_0}\le\epsilon$:

$$U_\epsilon(\omega):=\{\sigma:d(\sigma,\omega)<\epsilon\}\supset\left\{\sigma:d(\sigma,\omega)<\frac{1}{(N+1)^{i_0}}\right\}$$

und dies ist nach dem vorigen Lemma eine τ_1–Umgebung.

Daraus ergibt sich nun die Behauptung von Satz 1.1, da ein kompakter metrischer Raum vollständig ist.

Eine Teilmenge $S \subset X$ des metrischen Raumes (X, d) heißt dabei *kompakt*, falls jede unendliche Folge $\{x_n \in S\}_1^\infty$ eine in S konvergente Teilfolge besitzt. In einem metrischen Raum ist dies bekanntlich gleichbedeutend mit der Tatsache, daß jede offene Überdeckung von S eine endliche Teilüberdeckung enthält.

Eine Teilmenge $S \subset X$ heißt *beschränkt*, falls

$$\bigvee_{a \in X} \bigvee_{R>0} \bigwedge_{x \in S} \qquad d(a, x) \leq R.$$

Eine Teilmenge $S \subset X$ heißt *totalbeschränkt*, falls

$$\bigwedge_{\epsilon>0} \quad \bigvee_{\{y_1,\ldots,y_n\} \subset S} \bigwedge_{x \in S} \bigvee_{i} \qquad d(y_i, x) \leq \epsilon.$$

Anders ausgedrückt: zu jedem $\epsilon > 0$ existiert ein endliches ϵ–Netz. Der *Durchmesser* einer Menge S ist definiert als

$$diam(S) := |S| := sup\{d(x, y) : x, y \in S\}.$$

SATZ 1.3. *Sei (X, d) ein vollständiger, metrischer Raum. Eine Teilmenge $S \subset X$ ist genau dann kompakt, wenn S abgeschlossen und totalbeschränkt ist.*

BEWEIS. Sei S abgeschlossen und totalbeschränkt und $\{x_i\}_1^\infty$ sei eine unendliche Folge in S. Da S total beschränkt ist, gibt es eine endliche Menge von abgeschlossenen Kugeln vom Radius $1/2$, deren Vereinigung $\{x_i\}_1^\infty$ enthält. Wenigstens eine Kugel, etwa B_1 enthält nach dem Schubfachprinzip unendlich viele der x_i. Wähle $x_{N_1} \in S_1 := B_1 \cap S$. S_1 ist als Teilmenge einer total beschränkten Menge ebenfalls total beschränkt. Wir können daher S_1 mit endlich vielen abgeschlossenen Kugeln vom Radius $1/4$ überdecken. Eine von diesen, etwa B_2 enthält wieder unendlich viele der x_i. Wähle $N_2 > N_1$ mit $x_{N_2} \in S_2 := B_2 \cap S$. So fortfahrend konstruieren wir eine geschachtelte Folge $S_1 \supset S_2 \supset \cdots \supset S_N \supset \ldots$ von abgeschlossenen Teilmengen von S mit Durchmessern $\leq 1, 1/2, \ldots, 1/2^{N-1}, \ldots$ und eine Teilfolge $x_{N_i} \in S_i$. Dann ist x_{N_i} eine Cauchyfolge und hat, da X vollständig ist, einen Limes $\{x\} = \cap S_i$, der wegen der Abgeschlossenheit von S zu S gehört.

Sei umgekehrt S kompakt, im Widerspruch zur Behauptung aber nicht totalbeschränkt. Dann gibt es ein $\epsilon > 0$, zu dem kein endliches ϵ–Netz existiert. Dann aber existiert eine unendliche Folge $\{x_i\} \subset S$, sodaß für alle $i \neq j \quad d(x_i, x_j) > \epsilon$. Andererseits enthält $\{x_i\}$ eine konvergente Teilfolge $x_{N_1}, x_{N_2}, \cdots \to x \epsilon S$, d.h. ab einem gewissen Index gilt $d(x_{N_i}, x_{N_j}) < \epsilon$, im Widerspruch zu $d(x_i, x_j) \geq \epsilon$.

$\square$

Wir werden häufig Abbildungen $f : X \to X$ betrachten. Wir nennen sie *Transformationen* auf X. Zu $S \subset X$ ist das Bild von S unter der Abbildung f definert als $f(S) := \{f(x) : x \in S\}$. f heißt *injektiv*, falls verschiedene Punkte in verschiedene Punkte abgebildet werden. f heißt *surjektiv*, falls $f(X) = X$. f heißt *bijektiv*, oder *umkehrbar*, falls f injektiv und surjektiv ist. Die *Vorwärtsiterierte* von f ist definiert durch $f^0 := 1_X$ (identische Abbildung $X \to X$) und $f^{n+1} := f \circ (f^n)$. Zu $x \in X$ heißt $\{f^n(x) : n = 0, 1, 2, \ldots\}$ der *Orbit* von x.

Sei $f : X \to X$ eine Transformation auf dem metrischen Raum (X, d). Die *Lipschitz-Konstante* von f ist definiert als

$$Lip\, f := sup_{x \neq y} \frac{d(f(x), f(y))}{d(x, y)}.$$

Falls $Lip\, f = \lambda$, dann gilt offenbar für alle $x, y \in X$ $d(f(x), f(y)) \leq \lambda d(x, y)$. $Lip\, f$ ist die kleinste der Zahlen λ mit dieser Eigenschaft. f heißt *Lipschitz-Funktion* falls $Lip\, f < \infty$. f heißt *Kontraktion*, falls $Lip\, f < 1$. Wir bemerken, daß eine Lipschitz-Funktion beschränkte Mengen in beschränkte Mengen abbildet.

Von besonderer Wichtigkeit ist für uns im Folgenden der

Satz 1.4 (Fixpunktsatz von S. Banach). *Sei $f : X \to X$ eine Kontraktion auf dem vollständigen metrischen Raum (X, d). Dann hat f einen eindeutig bestimmten Fixpunkt x_f, d.h. $f(x_f) = x_f$ und es gilt*

$$\bigwedge_{x \in X} \qquad \lim_{n \to \infty} f^n(x) = x_f.$$

Beweis. In fast jedem Standardtext über Analysis. □

Definition 1.4. Eine Transformation $S : X \to X$ heißt Ähnlichkeit, falls ein $0 < r < \infty$ existiert, sodaß

$$\bigwedge_{x,y \in X} \qquad d(S(x), S(y)) = r.d(x, y).$$

$h_r : \mathbf{R}^m \to \mathbf{R}^m$ heißt Homothetie, falls $h_r(x) = r\,x$. $t_b : \mathbf{R}^m \to \mathbf{R}^m$ heißt Translation, falls $t_b(x) = x - b$. $w : \mathbf{R}^m \to \mathbf{R}^m$ heißt affine Transformation, falls $w(x) = Ax + b$. Dabei ist $A : \mathbf{R}^m \to \mathbf{R}^m$ linear und $b \in \mathbf{R}^m$.

Satz 1.5. *$S : \mathbf{R}^m \to \mathbf{R}^m$ ist Ähnlichkeit genau dann, wenn S die Zusammensetzung einer orthogonalen Transformation g, einer Translation t_b und einer Homothetie h_r ist: $S = h_r \circ t_b \circ g$.*

Beweis. Es ist klar, daß die Bedingung hinreichend ist. Um einzusehen, daß sie notwendig ist, sei S eine Ähnlichkeit und $Lip\, S = r \neq 0$. Wir setzen $g(x) = \frac{1}{r}(S(x) - S(0))$. Dann ist g eine Isometrie mit dem Fixpunkt 0. Nun ist

$$\begin{aligned} < x, y > &= \frac{1}{2}(|x|^2 + |y|^2 - |x - y|^2) \\ &= \frac{1}{2}\left(d(0, x)^2 + d(0, y)^2 - d(x, y)^2\right) \\ &= \frac{1}{2}\left(|g(x)|^2 + |g(y)|^2 - |g(x) - g(y)|^2\right) =< g(x), g(y) > . \end{aligned}$$

g erhält also innere Produkte. Sei $\{e_i : 1 \leq i \leq m\}$ eine ON-Basis für $\mathbf{R}^m$. Dann ist nach dem vorigen auch $\{g(e_i) : 1 \leq i \leq m\}$ eine ON-Basis und daher

$$g(x) = \sum_1^m < g(x), g(e_i) > g(e_i) = \sum_1^m < x, e_i > g(e_i).$$

Daher ist g linear und somit orthogonal. Es ist

$$S(x) = rg(x) + S(0) = r\left(g(x) + \frac{1}{r}S(0)\right).$$

Das heißt $S = h_r \circ t_{-\frac{1}{r}S(0)} \circ g$. $\square$

Sei $w : \mathbf{R}^2 \to \mathbf{R}^2$ eine affine Transformation. Das bedeutet also: es gibt $a, b, c, d, e, f \in \mathbf{R}$, sodaß für $(x_1, x_2) \in \mathbf{R}^2$

$$w(x_1, x_2) = (ax_1 + bx_2 + e, cx_1 + dx_2 + f)$$

oder in Matrixschreibweise

$$w(x) = w \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} e \\ f \end{bmatrix} = Ax + t.$$

Dabei sind unter der Wirkung von w $(a, c), (b, d), (e, f)$ die Bilder von resp. $(1, 0)$, $(0, 1)$, $(0, 0)$. Sind (r_1, θ_1) und $(r_2, \theta_2 + \pi/2)$ die Polarkoordinaten von resp. (a, c) und (b, d), dann gilt

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} r_1 \cos\theta_1 & -r_2 \sin\theta_2 \\ r_1 \sin\theta_1 & r_2 \cos\theta_2 \end{bmatrix}.$$

Umgekehrt kann man A und t einer affinen Transformation $w(x) = Ax + t$ auf $\mathbf{R}^2$ aus der Kenntnis der Bilder von 3 nicht kollinearen Punkten berechnen. Aus

$$\begin{aligned} (x_1, x_2) &\longmapsto (\tilde{x}_1, \tilde{x}_2) \\ (y_1, y_2) &\longmapsto (\tilde{y}_1, \tilde{y}_2) \\ (z_1, z_2) &\longmapsto (\tilde{z}_1, \tilde{z}_2) \end{aligned}$$

folgen nämlich die Gleichungen

$$\begin{aligned} \tilde{x}_1 &= x_1 a + x_2 b + e \\ \tilde{x}_2 &= x_1 c + x_2 d + f \\ \tilde{y}_1 &= y_1 a + y_2 b + e \\ \tilde{y}_2 &= y_1 c + y_2 d + f \\ \tilde{z}_1 &= z_1 a + z_2 b + e \\ \tilde{z}_2 &= z_1 c + z_2 d + f. \end{aligned}$$

Dies sind 6 Gleichungen für die 6 Unbekannten a, b, c, d, e, f.

Die Bedeutung affiner Transformationen liegt bei unserem Thema in der Tatsache, daß sich zahlreiche Formen der Natur mit ihrer Hilfe ineinander überführen lassen (vgl. z.B. [**Th**] Chapter IX, *On the theory of transformations, or the comparison of related forms*). Zur Illustration versuche man eine affine Transformation zu bestimmen, die entsprechend der folgenden Skizze annäherungsweise das große Ahornblatt in das kleine überführt.

ABBILDUNG 1. Aus [Ba], p.54

KAPITEL 2

Fraktale Mengen, $\mathcal{H}(X)$

1. $(\mathcal{H}(X), h)$, Hausdoff–Metrik

Sei (X, d) ein metrischer Raum. Mit $\mathcal{H}(X)$ bezeichnen wir die Menge der nichtleeren, kompakten Teilmengen von X

$$\mathcal{H}(X) = \{A \subset X : \emptyset \neq A \text{ kompakt}\}.$$

Als nichttriviale Beispiele seien in Hinblick auf unser Thema die *Cantor–Menge* $\mathcal{C}$ und das *Sierpinski–Dreieck* $\mathcal{S}$ erwähnt. Beide Mengen entstehen als Grenzwert eines ähnlichen Iterationsprozesses:

Sei $A \subset \mathbf{R}$ eine Vereinigung von Intervallen endlicher Länge. Aus jedem dieser Intervalle entfernen wir das *offene* mittlere Drittel. Der Rest besteht also aus disjunkten Dritteln des ursprünglichen Intervalles. Die Vereinigung der so entstehenden Intervalle bezeichnen wir mit $c(A)$. Sei etwa $A_0 = [0, 1]$. Dann ist $c(A_0) = [0, 1/3] \cup [2/3, 1]$. Weiters gilt offenbar

$$A_0 \supset A_1 := c(A_0) \supset A_2 := c(A_1) \supset \dots .$$

Die Cantormenge ist definiert als $\mathcal{C} = \bigcap_{i=0}^{\infty} A_i$.

Ein zweidimensionales Analogon ergibt sich folgendermaßen: Zunächst bemerken wir, daß jedes Dreieck Vereinigung von vier zu diesem Dreieck ähnlichen Dreiecken halber Seitenlänge ist. Nicht wahr? Ist eine Menge A Vereinigung von abgeschlossenen Dreiecken und entfernen wir aus jedem dieser Dreiecke das *offene* mittlere Dreieck, so entsteht wieder eine Menge, die Vereinigung abgeschlossener Dreiecke ist. Wir bezeichnen sie mit $s(A)$. Sei A_0 ein abgeschlossenes Dreieck, etwa mit den Eckpunkten $(0, 0), (1, 0), (0, 1)$. Dann gilt offenbar

$$A_0 \supset A_1 := s(A_0) \supset A_2 := s(A_1) \supset \dots .$$

Das Sierpinski-Dreieck ist definiert als $\mathcal{S} = \bigcap_{i=0}^{\infty} A_i$.

Offenbar gilt $0 \in \mathcal{C}$ und $(0, 0) \in \mathcal{S}$ und beide Mengen sind als Durchschnitt einer abnehmenden Folge abgeschlossener, beschränkter Mengen kompakt, somit $\mathcal{C} \in \mathcal{H}(\mathbf{R})$ und $\mathcal{S} \in \mathcal{H}(\mathbf{R}^2)$. Beim Sierpinski-Dreieck sind diese Mengen sogar zusammenhängend, sodaß nach einem Satz der allgemeinen Topologie auch $\mathcal{S}$ zusammenhängend ist. $\mathcal{C}$ hingegen ist total unzusammenhängend, d.h. die Zusammenhangskomponente eines jeden Punktes ist einpunktig. Für die Punkte von $\mathcal{C}$ gibt es eine arithmetische Charakterisierung.

SATZ 1.1. *Die Cantor–Menge $\mathcal{C}$ besteht aus allen Punkten $x \in [0,1]$, deren triadische Entwicklung ohne Verwendung von 1 möglich ist.*

$$\mathcal{C} = \left\{ x \in [0,1] : x = \sum_{i}^{\infty} \frac{x_i}{3^i},\ x_i \in \{0,2\} \right\}$$

BEWEIS. Ein Querstrich über einem x_i bedeutet im Folgenden die periodische Fortsetzung, etwa $0,0\overline{2} = 0,0222\ldots$. Falls $x_o \in [0,1]$ eine triadische Darstellung $x_o = 0,x_1x_2x_3\ldots$ hat mit $x_i \in \{0,2\}$, dann gibt es 2 Möglichkeiten: ist $x_1 = 0$, dann gilt

$$0 = 0,00\cdots \leq x_o \leq 0,0\overline{2} = 1/3, \text{ also } x_o \in [0,1/3].$$

Ist $x_1 = 2$, dann gilt

$$2/3 = 0,2 \leq x_o \leq 0,\overline{2} = 1, \text{ also } x_o \in [2/3,1],$$

also $x_o \in c([0,1])$. In beiden Fällen gibt es wieder 2 Möglichkeiten. Ist $x_o \in [0,1/3]$ und $x_2 = 0$, dann gilt

$$0 = 0,\overline{0} \leq x_o \leq 0,00\overline{2} = 1/9, \text{ also } x_o \in [0,1/9].$$

Ist $x_o \in [2/3,1]$ und $x_2 = 2$, dann gilt

$$2/9 = 0,02 \leq x_o \leq 0,0\overline{2} = 1/3, \text{ also } x_o \in [2/9,1/3],$$

also $x_o \in c(c([0,1]))$. So fortfahrend erhalten wir $x_o \in \mathcal{C} = \bigcap_{i=0}^{\infty} A_i$.

Sei umgekehrt $x_o \in \mathcal{C} = \bigcap_{i=0}^{\infty} A_i$. Wegen $x_o \in A_0 = [0,1]$, erhalten wir eine triadische Darstellung $x_o = 0,x_1x_2\ldots$ mit $x_i \in \{0,1,2\}$. Wegen $x_o \in A_1 = [0,1/3] \cup [2/3,1]$ müssen, b.z.w. können wir $x_1 = 0$ oder $x_1 = 2$ wählen. Zur Beachtung: für $x_o = 2/3$ gibt es 2 triadische Darstellungen, nämlich $x_o = 0,2 = 0,1\overline{2}$. In analoger Weise müssen, b.z.w. können wir wegen

$$x_o \in A_2 = [0,1/9] \cup [2/9,1/3] \cup [6/9,7/9] \cup [8/9,1]$$

$x_2 = 0$ oder $x_2 = 2$ wählen. Durch Induktion ergibt sich $x_o \in \{x \in [0,1] : x = \sum_i^{\infty} \frac{x_i}{3^i},\ x_i \in \{0,2\}\}$. □

Im Folgenden sei (X,d) ein **vollständiger** metrischer Raum. Wir werden nun durch Einführung einer geeigneten Metrik auch $\mathcal{H}(X)$ zu einem vollständigen metrischen Raum machen.

Zu $x \in X$ und $B \in \mathcal{H}(X)$ definieren den Abstand des Punktes x zur Menge B durch

$$d(x,B) := Min\{d(x,b) :\ b \in B\}.$$

Wir bemerken, daß dieses Minimum existiert, da bei festem x die Abbildung $d_x : X \mapsto [0,\infty)$, definiert durch $d_x(b) := d(x,b)$, stetig und B kompakt ist.

Zu $A, B \in \mathcal{H}(X)$ definieren wir

$$d(A,B) := Max\{d(a,B) :\ a \in A\} = Max_{a\in A} Min_{b\in B} d(a,b).$$

Wir bemerken wieder, daß wegen der Stetigkeit von d in beiden Argumenten und wegen der Kompaktheit von A, B $a_0 \in A$ und $b_0 \in B$ existieren, für die $d(A,B) = d(a_0,b_0)$ gilt. d ist keine Metrik, da i.a. $d(A,B) \neq d(B,A)$.

Für $\alpha, \beta \in \mathbf{R}$ sei $\alpha \vee \beta := Max\{\alpha,\beta\}$ und $\alpha \wedge \beta := Min\{\alpha,\beta\}$.

SATZ 1.2. *Zu $A, B \in \mathcal{H}(X)$ ist durch*

$$h(A,B) := d(A,B) \vee d(B,A)$$

eine Metrik auf $\mathcal{H}(X)$ definiert, der sogenannte Hausdoff-Abstand zwischen A und B. Dadurch wird $(\mathcal{H}(X), h)$ zu einem metrischen Raum.

BEWEIS. Seien $A, B, C \in \mathcal{H}(X)$. Offenbar gilt $h(A,A) = d(A,A) \vee d(A,A) = Max\{d(a,A) : a \in A\} = 0$. Weiters ist $h(A,B) = d(a_0, b_0) < \infty$, somit $0 \leq h(A,B) < \infty$. Ist $A \neq B$, dann existiert $a \in A$, $a \notin B$. Somit gilt $h(A,B) \geq d(A,B) \geq d(a,B) > 0$. Zum Beweis der Dreiecksungleichung zeigen wir zunächst $d(A,B) \leq d(A,C) + d(C,B)$: für beliebiges $a \in A$ und $c \in C$ gilt

$$\begin{aligned} d(a,B) &= Min\{d(a,b) : b \in B\} \leq Min\{d(a,c) + d(c,b) : b \in B\} \\ &= d(a,c) + Min\{d(c,b) : b \in B\}. \quad \Rightarrow \end{aligned}$$

$$\begin{aligned} d(a,B) &\leq Min\{d(a,c) : c \in C\} + Max_{c \in C} Min_{b \in B} d(c,b) \\ &= d(a,C) + d(C,B) \leq d(A,C) + d(C,B). \end{aligned}$$

Daraus folgt $d(A,B) \leq d(A,C) + d(C,B)$. Analog $d(B,A) \leq d(B,C) + d(C,A)$, somit

$$\begin{aligned} h(A,B) &= d(A,B) \vee d(B,A) \leq (d(A,C) \vee d(C,A)) + (d(B,C) \vee d(C,B)) \\ &= h(A,C) + h(C,B). \end{aligned}$$

□

Eine geometrische Interpretation des Hausdorff–Abstandes ergibt sich aus folgendem

LEMMA 1.3. *Für $A, B \in \mathcal{H}(X)$ und $\epsilon > 0$ gilt*

$$(1) \quad h(A,B) \leq \epsilon \quad \Longleftrightarrow \quad \begin{cases} A \subset B_\epsilon := \bigcup_{b \in B} B(b,\epsilon) \\ B \subset A_\epsilon := \bigcup_{a \in A} B(a,\epsilon). \end{cases}$$

$$(2) \quad h(A,B) = \inf\{\epsilon : A \subset B_\epsilon, B \subset A_\epsilon\}$$

Hierbei ist $B(a,\epsilon) = \{x : d(a,x) \leq \epsilon\}$. A_ϵ heißt ϵ–Parallelmenge von A.

BEWEIS. Wir zeigen $d(A,B) \leq \epsilon \iff A \subset B_\epsilon$.

Sei dazu $d(A,B) \leq \epsilon$, also $Max\{d(a,B) : a \in A\} \leq \epsilon$. Dann gilt für alle $a \in A$ $d(a,B) \leq \epsilon$, d.h. $\bigwedge_{a \in A} \bigvee_{b \in B} d(a,b) \leq \epsilon$, also $A \subset \cup_{b \in B} B(b,\epsilon)$.
Sei umgekehrt $A \subset \cup_{b \in B} B(b,\epsilon)$. Dann gilt für alle $a \in A$ $d(a,B) \leq \epsilon$, somit $d(A,B) = Max_{a \in A}\{d(a,B)\} \leq \epsilon$.
Analog sieht man $d(B,A) \leq \epsilon \iff B \subset A_\epsilon$.

(2) folgt unmittelbar aus (1). □

2. Vollständigkeit von $(\mathcal{H}(X), h)$

LEMMA 2.1. *(Erweiterung von Cauchy–Folgen) Sei $\{A_n : n = 1, 2, \ldots\}$ eine Cauchy–Folge in $(\mathcal{H}(X), h)$. Sei weiter $\{n_j\}_1^\infty$ eine Folge in* $\mathbf{N}$ *mit $0 < n_1 < n_2 < \ldots$. Sei $\{a_{n_j}\}_{j=1}^\infty$ Cauchy–Folge von Punkten $a_{n_j} \in A_{n_j}$. Dann existiert eine Cauchy–Folge $\{\tilde{a}_n \in A_n\}$ mit $\tilde{a}_{n_j} = a_{n_j}$ $(j = 1, 2, \ldots)$.*

BEWEIS. Wir definieren die Folge $\{\tilde{a}_n\}$ induktiv. $j = 1$: für $n \in \{1, 2, \ldots, n_1\}$ wählen wir $\tilde{a}_n \in A_n$ so, daß $d(a_{n_1}, A_n) = d(a_{n_1}, \tilde{a}_n)$. Offenbar gilt dann $\tilde{a}_{n_1} = a_{n_1}$. Diese Wahl ist möglich wegen der Kompaktheit der A_n. Zum Schritt $j \to j+1$ wählen wir für $n \in \{n_j + 1, \ldots, n_{j+1}\}$ $\tilde{a}_n \in A_n$ so, daß $d(a_{n_{j+1}}, A_n) = d(a_{n_{j+1}}, \tilde{a}_n)$. Somit gilt für $n = 1, 2, \ldots : \tilde{a}_n \in A_n$ und für $j = 1, 2, \ldots : \tilde{a}_{n_j} = a_{n_j}$.

Wir zeigen nun, daß $\{\tilde{a}_n\}_1^\infty$ eine Cauchy–Folge ist. Sei dazu $\epsilon > 0$. Da sowohl $\{a_{n_j}\}$ als auch $\{A_n\}$ Cauchy–Folgen sind, gilt

$$\bigvee_{N_1} \bigwedge_{n_k, n_j \geq N_1} \quad d(a_{n_k}, a_{n_j}) \leq \epsilon/3 \quad \text{und} \quad \bigvee_{N_2} \bigwedge_{m,n \geq N_2} \quad h(A_m, A_n) \leq \epsilon/3.$$

Sei $N = N_1 \vee N_2$. Dann gilt für $m, n \geq N$

$$d(\tilde{a}_m, \tilde{a}_n) \leq d(\tilde{a}_m, a_{n_j}) + d(a_{n_j}, a_{n_k}) + d(a_{n_k}, \tilde{a}_n),$$

wobei $m \in \{n_{j-1} + 1, n_{j-1} + 2, \ldots, n_j\}$, $n \in \{n_{k-1} + 1, n_{k-1} + 2, \ldots, n_k\}$. m, n liegen dabei in bestimmten, durch $n_1 < n_2 < \ldots$ definierten Abschnitten von $\mathbf{N}$. Nun ist $h(A_m, A_{n_j}) \leq \epsilon/3$ und $h(A_{n_k}, A_n) \leq \epsilon/3$ und somit $d(\tilde{a}_m, a_{n_j}) \leq \epsilon/3$ und $d(a_{n_k}, \tilde{a}_n) \leq \epsilon/3$, also $d(\tilde{a}_m, \tilde{a}_n) \leq \epsilon$ für alle $m, n > N$. □

SATZ 2.2. *Der Raum $(\mathcal{H}(X), h)$ ist vollständig. Falls $\{A_n \in \mathcal{H}(X)\}_1^\infty$ Cauchy–Folge ist, dann existiert $\lim_{n\to\infty} A_n$, gehört zu $\mathcal{H}(X)$ und*

$$A := \left\{a \in X : a = \lim_{n\to\infty} a_n,\ a_n \in A_n\right\} = \lim_{n\to\infty} A_n.$$

BEWEIS. Sei $\{A_n\}$ Cauchy–Folge in $\mathcal{H}(X)$ und A wie oben definiert. Wir zeigen

1. $A \neq \emptyset$
2. A ist abgeschlossen und daher vollständig (da X vollständig ist).
3. Zu $\epsilon > 0$ existiert ein N, sodaß für $n \geq N$ $A \subset (A_n)_\epsilon$.
4. A ist totalbeschränkt, somit (wegen (2) und 1 1.3) kompakt.
5. $\lim_{n\to\infty} A_n = A$.

zu (1):
Wir zeigen die Existenz einer Cauchy–Folge $\{a_i \in A_i\}$ in X. Nach Voraussetzung existiert eine Folge positiver Zahlen $N_1 < N_2 < \ldots$ mit

$$\bigwedge_{m,n \geq N_i} \quad h(A_m, A_n) < \frac{1}{2^i}.$$

Wir wählen $x_{N_1} \in A_{N_1}$. Da $h(A_{N_1}, A_{N_2}) < 1/2$, gibt es ein $x_{N_2} \in A_{N_2}$ mit $d(x_{N_1}, x_{N_2}) < 1/2$. Die induktive Konstruktion sei durchgeführt bis k. Es gibt also $x_{N_i} \in A_{N_i}$ für $i = 1, 2, \ldots, k$ mit $d(x_{N_{i-1}}, x_{N_i}) < 1/2^{i-1}$. Nun ist wegen der Wahl der N_k $h(A_{N_k}, A_{N_{k+1}}) < 1/2^k$ und $x_{N_k} \in A_{N_k}$. Es existiert also $x_{N_{k+1}} \in A_{k+1}$ mit $d(x_{N_k}, x_{N_{k+1}}) < 1/2^k$. Wir erhalten daher eine unendliche Folge $\{x_{N_i} \in A_{N_i}\}$ mit

$d(x_{N_i}, x_{N_{i+1}}) < 1/2^i$. Dies ist eine Cauchy–Folge in X: zu $\epsilon > 0$ wählen wir N_ϵ so groß, daß $\sum_{i=N_\epsilon}^\infty 1/2^i < \epsilon$. Dann gilt für $n > m \geq N_\epsilon$

$$\begin{aligned} d(x_{N_m}, x_{N_n}) &\leq d(x_{N_m}, x_{N_{m+1}}) + d(x_{N_{m+1}}, x_{N_{m+2}}) + \cdots + d(x_{N_{n-1}}, x_{N_n}) \\ &< \sum_{i=N_\epsilon}^\infty 1/2^i < \epsilon. \end{aligned}$$

Nach dem Erweiterungslemma 2.1 existiert eine Cauchy–Folge $\{a_i \in A_i\}$ mit $a_{N_i} = x_{N_i}$. Diese Folge hat einen Limes, der definitionsgemäß in A liegt, also gilt $A \neq \emptyset$.

zu (2):
Sei $a = \lim_{i\to\infty} a_i$ mit $a_i \in A$. Wir zeigen $a \in A$. Zu jedem a_i existiert nach Definition von A eine Folge $\{x_{in} \in A_n\}_{n=1}^\infty$ mit $a_i = \lim_{n\to\infty} x_{in}$. Weiters finden wir eine *wachsende* Folge natürlicher Zahlen $\{N_i\}_1^\infty$ mit $d(a_{N_i}, a) < 1/i$ und zu jedem N_i ein m_i mit $d(x_{N_i m_i}, a_{N_i}) \leq 1/i$. Setzen wir $y_{m_i} := x_{N_i m_i}$, dann ist $y_{m_i} \in A_{m_i}$ und $d(y_{m_i}, a) \leq 2/i$. Somit gilt $a = \lim_{i\to\infty} y_{n_i}$. Nach Lemma 2.1 kann $\{y_{m_i}\}$ erweitert werden zu einer Cauchy–Folge $\{z_i \in A_i\}$ mit $\lim z_i = a$, somit gilt $a \in A$.

zu (3):
Sei $\epsilon > 0$. Dann existiert N, sodaß für $m, n \geq N$ $h(A_m, A_n) \leq \epsilon$. Wählen wir nun $n \geq N$. Dann gilt für alle $m \geq n$ $A_m \subset (A_n)_\epsilon$. Wir zeigen nun $A \subset (A_n)_\epsilon$. Sei dazu $a \in A$. Dann existiert eine Folge $\{a_m \in A_m\}$ mit $\lim_{m\to\infty} a_m = a$. Nun ist $a_m \in A_m \subset (A_n)_\epsilon$. $(A_n)_\epsilon$ ist abgeschlossen (Beweis?) und daher gilt auch $\lim a_m = a \in (A_n)_\epsilon$.

zu (4):
Nehmen wir im Widerspruch zur Behauptung an, A sei nicht totalbeschränkt. Dann gibt es zu einem gewissen $\epsilon > 0$ kein endliches ϵ–Netz. Daher existiert eine Folge $\{x_i\}_1^\infty$, sodaß für alle $i \neq j$ $d(x_i, x_j) \geq \epsilon$ $(*)$. Dies führt zu folgendem Widerspruch: nach (3) gilt für hinreichend große n $A \subset (A_n)_{\epsilon/3}$. Zu jedem x_i existiert also ein $y_i \in A_n$ mit $d(x_i, y_i) \leq \epsilon/3$. Da A_n kompakt ist, gibt es eine konvergente Teilfolge $\{y_{n_i}\} \subset \{y_i\}$. Die Punkte dieser Teilfolge liegen letzten Endes beliebig nahe beisammen. Es gilt also schließlich $d(y_{n_j}, y_{n_i}) < \epsilon/3$. Dann aber folgt

$$d(x_{n_i}, x_{n_j}) \leq d(x_{n_i}, y_{n_i}) + d(y_{n_i}, y_{n_j}) + d(y_{n_j}, x_{n_j}) < \epsilon/3 + \epsilon/3 + \epsilon/3$$

im Widerspruch zu $(*)$. A ist somit totalbeschränkt, nach (2) abgeschlossen und daher, nach 1 Satz 1.3, kompakt.

zu (5):
Nach (1) und (4) gilt jedenfalls $A \in \mathcal{H}(X)$. Wegen (3) und Lemma 1.3 genügt es zu zeigen

$$\bigwedge_{\epsilon>0} \bigvee_{N} \bigwedge_{n\geq N} \quad A_n \subset A_\epsilon.$$

Jedenfalls folgt aus der Tatsache, daß $\{A_n\}$ Cauchy–Folge ist

$$\bigvee_{N_0} \bigwedge_{r,s\geq N_0} \quad A_r \subset (A_s)_{\epsilon/2}, \quad \text{insbes.} \quad A_n \subset (A_{N_1})_{\epsilon/2} \quad \text{für} \quad n, N_1 > N_0.$$

$$\bigvee_{N_1>N_0} \bigwedge_{r,s\geq N_1} \quad A_r \subset (A_s)_{\epsilon/4} \quad \text{insbes.} \quad A_{N_1} \subset (A_{N_2})_{\epsilon/4} \quad \text{für} \quad N_2 > N_1.$$

$$\bigvee_{N_2>N_1} \bigwedge_{r,s\geq N_2} \quad A_r \subset (A_s)_{\epsilon/8} \quad \text{insbes.} \quad A_{N_2} \subset (A_{N_3})_{\epsilon/8} \quad \text{für} \quad N_3 > N_2.$$

... u.s.w.

Wir zeigen nun, daß für $n \geq N_0$ $A_n \subset A_\epsilon$. Jedenfalls existiert zu $y \in A_n$ ein $x_{N_1} \in A_{N_1}$ mit $d(y, x_{N_1}) \leq \epsilon/2$. Weiter existiert zu x_{N_1} ein $x_{N_2} \in A_{N_2}$ mit $d(x_{N_1}, x_{N_2}) \leq \epsilon/4$. So fortfahrend erhalten wir zu $x_{N_{j-1}}$ ein $x_{N_j} \in A_{N_j}$ mit $d(x_{N_{j-1}}, x_{N_j}) \leq \frac{\epsilon}{2^j}$. Daraus folgt

$$\begin{aligned} d(y, x_{N_j}) &\leq d(y, x_{N_1}) + d(x_{N_1}, x_{N_2}) \cdots + d(x_{N_{j-1}}, x_{N_j}) \\ &\leq \epsilon \left(\frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^j} \right) < \epsilon. \end{aligned}$$

$\{x_{N_j}\}$ ist offenbar Cauchy–Folge, hat somit nach dem Erweiterungslemma 2.1 einen Grenzwert $x \in A$. Somit gilt auch $d(y, x) \leq \epsilon$, d.h. $y \in A_\epsilon$. □

Wir werden später nicht nur Vollständigkeit von X verlangen, sondern auch die Existenz einer in X dicht liegenden abzählbaren Teilmenge. X heißt dann ein *polnischer Raum*. Insbesondere ist jeder metrische Raum, dessen beschränkte Teilmengen relativ kompakt sind, ein polnischer Raum (Wieso?). Jeder $\sigma-$kompakte metrische Raum ist z.B. polnisch.

SATZ 2.3 (Auswahlsatz, W. Blaschke). *Sei X ein metrischer Raum, dessen beschränkte Teilmengen relativ kompakt sind. Sei $\mathcal{F}$ eine unendliche Familie von nichtleeren kompakten Teilmengen von X, die alle in einem beschränkten Teil von X liegen. Dann enthält $\mathcal{F}$ eine aus verschiedenen Mengen bestehende Folge, die gegen eine nichtleere kompakte Teilmenge von X konvergiert.*

BEWEIS. Wegen Satz 2.2 genügt es zu zeigen, daß $\mathcal{F}$ eine aus verschiedenen Mengen bestehende Cauchy–Folge enthält.

Sei dazu $\mathcal{E}_1 = \{E_{1,i} : i \in \mathbf{N}\}$ eine aus verschiedenen Mengen bestehende Folge in $\mathcal{F}$. Nach der Voraussetzung über X können wir o.B.d.A. annehmen, daß die Elemente von $\mathcal{E}_1$ enthalten sind in einer kompakten Menge $B \subset X$. Wir wählen eine endliche Überlagerung von B mit Kugeln vom Durchmesser höchstens $1/2$. Dann trifft jede Menge $E_{1,i}$ eine Teilmenge dieser Kugeln. Es gibt nur endlich viele solche Teimengen. Nach dem Schubfachprinzip existiert eine gewisse Teilmenge $\mathcal{B}_2$ dieser Kugeln, die von unendlich vielen $E_{1,i}$ getroffen wird. Sei $\mathcal{E}_2 = \{E_{2,i} : i \in \mathbf{N}\}$ diese Teilfolge von $\mathcal{E}_1$ und $F_2 = \cup\mathcal{B}_2$. Die Konstruktion sei durchgeführt bis $k-1$. Dann finden wir analog zu einer endlichen Überlagerung von B mit Kugeln vom Durchmesser höchstens $1/k$ eine Teilfolge $\mathcal{E}_k \subset \mathcal{E}_{k-1}$ und eine gewisse Teilmenge $\mathcal{B}_k$ dieser Kugeln, die von allen $E_{k,i} \in \mathcal{E}_k$ getroffen wird. Sei $F_k = \cup\mathcal{B}_k$.

Wir erhalten also eine ineinander geschachtelte Folge von Teilfolgen

$$\mathcal{E}_1 \supset \mathcal{E}_2 \supset \cdots \supset \mathcal{E}_k = \{E_{k,i}\} \supset \ldots$$

sodaß für alle i bei beliebigem k gilt (wieso?):

$$E_{k,i} \subset F_k \subset (E_{k,i})_{1/k}.$$

Daraus folgt $h(E_{k,i}, F_k) \leq 1/k$ und $h(E_{k,i}, E_{k,j}) \leq 2/k$. Setzen wir $E_k := E_{k,k}$, dann folgt

$$h(E_k, E_l) \leq \frac{2}{\min\{k, l\}}.$$

$\{E_k\}$ ist also eine Cauchy–Folge in B.

□

Anmerkung: W. Blaschke hat seinen Auswahlsatz ursprünglich für konvexe Körper bewiesen (vgl. [**Bl**]). Die Sätze 2.2 und 2.3 sind Verallgemeinerungen auf $\mathcal{H}(X)$.

Wir nennen eine Mengenfolge $\{A_n\}_0^\infty$ *aufsteigend* b.z.w. *abnehmend*, falls gilt

$$\begin{aligned} &A_0 \subset A_1 \subset A_2 \subset \ldots \text{ b.z.w.} \\ &A_0 \supset A_1 \supset A_2 \supset \ldots . \end{aligned}$$

Man überlege sich folgendes

Lemma 2.4. *Jede abnehmende Folge $\{A_n \in \mathcal{H}(X)\}_0^\infty$ ist Cauchyfolge. Falls X die Voraussetzungen des Auswahlsatzes von Blaschke erfüllt, dann ist auch jede gleichmäßig beschränkte aufsteigende Folge $\{A_n \in \mathcal{H}(X)\}_0^\infty$ Cauchyfolge.*

Beweis. Übung. □

3. Iterierte Funktionensysteme in $\mathcal{H}(X)$

Von jetzt an nehmen wir den metrischen Raum (X, d) als vollständig an.

LEMMA 3.1. *Sei $w : X \to X$ eine kontrahierende Abbildung mit dem Kontraktionsfaktor (KF) s. Die durch $w(B) = \{w(x) : x \in B\}$ für $B \in \mathcal{H}(X)$ definierte Abbildung $w : \mathcal{H}(X) \to \mathcal{H}(X)$ ist eine Kontraktion auf $(\mathcal{H}(X), h)$ mit KF s.*

BEWEIS. $w : X \to X$ ist als kontrahierende Abbildung stetig. Das stetige Bild einer kompakten Menge ist kompakt, somit $w(B) \in \mathcal{H}(X)$. Seien nun $B, C \in \mathcal{H}(X)$. Dann gilt

$$\begin{aligned} d(w(B), w(C)) &\leq Max\{Min\{d(w(b), w(c)) : c \in C\} : b \in B\} \\ &\leq Max\{Min\{s\, d(b, c) : c \in C\} : b \in B\} = s\, d(B, C). \end{aligned}$$

In analoger Weise sieht man $d(w(C), w(B)) \leq s\, d(C, B)$. Daraus folgt

$$\begin{aligned} h(w(B), w(C)) &= d(w(B), w(C)) \vee d(w(C), w(B)) \\ &\leq s\,(d(B, C) \vee d(C, B) = s\, h(B, C). \end{aligned}$$

□

LEMMA 3.2. *Für beliebige $A, B, C, D \in \mathcal{H}(X)$ gilt*

$$h(A \cup B, C \cup D) \leq h(A, C) \vee h(B, D)$$

BEWEIS. Übung. □

LEMMA 3.3. *Sei $\{w_n : n = 1, 2, \ldots, N\}$ eine endliche Menge kontrahierender Abbildungen auf $(\mathcal{H}(X), h)$ mit Kontraktionsfaktoren $0 \leq s_n < 1$. Zu $B \in \mathcal{H}(X)$ definieren wir*

$$W(B) := w_1(B) \cup w_2(B) \cup \cdots \cup w_N(B) = \bigcup_{n=1}^{N} w_n(B).$$

Dann ist $W : \mathcal{H}(X) \to \mathcal{H}(X)$ eine Kontraktionsabbildung auf $(\mathcal{H}(X), h)$ mit Kontraktionsfaktor $0 \leq \; Lip\, W \leq \; s = Max\{s_n : n = 1, 2, \ldots, N\}$.

BEWEIS. Wir zeigen die Gültigkeit der Behauptung für $N = 2$. Die Weiterführung mit Induktion liegt auf der Hand. Seien also $B, C \in \mathcal{H}(X)$. Dann gilt nach dem vorigen Lemma

$$\begin{aligned} h(W(B), W(C)) &= h(w_1(B) \cup w_2(B), w_1(C) \cup w_2(C)) \\ &\leq h(w_1(B), w_1(C)) \vee h(w_2(B), w_2(C)) \\ &\leq s_1\, h(B, C) \vee s_2\, h(B, C) = s\, h(B, C). \end{aligned}$$

□

$W(B) = \bigcup w_i(B)$ ist eine "Collage" der $w_i(B)$, wir sprechen daher von einer *Collage*-Abbildung.

Definition 3.1. Ein **i**teriertes **F**unktionen**s**ystem (IFS) besteht aus einem vollständigen metrischen Raum (X, d) zusammen mit einer endlichen Menge von Abbildungen $w_n : X \to X$. Sind die $w_n : X \to X$ Kontraktionen (mit Kontraktionsfaktoren (KF) s_n), dann spricht man von einem hyperbolischen IFS. Sein KF ist definiert durch $s = Max\{s_n : n = 1, \ldots, N\}$. Wir schreiben für das IFS $\{X; w_n, n = 1, \ldots, N\}$ oder kurz $\{X; w_n\}$.

Wir fassen das bisherige zusammen und erhalten wegen der Vollständigkeit von $\mathcal{H}(X)$ Satz 2.2 und aus dem Fixpunktsatz von Banach (1 Satz 1.4) den folgenden

Satz 3.4. *Sei $\{X; w_n, n = 1, \ldots, N\}$ ein IFS mit KF s. Dann gilt: die Abbildung $W : \mathcal{H}(X) \to \mathcal{H}(X)$, definiert durch $W(B) = \bigcup_1^N w_n(B)$ ist kontrahierend auf dem vollständigen metrischen Raum $(\mathcal{H}(X), h)$ mit KF s, d.h.*

$$\bigwedge_{B,C \in \mathcal{H}(X)} h(W(B), W(C)) \le s\, h(B, C).$$

Für den eindeutig bestimmten Fixpunkt $A \in \mathcal{H}(X)$ gilt bei beliebigem $B \in \mathcal{H}(X)$

$$A = W(A) = \bigcup_{n=1}^{N} w_n(A) = \lim_{n\to\infty} W^n(B).$$

A heißt der Attraktor des IFS $\{X; w_n, n = 1, \ldots, N\}$. Wenn er eine "besondere" Struktur hat, werden wir ihn als *seltsam* oder als *deterministisches Fraktal* bezeichnen.

Beispiel 3.1.
Auf dem metrischen Raum $(\mathbf{R}, |.|)$ betrachten wir das IFS $\{\mathbf{R}; w_1, w_2\}$ mit den kontrahierenden Abbildungen $w_1(x) = x/2$ und $w_2(x) = x/2 + 1/2$. Offenbar ist $s_1 = s_2 = s = 1/2$. Wir bemerken daß für $A = [0, 1]$ gilt: $A \in \mathcal{H}(\mathbf{R})$,

$$W(A) = w_1(A) \cup w_2(A) = [0, 1/2] \cup [1/2, 1] = A.$$

A ist also Fixpunkt von W, somit der Attraktor des IFS. Wir können diesen aber auch durch Iteration von W bei beliebigem Startwert gewinnen: Sei etwa $B_0 = \{0\}$. Dann ist

$$\begin{aligned}
B_1 &:= W(B_0) = \{0, 1/2\} \subset \\
B_2 &:= W(B_1) = \{0, 1/4, 1/2, 3/4\} \subset \\
B_3 &:= W(B_2) = \{0, 1/8, 1/4, 3/8, 1/2, 5/8, 3/4, 7/8\} \subset \cdots \subset \cup B_i.
\end{aligned}$$

$\bigcup B_i$ ist die Menge der dyadisch rationalen Zahlen. A ist der Limes der B_i im Sinne der Hausdorff-Metrik. Zu $\bigcup B_i$ kommen also noch sämtliche Häufungspunkte hinzu. Somit gilt $A = [0, 1]$.

Beispiel 3.2.
Auf dem metrischen Raum $(\mathbf{R}, |.|)$ betrachten wir das IFS $\{\mathbf{R}; w_1, w_2\}$ mit den kontrahierenden Abbildungen $w_1(x) = x/3$ und $w_2(x) = x/3 + 2/3$. Offenbar ist

$s_1 = s_2 = s = 1/3$. Wir starten mit $B_0 = [0,1]$ und erhalten

$$\begin{aligned}
B_1 :=& W(B_0) = w_1(B_0) \cup w_2(B_0) = [0,1/3] \cup [2/3,3/3] \\
B_2 :=& W(B_1) = w_1(B_1) \cup w_2(B_1) \\
=& [0,1/9] \cup [2/9,3/9] \cup [6/9,7/9] \cup [8/9,9/9] \\
& \dots \\
B_n =& W(B_{n-1}).
\end{aligned}$$

B_n ensteht dabei aus B_{n-1} durch Weglassen des mittleren offenen Drittels eines jeden konstituierenden Intervalles von B_{n-1}. Nach 1 Satz 1.1 ist der Grenzwert die Cantor-Menge $\mathcal{C}$. Dabei verwenden wir folgende Tatsache (Beweis?): Eine abnehmende Folge $\{B_n \in \ \mathcal{H}(X)\}$

$$B_0 \supset B_1 \supset B_2 \supset \dots \supset B_n \supset \dots$$

ist Cauchy-Folge in $(\mathcal{H}(X), h)$ und konvergiert gegen $\bigcap B_n$. Somit ist $\mathcal{C}$ der "seltsame" Attraktor des obigen IFS.

Mit Hilfe von 1 Satz 1.1 können wir dies auch direkt einsehen: Setzen wir nämlich unter Zuhilfenahme der triadischen Entwicklung

$$\begin{aligned}
A_1 &= \{x \in \mathcal{C} : x = 0,0x_2x_3\dots,\ x_i \in \{0,2\}\} \\
A_2 &= \{x \in \mathcal{C} : x = 0,2x_2x_3\dots,\ x_i \in \{0,2\}\},
\end{aligned}$$

dann gilt offenbar $\mathcal{C} = A_1 \cup A_2$ und $A_1 \cap A_2 = \emptyset$. Weiters ist $w_1(\mathcal{C}) = A_1$ und $w_2(\mathcal{C}) = A_2$ und somit

$$\mathcal{C} = w_1(\mathcal{C}) \cup w_2(\mathcal{C}) = W(\mathcal{C}).$$

$\mathcal{C}$ ist also Fixpunkt von W und außerdem kompakt, somit Attraktor des IFS.

Es sei in diesem Zusammenhang angemerkt, daß nicht *jede*, nichtleere und unter W invariante Teilmege Y (d.h. $W(Y) = Y$) Attraktor eines IFS sein kann. Oft ist X invariant, aber sicher nicht immer Attraktor, wie im vorigen Beispiel $X = \mathbf{R}$ zeigt.

Beispiel 3.3.
Betrachten wir das IFS $\{\mathbf{R}; w_1(x) \equiv 0, w_2(x) \equiv \frac{2}{3}x + \frac{1}{3}\}$ mit Startmenge $B_0 = \{0\}$. Dann ist

$$\begin{aligned}
B_1 &= \{0, 1/3\} = \{0, x_1\} \\
B_2 &= \{0, x_1, 1/3 + 2/9\} = \{0, x_1, x_2\} \\
B_3 &= \{0, x_1, x_2, 1/3 + 2/9 + 4/27\} = \{0, x_1, x_2, x_3\} \dots .
\end{aligned}$$

Daraus sieht man $B_n = \{0, x_1, \dots, x_n\}$, wobei

$$\begin{aligned}
x_n &= \frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \dots + \frac{2^{n-1}}{3^n} \\
&= \frac{1}{3}\left(1 + \frac{2}{3} + \left(\frac{2}{3}\right)^2 + \dots + \left(\frac{2}{3}\right)^{n-1}\right) \to \frac{1}{3}\,\frac{1}{1-2/3} = 1.
\end{aligned}$$

Somit ist $A = \{0\} \cup \{x_n\}_1^\infty \cup \{1\}$.

ABBILDUNG 1. Aus [Ba], p.85

BEISPIEL 3.4.
Betrachten wir das IFS $\{\mathbf{R}^2; w_1, w_2, w_3\}$. Fassen wir $\mathbf{R}^2$ als komplexe Ebene auf, dann seien die w_i definiert durch

$$w_1(z) = \frac{1}{2}z,\ \ w_2(z) = \frac{1}{2}z + \frac{1}{2},\ \ w_3(z) = \frac{1}{2}z + \frac{i}{2}.$$

Als Startmenge A_0 der Iteration unter W wählen wir das volle Dreieck mit den Eckpunkten $(0,0), (1,0), (0,1)$. Mit anderen Worten: A_0 ist die konvexe Hülle dieser drei Punkte. Die folgende Skizze zeigt die ersten zwei Iterationsschritte. $A_0 \to A_1 \to A_2$.

Der Grenzwert $\lim_{n\to\infty} A_n = \mathcal{C}$, ist also nach den Ausführungen zu Beginn von Kapitel 2, das Siepinski-Dreieck. Wir hätten nach Satz 3.4 genausogut mit dem Einheitsquadrat, oder irgend einer anderen nichtleeren, kompakten Menge als Startmenge beginnen können; das Endergebnis bliebe gleich.

Der bisher angewandte Algorithmus zur Bestimmung des Attraktors eines IFS ergab sich aus dem Banach'schen Fixpunktsatz durch Iteration der Collage-Abbildung W. Ein ganz anderer Algorithmus zur bildmäßigen Erzeugung von Schwarz-Weiß-Bildern mit Grautönung ist in der folgenden Definition enthalten.

DEFINITION 3.2. (Zufalls–Iterations–Algorithmus) Sei

$$\{X; w_1, \ldots, w_N, p_1, \ldots, p_N\} \qquad p_i > 0.$$

ein IFS mit Wahrscheinlichkeiten p_i, d.h. für die p_i soll gelten $p_1 + \cdots + p_N = 1$. Zu einem vorgegebenen $x_0 \in X$ wählen wir für $n = 1, 2, \ldots$ die weiteren Punkte $x_n \in \{w_1(x_{n-1}), \ldots, w_N(x_{n-1})\}$ so, daß die Wahrscheinlichkeit für $x_n = w_i(x_{n-1})$ gleich p_i ist.

Das Ziel der nächsten Abschnitte ist, zu sehen, daß und in welchem Sinne die so konstruierte Folge $\{x_n\}_0^\infty$ gegen den Attraktor des IFS $\{X; w_n, n = 1, \ldots, N\}$ konvergiert.

4. Kondensationsmengen

Manche Bilder haben ein dominantes Strukturelement, das in immer kleineren Versionen vorkommt.

Abbildung 2. Aus [Ba], p.95

Solche Mengen ("Bilder") führen zu einer weiteren Art von Kontraktionsabbildungen auf $\mathcal{H}(X)$. Sei (X, d) metrischer Raum und $C \in \mathcal{H}(X)$. Die Abbildung $w_0 : \mathcal{H}(X) \to \mathcal{H}(X)$ definiert durch

$$w_0(B) = C \quad (B \in \mathcal{H}(X))$$

heißt *Kondensationsabbildung* mit C als *Kondensationsmenge*. Als konstante Abbildung ist w_0 trivialerweise kontrahierend mit KF 0 und hat genau einen Fixpunkt, nämlich C.

Definition 4.1. Sei $\{X; w_n, n = 1, \ldots, N\}$ ein IFS mit KF $0 \le s < 1$. Dann ist auch $\{X; w_0, w_1, \ldots, w_N\}$ ein IFS, das sogenannte *IFS mit Kondensation*. Dieses hat ebenfalls KF s.

Ein einfaches Beispiel ist auch 1.3. Es folgt nun die geometrische Deutung des zugehörigen Attraktors.

Sei nun $\{X; w_n, n = 1, \ldots, N\}$ ein IFS mit Kondensationsmenge C, also $w_0(B) \equiv C$. Wir definieren

$$W_0(B) = w_0(B) \cup w_1(B) \cup \cdots \cup w_N(B) = C \cup W(B).$$

Setzen wir für $n = 0, 1, 2, \ldots$ $C_n = W_0^n(C)$, wobei $W_0^0(C) = C$, dann erhalten wir die aufsteigende Folge

$$\begin{aligned} &C_0 = C \subset \\ &C_1 = W_0(C_0) = C \cup W(C) \subset \\ &C_2 = W_0(C_1) = C \cup W(C_1) = C \cup W(C) \cup W^2(C) \subset \\ &\quad \ldots \\ &C_n = W_0(C_{n-1}) = C \cup W(C_{n-1}) = C \cup W(C) \cup \cdots \cup W^n(C). \end{aligned}$$

Daraus und aus Satz 3.4 ergibt sich unmittelbar

Satz 4.1. *Der Attraktor $\tilde{C}$ des IFS $\{X; w_n, n = 0, 1, \ldots, N\}$ mit Kondensationsmenge C ist*

$$\begin{aligned} \tilde{C} &= \lim_{n\to\infty} (C \cup W(C) \cup \cdots \cup W^n(C)) \\ &= cl\left(C \cup W(C) \cup W^2(C) \cup \ldots\right). \end{aligned}$$

Hier bedeutet cl den topologischen Abschluß. Weiters gilt für den Attraktor A des ursprünglichen IFS $\{X; w_n, n = 1, \ldots, N\}$: $A \subset \tilde{C}$.

Beweis. Nach Satz 3.4 ist $\tilde{C}$ jedenfalls der Attraktor des IFS $\{X; w_n, n = 0, 1, \ldots, N\}$ und $\tilde{C} \in \mathcal{H}(X)$. Wir können aber auch direkt einsehen, daß $\tilde{C}$ der Fixpunkt von W_0 ist: W_0 ist kontrahierend, also stetig in $\mathcal{H}(X)$ und daher

$$\begin{aligned}\tilde{C} &= \lim_{n\to\infty} C_n = \lim_{n\to\infty} C_{n+1} \\ &= \lim_{n\to\infty} W_0(C_n) = W_0 \lim_{n\to\infty} C_n = W_0(\tilde{C}).\end{aligned}$$

□

Daraus ergibt sich, daß wir die Wirkung eines IFS auf eine gegebene kompakte Teilmenge C innerhalb eines kompakten Teilraumes von X studieren können.

Folgerung 4.2. *$\{X; w_1, \ldots, w_N\}$ sei IFS und $C \in \mathcal{H}(X)$. Dann existiert $\tilde{C} \in \mathcal{H}(X)$ mit $C \subset \tilde{C}$ und $w_n : \tilde{C} \to \tilde{C}$ $(n = 1, 2, \ldots, N)$ d.h.: $\{\tilde{C}; w_1, \ldots, w_N\}$ ist IFS und der zugrunde liegende Raum $\tilde{C}$ ist kompakt.*

Beweis. Dies ergibt sich unmittelbar aus dem vorigen Satz und der Bemerkung $C \subset \tilde{C}$ und $W(\tilde{C}) \subset \tilde{C}$ □

5. Erzeugung von Schwarz-Weiß-Bildern

Der folgende Satz zeigt, wie der Attraktor eines IFS mit der durch seine Transformationen erzeugten Collage eines vorgegebenen "Bildes" zusammenhängt.

Satz 5.1 (Collage-Theorem, M. Barnsley). *Sei (X, d) ein vollständiger metrischer Raum und $L \in \mathcal{H}(X)$ gegeben. Weiters nehmen wir an, $\{X; w_n, n = 1, \ldots, N\}$ sei ein IFS (mit oder ohne Kondensation) mit KF $0 \leq s < 1$ und Attraktor A. Dann gilt*

$$h(L, A) \leq \frac{1}{1-s} h\left(L, \bigcup_1^N w_n(L)\right).$$

Der Beweis ergibt sich unmittelbar aus folgendem

Lemma 5.2. *Sei (X, d) ein vollständiger metrischer Raum, $f : X \to X$ kontrahierend mit KF $0 \leq s < 1$ und Fixpunkt $x_f \in X$. Dann gilt*

$$d(x, x_f) \leq \frac{1}{1-s} d(x, f(x)) \qquad (x \in X).$$

BEWEIS. Die Distanzfunktion $d(a,b)$ ist bei festem a stetig in b. Daher gilt

$$\begin{aligned} d(x,x_f) &= d\left(x, \lim_{n\to\infty} f^n(x)\right) = \lim_{n\to\infty} d\left(x, f^n(x)\right) \\ &\le \lim_{n\to\infty} \left(d(x,f(x)) + d(f(x),f^2(x)) + \cdots + d(f^{n-1}(x), f^n(x))\right) \\ &\le \lim_{n\to\infty} \left(d(x,f(x))(1+s+s^2+\cdots+s^{n-1})\right) \\ &= \frac{1}{1-s} d(x,f(x)). \end{aligned}$$

□

Die nächsten Lemmata dienen dem Nachweis, daß der Attraktor eines IFS stetig von Parametern der w_i abhängt.

LEMMA 5.3. *Seien (P, d_P), (X,d) metrische Räume, (X,d) sei vollständig. Sei $w : P \times X \to X$ eine parametrisierte Familie von Kontraktionsabbildungen auf X mit gleichmäßig durch $0 \le s < 1$ beschränkten KF–s, d.h. für beliebige $p \in P$ ist $w(p,.) : X \to X$ Kontraktionsabbildung mit KF $\le s$. Für festes $x \in X$ hänge $w(p,x)$ stetig von p ab. Dann hängt auch der Fixpunkt von $w(p,.)$ stetig von p ab.*

BEWEIS. Bei festem p sei $x_f(p)$ der Fixpunkt von $w(p,.)$. Dann gilt für alle $q \in P$

$$\begin{aligned} d(x_f(p), x_f(q)) &= d\left(w(p, x_f(p)), w(q, x_f(q))\right) \\ &\le d\left(w(p, x_f(p)), w(q, x_f(p))\right) + d\left(w(q, x_f(p)), w(q, x_f(q))\right) \\ &\le d\left(w(p, x_f(p)), w(q, x_f(p))\right) + s\, d(x_f(p), x_f(q)). \end{aligned}$$

Daher gilt

$$d(x_f(p), x_f(q)) \le \frac{1}{1-s} d\left(w(p, x_f(p)), w(q, x_f(p))\right).$$

Wegen der Stetigkeit von w in der ersten Komponente, kann für $q \to p$ die rechte Seite beliebig klein gemacht werden. □

LEMMA 5.4. *Seien (X,d) ein metrischer Raum, (P, d_P) ein kompakter metrischer Raum, $w_n : X \to X$ $(n = 1, \ldots, N)$ stetig und stetig abhängig von einem Parameter $p \in P$. D.h. $w_n(p,x)$ hängt bei festem $x \in X$ stetig von p ab. Dann ist die Abbildung $W : \mathcal{H}(X) \to \mathcal{H}(X)$, definiert durch*

$$W(p,B) = \bigcup_{n=1}^{N} w_n(p,B)$$

stetig in p, d.h. $W(p,B) \in \mathcal{H}(X)$ hängt stetig bezüglich der Hausdorff-Metrik von p ab.

BEWEIS. In Hinblick auf Lemma 3.2 können wir o.B.d.A. $N = 1$ annehmen. Für $B \in \mathcal{H}(X)$ und $p, q \in P$ gilt

$$\begin{aligned} & d\left(w_1(p,B), w_1(q,B)\right) = \\ & \quad = Max_{x\in B} Min_{y \in B} \left\{d\left(w_1(p,x), w_1(q,y)\right)\right\} \\ & \quad \le Max_{x\in B} Min_{y \in B} \left\{d\left(w_1(p,x), w_1(p,y)\right) + d(w_1(p,y), w_1(q,y))\right\}. \end{aligned}$$

$P \times B$ ist kompakt und $w_1 : P \times B \to X$ stetig, ist also gleichmäßig stetig. Zu $\epsilon > 0$ existiert daher $\delta > 0$, sodaß $d(w_1(p,y), w_1(q,y)) < \epsilon$ für alle $y \in B$ und $d_P(p,q) < \delta$. Für solche p, q gilt dann

$$\begin{aligned} d\left(w_1(p,B), w_1(q,B)\right) &= Max_{x\in B} Min_{y\in B}\left\{d\left(w_1(p,x), w_1(q,y)\right) + \epsilon\right\} \\ &\leq d(w_1(p,B), w_1(p,B)) + \epsilon = \epsilon. \end{aligned}$$

Ähnlich zeigt man $d(w_1(q,B), w_1(p,B)) < \epsilon$ für $d_P(p,q) < \delta$. Somit gilt

$$h(w_1(p,B), w_1(q,B)) < \epsilon \;\text{ für } d_P(p,q) < \delta.$$

□

SATZ 5.5. $\{X; w_n, n = 1, \ldots, N\}$ *sei IFS (mit oder ohne Kondensation) mit KF s. Für* $n = 1, \ldots, N$ *sei* w_n *stetig abhängig von* $p \in P$, P *sei kompakter metrischer Raum. Dann hängt der Attraktor* $A(p) \in \mathcal{H}(X)$ *stetig bezüglich der Hausdorff-Metrik von* p *ab.*

BEWEIS. Dies ist eine unmittelbare Folgerung aus den vorigen Lemmata. □

Dieser Satz legt die Möglichkeit von Computer-Animation mit Hilfe von IFS-Code nahe.

BEISPIEL 5.1.
Betrachten wir das IFS $\{\mathbf{R}; w_1(x) = \frac{1}{2}x, w_2(x) = \frac{1}{2}x + \frac{1}{2}\}$. Das "Bild" sei $L = [0,1]$. Hier ist $w_1(L) \cup w_2(L) = L$, somit zugleich der Attraktor.

BEISPIEL 5.2.
Betrachten wir nun ein leicht "gestörtes " IFS $\{\mathbf{R}; w_1(x) = 0.51x - 0.01, w_2(x) = 0.47x + 0.53\}$. Wie weit ist dessen Attraktor von $L = [0,1]$ entfernt? Der KF des IFS ist $s = 0.51$ und weiters gilt

$$h([0,1], w_1(L) \cup w_2(L)) = h([0,1], [-0.01, 0.5] \cup [0.53, 1]) = 0.015.$$

Daher nach Satz 5.1 $h(A, L) \leq 0.015/0.49 < 0.04$.

BEISPIEL 5.3.
Die folgende Tabelle enthält den IFS-Code eines Farnblattes und kann direkt im Zufalls-Iterations-Algorithmus verwendet werden.

w	a	b	c	d	e	f	p
1	0	0	0	0.16	0	0	0.01
2	0.85	0.04	-0.04	0.85	0	1.6	0.85
3	0.2	-0.26	0.23	0.22	0	1.6	0.07
4	-0.15	0.28	0.26	0.24	0	0.44	0.07

Bezüglich des Auffindens dieser Zahlen sei auf [**Ba**], p 87 und pp 96–110 verwiesen. Es lohnt sich darüber nachzusinnen, was es bedeutet, die Form eines Blattes "codiert" zu sehen in einem Tableau von 28 Zahlen und einem, von diesen unabängigen, universellen Algorithmus. Wieder sei auf [**Th**], Chapter IX verwiesen, wo mathematische Vision und Einsicht zugleich einen literarischen Ausdruck gefunden haben.

BEISPIEL 5.4.
Als einparametrige Familie von IFS's betrachten wir $\{\mathbf{R}^2; w_1^p, w_2^p, w_3^p\}$ mit $p \in [0,1]$, wobei

$$w_1^p \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1+p}{4} & 0 \\ 0 & \frac{1+p}{4} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$w_2^p \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1+p}{4} & 0 \\ 0 & \frac{1+p}{4} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \frac{3+p}{8} \\ \frac{p}{2} \end{bmatrix}$$

$$w_3^p \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1+p}{4} & 0 \\ 0 & \frac{1+p}{4} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \frac{3-p}{4} \\ 0 \end{bmatrix}$$

Man überzeuge sich, daß $A(0)$ eine Cantor-Menge in $[0,1]$ ist und $A(1)$ ein Sierpinski-Dreieck. Die folgende Skizze ist aus [**Ba**], p.150.

KAPITEL 3

Adreß-Raum und Attraktor eines IFS

1. Adressen von Punkten auf dem Attraktor

Wir betrachten ein IFS $\{X; w_1, \dots, w_N\}$ mit Attraktor A und wählen einen Punkt $a \in A$. Es gilt: $A = w_1(A) \cup w_2(A) \cup \dots \cup w_N(A)$. Daher existiert (ein nicht notwendig eindeutiges) $\sigma_1 \in \{1, 2, \dots, N\}$ mit $a \in w_{\sigma_1}(A)$. Es gilt: $w_{\sigma_1}(A) = w_{\sigma_1} \circ w_1(A) \cup w_{\sigma_2} \circ w_2(A) \cup \dots \cup w_{\sigma_1} \circ w_N(A)$. Somit existiert (ein nicht notwendig eindeutiges) $\sigma_2 \in \{1, 2, \dots, N\}$ mit $a \in w_{\sigma_1} \circ w_{\sigma_2}(A)$.

Es gilt: $w_{\sigma_1} \circ w_{\sigma_2}(A) = \bigcup_1^N w_{\sigma_1} \circ w_{\sigma_2} \circ w_i(A)$. Somit existiert $\sigma_3 \in \{1, 2, \dots, N\}$ mit $a \in w_{\sigma_1} \circ w_{\sigma_2} \circ w_{\sigma_3}(A)$ usw. Solcherart fortfahrend erhalten wir durch induktive Definition die Folge $\sigma = \{\sigma_i\}$. Jede so konstruierte Folge

$$\sigma = \sigma_1 \sigma_2 \sigma_3 \dots \in \prod_1^\infty \{1, 2, \dots, N\} = \Sigma$$

heißt "Adresse" von a.

Bemerkung 1.1. Nach der "Anzahl" der Adressen eines Attraktorpunktes gibt es verschiedene Möglichkeiten:
a) Jeder Punkt aus A hat genau *eine* Adresse, wenn $w_i(A), w_j(A)$ paarweise disjunkt sind, also $\bigwedge_{i \neq j} \quad w_i(A) \cap w_j(A) = \emptyset$. Das IFS $\{X, w_1, \dots, w_N\}$ heißt in diesem Fall *total unzusammenhängend.* Ist dies nicht der Fall, dann gibt es "Überlappungen" der $w_i(A)$ und daher Punkte mit mehreren Adressen.
b) Wenn der Anteil der Punkte mit Mehrfachadressen "klein" ist, dann heißt das IFS *gerade berührend.*
c) Ist dieser Anteil "groß", dann heißt das IFS *überlappend.*

Die Präzisierung dieser Begriffe erfolgt in Definition 2.3 unten.

Beispiel 1.1.
Zu a) Wir betrachten das IFS $\{\mathbf{R}; w_1(x) = \frac{1}{3}x, w_2(x) = \frac{1}{3}x + \frac{2}{3}\}$. Der Attraktor ist die Cantor-Menge $A = \mathcal{C}$ und es gilt

$$w_1(A) \cap w_2(A) = \emptyset, \quad w_1(A) \cup w_2(A) = \mathcal{C}$$

Das IFS ist also total unzusammenhängend.
Zu b) $\{\mathbf{R}; w_1(x) = \frac{1}{2}x, w_2(x) = \frac{1}{2}x + \frac{1}{2}\}$. Es ist

$$A = [0, 1] = w_1(A) \cup w_2(A) = [0, 1/2] \cup [1/2, 1] \quad w_1(A) \cap w_2(A) = \{1/2\}.$$

Es gibt Punkte mit Mehrfachadressen, z.B. ist die Adresse von $1/2$

$$1\,2\,2\,2 \dots =: 1\bar{2} \qquad \text{oder} \quad 2\,1\,1\,1 \dots =: 2\bar{1}$$

Die Adresse von 0 ist eindeutig, nämlich $1\,1\,1 \dots = \bar{1}$. Man überzeuge sich, daß die Gesamtheit der Punkte mit mehreren Adressen die Menge der dyadisch rationalen

Zahlen ist, also Zahlen der Form $(a_0 + a_1 2^1 + \cdots + a_k 2^k)/2^l$, wobei $k < l$ und $a_i \in \{0,1\}$, $a_k = 1$. Dies ist eine abzählbare Menge, also "klein"im Verhältnis zu $[0,1]$.
Zu c) $\left\{\mathbf{R}; w_1(x) = \frac{1}{2}x, w_2(x) = \frac{3}{4}x + \frac{1}{4}\right\}$. Hier gilt

$$A = [0,1], \quad w_1(A) \cap w_2(A) = [1/4, 1/2]\,.$$

Jeder Punkt in $[1/4, 1/2]$ hat somit wenigstens 2 Adressen. $0, 1$ haben jeweils **nur eine** Adresse. Wir bemerken, daß *total unzusammenhängend, überlappend, gerade berührend* Eigenschaften des IFS sind, nicht des Attraktors (siehe b),c)): sie sind abhängig vom Adressierungsschema und somit vom IFS.

2. Stetige Abbildungen zwischen Adreß-Raum und Attraktor

DEFINITION 2.1. Sei $\{X; w_1, \ldots, w_N\}$ ein IFS. Der zugehörige Adreß-Raum (Σ, d) ist definiert durch $\Sigma = \prod_1^\infty \{1, 2, \ldots, N\}$ (vgl. Beispiel I.1.2) mit der Metrik

$$d(\omega, \sigma) = \sum_{n=1}^{\infty} \frac{|\omega_n - \sigma_n|}{(N+1)^n} \qquad (\omega, \sigma \in \Sigma)$$

Dabei ist $\omega = \omega_1 \omega_2 \omega_3 \ldots$.

Wir benützen im Folgenden das Korollar 2 4.2, wonach wir die Wirkung eines IFS auf eine gegebene kompakte Teilmenge innerhalb eines kompakten Teilraumes von X studieren können.

Die Verwendung desselben Buchstabens d für die Metrik von X und Σ sollte im Folgenden zu keinen Verwechslungen führen.

LEMMA 2.1. *$\{X; w_1, \ldots, w_N\}$ sei IFS mit KF s und Adreß-Raum (Σ, d).
Wir definieren $\phi : \Sigma \times \mathbf{N} \times X \to X$ durch*

$$\phi(\sigma, n, x) = w_{\sigma_1} \circ w_{\sigma_2} \circ \cdots \circ w_{\sigma_n}(x).$$

Sei $K \in \mathcal{H}(X)$. Dann existiert $\infty > D \geq 0$, sodaß

$$d\big(\phi(\sigma, m, x_1), \phi(\sigma, n, x_2)\big) \leq D\, s^{m \wedge n}$$

für alle $\sigma \in \Sigma$, $m, n \in \mathbf{N}$, $x_1, x_2 \in K$.

BEWEIS. Seien σ, m, n, x_1, x_2 wie in der Voraussetzung. $\tilde{K}$ sei konstruiert nach Korollar 4.2 und o.B.d.A. sei $m < n$. Wir beachten

$$\phi(\sigma, n, x_2) = \phi(\sigma, m, \phi(\omega, n - m, x_2))$$

wobei $\omega = \sigma_{m+1}\sigma_{m+2}\ldots\sigma_n\cdots \in \Sigma$. Setzen wir $x_3 := \phi(\omega, n-m, x_2)$, dann ist $x_3 \in \tilde{K}$ und somit

$$\begin{aligned}
d\big(\phi(\sigma,m,x_1),\phi(\sigma,n,x_2)\big) &= d\big(\phi(\sigma,m,x_1),\phi(\sigma,m,x_3)\big)\\
&\le s\,d\big(w_{\sigma_2}\circ\cdots\circ w_{\sigma_m}(x_1), w_{\sigma_2}\circ\cdots\circ w_{\sigma_m}(x_3)\big)\\
&\le s^2 d\big(w_{\sigma_3}\circ\cdots\circ w_{\sigma_m}(x_1), w_{\sigma_3}\circ\cdots\circ w_{\sigma_m}(x_3)\big)\\
&\ldots\\
&\le s^m d(x_1,x_3) \le s^m D
\end{aligned}$$

wobei $D = Max\{d(x,y) : x,y \in \tilde{K}\} =: |\tilde{K}|$ der "Durchmesser"von $\tilde{K}$ ist.
D ist endlich, da $\tilde{K}$ kompakt ist. □

Satz 2.2. *Sei (X,d) vollständiger metrischer Raum, $\{X; w_1,\ldots,w_N\}$ ein IFS mit Attraktor A und Adreß-Raum (Σ,d). Für $\sigma \in \Sigma$, $n \in \mathbf{N}$, $x \in X$ sei wie oben $\phi(\sigma,n,x) = w_{\sigma_1}\circ w_{\sigma_2}\circ\cdots\circ w_{\sigma_n}(x)$. Dann existiert $\phi(\sigma) := \lim_{n\to\infty}\phi(\sigma,n,x)$, gehört zu A und ist unabhängig von x. Ist $K \in \mathcal{H}(X)$, dann ist die Konvergenz gleichmäßig in $x \in K$. Die Abbildung $\phi : \Sigma \to A$ ist stetig und surjektiv. Ist darüber hinaus ϕ bijektiv, dann ist $\phi : \Sigma \to A$ sogar Homöomorphie.*

Beweis. Sei $x \in K \in \mathcal{H}(X)$, $\tilde{K}$ und $W : \mathcal{H}(X) \to \mathcal{H}(X)$ wie oben. W ist Kontraktion auf $(\mathcal{H}(X),h)$ und $A = \lim_{n\to\infty} W^n(K)$. Insbesondere: $\{W^n(K)\}$ ist Cauchyfolge in $(\mathcal{H}(X),h)$.
Man beachte $\phi(\sigma,n,x) \in W^n(K)$. Wenn also $\lim\phi(\sigma,n,x)$ existiert, dann liegt dieser in A (Satz II.2.2). Dieser Grenzwert existiert in der Tat:
Wegen Lemma 2.1 oben ist $d(\phi(\sigma,m,x),\phi(\sigma,n,x)) \le Ds^{m\wedge n} \to 0$ für $m,n\to\infty$ und gleichmäßig in x, da D unabhängig von x ist (nur abhängig von K!).

Wir zeigen nun, daß $\phi : \Sigma \to A$ stetig ist.
Zu $\epsilon > 0$, wählen wir m so groß, daß $s^m D < \epsilon$ und $\sigma,\omega \in \Sigma$ so, daß

$$d(\sigma,\omega) = \sum_{i=1}^{\infty}\frac{|\sigma_i-\omega_i|}{(N+1)^i} < \sum_{i=m+2}^{\infty}\frac{N}{(N+1)^i} = \frac{1}{(N+1)^{m+1}} =: \delta(\epsilon).$$

Dann gilt $\sigma_1 = \omega_1,\ldots,\sigma_m = \omega_m$ (vgl. Lemma I.1.2). Somit folgt für $n \ge m$ und gewisse $x_1,x_2 \in \tilde{K}$

$$\begin{aligned}
d\big(\phi(\sigma,n,x),\phi(\omega,n,x)\big) &= d\big(\phi(\sigma,m,x_1),\phi(\sigma,m,x_2)\big)\\
&\le s^m D < \epsilon \quad \text{nach vorigem Lemma.}
\end{aligned}$$

Also für $n\to\infty$ $\quad d\big(\phi(\sigma),\phi(\omega)\big) < \epsilon$. ϕ ist somit sogar gleichmäßig stetig.

Wir zeigen nun, daß ϕ surjektiv ist.
Sei dazu $a \in A$. Nun ist $A = \lim_{n\to\infty} W^n(\{a\})$, also Abschluß der Menge

$$\left(\bigcup_{i=1}^{N}\{w_i(a)\}\right) \cup \left(\bigcup_{i,j=1}^{N}\{w_i\circ w_j(a)\}\right) \cup \left(\bigcup_{i,j,k=1}^{N}\{w_i\circ w_j\circ w_k(a)\}\right) \cup \ldots.$$

Es existiert daher eine Folge $\{\omega^{(n)} \in \Sigma : n = 1,2,3,\ldots\}$
mit $\lim_{n\to\infty}\phi(\omega^{(n)},n,a) = a$. Dazu bemerken wir, daß $\omega^{(n)}$ auf jeden Fall in $\phi(\omega^{(n)},n,a)$ nach der n-ten Stelle abgeschnitten wird. Was also ab der $n+1$-Stelle kommt, ist belanglos.

Weiters ist (Σ, d) kompakt. Die Folge $\{\omega^{(n)}\}$ besitzt daher eine konvergente Teilfolge mit Limes $\omega \in \Sigma$.
o.B.d.A. nehmen wir an $\omega^{(n)} \to \omega$. Mit wachsendem n stimmt also $\omega^{(n)}$ mit ω an den ersten $\alpha(n)$ Stellen überein, wobei gilt $\alpha(n) \to \infty$ mit $n \to \infty$. Daher

$$\begin{array}{ccc} d(\phi(\omega, n, a), & \phi(\omega^{(n)}, n, a)) & \leq s^{\alpha(n)} D \\ \downarrow & \downarrow & \downarrow \\ \phi(\omega) & a & 0 \end{array}$$

d.h. $\phi(\omega) = a$, also ist $\phi : \Sigma \to A$ surjektiv. Die letzte Behauptung ergibt sich aus einem bekannten Satz der Topologie für bijektive stetige Abbildungen zwischen einem kompakten und einem Hausdorff–Raum. □

Definition 2.2. Sei $\{X; w_n \quad n = 1, 2, \ldots, N\}$ IFS mit Adreß-Raum Σ und obiger Abbildung $\phi : \Sigma \to A$. Zu $a \in A$ heißt $\phi^{-1}(a) = \{\sigma \in \Sigma : \phi(\sigma) = a\}$ die Menge der Adressen von a.

Bemerkung 2.1. Die so definierten Adressen sind identisch mit jenen vom Beginn des Kapitels.

Beweis. Bei der dort induktiv konstruierten Adresse σ gilt

$$a = w_{\sigma_1}(a_1) = w_{\sigma_1} \circ w_{\sigma_2}(a_2) = \cdots = w_{\sigma_1} \circ w_{\sigma_2} \circ \cdots \circ w_{\sigma_n}(a_n) = \ldots$$

Somit

$$\begin{aligned} & d\big(a, \phi(\sigma, n, a)\big) \\ & \quad = d\big(w_{\sigma_1} \circ \cdots \circ w_{\sigma_n}(a_n), w_{\sigma_1} \circ \cdots \circ w_{\sigma_n}(a)\big) \\ & \quad \leq s^n\, D \to 0. \end{aligned}$$

Nun ist $\phi(\sigma, n, a) \to \phi(\sigma)$, somit $d(a, \phi(\sigma)) = 0$ und daher $a = \phi(\sigma)$.

Ist umgekehrt

$$\begin{aligned} a = \phi(\sigma) &= \lim_{n\to\infty} \phi(\sigma, n, x) = \lim_{n\to\infty} \phi(\sigma, n, a) \\ &= \lim_{n\to\infty} w_{\sigma_1} \circ w_{\sigma_2} \circ \cdots \circ w_{\sigma_n}(a) \\ &= w_{\sigma_1}\big(\lim w_{\sigma_2} \circ \cdots \circ w_{\sigma_n}(a)\big) \\ &= w_{\sigma_1}(x_1) \in w_{\sigma_1}(A) \\ &= w_{\sigma_1} \circ w_{\sigma_2}\big(\lim w_{\sigma_3} \circ \cdots \circ w_{\sigma_n}(a)\big) = \ldots \end{aligned}$$

usw. Hier haben wir die Stetigkeit der w_i benützt. Somit $a \in w_{\sigma_1}(A) \cap w_{\sigma_1} \circ w_{\sigma_2}(A) \cap \ldots$, also ist $\sigma_1\sigma_2\sigma_3\ldots$ auch Adresse im früheren Sinn. □

Definition 2.3. Ein IFS heißt

a) *total unzusammenhängend*, wenn jeder Punkt genau eine Adresse hat (ϕ ist also bijektiv).

b) *gerade berührend*, wenn es nicht total unzusammenhängend ist und A die folgende Bedingung (*open set condition*) erfüllt: es existiert eine nichtleere, offene und beschränkte Menge O, für die

(i) $w_i(O) \cap w_j(O) = \emptyset \quad (i \neq j)$
(ii) $\bigcup_1^N w_i(O) \subset O$

c) *überlappend*, falls weder a) noch b) gelten.

SATZ 2.3. *Sei $\{X; w_1, \ldots, w_N\}$ ein IFS mit Attraktor A. Das IFS ist total unzusammenhängend genau dann, wenn*

$$\bigwedge_{i \neq j} \qquad w_i(A) \cap w_j(A) = \phi.$$

BEWEIS. Dies ergibt sich unmittelbar aus der vorigen Bemerkung und jener zu Beginn des Kapitels. □

Übung:
a) Man zeige $\{\mathbf{R}; \frac{1}{2}x, \frac{1}{2}x + \frac{1}{2}\}$ ist gerade berührend.
b) Man zeige $\{\mathbf{R}; \frac{1}{2}x, \frac{3}{4}x + \frac{1}{4}\}$ ist überlappend.
b) $\{\mathbf{R}; \frac{1}{2}x, 1\}$ ist was?

BEMERKUNG 2.2. $\phi : \Sigma \to A$ ist abgeschlossen und stetig. (Da Σ kompakt ist, gilt: $S \subset \Sigma$ ist abgeschlossen genau dann, wenn S kompakt ist. $\phi(S)$ als stetiges Bild ist somit kompakt und daher wieder abgeschlossen, da A kompakt ist.) Somit ist ϕ identifizierend und A homöomorph zu $\Sigma/\sim$. Hier ist $\sigma \sim \omega$, falls $\phi(\sigma) = \phi(\omega)$. Wir zeigen nun, daß Σ metrisch äquivalent ist zu einer klassischen Cantor-Menge.

SATZ 2.4. *Seien $d_1, d_2 : \Sigma \to [0, \infty)$ definiert durch*

$$d_1(\xi, \eta) := \sum_{i=1}^{\infty} \frac{|\xi_i - \eta_i|}{(N+1)^i}, \;\; d_2(\xi, \eta) := \left| \sum_{1}^{\infty} \frac{\xi_i - \eta_i}{(N+1)^i} \right|.$$

Dann sind (Σ, d_1) (Σ, d_2) äquivalente metrische Räume.

BEWEIS. Wir führen den Beweis für $N = 10$. Sei $\xi, \eta \in \Sigma$. Jedenfalls gilt $d_2(\xi, \eta) \leq d_1(\xi, \eta)$. Wir haben zu zeigen:

$$\bigvee_{C} \bigwedge_{\xi, \eta} \quad C\, d_1(\xi, \eta) \leq d_2(\xi, \eta).$$

Wir zeigen, daß $C = 1/19$ geeignet ist. Sei k der kleinste Index, sodaß $\xi_1 = \eta_1, \xi_2 = \eta_2, \ldots, \xi_{k-1} = \eta_{k-1}, \xi_k \neq \eta_k$. Dann

$$\begin{aligned} d_2(\xi, \eta) &= \left| \sum_{1}^{\infty} \frac{\xi_i - \eta_i}{11^i} \right| \geq \frac{|\xi_k - \eta_k|}{11^k} - \sum_{i=k+1}^{\infty} \frac{|\xi_i - \eta_i|}{11^i} \\ &\geq \frac{|\xi_k - \eta_k|}{11^k} - \sum_{i=k+1}^{\infty} \frac{9}{11^i} = \left(|\xi_k - \eta_k| - \frac{9}{10} \right) \frac{1}{11^k} = (*) \end{aligned}.$$

Wie groß ist a zu wählen, damit für alle $x \geq 1$.

$$(**) \quad x - \frac{9}{10} \geq \frac{1}{a} \left(x + \frac{9}{10} \right)?$$

Für $x \geq 1$ ist $(**) \Leftrightarrow a\big(x - \frac{9}{10}\big) \geq x + \frac{9}{10}$ also

$$a \geq \frac{x + \frac{9}{10}}{x - \frac{9}{10}}.$$

die rechte Seite ist fallend in x, hat also Maximum bei $x = 1$, also

$$a \geq \frac{1 + \frac{9}{10}}{1 - \frac{9}{10}} = 19$$

somit

$$(*) \geq \frac{1}{19}\left(|\xi_k - \eta_k| + \frac{9}{10}\right)\frac{1}{11^k} = \frac{1}{19}\left(\frac{|\xi_k - \eta_k|}{11^k} + \sum_{i=k+1}^{\infty}\frac{9}{11^i}\right) \geq$$
$$\geq \frac{1}{19}\left(\frac{|\xi_k - \eta_k|}{11^k} + \sum_{i=k+1}^{\infty}\frac{|\xi_i - \eta_i|}{11^i}\right) \geq \frac{1}{19}\sum_{n=1}^{\infty}\frac{|\xi_i - \eta_i|}{11^i} = \frac{1}{19}d_1(\xi, \eta).$$

□

Satz 2.5. *Der Adreß-Raum Σ über N Symbolen $1, \ldots, N$ ist metrisch äquivalent zu einer (im topologischen Sinne) total unzusammenhängenden Cantor-Menge von $[0, 1]$. Der Attraktor eines total unzusammenhängenden IFS ist somit homöomorph zu dieser Cantor–Menge.*

Beweis. Wir betrachten das IFS

$$\left\{[0, 1]; w_i(x) = \frac{1}{N+1}x + \frac{1}{N+1}i, \quad i = 1, \ldots, N\right\}.$$

Dann ist $w_i[0, 1] = [\frac{i}{N+1}, \frac{i+1}{N+1}]$. Es folgt eine arithmetische Charakterisierung des Attraktors $A = \lim_{k\to\infty} W^k([0, 1])$. Sei $[0, 1] \ni x = 0, \sigma_1\sigma_2\sigma_3 \ldots$ die Entwicklung zur Basis $N + 1$. Dann gilt:
$W([0, 1]) \ni w_i(x) = 0, i\sigma_1\sigma_2\sigma_3 \ldots$
$W^2([0, 1]) \ni w_j \circ w_i(x) = 0, j\, i\sigma_1\sigma_2\sigma_3 \ldots$ usw. A ist die Menge der Grenzwerte von Cauchyfolgen von Elementen aus $W^k[0, 1]$, besteht somit aus allen Zahlen, deren Darstellung zur Basis $N + 1$ die 0 **nicht enthält**. Weiters ist sofort ersichtlich $w_i(A) \cap w_j(A) = \emptyset$ für $i \neq j$ (die Elemente $\in w_i(A)$ haben an 1. Stelle i, die Elemente $\in w_j(A)$ aber j).
Somit ist das IFS total unzusammenhängend und für $\phi : \Sigma \to A$ gilt

$$\phi(\sigma) = \phi(\sigma_1\sigma_2\sigma_3 \ldots) = \lim_{n\to\infty} w_{\sigma_1} \circ w_{\sigma_2} \circ \cdots \circ w_{\sigma_n}(0) =$$
$$= 0, \sigma_1\sigma_2\sigma_3 \cdots \in [0, 1]$$

ist die Darstellung zur Basis $N + 1$.

$$\big|\phi(\sigma) - \phi(\tau)\big| = \left|\sum_{i=1}^{\infty}\frac{\sigma_i - \tau_i}{(N+1)^i}\right| = d_2(\sigma, \tau).$$

Somit ist (Σ, d_2) isometrisch zu $(A, |\,.\,|)$, also auch (Σ, d_1), $(A, |\,.\,|)$ aber ist total unzusammenhängend. Die Zusammenhangskomponente eines $a \in (A, |\,.\,|)$ ist ein Intervall. Ist dieses nicht 1-punktig, so enthält es Punkte, in deren Darstellung zur Basis $N + 1$ die 0 vorkommt, was unmöglich ist. Die letzte Behauptung ergibt sich aus Satz 2.2. □

Sei $\{X; w_1, \ldots, w_N\}$ IFS mit Attraktor A und Adreß-Raum Σ. Wir bemerken:
Ist $\omega = \omega_1\omega_2\omega_3 \ldots$ Adresse von $a \in A$, also $\omega \in \phi^{-1}(a)$, dann ist $\tilde{\omega} = j\omega_1\omega_2\omega_3 \ldots$ Adresse von $w_j(a)$ $\quad j \in \{1, 2, \ldots, N\}$.

DEFINITION 2.4. $a \in A$ heißt *periodischer Punkt* des IFS, falls es eine endliche Folge $\{\sigma(n) \in \{1, 2, \ldots, N\}_{n=1}^{P}$ gibt mit

$$a = w_{\sigma(P)} \circ w_{\sigma(P-1)} \circ \cdots \circ w_{\sigma(1)}(a).$$

Das kleinste solche P heißt *Periode* von a.

BEMERKUNG 2.3. Aus $w_{\sigma(P)} \circ \cdots \circ w_{\sigma(1)}(a) = a$ folgt für

$$\sigma := \sigma(P) \ldots \sigma(1)\sigma(P) \ldots \sigma(1)\sigma(P) \cdots == \overline{\sigma(P) \ldots \sigma(1)} :$$

$$\phi(\sigma) = a.$$

Es existiert nämlich $\lim_{n\to\infty} \phi(\sigma, n, a)$ und die Folge $\phi(\sigma, n, a)$ hat die konstante Teilfolge $\phi(\sigma, k.P, a) = a$.

DEFINITION 2.5. Ein Punkt im Adreß-Raum mit periodisch wiederkehrenden Symbolen heißt *periodische Adresse*. Er heißt *schließlich periodisch*, falls er periodisch ist nach Weglassen eines Anfangsabschnittes.

$12\,12\,12 \cdots = \overline{12}$ ist eine periodische Adresse.
$112\,111\,221\,1|12\,12\,\overline{12}$ ist schließlich periodisch.

Übung:
Sei $\{X; w_1, w_2, \ldots, w_N\}$ IFS mit Attraktor A. Dann sind folgende Aussagen äquivalent:
(1) $x \in A$ ist periodischer Punkt.
(2) $x \in A$ hat periodische Adresse.
(3) $x \in A$ ist Fixpunkt eines Elementes der von $\{w_1, \ldots, w_N\}$ erzeugten Halbgruppe von Transformationen.

SATZ 2.6. *Der Attraktor eines IFS $\{X; w_n, n = 1, \ldots, N\}$ ist der Abschluß seiner periodischen Punkte. Daraus folgt: A ist der Abschluß der Menge der Fixpunkte der Abbildungen $w = w_{j_1} \circ w_{j_2} \circ \cdots \circ w_{j_k}$ für alle endlichen Folgen $\{j_1, j_2, \ldots, j_k\}$ $1 \leq j_i \leq N$.*

BEWEIS. Der Adreß-Raum ist der Abschluß der Menge S der periodischen Adressen, also $\bar{S} = \Sigma$. Nun ist $\phi : \Sigma \to A$ stetig und surjektiv, somit

$$A = \phi\,(\Sigma) = \phi(\bar{S}) \subset \overline{\phi(S)} \subset A.$$

Daher $\overline{\phi(S)} = A$. A ist also der Abschluß seiner periodischen Punkte, somit von Punkten mit periodischer Adresse, d.h. von solchen der Form $\lim_{n\to\infty} w^n(x)$. Dies ist aber Fixpunkt von w. □

3. Dynamische Systeme

DEFINITION 3.1. Ein dynamisches System (DS) $\{X; f\}$ besteht aus einer Menge X und einer Abbildung $f : X \to X$.
Als *Orbit* von $x \in X$ bezeichnet man die Folge $\{f^n(x)\}_{n=0}^{\infty}$. Dabei ist $f^0(x) = x$.

BEISPIEL 3.1.
a) $\{\Sigma, T\}$ mit $T : \Sigma \to \Sigma$ definiert durch $T(\sigma_1\sigma_2\sigma_2 \ldots) := \sigma_2\sigma_3\sigma_4 \ldots$.
b) $\{[0, 1]; f(x) = \lambda x(1 - x)\}$ $(\lambda \in [0, 4])$ ist eine 1-parametrige Familie dynamischer Systeme. Dies ist ein Wachstumsmodell für eine Population mit beschränkten

Ressourcen. Ist bei zu 1 normierter Maximalgröße x_n der Anteil der Individuen einer Population zum Zeitpunkt n, dann ergibt sich der Anteil zum Zeitpunkt $n+1$ durch $x_{n+1} = \lambda\, x_n(1-x_n)$. Dies ist das sogenannte *Verhulst*-Modell des beschränkten Wachstums.
c) Bäcker-Transformation: das Kneten von Teig. Hier folgt ein 1-dimensionales mathematisches Modell $\{[0,1]; S\}$. Je nach der Wahl des Bildes von $\frac{1}{2}$ gibt es 2 Möglichkeiten für S: $S\left(\frac{1}{2}\right) = 0, 1$ und

$$x \mapsto S(x) = \begin{cases} 2x & x \in [0, \frac{1}{2}) \\ 2x-1 & x \in (\frac{1}{2}, 1] \end{cases}$$

Definition 3.2. Sei $\{X; f\}$ dynamisches System.

x heißt *periodischer Punkt* von f, falls ein $n \in \mathbf{N}$ existiert mit $f^n(x) = x$. n heißt dann *Periode* von x. Als *minimale* Periode bezeichnet man das kleinste derartige n.

Der Orbit eines periodischen Punktes heißt *Zyklus*. Periode eines Zyklus = Periode eines darin vorkommenden Punktes.
Minimale Periode eines Zyklus = Anzahl der **verschiedenen** Punkte des Zyklus.

x_f sei FP (Fixpunkt) von f. x_f heißt *attrahierender* FP, falls für alle $\epsilon > 0$ $f|B(x_f,\epsilon)$ kontrahierend ist. Dabei ist $B(x,\epsilon) = \{y : d(x,y) \leq \epsilon\}$. Daraus folgt leicht $f\left(B(x_f,\epsilon)\right) \subset B(x_f,\epsilon)$.

x_f heißt *abstoßender* FP, falls

$$\bigvee_{C>1} \bigvee_{\epsilon>0} \bigwedge_{y\in B(x_f,\epsilon)} d\big(f(x_f), f(y)\big) \geq C d(x_f, y).$$

Ein periodischer Punkt x von f mit Periode n heißt *attrahierend*, wenn x attrahierender FP von f^n ist.

Ein Zyklus der Periode n heißt *attrahierender Zyklus* von f, wenn der Zyklus einen attrahierenden FP der Periode n enthält. Analoge Definition für *abstoßende* Zyklen usw..

$x \in X$ heißt *schließlich periodischer* Punkt von f, falls für ein natürliches m $f^m(x)$ periodisch ist.

Beispiel 3.2.
$\{\mathbf{R}; \frac{1}{2}x\}$ $x_f = 0$ ist attrahierender FP.
$\{\mathbf{R}; 2x\}$ $x_f = 0$ ist abstoßender FP.
$\{\hat{\mathbf{C}}; 0.9e^{i\alpha}z\}$ $z = 0$ ist attrahierender FP, $z = \infty$ ist abstoßender FP. $x_f = 11\cdots = \overline{1}$ ist abstoßender FP von $\{\Sigma; s\}$, wobei $s : \Sigma \to \Sigma$ definiert ist durch $s(x_1, x_2, \dots) = x_2x_3\dots$.
$x = 121\overline{212}$ ist abstoßender FP der Periode 2.
$\{\overline{12}, \overline{21}\}$ ist abstoßender Zyklus der Periode 2. Man überlege sich, daß hier alle Zyklen abstoßend sind.
$\{[0,1];\ 2x(1-x)\}$ hat attrahierenden FP $x_f = \frac{1}{2}$. Man finde einen abstoßenden FP.

Bei einem 1-dimensionalen DS $\{\mathbf{R}; f\}$ kann man unter Zuhilfenahme des Graphen von f in einem sogenannten *Netz-Diagram* den Orbit eines Punktes visualisieren:

Betrachten wir der Einfachheit halber den Fall $f : [0,1] \to [0,1]$. Wir zeichnen das Quadrat $\{(x,y) : 0 \leq x \leq 1, 0 \leq y \leq 1\}$, skizzieren den Graphen von $y = f(x)$ und die Diagonale $\{(x,x) : x \in [0,1]\}$. Wir starten beim Punkt (x_0, x_0) und verbinden ihn geradlinig mit dem Punkt $(x_0, x_1 := f(x_0))$. Nun verbinden wir diesen Punkt geradlinig mit (x_1, x_1), verbinden diesen Punkt geradlinig mit $(x_1, x_2 = f(x_1))$, u.s.w. . Der Orbit von x_0 erscheint nun auf der Diagonale als die Folge $(x_0, x_0), (x_1, x_1), (x_2, x_2), \ldots$. Die folgende Skizze ist aus [**Ba**], p.139.

KAPITEL 4

Shift-Dynamik

1. Total unzusammenhängendes IFS

LEMMA 1.1. *Sei $\{X; w_n, n = 1, \ldots, N\}$ ein IFS mit Attraktor A. Falls das IFS total unzusammenhängend ist, dann gilt für alle $n \in \{1, 2, \ldots, N\}$:*
$w_n|A : A \to A$ ist injektiv.

BEWEIS. Wir benützen ein Adreß-Raum Argument und führen den Beweis indirekt. Sei also für ein $n \in \{1, 2, \ldots, N\}$ und $a_1 \neq a_2 \in A \quad w_n(a_1) = w_n(a_2) = a \in A$. Seien $\sigma \neq \omega$ die Adressen von resp. a_1, a_2. Dann hat a zwei verschiedene Adressen $n\omega, n\sigma$, was unmöglich ist, da das IFS total unzusammenhängend ist. □

Daher ist folgende Definition sinnvoll.

DEFINITION 1.1. Sei $\{X; w_1, \ldots, w_N\}$ total unzusammenhängendes IFS mit Attraktor A. Die Verschiebungs- (Shift-)transformation $S : A \to A$ ist definiert durch

$$S(a) = w_n^{-1}(a) \quad \text{für} \quad a \in w_n(A)$$

Das dynamische System $\{A; S\}$ nennen wir Deterministisches Schift-Dynamisches System (DSDS), zugeordnet dem IFS.

BEMERKUNG 1.1. Sei $a \in w_{\sigma_1}(A)$, dann ist die Adresse von a von der Form $\sigma_1\sigma_2\sigma_3\sigma_4 \ldots$. Die Adresse von $S(a)$ ist daher $\sigma_2\sigma_3\sigma_4 \ldots$. Man sieht also eine Verschiebung nach links.

Zur Erläuterung betrachten wir das IFS

$$\left\{\mathbf{R}^2; 0.47 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, 0.47 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}, 0.47 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}.$$

Die folgende Abbildung zeigt den Attraktor und einen schließlich periodischen Orbit $\{a_n = S^{\circ n}(a_0)\}_{n=0}^{\infty}$ für das zugeordnete SDS. Dieser Orbit endet schließlich im Fixpunkt $\phi(22\overline{2})$. Der Orbit lautet $a_0 = \phi(12313222\overline{2})$, $a_1 = \phi(2313222\overline{2})$, $a_2 = \phi(313222\overline{2}), a_3 = \phi(13222\overline{2})$, $a_4 = \phi(3222\overline{2}), a_5 = \phi(222\overline{2})$, wobei $\phi : \Sigma \to A$ die zugehörige Adreß-Raum Abbildung ist. Der Punkt $a_5 \in A$ ist offenbar abstoßender FP des dynam. Systems. Man beachte, wie man den Orbit eines Punktes von seiner Adresse ablesen kann. Man sieht auch, daß die Dynamik nicht nur von A, sondern auch vom IFS selbst abhängt. Ein verschiedenes IFS mit demselben Attraktor hat i.a. eine verschiedene Shift-Dynamik.

Abbildung 1. Aus [Ba], p.143

Betrachten wir nochmals die Shift-Abbildung: für $\sigma = \sigma_1\sigma_2\sigma_3\cdots \in \phi^{-1}(a_0)$ gilt (σ ist eindeutig bei total unzusammenhängenden IFS!)

$$\begin{array}{cccl} \sigma_1\sigma_2\sigma_3\sigma_4\ldots & \stackrel{\phi}{\longrightarrow} & \phi(\sigma) & \quad = a_0 \\ \downarrow T & & \downarrow S & \\ \sigma_2\sigma_3\sigma_4\ldots & \longrightarrow & \phi(T\sigma) & = a_1 = S(a_0) \\ \downarrow T & & \downarrow S & \\ \sigma_3\sigma_4\sigma_5\ldots & \longrightarrow & \phi(T^2\sigma) & = a_2 = S^2(a_0) \end{array}$$

Bei inventierbarem ϕ erhalten wir also

$$S(a_0) = \phi \circ T \circ \phi^{-1}(a_0),$$

wobei $\sigma \in \phi^{-1}(a_0)$.

Ansonsten können wir, allerdings abhängig von der Wahl von $\sigma \in \phi^{-1}(a_0)$, definieren:

$$S(a_0) = \phi\big(T(\sigma)\big) \quad \text{wobei} \quad \sigma \in \phi^{-1}(a_0).$$

Wegen $a_0 = \lim_{n\to\infty} w_{\sigma_1} \circ w_{\sigma_2} \circ w_{\sigma_3} \circ \cdots \circ w_{\sigma_n}(x)$ ist $a_0, a_1, a_2, a_3, \ldots$ der "Rückwärts"Orbit von a_0 (bei invertiebarem w_i klar wegen

$$a_1 = w_{\sigma_1}^{-1} \lim_{n\to\infty} w_{\sigma_1} \circ w_{\sigma_2} \circ \ldots (x) = \lim w_{\sigma_2} w_{\sigma_3} \ldots (x) = \phi\big(T(\sigma)\big)$$

usw., also

$$\begin{aligned} a_0 &= w_{\sigma_1} \circ w_{\sigma_2} \circ w_{\sigma_3} \circ \cdots \circ w_{\sigma_n} \circ \ldots (a_0) \\ a_1 &= w_{\sigma_1}^{-1}(a_0) \\ a_2 &= w_{\sigma_2}^{-1} \circ w_{\sigma_1}^{-1}(a_0) = w_{\sigma_2}^{-1}(a_1) \\ &\vdots \\ a_n &= w_{\sigma_n}^{-1}(a_{n-1}). \end{aligned}$$

Wir erhalten also folgende allgemeinere Definition(vgl. unten Definition 2.1)

$$\{A; S\} = \left\{ \Big(\phi\big(T^n(\sigma)\big)\Big)_{n=0}^{\infty} : \sigma \in \phi^{-1}(a), a \in A \right\}.$$

DEFINITION 1.2. Zwei metrische Räume $(X_1, d_1), (X_2, d_2)$ heißen *topologisch äquivalent*, falls sie bezüglich ihrer metrischen Topologien homöomorph sind. Allgemeiner heißen Teilmengen metrischer Räume $S_1 \subset X_1, S_2 \subset X_2$ topologisch äquivalent, falls (S_1, d_1) homöomorph ist zu (S_2, d_2).
S_1, S_2 heißen *metrisch äquivalent*, falls $(S_1, d_1), (S_2, d_2)$ äquivalente metrische Räume sind.

Z.B. ist der Attraktor eines total unzusammenhängenden IFS topologisch äquivalent zu einer klassischen Cantor-Menge (3 Satz 2.5). Die topologische Äquivalenz erlaubt mehr "Streckung und Kompression" als metrische Äquivalenz. Die Fraktale Dimension wird sich später als invariant unter metrischer Äquivalenz herausstellen. Wenn es auf das Aussehen von Bildern ankommt, verwenden wir den eingeschränkten Begriff der metrischen Äquivalenz, um Begriffe wie "Ähnlichkeit" von Bildern zu definieren. In der dynamischen System-Theorie sind wir an der Art der Punkt-Bewegung, an der Existenz von Orbits mit bestimmten Eigenschaften (Periodizität, asymptotisches Verhalten usw.) interessiert. Diese sind meist invariant unter Homöomorphie. Daher ist dies für die Äquivalenz von dynamischen Systemen der adäquate Begriff.

DEFINITION 1.3. Zwei dynamische Systeme $\{X_1; f_1\}$, $\{X_2; f_2\}$ heißen äquivalent oder topologisch konjungiert, falls eine Homöomorphie $\Theta : X_1 \to X_2$ existiert, für die gilt

$$f_1(x_1) = \Theta^{-1} \circ f_2 \circ \Theta(x_1) \quad \forall x_1 \in X_1$$
$$f_2(x_2) = \Theta \circ f_1 \circ \Theta^{-1}(x_2) \quad \forall x_2 \in X_2$$

Wir können dies auch mit folgendem kommutativen Diagram ausdrücken:

$$\begin{array}{ccc} X_1 & \xrightarrow{f_1} & X_1 \\ \Theta\big\downarrow & & \Theta\big\downarrow \\ X_2 & \xrightarrow{f_2} & X_2 \end{array}$$

SATZ 1.2. *$\{X; w_1, w_2, \ldots, w_N\}$ sei ein total unzusammenhängendes IFS mit zugeordnetem DSDS $\{A; S\}$ und zugeordnetem Adreß-Raum Σ über N Symbolen und dem Verschiebungsoperator $T : \Sigma \to \Sigma$, definiert durch $T(\sigma_1\sigma_2\sigma_3\ldots) = \sigma_2\sigma_3\ldots$ für alle $\sigma = \sigma_1\sigma_2\sigma_3\cdots \in \Sigma$.*
Dann sind via der Abbildung $\phi : \Sigma \to A$ die dynamischen Systeme $\{\Sigma; T\}$ und $\{A; S\}$ äquivalent. Weiters gilt $\{a_1, a_2, \ldots, a_p\}$ ist abstoßender Zyklus der Periode p für S genau dann, wenn $\{\phi^{-1}(a_1), \phi^{-1}(a_2), \ldots, \phi^{-1}(a_p)\}$ abstoßender Zyklus der Periode p für T ist.

BEWEIS. Übung. □

2. Shiftdynamik bei überlappendenen IFS

Sei $\{X; w_1, w_2, \ldots, w_N\}$ ein IFS mit Attraktor A und Adreß-Raum Σ. Gemäß den Ausführungen nach Definition 5.2 definieren wir zu $\sigma \in \Sigma \quad \sigma = \sigma_1\sigma_2\sigma_3\ldots$ und

$a_0 = \phi(\sigma) \in A$ den Orbit $(a_n)_0^\infty$ durch

$$\begin{aligned} a_1 &= S(a_0) = \phi(\sigma_1\sigma_2\dots) = \phi\left(T(\sigma)\right) \\ &\dots \\ a_n &= S^n(a_0) = \phi\big(T^n(\sigma)\big). \end{aligned}$$

Umgekehrt: Jedes $a \in A$ hat eine Menge von Adressen $\phi^{-1}(a)$. Diese Menge ist einpunktig, wenn das IFS total unzusammenhängend ist. Und mit jeder dieser Adressen können wir einen, bei a beginnenden Orbit definieren.

Definition 2.1. Die Gesamtheit der wie oben gebildeten Orbits zu allen Adressen aller Punkte $a \in A$ heißt Zufalls-Shift Dynamisches System (Random Shift Dynamical System RSDS) $\{A; S\}$

Diese Definition ist unabhängig davon, ob die w_i auf A invertierbar sind. Wenn diese invertierbar sind, dann ist $a_{n+1} = w_i^{-1}(a_n)$ falls $a_n \in w_i(A)$ und $a_n \notin w_j(A)$ für $j \neq i$. Ist hingegen $a_n \in w_{i_1}(A) \cap w_{i_2}(A) \cap \dots \cap w_{i_k}(A)$ mit paarweise verschiedenen $i_1, i_2, \dots, i_k$, dann ist $a_{n+1} = w_j^{-1}(a_n)$ mit $j \in \{i_1, i_2, \dots, i_k\}$ willkürlich gewählt. Wir zeigen nun, daß zu jedem RSDS ein DSDS existiert, dessen Projektion das RSDS ist. Der Einfachheit halber beschränken wir uns auf den Fall eines IFS mit nur $N = 2$ Transformationen.

Zunächst eine Vorbemerkung. Einem Adreß-Raum Σ über N Symbolen können wir durch die Definition

$$t_n(\sigma) := n\sigma \qquad (n \in \{1, 2, \dots, N\},\ \sigma \in \Sigma)$$

ein IFS $\{\Sigma; t_1, \dots, t_N\}$ zuordnen. Es ist nicht schwer zu zeigen:
1) alle t_i sind kontrahierend.
2) $\Sigma = t_1\Sigma \cup t_2\Sigma \cup \dots \cup t_N\Sigma$ ist disjunkte Vereinigung.
Das IFS $\{\Sigma; t_1, \dots, t_N\}$ ist also total unzusammenhängend, sein Attraktor ist Σ und sein zugehöriges DSDS ist $\{\Sigma; T\}$.

Sei also $\{X; w_1, w_2\}$ ein IFS mit Adreß-Raum Σ. $\{\Sigma; t_1, t_2\}$, $\quad t_1, t_2 : \Sigma \to \Sigma$ seien definiert wie oben, nämlich $t_1(\sigma) = 1\sigma$ und $t_2(\sigma) = 2\sigma$. Damit definieren wir ein sogenanntes *angehobenes (lifted)* IFS $\{X \times \Sigma; \tilde{w}_1, \tilde{w}_2\}$: für $(x, \sigma) \in X \times \Sigma$ sei

$$\begin{aligned} \tilde{w}_1(x, \sigma) &:= \big(w_1(x), t_1(\sigma)\big) = \big(w_1(x), 1\sigma\big) \\ \tilde{w}_2(x, \sigma) &:= \big(w_2(x), t_2(\sigma)\big) = \big(w_2(x), 2\sigma\big). \end{aligned}$$

Dieses hat also ebenfalls den Adreß-Raum Σ.

Wir bestimmen nun den Attraktor $\tilde{A}$ von $\{X \times \Sigma; \tilde{w}_1, \tilde{w}_2\}$:
Nach 3 Satz 2.2 ist $\tilde{A} = \tilde{\phi}(\Sigma)$, wobei $\tilde{\phi} : \Sigma \to X \times \Sigma$ definiert ist durch

$$\begin{aligned} \tilde{\phi}(\sigma) = \tilde{\phi}(\sigma_1\sigma_2\dots\sigma_n\dots) &= \lim_{n\to\infty} \tilde{w}_{\sigma_1} \circ \tilde{w}_{\sigma_2} \circ \dots \circ \tilde{w}_{\sigma_n}(x, \tau) = \\ &= \lim_{n\to\infty} \big(w_{\sigma_1} \circ w_{\sigma_2} \circ \dots \circ w_{\sigma_n}(x), \sigma_1\sigma_2\dots\sigma_n\tau\big) = \big(\phi(\sigma), \sigma\big). \end{aligned}$$

Also

$$\tilde{A} = \Big(\Big(\phi(\sigma), \sigma\Big) : \sigma \in \Sigma\Big\} = \text{Graph von} \quad \phi : \Sigma \to X.$$

Wir beachten

$$\tilde{w}_1(\tilde{A}) = \Big\{ \Big(w_1\big(\phi(\sigma)\big), 1\sigma \Big) : \sigma \in \Sigma \Big\} = \big\{ \big(\phi(1\sigma), 1\sigma\big) : \sigma \in \Sigma \big\}$$

$$\tilde{w}_2(\tilde{A}) = \Big\{ \Big(w_1\big(\phi(\sigma)\big), 1\sigma \Big) : \sigma \in \Sigma \Big\} = \big\{ \big(\phi(2\sigma), 2\sigma\big) : \sigma \in \Sigma \big\}$$

Daher $\tilde{w}_1(\tilde{A}) \cup \tilde{w}_2(\tilde{A}) = \tilde{A}, \tilde{w}_1(\tilde{A}) \cap \tilde{w}_2(\tilde{A}) = \emptyset$.
$\{X \times \Sigma; \tilde{w}_1, \tilde{w}_2\}$ ist somit total unzusammenhängend. A ist also Projektion von $\tilde{A}$ auf die erste Komponente.

Definition 2.2. Das zum "angehobenenÏFS $\{X \times \Sigma; \tilde{w}_1, \tilde{w}_2\}$ gehörige DSDS $\{\tilde{A}; \tilde{S}\}$ heißt "Angehobenes" ("Lifted") Shift-Dynamisches System.

Wir bestimmen nun die zugehörige Shift-Abbildung $\tilde{S} : \tilde{A} \to \tilde{A}$.
Sei $\tilde{A} \ni \tilde{a} = (\phi(\sigma), \sigma) = (a, \sigma)$. Dann ist $\phi(\sigma) = a \in A, \quad \Sigma \ni \sigma = \sigma_1\sigma_2\sigma_3\ldots$ mit $\sigma_i \in \{1, 2\}$. Es ist

$$\tilde{S}(\tilde{a}) = \Big(\phi\big(T(\sigma)\big), t\sigma \Big) = \big(\phi(\sigma_2\sigma_3\ldots), \sigma_2\sigma_3\ldots \big) = \tilde{w}_{\sigma_1}^{-1}(\tilde{a}).$$

Für den Fall, daß die w_i invertierbar sind (oder jedenfalls a ein eindeutiges Urbild hat), gilt

$$\tilde{S}(\tilde{a}) = \tilde{w}_{\sigma_1}^{-1}(\tilde{a}) = \big(w_{\sigma_1}^{-1}(a), T(\sigma) \big),$$

wobei wieder $T\sigma = T(\sigma_1\sigma_2\sigma_3\ldots) = \sigma_2\sigma_2\ldots$. Wir fassen zusammen:

Satz 2.1 (Schatten-Theorem). *Sei $\{X; w_1, w_2\}$ ein IFS mit (nicht notwendig invertierbaren w_1, w_2 und) Attraktor A.*
Sei $\{x_n\}_{n=0}^{\infty}$ ein Orbit des RSDS $\{A; S\}$. Dann gibt es einen Orbit $\{\tilde{x}_n\}_{n=0}^{\infty}$ des angehobenen Systems $\{\tilde{A}; \tilde{S}\}$, sodaß die erste Komponente von $\tilde{x}_n$ gleich x_n ist für alle n.
m.a.W.: Jeder Zufalls-Orbit von $\{A; S\}$ ist Schatten (Projektion) eines deterministischen Orbit in $\{\tilde{A}; \tilde{S}\}$.

Für einen Zufallsorbit zu $\sigma \in \Sigma$, beginnend bei $x_0 = \phi(T^0\sigma) = \phi(\sigma)$, gilt also

$$\begin{aligned} x_n &= \phi(T^n\sigma) = \pi_1(\tilde{x}_n) \quad \text{wobei} \\ \tilde{x}_n &= \left(\phi\left(T^n\sigma\right), T^n\sigma\right). \end{aligned}$$

Beispiel 2.1.
Wir betrachten das gerade berührende IFS $\big\{[0,1]; w_1(x) = \frac{1}{2}x, w_2(x) = \frac{1}{2}x + \frac{1}{2}\big\}$ Es gilt: w_1, w_2 sind injektiv, $w_1^{-1}(x) = 2x$, $w_2^{-1}(x) = 2x - 1$ und $w_1[0,1] \cap w_2[0,1] = \{1/2\}$. Daher

$$S(x) = \begin{cases} 2x & x \in [0, \frac{1}{2}[\\ 2x - 1 & x \in]\frac{1}{2}, 1] \end{cases}$$

mit der Wahlmöglichkeit $S(1/2) = 0, 1$ im zugehörigen RSDS. Der Punkt $x_0 = 1/8$ hat 2 Adressen: $\sigma_1 = 00100\overline{0}, \quad \sigma_2 = 0001\overline{1}$. Dies ergibt respektive die beiden Orbits Die Punkte, deren Orbits durch die Definition von S betroffen sind, sind jene mit 2 Adressen, $\ldots 011\overline{1}$, bzw. $\ldots 100\overline{0}$, also die dyadisch rationalen.

ABBILDUNG 2. Aus [Ba], p.153

BEISPIEL 2.2.
Wir betrachten das überlappende IFS $\left\{[0,1]; w_1(x) = \frac{1}{2}x, w_2(x) = \frac{3}{4}x + \frac{1}{4}\right\}$. Es gilt: w_1, w_2 sind injektiv, $w_1^{-1}(x) = 2x,\ w_2^{-1}(x) = \frac{4}{3}x - \frac{1}{3}$ und $w_1[0,1] \cap w_2[0,1] = \{1/2\}$. Entscheiden wir uns im überlappenden Bereich $[1/4, 1/2]$ für eine der beiden Umkehrfunktionen bei der Definition von S, so gibt es entsprechend 2 mögliche DSDS bzw. 1 RSDS, entsprechend dem Schaubild.

Hier sind im schraffierten Bereich beide Umkehrfunktionen gleichberechtigt.

BEISPIEL 2.3.
Im IFS $\left\{\mathbf{R}; w_1(x) = \frac{1}{2}x, w_2(x) = \frac{3}{4}x + \frac{1}{4}\right\}$ ist $\Sigma = \{1,2\}^{\mathbf{N}}$ homöomorph zur Cantormenge $\mathcal{C} = \{0, \sigma_1\sigma_2 \cdots :$ triadisch $\sigma_i \in \{0,2\}\}$. $\mathcal{C}$ ist erzeugt vom IFS

$$\left\{\mathbf{R}; t_1(x) = \frac{1}{3}x, t_2(x) = \frac{1}{3}x + \frac{2}{3}\right\}$$

mit den Umkehrabbildungen $t_1^{-1}(x) = 3x,\ t_2^{-1}(x) = 3x - 2$. Das geliftete IFS hat somit die Darstellung

$$\left\{\mathbf{R} \times \mathcal{C} : \tilde{w}_1 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{3} \end{bmatrix}, \tilde{w}_2 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{3}{4} & 0 \\ 0 & \frac{1}{3} \end{bmatrix} + \begin{bmatrix} \frac{1}{4} \\ \frac{2}{3} \end{bmatrix}\right\}.$$

BEISPIEL 2.4.
Verwenden wir entsprechend Beispiel 2.2 das DSDS (I), so bedeutet dies, daß wir im dortigen RSDS bei der Bestimmung der Adresse eines Attraktorpunktes $x \in [0,1]$ bei gegebener Wahlmöglichkeit *immer* w_2 verwenden. Bestimmen wir also den Orbit eines x_0 unter Zugrundelegung von S gemäß (I), so gehen wir bei den möglichen Adressen von x_0 von jener (eindeutig bestimmten) Adresse σ aus, bei deren Bestimmung wir immer w_2 verwendet haben. Als Punkt der Cantormenge ist dieses σ also maximal unter allen Punkten, die Adressen von x_0 entsprechen. Sei

$$\tilde{A}_{top} := \left\{(x,y) \in \tilde{A} : (x,z) \in A \text{ (d.h. } z \text{ Adresse von } x) \Rightarrow z \leq y\right\},$$

dann erhalten wir
$\{x_n\}_0^\infty$ ist Orbit zum System (I) $\Leftrightarrow$
$\{(x_n, y_n)\}_0^\infty$ ist Orbit des *gehobenen* Systems $(*)$ in Beispiel 2.3 mit $(x_0, y_0) \in \tilde{A}_{top}$.

3. Sinnhaftigkeit berechneter Orbits

Wir haben Orbits in einem SDS mit Hilfe der Inversen von kontrahierenden Abbildungen berechnet. Diese sind aber expandierend, sodaß bei den Rechnungen große Fehlerfortpflanzungen zu erwarten sind. Es erhebt sich daher die Frage, wieweit ein berechneter Orbit überhaupt etwas mit einem "wirklichen" Orbit zu tun hat. Sei $\{X; w_1, w_2, \ldots, w_N\}$ IFS mit KF $0 < s < 1$ und Attraktor A und SDS $\{A; S\}$. Wir betrachten folgendes idealisierte Modell für die Beziehung zwischen dem exakten und dem berechneten Orbit eines $x_0 \in A$ mit einer Adresse $\sigma_1\sigma_2\sigma_3 \ldots \sigma_n \ldots$. Wir nehmen an, die $w_i : A \to A$ seien injektiv, sodaß exakt

$$x_1 = S(x_0) = w_{\sigma_1}^{-1}(x_0), \ldots x_n = S^n(x_0) = w_{\sigma_n}^{-1} \circ w_{\sigma_{n-1}}^{-1} \circ \cdots \circ w_{\sigma_1}^{-1}(x_0).$$

Bei der Computer-Berechnung treten Fehler auf

$$\tilde{x}_1 = \left(w_{\sigma_1}^{-1}\right)_c(x_0), \ldots \tilde{x}_n = \left(w_{\sigma_n}^{-1}\right)_c \circ \cdots \circ \left(w_{\sigma_1}^{-1}\right)_c(x_0) = S_c^n x_0.$$

Sei der exakte Orbit $\{x_n\}_{n=0}^{\infty}$ mit $x_n = S^n x_0$, der berechnete (=approximative) Orbit sei $\{\tilde{x}_n\}_{n=0}^{\infty}$ mit $\tilde{x}_n = S_c^n x_0 \quad x_0 = \tilde{x}_0$. Um überhaupt analysieren zu können wollen wir (optimistischerweise) annehmen, daß wenigstens bis zu einem gewissen M gilt $\tilde{x}_n, x_n \in w_{\sigma_n}(A) \quad (n = 0, 1, 2, \ldots, M)$ und daß lediglich die erste Berechnung (von $\tilde{x}_1$) fehlerbehaftet ist, bei den weiteren Berechnungen also **kein** Fehler mehr gemacht wird:

Sei also $d(\tilde{x}_1, x_1) = \vartheta > 0$ und für $n = 1, 2, \ldots, M$

$$\tilde{x}_{n+1} = w_{\sigma_n}^{-1}(\tilde{x}_n), x_{n+1} = w_{\sigma_n}^{-1}(x_n).$$

S wird duch die Inversen der w_i bestimmt, ist also expandierend, und zwar mindestens vom Faktor $1/s$ (ist doch $d\big(w_u(x), w_v(y)\big) \leq s\, d(x, y) \Leftrightarrow (1/s)d(u, v) \leq d\big(w^{-1}(u), w^{-1}(v)\big)$.)

Dann aber gilt $d(x_{n+1}, \tilde{x}_{n+1}) \geq s^{-n}\vartheta \quad$ für $\quad n = 0, 1, 2, \ldots, M$. D.h.: Lediglich unter der (ohnehin idealisierten) Annahme, daß nur ein einziges mal ein Rechenfehler auftritt, ergibt sich bei den ersten M Iterationen ein "exponentielles Auseinanderdriften". Es ist daher zu erwarten, daß für ein $J > M$ $x_J \in w_{\sigma_J}(A) \quad$ aber $\quad \tilde{x}_J \notin w_{\sigma_J}(A)$. Der berechnete Orbit hat also nichts mehr mit dem exakten zu tun. Einzig die triviale Fehlerabschätzung $d(x_n, \tilde{x}_n) \leq \operatorname{diam}(A)$ gilt immer. Sind also alle berechneten Orbits wetlos? Nein. Es gilt:

Satz 3.1. *Sei $\{X; w_1, w_2, \ldots, w_N\}$ ein IFS mit KF $\quad 0 < s < 1$ und Attraktor A; $w_n : A \to A$ sei injektiv, $\{A; S\}$ sei DSDS oder RSDS.*
Sei $\{\tilde{x}_n\}_{n=0}^{\infty} \subset A$ ein approximativer Orbit von S, mit $d(\tilde{x}_{n+1}, S(\tilde{x}_n)) \leq \vartheta$, wobei $0 \leq \vartheta \leq$ diam(A). Dann existiert $x_0 \in A$, für dessen exakten Orbit $\{x_n = S^n(x_0)\}_{n=0}^{\infty}$ gilt:

$$d(\tilde{x}_n, x_n) \leq \frac{s}{1-s}\vartheta \qquad \text{für alle} \quad n = 1, 2, \ldots \,.$$

Beweis. Wir führen den Beweis in einer Form, der das folgende Beispiel verständlich werden läßt und die tatsächliche Berechnung eines exakten Orbits sofort ermöglicht. Dazu sei $\sigma_n \in \{1, 2, \ldots, N\}$ so gewählt, wie es der tatsächlichen Iteration von S entspricht, sodaß also

$$S(\tilde{x}_0) = w_{\sigma_1}^{-1}(\tilde{x}_0),\ S(\tilde{x}_1) = w_{\sigma_2}^{-1}(\tilde{x}_1),\ S(\tilde{x}_2) = w_{\sigma_3}^{-1}(\tilde{x}_2),\ \ldots\,.$$

Wir betrachten zunächst einen endlichen Abschnitt $\tilde{x}_0, \tilde{x}_1, \dots, \tilde{x}_N$ des approx. Orbit. Nach Voraussetzung gilt

$$d\big(\tilde{x}_{n+1}, w_{\sigma_{n+1}}^{-1}(\tilde{x}_n)\big) \le \vartheta \qquad \text{für} \quad n = 1, 2, \dots, N.$$

Wir setzen

$$x_N^N := \tilde{x}_N,\ x_{N-1}^N := w_{\sigma_N}(\tilde{x}_N),\ \dots, x_0^N := w_{\sigma_1}(x_1) = w_{\sigma_1} \circ \cdots \circ w_{\sigma_N}(\tilde{x}_N).$$

Dann gilt $d(x_N^N, \tilde{x}_N) = 0$ und weiter

$$\begin{aligned}
&d(x_{N-1}^N, \tilde{x}_{N-1})\\
&= d\big(w_{\sigma_N}(x_N^N), \tilde{x}_{N-1}\big) \le s\, d\big(\tilde{x}_N, w_{\sigma_N}^{-1}(\tilde{x}_{N-1})\big) \le s\vartheta \le \frac{s}{1-s}\vartheta.\\
&d(x_{N-2}^N, \tilde{x}_{N-2})\\
&= d\big(w_{\sigma_{N-1}}(x_{N-1}^N), \tilde{x}_{N-2}\big) \le s\, d\big(x_{N-1}^N, w_{\sigma_{N-1}}^{-1}(\tilde{x}_{N-2})\big)\\
&\le s\big(d(x_{N-1}^N, \tilde{x}_{N-1}) + d\big(\tilde{x}_{N-1}, w_{\sigma_{N-1}}^{-1}(\tilde{x}_{N-2})\big)\big)\\
&\le s(s\vartheta + \vartheta) \le s\vartheta(1+s) \le \frac{s}{1-s}\vartheta.
\end{aligned}$$

So fortfahrend erhalten wir

$$d(x_1^N, \tilde{x}_1) \le \vartheta s + \vartheta s^2 + \cdots + \vartheta s^{N-1} = \vartheta s\big(1 + s + \cdots + s^{N-2}\big) \le \frac{s}{1-s}\vartheta.$$

Wie immer, sei nun $\phi : \Sigma \to A$ die Adreß-Raum Abbildung zum IFS. Wir setzen

$$x_0 = \phi(\sigma_1\sigma_2\sigma_3\dots) = \lim_{n\to\infty} w_{\sigma_1} \circ w_{\sigma_2} \circ \dots w_{\sigma_n}(x) \qquad (x \in A)$$

und vergleichen den fehlerhaften Orbit $\{\tilde{x}_n\}_{n=0}^{\infty}$ mit dem exakten Orbit $\{x_n = S^n(x_0) = \phi(\sigma_{n+1}\sigma_{n+2}\dots)\}_{n=0}^{\infty}$. Sei nun $N > n$.
Dann gilt $d(x_n, \tilde{x}_n) \le d(x_n, x_n^N) + d(x_n^N, \tilde{x}_n)$. Nun ist

$$\begin{aligned}
&d\big(x_n, x_n^N\big)\\
&= d\big(w_{\sigma_n} \circ \cdots \circ w_{\sigma_N}\big(\lim_{k\to\infty} w_{\sigma_{N+1}} \circ \cdots \circ w_{\sigma_k}(x)\big), w_{\sigma_n} \circ \cdots \circ w_{\sigma_N}(\tilde{x}_N)\big)\\
&\le s^{N-n+1}|A|.
\end{aligned}$$

Zusammenfassend erhalten wir

$$d(x_n, \tilde{x}_n) \le s^{N-n+1}|A| + \frac{s}{1-s}\vartheta$$

gültig für alle $N > n$. Es folgt für $N \to \infty$ und für $n = 1, 2, \dots$

$$d(x_n, \tilde{x}_n) \le \frac{s}{1-s}\vartheta.$$

□

Beispiel 3.1.
Wir betrachten das gerade berührende IFS $\big\{\mathbf{C}; w_1(z) = \frac{1}{2}z, w_2(z) = \frac{1}{2}z + \frac{1}{2}, w_3(z) = \frac{1}{2}z + \frac{i}{2}\big\}$ mit dem Sierpinski-Dreieck $\mathcal{S}$ als Attraktor. Man überlege sich, daß, abgesehen von dem "kleinen" Bereich, wo das IFS gerade berührend ist, die Shift-Transformation $S : \mathcal{S} \to \mathcal{S}$ gegeben ist durch

$$S(z) = S(x + i\,y) = (2x \, mod 1) + i(2y \, mod 1).$$

Der Vergleich eines "fehlerhaften" Orbits mit einem exakten Orbit gemäß Satz 3.1 findet sich am Kapitelende und ist entnommen aus [**Ba**], p.162. Man verifiziere, daß das dort angegebene x_0 identisch ist mit dem x_0^{10}, gemäß dem vorhergehenden Beweis.

4. Chaotische Dynamik auf Fraktalen

Wir betrachten das DSDS $\{A; S\}$ zu einem *total unzusammenhängenden* IFS. Dieses ist äquivalent zum DSDS $\{\Sigma; T\}$ über dem Adreß-Raum. D.h. beide Systeme haben einiges gemeinsam.
Z.B. haben beide Systeme dieselbe Anzahl von Zyklen einer vorgegebenen minimalen Periode und beide sind "chaotisch". Aber vom geometrischen Standpunkt aus können sie total verschieden sein, was die Wechselwirkung zwischen der Dynamik und den zugrunde liegenden Räumen betrifft. Z.B. haben alle IFS mit 3 Transformationen dasselbe $\{\Sigma; T\}$, ihre Attraktoren jedoch können verschieden sein.

Definition 4.1. Sei (X, d) metrischer Raum. $B \subset X$ heißt *dicht* in X, wenn $\bar{B} = X$. $\{x_n\}_{n=0}^{\infty}$ heißt *dicht in* $X :\Leftrightarrow$ zu jedem $x \in X$ existiert eine Teilfolge $\{x_{\sigma_n}\}_{n=0}^{\infty}$ mit $x_{\sigma_n} \to x$.

In diesem Sinne heißt ein Orbit des DS $\{X; f\}$ dicht in X. Experimente zeigen, daß der Random-Iterations- Algorithmus für $a_0 \in A$ i.a. einen dicht liegenden Orbit produziert.

Definition 4.2. Ein Dynamisches System $\{X; f\}$ heißt *transitiv*, falls zu jedem Paar U, V in X offener Mengen ein n existiert, mit $U \cap f^n(V) \neq \phi$.

Man überlegt sich leicht, daß diese Eigenschaft invariant ist unter topologischer Äquivalenz. Mit Hilfe des Baire'schen Kategoriensatzes kann man auch zeigen, daß jedenfalls bei Dynamischen Systemen über einem polnischen Raum (metrisch, vollständig, abzählbare Basis) dies gleichbedeutend ist mit der Eigenschaft, daß ein dicht liegender Orbit existiert.

Übung. Man zeige: $\{[0, 1]; f(x) = Min\{2x, 2 - 2x\}\}$ ist transitiv.

Definition 4.3. Ein DS $\{X; f\}$ heißt *sensitiv gegenüber Anfangsbedingungen*, wenn gilt:

$$\bigvee_{\delta>0} \bigwedge_{\substack{x\in X\\ \epsilon>0}} \bigvee_{\substack{y\in B(x,\epsilon)\\ n\geq 0}} \qquad d\big(f^n(x), f^n(y)\big) > \delta.$$

"Orbits, die nahe beeinander beginnen, werden also durch die Wirkung von f auseinander getrieben."

Beispiel 4.1.
1) $\{[0, 1]; 2x \mod 1\}$ ist sensitiv gegenüber Anfangsbedingungen und es gibt Punkte, deren Orbit dicht liegt in $[0, 1]$. (Sie sind sicher nicht dyadisch rational).

2) $\{[0, \infty); f(x) = 2x\}$ ist sensitiv gegenüber Anfangsbedingungen.

3) $\{[0, \infty); f(x) = \frac{1}{2}x\}$ ist nicht sensitiv gegenüber Anfangsbedingungen.

Definition 4.4. Ein Dynamisches System $\{X; f\}$ heißt *chaotisch*, falls gilt:
1) es ist transitiv,

2) es ist sensitiv gegenüber Anfangsbedingungen,
3) die Menge der periodischen Orbits ist dicht in X. (Ordnung im Chaos!)

SATZ 4.1. *Das DSDS, zugeordnet einem total unzusammenhängenden IFS mit wenigstens 2 Transformationen ist chaotisch.*

BEWEIS. Wir verwenden ein Adreß-Raum Argument:
$\phi : \Sigma \to A$ liefert die topologiche Äquivalenz zwischen $\{\Sigma; T\}$ und $\{A; S\}$. Da jedenfalls die Bedingungen 1) und 3) der vorigen Definition topologisch invariante Eigenschaften sind (Beweis?), genügt es zu zeigen, daß sie für $\{\Sigma; T\}$ gelten. Sei dazu

$$\Sigma \ni \sigma = \sigma_1\sigma_2\sigma_3\ldots \qquad T\sigma = \sigma_2\sigma_3\ldots \qquad d(\sigma,\tau) = \sum_{i=1}^{\infty} \frac{|\sigma_i - \tau_i|}{(N+1)^i}.$$

1) Zum Nachweis der Transitivität seien $\emptyset \neq U$, V offen in Σ und $\{1, 2, \ldots, N\} = X$.
V ist von der Form $V = \prod_1^\infty O_n$, wobei für höchstens endlich viele $n \quad O_n \neq X$. Sei $n_0 = Max\{n_i : 0_{n_i} \neq X\}$. Dann ist $T^{n_0}(V) = \Sigma$ und somit $U \cap T^{n_0}(V) = U \neq \emptyset$.
Für $N = 2$ ist hier ein Beispiel eines Punktes mit dicht liegendem Orbit

$$\begin{array}{llll} \sigma := 01| & 00\,01\,10\,11| & 000\,001\,010\,100\ldots 111| \ldots \\ \text{1. Block} & \text{2. Block} & \qquad\text{3. Block} \end{array}$$

Es handelt sich also um ein Nebeneinanderschreiben aller möglichen endlichen 01-Folgen fixer Länge. Eine geeignete Verschiebung stimmt mit jeder Folge an beliebig vielen Stellen überein.
2) Σ ist der Abschluß der periodischen Punkte, also ist die Gesamtheit der periodischen Orbits dicht in Σ. (Wir bemerken, daß in jedem Attraktor die periodischen Orbits dicht liegen.)
3) sensitive Abhängigkeit von den Anfangsbedingungen (für $\{\Sigma; T\}$):
Die Punkte einer ϵ-Umgebung $B(\sigma, \epsilon)$ von σ haben bis zu einem bestimmten (von ϵ abhängigen) N_0 gleiche Anfangskomponenten wie σ.
Zu $\sigma = \sigma_1\sigma_2\ldots\sigma_{N_0}\sigma_{N_0+1}\ldots$ sei $\sigma \neq \tau \in B(\sigma, \epsilon)$ mit $\tau = \sigma_1\sigma_2\ldots\sigma_{N_0}\tau_{N_0+1}\ldots$ und $\tau_{N_0+1} \neq \sigma_{N_0+1}$. Dann ist

$$T^{N^0}\sigma = \sigma_{N_0+1}\ldots$$
$$T^{N_0}\tau = \tau_{N_0+1}\ldots$$

und daher

$$d\left(T^{N_0}\sigma, T^{N_0}\tau\right) = \frac{|\tau_{N_0+1} - \sigma_{N_0+1}|}{N+1} + \cdots \geq \frac{1}{N+1} =: \delta.$$

□

Übung: Man zeige die sensitive Abhängigkeit allgemein für $\{A; S\}$. Tatsächlich ist $S : A \mapsto A$ sogar *expansiv*, d.h. die Orbits von 2 **beliebigen** verschiedenen Punkten liegen irgeneinmal um mehr als δ auseinander.

5. Warum funktioniert der Zufalls-Iterations-Algorithmus?

Für die folgende "intuitive" Begründung sei $\{X; w_1, \ldots, w_N; p_1, \ldots, p_N\}$ ein IFS mit Wahrscheinlichkeiten $0 < p_i < 1 \quad \sum_1^N p_i = 1$. Wir beschränken uns auf den Fall $\{\mathbf{R}^2; w_1, w_2\}$. Sei $a \in A$ mit Adresse $\sigma \in \Sigma$, also $a = \phi(\sigma)$. Wir wählen eine zufällige $1, 2$-Folge etwa der Länge 10^6. z.B. $2\,1 \ldots 1\,2\,1\,1\,1\,2\,2\,1\,2\,1\,2\,1\,1\,2\,2\,2\,1\,1$.
Dann wird die folgende Punktfolge auf A berechnet:

$$
\begin{aligned}
&a = \phi(\sigma)\\
&w_1(a) = \phi(1\sigma)\\
&w_1 \circ w_1(a) = \phi(11\sigma)\\
&w_2 \circ w_1 \circ w_1(a) = \phi(211\sigma)\\
&w_2 \circ w_2 \circ w_1 \circ w_1(a) = \phi(22\,11\sigma)\\
&w_2 \circ w_2 \circ w_2 \circ w_1 \circ w_1(a) = \phi(222\,11\sigma)\\
&\ldots\\
&\phi(21 \ldots 12\,11 \ldots 222\,11\sigma)
\end{aligned}
$$

Werden die Bildpunkte gespeichert und in umgekehrter Reihenfolge aufgelistet, so erhalten wir 1 Million Punkte auf dem Orbit des SDS $\{A; S\}$, nämlich

$$\left\{ S^n \left(\phi(21 \ldots 222\,11\sigma) \right) \right\}_{n=0}^{10^6}.$$

Wir sehen am Bildschirm den Teil eines "zufällig gewähltenÖrbits des SDS und erwarten, daß er dicht auf dem Attraktor liegt.

Lemma 5.1. *Sei $\{A; S\}$ das SDS zu einem total unzusamenhängenden IFS $\{X; w_1, \ldots, w_N\}$.*

Sei $\mathcal{N}(p)$ die Zahl der verschiedenen Zyklen minimaler Periode $p \in \mathbf{N}$.
Dann gilt:

$$p\mathcal{N}(p) = N^p - \sum_{\substack{k=1\\ k|p}}^{p-1} k\mathcal{N}(k).$$

Dies ist die Gesamtzahl der Punkte auf den veschiedenen Zyklen mit minimaler Periode $p \in \mathbf{N}$.

Beweis. Wir benützen die dynamische Identität zwischen Attraktor und Adreß-Raum. Ein Zyklus mit minimaler Periode k ist FP von T^k, enthält also ein Element σ mit $T^k\sigma = \sigma$, also $\sigma = \overline{\sigma_1 \ldots \sigma_k}$ und alle weiteren Elemente des Zyklus sind gegeben durch

$$\overline{\sigma_2 \ldots \sigma_k \sigma_1}, \overline{\sigma_3 \ldots \sigma_k \sigma_1 \sigma_2}, \ldots, \overline{\sigma_k \sigma_1 \ldots \sigma_{k-1}}.$$

Somit bestimmen k Folgen $\in \{1, 2, \ldots, N\}^k$ denselben Zyklus mit minimaler Periode k. Ist eine p-periodische Adresse auch k-periodisch mit $k < p$, so muß gelten $k|p$ (i.W.: k ist Teiler von p). Somit

$$p\mathcal{N}(p) = N^p - \sum_{\substack{k=1\\ k|p}}^{p-1} k\mathcal{N}(k).$$

□

z.B. $N = 2$
$p = 1$: Ein Zyklus der Periode 1 ist FP von T. Das sind $\bar{1}, \bar{2}$, also $\mathcal{N}(1) = 2$
$p = 2$: Ein Zyklus der Periode 2 ist FP von T^2. Das ergibt $\overline{11}, \overline{12}, \overline{21}, \overline{22}$, also ist $2\mathcal{N}(2) = 2^2 - 1\mathcal{N}(1) = 2$. Weiters gilt

$$\mathcal{N}(3) = 2, \mathcal{N}(4) = 3, \ldots, \mathcal{N}(20) = 52.377$$

Wieviele Punkte auf Zyklen der Periode 20 liegen auf Zyklen mit minimaler Periode 20? Es sind $20\,\mathcal{N}(20) = 1.047.540$.
Wieviele Punkte auf Zyklen der Periode 20 liegen auf Zyklen mit kleinerer minimaler Periode? Es sind

$$\begin{aligned} &1\,\mathcal{N}(1) + 2\,\mathcal{N}(2) + 4\,\mathcal{N}(4) + 5\,\mathcal{N}(5) \\ &\quad + 10\,\mathcal{N}(10) = 2 + 2.1 + 4.3 + 5.6 + 10.99 = 1046. \end{aligned}$$

D.h. $99,8\%$ aller Punkte auf Zyklen der Periode 20 liegen auf solchen mit minimaler Periode 20. Die periodischen Punkte liegen dicht im Attraktor A eines IFS. Wir erwarten daher, daß die Gesamtheit $\tilde{A}$ der periodischen Punkte einer festen endlichen Periode p (das sind 2^p) eine Approximaxion von A ergibt. Wählen wir einen dieser Punkte per Zufall. Dann liegt er sehr wahrscheinlich auf einem Zyklus mit **minimaler** Periode p. Sein Orbit besteht somit aus p **verschiedenen** Punkten und wir erwarten, daß diese über den ganzen Attraktor verteilt sind.

Eine klare Einsicht ergibt sich im Ramen der Ergodentheorie in Kap. 6

KAPITEL 5

W–Maße auf Fraktalen, Fraktale Dimension

1. $(\mathcal{P}(X), d_H)$, Hutchinson–Metrik

Sei (X, d) ein metrischer Raum und $\mathcal{B}(X)$ die von den offenen Teilmengen von X erzeugte σ-Algebra, also die σ-Algebra der Borel–Mengen. $C(X)$ sei der Raum aller reellwertigen, beschränkten, stetigen Funktionen auf X, $U(X)$ sei der Raum aller reellwertigen, beschränkten, gleichmäßig stetigen Funktionen auf X und $\mathfrak{Lip}_1(X)$ sei die Menge der reellwertigen Funktionen mit $Lipf \leq 1$, d.h.

$$\mathfrak{Lip}_1(X) := \{f : \mid f(x) - f(y) \mid \leq d(x, y)\}.$$

Referenz für die folgenden Sätze und Definitionen ist [**Pa**].

Satz 1.1. *a) Für $A \subset X$ ist die durch $x \mapsto d(x, A)$ definierte Funktion Element von $\mathfrak{Lip}_1(X)$. Es gilt also $|d(x, A) - d(y, A)| \leq d(x, y)$.*

b) Sind A, B abgeschlossene, disjunkte Teilmengen von X, dann gibt es eine stetige Funktion $f : X \to \mathbf{R}$ mit den Eigenschaften:

(1) $0 \leq f(x) \leq 1$;
(2) $f(x) = 0$ *für* $x \in A$ *und* $f(x) = 1$ *für* $x \in B$.

Falls $\inf\{d(a, b) : a \in A, b \in B\} = \delta > 0$, *dann kann f gleichmäßig stetig angenommen werden.*

Beweis. zu a) Sei $z \in A$ und $x, y \in X$. Dann gilt $d(x, A) \leq d(x, z) \leq d(x, y) + d(y, z)$. Daraus folgt $d(x, A) \leq d(x, y) + d(y, A)$. Aus $d(x, y) = d(y, x)$ folgt analog $d(y, A) \leq d(x, y)+d(x, A)$. Aus beidem zusammen daher $|d(x, A)-d(y, A)| \leq d(x, y)$.

zu b) Wähle $f(x) = d(x, A)/(d(x, A) + d(x, B))$. (1) und (2) folgen aus der Definition von f. Aus der letzten Bedingung folgt $d(x, A) + d(x, B) \geq \delta$ und damit die gleichmäßige Stetigkeit von f. □

Definition 1.1. $\mathcal{P}(X)$ bezeichne die Menge der nichtnegativen normalisierten Borelmaße auf X. $\mu \in \mathcal{P}(X)$ bedeutet also: μ ist ein nichtnegatives Borel–Maá auf X und $\mu(X) = 1$, kurz: ein Borel'sches W–Maß.

Satz 1.2. *Sei X ein separabler metrischer Raum und μ ein Borel'sches W–Maß auf X. Dann existiert eine eindeutig bestimmte, abgeschlossene Menge C_μ mit den Eigenschaften:*

(1) $\mu(C_\mu) = 1$
(2) *Falls für D abgeschlossen $\mu(D) = 1$, dann ist $C_\mu \subset D$.*

Die Menge C_μ heißt der Träger *oder das* Spektrum *von μ, in Zeichen $\operatorname{supp}\mu$ und ist auch charakterisiert durch die Eigenschaft: $\operatorname{supp}\mu$ ist die Menge der Punkte x, sodaß $\mu(U) > 0$ für alle offenen Umgebungen U von x.*

Falls X kompakt ist, dann gilt $\operatorname{supp}\mu \in \mathcal{H}(X)$.

Beweis. Sei $\mathcal{U} := \{U : U \text{ offen}, \mu(U) = 0\}$. Wegen der Separabilität von X gibt es offene $U_1, U_2 \dots$, sodaß $U_\mu := \bigcup U_n = \bigcup\{U : U \in \mathcal{U}\}$. Dann ist $C_\mu = X \backslash U_\mu = U'_\mu$ die behauptete Menge:

Jedenfalls ist C_μ abgeschlossen und wegen $\mu(U_\mu) = \mu(\bigcup U_n) \leq \sum \mu(U_n) = 0$ gilt $\mu(C_\mu) = 1$. Ist nun D eine abgeschlossene Menge mit $\mu(D) = 1$, dann ist $X \backslash D = D' \in \mathcal{U}$ und daher $D' \subset U_\mu$, also $C_\mu = U'_\mu \subset D$. Damit ist auch die Eindeutigkeit klar.

Falls $x \notin \operatorname{supp}\mu$, dann $x \in U_\mu$. U_μ ist aber eine offene x–Umgebung mit $\mu(U_\mu) = 0$. Falls umgekehrt $x \in \operatorname{supp}\mu$ und U eine offene x–Umgebung ist, dann muß $\mu(U) > 0$ gelten, da sonst definitionsgemäß $U \subset U_\mu$.

Falls X kompakt ist, dann ist $\operatorname{supp}\mu \in \mathcal{H}(X)$: Wäre nämlich $\operatorname{supp}\mu = \emptyset$, dann wäre das μ–Maß jeder Umgebung eines jeden Punktes aus X gleich 0, insbesondere wäre also $\mu(X) = 0$, im Widerspruch zu $\mu(X) = 1$, somit $\operatorname{supp}\mu \in \mathcal{H}(X)$. □

Sei $\mu \in \mathcal{P}(X)$. Eine Menge $A \in \mathcal{B}(X)$ heißt μ–regulär, falls

$$\begin{aligned}\mu(A) &= \sup\{\mu(C) : C \subset A, C \text{abgeschlossen}\}\\ &= \inf\{\mu(U) : U \supset A, U \text{offen}\}\end{aligned}$$

μ heißt *regulär*, falls alle $A \in \mathcal{B}(X)$ μ–reglär sind.

Satz 1.3. *Sei $\mu \in \mathcal{P}(X)$. Eine Menge $A \in \mathcal{B}(X)$ ist genau dann $\mu-$regulär, wenn zu jedem $\epsilon > 0$ abgeschlossene $C(\epsilon)$ und offene $U(\epsilon)$ existieren, sodaß*

(1) $C(\epsilon) \subset A \subset U(\epsilon)$
(2) $\mu(U(\epsilon) \backslash C(\epsilon)) < \epsilon$

Beweis. Übung. □

Satz 1.4. *Jedes Borel'sche W–Maß ist regulär.*

Beweis. Sei $\mathcal{R}$ die Menge aller μ–regulären Borel–Mengen in X. Dann ist $\mathcal{R} \subset \mathcal{B}(X), \emptyset, X \in \mathcal{R}$ (beide Mengen sind offen und abgeschlossen!)

$\mathcal{R}$ ist abgeschlossen gegenüber Komplementbildung:
Sei $A \in \mathcal{R}$ und $\epsilon > 0$. Dann existieren $C(\epsilon)$ abgeschlossen und $U(\epsilon)$ offen mit $C(\epsilon) \subset A \subset U(\epsilon)$ und $\mu(U(\epsilon) \backslash C(\epsilon)) < \epsilon$. Für die Komplemente gilt:
$U'(\epsilon) \subset A' \subset C'(\epsilon)$, $\quad C'(\epsilon) \backslash U'(\epsilon) = U(\epsilon) \backslash C(\epsilon)$; $U'(\epsilon)$ ist abgeschlossen und $C'(\epsilon)$ ist offen.

$\mathcal{R}$ ist abgeschlossen gegenüber abzählbaren Vereinigungen:
Seien $A_1, A_2, \dots \in \mathcal{R}$ und $A = \bigcup A_i$ und $\epsilon > 0$. Wegen $A_n \in \mathcal{R}$ existieren offene $U_n(\epsilon)$ und abgeschlossene $C_n(\epsilon)$ mit

$$C_n(\epsilon) \subset A_n \subset U_n(\epsilon) \quad \text{und} \quad \mu(U_n(\epsilon) \backslash C_n(\epsilon)) < \epsilon/3^n.$$

Sei $U(\epsilon) = \bigcup_n U_n(\epsilon)$ und $C = \bigcup_n C_n(\epsilon)$. Aus den Stetigkeitseigenschaften des Maßes μ folgt: es existiert ein k so groß, daß

$$\mu(C \backslash \bigcup_n^k C_n(\epsilon)) < \epsilon/2.$$

Sei $C(\epsilon) = \bigcup_n^k C_n(\epsilon)$. Dann ist $U(\epsilon)$ offen, $C(\epsilon)$ abgeschlossen und $C(\epsilon) \subset A \subset U(\epsilon)$. Weiters gilt

$$\begin{aligned}\mu(U(\epsilon) \backslash C(\epsilon)) &\leq \mu(U(\epsilon) \backslash C) + \mu(C \backslash C(\epsilon)) \\ &\leq \sum_{n=1}^{\infty} \mu(U_n(\epsilon) \backslash C_n(\epsilon)) + \epsilon/2 \\ &< \sum_{n=1}^{\infty} \epsilon/3^n + \epsilon/2 = \epsilon.\end{aligned}$$

$\mathcal{R}$ ist somit eine σ–Algebra. Wir zeigen nun, daß $\mathcal{R}$ alle abgeschlossenen Mengen enthält. Daraus folgt dann $\mathcal{R} = \mathcal{B}(X)$. Sei dazu C eine abgeschlossene Teilmenge von X und $\epsilon > 0$. Nun gibt es eine abnehmende Folge offener Obermengen G_n von C mit $C = \cap_{n=1}^{\infty} G_n$ (etwa $G_n = \{x : d(x, C) < 1/n\}$). Es gilt $\mu(G_n) \to \mu(C)$. Daher existiert ein n_0 mit $\mu(G_{n_0} \backslash C) < \epsilon$. Setze $C(\epsilon) = C$, $U(\epsilon) = G_{n_0}$. Dann gilt also $\mu(U(\epsilon) \backslash C(\epsilon)) < \epsilon$ und C ist somit μ–regulär. □

Der folgende Satz ist ein nützliches Kriterium für die Gleichheit zweier Maße.

SATZ 1.5. *Sei (X, d) metrischer Raum. Zwei Maße $\mu, \nu \in \mathcal{P}(X)$ sind genau dann gleich, wenn für alle $f \in U(X)$*

$$\int_X f d\mu = \int_X f d\nu$$

BEWEIS. Sei C eine beliebige abgeschlossene Menge in X und sei wieder $G_n = \{x : d(x, C) < 1/n\}$. G_n ist offen und $C = \cap_{n=1}^{\infty} G_n$. C und G_n' sind disjunkte abgeschlossene Mengen und $\inf\{d(x, y) : x \in C, y \in G_n'\} \geq 1/n$. Nach Satz 1.1 existieren daher $f_n \in U(X)$ mit $f_n(x) = 0$ für $x \in G_n'$ $f_n(x) = 1$ für $x \in C$ und $0 \leq f_n(x) \leq 1$. Daraus folgt $1_C(x) \leq f_n(x) \leq 1_{G_n}(x)$. Also

$$\mu(C) \leq \int f_n d\mu = \int f_n d\nu \leq \nu(G_n).$$

Mit $n \to \infty$ folgt daraus $\mu(C) \leq \nu(C)$. Mit vertauschten Rollen folgt ebenso $\nu(C) \leq \mu(C)$ und daher $\mu(C) = \nu(C)$ für alle abgeschlossenen C. Aus der Regularität folgt daher $\mu = \nu$. □

Wir versehen nun $\mathcal{P}(X)$ mit der von $C(X)$ induzierten schwachen Topologie: Seien $\mu_\alpha, \mu \in \mathcal{P}(X)$. Eine verallgemeinerte Folge (Netz) (μ_α) konvergiert *schwach* (oder *in Verteilung*) gegen μ, in Zeichen $\mu_\alpha \to \mu$ (w), falls

$$\bigwedge_{f \in C(X)} \lim_\alpha \int_X f d\mu_\alpha = \int_X f d\mu$$

Wir können $\mathcal{P}(X)$ auch als Teilmenge des Dualraumes von $C(X)$, versehen mit der w^*–Topologie, auffassen. Die obige Topologie ist dann die Relativtopologie und $\mathcal{P}(X)$ damit ein Hausdorff–Raum. Oder direkt: die schwache Topologie auf $\mathcal{P}(X)$ ist die gröbste Topologie, sodaß alle Abbildungen $p_f : \mathcal{P}(X) \to \mathbf{R}$, definiert durch $\mu \mapsto p_f(\mu) = \int_X f\, d\mu$, stetig sind. $\{p_f : f \in C(X)\}$ ist aber nach dem vorigen Satz separierend.

Der folgende Satz enthält wichtige äquivalente Definitionen für die schwache Konvergenz.

Satz 1.6. *Sei (μ_α) ein Netz in $\mathcal{P}(X)$. Folgende Aussagen sind äquivalent:*

(1) $\mu_\alpha \to \mu(w)$
(2) $\lim_\alpha \int_X f d\mu_\alpha = \int_X f d\mu$ für alle $f \in U(X)$.
(3) $\limsup_\alpha \mu_\alpha(C) \le \mu(C)$ für alle abgeschlossenen C.
(4) $\liminf_\alpha \mu_\alpha(D) \ge \mu(D)$ für alle offenen D.
(5) $\lim_\alpha \mu_\alpha(A) = \mu(A)$ für alle Borelmengen A *mit* $\mu(\partial A) = 0$.

Beweis. (1)⇒(2): dies ist klar wegen $U(X) \subset C(X)$.

(2)⇒(3): Sei C abgeschlossen. $G_n := \{x : d(x, C) < 1/n\}$ ist offfene Umgebung von C. C und G_n' sind disjunkt und abgeschlossen und $\inf\{d(x, y) : x \in C, y \in G_n'\} \ge 1/n$. Es gibt somit nach Satz 1.1 $f_n \in U(X)$ mit $0 \le f_n(x) \le 1$, $f_n(x) = 1$ für $x \in C$, $f_n(x) = 0$ für $x \in G_n'$. Weiters gilt $G_1 \supset G_2 \supset \cdots \to \bigcap G_n = C$. Daher

$$\begin{aligned}\limsup \mu_\alpha(C) &\le \limsup \int f_n d\mu_\alpha \\ &= \int f_n d\mu \le \mu(G_n) \quad (\text{da}\, f_n \le 1_{G_n})\end{aligned}$$

Wegen $\mu(G_n) \to \mu(C)$ folgt damit $\limsup \mu_\alpha(C) \le \mu(C)$.

(3)⇔(4): dies ist klar wegen der Komplementarität von *offen* und *abgeschlossen* und $\mu_\alpha(X) = \mu(X) = 1$.

(3),(4)⇒(5): Es ist $A^\circ \subset A \subset A^-$. Dabei ist A° der offene Kern und A^- der Abschluß von A. Sei nun $A \in \mathcal{B}(X)$ mit $\mu(A^- \setminus A^\circ) = \mu(\partial A) = 0$. Dann

$$\begin{aligned}\mu(A) = \mu(A^\circ) &\le \liminf \mu_\alpha(A^\circ) \le \liminf \mu_\alpha(A) \\ &\le \limsup \mu_\alpha(A) \le \limsup \mu_\alpha(A^-) \le \mu(A^-) = \mu(A).\end{aligned}$$

Somit $\mu(A) = \liminf \mu_\alpha(A) = \limsup \mu_\alpha(A)$. Daher existiert $\lim \mu_\alpha(A)$ und ist gleich $\mu(A)$.

(5)⇒(1): Sei $g \in C(X)$ und $\lim \mu_\alpha(A) = \mu(A)$ für alle $A \in \mathcal{B}(X)$ mit $\mu(A^- \setminus A^\circ) = 0$. Sei μ^g dass Bildmaß von μ unter g auf $\mathbf{R}$, d.h. die μ–Verteilung von $g : X \to \mathbf{R}$, definiert durch

$$\mu^g(E) = \mu\{g \in E\} = \mu\{x : g(x) \in E\} \quad (E \in \mathcal{B}(\mathbf{R})).$$

Da g beschränkt ist, ist μ^g konzentriert in einem beschränkten Intervall $[a, b]$. Die Verteilungsfunktion von μ^g hat höchstens abzählbar viele Sprungstellen. Zu $\epsilon > 0$

gibt es daher $t_0, t_1, \dots, t_m$ mit

$$\begin{aligned}
&1) \quad a = t_0 < t_1 \cdots < t_m = b\\
&2) \quad a < g(x) < b \quad \text{für alle} \quad x \in X\\
&3) \quad t_j - t_{j-1} < \epsilon \quad j = 1, \dots, m\\
&4) \quad \mu\{x : g(x) = t_j\} = 0 \quad j = 1, \dots, m
\end{aligned}$$

Sei $A_j = \{x : t_{j-1} \leq g(x) < t_j\}$. Dann gilt: $A_1, \dots, A_m$ sind disjunkte Borel–Mengen mit $X = \bigcup_1^m A_j$ und $A_j^- \backslash A_j^\circ \subset \{x : g(x) = t_{j-1}\} \bigcup \{x : g(x) = t_j\}$ und daher $\mu(A_j^- \backslash A_j^\circ) = 0$. Daraus folgt: $\lim \mu_\alpha(A_j) = \mu(A_j)$ für $j = 1, \dots, m$. Sei $g^* = \sum_1^m t_{j-1} 1_{A_j}$. Dann ist $|g^*(x) - g(x)| < \epsilon$ für alle $x \in X$ und somit

$$\begin{aligned}
&\left| \int g d\mu_\alpha - \int g d\mu \right| \leq\\
&\leq \int |g - g^*| d\mu_\alpha + \left| \int g^* d\mu_\alpha - \int g^* d\mu \right| + \int |g^* - g| d\mu\\
&\leq 2\epsilon + \sum_{j=1}^n |\mu_\alpha(A_j) - \mu(A_j)||t_{j-1}|.
\end{aligned}$$

Daher $\limsup_\alpha \left| \int g d\mu_\alpha - \int g d\mu \right| \leq 2\epsilon$ und somit $\lim_\alpha \int g d\mu_\alpha = \int g d\mu$. □

Wir untersuchen nun die Frage der Metrisierbarkeit und Kompaktheit von $\mathcal{P}(X)$ in Abhängigkeit von entsprechenden Eigenschaften von X.

Satz 1.7. *Vermöge der Abbildung $X \ni x \mapsto \delta_x \in \mathcal{P}(X)$ ist X homöomorph eingebettet in $\mathcal{P}(X)$.*

Beweis. Für $x \in X, g \in C(X)$ ist $\int g d\delta_x = g(x)$. Falls $x_\alpha \to x_0$, dann (wegen der Stetigkeit von g) $g(x_\alpha) \to g(x_0)$, also $\int g d\delta_{x_\alpha} \to \int g d\delta_{x_0}$. Das bedeutet aber $\delta_{x_\alpha} \to \delta_{x_0} \quad (w)$.

Sei umgekehrt $\delta_{x_\alpha} \to \delta_{x_0} \quad (w)$, aber $x_\alpha \not\to x_0$. Dann existiert ein Teilnetz $(x_\beta) \subset (x_\alpha)$ und eine offene Umgebung U von x_0, sodaß $(x_\beta) \subset U'$. U' ist abgeschlossen und daher existiert eine stetige Abbildung $f : X \to [0, 1]$ mit $f(x_0) = 0$ und $f(x) = 1$ auf U'. Daher $\int f d\delta_{x_\beta} = 1$ aber $\int f d\delta_{x_0} = 0$, im Widerspruch zur schwachen Konvergenz. □

Satz 1.8. *$D := \{\delta_x : x \in X\}$ ist sequentiell abgeschlossen in $\mathcal{P}(X)$, d.h. der Grenzwert jeder (w)–konvergenten* Folge *von δ–Funktionen ist wieder eine δ–Funktion.*

Beweis. Sei $q \in \mathcal{P}(X)$ und $\delta_{x_n} \to q\,(w)$ für eine Folge $\{x_n\}$ in X. Falls es eine Teilfolge $\{x_{n_k}\}$ und ein $x \in X$ gibt mit $x_{n_k} \to x$, dann folgt nach dem vorigen Satz $\delta_{x_k} \to \delta_x\,(w)$. Da $\mathcal{P}(X)$ Hausdorff–Raum ist, ist der Grenzwert jeder konvergenten Folge eindeutig. Somit gilt $q = \delta_x$. Wir führen den Beweis indirekt und nehmen also an, daß $\{x_n\}$ keine konvergente Teilfolge enthält.

Sei $S = \{x_1, x_2, \dots\}$ die der Folge $\{x_n\}$ zugrunde liegende Menge. o.B.d.A. enthält S unendlich viele Punkte (sonst enthält $\{x_n\}$ eine konstante und damit konvergente Teilfolge). S sowie jede Teilmenge $C \subset S$ sind abgeschlossen in X (C^- hätte nämlich sonst HP-e, die nicht zu C gehören, jedoch durch Teilfolgen von $\{x_n\}$ approximierbar wären). Wegen Satz 1.6 folgt $q(C) \geq \limsup \delta_{x_n}(C)$ für beliebige $C \subset S$. Nun ist $\limsup \delta_{x_n}(S_1) = 1$ für jede *unendliche* Teilmenge $S_1 \subset S$, also $q(S_1) = 1$. Dann aber ist q kein W–Maß (S_1 kann in 2 disjunkte unendliche Teile zerlegt werden, woraus folgte $q(S_1) = 2$). Aus diesem Widerspruch folgt die Existenz einer gegen ein $x \in X$ konvergenten Teilfolge von $\{x_n\}$. □

Satz 1.9. *X ist kompakt, metrisch genau dann, wenn $\mathcal{P}(X)$ kompakt, metrisch (metriserbar) ist.*

Beweis. Sei X kompakt, metrisch. Dann ist $C(X) = U(X)$ separabel. (Man konstruiert unter Zuhilfenahme von Satz 1.1 eine abzählbare, separierende Familie von Funktionen $f_i \in C(X)$. Die Menge der Polynome

$$\sum_k a_k\, f_{1k}^{n_{1k}} \cdots f_{lk}^{n_{lk}}$$

mit rationalen Koeffizienten a_k und mit $f_{jk} \in \{1_X, f_i\}$ und $n_{ij} \in \mathbf{N}$ ist abzählbar und liegt nach dem Satz von Stone–Weierstrass dicht in $C(X)$.) Es gibt also $g_1, g_2, \dots \in C(X)$ mit $g_1 \equiv 1, ||g_n|| \leq 1$, und $\{g_n\}$ dicht in der Einheitskugel $B(0,1) = \{f \in C(X) : ||f|| \leq 1\}$ von $C(X)$.

Wir betrachten nun die Abbildung $T : \mathcal{P}(X) \to I^\infty = [-1,1]^{\mathbf{N}}$, definiert durch $\mu \mapsto \{\int g_1 d\mu, \int g_2 d\mu, \dots\}$. I^∞ ist kompakt und metrisierbar. Wir zeigen nun, daß T eine Homöomorphie von $\mathcal{P}(X)$ auf $T(\mathcal{P}(X))$ ist:

T ist injektiv: Aus $T(\mu) = T(\nu)$ folgt nämlich $\int g_r d\mu = \int g_r d\nu$, damit $\int g d\mu = \int g d\nu$ für alle $f \in B(0,1)$ und daher $\mu = \nu$.

T ist stetig: Sei $\mu_\alpha \to \mu\,(w)$, dann gilt für $r \in \mathbf{N}$ $\int g_r d\mu_\alpha \to \int g_r d\mu$. Somit folgt $T(\mu_\alpha) \to T(\mu)$.

T^{-1} ist stetig: Sei $\{\mu_\alpha\}$ ein Netz in $\mathcal{P}(X)$ und $T(\mu_\alpha) \to T(\mu)$, d.h. (Produkttopologie!) $\int g_r d\mu_\alpha \to \int g_r d\mu$, ($r \in \mathbf{N}$). Dann gilt für alle $g \in B(0,1)$

$$\begin{aligned}\left|\int g d\mu_\alpha - \int g d\mu\right| &= \left|\int (g - g_r) d\mu_\alpha + \int g_r d\mu_\alpha - \int (g - g_r) d\mu - \int g_r d\mu\right| \\ &\leq 2||g - g_r|| + \left|\int g_r d\mu_\alpha - \int g_r d\mu\right| \to 2||g - g_r||\end{aligned}$$

Daher $\limsup |\int g d\mu_\alpha - \int g d\mu| \leq 2||g - g_r||$. Dies kann beliebig klein gemacht werden. Daher $\lim_\alpha \int g d\mu_\alpha = \int g d\mu$ für alle $g \in B(0,1)$. Ist nun $g \in C(X)$ beliebig, dann existiert ein $c \neq 0$ mit $cg \in B(0,1)$. Dann aber folgt $\lim_\alpha \int cg d\mu_\alpha = \int cg d\mu$ und damit $\lim_\alpha \int g d\mu_\alpha = \int g d\mu$, somit $\mu_\alpha \to \mu\,(w)$.

$T(\mathcal{P}(X))$ ist abgeschlossen in I^∞: Sei dazu $\{\mu_n\}$ eine Folge von Borel'schen W–Maßen mit $T(\mu_n) \to (\alpha_1, \alpha_2, \dots) \in I^\infty, g \in B(0,1)$. Wie vorhin folgt

$$\int g d\mu_m - \int g d\mu_n| \leq 2||g - g_r|| + |\int g_r d\mu_m - \int g_r d\mu_n|, \quad \text{also}$$
$$\limsup |\int g d\mu_m - \int g d\mu_n| \leq 2||g - g_r||.$$

Die rechte Seite kann beliebig klein gemacht werden, $\int g d\mu_n$ ist also Cauchy–Folge in $\mathbf{R}$ für alle $g \in B(0,1)$, hat also einen Grenzwert $\Lambda(g) := \lim_{n\to\infty} \int g d\mu_n$. Ist nun

$f \in C(X)$, dann sei wieder $c \neq 0$ mit $cf \in B(0,1)$. Definiert man $\Lambda(f) := c\Lambda(f/c)$, dann ist $\Lambda : C(X) \to \mathbf{R}$ ein nicht negatives lineares Funktional mit $\Lambda(1) = 1$. Nach Riesz existiert ein eindeutig bestimmtes $\mu(\in \mathcal{P}(X))$ mit $\Lambda(g) := \int g d\mu$. Insbesondere gilt also $\int g_r d\mu = \alpha_r$, d.h. $T(\mu) = (\alpha_1, \alpha_2, \dots)$ und $T(\mathcal{P}(X))$ ist daher abgeschlossen, also kompakt in I^∞. Somit ist $\mathcal{P}(X)$ kompakt.

Ist umgekehrt $\mathcal{P}(X)$ kompakt, metrisch, dann ist $D = \{\delta_x : x \in X\}$ nach dem vorigen Satz abgeschlossen, also kompakt in $\mathcal{P}(X)$. D ist aber homöomorph zu X. □

Sei nun (X,d) ein kompakter, metrischer Raum. $\mathcal{P}(X)$ ist also ebenfalls kompakt und metrisierbar. Wir betrachten die folgende Metrik, die s.g. Hutchinson–Metrik.

Definition 1.2. Zu $\mu, \nu \in \mathcal{P}(X)$ sei

$$d_H(\mu,\nu) := \sup\left\{\int_X f d\mu - \int_X f d\nu : f \in \mathfrak{Lip}_1(X)\right\}.$$

Satz 1.10. *Sei (X,d) ein kompakter, metrischer Raum. Dann ist durch $d_H(\mu,\nu)$ auf $\mathcal{P}(X)$ eine Metrik definiert. Die davon induzierte Topologie ist jene der schwachen Konvergenz und $(\mathcal{P}(X), d_H)$ ist daher ein kompakter, metrischer Raum.*

Beweis. Wir zeigen zunächst $d_H < \infty$: $\operatorname{supp}\mu$, $\operatorname{supp}\nu$ sind als kompakte Mengen beschränkt, d.h. es existieren $a \in X$ und $r \in (0,\infty)$ mit $\operatorname{supp}\mu \bigcup \operatorname{supp}\nu \subset K(a,r)$. Dann gilt für $f \in \mathfrak{Lip}_1(X)$ wegen $\mu(X) = \nu(X) = 1$

$$\begin{aligned}
&\left|\int f d\mu - \int f d\nu\right| = \\
&= \left|\int (f(y) - f(a) + f(a)) d\mu(y) - \int (f(y) - f(a) + f(a)) d\nu(y)\right| \\
&= \left|\int (f(y) - f(a)) d\mu(y) - \int (f(y) - f(a)) d\nu(y)\right| \\
&\leq \int |f(y) - f(a)| d\mu(y) + \int |f(y) - f(a)| d\nu(y) \\
&\leq \int d(a,y) d\mu(y) + \int d(a,y) d\nu(y) \leq 2r < \infty
\end{aligned}$$

d_H ist eine Metrik (Übung). Wir zeigen nun, daß d_H die Topologie der schwachen Konvergenz induziert und dazu zunächst:

$$\int f d\mu_\alpha \to \int f d\mu \quad (f \in C(X)) \Leftrightarrow \int g d\mu_\alpha \to \int g d\mu \quad (g \in \mathfrak{Lip}_1(X))$$

Die Notwendigkeit folgt aus $\mathfrak{Lip}_1(X) \subset C(X)$. Zum Nachweis, daß die Bedingung hinreichend ist, benützen wir Satz 1.6(3). Sei A abgeschlossen in X. Wir betrachten $A_n = \{x \in X : d(x,A) < 1/n\}$. Dann ist A_n' abgeschlossen, disjunkt zu A und $A_1 \supset A_2 \supset \dots \to \bigcap A_n = A$. Sei $M = |X|$. Wir betrachten die Funktionen

$$g_n(x) := \frac{1}{2Mn^2} \frac{d(x, A_n')}{d(x,A) + d(x,A_n')}$$

Dann ist $g_n \in \mathfrak{Lip}_1(X), g_n(x) = 0$ für $x \in A_n^{'}$, $2Mn^2 g_n(x) = 1$ für $x \in A$ und $1_A(x) \leq 2Mn^2 g_n(x)$. Daraus folgt für alle α $\mu_\alpha(A) \leq 2Mn^2 \int g_n d\mu_\alpha$ und daher

$$\begin{aligned}\limsup_\alpha \mu_\alpha(A) &\leq 2Mn^2 \lim_\alpha \int g_n(x) d\mu_\alpha(x)\\ &= 2Mn^2 \int g_n(x) d\mu(x)\\ &= \int \frac{d(x, A_n^{'})}{d(x,A) + d(x, A_n^{'})} d\mu(x) \leq \mu(A_n) \downarrow \mu(A)\end{aligned}$$

Somit gilt auch $\limsup \mu_\alpha(A) \leq \mu(A)$ und daher $\mu_\alpha \to \mu$ (w).

Sei nun $a \in X$ fest. Zu $f \in \mathfrak{Lip}_1(X)$ sei $\tilde{f} := f - f(a)$. Dann hat $\tilde{f}$ in a eine 0–Stelle und die Menge $\{\tilde{f} : f \in \mathfrak{Lip}_1(X)\}$ ist gleichmäßig beschränkt, besteht aus gleichgradig stetigen Funktionen und ist daher kompakt. Wegen $d_H(\mu, \nu) = \sup \left\{ \int_X \tilde{f} d\mu - \int_X \tilde{f} d\nu : f \in \mathfrak{Lip}_1(X) \right\}$ folgt daher die Behauptung.

Aus der gezeigten Gleichheit der metrischen Topologie bezüglich d_H und der Topologie der schwachen Konvergenz und aus dem vorigen Satz ergibt sich nun die letzte Behauptung. □

Sei nun (X, d) ein kompakter, metrischer Raum, $\{X; w_n; p_n, n = 1, \dots, N\}$ ein IFS mit Wahrscheinlichkeiten, kurz ein *Zufalls*–IFS. Jedes $w_i : X \to X$ ist als Kontraktion stetig, also $(\mathcal{B}(X), \mathcal{B}(X))$–meßbar. Zu $\mu \in \mathcal{P}(X)$ist also auch das Bildmaß unter w_i, definiert durch $\mu \circ w_i^{-1}(A) = \mu(w_i^{-1}(A))$, $(A \in \mathcal{B}(X))$, wieder ein Element von $\mathcal{P}(X)$und damit ebenso jede Konvexkombination der $\mu \circ w_i^{-1}$.

Definition 1.3. Zu einem Zufalls–IFS $\{X; w_n; p_n\}$ ist der Markov–Operator $M{:}\mathcal{P}(X) {\to} \mathcal{P}(X)$ für $\nu \in \mathcal{P}(X)$ und $A \in \mathcal{B}(X)$ definiert durch

$$M(\nu)(A) := \sum_{i=1}^N p_i \nu \circ w_i^{-1}(A) = p_1 \nu \circ w_1^{-1}(A) + \dots + p_N \nu \circ w_N^{-1}(A).$$

Wir benötigen die folgende Formel. Sie folgt sofort aus dem Integraltransformationssatz.

Lemma 1.11. *Für alle $f \in C(X)$ gilt*

$$\int_X f dM(\nu) = \sum_{i=1}^N p_i \int_X f \circ w_i d\nu.$$

Satz 1.12. *Sei (X, d) ein kompakter, metrischer Raum und $\{X; w_n; p_n\}$ ein Zufalls–IFS mit KF $s \in [0, 1)$. Dann ist der Markov–Operator $M{:}\mathcal{P}(X) {\to} \mathcal{P}(X)$ kontrahierend bezüglich der Hutchinson–Metrik d_H auf $\mathcal{P}(X)$und hat ebenfalls den KF s. Es gilt also für alle $\mu, \nu \in \mathcal{P}(X)$*

$$d_H(M(\mu), M(\nu)) \leq s\, d_H(\mu, \nu).$$

Es existiert daher ein eindeutig bestimmtes W–Maß $\mu \in \mathcal{P}(X)$mit $M(\mu) = \mu$. Dieses μ heißt das invariante *Maß des Zufall–IFS $\{X; w_n; p_n\}$.*

BEWEIS. Nach dem vorigen Lemma gilt

$$\begin{aligned} d_H(M(\mu), M(\nu)) &= \sup\left\{\int_X f dM(\mu) - \int_X f dM(\nu) : f \in \mathfrak{Lip}_1(X)\right\} \\ &= \sup\left\{\int_X \sum_{i=1}^N p_i \int_X f\circ w_i d\mu - \int_X \sum_{i=1}^N p_i \int_X f\circ w_i d\nu : f \in \mathfrak{Lip}_1(X)\right\} \end{aligned}$$

Sei $\bar{f} = s^{-1}\sum_{i=1}^N p_i f\circ w_i$. Dann ist $\bar{f} \in \mathfrak{Lip}_1(X)$. Sei L definiert durch

$$L := \left\{\bar{f} = \frac{1}{s}\sum_{i=1}^N p_i f\circ w_i : f \in \mathfrak{Lip}_1(X)\right\}.$$

Dann gilt

$$d_H(M(\mu), M(\nu)) = \sup\left\{s\int_X \bar{f} d\mu - s\int_X \bar{f} d\nu : \bar{f} \in L\right\}$$

Wegen $L \subset \mathfrak{Lip}_1(X)$ folgt daraus $d_H(M(\mu), M(\nu)) \leq s\, d_H(\mu, \nu)$. □

SATZ 1.13. *Sei $\{X; w_n; p_n\}$ ein Zufalls–IFS mit dem zugehörigen invarianten Maß μ. Dann ist der Träger von μ gleich dem Attraktor des IFS $\{X; w_n\}$, in Zeichen*

$$\operatorname{supp}\mu = A.$$

BEWEIS. Wir zeigen zuerst $\operatorname{supp}\mu \subset A$. Sei dazu ν irgend ein, in A konzentriertes Borel'sches W–Maß, d.h. also $\nu(A) = 1$ (Jedes δ_a für $a \in A$ ist ein Beispiel). Dann folgt wegen $w_i^{-1}(A) \supset A$

$$M(\nu)(A) = \sum_1^N p_i \nu(w_i^{-1}(A)) \geq \sum_1^N p_i \nu(A) = 1$$

und daher ebenso

$$M^2(\nu)(A) = M(M(\nu))(A) = \sum_1^N p_i M(\nu)(w_i^{-1}(A)) \geq \sum_1^N p_i M(\nu)(A) = 1$$

und durch Induktion allgemein $M^n(\nu)(A) = 1$ und damit nach Satz 1.6
$1 = \lim M^n(\nu)(A) \leq \mu(A)$, also nach Satz 1.2 $\operatorname{supp}\mu \subset A$.

Sei umgekehrt $a \in A$ und U eine offene a–Umgebung und $\sigma \in \Sigma$ eine Adresse von a. Nach Satz 2.2 Kap.3 gilt $\lim_{n\to\infty} \Phi(\sigma, n, A) = a$ in der Hausdorff–Metrik von $\mathcal{H}(X)$. Daraus folgt: es gibt ein $n \in \mathbf{N}$ mit $\Phi(\sigma, n, A) \subset U$. Nun ist aber (Übung)

$$\mu(\Phi(\sigma, n, A)) \geq p_{\sigma_1} p_{\sigma_2} \cdots p_{\sigma_n} > 0$$

und daher $\mu(U) > 0$. Nach Satz 1.2 gilt daher $a \in \operatorname{supp}\mu$ und somit
$A \subset \operatorname{supp}\mu$. □

SATZ 1.14 (Collage–Theorem, M. Barnsley). *Sei $\{X; w_n; p_n, n = 1, \ldots, N\}$ ein Zufalls–IFS mit zugehörigem Markov–Operator $M: \mathcal{P}(X) \to \mathcal{P}(X)$ und invariantem Maß μ und KF s.*
Sei $\nu \in \mathcal{P}(X)$. Dann gilt

$$d_H(\nu, \mu) \leq \frac{d_H(\nu, M(\nu))}{1 - s}$$

BEWEIS. Wie bei 5.1 Kap.2. □

2. Fraktale Dimension

Wie groß ist ein Fraktal? Wie läßt sich Ähnlichkeit von Fraktalen erfassen? Mit dem Begriff der Fraktalen Dimension versucht man das subjektive Emproofinden zu quantifizieren, wie dicht ein Fraktal ($\in \mathcal{H}(X)$) einen Teil des metrischen Raumes X erfüllt. Warum ist die Einschränkung auf kompakte Mengen dabei legitim? In der Physik hat man mit beschränkten Mengen zu tun und dabei kann nicht unterschieden werden zwischen einer Menge und ihrem Abschluß. Somit sind kompakte Mengen sinnvolle Modelle von Mengen der realen physikalischen Welt.

2.1. Box–counting Dimension. Sei $A \in \mathcal{H}(X), \epsilon > 0 \quad B(x,\epsilon) = \{y \in X : d(x,y) \leq \epsilon\}$. Sei $\mathcal{N}(A,\epsilon)$ die kleinste Zahl solcher Kugeln, sodaß

$$A \subset \bigcup_{n=1}^{M} B(x_n,\epsilon) \quad (x_n \in X).$$

$\mathcal{N}(A,\epsilon)$ existiert wegen der Kompaktheit von A. Man beachte: Für die gewöhnliche Dimension D gilt bei Überlagerung eines Würfels A mit kleineren Würfeln der Länge $\epsilon \quad \mathcal{N}(A,\epsilon) \approx C\,(1/\epsilon)^D$, somit $\log \mathcal{N}(A,\epsilon) \approx \log C + D \log(1/\epsilon)$, also

$$D \approx \frac{\log \mathcal{N}(A,\epsilon) - \log C}{\log(1/\epsilon)}.$$

Da $\frac{\log C}{\log(1/\epsilon)} \to 0$ für $\epsilon \downarrow 0$, ist folgende Definition motiviert.

Definition 2.1. Sei $A \in \mathcal{H}(X)$. Falls

$$D := \lim_{\epsilon \to 0} \frac{\log \mathcal{N}(A,\epsilon)}{\log(1/\epsilon)} \tag{2.1}$$

existiert, dann heißt D fraktale Dimension (Box–counting dimension) von A.

Beispiel 2.1.
Sei $A = [0,1] \subset (\mathbf{R}^2, |\,.\,|)$. Zu $\epsilon > 0$ ist $\mathcal{N}(A,\epsilon) = -[-1/\epsilon]$; denn für $0 < \epsilon < 1$ gilt $1/\epsilon \leq -[-1/\epsilon] \leq 1/\epsilon + 1$. Daher

$$\begin{aligned} 1 = \frac{\log(1/\epsilon)}{\log(1/\epsilon)} \leq \frac{\log(-[-1/\epsilon])}{\log(1/\epsilon)} &= \frac{\log(\mathcal{N}(A,\epsilon))}{\log(1/\epsilon)} \leq \frac{\log(1/\epsilon+1)}{\log(1/\epsilon)} \\ &= \frac{\log(1+\epsilon)+\log(1/\epsilon)}{\log(1/\epsilon)} \to 1 \end{aligned}$$

Satz 2.1. *Sei* $A \in \mathcal{H}(X), \quad \epsilon_n = C\,r^n \quad (0 > r > 1,\, C > 0) \quad n = 1,2,\ldots$ *Falls*

$$D = \lim_{n\to\infty} \left\{ \frac{\log(\mathcal{N}(A,\epsilon_n))}{\log(1/\epsilon_n)} \right\} \tag{2.2}$$

existiert, dann hat A fraktale Dimension D.

Beweis. Wir setzen $f(\epsilon) := \max\{\epsilon_n : \epsilon_n \leq \epsilon\}$. Sei $\epsilon \leq r$. Dann $f(\epsilon) = C\,r^{n_0} \leq \epsilon < C\,r^{n_0-1} = f(\epsilon)/r$ und somit

$$\mathcal{N}\big(A, f(\epsilon)\big) \geq \mathcal{N}(A,\epsilon) \geq \mathcal{N}\big(A, f(\epsilon)/r\big). \quad \Rightarrow$$

$$\frac{\log(\mathcal{N}(A, f(\epsilon)/r)}{\log(1/f(\epsilon))} \leq \frac{\log(\mathcal{N}(A,\epsilon))}{\log(1/\epsilon)} \leq \frac{\log(\mathcal{N}(A, f(\epsilon))}{\log(r/f(\epsilon))}. \tag{2.3}$$

o.B.d.A. $\mathcal{N}(A,\epsilon) \to \infty$ mit $\epsilon \to 0$. (Falls nicht, sind beide Seiten von (2.2) gleich 0 und daher nichts zu beweisen.) Für die rechte Seite von (2.3) gilt

$$\begin{aligned}\lim_{\epsilon\downarrow 0} \frac{\log \mathcal{N}(A, f(\epsilon))}{\log(r/f(\epsilon))} &= \lim_{n\to\infty} \frac{\log \mathcal{N}(A,\epsilon_n)}{\log(r/\epsilon_n)} \\ &= \lim_{n\to\infty} \frac{\log \mathcal{N}(A,\epsilon_n)}{\log r + \log(1/\epsilon_n)} \\ &= \lim_{n\to\infty} \frac{\log \mathcal{N}(A,\epsilon_n)}{\log(1/\epsilon_n)} = D.\end{aligned}$$

Für die linke Seite von (2.3) gilt:

$$\begin{aligned}\lim_{\epsilon\downarrow 0} \frac{\log \mathcal{N}(A, f(\epsilon)/r)}{\log(1/f(\epsilon))} &= \lim_{n\to\infty} \frac{\log \mathcal{N}(A,\epsilon_{n-1})}{\log(1/\epsilon_n)} \\ &= \lim_{n\to\infty} \frac{\log \mathcal{N}(A,\epsilon_{n-1})}{\log 1/r + \log(1/\epsilon_{n-1})} \\ &= \lim_{n\to\infty} \frac{\log \mathcal{N}(A,\epsilon_{n-1})}{\log(1/\epsilon_{n-1})} = D.\end{aligned}$$

Daher konvergiert die Mitte von (2.3) ebenfalls gegen D. □

Wir beweisen damit im konkreten Fall $X = \mathbf{R}^m$

Satz 2.2. *Sei $A \in \mathcal{H}(\mathbf{R}^m)$ unter Benützung der euklidischen Metrik. $\mathbf{R}^m$ wird überdeckt mit gerade berührenden Würfeln der Länge $1/2^n$. Sei $\mathcal{N}_n(A)$ die Zahl dieser Würfel, die A schneiden. Falls*

$$D = \lim_{n\to\infty} \frac{\log \mathcal{N}_n(A)}{\log 2^n} \tag{2.4}$$

existiert, dann hat A fraktale Dimension D.

(Genauso könnten wir Würfel der Länge $C\,r^n \quad C > 0 \quad 0 < r < 1$ verwenden)

Beweis. Da eine Kugel von Radius $1/2^n$ höchstens 2^m Würfel der angegebenen Art schneidet, gilt jedenfalls

$$\mathcal{N}_{n-1}(A) \le 2^m \mathcal{N}(A, 1/2^n).$$

Andererseits benötigt man zur Überdeckung eines Würfels dieser Art höchstens $2^m + 1$ Kugeln mit Radius $1/2^n$. Somit $\mathcal{N}(A,1/2^n) \le \mathcal{N}_{n-1}(A)(2^m+1)$. Daher $2^{-m}\mathcal{N}_{n-1}(A) \le \mathcal{N}(A,1/2^n) \le \mathcal{N}_{n-1}(A)(2^m+1)$. Also

$$\frac{\log 2^{-m}\mathcal{N}_{n-1}(A)}{\log 2^n} \le \frac{\log \mathcal{N}(A,1/2^n)}{\log 2^n} \le \frac{\log(2^m+1)\mathcal{N}_{n-1}(A)}{\log 2^n}$$

$$\frac{\log 2^{-m} + \log \mathcal{N}_{n-1}(A)}{\log 2 + \log 2^{n-1}} \le \frac{\log \mathcal{N}(A,1/2^n)}{\log 2^n} \le \frac{\log(2^m+1) + \log \mathcal{N}_{n-1}(A)}{\log 2 + \log 2^{n-1}}.$$

Die linke und die rechte Seite, und damit der mittlere Term der letzten Zeile konvergieren gegen D. □

Satz 2.3. *$(X_1,d_1),(X_2,d_2)$ seien metrisch äquivalent via $\Theta : X_1 \to X_2$. Sei $A_1 \in \mathcal{H}(X_1)$ mit fraktaler Dimension D. Dann hat auch $A_2 = \Theta(A_1)$ fraktale Dimension D, also $D(A_1) = D(\Theta(A_1))$.*

In $\mathbf{R}^m$ ist daher die fraktale Dimension unabhängig von der Norm-Metrik.

Beweis. Es gibt nach Voraussetzung positive Konstante c_1, c_2, sodaß

$$\bigwedge_{x,y\in X_1} \qquad c_1 d_2\big(\Theta(x),\Theta(y)\big) \le d_1(x,y) \le c_2 d_2\big(\Theta(x),\Theta(y)\big). \tag{2.5}$$

Die metrische Äquivalenz bleibt gültig unter der Annahme $c_1 < 1 < c_2$. Somit einerseits

$$d_2\big(\Theta(x),\Theta(y)\big) \le \frac{1}{c_1} d_1(x,y) \qquad (x,y \in X_1).$$

Daher $\Theta(B(x,\epsilon)) \subset B(\Theta(x),\epsilon/c_1)$. Weiters: Es gibt $\{x_1,x_2,\dots,x_{\mathcal{N}}\} \subset X_1$ mit $\mathcal{N} := \mathcal{N}(A_1,\epsilon)$ mit

$$A_1 \subset \bigcup_{n=1}^{\mathcal{N}} B(x_n,\epsilon).$$

Daher $A_2 = \Theta(A_1) \subset \bigcup_{n=1}^{\mathcal{N}} B(\Theta(x_n),\epsilon/c_1)$, also $\mathcal{N}(A_2,\epsilon/c_1) \le \mathcal{N}(A_1,\epsilon)$. Somit für alle $\epsilon < 1$ (sodaß $\log 1/\epsilon \ne 0$)

$$\frac{\log \mathcal{N}(A_2,\epsilon/c_1)}{\log(1/\epsilon)} \le \frac{\log \mathcal{N}(A_1,\epsilon)}{\log(1/\epsilon)} \Rightarrow$$

$$\begin{aligned}\limsup_{\epsilon\downarrow 0} \frac{\log \mathcal{N}(A_2,\epsilon)}{\log(1/\epsilon)} &= \limsup_{\epsilon\downarrow 0} \frac{\log \mathcal{N}(A_2,\epsilon/c_1)}{\log(1/\epsilon)} \\ &\le \lim_{\epsilon\downarrow 0} \frac{\log \mathcal{N}(A_1,\epsilon)}{\log(1/\epsilon)} = D(A_1).\end{aligned}$$

Nun zur umgekehrten Ungleichung. Aus (2.5) folgt

$$d_1\big(\Theta^{-1}(x),\Theta^{-1}(y)\big) \le c_2 d_2(x,y) \quad (x,y \in X_2). \qquad \Rightarrow$$

$$\Theta^{-1}(B(x,\epsilon)) \subset B(\Theta^{-1}(x), c_2\epsilon) \qquad (x \in X_2).$$

Somit $\mathcal{N}(A_1,c_2\epsilon) \le \mathcal{N}(A_2,\epsilon)$ also für $\epsilon < 1$

$$\frac{\log \mathcal{N}(A_1,c_2\epsilon)}{\log(1/\epsilon)} \le \frac{\log \mathcal{N}(A_2,\epsilon)}{\log(1/\epsilon)}$$

Damit erhalten wir

$$\begin{aligned}D(A_1) = \lim_{\epsilon\downarrow 0} \frac{\log \mathcal{N}(A_1,\epsilon)}{\log(1/\epsilon)} &= \lim_{\epsilon\downarrow 0} \frac{\log \mathcal{N}(A_1,c_2\epsilon)}{\log(1/\epsilon)} \\ &\le \liminf_{\epsilon\downarrow 0} \frac{\log \mathcal{N}(A_2,\epsilon)}{\log(1/\epsilon)} \qquad \Rightarrow\end{aligned}$$

$$\limsup_{\epsilon\downarrow 0} \frac{\log \mathcal{N}(A_2,\epsilon)}{\log(1/\epsilon)} \le D(A_1) \le \liminf_{\epsilon\downarrow 0} \frac{\log \mathcal{N}(A_2,\epsilon)}{\log(1/\epsilon)}, \quad \Rightarrow$$

$$D(A_2) = \lim_{\epsilon \downarrow 0} \frac{\log \mathcal{N}(A_2, \epsilon)}{\log(1/\epsilon)} = D(A_1).$$

$\square$

BEISPIEL 2.2.
1) Einheitsquadrat in $\mathbf{R}^2$: $\mathcal{N}_1(\square) = 4, \mathcal{N}_2(\square) = 16, \dots \mathcal{N}_n(\square) = 4^n$ somit

$$D(\square) = \lim_{n \to \infty} \frac{\log \mathcal{N}_n(\square)}{\log(2^n)} = 2.$$

2) Sierpinski-Dreieck $\mathcal{S}$: $\mathcal{N}_1(\mathcal{S}) = 3 \quad \mathcal{N}_2(\mathcal{S}) = 9 \quad \mathcal{N}_3(\mathcal{S}) = 27 \dots \mathcal{N}(\mathcal{S}) = 3^n$. Daher

$$\lim_{n \to \infty} \frac{\log \mathcal{N}_n(\mathcal{S})}{\log 2^n} = \lim_{n \to \infty} \frac{\log 3^n}{\log 2^n} = \frac{\log 3}{\log 2}.$$

3) Cantor-Menge $\mathcal{C}$: Mit Intervallen der Länge $c\, r^n = 1\,(1/3)^n$ gilt $\mathcal{N}_1 = 2$, $\mathcal{N}_2 = 2^2, \dots, \mathcal{N}_n = 2^n$ und daher

$$D(\mathcal{C}) = \lim_{n \to \infty} \frac{\log 2^n}{\log 3^n} = \frac{\log 2}{\log 3}.$$

4) Für die Cantor-Menge, die als Attraktor des IFS $\{[0,3]; \frac{1}{3}x, \frac{1}{3}x + 2\}$ erscheint, erhalten wir wegen der Isometrie $\big([0,1], |\ .\ |\big) \xrightarrow{3x} \big([0,3], |\ .\ |\big)$ dieselbe fraktale Dimension wie in 3)

Die folgende Definition ermöglicht es, einer größeren Klasse von Mengen eine Dimension zuzuordnen.

DEFINITION 2.2. (X, d) sei vollständiger metrischer Raum. Zu $A \in \mathcal{H}(X)$ sei $\mathcal{N}(\epsilon)$ die kleinste Zahl abgeschlossener Kugeln vom Radius ϵ für eine Überdeckung von A. Falls

(2.6)
$$D := \lim_{\epsilon \downarrow 0} \left\{ \sup \left\{ \frac{\log \mathcal{N}(\tilde{\epsilon})}{\log(1/\tilde{\epsilon})} \right\} : \tilde{\epsilon} \in (0, \epsilon) \right\} = \inf_{\epsilon > 0} \left\{ \sup \left\{ \frac{\log \mathcal{N}(\tilde{\epsilon})}{\log(1/\tilde{\epsilon})} \right\} : \tilde{\epsilon} \in (0, \epsilon) \right\}$$

existiert, dann heißt $D (= D(A))$ fraktale Dimension von A.

BEMERKUNG 2.1. Für eine, in einer 0-Umgebung definierte Funktion $f(\epsilon)$ ist

$$\limsup_{\epsilon \downarrow 0} f(\epsilon) = \lim_{\epsilon \downarrow 0} \{\sup\{f(\tilde{\epsilon}) : \tilde{\epsilon} \in (0, \epsilon)\}\}.$$

Man beachte: $\sup\{f(\tilde{\epsilon}) : \tilde{\epsilon} \leq \epsilon\}$ ist abnehmend in ϵ.

Wenn D im früheren Sinne existiert, dann auch im neuen und beide Größen sind gleich.

SATZ 2.4. *Sei $m \in \mathbf{N}$, $\mathbf{R}^m$ versehen mit der euklidischen Metrik. Es gilt: Für alle $A \in \mathcal{H}(\mathbf{R}^m)$ existiert $D(A)$. Ist $B \in \mathcal{H}(\mathbf{R}^m)$ mit $A \subset B$, dann gilt $D(A) \leq D(B)$. Insbesondere: $0 \leq D(A) \leq m$.*

Beweis. Für $m = 2$: Sei $0 < \epsilon < 1$ und ohne Beschränkung der Allgemeinheit sei $A \subset B \subset Q$ mit einem Quadrat Q und $0 < \epsilon < 1$. Somit $\mathcal{N}(A,\epsilon) \leq \mathcal{N}(B,\epsilon) \leq \mathcal{N}(Q,\epsilon)$. Daher

$$0 \leq \frac{\log \mathcal{N}(A,\epsilon)}{\log(1/\epsilon)} \leq \frac{\log \mathcal{N}(B,\epsilon)}{\log(1/\epsilon)} \leq \frac{\log \mathcal{N}(Q,\epsilon)}{\log(1/\epsilon)}.$$

Also

$$\limsup_{\epsilon\downarrow 0} \frac{\log \mathcal{N}(A,\epsilon)}{\log(1/\epsilon)} \leq \limsup_{\epsilon\downarrow 0} \frac{\log \mathcal{N}(B,\epsilon)}{\log(1/\epsilon)} \leq \limsup_{\epsilon\downarrow 0} \frac{\log \mathcal{N}(Q,\epsilon)}{\log(1/\epsilon)} = 2.$$

Letzteres nach dem vorigen Beispiel und Satz 2.2 □

Zur Berechnung der fraktalen Dimension der Vereinigung zweier Mengen ist folgender Satz nützlich.

Satz 2.5. *Seien $A, B \in \mathcal{H}(\mathbf{R}^m)$. Sei $D(A) = \lim_{\epsilon\downarrow 0} \dfrac{\log \mathcal{N}(A,\epsilon)}{\log(1/\epsilon)}$ und $D(B) \leq D(A)$. Dann gilt $D(A \cup B) = D(A)$.*

Beweis. Beweis Zuerst nehmen wir an: $D(B) < D(A)$. Jedenfalls gilt $D(A \cup B) \geq D(A)$. Wir zeigen nun $D(A \cup B) \leq D(A)$. Zunächst gilt zu $\epsilon > 0$: $\mathcal{N}(A \cup B,\epsilon) \leq \mathcal{N}(A,\epsilon) + \mathcal{N}(B,\epsilon)$. Somit

$$\begin{aligned} D(A \cup B) &= \limsup_{\epsilon\downarrow 0} \frac{\log \mathcal{N}(A \cup B,\epsilon)}{\log(1/\epsilon)} \leq \limsup_{\epsilon\downarrow 0} \frac{\log(\mathcal{N}(A,\epsilon) + \mathcal{N}(B,\epsilon)}{\log(1/\epsilon)} \\ &\leq \limsup_{\epsilon\downarrow 0} \frac{\log \mathcal{N}(A,\epsilon)}{\log(1/\epsilon)} + \limsup_{\epsilon\downarrow 0} \frac{\log(1 + \mathcal{N}(B,\epsilon)/\mathcal{N}(A,\epsilon)}{\log(1/\epsilon)}. \end{aligned}$$

Wir zeigen nun: $\mathcal{N}(B,\epsilon)/\mathcal{N}(A,\epsilon) < 1$ für hinreichend kleines ϵ. Damit ist der 2. Summand der letzten Formel, den wir mit $\limsup(II)$ abkürzen, 0, der erste ist per def. $D(A)$. Wir bemerken: $\sup\left\{\dfrac{\log \mathcal{N}(B,\tilde{\epsilon})}{\log(1/\tilde{\epsilon})} : \tilde{\epsilon} < \epsilon\right\}$ ist abnehmend in ϵ. Somit $\dfrac{\log \mathcal{N}(B,\epsilon)}{\log(1/\epsilon)} < D(A)$ für hinreichend kleine $\epsilon > 0$. Wegen der vorausgesetzten Existenz des lim für $D(A)$ gilt

$$\frac{\log \mathcal{N}(B,\epsilon)}{\log(1/\epsilon)} < \frac{\log \mathcal{N}(A,\epsilon)}{\log(1/\epsilon)}$$

für hinreichend kleine $\epsilon > 0$. Somit für diese ϵ:

$$\frac{\mathcal{N}(B,\epsilon)}{\mathcal{N}(A,\epsilon)} < 1.$$

Sei nun $D(A) = D(B)$, also

$$D(B) = \lim_{\epsilon\downarrow 0} \left\{ \sup_{\tilde{\epsilon}\leq\epsilon} \frac{\log \mathcal{N}(B,\tilde{\epsilon})}{\log(1/\tilde{\epsilon})} \right\} = \lim_{\epsilon\downarrow 0} \frac{\log \mathcal{N}(A,\tilde{\epsilon})}{\log(1/\epsilon)} = \lim_{\epsilon\downarrow 0} \left\{ \inf_{\tilde{\epsilon}\leq\epsilon} \frac{\log \mathcal{N}(A,\tilde{\epsilon})}{\log(1/\tilde{\epsilon})} \right\}.$$

Es gibt 2 Möglichkeiten: 1) $\mathcal{N}(B,\tilde{\epsilon})/\mathcal{N}(A,\tilde{\epsilon})$ ist beschränkt für $\tilde{\epsilon} \downarrow 0$. Dann folgt $\limsup_{\epsilon\to 0}(II) = 0$. 2) $\mathcal{N}(B,\tilde{\epsilon})/\mathcal{N}(A,\tilde{\epsilon})$ ist unbeschränkt für $\tilde{\epsilon} \downarrow 0$. Dann folgt

$$
\begin{aligned}
0 \le \sup_{\tilde\epsilon\le\epsilon}(II) &\le \sup_{\tilde\epsilon\le\epsilon} \frac{\log 2\mathcal{N}(B,\tilde\epsilon)/\mathcal{N}(A,\tilde\epsilon)}{\log(1/\tilde\epsilon)} \\
&\le \sup_{\tilde\epsilon\le\epsilon} \frac{\log 2}{\log(1/\tilde\epsilon)} + \sup_{\tilde\epsilon\le\epsilon}\left(\frac{\log\mathcal{N}(B,\tilde\epsilon)}{\log(1/\tilde\epsilon)} - \frac{\log\mathcal{N}(A,\tilde\epsilon)}{\log(1/\tilde\epsilon)}\right) \\
&\le \frac{\log 2}{\log(1/\epsilon)} + \sup_{\tilde\epsilon\le\epsilon}\frac{\log\mathcal{N}(B,\tilde\epsilon)}{\log(1/\tilde\epsilon)} - \inf_{\tilde\epsilon\le\epsilon}\frac{\log\mathcal{N}(A,\tilde\epsilon)}{\log(1/\tilde\epsilon)}.
\end{aligned}
$$

Somit

$$
\begin{aligned}
0 \le \lim\sup_{\tilde\epsilon\le\epsilon}(II) &\le \lim_{\epsilon\downarrow 0}\frac{\log 2}{\log(1/\epsilon)} + \lim_{\epsilon\downarrow 0} f(\epsilon) - \lim_{\epsilon\downarrow 0} g(\epsilon) \\
&= 0 + D(B) - D(A) = 0.
\end{aligned}
$$

$f(\epsilon)$, $g(\epsilon)$ sind dabei die entsprechenden Ausdrücke der vorletzten Zeile. □

Die tatsächliche Berechnung ist in vielen Fällen möglich mit folgendem

Satz 2.6. *Sei $\{\mathbf{R}^m; w_1, \dots, w_N\}$ ein IFS mit Attraktor A. w_n sei Ähnlichkeit mit Skalierungsfaktor s_n $(n = 1, 2, \dots, N)$. Falls das IFS total unzusammenhängend oder gerade berührend ist, dann ist die fraktale Dimension $D(A)$ gegeben durch die eindeutige Lösung der Gleichung*

$$
\sum_{n=1}^{N} |s_n|^{D(A)} = 1 \qquad \big(D(A) \in [0, m]\big). \tag{2.7}
$$

Falls das IFS überlappend ist, dann gilt für die Lösung $\bar D\big(\in [0,\infty)\big)$ von $\sum_{n=1}^N |s_n|^{\bar D} = 1$ $D(A) \le \bar D$.

Bemerkung 2.2. Die Eindeutigkeit der Lösung sieht man folgendermaßen ein: Für $\gamma(t) = \sum_{n=1}^N s_n^t$ gilt $\gamma(0) = N, \quad \gamma(t) \downarrow 0$ für $t \to \infty$. Daher existiert genau ein D mit $\gamma(D) = 1$.

Hier eine Beweisidee (Beweis später!) für ein total unzusammenhängendes IFS. Sie gründet auf 2 Bemerkungen: Sei $s_n > 0$ für alle $n \in \{1, 2, \dots, N\}$:

1) Jedes w_n ist Ähnlichkeit mit Skalierungsfaktor s_n, bildet somit abgeschlossene Kugeln in abgeschlossene Kugeln ab:

$$
w_n\big(B(x,\epsilon)\big) = B\big(w_1(x), s_n\epsilon\big).
$$

Wegen $s_n > 0$ ist w_n invertierbar und ist ebenfalls Ähnlichkeit, also

$$
w_n^{-1}\big(B(x,\epsilon)\big) = B\big(w_n^{-1}(x), s_n^{-1}\epsilon\big)\big).
$$

Daraus ergibt sich: $\mathcal{N}(A,\epsilon) = \mathcal{N}(w_n(A), s_n\epsilon)$, oder, äquivalent dazu $\mathcal{N}(A, s_n^{-1}\epsilon)$ $= \mathcal{N}(w_n(A), \epsilon)$.

2) Für ein total unzusammenhängendes IFS ist $A = w_1(A) \cup w_2(A) \cup \cdots \cup w_N(A)$ eine disjunkte Vereinigung der kompakten Mengen $w_n(A)$. Wegen des positiven Randabstandes disjunkter kompakter Mengen zerfällt für hinreichend kleine $\epsilon > 0$ eine

Überdeckung von A mit Kugeln $B(x,\epsilon)$ in disjunkte Teile, die jeweils $w_n(A)$ überdecken. Somit gilt für alle hinreichend kleinen $\epsilon > 0$

$$\begin{aligned}\mathcal{N}(A,\epsilon) &= \mathcal{N}\big(w_1(A),\epsilon\big) + \mathcal{N}\big(w_2(A),\epsilon\big) + \cdots + \mathcal{N}\big(w_N(A),\epsilon\big)\\ &= \mathcal{N}\big(A, s_1^{-1}\epsilon\big) + \mathcal{N}\big(A, s_2^{-1}\epsilon\big) + \cdots + \mathcal{N}\big(A, s_N^{-1}\epsilon\big).\end{aligned}$$

Nun folgt ein heuristisches Argument für **kleine** $\epsilon > 0$: $\mathcal{N}(A,\epsilon) \sim C\epsilon^{-D(A)}$, daher

$$C\epsilon^{-D} \sim Cs_1^D\epsilon^{-D} + Cs_2^D\epsilon^{-D} + \cdots + Cs_N^D\epsilon^{-D}.$$

Also

$$1 = s_1^D + s_2^D + \cdots + s_N^D.$$

Beispiel 2.3.
Dimension des Sierpinski Dreieckes:

$$\big(\frac{1}{2}\big)^D + \big(\frac{1}{2}\big)^D + \big(\frac{1}{2}\big)^D = 1 \Rightarrow D = \frac{\log 3}{\log 2}.$$

2.2. Hausdorff-Dimension. Sei $E \subset \mathbf{R}^m$, mit eukl. Metrik. Zu $\emptyset \neq U \subset \mathbf{R}^m$ definieren wir $|U| = diam\, U = \sup\{|x-y| : x, y \in U\}$. Falls $E \subset \bigcup_1^\infty U_i$ und $0 < |U_i| \leq \delta$, dann heißt $\{U_i\}$ eine δ-Überlagerung von E. Zu $s \geq 0$ definieren wir

$$\mathcal{H}_\delta^s(E) = \inf\left\{\sum_{i=1}^\infty |U_i|^s : \{U_j\} \quad \delta - \text{Überlagerung von } E\right\}.$$

Man zeigt leicht:

$$\begin{aligned} &1) \quad \mathcal{H}_\delta^s(\phi) = 0 \\ &2) \quad E \subset F \Rightarrow \mathcal{H}_\delta^s(E) \leq \mathcal{H}_\delta^s(F) \\ &3) \quad \mathcal{H}_\delta^s\left(\bigcup_{i=1}^\infty E_i\right) \leq \sum_{i=1}^\infty \mathcal{H}_\delta^s(E_i) \end{aligned}$$

d.h. $\mathcal{H}_\delta^s$ ist ein äußeres Maß. Weiters ist $\mathcal{H}_\delta^s(E)$ wachsend für $\delta \downarrow$; denn für $\delta' < \delta$ ist jede δ'-Überlagerung auch δ-Überlagerung und die kleinere Menge von Zahlen hat das größere inf. Damit definieren wir weiters:

$$\mathcal{H}^s(E) := \lim_{\delta \to 0} \mathcal{H}_\delta^s(E) = \sup_{\delta > 0} \mathcal{H}_\delta^s(E). \tag{2.8}$$

$\mathcal{H}^s(E)$ kann auch ∞ sein. Es gilt: Auch $\mathcal{H}^s$ ist äußeres Maß und zugleich **metrisch**, d.h.: Für E, F mit $d(E,F) = \inf\{|x-y| : x \in E, y \in F\} > \delta$ gilt:

$$\mathcal{H}_\delta^s(E \cup F) = \mathcal{H}_\delta^s(E) + \mathcal{H}_\delta^s(F).$$

(Keine Menge U einer δ-Überlagerung von $E \cup F$ für $\delta < d(E,F)$ kann gleichzeitig E und F treffen!) Hier ein kurzer Nachweis, daß $\mathcal{H}^s$ äußeres Maß ist:
1) $\mathcal{H}_\delta^s(\phi) = 0$ für alle $\delta > 0$ also $\mathcal{H}^s(\phi) = 0$
2) $E \subset F \Rightarrow \mathcal{H}_\delta^s(E) \leq \mathcal{H}_\delta^s(F)$ für alle $\delta > 0$, somit

$$\mathcal{H}_\delta^s(E) \leq \sup_{\delta > 0} \mathcal{H}_\delta^s(F) = \mathcal{H}^s(F) \qquad \text{also} \qquad \mathcal{H}^s(E) \leq \mathcal{H}^s(F).$$

3) Jede δ-Überlagerung von $\bigcup_{j=1}^\infty E_j$ liefert δ-Überlagerung von jedem E_j

$$\bigcup_1^\infty E_j \subset \bigcup_1^\infty U_i = \bigcup_{E_1 \cap U_i \neq \phi} U_i \cup \bigcup_{E_2 \cap U_i \neq \phi} U_i \ldots$$

somit

$$\sum_1^\infty |U_i|^s \leq \sum_{E_1 \cap U_i \neq \phi} |U_i|^s + \sum_{E_2 \cap U_i \neq \phi} |U_i|^s + \ldots$$

Daher

$$\mathcal{H}_\delta^s\left(\bigcup_{j=1}^\infty E_j\right) \leq \sum_1^\infty \mathcal{H}_\delta^s(E_j) \leq \sum_1^\infty \sup_{\delta > 0} \mathcal{H}_\delta^s(E_j) = \sum_1^\infty \mathcal{H}^s(E_j).$$

Also

$$\mathcal{H}^s\left(\bigcup_1^\infty E_j\right) \leq \sum_1^\infty \mathcal{H}^s(E_j).$$

Definition 2.3. $\mathcal{H}^s$ eingeschränkt auf die σ-Algebra der $\mathcal{H}^s$-meßbaren Mengen, die alle Borel-Mengen umfaßt (da $\mathcal{H}^s$ metrisch ist) heißt *s*-**dimensionales Hausdorff-Maß**.

Zur Erinnerung: Sei ν äußeres Maß auf der Menge X. $X \supset E$ heißt ν-meßbar, falls gilt

$$\bigwedge_{A \subset X} \qquad \nu(E) = \nu(E \cap A) + \nu(E \setminus A).$$

Bemerkung 2.3.
1) Für jedes E ist $\mathcal{H}^s(E) \downarrow$ für $0 \le s \uparrow \infty$; denn für $0 < \delta < 1$ ist $|U_i|^s \downarrow 0$ für $s \uparrow \infty$.

2) Falls $s < t$, dann

$$\sum_1^\infty |U_i|^t = \sum_1^\infty |U_i|^{t-s} |U_i|^s \le \delta^{t-s} \sum_1^\infty |U_i|^s, \quad \text{also}$$

$$\sum_1^\infty |U_i|^s \ge \frac{1}{\delta^{t-s}} \sum_1^\infty |U_i|^t.$$

Daraus folgt:

$$0 < \mathcal{H}^{t_0}(E) < \infty \qquad \Rightarrow \qquad \begin{cases} \mathcal{H}^s(E) = \infty & \text{für} \quad s < t_0 \\ \mathcal{H}^s(E) = 0 & \text{für} \quad s > t_0. \end{cases}$$

Es gilt somit

Satz 2.7. *Für jedes $E \subset \mathbf{R}^m$ gibt es 3 Möglichkeiten bzgl. der Menge* $\{\mathcal{H}^s(E) : s \ge 0\}$

1) $\bigwedge_{s \ge 0} \qquad \mathcal{H}^s(E) = 0.$

2) $\bigwedge_{s \ge 0} \qquad \mathcal{H}^s(E) = \infty.$

3) Es gibt genau ein $t_0 \in [0, \infty)$ mit

$$\mathcal{H}^s(E) = \begin{cases} \infty & 0 \le s < t_0 \\ \textit{endlich} > 0 & s = t_0 \\ 0 & s > t_0. \end{cases} \tag{2.9}$$

$t_0 = D_H(E)$ heißt Hausdorff-Dimension von E und es gilt

$$D_H(E) = \inf\big\{s : \mathcal{H}^s(E) = 0\big\} = \sup\big\{s : \mathcal{H}^s(E) = \infty\big\}. \tag{2.10}$$

Satz 2.8. *Sei $A \in \mathcal{H}(\mathbf{R}^m)$. Dann gilt $0 \le D_H(A) \le m$.*

Beweis. Wir wissen bereits

$$D(A) = \inf_{\epsilon > 0} \left\{ \sup_{\tilde\epsilon \le \epsilon} \frac{\log \mathcal{N}(A, \tilde\epsilon)}{\log(1/\tilde\epsilon)} \right\} \le m.$$

Wir setzen die Abkürzung $D := D(A)$. Es genügt zu zeigen: Das D - dim. Hausdorff-Maß von A ist endlich, also $\mathcal{H}^D(A) < \infty$. Jedenfalls gilt

$$D \le \sup_{\tilde\epsilon \le \epsilon} \frac{\log \mathcal{N}(A, \tilde\epsilon)}{\log(1/\tilde\epsilon)} \qquad (\epsilon > 0).$$

Somit existiert zu jedem $\delta > 0$ ein $\epsilon_0(\delta)$, sodaß für alle $\epsilon < \epsilon_0(\delta)$

$$D \leq \frac{\log \mathcal{N}(A,\epsilon)}{\log(1/\epsilon)} + \delta.$$

somit für alle $\epsilon < \epsilon_0(\delta)$

$$\begin{aligned}
\mathcal{H}^D_\epsilon(A) &= \inf\left\{\sum_1^\infty |U_i|^D : \{U_i\} \quad \epsilon - \text{Überdeckung von } A\right\} \\
&\leq \sum_1^{\mathcal{N}(A,\epsilon)} \big|B(x_n,\epsilon)\big|^D = 2^D\epsilon^D\mathcal{N}(A,\epsilon) \\
&\leq 2^m \epsilon^{\frac{\log\mathcal{N}(A,\epsilon)}{\log(1/\epsilon)}+\delta}\mathcal{N}(A,\epsilon) \\
&\leq 2^m e^{-\frac{\log\mathcal{N}(A,\epsilon)}{\log\epsilon}\log\epsilon}\mathcal{N}(A,\epsilon) = 2^m.
\end{aligned}$$

Somit $\mathcal{H}^D(A) \leq 2^m < \infty$. □

Man beachte: Bei gleichem Wert der Hausdorff-Dimension ermöglicht das zugehörige Hausdorff-Maß immer noch eine Unterscheidung der Mengen nach ihrer "Größe".

Es folgen nun einige Vorbereitungen zum Beweis von Satz 2.6

Bemerkung 2.4. Nach Kap.3 Abschnitt 2 Satz 2.6 ist A der Abschluß der Menge der Fixpunkte der Abbildungen $w = w_{j_1} \circ w_{j_2} \circ \cdots \circ w_{j_k}$ für alle endlichen Folgen $\{j_1, j_2, \ldots, j_k\}$ mit $1 \leq j_i \leq N$.

Bemerkung 2.5. Sei D Lösung von $s_1^D + s_2^D + \cdots + s_N^D = 1$. Dann folgt durch k-te Potenzbildung die Identität

$$\sum_{j_1\ldots j_k} (s_{j_1}s_{j_2}\ldots s_{j_k})^D = 1.$$

Dies ist die Summe über alle k-Tupel $\{j_1,\ldots,j_k\}$ $1 \leq j_i \leq N$. In diesem Sinne sind im Folgenden ähnliche Summenzeichen zu verstehen.

Für eine beliebige Menge F schreiben wir

$$F_{j_1\ldots j_k} = w_{j_1} \circ \cdots \circ w_{j_k}(F).$$

A sei, wie immer, der Attraktor des IFS $\{\mathbf{R}^m; w_1, \ldots, w_N\}$.

Lemma 2.9. *Es existiert ein Borel-Maß μ mit Träger $\operatorname{supp}\mu \subset A$ und $\mu(\mathbf{R}^m) = 1$ so, daß für jede meßbare Menge F*

$$\mu(F) = \sum_1^N s_j^D \mu\big(w_j^{-1}(F)\big) \tag{2.11}$$

Beweis. Dies folgt unmittelbar aus Kapitel 5 Satz 1.12 für das Zufalls–IFS $\{X; w_n; p_n\}$ mit $p_1 = s_1^D, \ldots, p_n = s_n^D$. □

Hier ein weiterer Beweis (Nach Falconer [**Fa**] p.120): Wir wählen $x \in A$ und setzen $x_{j_1\ldots j_k} := w_{j_1} \circ \cdots \circ w_{j_k}(x)$ $\quad k = 1, 2, \ldots$. Auf der Menge $C_{00}(\mathbf{R}^m)$ der stetigen

Abbildung $f : \mathbf{R}^m \to \mathbf{C}$ mit kompaktem Träger definieren wir das positive lineare Funktional

$$\phi_k(f) := \sum_{j_1 \dots j_k} (s_{j_1} s_{j_2} \dots s_{j_k})^D f(x_{j_1 \dots j_k})$$

und zeigen, daß $\{\phi_k(f)$ Cauchy-Folge ist: Jedenfalls gilt

$$|A_{j_1 \dots j_k}| \leq S^k |A| \to 0 \qquad (S = \max\{s_j\})$$

für $k \to \infty$. Da f gleichmäßig stetig ist, gibt es zu $\epsilon > 0$ ein p, sodaß für $k \geq p$ die Wertänderung von f auf $A_{j_1 \dots j_k}$ kleiner ist als ϵ. Daher $\sup\{|f(u) - f(v)| : u, v \in A_{j_1 \dots j_k}\} < \epsilon$, unabhängig von der Wahl von $x \in A$. Weiters gilt für $k, k' \geq p$

$$x_{j_1 \dots j_k}, x_{j_1 \dots j_{k'}} \in A_{j_1 \dots j_p}.$$

Sei etwa $k' \geq k$. Aus

$$\sum (s_{j_1} \dots s_{j_{k'}})^D = \sum (s_{j_1} \dots s_{j_k})^D \sum (s_{j_{k+1}} \dots s_{j_{k'}})^D$$

folgt: $|\phi_k(f) - \phi_{k'}(f)|$ ist gleich dem Absolutbetrag von

$$\sum_{j_{k+1} \dots j_{k'}} (s_{j_{k+1}} \dots s_{j_{k'}})^D \sum_{j_1 \dots j_k} (s_{j_1} \dots s_{j_k})^D f(x_{j_1 \dots j_k}) -$$
$$- \sum (s_{j_1} \dots s_{j_{k'}})^D f(x_{j_1 \dots j_{k'}})$$

und somit $\leq \sum_{j_1 \dots j_k} (s_{j_1} \dots s_{j_{k'}})^D \sup_{u,v \in A_{j_1 \dots j_p}} |f(u) - f(v)| < \epsilon$ für alle $k, k' \geq p$. Somit ist $\{\phi_k(f)\}_k$ konvergent für alle $f \in C_{00}(\mathbf{R}^m)$ und der Limes definiert daher ein positives lineares Funktional auf $C_{00}(\mathbf{R}^m)$, das unabhängig vom Ausgangspunkt $x \in A$ ist. Nach dem Darstellungssatz von Riesz existiert ein positives Borel-Maß μ mit

$$\int f d\mu = \lim_{k \to \infty} \phi_k(f) \qquad f \in C_{00}(\mathbf{R}^m).$$

Sei $f \equiv 1 \in C_{00}(\mathbf{R}^m)$ mit kompaktem Träger $\supset A$. Dann gilt für alle k $\phi_k(f) = 1$, also $\int_{\mathbf{R}^m} d\mu = 1 = \mu(\mathbf{R}^m)$. Ist f stetig und $f|A \equiv 0$, dann $\phi_k(f) = 0$ für alle k, somit

$$\int f d\mu = 0, \quad \text{also} \quad supp\, \mu \subset A$$

Weiters gilt für $f \in C_{00}(\mathbf{R}^m)$

$$\begin{aligned}
\phi_k(f) &= \sum_{j_1} s_{j_1}^D \sum_{j_2 \dots j_k} (s_{j_2} \dots s_{j_k})^D f(x_{j_1} \dots x_{j_k}) \\
&= \sum_{j_1} s_{j_1}^D \sum_{j_2 \dots j_k} (s_{j_2} \dots s_{j_k})^D f\big(w_{j_1}(x_{j_2 \dots j_k})\big) \\
&= \sum_{j_1=1}^{N} s_{j_1}^D \phi_{k-1}(f \circ w_{j_1}).
\end{aligned}$$

Für $k \to \infty$ folgt daraus

$$\int f d\mu = \sum_{j=1}^{N} s_j^D \int f \circ w_j d\mu.$$

Nach dem Satz über monotone Konvergenz gilt diese Gleichung dann für alle nicht-negativen meßbaren f und daher für Indikatorfunktionen 1_A meßbarer Mengen A. Also erhalten wir

$$\mu(F) = \sum_{j=1}^{N} s_j^D \mu\big(w_j^{-1}(F)\big).$$

Man beachte dabei $\int 1_A(w(x))dx = \int 1_{w^{-1}(A)}(x)ds$. Damit ist das Lemma bewiesen.

Setze

$$\mu_{j_1\ldots j_k}(F) := \mu\big((w_{j_1} \circ \cdots \circ w_{j_k})^{-1}(F)\big) = \mu\big((w_{j_k}^{-1} \circ \cdots \circ w_{j_1}^{-1})(F)\big).$$

Daraus folgt $supp\, \mu_{j_1\ldots j_k} \subset A_{j_1\ldots j_k}$; denn

$$\begin{aligned}\mu_{j_1\ldots j_k}(F) &= \mu\big((w_{j_1} \circ \cdots \circ w_{j_k})^{-1}(F) \cap A\big) = \\ &= \mu\Big((w_{j_1} \circ \cdots \circ w_{j_k})^{-1}\big(F \cap (w_{j_1} \circ \cdots \circ w_{j_k})(A)\big)\Big).\end{aligned}$$

Bemerkung 2.6. Durch Einsetzen in (2.11) folgt

$$\mu_{j_1\ldots j_k}(F) = \sum s_j^D \mu\big(w_j^{-1}(w_{j_k}^{-1} \circ \cdots \circ w_{j_1}^{-1})(F)\big) = \sum s_j^D \mu_{j_1\ldots j_k j}(F)$$

für alle meßbaren F. Weiters überlegt man sich durch induktive Anwendung von (2.11) für beliebiges k die Gültigkeit von

$$\mu = \sum_{(j_1,\ldots,j_k)} \big(s_{j_1} \ldots s_{j_k}\big)^D \mu_{j_1,\ldots,j_k}. \tag{2.12}$$

Es sei erinnert an Kap.3 Definition 2.3: Ein IFS heißt gerade berührend, falls es die folgende Bedingung (open set condition) erfüllt:
Es existiert eine nichtleere, offene und beschränkte Menge V, für die gilt
$O_1)$ $W(V) = w_1(V) \cup \cdots \cup w_N(V) \subset V$
$O_2)$ $w_i(V) \cap w_j(V) = \emptyset \quad (i \neq j)$
Daraus und mit der eingeführten Schreibweise folgt durch Iteration

$$\bigcup_{j=1}^{N} V_{j_1\ldots j_k j} \subset V_{j_1\ldots j_k}$$

mit disjunkter Vereinigung.

Das Mengensystem $\{V_{j_1\ldots j_k} : k \in \mathbf{N}, 1 \leq j_i \leq N\}$ ist ein **Netz**, d.h. zwei Mengen daraus sind entweder disjunkt, oder eine davon ist in der anderen enthalten. Sei etwa $k \leq k'$, dann gilt genauer: $V_{j_1\ldots j_{k'}} \subset V_{i_1\ldots i_k}$ genau dann, wenn $(i_1, \ldots, i_k)$ das Anfangssegment von $(j_1, \ldots, j_{k'})$ d.h. $(j_1, \ldots, j_{k'}) = (i_1, \ldots, i_k, j_{k+1}, \ldots, j_{k'})$ ist. In Zeichen $(i_1, \ldots, i_k) \prec (j_1, \ldots, j_{k'})$.
Ist andererseits $(i_1, \ldots, i_k) \not\prec (j_1, \ldots, j_{k'})$, dann gilt $V_{j_1\ldots j_{k'}} \cap V_{i_1\ldots i_k} = \emptyset$. In diesem Fall existiert nämlich ein kleinstes $r \quad 1 \leq r \leq k$ mit $i_r \neq j_r$ und somit

$$w_{i_r}(V_{i_1} \ldots i_{r-1}) \cap w_{j_r}(V_{i_1\ldots i_{r-1}}) = \emptyset.$$

sodaß auch $V_{j_1\ldots j_{k'}} \cap V_{i_1\ldots i_k} = \emptyset$.
A sei nun Attraktor eines IFS $\{\mathbf{R}^m, w_i, \ldots, w_N\}$.

Definition 2.4. *A heißt selbstähnlich, falls die w_i Ähnlichkeitstransformationen sind und ein D existiert mit $\mathcal{H}^D(A) > 0$ aber $\mathcal{H}^D\big(w_i(A) \cap w_j(A)\big) = 0$ für $i \neq j$.*

Überlappungen sind also vergleichsweise klein, die Selbst-Ähnlichkeit geht dort daher nicht verloren. Da $A = w_1(A) \cup \cdots \cup w_N(A)$, ist also A „bis auf höchstens kleine Teile“ die disjunkte Vereinigung von verkleinerten Versionen von sich selbst.

BEMERKUNG 2.7. Ein total unzusammenhängendes IFS (mit Attraktor A) erüllt die open set condition: Da die $w_i(A)$ kompakt und disjunkt sind, haben sie positiven Abstand zueinander. Daher existiert $\epsilon > 0$, sodaß die offene Parallel-Menge wegen

$$A_\epsilon := \bigcup_{x \in A} U(x,\epsilon) = \bigcup_{i=1}^{N} \big(\bigcup_{x \in w_i(A)} U(x,\epsilon) \big)$$

ebenfalls disjunkte Vereinigung ist.

SATZ 2.10. *$\{\mathbf{R}^m; w_1 \ldots, w_N\}$ sei IFS mit Ähnlichkeiten w_j und Ähnlichkeitsfaktoren $s_1, \ldots, s_N$ und Attraktor A. Das IFS erfülle die open set condition. Dann hat A positives Hausdorffmaß $0 < \mathcal{H}^D(A) < \infty$, wobei $s_1^D + \cdots + s_N^D = 1$. A ist selbstähnlich, wobei $\mathcal{H}^D(w_i(A) \cap w_j(A)) = 0$ für $i \neq j$.*

BEMERKUNG 2.8. Wenn (O_1) gilt, dann folgt $A \subset \bar{V}$; denn aus der Stetigkeit der w_i folgt

$$\begin{aligned} W(\bar{V}) &= w_1(\bar{V}) \cup \cdots \cup w_N(\bar{V}) \subset \overline{w_1(V)} \cup \cdots \cup \overline{w_N(V)} \\ &\subset \overline{w_1(V) \cup \cdots \cup w_N(V)} \subset \bar{V}. \end{aligned}$$

Und daher

$$\bar{V} \supset W(\bar{V}) \supset W^2(\bar{V}) \supset \cdots \to \bigcap_{k=1}^{\infty} W^k(\bar{V}) = A$$

in der Hausdorff-Metrik. Daher kann kein Punkt von A außerhalb von $\bar{V}$ liegen. Daraus folgt $A_{j_1 \ldots j_k} \subset \bar{V}_{j_1 \ldots j_k}$ wegen

$$(w_{j_1} \circ \cdots \circ w_{j_k}(A) \subset w_{j_1} \circ \cdots \circ w_{j_k}(\bar{V}) \subset \overline{w_{j_1} \circ \cdots \circ w_{j_k}(V)}.$$

LEMMA 2.11. *$\{V_i\}$ sei Familie disjunkter offener Mengen in $\mathbf{R}^m$, sodaß jedes V_i eine Kugel vom Radius $c_1\rho$ enthält und enthalten ist in einer Kugel vom Radius $c_2\rho$. Dann trifft jede Kugel B vom Radius ρ höchstens $(1+2c_2)^m c_1^{-m}$ viele der Mengen $\bar{V}_i$.*

BEWEIS. Bezeichne κ das Volumen der m-dimensionalen Einheitskugel. Sei $\bar{V}_i \cap B \neq \phi$. Dann ist $\bar{V}_i$ enthalten in einer Kugel K, konzentrisch mit B, und mit einem Radius $(1+2c_2)\rho$. Wenn also q der $\{\bar{V}_i\}$ B treffen, dann gilt für die Summe der Volumina der disjunkten (sodaß Maßargument Sinn hat) inneren Kugeln

$$\kappa q \, (c_1\rho)^m \leq \kappa (1+2c_2)^m \rho^m.$$

□

Nun zum Beweis von Satz 2.6 bzw. Satz 2.10: Es ist

$$A = A_1 \cup A_2 \cup \cdots \cup A_N = \bigcup_{j_1 j_2} A_{j_1 j_2} = \cdots = \bigcup_{j_1 \ldots j_k} A_{j_1 j_2 \ldots j_k}.$$

Sei o.B.d.A. $o < s := s_1 \le s_2 \le \cdots \le s_N =: S < 1$. Es gilt für die Ähnlichkeiten $w_1, \ldots, w_N$

$$\sum_{j_1 \ldots j_k} |A_{j_1 \ldots j_k}|^D = \sum_{j_1 \ldots j_k} |A|^D (s_{j_1} s_{j_2} \ldots s_{j_k})^D = |A|^D.$$

Da $|A_{j_1 \ldots j_k}| = |A| s_{j_1} \ldots s_{j_k} \le |A| S^k \to 0$ für $k \to \infty$, ist zu $\delta > 0$ für hinreichend großes $k \quad |A| S^k \le \delta$ und somit $\{A_{j_1 \ldots j_k}\}$ eine δ-Überlagerung von A, daher $\mathcal{H}^D_\delta(A) \le |A|^D$, somit auch $\mathcal{H}^D(A) \le |A|^D < \infty$. Es bleibt also zu zeigen: $0 < \mathcal{H}^D(A)$. Sei dazu obiges V enthalten in einer Kugel vom Radius c_2 und V enthalte eine Kugel von Radius c_1. Sei $\rho > 0$ klein genug. Für jede unendliche Folge $\{j_1, j_2, \ldots\} \quad (1 \le j_i \le N)$ schneiden wir die Folge ab beim **kleinsten** $k \ge 1$, für das

$$s\,\rho < s_{j_1} \cdots s_{j_k} \le \rho \quad (*)$$

Sei $\mathcal{S}$ die Menge der so erhaltenen Folgen. Es gilt: Für keine zwei Folgen $\alpha \ne \beta \in \mathcal{S}$ gilt $\alpha \prec \beta$ oder $\beta \prec \alpha$. Daher ist $\{V_{j_1 \ldots j_k} : \{j_1, \ldots, j_k\} \in \mathcal{S}\}$ eine disjunkte Familie. Jede dieser Mengen enthält eine Kugel vom Radius $c_1 s_{j_1} \cdots s_{j_k}$ und somit eine von Radius $s\rho c_1$ und ist enthalten in einer Kugel vom Radius $c_2 s_{j_1} \cdots s_{j_k}$, also einer von Radius $c_2 \rho$. Nach vorigem Lemma trifft somit jede Kugel B vom Radius ρ höchstens $q = (1 + 2c_2)^m c_1^{-m} s^{-m}$ Mengen der Familie $\left\{\bar{V}_{j_1 \ldots j_k} : \{j_1, \ldots, j_k\} \in \mathcal{S}\right\}$.
Weiters gilt für jedes
$\{j_1, \ldots, j_k\} \quad \mu_{j_1 \ldots j_k}(\mathbf{R}^m) = 1$, und $supp(\mu_{j_1 \ldots j_k}) \subset A_{j_1 \ldots j_k} \subset \bar{V}_{j_1 \ldots j_k}$. Daher Es gilt

$$\mu = \sum_{\{j_1, \ldots, j_k\} \in \mathcal{S}} (s_{j_1} \cdots s_{j_k})^D \mu_{j_1 \ldots j_k} \quad (**)$$

Daher $\mu(B) \le \sum (s_{j_1} \ldots s_{j_k})^D \mu_{j_1 \ldots j_k}(\mathbf{R}^m)$, wobei die Summation über jene Indexfolgen $j_1 \ldots j_k \in \mathcal{S}$ geht, für die $B \cap \bar{V}_{j_1 \ldots j_k} \ne \emptyset$ (da $supp\, \mu_{j_1 \ldots j_k} \subset \bar{V}_{j_1 \ldots j_k}$!). Aus der Formel $(**)$ folgt für hinreichend kleine $\rho = |B|/2$:

$$\mu(B) \le q\,\rho^D = q \left(\frac{|B|}{2}\right)^D.$$

Wir können nun zu $\mathcal{H}^D(A)$ eine δ-Überlagerung $\{U_i\}$ von A finden, für die $\sum |U_i|^D$ beliebig nahe bei $\mathcal{H}^D(A)$ liegt. Andererseits können wir dann aber eine Überlagerung von A mit Kugeln B_i finden, für die $|B_i|/2 \le |U_i|$, sodaß

$$1 = \mu(A) \le \sum \mu(B_i) \le q \sum \left(\frac{|B_i|}{2}\right)^D \le q \sum |U_i|^D.$$

Daraus folgt $\mathcal{H}^D(A) \ge 1/q > 0$. Da die w_i Ähnlichkeiten sind, folgt

$$\sum_1^N \mathcal{H}^D\big(w_j(A)\big) = \sum_1^N s_j^D \mathcal{H}^D(A) = \mathcal{H}^D(A).$$

Wegen $\mathcal{H}^D(A) < \infty$ ist dies nur möglich, falls

$$\mathcal{H}^D\big(w_i(A) \cap w_j(A)\big) = 0 \quad (i \ne j).$$

Damit ist der Satz 2.6 bewiesen.

KAPITEL 6

Markov–Ketten, Stationäre Folgen und Zufalls–IFS

1. Markov–Ketten

Sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeits–Raum (kurz W–Raum), $(X, \mathcal{E})$ ein meßbarer Raum, $A \in \mathcal{F}$ und $\mathcal{G}$ sei eine Unter–σ–Algebra von $\mathcal{F}$. Sei weiters $\xi : \Omega \to \mathbf{R}$ eine relle ZV mit $\xi = \xi^+ - \xi^-$.

Definition 1.1. Die bedingte Erwartung einer nichtnegativen ZV ξ bezüglich $\mathcal{G}$ ist eine nichtnegative ZV $E[\xi|\mathcal{G}] : \Omega \to [0, \infty]$ mit den Eigenschaften

a) $E[\xi|\mathcal{G}]$ ist $\mathcal{G}$–meßbar.

b) Für alle $A \in \mathcal{G}$ gilt

$$\int_A \xi dP = \int_A E[\xi|\mathcal{G}] dP \tag{1.1}$$

Ist ξ eine beliebige ZV, dann definiert man $E[\xi|\mathcal{G}] = E[\xi^+|\mathcal{G}] - E[\xi^-|\mathcal{G}]$, falls $\min(E[\xi^+|\mathcal{G}], E[\xi^-|\mathcal{G}]) < \infty$ P–f.s..

Bemerkung 1.1. Die Existenz der bedingten Erwartung $E[\xi|\mathcal{G}]$ folgt aus dem Satz von Radon–Nykodym: für $\xi \geq 0$ und $A \in \mathcal{G}$ betrachten wir $Q(A) := \int_A \xi dP$. Q ist ein Maß auf $(\Omega, \mathcal{G})$ und Q ist absolut stetig bezüglich P (in Zeichen $Q \prec\prec P$, also $P(A) = 0 \Rightarrow Q(A) = 0$). Daher existiert eine nichtnegative, $\mathcal{G}$–meßbare ZV, bezeichnet als $E[\xi|\mathcal{G}]$, mit der Eigenschaft

$$Q(A) = \int_A \xi dP = \int_A E[\xi|\mathcal{G}] dP.$$

$E[\xi|\mathcal{G}]$ ist nur bis auf eine 0–Menge eindeutig bestimmt, d.h. jede $\mathcal{G}$–meßbare Funktion $f(\omega)$ mit $Q(A) = \int_A f(\omega) P(d\omega)$ $(A \in \mathcal{G})$ ist ebenso geeignet und heißt *Version* von $E[\xi|\mathcal{G}]$.

$E[\xi|\mathcal{G}]$ ist die Radon–Nykodym–Ableitung von Q nach P:

$$E[\xi|\mathcal{G}](\omega) = \frac{dQ}{dP}(\omega).$$

Definition 1.2. Zu $B \in \mathcal{F}$ heißt $E[1_B|\mathcal{G}]$ die bedingte Wahrscheinlichkeit des Ereignisses B bezüglich $\mathcal{G}$, geschrieben $P[B|\mathcal{G}]$, also

a) $P[B|\mathcal{G}]$ ist $\mathcal{G}$–meßbare ZV auf Ω

b) Für alle $A \in \mathcal{G}$ gilt

$$P(A \cap B) = \int_A P\left[B|\mathcal{G}\right](\omega) P(d\omega) \tag{1.2}$$

Von besonderer Bedeutung ist der Fall, wo $\mathcal{G}$ von einer meßbaren Abbildung erzeugt wird.

Definition 1.3. Sei ξ ZV auf Ω und $\eta : \Omega \to X$ ein Zufallselement, sei weiters $\sigma\{\eta\}$ die von η erzeugte σ–Algebra (sie besteht aus allen Mengen der Form $\{\eta \in B\} := \eta^{-1}(B) = \{\omega : \eta(\omega) \in A\}$, $B \in \mathcal{E}$). Dann heißt $E[\xi|\eta] := E[\xi|\sigma(\eta)]$ die bedingte Erwartung von ξ bezüglich η (oder: "unter der Hypothese η").

Wir benötigen noch eine Verfeinerung dieses Begriffes und erinnern uns zuvor an den Integraltransformationssatz:

Satz 1.1. *Sei P_η das Bild–Maß von P auf X unter der Abbildung η, also*

$$P_\eta(B) = P(\{\eta \in B\}) \quad (B \in \mathcal{E}). \tag{1.3}$$

Dann gilt für alle $B \in \mathcal{E}$ und alle $\mathcal{E}$–meßbaren ZV $g : X \to \mathbf{R}$

$$\int_B g(x) P_\eta(dx) = \int_{\eta^{-1}(B)} g(\eta(\omega)) P(d\omega) \tag{1.4}$$

in dem Sinne, daß aus der Existenz des einen Integrals die des anderen und die Gleichheit beider folgt.

Definition 1.4. Zu $A \in \mathcal{F}$ und $x \in X$ heißt jede ZV $g : X \to \mathbf{R}$ mit der Eigenschaft $P(A \cap \{\eta \in B\}) = \int_B g(x) P_\eta(dx)$ die bedingte Wahrscheinlichkeit von A unter der Hypothese $\eta = x$, geschrieben $P[A|\eta = x]$.

Bemerkung 1.2. 1) Wir benützen wieder den Satz von Radon–Nykodym zum Nachweis der Existenz eines solchen g: dazu betrachten wir für $B \in \mathcal{E}$ und $A \in \mathcal{F}$ $Q(B) := P(A \cap \{\eta \in B\})$. Dann ist $0 \leq Q(B) \leq P_\eta(B)$ und somit $Q \prec\prec P_\eta$ und daraus folgt die Existenz von g. Wir bemerken wieder, daß $P[A|\eta = x] = g(x)$ nur P_η–f.s. eindeutig bestimmt ist.

2) Unter Benützung des Satzes 1.1 können wir bei beliebigem $B \in \mathcal{E}$ schreiben:

$$\begin{aligned} P(A \cap \{\eta \in B\} &= \int_B P[A|\eta = x] P_\eta(dx) \\ &= \int_B g(x) P_\eta(dx) = \int_{\{\omega : \eta(\omega) \in B\}} g(\eta(\omega)) P(d\omega). \end{aligned} \tag{1.5}$$

Somit gilt auch P–f.s.

$$g\left(\eta(\omega)\right) = P\left[A|\eta\right](\omega) = P\left[A|\eta = \eta(\omega)\right].$$

Wir können also aus $P\left[A|\eta\right] : \Omega \to \mathbf{R}$ die Funktion $P\left[A|\eta = .\right] : X \to \mathbf{R}$ gewinnen und umgekehrt. Ein für uns wichtiger Spezaialfall ergibt sich für
$A = \{\xi \in D\}$ mit $D \in \mathcal{E}$ und $(\mathcal{F}, \mathcal{E})$–meßbarem $\xi : \Omega \to X$.

Wann können wir mit den *bedingten* Wahrscheinlichkeiten rechnen wie mit *normalen* Wahrscheinlichkeiten? Es gilt (Beweis in [**Pa**])

Satz 1.2. *Sei X ein polnischer Raum (d.h. metrisch, vollständig, separabel), $\mathcal{E} = \mathcal{B} := \mathcal{B}(X)$ und $\xi, \eta : \Omega \to X$ $(\mathcal{F}, \mathcal{B})$–meßbar.*

Dann existiert eine Funktion $p(.\,,.) : X \times \mathcal{B} \to [0,1]$ mit den Eigenschaften

1) Für jedes $A \in \mathcal{B}$ ist $p(.\,,A) : X \to [0,1]$ eine Version von $P[\xi \in A|\eta = .\,]$.

2) Für jedes $x \in X$ ist $p(x,.\,) : \mathcal{B} \to [0,1]$ ein W–Maß auf $(X, \mathcal{B})$.

$p(.\,,.)$ heißt reguläre Version von $P[\xi \in .\,|\eta = .\,]$ und $p(x,A)$ wird als Übergangswahrscheinlichkeit von x in A interpretiert (abhängig von ξ, η).

Lemma 1.3. *Sei $p(.\,,.)$ eine reguläre Version von $P[\xi \in .\,|\eta = .\,]$, $g : X \to \mathbf{R}$ eine ZV mit $E|g \circ \xi| < \infty$. Dann gilt für P–fast alle $\omega \in \Omega$*

$$E\left[g \circ \xi|\eta\right](\omega) = \int_X g(x)p(\eta(\omega), dx). \tag{1.6}$$

Beweis. Die Formel gilt für Indikatorfunktionen 1_B, $B \in \mathcal{B}$; denn $1_B \circ \xi(\omega) = 1_{\{\xi \in B\}}(\omega)$ und

$$\begin{aligned} E\left[g \circ \xi|\eta\right](\omega) &= P\left[\xi \in B|\eta\right](\omega) = P\left[\xi \in B|\eta = \eta(\omega)\right] \\ &= p\left(\eta(\omega), B\right) = \int_X 1_B(x)p(\eta(\omega), dx) \end{aligned}$$

Damit gilt die Aussage auch für Stufenfunktionen und nach dem Satz über dominierte Konvergenz für beliebige g mit der angegebenen Voraussetzung. □

Für einen meßbaren Raum $(X, \mathcal{E})$ bezeichne $b\mathcal{E}$ den Raum der beschränkten, $(\mathcal{E}, \mathcal{B}(\mathbf{R}))$ meßbaren Zufallsvariablen $f : X \to \mathbf{R}$. Sei $(\mathcal{F}_n)$ eine Folge von σ–Algebren mit $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{F}$ und $\xi_n : \Omega \to X$ eine Folge von Zufallselementen, also $(\mathcal{F}, \mathcal{E})$–meßbaren Abbildungen. Falls die ξ_n $(\mathcal{F}_n, \mathcal{E})$–meßbar sind, dann heißt $(\xi_n, \mathcal{F}_n)$ eine *stochastische Folge.* Wir werden ausschließlich den Fall betrachten, wo $\mathcal{F}_n = \sigma\{\xi_0, \xi_1, \ldots, \xi_n\}$ (d.i. die von $\xi_0, \xi_1, \ldots, \xi_n$ erzeugte σ–Algebra).

Definition 1.5. Die stochastische Folge (ξ_n) heißt *Markov–Kette* (bezüglich P), falls

$$\bigwedge_{n \geq 0} \bigwedge_{B \in \mathcal{E}} P\left[\xi_{n+1} \in B|\mathcal{F}_n\right] = P\left[\xi_{n+1} \in B|\xi_n\right] \quad \text{(P–f.s.)}$$

Der Raum X heißt *Phasen*– oder *Zustandsraum* der Markov–Kette, falls alle einpunktigen Mengen $\{x\}$ zu $\mathcal{E}$ gehören (wie z.B. in einem Hausdorff–Raum).

Satz 1.4. *Für eine stochastische Folge $(\xi_n, \mathcal{F}_n)$ sind folgende Aussagen äquivalent.*

1. *Für alle $B \in \mathcal{E}$ und $n \geq 0$ gilt*
$$P\left[\xi_{n+1} \in B|\mathcal{F}_n\right] = P\left[\xi_{n+1} \in B|\xi_n\right] \quad \textit{(P–f.s.)}$$
2. *Für alle $B \in \mathcal{E}$ und $n \geq m \geq 0$ gilt*
$$P\left[\xi_n \in B|\mathcal{F}_m\right] = P\left[\xi_n \in B|\xi_m\right] \quad \textit{(P–f.s.)}$$
3. *Für alle $g \in b\mathcal{E}$ und $n \geq m \geq 0$ gilt*
$$E\left[g(\xi_n)|\mathcal{F}_m\right] = E\left[g(\xi_n)|\xi_m\right] \quad \textit{(P–f.s.)}$$

4. *Für jede beschränkte und $\sigma\{\xi_i : i \geq m\}$–meßbare Zufallsvariable η und alle $m \geq 0$ gilt*
$$E[\eta|\mathcal{F}_m] = E[\eta|\xi_m] \quad (P\text{–f.s.})$$
5. *Für alle $Z \in \sigma\{\xi_i : i \geq m\}$, $V \in \mathcal{F}_l$, $(l \leq m)$ und $m \geq 0$ gilt*
$$P[Z \cap V|\xi_m] = P[Z|\xi_m]\,P[V|\xi_m] \quad (P\text{–f.s.})$$
6. *Für alle $Z \in \sigma\{\xi_i : i \geq m\}$ und $m \geq 0$ gilt*
$$E[1_Z|\mathcal{F}_m] = E[1_Z|\xi_m] \quad (P\text{–f.s.})$$
7. *Für alle $n \geq 0$ und $h \in b\sigma\{\xi_{n+1}\}$ gilt*
$$E[h|\mathcal{F}_n] = E[h|\xi_n] \quad (P\text{–f.s.})$$

Intuitiv bedeutet die Aussage (5): Gegeben die Gegenwart ξ_m, sind bei einem Markov'schen Prozeß die **Z**ukunft Z und die **V**ergangenheit V unabhängig.

Beweis. (1) $\Rightarrow$ (7) : Aus (1) und den Linearitäts–Eigenschaften der bedingten Erwartung folgt (7) zunächst für alle $\sigma(\xi_{n+1})$–meßbaren Stufenfunktionen h und aus den Grenzwert–Eigenschaften der bedingten Erwartung folgt (7) für alle $\sigma(\xi_{n+1})$–meßbaren beschränkten ZufallsVariablen h.

(7) $\Rightarrow$ (6) : Wir setzen jetzt zur Abkürzung $I_i := 1_{\{\xi_i \in A_i\}}$ und verwenden im Folgenden zum Nachweis der Beschränktheit auftretender Funktionen häufig die Monotonie–Eigenschaft der bedingten Erwartung:
$$\xi \leq \eta \Rightarrow E[\xi|\mathcal{G}] \leq E[\eta|\mathcal{G}].$$
Insbesondere nehmen also bedingte Erwartungen von Indikatorfunktionen (=bedingte Wahrscheinlichkeiten) f.s. nur Werte $\in [0,1]$ an.

Wir bemerken, daß es für (7) genügt, bei $m < n$ zu zeigen
$$E[I_m I_{m+1} \dots I_n|\xi_{m-1},\dots,\xi_0] = E[I_m I_{m+1} \dots I_n|\xi_{m-1}].$$

Es ist einerseits
$$\begin{aligned}
&E[I_m I_{m+1}\cdots I_{n-1}I_n|\,\mathcal{F}_{m-1}]\\
&\quad= E[E[I_m I_{m+1}\cdots I_{n-1}I_n|\,\mathcal{F}_{n-1}]|\,\mathcal{F}_{m-1}]\\
&\quad= E[I_m I_{m+1}\cdots I_{n-1}E[I_n|\,\mathcal{F}_{n-1}]|\,\mathcal{F}_{m-1}]\\
&\quad= E[I_m I_{m+1}\cdots I_{n-1}E[I_n|\,\xi_{n-1}]|\,\mathcal{F}_{m-1}]\\
&\quad= E[E[I_m I_{m+1}\cdots I_{n-1}E[I_n|\,\xi_{n-1}]|\,\mathcal{F}_{n-2}]|\,\mathcal{F}_{m-1}]\\
&\quad= E[I_m I_{m+1}\cdots E[I_{n-1}E[I_n|\,\xi_{n-1}]|\,\mathcal{F}_{n-2}]|\,\mathcal{F}_{m-1}]\\
&\quad= E[I_m I_{m+1}\cdots E[I_{n-1}E[I_n|\,\xi_{n-1}]|\,\xi_{n-2}]|\,\mathcal{F}_{m-1}]\\
&\quad\dots\\
&\quad= E[I_m E[I_{m+1}E[\cdots E[I_n|\,\xi_{n-1}]|\,\xi_{n-2}]|\cdots|\,\xi_m]|\,\mathcal{F}_{m-1}]\\
&\quad= E[I_m E[I_{m+1}E[\cdots E[I_n|\,\xi_{n-1}]|\,\xi_{n-2}]|\cdots|\,\xi_m]|\,\xi_{m-1}].
\end{aligned}$$

Damit folgt

$$\begin{aligned}
&E\,[I_m I_{m+1}\cdots I_{n-1}I_n|\,\xi_{m-1}]\\
&\quad= E\,[E\,[I_m I_{m+1}\cdots I_{n-1}I_n|\,\mathcal{F}_{m-1}]|\,\xi_{m-1}]\\
&\quad= E\,[I_m E\,[I_{m+1}E\,[\cdots E\,[E\,[I_n|\,\xi_{n-1}]|\,\xi_{n-2}]|\cdots|\,\xi_m]|\,\xi_{m-1}]|\,\xi_{m-1}]\\
&\quad= E\,[I_m E\,[I_{m+1}E\,[\cdots E\,[I_n|\,\xi_{n-1}]|\,\xi_{n-2}]|\cdots|\,\xi_m]|\,\xi_{m-1}]\,.
\end{aligned}$$

Dies zeigt die Behauptung.

Wegen $\mathcal{F}_0 = \sigma\{\xi_0\}$ erhalten wir für $m = 1$ damit auch die für den folgenden Satz wichtige Beziehung

$$\begin{aligned}
&P(\{\xi_0 \in A_0\} \cap \{\xi_1 \in A_1\} \cap \cdots \cap \{\xi_n \in A_n\})\\
&\quad= E\,\{I_0 E\,[I_1\cdots I_n|\,\xi_0]\}\\
&\quad= E\,\{I_0 E\,[I_1 E\,[\ldots E\,[I_n|\xi_{n-1}]\,|\ldots|\xi_1]\,|\xi_0]\}
\end{aligned} \tag{1.7}$$

(6) $\Rightarrow$ (5) :

$$\begin{aligned}
E\,[1_Z 1_V|\xi_m] &= E\,[E\,[1_Z 1_V|\mathcal{F}_m]\,|\xi_m]\\
&= E\,[1_V E\,[1_Z|\mathcal{F}_m]\,|\xi_m]\\
&= E\,[1_V E\,[1_Z|\xi_m]\,|\xi_m]\\
&= E\,[1_Z|\xi_m]\,E\,[1_V|\xi_m]
\end{aligned}$$

(5) $\Rightarrow$ (4) : Wir zeigen zunächst die Gültigkeit von (4), falls η von der Form 1_Z mit $Z \in \sigma\{\xi_i : i \geq m\}$ ist. Dann gilt für alle $V \in \mathcal{F}_m$

$$\begin{aligned}
\int_V E\,[1_Z|\xi_m]\,dP &= E\,\{1_V E\,[1_Z|\xi_m]\}\\
&= E\,\{E\,[1_V E\,[1_Z|\xi_m]\,|\xi_m]\} = E\,\{E\,[1_Z|\xi_m]\;E\,[1_V|\xi_m]\}\\
&= E\,\{E\,[1_V 1_Z|\xi_m]\} = E\,\{1_V 1_Z\}\\
&= P(V\cap Z) = \int_V E\,[1_Z|\mathcal{F}_m]\,dP
\end{aligned}$$

Daraus folgt $E\,[1_Z|\xi_m] = E\,[1_Z|\mathcal{F}_m]$ (P–f.s.). Die Gültigkeit der Behauptung für beliebige $\eta \in b\sigma\{\xi_i : i \geq m\}$ folgt durch die Grenzwerteigenschaften der bedingten Erwartung.

(4) $\Rightarrow$ (3) : Falls $g \in b\mathcal{E}$, dann ist $g(\xi_n) \in b\sigma\{\xi_i : i \geq m\}$ für jedes $n \geq m$.

(3) $\Rightarrow$ (2) : Für $B \in \mathcal{E}$ ist $1_B \in b\mathcal{E}$ und wegen $1_B(\xi_n(\omega)) = 1 \Leftrightarrow \xi_n(\omega) \in B$ folgt aus (3)

$$\begin{aligned}
P\,[\xi_n \in B|\mathcal{F}_m] &= E\left[1_{\{\xi_n\in B\}}|\mathcal{F}_m\right] = E\,[1_B(\xi_n)|\mathcal{F}_m]\\
&= E\,[1_B(\xi_n)|\xi_m] = P\,[\xi_n \in B|\xi_m]\,.
\end{aligned}$$

(2) $\Rightarrow$ (1) : klar. $\square$

Wir wollen von nun an X als polnischen Raum und $\mathcal{E} = \mathcal{B}(X)$ annehmen. Dann sind natürlich alle 1–punktigen Mengen Borel–Mengen und außerdem existieren reguläre bedingte Verteilungen $p(\xi_n(\omega), B)$, und damit verbunden, die 1–Schritt–Übergangswahrscheinlichkeiten $p_n(x, B) = P[\xi_{n+1} \in B | \xi_n = x]$. Sind die Übergangswahrscheinlichkeiten $p_n(x, B)$ *zeitunbhängig*, d.h. $p_0 = p_1 = \cdots$, dann heißt die Markov–Kette *homogen*. Von besonderer Bedeutung ist die sogenannte Anfangsverteilung $\nu(B) := P[\xi_0 \in B]$.

SATZ 1.5. *Für die endlichdimensionalen Verteilungen der Markov–Kette* $\{\xi_0, \xi_1, \dots\}$ *gilt bei beliebigem* $A \in \mathcal{B}(X^{n+1})$

$$\begin{aligned} &P\left((\xi_0, \dots, \xi_n) \in A\right) \\ &= \int_X \nu(dx_0) \int_X p_0(x_0, dx_1) \cdots \int_X 1_A(x_0, \dots, x_n) p_{n-1}(x_{n-1}, dx_n). \end{aligned} \tag{1.8}$$

Für eine homogene Markov–Kette gilt entsprechend

$$\begin{aligned} &P\left((\xi_0, \dots, \xi_n) \in A\right) = \\ &= \int_X \nu(dx_0) \int_X p(x_0, dx_1) \cdots \int_X 1_A(x_0, \dots, x_n) p(x_{n-1}, dx_n). \end{aligned} \tag{1.9}$$

BEWEIS. Nur für $n = 3$. Sei $(x_0, x_1, x_2) \in X^3$. Sei o.B.d.A. $A = A_0 \times A_1 \times A_2$. Dann gilt nach (1.7)

$$\begin{aligned} &P\left(\xi_0 \in A_0, \xi_1 \in A_1, \xi_2 \in A_2\right) \\ &\quad = E\left\{1_{\{\xi_0 \in A_0\}} E\left[1_{\{\xi_1 \in A_1\}}\, E\left[1_{\{\xi_2 \in A_2\}} \middle|\, \xi_1\right] \middle|\, \xi_0\right]\right\}. \end{aligned}$$

Wegen $1_{\{\xi_i \in A_i\}} = 1_{A_i} \circ \xi_i$ und nach Lemma 1.3 gilt P–f.s.

$$0 \le E\left[1_{\{\xi_2 \in A_2\}} | \xi_1\right](\omega) = \int_X 1_{A_2}(x_2) p_1(\xi_1(\omega), dx_2) =: h(\xi_1(\omega)) \le 1,$$

wobei nach dem Faktorisierungssatz $h : X \to \mathbf{R}$ $\mathcal{E}$–meßbar ist und $0 \le h \le 1$. Nach nochmaliger Anwendung des Lemmas folgt

$$\begin{aligned} &E\left[1_{\{\xi_1 \in A_1\}}(.) \int_X 1_{A_2}(x_2) p_1(\xi_1(.), dx_2) \middle|\, \xi_0\right](\omega) \\ &= E\left[1_{A_1} \circ \xi_1 \,.\, h \circ \xi_1 |\, \xi_0\right](\omega) = E\left[(1_{A_1} \,.\, h) \circ \xi_1 |\, \xi_0\right](\omega) \\ &= \int_X 1_{A_1}(x_1) \int_X 1_{A_2}(x_2) p_1(x_1, dx_2) p_0(\xi_0(\omega), dx_1) \end{aligned}$$

Eine Anwendung der Integraltransformations–Formel liefert schließlich

$$\begin{aligned} &P\left(\xi_0 \in A_0, \xi_1 \in A_1, \xi_2 \in A_2\right) = \\ &\int_\Omega 1_{\{\xi_0 \in A_0\}}(\omega) \left(\int_X 1_{A_1}(x_1) \int_X 1_{A_2}(x_2) p_1(x_1, dx_2) p_0(\xi_0(\omega), dx_1)\right) P(d\omega) \\ &= \int_X \nu(dx_0) \int_X p_0(x_0, dx_1) \int_X 1_{A_0 \times A_1 \times A_2}(x_0, x_1, x_2) p_1(x_1, dx_2). \end{aligned}$$

Die Verallgemeinerung auf beliebige n ist naheliegend. □

Mit dem üblichen Approximationprozeß folgt daraus für alle $\mathcal{B}^{n+1}$–meßbaren, reellwertigen Funktionen $g(x_0, \dots, x_n)$, die entweder konstantes Vorzeichen haben oder beschränkt sind:

$$
\begin{aligned}
Eg(\xi_0, \dots, \xi_n) &= \int_\Omega g \circ (\xi_0, \dots, \xi_n)(\omega) P(d\omega) \\
&= \int_X \nu(dx_0) \int_X p(x_0, dx_1) \cdots \int_X g(x_0, \dots, x_n) p(x_{n-1}, dx_n).
\end{aligned} \tag{1.10}
$$

Wir gehen jetzt den umgekehrten Weg und konstruieren, ausgehend von Übergangswahrscheinlichkeiten und einer Anfangsverteilung eine Markov–Kette. Sei wieder $(X, \mathcal{B})$ unser Standardraum mit der σ–Algebra der Borelmengen.

Satz 1.6. *Sei $p_n : X \times \mathcal{B} \to [0,1]$ eine Folge von Übergangswahrscheinlichkeiten, d.h.*

1) Für alle $A \in \mathcal{B}$ ist $p_n(.\,, A) : X \to [0,1]$ $(\mathcal{B}(X), \mathcal{B}(\mathbf{R}))$–meßbar.

2) Für alle $x \in X$ ist $p_n(x, .\,) : \mathcal{B} \to [0,1]$ ein W–Maß

Sei weiters $\nu : \mathcal{B} \to [0,1]$ ein W–Maß. Dann existiert ein W–Raum $(\Omega, \mathcal{F}, P)$ und eine Markov–Kette (ξ_n) auf $(\Omega, \mathcal{F}, P)$ mit der Anfangsverteilung ν und den Übergangswahrscheinlichkeiten p_n. Sind die p_n alle gleich, so ist die Markov–Kette homogen.

Beweis. Sei $\Omega = \Pi_1^\infty X, \mathcal{F} = \mathcal{B}(X^\infty)$. $\mathcal{F}$ ist die kleinste σ–Algebra, sodaß die Projektionen $\xi_n : \Omega \to X$, definiert durch $\omega = (x_0, x_1, \dots) \mapsto \xi_n(\omega) = x_n$ $(\mathcal{F}, \mathcal{B})$–meßbar sind. Dann sind auch die Projektionen $\pi_n(\omega) := (x_0, \dots, x_n) \in X^{n+1}$ meßbar. Motiviert durch Satz 1.5 definieren wir nun ein Maß $P_\nu := P$ zunächst auf den *endlichdimensionalen Zylinder–Mengen*: zu $A \in \mathcal{B}(X^{n+1})$ sei

$$
\begin{aligned}
&P(\{\omega : (x_0, \dots, x_n) \in A\}) = \\
&\quad = \int_X \nu(dx_0) \int_X p_0(x_0, dx_1) \cdots \int_X 1_A(x_0, \dots, x_n) p_{n-1}(x_{n-1}, dx_n).
\end{aligned}
$$

Wir sehen, daß dies eine bezüglich der Einbettung von A in höhere Dimensionen konsistente Definition von P ist (Wieso?). Nach dem Erweiterungssatz von Kolmogorov ist P damit auf $\mathcal{F}$ eindeutig zu einem W–Maß erweiterbar. Es gilt daher wieder Formel (1.10). Wir betrachten nun die Koordinatenabbildungen $\xi_n : \Omega \to X$, definiert durch $\xi_n(\omega) := x_n$. Klarerweise ist $P\left(\xi_0^{-1}(A_0)\right) = \int_X 1_{A_0}(x_0)\nu(dx_0) = \nu(A_0)$, somit ist ν die Anfangsverteilung. Sei $C \in \mathcal{B}(X^{n+1})$ und $B \in \mathcal{B}(X)$. Es gilt nach Definition der bedingten Wahrscheinlichkeit (vgl. (1.2))

$$
\begin{aligned}
&P(\xi_{n+1} \in B, (\xi_0, \dots, \xi_n) \in C) \\
&= \int_{\{\omega : (\xi_0, \dots, \xi_n) \in C\}} P\left[\xi_{n+1} \in B | \xi_0, \dots, \xi_n\right](\omega) P(d\omega) \\
&= \int_\Omega 1_C \circ (\xi_0, \dots, \xi_n)(\omega) 1_B \circ \xi_{n+1}(\omega) P(d\omega) = \quad \text{nach} \quad (1.10)
\end{aligned}
$$

$$= \int_X \nu(dx_0) \int_X p_0(x_0, dx_1) \cdots$$
$$\cdots \int_X p_{n-1}(x_{n-1}, dx_n) \int_X 1_B(x_{n+1}) 1_C(x_0, \dots, x_n) p_n(x_n, dx_{n+1})$$
$$= \int_X \nu(dx_0) \int_X p_0(x_0, dx_1) \cdots \int_X p_n(x_n, B) 1_C(x_0, \dots, x_n) p_{n-1}(x_{n-1}, dx_n)$$
$$= \int_{\{\omega:(\xi_0,\dots,\xi_n)\in C\}} p_n\left(\xi_n(\omega), B\right) P(d\omega).$$

Somit gilt P–f.s. $P\left[\xi_{n+1} \in B | \xi_0, \dots, \xi_n\right](\omega) = p_n\left(\xi_n(\omega), B\right)$. Andererseits gilt

$$P(\xi_{n+1} \in B, \xi_n \in A_n) = \int_{\{\omega:\xi_n \in A_n\}} P\left[\xi_{n+1} \in B | \xi_n\right](\omega) P d(\omega)$$
$$= \int_\Omega 1_B \circ \xi_{n+1}(\omega) 1_{A_n} \circ \xi_n(\omega) P(d\omega)$$
$$= \int_\Omega 1_B \circ \xi_{n+1}(\omega) 1_{X\times\cdots\times X\times A_n} \circ (\xi_0, \dots, \xi_n)(\omega) P(d\omega) = \quad \text{nach (1.10)}$$
$$= \int_X \nu(dx_0) \int_X p_0(x_0, dx_1) \cdots$$
$$\cdots \int_X p_{n-1}(x_{n-1}, dx_n) \int_X 1_B(x_{n+1}) 1_{X\times\cdots\times X\times A_n)}(x_0, \dots, x_n) p_n(x_n, dx_{n+1})$$
$$= \int_X \nu(dx_0) \int_X p_0(x_0, dx_1) \cdots \int_X p_n(x_n, B) 1_{A_n}(x_n) p_{n-1}(x_{n-1}, dx_n)$$
$$= \int_{\{\omega:\xi_n \in A_n\}} p_n\left(\xi_n(\omega), B\right) P(d\omega).$$

Somit: $P\left[\xi_{n+1} \in B | \xi_n\right](\omega) = p_n\left(\xi_n(\omega), B\right) = P\left[\xi_{n+1} \in B | \xi_0, \dots, \xi_n\right](\omega)$ P–f.s. (ξ_n) ist daher eine Markov–Kette mit der Anfangsverteilung ν. Die Übergangswahrscheinlichkeiten sind p_n. □

Bemerkung 1.3. Ausgehend von einem W–Maß ν auf X und einer Übergangswahrscheinlichkeit $p(x, B)$, erhält man also den W–Raum $(X^\infty, \mathcal{B}(X^\infty))$ und ein W–Maß P_ν darauf und der durch die Koordinatenabbildungen ξ_n darauf definierte stochastische Prozeß ist eine homogene Markov–Kette mit Anfangsverteilung ν. Ist insbesondere ν konzentriert auf einen Punkt x, also $\nu = \delta_x$, dann schreiben wir kurz $P_x = P_{\delta_x}$ und nennen die damit definierte Markov–Kette die *Realisierung des Prozesses mit dem Startpunkt* x. Für beliebiges $B \in \mathcal{B}(X^\infty)$ und $\nu \in \mathcal{P}(X)$ ist die "absolute Verteilung" $P_x\left((\xi_0, \xi_1, \dots) \in B\right)$ eine Version der bedingten Verteilung $P_\nu\left[(\xi_0, \xi_1, \dots) \in B | \xi_0 = x\right]$. Dies bedeutet, daß nach obiger Konstruktion aus $p(x, B)$ die Verteilungen von P_ν, P_x und ν in folgender Beziehung stehen:

$$\begin{aligned} P_\nu\left((\xi_0, \xi_1, \dots) \in B\right) &= \int_X P_\nu\left[(\xi_0, \xi_1, \dots) \in B | \xi_0 = x\right] \nu(dx) \\ &= \int_X P_x\left((\xi_0, \xi_1, \dots) \in B\right) \nu(dx). \end{aligned} \tag{1.11}$$

Mit jeder Übergangswahrscheinlichkeit $p(x, B)$ ist der sogenannte Markov–Operator $M : \mathcal{P}(X) \to \mathcal{P}(X)$ verknüpft:

$$M(\nu)(B) := \int_X p(x, B)\, \nu(dx) \quad (\nu \in \mathcal{P}(X)\,, B \in \mathcal{F}). \tag{1.12}$$

Definition 1.6. Ein Maß $\mu \in \mathcal{P}(X)$ heißt invariantes Maß (rel. p), falls

$$M(\mu) = \mu \tag{1.13}$$

In Bezug zur Markov–Kette nennen wir μ dann *invariante Anfangsverteilung.*

Im allgemeinen ist weder die Existenz noch Eindeutigkeit von invarianten Anfangsverteilungen gesichert. Falls es welche gibt, so ist deren Gesamtheit offensichtlich eine konvexe Teilmenge von $\mathcal{P}(X)$.

Beispiel 1.1. (Übung) Seien $\eta_1, \eta_2, \dots : \Omega \to \mathbf{R}$ unabhängige ZV und sei $A \in \mathcal{B}(\mathbf{R})$. Sei $S_0 = 0, S_{n+1} = S_n + \eta_{n+1}$. Dann gilt

$$P\left[S_{n+1} \in A | S_n, \dots, S_0\right] = P\left[S_{n+1} \in A | S_n\right] \quad \text{P–f.s.}$$

d.h. $\{S_n\}$ ist eine Markov–Kette auf $(\Omega, \mathcal{F}, P)$. Falls die η_i außerdem identisch verteilt sind, dann ist die Markov–Kette homogen.

Beispiel 1.2. Nach J.H. Elton (vgl. [**El**]) betrachten wir eine wesentliche Verallgemeinerung eines Zufalls–IFS. Sei $\{X; w_n; p_n\}$ ein Zufalls–IFS auf dem metrischen Raum X, von dem wir annehmen, daß die beschränkten Mengen relativ kompakt sind. Wir bemerken, daß ein solcher Raum σ–kompakt und daher auch ein polnischer Raum ist. Die Wahrscheinlichkeiten p_i seien ortsabhängig, also Funktionen $p_i : X \to [0, 1]$ mit der Eigenschaft $\sum_{i=1}^N p_i(x) = 1$ für alle $x \in X$. Wir nehmen die p_i als stetig an. Sie seien weiters von 0 weg beschränkt und ihre Stetigkeitsmoduli

$$\phi_i(t) := \sup_{d(x,y) \le t} |p_i(x) - p_i(y)|$$

erfüllen die Dini–Bedingung, d.h. es existiert ein $\alpha > 0$, sodaß $\phi_i(t)/t$ über $(0, \alpha)$ integrierbar ist. Die $w_i : X \to X$ seien Lipschitz–Abbildungen, nicht notwendig kontrahierend, jedoch *kontrahierend im Mittel* zwischen zwei Punkten, d.h. es existiert ein $r < 1$, sodaß für beliebige $x, y \in X$ gilt

$$\prod_{i=1}^N d(w_i x, w_i y)^{p_i(x)} \le r\, d(x, y)$$

Eine stationäre (= zeitunabhängige) Übergangswahrscheinlichkeit sei definiert durch

$$p(x, B) = \sum_{i=1}^N p_i(x) 1_B(w_i(x)) = \sum_{i=1}^N p_i(x) \delta_{w_i(x)}(B). \tag{1.14}$$

Intuitiv heißt dies: ausgehend von x, wähle man ein $i \in \{1, \dots, N\}$ entsprechend der Verteilung $p_i(x)$ und gehe nach $w_i(x)$.

Aus Kap. 5 Satz 1.12 wissen wir:

Satz 1.7. *Sei (X, d) ein kompakter, metrischer Raum und $\{X; w_n; p_n\}$ ein Zufalls–IFS mit kontrahierenden w_i und konstanten p_i. Dann ist das zugehörige invariante Maß μ für die dem Zufalls–IFS zugeordnete Markov–Kette mit der gemäß Formel*

(1.14) definierten Übergangswahrscheinlichkeit eine eindeutig bestimmte invariante Anfangsverteilung.

Dieser Satz gilt aber auch unter den viel allgemeineren Bedingungen des letzten Beispiels. Auch in diesem Fall läßt sich zeigen, daß der Prozeß eine eindeutige, attrahierende Anfangsverteilung μ hat, d.h.: für alle Borel–Mengen $B \in \mathcal{B}(X)$ gilt

$$M(\mu)(B) = \int_X p(x,B)\mu(dx)$$

und für jede Anfangsverteilung ν konvergiert $M^n(\nu)$ in Verteilung (= schwach) gegen μ.

Wir wollen im letzten Beispiel noch das Wahrscheinlichkeitsmaß P_x für die Koordinatendarstellung des Prozesses mit Startpunkt x berechnen. Wir betrachten dazu den Adreß–Raum $\Sigma = N^\infty$ mit der σ–Algebra $\mathcal{A}$, die von den endlichdimensionalen Zylindermengen in Σ erzeugt wird. Zu jedem $x \in X$ sei das W–Maß $\tilde{P}_x$ auf $\mathcal{A}$ für die Zylinder definiert durch die Angabe auf den "Atomen"dieser Zylindermengen

$$\begin{aligned}
&\tilde{P}_x\left((i_1,i_2,\dots,i_n)\right) := \\
&= \tilde{P}_x\left(\{(i_1,i_2,\dots,i_n)\}\times N\times N\times\cdots\right) \\
&= p_{i_1}(x)\,p_{i_2}(w_{i_1}x)\,p_{i_3}(w_{i_2}w_{i_1}x)\,\cdots\,p_{i_n}(w_{i_{n-1}}\cdots w_{i_1}x).
\end{aligned} \tag{1.15}$$

Seien wieder $\xi_n : X^\infty \to X$ die Koordinatenabbildungen (=Projektionen), also

$$X^\infty \ni \omega = (x_0,x_1,\dots) \mapsto \xi_n(\omega) = x_n \in X.$$

Wir wählen $C \in \mathcal{B}^\infty = \mathcal{B}(X^\infty)$ von der Form $C = B_0\times B_1\times B_2\times X\times X\times\cdots$ und erhalten andererseits

$$\begin{aligned}
&P_x\left((\xi_0,\xi_1,\xi_2,\dots)\in C\right) \\
&= \int_X 1_{B_0}(x_0)\delta_x(dx_0)\int_X 1_{B_1}(x_1)p(x_0,dx_1)\int_X 1_{B_2}(x_2)p(x_1,dx_2) \\
&= \int 1_{B_0}(x_0)\delta_x(dx_0)\int 1_{B_1}(x_1)\sum p_{i_1}(x_0)\delta_{w_{i_1}x_0}(dx_1)\left(\sum p_{i_2}(x_1)1_{B_2}(w_{i_2}x_1)\right) \\
&= \int 1_{B_0}(x_0)\delta_x(dx_0)\int\left(\sum p_{i_1}(x_0)1_{B_1}(x_1)\sum p_{i_2}(x_1)1_{B_2}(w_{i_2}x_1)\right)\delta_{w_{i_1}x_0}(dx_1) \\
&= \int_X 1_{B_0}(x_0)\delta_x(dx_0)\left(\sum p_{i_1}(x_0)1_{B_1}(w_{i_1}x_0)\sum p_{i_2}(w_{i_1}x_0)1_{B_2}(w_{i_2}w_{i_1}x_0)\right) \\
&= 1_{B_0}(x)\sum_{i_1=1}^N p_{i_1}(x)1_{B_1}(w_{i_1}x)\sum_{i_2=1}^N p_{i_2}(w_{i_1}(x)1_{B_2}(w_{i_2}w_{i_1}x)\,.
\end{aligned}$$

Die Verallgemeinerung auf höherdimensionale Zylinder ist klar. Wir erhalten daher für alle $C \in \mathcal{B}^\infty$

$$\begin{aligned}
&P\left[(\xi_0,\xi_1,\dots)\in C|\xi_0 = x\right] \\
&= P_x\left((\xi_0,\xi_1,\xi_2,\dots)\in C\right) \\
&= \tilde{P}_x\left((i_1,i_2,\dots) : (x, w_{i_1}x, w_{i_2}w_{i_1}x,\dots)\in C\right)
\end{aligned}$$

Das oben definierte W–Maß $\tilde{P}_x$ auf $\mathcal{A}$ entspricht also gerade dem Maß P_x auf $\mathcal{B}^\infty$, das den bei x startenden Prozeß in X steuert.

Wir werden sehen, daß aus der Eindeutigkeit von μ die *Ergodizität* der zugehörigen Markov–Kette folgt.

2. Stationäre Folgen und Ergodensätze

2.1. Stationäre Folgen. Sei $\xi = (\xi_1, \xi_2, \dots)$ eine stochastische Folge auf dem W–Raum $(\Omega, \mathcal{F}, P)$ mit dem polnischen Zustandsraum $(X, \mathcal{E})$. Oft wird zur Beschreibung des Prozesses die *Koordinatendarstellung* benützt. Wir verwenden dazu X^∞ mit der σ–Algebra $\mathcal{E}^\infty = \mathcal{B}(X^\infty)$ und den Projektionen $\hat{\xi}_i$,

definiert durch

$$X^\infty \ni x = (x_1, x_2, \dots) \mapsto \hat{\xi}_i(x) = x_i \in X.$$

Sei $\hat{\xi} = \left(\hat{\xi}_1, \hat{\xi}_2, \dots\right)$. Wir identifizieren die Folgen ξ und $\hat{\xi}$ mit den Abbildungen $\xi : \Omega \to X^\infty$ und $\hat{\xi} : X^\infty \to X^\infty$, definiert durch

$$\xi(\omega) = (\xi_1(\omega), \xi_2(\omega), \dots)$$
$$\hat{\xi}(x) = x = (x_1, x_2, \dots).$$

Dann ist ξ $(\mathcal{F}, \mathcal{E}^\infty)$–meßbar und $\hat{\xi}$ ist $(\mathcal{E}^\infty, \mathcal{E}^\infty)$–meßbar. Sei $\hat{P} = P_\xi$ das Bildmaß von P auf $\mathcal{E}^\infty$ unter der Abbildung ξ. Es ist klar, daß die Folgen ξ und $\hat{\xi}$ die gleiche Verteilung haben, d.h.

$$\bigwedge_{B \in \mathcal{E}^\infty} P(\xi \in B) = P(\{\omega : (\xi_1(\omega), \xi_2(\omega), \dots) \in B\})$$
$$= \hat{P}(\{x : (x_1, x_2, \dots) \in B\}) = \hat{P}\left(\hat{\xi} \in B\right).$$

Zu $k \in \mathbf{N}$ setzen wir $\sigma^k \xi = (\xi_{k+1}, \xi_{k+2}, \dots)$.

Definition 2.1. Die Folge ξ heißt stationär, falls ξ und $\sigma^k \xi$ für jedes $k \geq 1$ dieselbe Verteilung haben, d.h.

$$\bigwedge_{B \in \mathcal{E}^\infty} P((\xi_1, \xi_2, \dots) \in B) = P((\xi_{k+1}, \xi_{k+2}, \dots) \in B) \tag{2.1}$$

Bemerkung 2.1. Jede unabhängige, identisch verteilte Folge von ZV–en $\xi_i : \Omega \to \mathbf{R}$ ist stationär. Für eine solche Folge gilt Kolmogoroff's *Starkes Gesetz der Großen Zahlen*: Falls $E|\xi_1| < \infty$ und $E\,\xi_1 = m$, dann

$$\lim_n \frac{1}{n}(\xi_1(\omega) + \xi_2(\omega) + \dots + \xi_n(\omega)) = m \quad \text{P–f.s.}$$

Birkhoff's Ergodentheorem, das wir nun ableiten wollen, ist eine weitreichende Verallgemeinerung auf stationäre Folgen.

Bemerkung 2.2. (Übung)

(1) Eine Folge $\xi = (\xi_1, \xi_2, \dots)$ ist genau dann stationär, wenn $(\xi_1, \xi_2, \dots)$ und $(\xi_2, \xi_3, \dots)$ dieselbe Verteilung haben.

(2) Bei einer stationären Folge haben alle ξ_i dieselbe Verteilung.

(3) Aus einer stationären Folge können weitere stationäre Folgen konstruiert werden: Sei $\Phi : X^\infty \to X$ $(\mathcal{E}^\infty, \mathcal{E})$–meßbar, d.h. $\Phi^{-1}(B) \in \mathcal{E}^\infty$ für alle $B \in \mathcal{E}$. Setzen wir $\eta_i = \Phi(\xi_i, \xi_{i+1}, \dots)$. Dann ist auch $\eta = (\eta_i)$ stationär (vgl. Satz 2.16 unten).

BEISPIEL 2.1. Ist die Folge reeller ZV-en $\xi = (\xi_i)$ stationär, dann gilt dies auch für die Folge η der *gleitenden Mittelwerte* ($\mathbf{N} \ni m \geq 1$):

$$\eta = \left(\eta_i := \frac{1}{m} \sum_{k=i}^{i+m-1} \xi_k \right).$$

DEFINITION 2.2. Eine $(\mathcal{F}, \mathcal{F})$–meßbare Abbildung $T : \Omega \to \Omega$ heißt maßerhaltend, falls

$$\bigwedge_{A \in \mathcal{F}} \quad P\left(T^{-1}(A)\right) = P(A).$$

SATZ 2.1. *Sei $T : \Omega \to \Omega$ maßerhaltend. Zu $\xi_1 : \Omega \to X$ definieren wir $\xi_n : \Omega \to X$ durch $\xi_n(\omega) = \xi_1(T^{n-1}(\omega))$ $(n \geq 2)$. Dann ist $\xi = (\xi_1, \xi_2, \dots)$ stationär.*

BEWEIS. Sei $B \in \mathcal{E}^\infty$ und $A = \{\omega : \xi(\omega) \in B\}$, $A_1 = \{\omega : \sigma\,\xi(\omega) \in B\}$. Dann ist

$$A = \{\omega : (\xi_1(\omega), \xi_1(T\omega), \dots) \in B\}\} = \xi^{-1}(B) \text{ und}$$
$$A_1 = \left\{\omega : \left(\xi_1(T(\omega)), \xi_1(T^2(\omega)), \dots\right) \in B\right\}.$$

Damit folgt $T^{-1}(A) = T^{-1}\xi^{-1}(B) = (\xi \circ T)^{-1}(B) = A_1$ und daher

$$\begin{aligned} P\left((\xi_1, \xi_1 T, \dots) \in B\right) &= P(A) = P\left(T^{-1}A\right) \\ &= P(A_1) = P\left(\left(\xi_1 T, \xi_1 T^2, \dots\right) \in B\right) \end{aligned}$$

d.h. ξ ist stationär. □

SATZ 2.2. *Sei $(\xi_i : i \geq 1)$ die Koordinatendarstellung einer stationären Folge. Dann existiert eine maßerhaltende Transformation σ auf $(X^\infty, \mathcal{E}^\infty, P)$, sodaß für alle $\omega \in X^\infty$*

$$\xi_1(\omega) = \xi_1(\omega),\ \xi_2(\omega) = \xi_1(\sigma\omega),\ \dots,\ \xi_n(\omega) = \xi_1(\sigma^{n-1}\omega),\ \dots.$$

BEWEIS. Sei wie oben in Satz 2.16 $\sigma(\omega) = \sigma(x_1, x_2, \dots) = (x_2, x_3, \dots)$ die Verschiebungstransformation. Dann gilt für beliebiges $n \geq 1$ und alle $\omega \in X^\infty$

$$\begin{aligned} \xi_1\left(\sigma^{n-1}(x_1, x_2, \dots)\right) &= \xi_1\left(x_n, x_{n+1}, \dots\right) \\ &= x_n = \xi_n\left(x_1, x_2, \dots\right). \end{aligned}$$

Wir zeigen nun, daß σ maßerhaltend ist. Sei dazu C ein endlichdimensionaler Produktzylinder. Dann ist auch $\sigma^{-1}(C)$ wieder eine solche Zylindermenge. Diese Mengen erzeugen $\mathcal{E}^\infty$ und daher ist σ meßbar. Sei $C \in \mathcal{E}^\infty$.

$$\begin{aligned} P\left(\sigma^{-1}C\right) &= P\left(\omega : \sigma\omega \in C\right) \\ &= P\left((x_1, x_2, \dots) : (x_2, x_3, \dots \in C\right) \\ &= P\left((\xi_2, \xi_3, \dots) \in C\right) \\ &= P\left((\xi_1, \xi_2, \dots) \in C\right) \end{aligned}$$

letzteres wegen der vorausgesetzten Stationarität. Die Behauptung folgt aus

$$P\left((\xi_1, \xi_2, \dots) \in C\right) = P(C).$$

□

SATZ 2.3. *Eine homogene Markov–Kette $(\xi_i : i \geq 0)$ ist genau dann stationär, wenn ξ_0 und ξ_1 dieselbe Verteilung haben.*

Beweis. Die Notwendigkeit folgt unmittelbar aus der Bemerkung 2.2. Nehmen wir umkehrt an, daß ξ_0 und ξ_1 dieselbe Verteilung haben. Um zu zeigen, daß $(\xi_0, \xi_1, \dots)$ und $(\xi_1, \xi_2, \dots)$ dieselbe Verteilung haben, genügt es, dies auf den Zylindermengen der Form $B_0 \times B_1 \times \cdots \times B_n \times X \times X \times \cdots$ nachzuweisen. Nach Satz 1.5 gilt

$$\begin{aligned} & P\left(\xi_0 \in B_0, \xi_1 \in B_1, \dots, \xi_n \in B_n\right) \\ & = \int_X \cdots \int_X 1_{B_0 \times \cdots \times B_n} p(x_{n-1}, dx_n) \cdots p(x_0, dx_1) P_{\xi_0}(dx_0). \end{aligned}$$

Wegen

$$P_{\xi_1}(dx_1) = \int_X p(x_0, dx_1) P_{\xi_0}(dx_0) \tag{2.2}$$

gilt andererseits

$$\begin{aligned} & P\left(\xi_1 \in B_0, \xi_2 \in B_1, \dots, \xi_{n+1} \in B_n\right) \\ & = P\left(\xi_0 \in X, \xi_1 \in B_0, \xi_2 \in B_1, \dots, \xi_{n+1} \in B_n\right) \\ & = \int_X \int_X \cdots \int_X 1_{X \times B_0 \times \cdots \times B_n} p(x_n, dx_{n+1}) p(x_{n-1}, dx_n) \cdots p(x_0, dx_1) P_{\xi_0}(dx_0) \\ & = \int_X \cdots \int_X 1_{B_0 \times \cdots \times B_n} p(x_n, dx_{n+1}) p(x_{n-1}, dx_n) \cdots p(x_1, dx_2) P_{\xi_1}(dx_1). \end{aligned}$$

Wegen $P_{\xi_0} = P_{\xi_1}$ folgt daraus

$$\begin{aligned} & P\left(\xi_1 \in B_0, \xi_2 \in B_1, \dots, \xi_{n+1} \in B_n\right) \\ & = \int_X \cdots \int_X 1_{B_0 \times \cdots \times B_n} p(x_n, dx_{n+1}) p(x_{n-1}, dx_n) \cdots p(x_1, dx_2) P_{\xi_0}(dx_1) \\ & = P\left(\xi_0 \in B_0, \xi_1 \in B_1, \dots, \xi_n \in B_n\right). \end{aligned}$$

□

Aus der Formel (2.2) und der Definition 1.6 dieses Kapitels folgt sofort

Folgerung 2.4. *Sei $(\xi_i : i \geq 0)$ eine homogene Markov–Kette mit Anfangsverteilung μ. Dann gilt: $(\xi_i : i \geq 0)$ ist genau dann stationär, wenn μ inviariant ist.*

Definition 2.3. Ein Ereignis $A \in \mathcal{F}$ heißt invariant, falls $T^{-1}(A) = A$. Das Ereignis A heißt fast–invariant, falls $P\left(A \triangle T^{-1}(A)\right) = 0$. Eine ZV $\xi : \Omega \to \mathbf{R}$ heißt invariant, falls $\xi = \xi \circ T$. Eine ZV $\xi : \Omega \to \mathbf{R}$ heißt fast–invariant, falls $\xi = \xi \circ T$ P–f.s.

Bemerkung 2.3. (1) A ist genau dann invariant, wenn 1_A invariant ist.

(2) A ist genau dann fast–invariant, wenn 1_A fast–invariant ist. Dies folgt sofort aus

$$A \triangle T^{-1}(A) = \{\omega : 1_A(\omega) \neq 1_{T^{-1}A}(\omega)\} = \left\{\omega : \left(1_A - 1_{T^{-1}A}\right)^2(\omega) = 1\right\}.$$

Lemma 2.5. *A ist fast invariant, falls*

$$P\left(A \cap \left(T^{-1}A\right)'\right) = 0 \text{ oder } P\left(A' \cap \left(T^{-1}A\right)\right) = 0.$$

Beweis. Es ist $\left(A \cap (T^{-1}A)'\right) \cup (A \cap T^{-1}A) = A$, also

$$P\left(A \cap \left(T^{-1}A\right)'\right) = P[A] - P\left(A \cap \left(T^{-1}A\right)\right)$$
$$= P\left(T^{-1}A\right) - P\left(A \cap \left(T^{-1}A\right)\right) = P\left(A' \cap T^{-1}A\right).$$

Somit

$$P\left(A \triangle \left(T^{-1}A\right)\right) = P\left(A \cap \left(T^{-1}A\right)'\right) + P\left(A' \cap \left(T^{-1}A\right)\right)$$
$$= 2P\left(A \cap \left(T^{-1}A\right)'\right) = 2P\left(A' \cap \left(T^{-1}A\right)\right).$$

Daraus folgt die Behauptung. □

Satz 2.6. *Sei $T : \Omega \to \Omega$ maßerhaltend. Die Menge $\mathcal{J}_T$ der T–invarianten Ereignisse ist eine σ–Algebra. Auch die Menge $\mathcal{J}^*$ der fast–invarianten Ereignisse ist eine σ–Algebra. $\mathcal{J}^*$ ist die Vervollständigung von $\mathcal{J}_T$ bezüglich P in $\mathcal{F}$, d.h. jede fast–invariante Menge unterscheidet sich von einer invarianten Menge höchstens durch eine Menge vom P–Maß 0.*

Beweis. Übung mit Hinweis für die letzte Behauptung: Zu $A \in \mathcal{J}^*$ betrachte man $B := \limsup T^{-n}A$ und zeige: $B \in \mathcal{J}$ und $P(A \triangle B) = 0$. □

Anstelle von $\mathcal{J}_T$ schreiben wir im Folgenden kurz $\mathcal{J}$.

Lemma 2.7. *Eine Zufallsvariable $\xi : \Omega \to \mathbf{R}$ ist genau dann (fast–) invariant, falls sie $(\mathcal{J}^{(*)}, \mathcal{B}(\mathbf{R}))$–meßbar ist.*

Beweis. Sei ξ invariant und $B \in \mathcal{B}(\mathbf{R})$. Dann ist $\xi^{-1}(B) \in \mathcal{F}$ und weiter

$$T^{-1}\left(\xi^{-1}(B)\right) = (\xi \circ T)^{-1}(B) = \xi^{-1}(B),$$

d.h. $\xi^{-1}(B) \in \mathcal{J}$ und ξ daher $(\mathcal{J}, \mathcal{B}(\mathbf{R}))$–meßbar.

Umgekehrt ist zunächst für jedes $A \in \mathcal{J}$ die Indikatorfunktion 1_A wegen

$$1_A(T\omega) = 1_{T^{-1}A}(\omega) = 1_A(\omega)$$

invariant, damit auch jede $\mathcal{J}$–meßbare Stufenfunktion und nach dem üblichen Approximationsargument jede $\mathcal{J}$–meßbare ZV ξ. □

2.2. Ergodensätze.

Definition 2.4. Eine maßtreue Transformation $T : \Omega \to \Omega$ heißt ergodisch, falls jedes invariante Ereignis entweder Maß 0 oder 1 hat:

$$\bigwedge_{A \in \mathcal{J}} P(A) = 0 \text{ oder } P(A) = 1$$

Satz 2.8. *Sei $T : \Omega \to \Omega$ maßtreu. Dann ist T ergodisch genau dann, wenn*

$$\bigwedge_{A \in \mathcal{J}^*} P(A) = 0 \textit{ oder } P(A) = 1$$

Beweis. Die Bedingung ist hinreichend wegen $\mathcal{J} \subset \mathcal{J}^*$. Die Bedingung ist notwendig: Sei T ergodisch und $A \in \mathcal{J}^*$. Dann ist $A = B \cup N$ mit $B \in \mathcal{J}$ und $P(N) = 0$. Es ist $P(B) = 0$ oder $P(B) = 1$. Daraus folgt $P(A) = 0$ oder $P(A) = 1$. □

Ergodizität ist i.a. schwierig nachzuweisen. Beispiele für ergodische Abbildungen sind die *mischenden.*

DEFINITION 2.5. Sei $T : \Omega \to \Omega$ maßtreu. T heißt mischend, falls

$$\bigwedge_{A,B\in\mathcal{F}} \lim_{n\to\infty} P\left(A \cap T^{-n}\, B\right) = P(A)\, P(B).$$

SATZ 2.9. *Eine mischende Abbildung ist ergodisch.*

BEWEIS. Sei $A \in \mathcal{F}$ und $B \in \mathcal{J}$ und T mischend. Es ist $B = T^{-n}B$ und daher

$$P(A \cap B) = P\left(A \cap T^{-n}B\right) \to P(A)\, P(B).$$

Speziell für $A = B$ folgt $P(B) = P^2(B)$ und daher $P(B) = 0$ oder $P(B) = 1$. □

SATZ 2.10. *Sei $T : \Omega \to \Omega$ maßtreu. Dann sind folgende Aussagen äquivalent:*

(1) T ist ergodisch.

(2) Jede fast–invariante ZV ist P–f.s. konstant.

(3) Jede invariante ZV ist P–f.s. konstant.

BEWEIS. ($1 \Rightarrow 2$): Sei T ergodisch und ξ fast–invariant. Es gilt also P–f.s. $\xi(\omega) = \xi(T\omega)$. Dann ist nach Lemma 2.7 für jedes $c \in \mathbf{R}$ $A_c := \{\omega : \xi(\omega) \leq c\} \in \mathcal{J}^*$ und wegen der Ergodizität $P\left(A_c\right) \in \{0, 1\}$. Sei $C = \sup\left\{c : P\left(A_c\right) = 0\right\}$. Falls $|C| = \infty$, dann gilt P–f.s. $\xi = \pm\infty$ und es ist nichts zu beweisen. Sei also $|C| < \infty$. Dann gilt

$$\left\{\omega : \xi(\omega) < C - \frac{1}{n}\right\} \uparrow \{\omega : \xi(\omega) < C\}$$
$$\left\{\omega : \xi(\omega) > C + \frac{1}{n}\right\} \uparrow \{\omega : \xi(\omega) > C\}.$$

Jede der links stehenden Mengen hat Maß 0 und wir erhalten daher

$$P\left(\{\omega : \xi(\omega) < C\}\right) = \lim P\left(\left\{\omega : \xi(\omega) < C - \frac{1}{n}\right\}\right) = 0$$
$$P\left(\{\omega : \xi(\omega) > C\}\right) = \lim P\left(\left\{\omega : \xi(\omega) > C - \frac{1}{n}\right\}\right) = 0.$$

Daraus folgt $P\left(\{\omega : \xi(\omega) = C\}\right) = 1$.

($2 \Rightarrow 3$): klar, da jede invariante Funktion fast–invariant ist.

($3 \Rightarrow 1$): Sei $A \in \mathcal{J}$. Dann ist 1_A invariant und nach (3) $1_A = 0$ oder $1_A = 1$ P–f.s., d.h. $P(A) \in \{0, 1\}$. T ist also ergodisch. □

BEMERKUNG 2.4. Der Beweis zeigt, daß anstelle beliebiger ZV auch schon die beschränkten ausreichen.

SATZ 2.11 (Maximaler Ergodensatz, Riesz 1945). *Sei $\xi : \Omega \to \mathbf{R}$ eine ZufallsVariable mit $E|\xi| < \infty$ und $T : \Omega \to \Omega$ maßerhaltend. Sei $S_0(\omega) = 0$ und für $k \geq 1$,*

$$S_k(\omega) = \xi(\omega) + \xi(T\omega) + \cdots + \xi\left(T^{k-1}\omega\right).$$

Sei $M_n(\omega) = \max\{S_k(\omega) : 0 \le k \le n\}$ $(n \ge 1)$. *Dann gilt für alle* $n \ge 1$

$$\int_{\{\omega : M_n(\omega) > 0\}} \xi(\omega) P(d\omega) \ge 0.$$

Mit anderen Worten: $P(M_n > 0) > 0 \Rightarrow E[\xi | M_n > 0] \ge 0$.

Beweis. Sei $n \ge 1$ fixiert. Für jedes $0 \le k \le n$ gilt $M_n(T\omega) \ge S_k(T\omega)$ und daher $\xi(\omega) + M_n(T\omega) \ge \xi(\omega) + S_k(T\omega) = S_{k+1}(\omega)$, also für $1 \le k \le n$: $\xi(\omega) \ge S_k(\omega) - M_n(T\omega)$ und $\xi(\omega) \ge \max\{S_1(\omega), \dots, S_n(\omega)\} - M_n(T\omega)$. Daher

$$\begin{aligned} E\left(\xi 1_{\{M_n > 0\}}\right) &= \int_{\{M_n > 0\}} \xi dP \\ &\ge \int_{\{M_n > 0\}} \left(\max S_k(\omega) - M_n(T\omega)\right) P(d\omega). \end{aligned}$$

Auf $\{M_n > 0\}$ ist aber $\max\{S_1, \dots, S_n\} = M_n$. Somit

$$\int_{\{M_n > 0\}} \xi dP \ge \int_{\{M_n > 0\}} \left(M_n(\omega) - M_n(T\omega)\right) P(d\omega).$$

Nun ist $M_n T \ge 0$ P–f.s. und daher

$$\int_{\{M_n > 0\}} M_n dP - \int_{\{M_n > 0\}} M_n T dP \ge \int_\Omega M_n dP - \int_\Omega M_n T dP = 0$$

nach dem Integraltransformationssatz, da T maßtreu und Ω invariant ist. $\square$

Satz 2.12 (Punktweiser Ergodensatz, Birkhoff 1931). *Sei*
$T : \Omega \to \Omega$ *maßerhaltend und* $E|\xi| < \infty$. *Dann gilt*

$$P\left(\left\{\omega : \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \xi\left(T^k(\omega)\right) = E[\xi|\mathcal{J}^*](\omega)\right\}\right) = 1.$$

Ist T *darüber hinaus ergodisch, dann gilt*

$$P\left(\left\{\omega : \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \xi\left(T^k(\omega)\right) = E\xi\right\}\right) = 1.$$

Anders ausgedrückt: die Zeitmittel konvergieren P*–f.s. gegen das Phasenmittel:*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \xi\left(T^k(\omega)\right) = E\xi \quad P\text{–f.s}$$

Beweis. $E[\xi|\mathcal{J}^*]$ ist per def. $\mathcal{J}^*$–meßbar und daher nach Lemma 2.7 fast–invariant. Es folgt für alle $n \ge 1$ und fast alle $\omega \in \Omega$

$$\frac{1}{n} \sum_{k=0}^{n-1} E[\xi|\mathcal{J}^*]\left(T^k(\omega)\right) = E[\xi|\mathcal{J}^*](\omega).$$

Es genügt daher, den Satz zu beweisen für $\xi - E[\xi|\mathcal{J}^*]$, d.h. unter der Annahme $E[\xi|\mathcal{J}^*](\omega) = 0$ f.s. und zu zeigen $\frac{1}{n}\sum_{k=0}^{n-1} \xi \circ T^k(\omega) \to 0$ f.s.

Sei $n \ge 1$, $S_n(\omega) = \sum_{k=0}^{n-1} \xi(T^k\omega)$, $A_\epsilon = \{\omega : \limsup S_n(\omega)/n > \epsilon\}$ für $\epsilon > 0$. Wir zeigen $P(A_\epsilon) = 0$:

Es ist $\limsup S_k(\omega)/k = \limsup S_k(T\omega)/k$ P–f.s. $\limsup S_k/k$ ist also eine fast–invariante ZV und somit ist A_ϵ fast–invariant nach Lemma 2.7.
Für $\xi^*(\omega) = (\xi(\omega) - \epsilon)1_{A_\epsilon}(\omega)$ setzen wir

$$S_n^*(\omega) = \xi^*(\omega) + \cdots + \xi^*\left(T^{n-1}\omega\right), \; M_n^*(\omega) = \max\left\{0, S_1^*(\omega), \ldots, S_n^*(\omega)\right\}.$$

Nach dem Maximalen Ergodensatz gilt für jedes $n \geq 1$

$$\int_{\{M_n^*>0\}} \xi^* dP \geq 0.$$

Für $n \to \infty$ aber gilt

$$\begin{aligned}\{M_n^* > 0\} = \left\{\max_{1\leq k\leq n} S_k^* > 0\right\} \uparrow \\ \uparrow \left\{\sup_{k\geq 1} S_k^* > 0\right\} &= \left\{\sup_{k\geq 1} S_k^*/k > 0\right\} \\ &= \left\{\sup_{k\geq 1} S_k/k > 0\right\} \cap A_\epsilon = A_\epsilon,\end{aligned}$$

letzteres, da $A_\epsilon \subset \left\{\sup_{k\geq 1} S_k/k > \epsilon\right\}$.
Somit $\{M_n^* > 0\} \uparrow A_\epsilon$ und außerdem $E|\xi^*| \leq E|\xi| + \epsilon < \infty$. Nach Lebesgue's Satz über dominierte Konvergenz folgt

$$0 \leq \int_{\{M_n^*>0\}} \xi^* dP \to \int_{A_\epsilon} \xi^* dP$$

und somit auch $\int_{A_\epsilon} \xi^* dP \geq 0$. Nun ist aber

$$\begin{aligned}0 \leq \int_{A_\epsilon} \xi^* dP &= \int_{A_\epsilon} \xi dP - \epsilon P(A_\epsilon) \\ &= \int_{A_\epsilon} E[\xi|\mathcal{J}^*] dP - \epsilon P\left(A_\epsilon\right) = -\epsilon P\left(A_\epsilon\right).\end{aligned}$$

Für $\epsilon > 0$ folgt daher $P\left(A_\epsilon\right) = 0$ und somit $\limsup S_n(\omega)/n \leq 0$ P–f.s. Wenden wir das Gezeigte auf $-\xi$ anstelle von ξ an und beachten $\limsup(-S_n/n) = -\liminf(S_n/n)$, so folgt $-\liminf(S_n/n) \leq 0$, also insgesamt P–f.s.

$$0 \leq \liminf S_n(\omega)/n \leq \limsup S_n(\omega)/n \leq 0.$$

Wenn T ergodisch ist, hat jedes $A \in \mathcal{J}^*$ Maß 0 oder 1 und nach einer bekannten Eigenschaft der bedingten Erwartung ist $E[\xi|\mathcal{J}^*]$ konstant f.s. $= E\xi$. □

Bemerkung 2.5. In der Literatur findet man die erste Aussage des Satzes mit $E[\xi|\mathcal{J}]$ anstelle von $E[\xi|\mathcal{J}^*]$. Für die im Beweis verwendete Funktion $\limsup S_k/k$ wird dabei Invarianz behauptet. Meiner Meinung nach ist jedoch i.a. nur für *beschränkte* ZV ξ richtig. Für die Aussage des Satzes ist dies aber unwichtig, da $E[\xi|\mathcal{J}] = E[\xi|\mathcal{J}^*]$ P–f.s. (Übung).

Folgerung 2.13. *Eine maßtreue Transformation $T : \Omega \to \Omega$ ist genau dann ergodisch, wenn*

$$\bigwedge_{A,B\in\mathcal{F}} \lim_n \frac{1}{n} \sum_{k=0}^{n-1} P\left(A \cap T^{-k}B\right) = P(A)\,P(B). \tag{2.3}$$

Beweis. Die Bedingung ist hinreichend: Wir setzen dazu $A = B$ und nehmen B invariant an. Dann ist $A \cap T^{-k}B = B$ und 2.3 impliziert

$$\frac{1}{n}\sum_{k=0}^{n-1} P\left(A \cap T^{-k}B\right) = P(B) = P(B)^2.$$

Daher $P(B) \in \{0,1\}$ und somit T ergodisch. Die Bedingung ist notwendig: Wir verwenden den Ergodensatz 2.12 für $\xi(\omega) = 1_B(\omega)$. Dann gilt

$$\lim_n \frac{1}{n}\sum_{k=0}^{n-1} 1_{T^{-k}B}(\omega) = P(B) \quad P\text{–f.s.}$$

Andererseits ist $0 \leq \frac{1}{n}\sum_{k=0}^{n-1} 1_{T^{-k}B}(\omega) \leq 1$ und daher nach Lebesgue's Satz über dominierte Konvergenz

$$\int_A \frac{1}{n}\sum_{k=0}^{n-1} 1_{T^{-k}B}(\omega)P(d\omega) = \frac{1}{n}\sum_{k=0}^{n-1} P\left(A \cap T^{-k}B\right) \to$$
$$\to P(B)\int_A dP = P(A)\,P(B).$$

□

Satz 2.14 (L^1–Ergodensatz). *Sei $T : \Omega \to \Omega$ maßtreu, $\xi : \Omega \to \mathbf{R}$ mit $E|\xi| < \infty$. Dann gilt für $n \to \infty$*

$$E\left|\frac{1}{n}\sum_{k=0}^{n-1} \xi \circ T^k - E[\xi|\mathcal{J}^*]\right| \to 0.$$

Ist T ergodisch, dann

$$E\left|\frac{1}{n}\sum_{k=0}^{n-1} \xi \circ T^k - E\xi\right| \to 0.$$

Beweis. Zu jedem $\epsilon > 0$ existiert eine beschränkte ZV η $(\eta(\omega) \leq M)$ mit $E|\xi - \eta| \leq \epsilon$. Dann

$$E\left|\frac{1}{n}\sum_{k=0}^{n-1} \xi \circ T^k - E[\xi|\mathcal{J}^*]\right| \leq E\left|\frac{1}{n}\sum_{k=0}^{n-1} \xi \circ T^k - \eta \circ T^k\right| +$$
$$E\left|\frac{1}{n}\sum_{k=0}^{n-1} \eta \circ T^k - E[\eta|\mathcal{J}^*]\right| + E\left|E[\xi|\mathcal{J}^*] - E[\eta|\mathcal{J}^*]\right|.$$

Der erste und der letzte Summand sind durch ϵ nach oben beschränkt, der mittlere Summand geht gegen 0 nach dem Satz über dominierte Konvergenz und dem Punktweisen Ergodensatz.

Ist T ergodisch, dann besteht $\mathcal{J}^*$ nur aus Ereignissen mit Wahrscheinlichkeit 0 oder 1. Nach dem bekannten Satz über die bedingte Erwartung ist dann
$E[\xi|\mathcal{J}^*] = E\xi \quad P$–f.s.. □

Wir kommen nun zum Ergodensatz für stationäre Folgen. Zur Motivation der folgenden Definition beweisen wir

Lemma 2.15. *Sei $\{\hat{\xi}_i : i \geq 1\}$ die Koordinatendarstellung einer stationären Folge reeller ZV–en und σ die zugehörige (maßerhaltende) Verschiebungstransformation. Ein Ereignis $A \in \mathcal{B}(\mathbf{R}^\infty)$ ist invariant genau dann, wenn ein $C \in \mathcal{B}(\mathbf{R}^\infty)$ existiert mit der Eigenschaft: für alle $k \geq 1$*

$$A = \Big\{(\hat{\xi}_k, \hat{\xi}_{k+1}, \dots) \in C\Big\}.$$

Beweis. Sei $A \in \mathcal{B}(\mathbf{R}^\infty)$ invariant, d.h. $\sigma^{-1}A = A$. Bedingt durch die Koordinatendarstellung und die Definition von σ gilt: $A = \Big\{(\hat{\xi}_1, \hat{\xi}_2, \dots) \in A\Big\}$ und $\sigma^{-1}A = \Big\{(\hat{\xi}_2, \hat{\xi}_3, \dots) \in A\Big\}$.

Aus der Invarianz von A folgt daher durch Iteration

$$\begin{aligned} A &= \Big\{(\hat{\xi}_1, \hat{\xi}_2, \dots) \in A\Big\} \\ &= \Big\{(\hat{\xi}_2, \hat{\xi}_3, \dots) \in A\Big\} \\ &= \Big\{(\hat{\xi}_k, \hat{\xi}_{k+1}, \dots) \in A\Big\}. \end{aligned}$$

Umgekehrt folgt aus $A = \Big\{(\hat{\xi}_k, \hat{\xi}_{k+1}, \dots) \in C\Big\}$ für $k \geq 1$:

$$A = \Big\{(\hat{\xi}_1, \hat{\xi}_2, \dots) \in C\Big\} = \Big\{(\hat{\xi}_2, \hat{\xi}_3, \dots) \in C\Big\} = \sigma^{-1}A.$$

□

Definition 2.6. Sei $\{\xi_i : \Omega \to X : i \geq 1\}$ stationär und $A \in \sigma\{\xi_i : i \geq 1\}$ (das ist die von allen ξ_i erzeugte σ–Algebra), also $A = \{(\xi_1, \xi_2, \dots) \in C\}$ für ein $C \in \mathcal{B}(X^\infty)$. Dann heißt A invariant (bzgl. ξ), falls für alle $k \geq 1$

$$A = \{(\xi_k, \xi_{k+1}, \dots) \in C\}.$$

A heißt fast–invariant, falls es ein invariantes B gibt mit $P(A \triangle B) = 0$. Eine ZV $\eta : \Omega \to \mathbf{R}$ heißt invariant, falls eine meßbare ZV $\Phi : \mathbf{R}^\infty \to \mathbf{R}$ existiert, sodaß für alle $k \geq 1$ gilt $\eta(\omega) = \Phi(\xi_k, \xi_{k+1}, \dots)(\omega)$.

Bemerkung 2.6. Die Invarianz von A bedeutet insbesondere, daß

$$A = \{(\xi_1, \xi_2, \dots) \in C\} = \{(\xi_2, \xi_3, \dots) \in C\}.$$

Andererseits folgt aus dieser Gleichheit wegen der Invarianz der definierenden Eigenschaft unter σ auch die Invarianz von A.

Sei $\mathcal{J}_\xi$ ($\mathcal{J}_\xi^*$) die σ–Algebra der bezüglich einer stationären Folge $\xi = \{\xi_i\}$ (fast–)invarianten Ereignisse.

Definition 2.7. Eine stationäre Folge $\{\xi_i\}$ heißt ergodisch, falls jedes fastinvariante Ereignis die Wahrscheinlichkeit 0 oder 1 hat.

Satz 2.16. *Sei $\xi = \{\xi_i : \Omega \to X\}$ eine stochastiche Folge, $\Phi : X^\infty \to X$ sei $(\mathcal{E}^\infty, \mathcal{E})$–meßbar, d.h. $\Phi^{-1}(B) \in \mathcal{E}^\infty$ für alle $B \in \mathcal{E}$. Setzen wir $\eta_i = \Phi(\xi_i, \xi_{i+1}, \dots)$. Dann gilt: Ist ξ stationär, dann ist auch $\eta = (\eta_i)$ stationär. Ist ξ darüberhinaus ergodisch, dann auch η.*

BEWEIS. Zu Φ betrachten wir Φ_i, definiert durch

$$(x_1, x_2, \dots) = x \mapsto \Phi_i(x) = \Phi(x_i, x_{i+1}, \dots)$$

Dies ist eine $(\mathcal{E}^\infty, \mathcal{E})$–meßbare Funktion.
Zu $B \in \mathcal{E}^\infty$ sei $A = \{x : (\Phi_1(x), \Phi_2(x), \dots) \in B\}$. Dann ist $A \in \mathcal{E}^\infty$ und

$$\{\omega : (\eta_1, \eta_2, \dots) \in B\} = \{\omega : (\xi_1, \xi_2, \dots) \in A\}$$
$$\{\omega : (\eta_2, \eta_3, \dots) \in B\} = \{\omega : (\xi_2, \xi_3, \dots) \in A\}$$

Daraus sieht man: a) ist ξ stationär, dann auch η. b) Jedes η–invariante Ereignis stimmt mit einem ξ–invarianten Ereignis überein. □

Setzen wir speziell $\Phi(x) = \Phi(x_1, x_2, \dots) = f(x_1)$, so erhalten wir

FOLGERUNG 2.17. *Sei $\{\xi_i : \Omega \to X : i \geq 1\}$ stationär und $f \in \mathcal{B}(X)$, d.h. $f : X \to \mathbf{R}$ ist $\mathcal{B}(X)$–meßbar. Dann ist auch die Folge $f \circ \xi_i : \Omega \to \mathbf{R}$ stationär. Ist überdies $\{\xi_i : \Omega \to X : i \geq 1\}$ ergodisch, dann ist auch $\{f \circ \xi_i : \Omega \to \mathbf{R}\}$ ergodisch.*

SATZ 2.18 (Ergodentheorem). *Sei $\xi = (\xi_1, \xi_2, \dots)$ eine stationäre Folge von ZV–en $\xi_i : \Omega \to \mathbf{R}$ mit $E|\xi_1| < \infty$. Dann gilt P–f.s. und im L^1–Mittel*

$$\lim \frac{1}{n} \sum_{k=1}^n \xi_k(\omega) = E\left[\xi_1 | \mathcal{J}_\xi\right](\omega).$$

Wenn zusätzlich ξ ergodisch ist, dann gilt P–f.s. und im L^1–Mittel

$$\lim \frac{1}{n} \sum_{k=1}^n \xi_k(\omega) = E\xi_1.$$

BEWEIS. Ausgehend vom stationären Prozeß $\xi = (\xi_1, \xi_2, \dots)$ betrachten wir auf dem Raum $(\mathbf{R}^\infty, \mathcal{B}^\infty, \hat{P})$ zunächst die zu Beginn dieses Abschnittes beschriebene Koordinatendarstellung $\hat{\xi} = \left(\hat{\xi}_1, \hat{\xi}_2, \dots\right)$ mit der maßerhaltenden Verschiebungstransformation σ. Nach Satz 2.12 und Satz 2.14 folgt die Existenz einer ZV $\hat{\eta} : \mathbf{R}^\infty \to \mathbf{R}$, sodaß

$$\frac{1}{n} \hat{S}_n(\omega) = \frac{1}{n} \sum_{k=1}^n \hat{\xi}_1\left(\sigma^{k-1}(\hat{\omega})\right) \to \hat{\eta}(\hat{\omega})$$

$\hat{P}$–f.s. und L^1. Wegen der Gleichmeßbarkeit von ξ und $\hat{\xi}$ sind auch die Folgen $(S_1, S_2, \dots)$ und $\left(\hat{S}_1, \hat{S}, \dots\right)$ gleichmeßbar und daher existiert auch eine mit $\hat{\eta}$ gleichmeßbare (identisch verteilte) ZV $\eta : \Omega \to \mathbf{R}$, sodaß auch

$$\frac{1}{n} S_n(\omega) = \frac{1}{n} \sum_{k=1}^n \xi_k(\omega) \to \eta(\omega)$$

P–f.s. und L^1. Wir zeigen nun, daß $\eta = E\left[\xi_1 | \, \mathcal{J}_\xi\right]$ P–f.s.

Zunächst folgt aus

$$E\left|\frac{1}{n} \sum_{k=1}^n \xi_k - \eta\right| \to 0$$

für beliebige $A \in \mathcal{F}$

$$\frac{1}{n}\sum_{k=1}^{n}\int_A \xi_k dP \to \int_A \eta dP. \tag{2.4}$$

Falls $A \in \mathcal{J}_\xi$, dann existiert ein $B \in \mathcal{B}(\mathbf{R}^\infty)$, sodaß $A = \{\omega : (\xi_k, \xi_{k+1} \dots) \in B\}$ für alle $k \geq 1$. Aus der Stationarität von ξ und Satz 2.16 folgt, daß auch $\xi_k 1_B((\xi_k, \xi_{k+1} \dots))$ stationär ist und damit

$$\begin{aligned}\int_A \xi_k dP &= \int_{\{(\xi_k,\xi_{k+1}\dots)\in B\}} \xi_k dP \\ &= \int_\Omega \xi_k 1_B((\xi_k, \xi_{k+1} \dots)(\omega))\, P(d\omega) \\ &= \int_\Omega \xi_1 1_B((\xi_1, \xi_2 \dots)(\omega))\, P(d\omega) \\ &= \int_{\{(\xi_1,\xi_2\dots)\in B\}} \xi_1 dP \\ &= \int_A \xi_1 dP.\end{aligned}$$

Aus (2.4) folgt daher nach Definition der bedingten Erwartung $\eta = E[\xi_1|\mathcal{J}_\xi]$. Falls ξ ergodisch ist, dann reduziert sich $\mathcal{J}_\xi$ auf Mengen vom Maß 0 oder 1 und $E[\xi_1|\mathcal{J}_\xi]$ ist f.s konstant gleich $E\xi_1$. □

Die invarianten Ereignisse einer stationären Markov–Kette haben eine einfache Charakterisierung. Zum Beweis benötigt man allerdings Martingaltheorie.

Satz 2.19. *Ein Ereignis $A \in \sigma\{\xi_i : i \geq 0\}$ einer stationären Markov–Kette $\{\xi_i : \Omega \to X\}$ ist ein fast–invariantes Ereignis genau dann, wenn ein $C \in \mathcal{B}(X)$ existiert, sodaß im Sinne der Gleichheit f.ü. für alle $n \geq 0$ gilt*

$$A = \{\omega : \xi_n(\omega) \in C\}.$$

Beweis. vgl. [**St**], pp 196, 197 □

Damit sehen wir die Ergodizität spezieller Markov–Ketten.

Satz 2.20 (Elton, 1987). *Falls μ ein Extremalpunkt in der Menge der stationären Anfangsverteilungen oder falls μ die einzige stationäre Anfangsverteilung ist, dann ist der Markov–Prozeß $\{\xi_i : i \geq 0\}$ mit μ als Verteilung von ξ_0 ergodisch.*

Beweis. Zunächst bemerken wir: $\{\xi_i\}$ ist jedenfalls stationär (Folgerung 2.4 dieses Abschnittes) und die Menge der stationären Anfangsverteilungen ist konvex.

Sei im Widerspruch zur Behauptung $\{\xi_i\}$ nicht ergodisch. Dann existiert ein invariantes Ereignis A mit $0 < P(A) < 1$. Nach dem vorigen Satz existiert ein $C \in \mathcal{B}(X)$, sodaß für alle $n \geq 0$ $A = \{\xi_n \in C\}$. Für $n = 0$ erhalten wir damit insbesondere $P(A) = P_{\xi_0}(C) = \mu(C)$ und somit auch $0 < \mu(C), \mu(C') < 1$. Wir definieren für $B \in \mathcal{B}(X)$

$$\nu(B) = \frac{1}{\mu(C)}\mu(B \cap C) \quad \text{und} \quad \lambda(B) = \frac{1}{\mu(C')}\mu(B \cap C').$$

Dann gilt $\mu = \mu(C) \cdot \nu + \mu(C') \cdot \lambda$, $\nu(\lambda)$ ist absolut stetig bezüglich μ und $d\nu/d\mu = 1_C/\mu(C)$. Da offenbar $\nu \neq \lambda$, ist der Widerspruchsbeweis geführt, sobald wir zeigen, daß ν (und damit auch λ) invariant ist: μ ist dann nämlich entweder nicht eindeutig oder nicht Extremalpunkt in der Menge der Anfangsverteilungen. Nun gilt

$$\begin{aligned}
\nu(B) &= P\left(\{\xi_0 \in B \cap C\}\right)/\mu(C)\\
&= P\left(\{\xi_1 \in B \cap C\}\right)/\mu(C)\\
&= P\left(\{\xi_1 \in B\} \cap \{\xi_1 \in C\}\right)/\mu(C)\\
&= P\left(\{\xi_1 \in B\} \cap \{\xi_0 \in C\}\right)/\mu(C)\\
&= \frac{1}{\mu(C)} \int_C P\left[\xi_1 \in B | \xi_0 = z\right] d\mu(z)\\
&= \frac{1}{\mu(C)} \int_C p(z, B) d\mu(z)\\
&= \int_X p(z, B) d\nu(z).
\end{aligned}$$

Damit ist also ν invariant und daher auch λ. □

2.3. Anwendungen. Wir betrachten das Zufalls–IFS $\{X; w_n; p_n\}$ auf dem kompakten, metrischen Raum (X, d) und die kanonischen Koordinatendarstellung $\xi_i : i \geq 0$ des zugeordneten Markov–Prozesses und dessen Anfangsverteilung das invariante Maß μ sei (vgl. 1.2 und Schluß des vorigen Abschnittes). Der Prozeß ist stationär, da μ stationäre Anfangsverteilung ist. Nach Satz 2.20 ist der Prozeß auch ergodisch, da μ eindeutig ist. Für $f \in C(X)$ ist dann auch $\{f \circ \xi_i\}$ stationär und ergodisch (Folgerung 2.17). Da wir die p_i als konstant annehmen, ist das dort konstruierte Maß $\hat{P}_x =: \hat{P}$ von x unabhängig. Sei

$$B = \left\{ (x_0, x_1, \dots) \in X^\infty : \frac{1}{n+1} \sum_{k=0}^{n} f(x_k) \to \int_X f d\mu \right\}.$$

Nach dem ErgodenTheorem gilt $P\left(\{(\xi_0, \xi_1, \dots) \in B\}\right) = 1$. Nun ist aber

$$\begin{aligned}
P\left(\{(\xi_0, \xi_1, \dots) \in B\}\right) &= \int_X P\left[(\xi_0, \xi_1, \dots) \in B | \xi_0 = x\right] d\mu(x)\\
= \int_X \hat{P}\left(\{(\sigma_1, \sigma_2, \dots) : (x, w_{\sigma_1}x, w_{\sigma_1}w_{\sigma_2}x, \dots) \in B\}\right) d\mu(x).
\end{aligned}$$

Dies ist nur möglich, falls $\hat{P}\left(\{(\sigma_1, \sigma_2, \dots) : (x, w_{\sigma_1}x, w_{\sigma_2}w_{\sigma_1}x, \dots) \in B\}\right) = 1$. Wir erhalten zusammenfassend

Satz 2.21 (Elton, 1987). *Sei (X, d) ein kompakter, metrischer Raum und $\{X; w_n; p_n, n = 1, \dots, N\}$ ein hyperbolisches Zufalls–IFS. Sei $\{x_n\}$ ein bei Anwendung des Zufalls–Iterations–Algorithmus erzeugter Orbit (= Trajektorie), beginnend bei x_0, also für eine Indexfolge $\sigma_1, \sigma_2, \dots$ berechnet nach*

$$x_0, x_1 = w_{\sigma_1}x_0, x_2 = w_{\sigma_2}w_{\sigma_1}x_0, \dots, x_n = w_{\sigma_n} \cdots w_{\sigma_2}w_{\sigma_1}x_0,$$

wobei die die σ_i unabhängig voneinander nach der Verteilung p_i $(1 \leq i \leq N)$ gewählt werden. Sei weiters μ das eindeutig bestimmte invariante Maß des IFS. Sei $\hat{P}$ das vorhin erwähnte W–Maß auf dem Adreßraum $\{1, 2, \dots, N\}^\infty$. Dann gilt für jeden

Startpunkt x_0 *und jede stetige Abbildung* $f : X \to \mathbf{R}$*: für* $\hat{P}$*–fast alle Folgen* $(\sigma_1, \sigma_2, \dots) \in \{1, 2, \dots, N\}^\infty$

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{k=0}^{n} f(x_k) = \int_X f(x) d\mu(x). \tag{2.5}$$

Aus Satz 1.6 Kapitel 5 ergibt sich weiter

FOLGERUNG 2.22. *Sei* B *eine Borel–Menge in* X *mit* $\mu(\partial B) = 0$*, dann existiert zu jedem* $x \in X$ *eine Menge von Folgen* $(\sigma_1, \sigma_2, \dots) \in \{1, 2, \dots, N\}^\infty$ *mit* $\hat{P}$*–Maß 1, sodaß die mittlere Verweilzeit jeder entsprechenden Trajektorie* $\{x_n\}$ *in* B *gegen* $\mu(B)$ *konvergiert. In Zeichen*

$$\lim_{n\to\infty} \frac{|\{i : 0 \le i \le n, x_i \in B\}|}{n+1} = \mu(B).$$

Die letzte Formel läßt sich auch folgendermaßen interpretieren: Für jeden Startpunkt x_0 konvergiert fast jede durch Anwendung des Zufalls–Iterations–Algorithhmus erzeugte Trajektorie $x_0, x_1, x_2, \dots$ in Verteilung gegen μ. Oder nochmal anders: Die empirische Verteilung

$$\nu_n = \frac{1}{n+1} \sum_{k=0}^{n} \delta_{x_k}$$

der ersten $n+1$ Punkte fast jeder, bei x_0 beginnenden, Trajektorie konvergiert schwach gegen das invariante Maß μ.

BEMERKUNG 2.7. In Elton's Originalarbeit wird gezeigt, daß der vorige Satz samt Folgerung auch gültig bleibt, falls X ein lokalkompakter, separabler metrischer Raum ist, die p_i ortsabhängig und die w_i nur Kontraktionen im Mittel sind (vgl. Beispiel 1.2 dieses Kapitels).

BEISPIEL 2.2 (Übung). Für das IFS

$$\left\{[0,1]; \frac{1}{2}x, \ \frac{1}{2}x + \frac{1}{2}; \frac{1}{2}, \ \frac{1}{2}\right\}$$

ist das invariante Maß $d\mu$ das gewöhnliche Lebesgue–Maß dx. Es gilt daher für jedes $f \in C([0,1]$

$$\int_0^1 f(x)dx = \int_{[0,1]} f d\mu.$$

Man vergleiche den exakten Wert, etwa für $f(x) = 1 + x^2$ mit Approximationen aus dem Zufalls–Iterations–Algorithmus, entsprechend Elton's Ergodensatz (*Monte–Carlo–Methode* zur Berechnung bestimmter Integrale).

Als Anwendung des Ergodensatzes folgt noch ein Beweis des Starken Gesetzes der Großen Zahlen. Sei $(\Omega, \mathcal{F}, P)$ ein W–Raum und $\xi_1, \xi_2, \dots$ eine Folge von Zufallsvariablen. Die σ–Algebra

$$\mathcal{X} = \cap_{n=1}^{\infty} \sigma\{\xi_n, \xi_{n+1}, \dots\}$$

heißt Algebra der terminalen Ereignisse.

SATZ 2.23 (Null–Eins–Gesetz von Kolmogoroff). *Sei* $\xi_1, \xi_2, \dots$ *eine Folge von unabhängigen Zufallsvariablen und* A *ein bezüglich dieser Folge terminales Ereignis. Dann gilt* $P(A) \in \{0, 1\}$.

Beweis. Wir zeigen, daß unter den gegebenen Voraussetzungen ein terminales Ereignis A unabhängig von sich selbst ist und daher $P(A \cap A) = P(A) = (P(A))^2$, also $P(A) = 0$ oder 1.

Sei also $A \in \mathcal{X}$. Dann gilt wegen $\mathcal{X} = \cap_{n=1}^{\infty} \sigma\{\xi_n, \xi_{n+1}, \dots\}$ auch

$$A \in \mathcal{F}_1^{\infty} = \sigma\{\xi_1, \xi_2, \dots\} = \sigma\{\cup_{n=1}^{\infty} \sigma\{\xi_1, \xi_2, \dots, \xi_n\}\}.$$

$\cup_{n=1}^{\infty} \sigma\{\xi_1, \xi_2, \dots, \xi_n\}$ ist eine Algebra und somit gilt für die davon erzeugte σ–Algebra $\mathcal{F}_1^{\infty}$: zu jedem $\epsilon > 0$ und $A \in \mathcal{F}_1^{\infty}$ existieren ein $n \in \mathbf{N}$ und $A_n \in \sigma\{\xi_1, \xi_2, \dots, \xi_n\}$ mit $P(A \triangle A_n) \leq \epsilon$.

Daher existiert also eine Folge $A_n \in \sigma\{\xi_1, \xi_2, \dots, \xi_n\}$ mit $P(A \triangle A_n) \to 0$. Daraus folgt dann auch $P(A) - P(A \cap A_n) \to 0$ und $P(A) - P(A_n) \to 0$. Da A, A_n unabhängig sind, folgt damit

$$P(A) = \lim P(A \cap A_n) = P(A)^2$$

und somit $P(A) \in \{0, 1\}$. □

Da jedes invariante Ereignis einer stationären Folge ein terminales Ereignis ist, erhalten wir sofort

Folgerung 2.24. *Sei $\xi = \{\xi_i\}$ eine unabhängige und identisch verteilte reelle stochastische Folge. Dann ist ξ ergodisch.*

Folgerung 2.25 (Starkes Gesetz der Großen Zahl). *Sei $\xi = \{\xi_i : \Omega \to \mathbf{R}\}$ eine unabhängige und identisch verteilte reelle stochastische Folge auf dem W–Raum $(\Omega, \mathcal{F}, P)$ mit $E|\xi_1| < \infty$. Dann gilt P–f.s.*

$$\frac{1}{n} \sum_{i=1}^{n} \xi_i(\omega) \to m (= E\xi_1 = E\xi_i).$$

Zum Abschluß betrachten wir noch Gleichverteilung unter maßerhaltenden Transformationen auf $\mathbf{R}$.

Sei $\Omega = [0, 1[$ mit $\mathcal{F}$ = Borelmengen und P= Lebesgue–Maß. Für festes $\lambda \in \mathbf{R}$ sei $T : \Omega \to \Omega$ definiert durch $T(\omega) = \omega + \lambda \pmod 1$. T ist als Translation maßtreu. Sei $J \subset [0, 1]$ ein Intervall. Nach dem Punktweisen Ergodensatz gilt

$$\frac{1}{n} \sum_{n=1}^{n} 1_J \left(T^{i-1}(\omega)\right) \to E[1_J | \mathcal{J}](\omega) \quad P\text{–f.s.}$$

Die linke Seite gibt den Anteil der Punkte der Folge

$$\omega, \omega + \lambda, \omega + 2\lambda, \dots, \omega + (n-1)\lambda$$

an, die in J fallen. Wann ist diese Folge gleichverteilt, die rechte Seite also gleich $P(J)$, für beliebiges $J \in \mathcal{F}$?

Satz 2.26. *T ist ergodisch und damit*

$$\lim \frac{1}{n} \sum_{n=1}^{n} 1_J \left(T^{i-1}(\omega)\right) = P(J)$$

genau dann, wenn λ irrational ist.

Beweis. Nach Satz 2.10 genügt es zu zeigen, daß eine beschränkte, T–invariante ZV $\xi : \Omega \to \mathbf{R}$ fast sicher konstant ist:

Ein solches ξ ist quadratisch integrierbar, die zugehörige Fourier–Reihe ist daher L_2–konvergent, d.h.

$$\xi(\omega) = \sum_{n \in \mathbf{Z}} \hat{\xi}(n) e^{2\pi i n \omega} \quad \text{mit} \quad \sum_{n \in \mathbf{Z}} |\hat{\xi}(n)|^2 < \infty.$$

Nun ist nach Voraussetzng $\xi(\omega) = \xi(T(\omega))$ und weiters $\widehat{(\xi \circ T)}(n) = \hat{\xi}(n) e^{2\pi i n \lambda}$, also auch $\xi \circ T \in L_2[0,1]$.
Somit gilt für die Fourier–Koeffizienten von $\xi - \xi \circ T$ $(\equiv 0)$:

$$\hat{\xi}(n) \left(e^{2\pi i n \lambda} - 1 \right) = 0.$$

Wegen der Voraussetzung über λ folgt daraus $\hat{\xi}(n) = 0$ für $n \neq 0$ und daher $\xi(\omega) = c_0 = \hat{\xi}(0)$.

Falls andererseits λ rational, etwa $= k/m$ ist, dann ist

$$A := \bigcup_{i=0}^{m-1} \left\{ \omega : \frac{2i}{2m} \leq \omega < \frac{2i+1}{2m} \right\}$$

invariant, aber $P(A) = 1/2$. Also ist T nicht ergodisch. □

finis

KAPITEL 7

Anhang

1. Eigenschaften der Bedingten Erwartung

SATZ 1.1. *Sei* $(\Omega, \mathcal{F}, P)$ *ein W–Raum und* $\mathcal{G}$ *eine Unter–*σ*–Algebra von* $\mathcal{F}$*,* ξ, η *seien ZufallsVariable. Dann gelten folgende Aussagen:*

1. *Falls* C *konstant und* $\xi = C$ P*–f.s, dann* $E[\xi|\mathcal{G}] = C$ P*–f.s.*
2. *Falls* $\xi \leq \eta$ P*–f.s., dann* $E[\xi|\mathcal{G}] \leq E[\eta|\mathcal{G}]$.
3. $|E[\xi|\mathcal{G}]| \leq E[|\xi||\mathcal{G}]$ P*–f.s.*
4. *Falls für Konstante* a, b $aE\xi + bE\eta$ *definiert ist, dann*
$$E[a\xi + b\eta|\mathcal{G}] = aE[\xi|\mathcal{G}] + bE[\eta|\mathcal{G}] \quad Pf.s.$$
5. *Sei* $\mathcal{F}_*$ *eine* σ*–Algebra, die nur aus Mengen vom Maß 0 oder 1 besteht, insbesondere etwa* $\mathcal{F}_* = \{\emptyset, \Omega\}$ *die triviale* σ*–Algebra, dann*
$$E[\xi|\mathcal{F}_*] = E\xi \quad Pf.s.$$
6. *Falls* ξ $\mathcal{G}$*–meßbar ist, dann gilt* $E[\xi|\mathcal{G}] = \xi \quad Pf.s.$
7. $E\{E[\xi|\mathcal{G}]\} = E\xi$.
8. *Falls* $\mathcal{G}_1 \subset \mathcal{G}_2$*, dann*
$$E[E[\xi|\mathcal{G}_2]|\mathcal{G}_1] = E[\xi|\mathcal{G}_1].$$
9. *Falls* $\mathcal{G}_1 \supset \mathcal{G}_2$*, dann*
$$E[E[\xi|\mathcal{G}_2]|\mathcal{G}_1] = E[\xi|\mathcal{G}_2].$$
10. *Sei für xi* $E\xi$ *definiert und* ξ *unabhängig von* $\mathcal{G}$ *(d.h. für alle* $B \in \mathcal{G}$ *sind* $\xi, 1_B$ *unabhängig). Dann*
$$E[\xi|\mathcal{G}] = E\xi \quad Pf.s.$$
11. *Sei* η *ein* $\mathcal{G}$*–meßbare ZV mit* $E|\eta| < \infty$ *und* $E|\xi\eta| < \infty$*, dann*
$$E[\xi\eta|\mathcal{G}] = \eta E[\xi|\mathcal{G}] \quad Pf.s.$$

Sei $\{\xi_n\}_{n\geq 1}$ *eine Folge von erweiterten ZufallsVariablen. Bezüglich der Vertauschbarkeit von Limes und Erwartung gelten folgende Aussagen:*

1. *Falls* $|\xi_n| < \eta, E\eta < \infty$ *und* $\xi_n \to \xi$ *(f.s.), dann*
$$E\left[\xi_n|\,\mathcal{G}\right] \to E\left[\xi|\,\mathcal{G}\right] \quad \text{f.s. und}$$
$$E\left[|\xi_n - \xi||\,\mathcal{G}\right] \to 0 \quad \text{f.s.}$$
2. *Falls* $|\xi_n| \geq \eta, E\eta > -\infty$ *und* $\xi_n \uparrow \xi$ *(f.s.), dann*
$$E\left[\xi_n|\,\mathcal{G}\right] \uparrow E\left[\xi|\,\mathcal{G}\right] \quad \text{f.s.}$$
3. *Falls* $|\xi_n| \leq \eta, E\eta < \infty$ *und* $\xi_n \downarrow \xi$ *(f.s.), dann*
$$E\left[\xi_n|\,\mathcal{G}\right] \downarrow E\left[\xi|\,\mathcal{G}\right] \quad \text{f.s.}$$

4. *Falls* $|\xi_n| \geq \eta, E\eta > -\infty$, *dann*
$$E\left[\liminf \xi_n \middle| \mathcal{G}\right] \leq \liminf E\left[\xi_n \middle| \mathcal{G}\right] \quad f.s.$$
5. *Falls* $|\xi_n| \leq \eta, E\eta < \infty$, *dann*
$$\limsup E\left[\xi_n \middle| \mathcal{G}\right] \leq E\left[\limsup \xi_n \middle| \mathcal{G}\right] \quad f.s.$$
6. *Falls* $\xi_n \geq 0$, *dann*
$$E\left[\sum \xi_n \middle| \mathcal{G}\right] = \sum \left[\xi_n \middle| \mathcal{G}\right] \quad f.s.$$

BEWEIS. Siehe [**Sh**], Seite 214–217. □

SATZ 1.2 (Faktorisierungssatz). *Seien* $(\Omega, \mathcal{F})$ *und* $(X, \mathcal{E})$ *zwei Meßräume und* $\xi : \Omega \to X$ *meßbar. Sei* $\eta : \Omega \to \mathbf{R}$ *eine* $(\sigma(\xi), \mathcal{B}(\mathbf{R}))$*–meßbare ZV. Dann existiert eine* $(\mathcal{E}, \mathcal{B}(\mathbf{R}))$*–meßbare ZV* $\varphi : X \to \mathbf{R}$ *mit* $\varphi \circ \xi = \eta$*. Falls* η *beschränkt ist, dann kann auch* φ *beschränkt gewählt werden.*

BEWEIS. Sei Φ_ξ die Menge der $\sigma(\xi)$–meßbaren ZV–en η und $\tilde{\Phi}_\xi$ die Menge der $\sigma(\xi)$–meßbaren ZV–en η, die sich in der Form $\varphi \circ \xi$ mit $(\mathcal{E}, \mathcal{B})$–meßbarem φ darstellen lassen. Klarerweise gilt $\tilde{\Phi}_\xi \subset \Phi_\xi$.

Sei nun $A \in \sigma(\xi)$, also $A = \xi^{-1}(E)$ (mit $E \in \mathcal{E}$) und $\eta = 1_A(\omega) \in \Phi_\xi$. Sei $\varphi(x) = 1_E(x)$. Dann ist φ eine $(\mathcal{E}, \mathcal{B})$–meßbare Abbildung $X \to \mathbf{R}$ und $\varphi \circ \xi(\omega) = \eta(\omega)$, d.h. $1_E(\xi(\omega)) = 1_A(\omega)$, somit $1_A \in \tilde{\Phi}_\xi$. Daher ist auch jede einfache $\sigma(\xi)$–meßbare Funktion Element von $\tilde{\Phi}_\xi$. Sei $\eta : \Omega \to \mathbf{R} \in \Phi_\xi$. Dann existiert eine Folge von einfachen $\sigma(\xi)$–meßbaren Funktionen $\eta_n : \Omega \to \mathbf{R}$ mit $\eta_n(\omega) \to \eta(\omega)$ auf Ω. Dazu existieren dann $\varphi_n : X \to \mathbf{R}$ mit $\varphi_n(\xi(\omega)) \to \eta(\omega)$. Sei $B = \{x \in X : \lim \varphi_n(x) \text{ ex.}\}$. Offenbar $B \supset \xi(\Omega)$. Setzen wir $\varphi(x) = \lim \varphi_n(x)$ für $x \in B$ und $\varphi(x) = 0$ für $x \notin B$, also $\varphi(x) = 1_B(x) \lim \varphi_n(x)$. Nun ist B eine Borelmenge in X; denn
$$B = \{\limsup \varphi_n(x) = \liminf \varphi_n(x)\}\,.$$
$\limsup \varphi_n$ und $\liminf \varphi_n$ sind ZV und daher $B \in \mathcal{E}$. (vgl. [**HS**], S. 154, 11.12 und (11.9))

Somit ist φ eine $(\mathcal{E}, \mathcal{B})$–meßbare Funktion und für alle $\omega \in \Omega$ gilt $\xi(\omega) \in B$ und $\varphi(\xi(\omega)) = \eta(\omega)$, also $\eta \in \tilde{\Phi}_\xi$, d.h. $\tilde{\Phi}_\xi = \Phi_\xi$.

Sei $0 \leq \eta \leq 1$. Dann betrachte $\psi(x) := 1 \wedge \varphi(x)$. Dies ist ebenfalls Borelfunktion und es gilt $0 \leq \varphi \leq 1$ und $\psi(\xi(\omega)) = 1 \wedge \varphi(\xi(\omega)) = 1 \wedge \eta(\omega) = \eta(\omega)$. □

Literaturverzeichnis

[Ba] M. F. Barnsley *Fractals Everywhere* Academic Press, Boston, London, 1993

[Bl] W. Blaschke *Kreis und Kugel* Verlag von Veit & Comp., Leipzig, 1916

[Ed] G. A. Edgar *Measure, Topology, and Fractal Geometry*, Springer, Undergraduate Texts in Mathematics, Berlin, Heidelberg, New York, 1990

[El] J.H. Elton *An ergodic theorem for iterated maps*, Ergod. Th. & Dynam. Sys. Vol 7, 1987 pp 481–488

[Fa] K. F. Falconer *The Geometry of Fractal Sets*, Cambridge Tracts in Mathematics, Cambridge University Press, Cambridge, 1987

[Hu] J.E. Hutchinson *Fractals and Self Similarity*, Indiana Univ. Math. Jour. Vol 30, No. 5, 1981 pp 713–747

[HS] E. Hewitt, K. Stromberg *Real and Abstract Analysis*, Springer, N.Y., Berlin, Heidelberg, 1969

[Pa] K.R. Parthasarathy *Probability Measures on Metric Spaces*, Academic Press, N.Y., London, 1967

[Sh] A.N. Shiryayev *Probability*, Graduate Texts in Mathematics, Springer, N.Y., Berlin, Heidelberg, 1984

[St] W.F. Stout *Almost Sure Convergence*, Academic Press, N.Y., London, 1974

[Th] D'Arcy Thompson *On Growth and Form*, An abridged edition, paperback, Cambridge University Press, Cambridge, 1971