Einführung in die Ergodentheorie - Notizen [1]

Jena - Sommersemester 2014

Daniel Lenz

[1] Es handelt sich lediglich um Notizen und nicht um ein Skriptum zur Vorlesung. Konstruktive Kommentare sind willkommen.

Inhaltsverzeichnis

KAPITEL 0

Einführung

Wir betrachten ein physikalisches System bestehend aus N (mit N groß) Massepunkten (Teilchen) in einer Menge $Q \subset \mathbb{R}^3$. Der Zustand des Systems zur Zeit $t \in \mathbb{R}$ ist dann beschrieben durch die

- Positionen $q^{(1)}(t), \ldots, q^{(N)}(t)$ der Massepunkte,
- Impulse $p^{(1)}(t), \ldots, p^{(N)}(t)$ der Massepunkte (Impuls = Masse mal Geschwindigkeit).

Damit wird das System dann zu jedem Zeiptunkt t durch einen Vektor

$$\omega(t) = (q^{(1)}, \ldots, q^{(N)}, p^{(1)}, \ldots, p^{(N)} \in Q^N \times (\mathbb{R}^3)^N =: \Omega \subset \mathbb{R}^{6N}$$

beschrieben. Es heißt Ω der Phasenraum des Systems. Die physikalischen Größen (etwa Ort, Geschwindigkeit, Energie...) werden dann durch Funktionen

$$f : \Omega \longrightarrow \mathbb{R}$$

auf Ω beschrieben.

Grundlegendes Ziel: Beschreibe Verhalten von $t \mapsto \omega(t)$ bzw. das Verhalten von $t \mapsto f(T_t\omega)$ fuer $f : \Omega \longrightarrow \mathbb{R}$.

Es ist die Gesamtenergie des Systems gegeben durch die Hamiltonfunktion

$$H : \Omega \longrightarrow \mathbb{R}.$$

(Ueblicherweise gilt

$$H(\omega) = H_{kin}(\omega) + H_{pot}(\omega) = \frac{1}{2}\sum_{j=1}^{N} \frac{\|p^{(j)}\|^2}{m_j} + H_{pot}(\omega)$$

mit einer geeigneten Funktion H_{pot}.) Die Zeitentwicklung des Systems ist gegeben durch die Hamiltongleichungen

$$\frac{dp_i^{(j)}}{dt} = \frac{\partial H}{\partial q_i^{(j)}}(\omega(t)), \quad \frac{dq^{(j)}{}_i}{dt} = \frac{\partial H}{\partial p_i^{(j)}}(\omega(t)),$$

$j = 1, \ldots, N$, $i = 1, 2, 3$. Es handelt sich um eine gewöhnliche Differentialgleichung

$$(HE) \quad \omega'(t) = \widetilde{H}(\omega(t)).$$

Nimmt man Q als beschraenkt an und H als stetig differenzierbar, so existieren dann also nach der bekannten Lösungstheorie fuer gewöhnliche Differentialgleichungen zu jedem Anfangwert $\omega_0 \in \Omega$ eine eindeutige Funktion

$$\omega : \mathbb{R} \longrightarrow \Omega$$

mit

$$\omega'(t) = \widetilde{H}(\omega(t)) \text{ und } \omega(0) = \omega_0.$$

Diese Loesung heißt auch der Orbit (die Bahn) des Systems zum Anfangswert ω_0. Damit erhaelt man dann zu jedem $t \in \mathbb{R}$ eine eindeutige Abbildung

$$T_t : \Omega \longrightarrow \Omega,$$

sodaß

$$\omega(t) := T_t \omega_0$$

gerade die Loesung von (HE) zum Anfangswert ω_0 fuer $t = 0$ liefert. Damit besteht also das grundlegende Ziel in der Beschreibung des Verhalten der Abbildungen $t \mapsto T_t$ bzw. das Verhalten von $t \mapsto f(T_t\omega)$ fuer $f : \Omega \longrightarrow \mathbb{R}$.

Mit dem Eindeutigkeitssatz fuer Loesungen sieht man leicht, daß folgendes gilt:

- $T_0 = \mathrm{id}$,
- $T_{t+s} = T_t \circ T_s$.

Die Differentialgleichung liefert also eine Gruppenwirkung (Operation) der Gruppe $\mathbb{R}$ auf Ω. Die Gruppenwirkung hat einige bemerkenswerte Eigenschaften:

Betrachtet man nun fuer festes $\omega \in \Omega$ die Abbildung

$$\mathbb{R} \longrightarrow \mathbb{R},\ t \mapsto H(T_t\omega),$$

so liefert eine kleine Rechnung

$$\frac{dH}{dt} = 0.$$

Es ist also die Energie

$$\mathbb{R} \longrightarrow \mathbb{R},\ t \mapsto H(T_t\omega),$$

auf der Bahn des Systems konstant. Definiert man fuer $E \in \mathbb{R}$ die Energieschale (Zeichung)

$$X := \Omega_E := \{\omega \in \Omega : H(\omega) = E\},$$

so ist X also invariant unter der Gruppenwirkung T_t. In gaengigen Faellen ist darueberhinaus auch X kompakt.

Grundlegendes Ziel - Neuformulierung: Beschreibe Verhalten der Gruppenoperation $t \mapsto T_t$ auf der Energieschale X bzw. das Verhalten von $t \mapsto f(T_t\omega)$ fuer $f : X \longrightarrow \mathbb{R}$.

Bei der weiteren Untersuchung gibt es folgende grundsaetzliche **Probleme:**

- Es ist im allgemeinen nicht moeglich die T_t explizit auszurechnen (fuer grosse N z.B. $N \geq 3$, tatsaechlich $N = 10^{23}$).
- Die Anfangsbedingungen sind nicht genau bekannt.

Als Loesung bietet sich an, sich auf 'statistische' Betrachtungen zu beschraenken, d.h. sich um gemittelte Groessen zu bemuehen. Dabei bieten sich zwei Arten von Mittelungen an:

Zeitmittel. Man ersetzt $f(T_t\omega)$ durch $\frac{1}{S}\int_0^S f(T_t\omega)dt$ fuer ein großes S. Zur weiteren Vereinfachung kann man dann sogar

$$\lim_{S\to\infty} \frac{1}{S}\int_0^S f(T_t\omega)dt$$

betrachten (falls dieser Grenzwert existiert).

Raummittel: Nach dem Liouville Theorem erhaelt die Abbildung $T_t : \Omega \longrightarrow \Omega$ das Standardvolumen im $\Omega \subset \mathbb{R}^{6N}$. Nach einem Argument von Khinchin '49 gibt es weiterhin auch ein invariantes Maß λ auf $X := \Omega_E$. Damit definiert man dann das Raummittel als

$$\frac{1}{\lambda(X)} \int_X f(\omega) d\lambda(\omega).$$

Entsprechende Betrachungen gehen (fuer Behandlung von Gasen) auf Boltzmann 1887 zurueck. Boltzmann machte (fuer das von ihm untersuchte Beispiel) die Annahme, 'Ergodenhypthese',

$$\{T_t\omega : t \in \mathbb{R}\} = X$$

fuer ein $\omega \in X$. Er 'folgerte' daraus die Gleichheit von Raum- und Zeitmittel, d.h.

$$\lim_{S\to\infty} \frac{1}{S} \int_0^S f(T_t\omega) dt = \frac{1}{\lambda(X)} \int_X f(\omega) d\lambda(\omega).$$

Heute weiss mann, dass die getroffene Annahme im allgemeinen nicht gilt. Stattdessen haben die Betrachtungen von Boltzmann haben ein ganzes Gebiet angestossen:

- Voraussetzungen und Konsequenzen der Gleichheit von Raum- und Zeitmittel. (Ergodensaetze)
- Untersuchung der 'Lage' von $\{T_t\omega : t \in \mathbb{R}\}$ in X. (Rekurrenz)

Darum geht es in dieser Vorlesung.

Bemerkung. Fuer Boltzmann's urspruengliches Beispiel ist es immer noch nicht bekannt, ob die Gleichheit von Raum- und Zeitmittel gilt.

KAPITEL 1

Der Rahmen: Grundlegende Konzepte und Beispiele

In diesem Abschnitt fuehren wir die grundlegenden Konzepte ein und stellen einige Beispiele vor.

1. Grundlegende Konzepte

Erinnerung. Sei X eine nichtleere Menge. Eine Familie $\mathcal{B}$ von Teilmengen von X heißt σ-Algebra, wenn folgendes gilt:

- $\emptyset \in \mathcal{B}$
- $A \in \mathcal{B} \longrightarrow A^c \in \mathcal{B}$
- $A_n \in \mathcal{B}, n \in \mathbb{N} \longrightarrow \bigcup A_n \in \mathcal{B}$.

Eine Abbildung $m : \mathcal{B} \longrightarrow [0, \infty]$ heißt Maß, wenn folgendes gilt:

- $m(\emptyset) = 0$,
- $m(\bigcup_n A_n) = \sum_n (A_n)$ fuer A_n paarweise disjunkte Element aus $\mathcal{B}$.

Ein Tripel $(X, \mathcal{B}, m)$ bestehend aus einer nichtleeren Menge X, einer σ-Algebra $\mathcal{B}$ und einem Maß m heißt meßbarer Raum.

Definition (Messbare Transformationen). *Seien $(X_j, \mathcal{B}_j, m_j)$, $j = 1, 2$ Maßraüme und $T : X_1 \longrightarrow X_2$ gegeben.*
(a) Es heißt T eine meßbare Transformation, wenn $T^{-1}B$ zu $\mathcal{B}_1$ gehoert fuer jedes $B \in \mathcal{B}_2$.
(b) Es heißt T eine maßerhaltende Transformation, wenn T meßar ist und

$$m_1(T^{-1}B) = m_2(B)$$

fuer alle $B \in \mathcal{B}_2$ gilt.
(c) Es heißt T eine invertierbare maßerhaltende Transformation, wenn T bijektiv ist und sowohl T als auch seine Inverse maßerhaltende Transformationen sind.

Im folgenden werden wir meist den Fall $(X_1, \mathcal{B}_1, m_1) = (X_2, \mathcal{B}_2, m_2)$ betrachten. Dann schreiben wir auch nur $(X, \mathcal{B}, m)$. In diesem Fall induziert die maßerhaltende Abbildung $T : X \longrightarrow X$ fuer jedes $n \in \mathbb{N}$ eine maßerhaltende Abbildung

$$T_n : X \longrightarrow X, T_n x = T^n x.$$

Setzt man noch $T_0 := id$, so gilt dann offenbar

- $T_0 = id$,
- $T_{n+m} = T_n \circ T_m$ fuer alle $n, m \in \mathbb{N}_0$.

Ist darueberhinaus T noch invertierbar, so kann man fuer $n \in \mathbb{N}$ definieren

$$T_{-n} := (T^{-1})^n.$$

Dann gelten

- $T_0 = id$,
- $T_{n+m} = T_n \circ T_m$ fuer alle $n, m \in \mathbb{Z}$.

Damit stellen also (invertierbare) maßerhaltende Transformationen diskrete Analoga zu den Gruppenwirkungen des vorigen Abschnittes dar.

Definition (Messbare dynamische Systeme). *Ein Paar bestehend aus einem Maßraum $(X, \mathcal{B}, m)$ und einer maßerhaltenden Abbildung T heißt ein meßbares dynamisches System.*

Notation. Man schreibt oft (X, T) statt $((X, \mathcal{B}, m), T)$.

Maßerhaltende Transformationen haben folgende grundlegende Eigenschaft.

Proposition (Invarianz des Integrals nichtnegativer Funktionen). *Sei $(X, \mathcal{B}, m)$ ein beliebiger Maßraum. Sei $T : X \longrightarrow X$ eine meßbare Transformation. Dann sind aequivalent:*

(i) *T ist maßerhaltend.*

(ii) *Fuer jedes meßbare nichtnegative Funktion $f : X \longrightarrow [0, \infty]$ ist auch $f \circ T$ meßbar und nichtnegativ mit*

$$\int f dm = \int f \circ T dm.$$

Beweis. (i)$\Longrightarrow$ (ii): Ist B eine Teilmenge von X und 1_B die characteristische Funktion von B, so gilt

$$1_{T^{-1}B} = 1_B \circ T.$$

Nach Voraussetzung der Maßerhaltung von T gilt also

$$(*) \quad \int_X 1_B dm = \int_X 1_B \circ T dm$$

fuer alle meßbaren Teilmengen B. Offenbar ist die Menge $\mathcal{F}$ der meßbaren nichtnegativen Funktionen f fuer die gilt

$$\int_X f dm = \int_X f \circ T dm$$

abgeschlossen unter

- Bilden von Linearkombinationen mit nichtnegativen Koeffizienten und
- Bilden von monotonen Limiten von Funktionen.

Aufgrund von $(*)$ enthaelt sie außerdem alle charakteristischen Funktionen von meßbaren Mengen. Damit ist $\mathcal{F}$ dann gerade die Menge aller nichtnegativen meßbaren Funktionen.

(ii)$\Longrightarrow$ (i): Das folgt einfach. (Waehle $f = 1_A$ mit einer meßbaren Teilmenge A von X).

$\square$

Als Konsequenz aus dem vorigen Proposition halten wir fest:

FOLGERUNG. *Sei $(X, \mathcal{B}, m)$ ein beliebiger Maßraum. Sei $T : X \longrightarrow X$ eine maßerhaltende Transformation. Dann ist fuer jedes meßbares Funktion $f : X \longrightarrow \mathbb{R}$ in L^p auch $f \circ T$ in L^p mit gleicher Norm.*

Um nachzuweisen, dass eine Transformation meßbar und maßerhaltend ist, reicht es oft, die entsprechende Eigenschaft fuer (kleine) Teilmengen der σ-Algebra nachzuweisen. Das ist bei konkreten Untersuchunen oft nuetzlich. Den noetigen Hintergrund diskutieren wir nun.

DEFINITION. *Sei $\mathcal{R}$ eine Familie von Teilmengen von X.*
(a) Es heißt $\mathcal{R}$ eine Semialgebra, wenn gilt:

- $\emptyset \in \mathcal{R}$.
- $R_1, R_2 \in \mathcal{R}$ *impliziert* $R_1 \cap R_2 \in \mathcal{R}$.
- $R \in \mathcal{R}$ *impliziert* $X \setminus R = \bigcup_{j=1}^{N} R_j$ *mit* $N \in \mathbb{N}$ *und paarweise disjunkten* $R_1, \ldots, R_N \in \mathcal{R}$.

(b) Eine Seminalgebra $\mathcal{R}$ heißt Algebra, wenn zu jedem $R \in \mathcal{R}$ auch $X \setminus R$ zu $\mathcal{R}$ gehoert.
(c) Es heißt $\mathcal{R}$ eine monotone Klasse, wenn gilt:

- $\bigcap_{j=1}^{\infty} R_j \in \mathcal{R}$ *falls* $R_j \in \mathcal{R}$ *mit* $R_1 \supset R_2 \supset \ldots$
- $\bigcup_{j=1}^{\infty} R_j \in \mathcal{R}$ *falls* $_j \in \mathcal{R}$ *mit* $R_1 \subset R_2 \subset \ldots$

Bemerkung. Die von einer Semialgebra erzeugte Algebra besteht gerade aus den endlichen Vereinigungen disjunkter Mengen aus der Semialgebra.

Wir erinnern nun an folgendes Resultat aus der Maßtheorie.

LEMMA (Monotone Klassen Lemma). *Sei $\mathcal{A}$ eine Algebra von Teilmengen einer Menge X. Dann ist die von $\mathcal{A}$ erzeugte σ-Algebra gerade die von $\mathcal{A}$ erzeugte monotone Klasse.*

Ende der 1. Vorlesung

Darauf aufbauend erhalten wir nun folgendes Kriterium zum Nachweis der Maßerhaltung einer Transformation.

THEOREM (Kriterium Maßerhaltung von T). *Seien $(X_j, \mathcal{B}_j, m_j)$ Maßraeume und $T : X_1 \longrightarrow X_2$ eine Abbildung. Sei $\mathcal{R}_2$ eine Semialgebra, die $\mathcal{B}_2$ erzeugt. Gilt fuer jedes $R \in \mathcal{R}_2$*

$$T^{-1}B \in \mathcal{B}_1 \ \text{ und } \ m_1(T^{-1}R) = m_2(R),$$

so ist T eine maßerhaltende Transformation.

Beweis. Set

$$\mathcal{C} := \{B \in \mathcal{B}_2 : T^{-1}(B) \in \mathcal{B}_1 \text{ and } m_1(T^{-1}B) = m_2(B)\}.$$

Zu zeigen: $\mathcal{C} = \mathcal{B}_2$.

Man sieht leicht:

- $\mathcal{C}$ ist abgeschlossen unter dem Bilden endlicher disjunkter Vereinigungen.
- $\mathcal{C}$ ist eine monotone Klasse.

Nach Voraussetzung gilt $\mathcal{R}_2 \subset \mathcal{C}$.
Nach dem ersten Punkt enthaelt $\mathcal{C}$ dann die von $\mathcal{R}_2$ erzeugte Algebra, $\mathcal{A}_2$. Nach dem zweiten Punkt enthaelt es dann die von $\mathcal{A}_2$ erzeugte monotone Klasse. Damit enthaelt aber $\mathcal{C}$ nach dem monotonen Klassen Lemma die von $\mathcal{A}_2$ erzeugte σ-Algebra. Nach Voraussetzung ist dies aber gerade $\mathcal{B}_2$. □

2. Zwei Fragestellungen und eine Eigenschaft

Sei (X, T, m) ein meßbares dynamisches System. Dann koennen wir zwei Fragestellungen angeben, um die es im folgenden viel gehen wird:

Grenzwerte ergodischer Mittel. Sei $f : X \longrightarrow \mathbb{R}$ meßbar. Dann definieren wir das *N-te ergodische Mittel* von f als die Funktion

$$A_N f : X \longrightarrow \mathbb{R}, A_N f(x) = \frac{1}{N} \sum_{k=0}^{N-1} f(T^k x).$$

Ein grundlegende Frage betrifft dann die Konvergenz der ergodischen Mittel in einem geeigneten Sinne.

Rekurrenz. Es geht um die Frage, wie der *Orbit* $\{T^n x : n \in \mathbb{N}_0\}$ von x in X liegt. Stattdessen kann man auch die Frage untersuchen, wie sich die Urbilder $T^{-n} A$, $n \in \mathbb{N}_0$, fuer ein meßbares A im Raum verteilen.

Bei den Betrachtungen werden immer wieder meßbare Funktionen $f : X \longrightarrow \mathbb{R}$ auftauchen, die T-invariant sind d.h. $f = f \circ T$ erfuellen. Systeme, in denen jede solche Funktion fast sicher konstant ist, heißen *ergodisch*. In diesem Zusammenhang halten wir fest.

Lemma (Erodizitaet via invariante Funktionen). *Sei (X, T, m) ein meßbares dynamisches System mit $m(X) < \infty$. Dann sind die folgenden Aussagen aequivalent:*

(i) *Jede meßbare Funktion $f : X \longrightarrow \mathbb{R}$ mit $f(x) = f(Tx)$ fuer fast alle $x \in X$ ist fast sicher konstant.*
(ii) *Jede meßbare Funktion $f : X \longrightarrow \mathbb{R}$ mit $f(x) = f(Tx)$ fuer alle $x \in X$ ist fast sicher konstant.*

Beweis. (i)$\Longrightarrow$ (ii): Das ist klar.
(ii)$\Longrightarrow$ (i): Sei $f : X \longrightarrow \mathbb{R}$ mit $f(x) = f(Tx)$ fuer fast alle $x \in X$. Es reicht ein $g : X \longrightarrow \mathbb{R}$ zu finden mit

- $g(x) = g(Tx)$ fuer alle $x \in X$.
- $g = f$ fast ueberall.

Definiere

$$g : X \longrightarrow \mathbb{R}, g(x) := \limsup_N \frac{1}{N} \sum_{k=0}^{N-1} f(T^k x).$$

Dann gilt offenbar $g(x) = g(Tx)$ fuer alle $x \in X$. Sei weiterhin N eine Nullmenge mit $f(x) = f(Tx)$ fuer alle $x \notin N$. Sei

$$\widetilde{N} := \bigcup_{n \in \mathbb{N}_0} T^{-n} N = \{x : T^n x \in N \text{ fuer ein } n \in \mathbb{N}_0\}.$$

Dann ist $\widetilde{N}$ eine Nullmenge (als abzaehlbare Vereinigung von Nullmengen) und fuer $x \notin \widetilde{N}$ gilt

$$f(x) = f(Tx) = f(T^2 x) = ... = f(T^k x) = ...$$

Damit gilt dann fuer diese x auch $f(x) = g(x)$. □

3. Einige Beispiele

In diesem Abschnitt stellen wir einige Beispiele von meßbaren dynamischen Systemen vor.

3.1. Die Identitaet. Sei $(X, \mathcal{B}, m)$ ein beliebiger Maßraum. Sei $T : X \longrightarrow X$ die Identitaet.

3.2. Rotation auf dem Einheitskreis. Sei $\mathbb{T} := \mathbb{R}/\mathbb{Z}$ und $\pi : \mathbb{R} \longrightarrow \mathbb{T}$ die kanonische Projektion. Dann gibt es eine eindeutige σ-Algebra $\mathcal{B}$ auf $\mathbb{T}$, so daß $\pi^{-1}(B)$ zur Borel-σ-Algebra auf $\mathbb{R}$ gehoert fuer jedes $B \in \mathcal{B}$. Auf dieser σ-Algebra gibt es ein eindeutiges Maß m, so daß

$$m(B) = \lambda(\pi^{-1}(B) \cap [0,1])$$

gilt fuer jedes $B \in \mathcal{B}$. Die Translationsinvarianz des Lebesguemaß λ liefert, daß m fuer jedes $\alpha \in \mathbb{R}$ invariant ist unter den Abbildung

$$T = T_\alpha : \mathbb{T} \longrightarrow \mathbb{T},\ T(x + \mathbb{Z}) := x + \alpha + \mathbb{Z}.$$

Tatsaechlich ist T sogar invertierbar und das Inverse ist gegeben durch die Abbildung $T_{-\alpha}$.

Eine alternative Darstellung dieses dynamisschen System ergibt sich wie folgt.

$$S^1 := \{z \in \mathbb{C} : |z| = 1\}$$

zusammen mit der Borel-σ-Algebra und dem kanonischen Maß auf S^1. Sei $\alpha \in \mathbb{R}$ gegeben und

$$T = T_\alpha : S^1 \longrightarrow S^1, Tz = e^{2\pi i \alpha} z.$$

Dann ist T eine maßerhaltende Transformation.

Erinnerung. Ist G eine lokalkompakte topologische Gruppe, so gibt es auf G ein (bis auf Normierung) eindeutiges linksinvariantes Maß λ auf den Borel-Mengen. Dieses Maß heißt das Haarmaß auf G.

Das Beispiel kann verallgemeinert werden wie folgt. Sei G eine beliebige kompakte Gruppe zusammen mit der Borel-σ-Algebra und dem Haarmaß. Sei $a \in G$ beliebig. Dann ist $T = T_a$ mit

$$T : G \longrightarrow G, Tx := ax$$

eine meßbare Transformation. Tatsaechlich ist T invertierbar mit inverser gegeben durch $T_{a^{-1}}$.

3.3. Homomorphismen auf dem Einheitskreis. Definiere fuer $p \in \mathbb{Z}$ die Abbildung

$$T : \mathbb{T} \longrightarrow \mathbb{T}, T(x + \mathbb{Z}) = px + \mathbb{Z}.$$

Dann ist T eine meßbare Transformation. Eine alternative Darstellung ist gegeben durch

$$T : S^1 \longrightarrow S^1, Tz = z^p.$$

Man kann direkt nachrechnen, daß T maßerhaltend ist (z.b. auf Intervallen). Es ist T fuer $p \neq \pm 1$ nicht invertierbar.

Das Beispiel kann verallgemeinert werden wie folgt. Sei G eine beliebige kompakte Gruppe zusammen mit der Borel-σ-Algebra und dem Haarmaß. Sei $A : G \longrightarrow G$ ein stetiger surjektiver Gruppenhomomorphismus. Dann ist T gegeben durch

$$T : G \longrightarrow G, Tx = A(x)$$

eine maßerhaltende Transformation. (Bew. Definiere das Maß μ durch $\mu(B) := \lambda(T^{-1}B)$. Zeige, daß μ normiert und linksinvariant ist.)

3.4. Der Bernoulli Shift. Sei $Y := \{0, 1\}$ mit der σ-Algebra bestehend aus allen Teilmengen und sei μ das Maß auf Y mit $\mu(\{0\}) = \mu(\{1\}) = 1/2$. Sei

$$X := Y^{\mathbb{Z}} = \{f : \mathbb{Z} \longrightarrow Y\} = \prod_{j=-\infty}^{\infty} Y$$

mit der Produkt σ-Algebra und dem Produktmaß m. Sei

$$T : X \longrightarrow X, (Tx)(n) = x(n + 1).$$

Es heißt T Shiftoperator und (X, T) der Bernoullishift. Wir zeigen nun, daß T maßerhaltende Transformation ist. Dazu erinnern wir an die Definition von Produktmaß und Produkt-σ-Algebra. Fuer $r, s \in \mathbb{Z}$ mit $r \leq s$ und $A_j \subset \{0, 1\}$, $j = 0, \ldots, s - r$, definiert man das *meßbare Rechteck*

$$[A_0, \ldots, A_{s-r}]_{r,s} \in X$$

durch

$$[A_0, \ldots, A_{s-r}]_{r,s} := \{x \in X : x_{r+j} \in A_j, j = 0, \ldots, s-r\} = \prod_{-\infty}^{r-1} Y \times \prod_{j=0}^{s-r} A_{j+r} \times \prod_{s-r+1}^{\infty} Y.$$

Dann ist die Produkt-σ-Algebra die von den meßbaren Rechtecken erzeugte σ-Algebra und das Produktmaß ist das eindeutige Maß m mit der (charakteristischen) Eigenschaft

$$m([A_0, \ldots, A_{s-r}]_{r,s}) = \prod_{j=0}^{s-r} \mu(A_j)$$

fuer alle meßbaren Rechtecke. Eine direkte Rechnung zeigt nun

$$T^{-1}([A_0, \ldots, A_{s-r}]_{r,s} = [A_0, \ldots, A_{s-r}]_{r+1,s+1}.$$

Aus der charakteristischen Eigenschaft des Maßes folgt nun

$$m(T^{-1}([A_0, \ldots, A_{s-r}]_{r,s})) = m([A_0, \ldots, A_{s-r}]_{r+1,s+1})) = m([A_0, \ldots, A_{s-r}]_{r,s}).$$

Damit ist dann also $T^{-1}R$ meßbar mit $m(T^{-1}R) = m(R)$ fuer alle R aus der Familie $\mathcal{R}$ der meßbaren Rechtecke. Diese Familie erfuellt nun - wie man leicht sieht - die folgenden Eigenschaften:

- $\emptyset \in \mathcal{R}$.
- $R_1, R_2 \in \mathcal{R}$ impliziert $R_1 \cap R_2 \in \mathcal{R}$.
- $R \in \mathcal{R}$ impliziert $X \setminus R = \bigcup_{j=1}^{n} R_j$ mit geeignetem $n \in \mathbb{N}$ und paarweise disjunkten $R_j \in \mathcal{R}$.

Des weiteren erzeugt diese Familie die Produkt-σ-Algebra. Damit folgt aus dem Kriterium zur Maßerhaltenden Transformation die gewuenschte Aussage.
Man ueberlegt sich leicht, daß T invertierbar ist und die inverse gegeben ist durch

$$S : X \longrightarrow X, (Sx)(n) = x(n-1).$$

3.5. Der $(p_1, \ldots, p_k)$ **- Shift.** . Seien $k \geq 2$ und $p_1, \ldots, p_k > 0$ mit $\sum_{j=1}^{k} p_j = 1$ gegeben. Sei $Y := \{0, \ldots, k-1\}$ mit der σ-Algebra aller meßbaren Teilmengen ausgestattet und sei μ auf Y das eindeutige Maß mit

$$\mu(\{j\}) = p_{j+1}$$

fuer $j = 0, \ldots, k-1$. Sei $X := Y^{\mathbb{Z}}$ mit der Produkt-σ-Algebra und dem Produktmaß. Sei

$$T : X \longrightarrow X, (Tx)(n) = x(n+1).$$

Dann ist T eine invertierbare maßerhaltende Transformation.
Das folgt durch einfache Adaption des Beweises im vorigen Beispiel.

3.6. Der einseitige $(p_1, \ldots, p_k)$ **- Shift.** Seien $k \geq 2$ und $p_1, \ldots, p_k > 0$ mit $\sum_{j=1}^{k} p_j = 1$ gegeben. Sei $Y := \{0, \ldots, k-1\}$ mit der σ-Algebra aller meßbaren Teilmengen ausgestattet und sei μ auf Y das eindeutige Maß mit

$$\mu(\{j\}) = p_{j+1}$$

fuer $j = 0, \ldots, k-1$. Sei $X := Y^{\mathbb{N}}$ mit der Produkt-σ-Algebra und dem Produktmaß. Sei

$$T : X \longrightarrow X, (Tx)(n) = x(n+1).$$

Dann ist T eine maßerhaltende Transformation. Das folgt durch einfache Adaption des Beweises im vorigen Beispiel. Es operiert T durch 'Abschneiden' der ersten Koordinate. Entsprechend ist T nicht invertierbar. (Denn es ist auf einer 'grossen' Menge nicht injektiv. Tatsaechlich hat jedes ω genau k-Urbilder.)

KAPITEL 2

Koopmanismus und von Neumanns Ergodensatz

Sei (X, T, m) ein meßbares dynamisches System. Eine grundlegende Beobachtung von Koopman ist nun, daß T einen isometrischen Operator

$$U_T : L^2(X, m) \longrightarrow L^2(X, m), U_T f = f \circ T$$

erzeugt. Damit lassen sich Eigenschaften von T mittels Hilbertraumtheorie studieren. Dieser Zugang ist als Koopmanismus bekannt.

1. Isometrien im Hilbertraum

In diesem Abschnitt wiederholen wir einige grundlegende Sachverhalte aus der Theorie der Hilbertraeume.

Ein vollstaendiger normierter Raum $\mathcal{H}$ mit Norm $\|\cdot\|$ heißt Hilbertraum, wenn die Norm induziert wird durch ein Skalarprodukt

$$\langle\cdot,\cdot\rangle : \mathcal{H} \times \mathcal{H} \longrightarrow \mathbb{C}$$

via

$$\|f\| = \langle f, f\rangle^{1/2}.$$

Das Skalarprodukt ist linear im zweiten Argument.

Zu einer beliebigen Teilmenge C von $\mathcal{H}$ definiert man das orthogonale Komplement $C^\perp$ durch

$$C^\perp := \{f \in \mathcal{H} : \langle f, g\rangle = 0 \text{ fuer alle } g \in C\}.$$

Dann ist C ein abeschlossener Unterraum. Weiterhin gilt fuer einen beliebigen Unterraum C in $\mathcal{H}$ nach dem Projektionssatz

$$C^{\perp\perp} = \overline{C}$$

und fuer eine abgeschlossenen Unterraum C gilt

$$\mathcal{H} = C \oplus C^\perp.$$

Eine linearer Operator $A : \mathcal{H} \longrightarrow \mathcal{H}$ heißt beschraenkt, wenn ein $C \geq 0$ existiert mit

$$\|Af\| \leq C\|f\|$$

fuer alle $f \in \mathcal{H}$. Zu jedem beschraenkten Operator A existiert ein eindeutiger Opertor A^* - der *adjungierte* von A - mit

$$\langle Af, g\rangle = \langle f, A^* g\rangle$$

fuer alle $f, g \in \mathcal{H}$. Ist A ein beschraenkter Operator, so definiert man seinen Kern $Ker(A)$ durch $Ker(A) = \{f \in \mathcal{H} : Af = 0\}$ und sein Bild $Ran(A)$ durch $Ran(A) = \{Af : f \in \mathcal{H}\}$

PROPOSITION. *Sei A ein beschraenkter Operator im Hilbertraum $\mathcal{H}$. Dann gilt*

$$Ker(A) = Ran(A^*)^\perp.$$

Beweis. Das folgt durch direkte Rechnung:

$$\begin{aligned} Ker(A) &= \{f : Af = 0\} \\ &= \{f : \langle Af, g\rangle = 0 \text{ fuer alle } g \in \mathcal{H}\} \\ &= \{f : \langle f, A^*g\rangle = 0 \text{ fuer alle } g \in \mathcal{H}\ \} \\ &= \{A^*g : g \in \mathcal{H}\}^\perp. \end{aligned}$$

□

Ein Operator $U : \mathcal{H} \longrightarrow \mathcal{H}$ heißt Isometrie, wenn eine der folgenden drei aequivalenten Eigenschaften gilt:

- $\|Uf\| = \|f\|$ fuer alle $f \in \mathcal{H}$.
- $\langle Uf, Ug\rangle = \langle f, g\rangle$ fuer alle $f, g \in \mathcal{H}$.
- $U^*U = id$.

Ende der 2. Vorlesung

Ein Operator $U : \mathcal{H} \longrightarrow \mathcal{H}$ heißt unitaer, wenn eine der folgenden drei aequivalenten Eigenschaften gilt:

- U ist eine surjektive Isometrie.
- $U^*U = id = UU^*$.
- $U^* = U^{-1}$.

PROPOSITION (Struktur Kern und Bild von Isometrien). *Sei $U : \mathcal{H} \longrightarrow \mathcal{H}$ eine Isometrie. Dann gilt:*

- $Ker(U - id) = Ker(U^* - id)$.
- $Ker(U - id)^\perp = \overline{Ran(U - id)}$.

Insbesondere gilt dann also

$$\mathcal{H} = Ker(U - id) \oplus \overline{Ran(U - id)}.$$

Beweis. Zum ersten Punkt: Es sind zwei Inklusionen zu zeign.
$\subset$: $(U - id)f = 0$ impliziert $0 = U^*(U - id)f = U^*Uf - U^*f = f - U^*f$.
$\supset$: Sei $(U^* - id)f = 0$. Dann gilt also $f = U^*f$. Damit folgt

$$\|(U - id)f\|^2 = \langle Uf, Uf\rangle - \langle Uf, f\rangle - \langle f, Uf\rangle + \langle f, f\rangle = 0.$$

Zum zweiten Punkt: Unter Nutzen des ersten Punktes und der allgemeinen Struktur folgt:

$$\begin{aligned} Ker(U - id)^* &= Ker(U^* - id)^\perp \\ (KerA, RanA^*) &= (Ran(U^{**} - id)^\perp)^\perp \\ &= (Ran(U - id))^{\perp\perp} \\ (Projektionssatz) &= \overline{Ran(U - id)}. \end{aligned}$$

□

2. Der von Neumannsche Ergodensatz

Mit den im vorigen Abschnitt wiederholten Eigenschaften von Isometrien ist es sehr leicht einen ersten Ergodensatz zu beweisen. Die Bezeichnung Ergodensatz ist an dieser Stelle noch nicht verstaendlich.

THEOREM (von Neumannscher Ergodensatz - Isometrie im Hilbertraum). *Sei U ein Isometrie auf dem Hilbertraum $\mathcal{H}$. Sei P die orthogonale Projektion auf $Ker(U - id)$. Dann gilt*

$$\frac{1}{N}\sum_{k=0}^{N-1} U^k\xi \to P\xi, N \to \infty,$$

fuer jedes $\xi \in \mathcal{H}$.

Bemerkung. Die Behauptung gilt auch wenn $\frac{1}{N}\sum_{k=0}^{N-1}$ durch $\frac{1}{N}\sum_{k=p}^{N+q}$ mit beliebigen aber festen $p, q \in \mathbb{N}$ ersetzt werden (Aenderung der Summe in endlich vielen Termen).

Beweis. Definiere fuer $N \in \mathbb{N}_0$ den Mittelungsoperator $A_N := \frac{1}{N}\sum_{k=0}^{N-1}$. Dann gilt - wie man leicht sieht - $\|A_N\| \leq 1$ fuer alle $N \in \mathbb{N}_0$. Zu zeigen:

$$A_N\xi \to P\xi, N \to \infty,$$

fuer alle $\xi \in \mathcal{H}$. Nach dem Strukurproposition zu Bild und Kern von Isometrien aus dem vorigen Abschnitt gilt

$$\mathcal{H} = Ker(U - id) \oplus \overline{Ran(U - id)}.$$

Wir zeigen die entsprechende Konvergenz auf $Ker(U-id)$ und auf $\overline{Ran(U - id)}$.

Es gilt $A_N\xi \to P\xi$, $N \to \infty$, fuer $\xi \in Ker(U - id)$: Sei $\xi \in Ker(U - id)$. Dann gilt also $U\xi = \xi$ und damit $U^k\xi = \xi$ fuer alle $k \in \mathbb{N}_0$. Damit folgt die gewuenschte Aussage sofort.

Es gilt $A_N\xi \to P\xi$, $N \to \infty$, fuer $\xi \in \overline{Ran(U - id)}$: Aufgrund der Zerlegung des Hilbertraumes gilt

$$P\xi = 0$$

fuer alle $\xi \in \overline{Ran(U - id)}$. Wegen $\|A_N\| \leq 1$ fuer alle $N \in \mathbb{N}_0$ reicht es $\xi \in Ran(U - id)$ zu betrachten. Sei also

$$\xi = (U - id)\eta.$$

Dann wird $A_N\xi$ eine Teleskopsumme und die gewuenschte Aussage folgt einfach:

$$\begin{aligned} A_N\xi &= A_N(U - id)\eta \\ &= A_NU\eta - A_N\eta \\ &= \frac{1}{N}\sum_{k=1}^{N} U^k\eta - \frac{1}{N}\sum_{k=0}^{N-1} U^k\eta \\ &= \frac{1}{N}(U^N\eta - \eta) \\ &\to 0, N \to \infty. \end{aligned}$$

Das ist die gewuenschte Aussage. □

Bemerkung. Die Aussage läßt sich - mit praktisch dem gleichen Beweis - sehr verallgemeinern. Das diskutieren wir in einer Variante: Sei H eine (abzaehlbare) Halbgruppe mit Kuerzung (d.h. $mn = km \Longrightarrow m = k$) und

$$U : H \longrightarrow \text{Isometrien auf } \mathcal{H}$$

eine Abbildung mit $U_{nm} = U_n U_m$ (Darstellung der Halbgruppe). Es gebe eine Folge I_N von endlichen Teilmengen von H mit

$$\frac{\sharp m I_N \triangle I_N}{\sharp I_N} \to 0, N \to \infty$$

fuer jedes $m \in H$. Sei

$$\mathcal{R} := \overline{\sum_n (Ran(U_n - id)}, \text{ und } \mathcal{K} := \cap_{n \in H} Ker(U_n - id).$$

Dann gilt

$$\mathcal{H} = \mathcal{K} \oplus \mathcal{R}$$

und mit der Projektion P auf $\mathcal{K}$ und dem Mittelungsoperator

$$A_N := \frac{1}{I_N} \sum_{k \in I_N} U_k$$

gilt

$$A_N \xi \to P\xi, N \to \infty.$$

Fuer unitaere Abbildungen erhalten wir noch folgende Variante des von Neumannschen Ergodensatzes.

Corollary. *Sei U unitaer auf dem Hilbertraum $\mathcal{H}$. Sei P die orthogonale Projektion auf $Ker(U - id)$. Dann gilt*

$$\frac{1}{2N+1} \sum_{k=-N}^{N-1} U^k \xi \to P\xi, N \to \infty,$$

fuer jedes $\xi \in \mathcal{H}$.

Beweis. Da U unitaer ist, sind sowohl U als auch sein Inverses $V := U^{-1}$ Isometrien und es gilt

$$Ker(V - id) = Ker(U - id) =: \mathcal{K}.$$

Sei P die Projektion auf $\mathcal{K}$. Dann liefert der von Neumannsche Ergodensatz (angewendet auf U bzw V) also

$$\frac{1}{N} \sum_{k=0}^{N-1} U^k \xi \to P\xi, N \to \infty,$$

bzw.

$$\frac{1}{N} \sum_{k=0}^{N-1} V^k \xi \to P\xi, N \to \infty,$$

fuer jedes $\xi \in \mathcal{H}$. Mit

$$\frac{1}{2N+1} \sum_{k=-N}^{N} U^k \xi = \frac{N}{2N+1} \frac{1}{N} \sum_{k=0}^{N-1} U^k \xi + \frac{N}{2N+1} \sum_{k=0}^{N-1} V^k \xi + \frac{1}{2N+1} U^N \xi - \frac{1}{2N+1} \xi$$

folgt dann leicht die gewuenschte Aussage. □

3. Koopmanismus - ein erster Blick

In diesem Abschnitt betrachten wir ein maßerhaltenden Transformation T auf $(X, \mathcal{B}, m)$. Zu dieser Transformation assoziieren wir die Abbildung

$$U_T : L^2(X, m) \longrightarrow L^2(X, m), U_T f = f \circ T.$$

Lemma. *Sei $(X, \mathcal{B}, m)$ ein beliebiger Maßraum. Sei $T : X \longrightarrow X$ eine maßerhaltende Transformation. Dann ist die Abbildung*

$$U_T : L^2(X, m) \longrightarrow L^2(X, m), f \mapsto f \circ T,$$

wohldefiniert und eine Isometrie auf $L^2(X, m)$. Ist T sogar invertierbar, so ist U_T unitaer.

Beweis. Eine vorangegangene Proposition zur Invarianz des Integrals liefert leicht die Wohldefiniertheit und Isometrie von U_T.
Ist T invertierbar, so ist sein Inverses S ebenfalls eine maßerhaltenden Transformation und man sieht leicht, dass U_T und U_S zueinander inverse Isometrien sind. Damit sind dann U_T und U_S unitaer. □

Bemerkung. Der einseitige $(1/2, 1/2)$ - Shift liefert ein Beispiel fuer eine isometrische nicht unitaere Transformation. Tatsaechlich gilt (mit Notation von oben) nach Definition des Produktmaßes

$$\int_X g f \circ T dm = \int_X g dm \int_X f dm$$

fuer jede Funktion g die nur von der ersten Variablen abhaengt. Insbesondere ist also jedes g mit $\int_X g dm = 0$, das nur von der ersten Variablen abhaengt senkrecht auf dem Bild von U_T. Damit ist U_T also nicht surjektiv.

Aus dem Lemma und dem von Neumannschen Ergodensatz erhaelt man sofort folgenden Ergodensatz.

Theorem (von Neumannscher Ergodensatz - maßerhaltende Transformation). *Sei $(X, \mathcal{B}, m)$ ein beliebiger Maßraum. Sei $T : X \longrightarrow X$ eine maßerhaltende Transformation. Sei*

$$\mathcal{K} := \{f \in L^2(X, m) : f = f \circ T\}$$

und P die Projektion auf $\mathcal{K}$. Dann gilt

$$\frac{1}{N} \sum_{k=0}^{N-1} f \circ T^k \to Pf, N \to \infty,$$

fuer jedes $f \in L^2(X, m)$.

In dieser Fassung des Ergodentheoremes wird der Bezug zur Einleitung schon deutlicher.
Wir betrachten nun den Fall $m(X) < \infty$. Dann koennen wir auch ohne Einschraenkung annehmen $m(X) = 1$. Es ist dann die konstante Funktion 1 (und ihre Vielfache) auf jeden Fall ein Element von $\mathcal{K}$. Von besonderem

Interesse wird die Situation sein, daß $\mathcal{K}$ gerade aus den Vielfachen der konstanten Funktion besteht:

(E) $\mathcal{K} = Lin\{1\}$ (wobei 1 die konstante Funktion bezeichnet).

Zerlegt man dann ein $f \in L^2(X, m)$ in $f = f^\perp + f_0$ mit $f_0 \in \mathcal{K}$ und $f^\perp \perp \mathcal{K}$, so gilt also

$$f_0 = \langle 1, f \rangle 1 = \int_X f dm 1 \text{ und } 0 = \langle 1, f^\perp \rangle = \int_X f^\perp dm.$$

Folgerung. *Sei die Situation wie im vorigen Satz. Es gelte zusaetzlich noch (E). Dann konvergieren fuer jedes $f \in L^2(X, m)$ die Mittel $\frac{1}{N} \sum_{k=0}^{N-1} f \circ T^k$ gegen die konstante Funktion $\int_X f dm$.*

Nach dem Korollar stellt die Bedingung (E) also gerade die Konvergenz der diskreten Zeitmittel gegen das Raummitel sicher. Wir werden die Bedingung (E) im kommenden Abschnitt weiter untersuchen.

Ende der 3. Vorlesung

4. Eine Bemerkung zu Ergodizitaet

Im vorigen Abschnitt haben wir die Ergodizitaetsbedingung (E) im L^2 Kontext kennengelernt:

$$\{f \in L^2(X, m) : f = f \circ T\} = Lin\{1\}.$$

Hierbei ist $(X, \mathcal{B}, m)$ ein beliebiger Maßraum mit $m(X) < \infty$ und $T : X \longrightarrow X$ eine maßerhaltende Transformation. Hier lernen wir einige aequivalente Formulierungen kennnen:

Lemma (Erodizitaet via invariante Funktionen). *Ein Sei (X, T, m) ein meßbares dynamisches System mit $m(X) < \infty$. Dann sind die folgenden Aussagen aequivalent:*

(i) *Jede Funktion $f \in L^2(X, m)$ mit $U_T f = f$ ist fast sicher konstant.*
(ii) *Jede meßbare Funktion $f : X \longrightarrow \mathbb{R}$ mit $f(x) = f(Tx)$ fuer fast alle $x \in X$ ist fast sicher konstant.*
(iii) *Jede meßbare Funktion $f : X \longrightarrow \mathbb{R}$ mit $f(x) = f(Tx)$ fuer alle $x \in X$ ist fast sicher konstant.*

Beweis. (i)$\Longrightarrow$ (ii): Das folgt leicht durch Abschneiden.

(ii)$\Longrightarrow$ (i): Das ist klar.

(ii)$\Longleftrightarrow$ (iii): Das haben wir oben schon gesehen. $\square$

KAPITEL 3

Das Birhkoffsche Ergodentheorem

Sei (X, T, m) ein meßbares dynamisches System. Im vorigen Abschnitt haben wir fuer $f \in L^2(X, m)$ die Konvergenz von

$$\frac{1}{N}\sum_{k=0}^{N-1} f \circ T^k \stackrel{L^2}{\to} Pf$$

gezeigt mit der Projektion P auf die invarianten Funktionen in L^2 (von Neumannscher Ergodensatz). In diesem Abschnitt geht es fuer $f \in L^1(X, m)$ die Konvergenz von

$$\frac{1}{N}\sum_{k=0}^{N-1} f \circ T^k(x) \stackrel{punktweise f.u.}{\to} f^*(x)$$

mit einem geeigneten f^* (Birkhoffscher Ergodensatz).

Achtung. Das ist eine tiefliegende Aussage, da es keine Topologie der punktweisen fast sicheren Konvergenz gibt. Es handelt sich also nicht um eine topologisch / geometrische Aussage.

Entscheidend wird die Kontrolle ueber die Maximalfunktion

$$\sup_N \frac{1}{N}\sum_{k=0}^{N-1} f(T^k x)$$

sein (Maximaler Ergodensatz).

1. Der Birkhoffsche Ergodensatz

Wir brauchen etwas Vorbereitung.

LEMMA. *Sei $(X, \mathcal{B}, m)$ ein Maßraum und $U : L^1(X, m) \longrightarrow L^1(X, m)$ linear und positivitaetserhaltend (d.h. $Uf \geq 0$ f. ü. falls $f \geq 0$ f. ü.) mit $\|U\| \leq 1$. Sei $\in L^1(X, m)$ beliebig. Sei $f_0 := 0$ und $f_n := f + Uf + \ldots + U^{n-1}f$ fuer $n \in \mathbb{N}$. Sei*

$$F_N := \max\{f_0, \ldots, f_N\}.$$

Dann gilt

$$\int_{\{x: F_N(x) > 0\}} f dm \geq 0$$

fuer jedes $N \in \mathbb{N}$.

Bemerkungen.

- Es gilt $F_N(x) > 0$ genau dann, wenn
$$\sum_{k=1}^{n} U^n f(x) > 0$$
gilt fuer ein $n \in \{1, \dots, N\}$. (Hier ist f_0 nicht noetig!).
- Aus $F_N(x) > 0$ folgt im allgemeinen nicht $f(x) > 0$. Entsprechend ist die Aussage sehr bemerkenswert.

Beweis. Fuer alle $0 \leq k \leq N$ gilt $F_N \geq f_k$ (nach Definition von F_N als Maximum.) Da U positivitaetserhaltend ist folgt dann $UF_N \geq Uf_k$ und damit

$$UF_n f + f \geq Uf_k + f = f_{k+1}.$$

Wir koennen dies fuer alle k mit $0 \leq k \leq N-1$ anwenden und erhalten also

$$\begin{aligned} UF_N(x) + f(x) &\geq \max_{1 \leq k \leq N} f_k(x) \\ (\text{falls } F_N(x) > 0) \quad &= \max_{0 \leq k \leq N} f_k(x) \\ &= F_N(x). \end{aligned}$$

Auf der Menge

$$A := \{x : F_N(x) > 0\}$$

gilt also die Ungleichung

$$f(x) \geq F_N(x) - UF_N(x).$$

Integrieren liefert dann

$$\begin{aligned} \int_A f dm &= \int_A F_N dm - \int_A UF_N dm \\ (F_N = 0 \ \text{ auf } X \setminus A) \quad &= \int_X F_N dm - \int_A F_N dm \\ (F_N \geq 0, \text{ U positiverh}) \quad &\geq \int_X F_N - \int_X UF_N dm \\ &= \|F_N\|_1 - \|UF_N\|_1 \\ (\|U\| \leq 1) \quad &\geq 0. \end{aligned}$$

Das beendet den Beweis. □

Folgerung. *Sei $(X, \mathcal{B}, m)$ ein Maßraum und $T : X \longrightarrow X$ eine maßerhaltende Transformation. Sei $f \in L^1(X, m)$ reellwertig und sei $\alpha \in \mathbb{R}$ gegeben. Sei*

$$B_\alpha := \{x : \sup_N \frac{1}{N} \sum_{k=0}^{N-1} f \circ T^k(x) > \alpha.$$

Dann gilt fuer jede meßbare Teilmenge A von X mit $T^{-1}A = A$ und $m(A) < \infty$ die Abschaetzung

$$\int_{B_\alpha \cap A} f dm \geq \alpha m(B_\alpha \cap A).$$

Beweis. Ohne Einschraenkung sei $A = X$ und $m(X) < \infty$. (Andernfalls ersetze man einfach T durch die Einschraenkung von T auf A.) Sei $f := g - \alpha$. Sei F_N zu f wie in der vorangehenden Proposition assoziiert (mit U der Abbildung $f \mapsto f \circ T$). Dann gilt also

$$F_N(x) = \max\{0, \max_{0 \le n \le N-1} \sum_{k=0}^{n} f(T^k x)\} = \max\{0, \max_{0 \le n \le N-1} \sum_{k=0}^{n} (f(T^k x) - \alpha)\}.$$

Insbesondere gilt

$$F_N(x) > 0 \Longleftrightarrow \max_{0 \le n \le N-1} \sum_{k=0}^{n} (f(T^k x) - \alpha) > 0 \Longleftrightarrow \max_{0 \le n \le N-1} \frac{1}{n} \sum_{k=0}^{n} f(T^k x) - \alpha > 0.$$

Damit folgt also

$$B_\alpha = \bigcup_{N=0}^{\infty} \{x : F_N(x) > 0\}.$$

Aus dem vorangehenden maximalen Ergodentheorem ergibt sich dann

$$\int_{B_\alpha} f dm \ge 0.$$

Einsetzen der Definition von f liefert die gewuenschte Behauptung

$$\int_{B_\alpha} g dm \ge \alpha m(B_\alpha).$$

Das beendet den Beweis. □

Damit koennen wir nun zum Birkhoffschen Ergodensatz kommen.

Theorem (Birkhoffscher Ergodensatz - endliche Maßraeume). *Sei $(X, \mathcal{B}, m)$ ein endlicher Maßraum und $T : X \longrightarrow X$ maßerhaltend. Dann konvergiert fuer jedes $f \in L^1(X, m)$ die Folge der Mittel*

$$\frac{1}{N} \sum_{k=0}^{N-1} f \circ T^k$$

punktweise fast ueberall gegen eine Funktion $f^* \in L^1(X, m)$ mit $f^* = f^* \circ T$ und $\int_X f dm = \int_X f^* dm$.

Beweis. Es reicht reellwertige f zu betrachten. Wir definieren

$$f^*(x) = \limsup_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} f(T^k x)$$

und

$$f_*(x) = \liminf_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} f(T^k x).$$

Dann gilt offenbar

$$f^* = f^* \circ T, \quad f_* = f_* \circ T$$

und $f_* \le f^*$.

Ende der 4. Vorlesung

Wir zeigen:

Es gehoeren f^* und f_* zu $L^1(X,m)$ mit $\int f_* \leq \int f dm \leq \int f^* dm$. (Wegen $f_* \leq f^*$ folgt dann sofort $f_* = f^*$ fast ueberall. Damit folgt dann die punktweise fast sichere Konvergenz der Mittel gegen $f* = f_*$ sowie die gewuenschte Aussage ueber das Integral.)

Bew. Definiere fuer $n \in \mathbb{N}$ und $k \in \mathbb{Z}$

$$D_k^n := \{x \in X : \frac{k}{n} \leq f^* < \frac{k+1}{n}\}.$$

Damit ist dann also fuer festes $n \in \mathbb{N}$ die Menge X die disjunkte Vereinigung der D_k^n, $k \in \mathbb{Z}$. Insbesondere gilt also

$$m(X) = \sum_k m(D_k^n).$$

Weiterhin gilt fuer alle $\varepsilon > 0$

$$T^{-1} D_k^n = D_k^n, \;\; D_k^n \cap B_{\frac{k}{n}-\varepsilon} = D_k^n.$$

Damit folgt aus der vorangehenden Folgerung also

$$\int_{D_k^n} f dm \geq \left(\frac{k}{n} - \varepsilon\right) m(D_k^n).$$

Da $\varepsilon > 0$ beliebig war, ergibt sich

$$(*) \;\; \int_{D_k^n} f dm \geq \frac{k}{n} m(D_k^n).$$

Mit der Definition von D_k^n und $(*)$ folgt sofort

$$(*) \;\; \int_{D_k^n} f^* dm \leq \frac{k+1}{n} m(D_k^n) \leq \frac{1}{n} m(D_k^n) + \int_{D_k^n} f dm.$$

Fuer $f \geq 0$ gilt offenbar $f^* \geq 0$ und Summierbarkeit der rechten Seite liefert dann $f^* \in L^1(X,m)$. Fuer beliebiges reelles f mit $f = f_+ - f_-$ mit $f_\pm \geq 0$ fuer $f_\pm \in L^1$ gilt nun offenbar

$$-(f_-)^* = (-f_-)_* \leq f_* \leq f^* \leq (f_+)^*.$$

Da nach dem schon gezeigten $(f_-)^*$ und $(f_+)^*$ zu L^1 gehoeren, folgt dann $f^* \in L^1$. Nach dieser Zwischenueberlegung koennen wir nun in $(*)$ Summieren undd die Disjunktheit der D_k^n, $k \in \mathbb{Z}$, (bei festem n) nutzen, um

$$\int_X f^* dm \leq \frac{m(X)}{n} + \int_X f dm$$

zu erhalten. Da $n \in \mathbb{N}$ beliebig war, folgt

$$\int_X f^* dm \leq \int_X f dm.$$

Ersetzen von f durch $-f$ und Nutzen von $(-f)^* = -f_*$ liefert dann

$$\int_X f_* dm \geq \int_X f dm.$$

Das beendet den Beweis. □

Falls die Ergodizitaetsvoraussetzung gilt, erhalten wir aus dem Birkhoffschen Theorem eine starke Version des in der Einleitung als gewuenscht diskutierten.

Folgerung. *Sei die Situation wie im vorigen Satz. Es gelte noch $m(X) = 1$ und die Ergodizitaetsvoraussetzung (E). Dann ist f^* fast sicher konstant und zwar gleich $\int_X f^* dm = \int_X f dm$ und es folgt aus dem Birkhoffschen Ergodensatz*

$$\frac{1}{N}\sum_{k=0}^{N-1} f(T^k x) \longrightarrow \int_X f dm$$

fuer fast alle $x \in X$.

Tatsaechlich gelten die Aussagen des Birkhoffschen Ergodensatzes (bis auf die Gleichheit von $\int f dm$ mit $\int f^* dm$) auch fuer σ-endliche Maßraeume. Das diskutieren wir nun kurz.

Theorem (Birkhoffscher Ergodensatz - σ-endliche Maßraeume). *Sei $(X, \mathcal{B}, m)$ ein σ- endlicher Maßraum und $T : X \longrightarrow X$ maßerhaltend. Dann konvergiert fuer jedes $f \in L^1(X, m)$ die Folge der Mittel*

$$\frac{1}{N}\sum_{k=0}^{N-1} f \circ T^k$$

punktweise fast ueberall gegen eine Funktion $f^* \in L^1(X, m)$ mit $f^* = f^* \circ T$.

Beweis. Es reicht reellwertige f zu betrachten. Wir definieren

$$f^*(x) = \limsup_{N\to\infty} \frac{1}{N}\sum_{k=0}^{N-1} f(T^k x)$$

und

$$f_*(x) = \liminf_{N\to\infty} \frac{1}{N}\sum_{k=0}^{N-1} f(T^k x).$$

Dann gilt offenbar

$$f^* = f^* \circ T, \;\; f_* = f_* \circ T$$

und $f_* \leq f^*$.

Wir zeigen zunaechst $f^* = f_*$. (Das liefert dann sofort die punktweise fast sichere Konvergenz.)

Setze fuer $\beta, \alpha \in \mathbb{R}$

$$E_{\alpha,\beta} := \{x \in X : f_*(x) < \beta, \alpha < f^*(x)\}.$$

Wegen

$$\{x \in X : f_*(x) \neq f^*\} = \bigcup_{\beta<\alpha,\beta,\alpha\in\mathbb{Q}} E_{\alpha,\beta}$$

reicht es zu zeigen $m(E_{\alpha,\beta}) = 0$ fuer alle $\beta < \alpha$. Dem werden wir uns nun widmen. Wir gehen in zwei Schritten vor.

Jedes $E_{\alpha,\beta}$ hat endliches Maß.

Bew. (Uebung ;-)

Jedes $E_{\alpha,\beta}$ hat Maß 0.

Bew. Offenbar gilt

$$T^{-1}E_{\alpha,\beta} = E_{\alpha,\beta}, \; B_\alpha \cap E_{\alpha,\beta} = E_{\alpha,\beta}$$

mit $B_\alpha = \{x : \sup_N \frac{1}{N}\sum_{k=0}^{N-1} f(T^k x) > \alpha\}$. (Erste Gleichung folgt aus der T-invarianz von f_* und f^*. Die zweite Gleichung folgt aus $\limsup a_N \leq \sup a_N$.) Da jedes $E_{\alpha,\beta}$ endliche Maß hat, koennen wir die vorige Folgerung anwenden und erhalten

$$\int_{E_{\alpha,\beta}} f dm = \int_{E_{\alpha,\beta}\cap B_\alpha} f dm \geq \alpha m(E_{\alpha,\beta}\cap B_\alpha) = \alpha m(E_{\alpha,\beta}).$$

Damit folgt dann

$$\int_{E_{\alpha,\beta}} f dm \geq \alpha m(E_{\alpha,\beta}).$$

Ersetzt man nun f durch $-f$, α durch $-\alpha$, β durch $-\beta$, so erhaelt man wegen $(-f)^* = -f_*$ und $(-f)_* = -f^*$ genau die gleichen Mengen $E_{\alpha,\beta}$ sowie

$$\int_{E_{\alpha,\beta}} f dm \leq \beta m(E_{\alpha,\beta}).$$

Insgesamt folgt also

$$\alpha m(E_{\alpha,\beta}) \leq \int_{E_{\alpha,\beta}} f dm \leq \beta m(E_{\alpha,\beta}).$$

Mit $\beta < \alpha$ folgt nun das gewuenschte

$$m(E_{\alpha,\beta}) = 0.$$

Wir zeigen nun $f^* \in L^1(X, m)$.
Das folgt aus dem Lemma von Fatou auf folgende Weise.

$$\begin{aligned}
\int_X |f^*| dm &= \int_X \left| \lim_{N\to\infty} \frac{1}{N}\sum_{k=0}^{N-1} f(T^k x)\right| dm \\
(Fatou) \quad &\leq \liminf_{N\to\infty} \int_X \left|\frac{1}{N}\sum_{k=0}^{N-1} f(T^k x)\right| dm \\
(Dreiecksungl) \quad &\leq \|f\|_1 < \infty.
\end{aligned}$$

Das beendet den Beweis. □

2. Der Ergodensatz in L^p

Wir ziehen nun als eine erste Folgerungen aus dem Birkhoffschen Ergodensatz eine Variante des von Neumannschen Ergodensatz fuer L^p. Das liefert insbesondere die Konvergenz in L^1 im Birkhoffschen Ergodesatz. Die Konvergenz in L^1 im Birkhoffschen Ergodensatz ist bemerkenswert. Der Satz selber sagt ja zunaechst lediglich etwas zu punktweiser Konvergenz aus. Als Voraussetzung fuer die L^p Konvergenz brauchen wir, daß das Gesamtmaß endlich ist.

Theorem (L^p - Ergodensatz). *Sei $1 \leq p < \infty$. Sei T eine maßerhaltende Transformation des Wahrscheinlichkeitsraumes $(X, \mathcal{B}, m)$. Dann existiert zu jedem $f \in L^p(X, m)$ ein $f^* \in L^p(X, m)$ mit den folgenden drei Eigenschaften:*

- $f^* = f^* \circ T$.
- $\frac{1}{N}\sum_{k=0}^{N-1} f(T^k x) \to f^*(x)$ *fuer fast alle* $x \in X$.

- $\|\frac{1}{N}\sum_{k=0}^{N-1} f(T^k x) - f\|_p \to 0, N \to \infty.$

Bemerkung. Der Satz liefert insbesondere die Konvergenz in L^1 im Birkhoffschen Ergodensatz.

Beweis. Wegen $m(X) < \infty$ gilt $L^p(X,m) \subset L^1(X,m)$. Damit folgt aus dem Birkhoffschen Ergodensatz direkt die Existenz eines f^*, das die ersten beiden Punkte erfuellt. Es bleibt die Konvergenz in L^p gegen dieses f^* zu zeigen. Dazu reicht es zu zeigen, daß die Folge der Mittel

$$A_N f := \frac{1}{N}\sum_{k=0}^{N-1} f \circ T^k$$

ueberhaupt in L^p konvergiert. (Denn dann hat sie eine Teilfolge, die fast ueberall konvergiert, und aufgrund der schon gezeigten punktweise fast sicheren Konvergenz gegen $f*$ muss dann f^* der Grenzwert sein.)
Offenbar ist jedes A_N ein linearer beschraenkter Operator auf $L^p(X,m)$ mit $\|A_N\| \leq 1$. Daher reicht es die Konvergenz von $(A_N f)_N$ fuer alle f aus einer dichten Teilmenge von L^p zu zeigen. Als diese Teilmengen kann $L^p(X,m) \cap L^\infty(X,m)$ gewaehlt werden. (Denn diese Teilmenge ist offenbar dicht und auf ihr gilt die gewuenschte Konvergenz aufgrund des Satzes von der dominierten Konvergenz.) □

3. Zwei Charakterisierungen von Ergodizitaet

Als naechstes koennen wir den Birkhoffschen Satz nutzen, um die Ergodizitaetsbedingung zu charakterisieren. Wir werden hier zwei Charakterisierungen kennenlernen: eine Charakterisierung via Existenz von Frequenzen und eine Charakterisierung via Ausschöpfen. Beiden Charakterisierungen gemeinsam ist, dass sie eine praezise Fassung davon geben, dass die Iteration der Transformation den *Raum gleichmäßig ausschöpft.* Das ist dann eine gültige Fassung der im allgemeinen nicht gueltigen Formulierung $\{T^n x : n \in \mathbb{N}_0\} = X$.

Erinnerung. Ein meßbares dynamisches System heißt ergodisch, wenn jede meßbare Funktion $f : X \longrightarrow \mathbb{R}$ mit $f = f \circ T$ fast sicher konstant ist.

Lemma (Charakterisierung via Frequenzen bzw. Auschoepfen). *Sei $(X, \mathcal{B}, m)$ ein Maßraum mit $m(X) = 1$ und $T : X \longrightarrow X$ eine maßerhaltende Transformation. Dann sind aequivalent:*

(i) *Es ist T ergodisch (d.h. jede meßbare invariante Funktion f ist fast sicher konstant).*
(ii) *Es existieren die Frequenzen der meßbaren Mengen d.h. fuer jede meßbare Menge E gilt*

$$\frac{1}{N}\sharp\{k \in \{0, \ldots, N-1\} : T^k x \in E\} \to m(E)$$

fuer fast alle $x \in X$.

(iii) *Die Urbilder jeder meßbaren Menge A schoepfen den Raum gleichmaeßig aus in dem Sinne, daß fuer jede meßbare Menge B gilt*

$$\frac{1}{N}\sum_{k=0}^{N-1} m(T^{-k}A\cap B)\to m(A)m(B).$$

Beweis. (i)$\Longrightarrow$ (ii): Mit der charakteristischen Funktion 1_E von E gilt

$$\frac{1}{N}\sharp\{k\in\{0,\dots,N-1\}:T^kx\in E\}=\sum_{k=0}^{N-1}1_E(T^kx).$$

Damit folgt die gewuenschte Aussage direkt aus dem Birkhoffschen Theorem bzw. seiner Folgerung fuer den ergodischen Fall.

(ii)$\Longrightarrow$ (iii): Seien 1_A und 1_B die charakteristischen Funktionen von A bzw. B. Dann gilt nach (ii) also

$$1_B(x)\frac{1}{N}\sum_{k=0}^{N-1}1_A(T^kx)\to m(A)1_B(x)$$

fuer fast alle $x\in X$. Da die linken Seiten offenbar gleichmaeßig beschraenkt sind, folgt nach Integration dann

$$\frac{1}{N}\sum_{k=0}^{N-1}m(T^{-k}A\cap B)=\int_X 1_B(x)\frac{1}{N}\sum_{k=0}^{N-1}1_A(T^kx)dm\to\int_X m(A)1_Bdm=m(A)m(B).$$

(iii)$\Longrightarrow$ (i): Wir zeigen zunaechst, dass jedes meßbare E mit $T^{-1}E=E$ Maß 0 oder 1 hat. Setze $A=B=E$ in (iii). Dann folgt

$$m(E)=\frac{1}{N}\sum_{k=0}^{N}m(T^{-k}E\cap E)\to m(E)m(E).$$

Damit folgt $m(E)=m(E)^2$ und daher $m(E)\in\{0,1\}$.

Sei nun f eine meßbare Funktion mit $f=f\circ T$. Waere f nicht konstant fast sicher, so gaebe es ein $\alpha\in\mathbb{R}$ (z.B. $\alpha=\int_X fdm$) so daß sowohl die Menge

$$A_\alpha=\{x:f(x)>\alpha\}$$

als auch die Menge

$$B_\alpha=\{x:f(x)<\alpha\}$$

positives Maß haben. Da die Mengen komplementaer sind, hat dann keine von ihnen das Maß 1. Da diese beiden Mengen invariant und meßbar sind, muessen sie aber nach dem schon gezeigten beide Maß in $\{0,1\}$ haben. Das ist ein Widerspruch. $\square$

KAPITEL 4

Ergodizitaet: Charakterisierungen und Beispiele

In diesem Kapitel untersuchen wir das grundlegendes Konzept der Ergodizitaet einer maßerhaltenden Transformation. Dabei werden wir uns auf Raeume mit endlichem Maß einschraenken. Gueltigkeit von Ergodizitaet ist aequivalent zur Konvergenz gegen eine konstante Funktion im von Neumannscher Ergodensatz. Tatsaechlich hat Ergodizitaet noch viele weitere Konsequenzen.

1. Die Ergodizitaet einer Transformation

Ist T eine meßbare Transformation auf einem Maßraum $(X, \mathcal{B}, m)$, so heißt eine meßbare Teilmenge B von X invariant unter T (kurz: T-invariant), wenn gilt $T^{-1}B = B$. Ist B eine meßbare T-invariante Menge, so ist dann auch $X \setminus B$ eine meßbare invariante Menge. Damit wird durch ein solches B dann die Transformation T 'zerlegt' in zwei Teile

$$T_B : B \longrightarrow B, x \mapsto Tx, \text{ und } T_{X \setminus B} : X \setminus B \longrightarrow X \setminus B, x \mapsto Tx,$$

in dem Sinne, daß sowohl T_B als auch $T_{X \setminus B}$ maßerhaltende Transformationen sind. Uns wird es um den Fall gehen, daß es solche Zerlegungen eigentlich nicht gibt.

Definition (Zerlegbarkeit). *Sei $(X, \mathcal{B}, m)$ ein beliebiger Maßraum. Eine maßerhaltende Transformation $T : X \longrightarrow X$ heißt zerlegbar, wenn es eine meßbare Teilmenge $B \subset X$ mit $T^{-1}B = B$ und sowohl $0 < m(B)$ als auch $0 < m(X \setminus B)$ gibt. Andernfalls heißt die Transformation unzerlegbar.*

Bemerkungen. Ist m ein Wahrscheinlichkeitsmaß auf X, so bedeutet Unzerlegbarkeit gerade $m(B) \in \{0, 1\}$ fuer alle meßbaren T-invarianten Mengen B.

Ende der 5. Vorlesung.

!!! Ein Teil der folgenden Betrachtungen gilt fuer endliche und unendliche Maßraeume gleichermaßen. Wir beschraenken uns hier aber auf den Fall endlicher Maßraeume, da unsere Hauptanwendungen in diese Klasse fallen.

Theorem (Charakterisierung Zerlegbarkeit). *Sei $(X, \mathcal{B}, m)$ ein Maßraum mit $m(X) = 1$ und $T : X \longrightarrow X$ eine maßerhaltende Transformation. Dann sind die folgenden Aussagen aequivalent:*

(i) *T ist unzerlegbar (d.h. jedes meßbare $B \subset X$ mit $T^{-1}B = B$ erfüllt $m(B) \in \{0, 1\}$).*
(ii) *Die meßbaren Mengen B mit $m(B \triangle T^{-1}B) = 0$ erfüllen $m(B) \in \{0, 1\}$.*
(iii) *Fuer jedes meßbare A mit $m(A) > 0$ gilt $m(X \setminus \bigcup_{n=0}^{\infty} T^{-n}A) = 0$.*

(iv) *Fuer alle meßbaren A und B mit $m(A), m(B) > 0$ gibt es ein $N \in \mathbb{N}$ mit $m(T^{-N}A \cap B) > 0$.*

Bemerkungen.

- Bedingungen (iii) und (iv) lieferen 'aufgedickte' Varianten von Boltzmann's Idee $\overline{\{T^n x : n \in \mathbb{N}_0\}} = X$.
- In (iii) und (iv) kann man A (oder B) ersetzen durch $T^{-k}A$ (bzw $T^{-k}B$).

Beweis. (i)$\Longrightarrow$ (ii): Sei B mit $m(B \triangle T{-}1B) = 0$ gegeben. Es reicht ein $\widetilde{B}$ zu finden mit folgenden beiden Eigenschaften:

- $T^{-1}\widetilde{B} = \widetilde{B}$.
- $m(B \triangle \widetilde{B}) = 0$.

Setze

$$\widetilde{B} := \bigcap_{n=0}^{\infty} \bigcup_{j=n}^{\infty} T^{-j}B = \{x : T^j x \in B \text{ fuer unendliche viele } j\}.$$

Dann gilt offenbar

$$T^{-1}\widetilde{B} = \{x : T^j(Tx) \in B \text{ fuer unendlich viele } j\} = \widetilde{B}.$$

Weiterhin gilt fuer jedes $k \in \mathbb{N}_0$ offenbar

$$m(B \triangle T^{-k}B) \leq \sum_{j=0}^{k-1} m(T^{-j}B \triangle T^{-j-1}B) = \sum_{j=0}^{k-1} m(T^{-j}(B \triangle T^{-1}B)) = km(B \triangle T^{-1}B) = 0.$$

Damit folgt dann

$$\begin{aligned} m(B \triangle \widetilde{B}) &= \lim_{n\to\infty} m(B \triangle (\bigcup_{j=n}^{\infty} T^{-j}B)) \\ &\leq \liminf_{n\to\infty} m(\bigcup_{j=n}^{\infty} B \triangle T^j B) \\ &\leq \liminf_{n\to\infty} \sum_{j=n}^{\infty} m(B \triangle T^{-j}B) \\ &= 0. \end{aligned}$$

(ii)$\Longrightarrow$ (iii): Setze

$$B := X \setminus \bigcup_{n=0}^{\infty} T^{-n}A = \bigcap_{n=0}^{\infty} X \setminus T^{-n}A = \bigcap_{n=0}^{\infty} T^{-n}(X \setminus A).$$

Damit folgt

$$T^{-1}B = \bigcap_{j=1}^{\infty} T^{-j}(X \setminus A) \supset B.$$

Weiterhin gilt - da T maßerhaltend ist - auch

$$m(T^{-1}B) = m(B).$$

Insgesamt ergibt sich also

$$m(B \triangle T^{-1}B) = 0.$$

Aus (ii) folgt dann $m(B) = 0$ oder $m(X \setminus B) = 0$. Mit
$$m(X \setminus B) \geq m(A) > 0$$
folgt dann $m(B) = 0$. Das ist die gewuenschte Aussage (iii).

(iii)$\Longrightarrow$ (iv): Seien meßbare A und B mit $m(A) > 0$ gegeben und $m(T^{-n}A \cap B) = 0$ fuer alle $n \in \mathbb{N}_0$ gegeben.
Dann gilt
$$m(\left(\bigcup_{n=0}^{\infty} T^{-n}A\right) \cap B) = 0.$$
Mit (iii) folgt dann sofort $m(B) = 0$. Das liefert die gewuenschte Behauptung.

(iv)$\Longrightarrow$ (i): Sei B eine meßbare Menge mit $B = T^{-1}B$. Dann folgt offenbar
$$T^{-n}B = B$$
fuer alle $n \in \mathbb{N}_0$. Damit folgt dann
$$m((X \setminus B) \cap T^{-n}B) = m((X \setminus B) \cap B) = 0$$
fuer alle $n \in \mathbb{N}_0$. Aus (iv) folgt nun $m(B) = 0$ oder $m(X \setminus B) = 0$. □

Theorem (Aequivalenz von Zerlegbarkeit und Ergodizitaet). *Sei* $(X, \mathcal{B}, m)$ *ein Maßraum mit* $m(X) = 1$ *und* $T : X \longrightarrow X$ *eine maßerhaltende Transformation. Dann sind die folgenden Aussagen aequivalent:*

(i) T *ist unzerlegbar (d.h. die meßbaren Mengen* B *mit* $T^{-1}B = B$ *erfüllen* $m(B) \in \{0, 1\}$*).*
(ii) T *ist ergodisch (d.h. jedes meßbare* $f : X \longrightarrow \mathbb{R}$ *mit* $f = f \circ T$ *ist fast ueberall konstant.)*

Beweis. (ii)$\Longrightarrow$ (i): Ist B eine invariante Menge, so gilt fuer die charakteristische Funktion 1_B von B offenbar $1_B = 1_B \circ T$. Damit folgt (i) einfach aus (ii).

(i)$\Longrightarrow$ (ii): Sei $f : X \longrightarrow \mathbb{R}$ invariant. Definiere fuer $k \in \mathbb{Z}$ und $n > 0$
$$X(k, n) := \{x \in X : \frac{k}{2^n} \leq f(x) < \frac{k+1}{2^n}\}.$$
Dann gilt
$$T^{-1}X(k, n) \triangle X(k, n) \subset N.$$
Damit folgt aus (i) dann also $m(X(k, n)) = 0$ oder $m(X(k, n)) = 1$. Weiterhin gilt offenbar
$$X = \bigcup_k X(k, n),$$
wobei die Vereinigung disjunkt ist. Daher existiert also ein eindeutiges k_n mit
$$m(X(k_n, n)) = 1.$$
Setze
$$Y := \bigcap_n X(k_n, n).$$
Dann hat Y volles Maß und f ist auf Y konstant. □

2. Die Menge der ergodischen Maße

In diesem Abschnitt studieren wir die Struktur der Menge der ergodischen Wahrscheinlichkeitsmaße bezueglich einer gegebenen meßbaren Transformation.

Definition. *Sei $(X, \mathcal{B})$ ein meßbarer Raum und $T : X \longrightarrow X$ eine meßbare Transformation. Dann heißt ein Maß m auf X ergodisch (bzgl. T), wenn T eine ergodische Transformation auf $(X, \mathcal{B}, m)$ ist.*

Lemma. *Sei $(X, \mathcal{B})$ ein meßbarer Raum und $T : X \longrightarrow X$ meßbar. Seine m_1, m_2 T-invariante Wahrscheinlichkeitsmaße auf X. Ist m_1 absolut stetig bezueglich m_2 und m_2 ergodisch bzgl T, so gilt $m_1 = m_2$.*

Bemerkung. Fuer invertierbare T ist der Beweis der Aussage eine recht einfache Uebung.

Beweis. Da m_1 absolut stetig bezueglich m_2 ist existiert eine $f \in \mathcal{L}^1(X, m_2)$ mit

$$m_1(A) = \int 1_A(x) f(x) dm_2(x)$$

fuer alle meßbaren Mengen A. Setze

$$C_- := \{x \in X : f(x) < 1\}, \ C_+ := \{x \in X : f(x) > 1\}.$$

Ziel. Zeige $m_2(C_+) = m_2(C_-) = 0$. (Dann gilt also $f = 1$ fast ueberall und $m_1 = m_2$ folgt).

Wir zeigen nur die Aussage fuer C_-. Die Aussage fuer $C_+ 0$ folgt analog. Aufgrund der Ergodizitaet von m_2 reicht es zu zeigen:

- $m_2(C_- \triangle T^{-1} C_-) = 0$
- $m_2(C_-) \neq 1$.

Wir zeigen zunaechst die zweite Aussage: Angenommen $m_2(C_-) = 1$. Dann ist also $X \setminus C_-$ eine Nullmenge bezueglich m_2. Weiterhin gilt

$$m_1(C_-) = \int_{C_-} f(x) dm_2(x) < \int_{C_-} dm_2(x) = 1.$$

Damit ist also $X \setminus C_-$ eine Menge von positivem Maß bzgl. m_1. Insgesamt ergibt sich ein Widerspruch dazu, daß m_1 absolutstetig bzgl. m_2 ist.

Wir zeigen nun die erste Aussage: Dazu zunaechst eine Hilfsaussage: Ist T eine maßerhaltende Transformation bzgl. eines Maßes m so gilt

$$(*) \ \ m(A \setminus T^{-1} A) = m(T^{-1} A \setminus A)$$

fuer alle meßbaren Mengen A.
(Bew. Es gilt

$$A = (A \cap T^{-1} A) \cup (A \setminus T^{-1} A)$$

und

$$T^{-1} A = (T^{-1} A \cap A) \cup (T^{-1} A \setminus A).$$

Mit $m(A) = m(T^{-1} A)$ folgt dann die Aussage.)

Angenommen: $m_2(C_- \setminus T^{-1} C_-) > 0$.

Dann liefern kurze Rechnungen

$$m_1(C_- \setminus T^{-1}C_-) < m_2(C_- \setminus T^{-1}C_-)$$

und

$$m_1(T^{-1}C_- \setminus C_-) \geq m_2(T^{-1}C_- \setminus C_-).$$

Mit $(*)$ ergibt sich sofort ein Widerspruch. Damit folgt also $m_2(C_- \setminus T^{-1}C_-) = 0$. Aus $(*)$ folgt dann auch $m_2(T^{-1}C_- \setminus C_-) = 0$. Das liefert die gewuenschte Aussage. □

Theorem (Charakterisierung ergodischer Maße innerhalb der invarianten Maße). *Sei $(X, \mathcal{B})$ ein meßbarer Raum und $T : X \longrightarrow X$ eine meßbare Transformation. Sei*

$$\mathcal{M}^1(X,T) := \{m : \ m \textit{ ist } T\textit{-invariantes Wahrscheinlichkeitsmaß}\}.$$

Dann ist $\mathcal{M}^1(X,T)$ konvex und die Extrempunkte von $\mathcal{M}^1(X,T)$ sind gerade die ergodischen Maße.

Erinnerung. Ein Element c einer konvexen Menge C heißt Extrempunkt, wenn aus $c = \lambda q + (1-\lambda)p$ mit $p, q \in C$ und $\lambda \in (0,1)$ folgt $q = p = c$.

Ende der 6. Vorlesung

Beweis. Konvexitaet von $\mathcal{M}^1(X,T)$ ist klar. Es bleibt die Charakterisierung der ergodischen Maße als Extrempunkte zu zeigen. Es sind zwei Richtungen zu zeigen.

Sei m nicht ergodisch. Zu zeigen: m ist kein Extrempunkt.

Da m nicht ergodisch ist, gibt es ein meßbares T-invariantes B mit $0 < m(B) < 1$. Dann gilt

$$m = m(B)\left(\frac{1}{m(B)} m|_B\right) + m(X \setminus B)\left(\frac{1}{m(X \setminus B)} m|_{X\setminus B}\right).$$

Damit ist also m eine nichttriviale Linearkombination von zwei verschiedenen (!) T-invarianten Maßen.

Sei m ein ergodisches Wahrscheinlichkeitsmaß. Zu zeigen: m ist ein Extrempunkt.

Sei

$$m = \lambda m_1 + (1-\lambda) m_2$$

mit T-invarianten Wahrscheinlichkeitsmaßen m_1 und m_2 und $0 < \lambda < 1$. Dann sind also sowohl m_1 als auch m_2 absolut stetig bezueglich m. Damit stimmen sie nach dem vorigen Lemma aber mit m ueberein. Damit ist also m ein Extrempunkt. □

Theorem (Gegenseitige Singularitaet ergodischer Maße). *Sei $(X, \mathcal{B})$ ein meßbarer Raum und $T : X \longrightarrow X$ eine meßbare Transformation. Seien m, n ergodische Wahrscheinlichkeitsmaße bezueglich T. Gilt $m \neq n$, so sind m und n gegenseitig singulaer.*

Beweis. Nach dem Lebesgueschen Zerlegungssatz (angewendet auf n bezueglich des Maßes m) kann man n schreiben als

$$n = p\mu + (1-p)\nu$$

mit Wahrscheinlichkeitsmaßen μ, ν mit

$$\mu \prec m, \;\; \nu \perp m.$$

Da μ und ν Wahrscheinlichkeitsmaße sind mit $\mu \prec m$ und $\nu \perp m$ sind μ und ν verschieden. Da die Zerlegung eindeutig ist und n aber T-invariant ist, ergibt sich, daß auch μ und ν invariant unter T sind. Da n als ergodisches Maß ein Extrempunkt ist und $\nu \neq \mu$ folgt also $p = 1$ oder $p = 0$.
Im Fall $p = 1$ waere aber $n = \mu$ und, weiterhin, ist nach dem Lemma aber $\mu = m$. Damit ergaebe sich der Widerspruch $n = m$. Es bleibt also nur der Fall $p = 0$. Dann ist aber $n = \nu$ singulaer bezueglich m. □

3. Beispiele fuer ergodische Systeme

In diesem Abschnitt untersuchen wir die oben angegebenen Beispiele auf Ergodizitaet.

3.1. Die Identitaet. Sei $(X, \mathcal{B}, m)$ ein beliebiger Maßraum mit $m(X) = 1$. Sei $T : X \longrightarrow X$ die Identitaet. Dann ist T im allgemeinen nicht ergodisch. Genauer ist T genau dann ergodisch wenn es keine meßbare Menge B mit $0 < m(B) < 1$ gibt.

3.2. Rotation auf dem Einheitskreis. Sei $\mathbb{T} := \mathbb{R}/\mathbb{Z}$ mit dem Lebesguemass λ und sei fuer $\alpha \in RR$ die Transformation T_α definiert durch

$$T_\alpha : \mathbb{T} \longrightarrow \mathbb{T}, T_\alpha(x + \mathbb{Z}) = \alpha + x + \mathbb{Z}.$$

Theorem. *Fuer $\alpha \in \mathbb{R}$ sind die folgenden Aussagen aequivalent:*
- *Es ist T_α ergodisch.*
- *Es ist α irrational.*
- *Es ist der Orbit $O(0) := \{T_\alpha^n 0 : n \in \mathbb{Z}\}$ von 0 dicht.*

Beweis. Ist $n \in \mathbb{Z}$ beliebig und $x \in \mathbb{T}$ so haengt

$$\exp(2\pi i n x^*)$$

nicht vom Representanten x^* von x ab. Wir schreiben dann $\exp(inx)$ fuer diesen Wert.

(i)$\Longrightarrow$ (ii): Angenommen $\alpha = p/q$ mit $p \in \mathbb{Z}$ und $q \in \mathbb{N}$. Betrachte die (wohldefinierte) Funktion $\mathbb{T} \longrightarrow S^1, x \mapsto \exp(2\pi i p x)$. Diese Funktion ist offenbar messbar, nicht konstant und invariant unter T_α. Das ist ein Widerspruch zur Ergodizitaet (i).

(ii)$\Longrightarrow$ (iii): Es reicht zu zeigen, daß der Orbit Elemente enthaelt die nicht 0 sind aber beliebig nahe an 0 sind. Dazu reicht es zu zeigen, daß der Orbit ueberhaupt Elemente enthaelt die beliebig nahe aneinander sind (ohne gleich sein). Das folgt sofort, da - aufgrund der Irrationalitaet von α - alle Elemente $T_\alpha^n 0$ verschieden sind.

(iii)$\Longrightarrow$ (i): Wir geben zwei Beweise:

Erster Beweis: Ist das Maß m ein unter T_α invariantes nichttriviales Maß, so ist m unter allen Verschiebungen T_β fuer $\beta \in O(0)$ invariant. Damit ist m dann unter allen Elementen von U invariant. Damit ist m das Haarmaß. Ist nun A eine T-invariante Menge mit positivem Maß, so ist die Einschraenkung m des Haarmaßes auf A ein T_α invariantes nichttriviales Maß. Damit muss es sich um das Haarmaß handeln. Damit ist dann A bis auf eine Nullmenge der gesamte Raum.

Zweiter Beweis: Sei $f \in L^2(\mathbb{T}, m)$ mit $f = f \circ T$ gebeben. Betrachte die Fouriertransformation

$$F(f)(n) = \int_{\mathbb{T}} \exp(2\pi inx) f(x) dm$$

von f.
Dann gilt aufgrund der Invarianz von f also

$$\begin{aligned} F(f)(n) &= F(f \circ T_\alpha)(n) \\ &= \int_{\mathbb{T}} \exp(2\pi inx) f(\alpha + x) dm \\ &= \int_{\mathbb{T}} \exp(2\pi in(x-\alpha)) f(x) dm \\ &= \exp(-2\pi in\alpha) \int_{\mathbb{T}} \exp(2\pi inx) f(x) dm \\ &= \exp(-2\pi in\alpha) F(f)(n). \end{aligned}$$

Aufgrund der Dichtheit des Orbits (iii), muss aber nun gelten $\exp(-2\pi in\alpha) \neq 1$ fuer alle $n \neq 0$ (sonst 'Orbit periodisch'...). Damit folgt aus der Invarianz von f also

$$Ff(n) = 0$$

fuer alle $n \neq 0$. Damit ist f dann konstant. □

Das Beispiel kann verallgemeinert werden wie folgt. Sei G eine beliebige kompakte Gruppe zusammen mit der Borel-σ-Algebra und dem Haarmaß. Sei $a \in G$ beliebig. Dann ist $T = T_a$ mit

$$T : G \longrightarrow G, Tx := ax$$

eine meßbare Transformation. Tatsaechlich ist T invertierbar mit inverser Transformation gegeben durch $T_{a^{-1}}$.

Theorem. *Sei G eine kompakte Gruppe mit Haarmaß λ und $T = T_a$ die Translation um $a \in G$. Dann ist T_a genau dann ergodisch, wenn*

$$O(e) := \{a^n : n \in \mathbb{Z}\}$$

dicht in G ist. Insbesondere ist dann also G abelsch.

Beweis. Sei $U := \overline{O(e)}$. Dann ist U eine abgeschlossene Untergruppe von G. Wir zeigen zwei Richtungen:

Sei $O(e)$ dicht in G. Dann ist also $U = G$ und da $O(e)$ offenbar abelsch ist, ist auch G eine abelsche Gruppe. Nun Nun kann man den zweiten obigen Beweis von (iii)$\Longrightarrow$ (ii) leicht adaptieren.

Sei T ergodisch. Es ist U eine T_a invariante Menge. Damit hat dann also U das Maß 0 oder das Maß 1. Hat U das Maß 1, so ist offenbar U dicht in G und stimmt damit als abgeschlossene Menge mit G ueberein.

□

3.3. Homomorphismen auf dem Einheitskreis. Definiere fuer $p \in \mathbb{Z}$ die Abbildung

$$T : \mathbb{T} \longrightarrow \mathbb{T}, T(x + \mathbb{Z}) = px + \mathbb{Z}.$$

Dann ist T eine meßbare Transformation. Eine alternative Darstellung ist gegeben durch

$$T : S^1 \longrightarrow S^1, Tz = z^p.$$

Theorem. *Fuer jedes $p \in \mathbb{Z}$ mit $p \notin \{-1, 0, 1\}$ ist T ergodisch.*

Beweis. Sei $f \in L^2(\mathbb{T}, m)$ mit $f = f \circ T$ gebeben. Betrachte die Fouriertransformation

$$F(f)(n) = \int_{\mathbb{T}} \exp(2\pi inx) f(x) dm$$

von f. Dann gilt aufgrund der Invarianz von f also

$$\begin{array}{rcl} F(f)(pn) & = & F(f \circ T)(pn) \\ & = & \int_{\mathbb{T}} \exp(2\pi ipnx) f(px) dm \\ & = & \int_{\mathbb{T}} \exp(2\pi in\cdot)(px) f(px) dm \\ (m \;\; inv) & = & \int_{\mathbb{T}} \exp(2\pi inx) f(x) dm \\ & = & F(f)(n). \end{array}$$

Damit folgt dann also

$$F(f)(n) = F(f)(pn) = F(f)(p^2 n) = ...$$

Mit

$$\sum_{n \in \mathbb{Z}} |F(f)(n)|^2 < \infty$$

ergibt sich dann $F(f)(n) = 0$ fuer alle $n \neq 0$. Damit ist f konstant. □

Das Beispiel kann verallgemeinert werden wie folgt. Sei G eine beliebige kompakte Gruppe zusammen mit der Borel-σ-Algebra und dem Haarmaß. Sei $A : G \longrightarrow G$ ein stetiger surjektiver Gruppenhomomorphismus. Dann ist T gegeben durch

$$T : G \longrightarrow G, Tx = A(x)$$

eine maßerhaltende Transformation. (Bew. Definiere das Maß μ durch $\mu(B) := \lambda(T^{-1}B)$. Zeige, daß μ normiert und linksinvariant ist.)

Ende der 7. Vorlesung

3.4. Der Bernoulli und der $(p_1,\ldots,p_k)$ Shift. . Seien $k \geq 2$ und $p_1,\ldots,p_k > 0$ mit $\sum_{j=1}^k p_j = 1$ gegeben. Sei $Y := \{0,\ldots,k-1\}$ mit der σ-Algebra aller meßbaren Teilmengen ausgestattet und sei μ auf Y das eindeutige Maß mit

$$\mu(\{j\}) = p_{j+1}$$

fuer $j = 0,\ldots,k-1$. Sei $X := Y^{\mathbb{Z}}$ mit der Produkt-σ-Algebra und dem Produktmaß. Sei

$$T : X \longrightarrow X, (Tx)(n) = x(n+1).$$

Dann ist T eine invertierbare maßerhaltende Transformation.

Theorem. *Der $(p_1,\ldots,p_k)$ Shift ist ergodisch.*

Beweis. Sei B eine meßbare invariante Menge.

Idee. Gehoert B zu den Rechtecken, so gibt es aufgrund der Shift-Wirkung ein $N \in \mathbb{N}$, so daß B und $T^{-N}B$ 'unabhaengig' sind. Dann gilt also

$$m(B) = m(T^{-N}B \cap B) = m(T^{-N}B)m(B) = m(B)^2.$$

Das impliziert $m(B) \in \{0,1\}$.

Wir nutzen nun folgenden Sachverhalt: Ist $(X,\mathcal{B},m)$ ein Maßraum und $\mathcal{R}$ eine Algebra, die $\mathcal{B}$ erzeugt, so gibt es fuer jedes $\varepsilon > 0$ und jedes $B \in \mathcal{B}$ ein $A \in \mathcal{R}$ mit $m(A \triangle B) < \varepsilon$.

Waehle nun zu dem invarianten B und zu $\varepsilon > 0$ ein Menge A aus den Rechtecken mit $m(A\triangle B) < \varepsilon$. Waehle nun ein $N \in \mathbb{N}$, so daß $T^{-N}A$ und A von verschiedenen Koordinaten abhaengen. Dann gilt also

$$(*) \;\; m(A \cap T^{-N}A) = m(A)m(T^{-N}A) = m(A)^2.$$

Weiterhin gilt

$$\begin{aligned}(**) \;\; m(B\triangle(A\cap T^{-N}A)) &\leq m(B\triangle A \cup B\triangle T^{-N}A)\\ &\leq m(B\triangle A) + m(T^{-N}B\triangle T^{-N}A)\\ &= 2m(B\triangle A)\\ &< 2\varepsilon.\end{aligned}$$

Damit folgt dann

$$\begin{aligned}|m(B) - m(B)^2| &\leq |m(B) - m(A\cap T^{-N}A)| + |m(A\cap T^{-N}A) - m(B)^2|\\ (*) \quad &= |m(B) - m(A\cap T^{-N}A)| + |m(A)^2 - m(B)^2|\\ (**) \quad &\leq 2\varepsilon + |m(A) - m(B)||m(A) + m(B)|\\ &\leq 4\varepsilon.\end{aligned}$$

Da $\varepsilon > 0$ beliebig war, folgt $m(B) = m(B)^2$. Das liefert dann die gewuenschte Aussage $m(B) \in \{0,1\}$. $\square$

KAPITEL 5

Mischungseigenschaften und Spektraltheorie

In einem vorangehenden Kapitel haben wir gesehen, daß ein meßbares dynamisches System (X, T, m) mit endlichem Maß ergodisch ist, wenn fuer alle meßbaren Mengen $A, B \subset X$ gilt

$$\frac{1}{N}\sum_{j=0}^{N-1} m(T^{-j}A \cap B) \to m(A)m(B).$$

In diesem Kapitel untersuchen wir zwei Verschaerfungen dieser Konvergenz.

1. Grundlegendes zu Mischungseigenschaften

DEFINITION. *Sei $(X, \mathcal{B}, m)$ ein Wahrscheinlichkeitsraum und $T : X \longrightarrow X$ eine maßerhaltende Transformation.*
(a) Es heißt T (bzw. das dynamische System (X, T, m)) schwach mischend, wenn fuer alle meßbaren A, B in X gilt

$$\frac{1}{N}\sum_{k=0}^{N-1} |m(T^{-k}A \cap B) - m(A)m(B)| \to 0, N \to \infty.$$

(b) Es heißt T (bzw. das dynamische System (X, T, m)) stark mischend, wenn fuer alle meßbaren A, B in X gilt

$$m(T^{-N}A \cap B) \to m(A)m(B), N \to \infty.$$

Bemerkungen.

- Offenbar gilt: T *stark mischend* $\Longrightarrow$ T *schwach mischend* $\Longrightarrow$ T *ergodisch.*
- Aus Ergodizitaet folgt nicht schwaches Mischen: Eine irrationale Rotation auf dem Einheitskreist ist ergodisch (s.o.), aber nicht schwach mischend (s.u. bzw. Uebung: Betrachte zwei kleine Intervalle A, B mit positiver Laenge; zeige $T^{-k}A \cap B \neq \emptyset \Longrightarrow T^{-k-1}A \cap B = \emptyset$ und schließe damit schwache Mischung aus).
- Aus schwachem Mischen folgt nicht starkes Mischen: Dazu gibt es verschiedene Beispiele. Tatsaechlich sind im generischen Sinne die meisten Transformationen schwach mischend aber nicht stark mischend.
- Ergodizitaet und schwaches Mischen moegen sehr aehnlich wirken. Es besteht aber ein grosser Unterschied. In gewisser Weise ist schwaches Mischen sogar dem starken Mischen naeher als der Ergodiziaet. Tatsaechlich sind (s.u.) fuer eine beschraenkte Folge (a_n) in $\mathbb{R}$ die folgenden beiden Aussage aequivaelent:
  (i) $\frac{1}{N}\sum_{k=0}^{N-1} |a_k| \to 0, N \to \infty.$

(ii) Es gibt eine Menge J in $\mathbb{N}_0$ mit Dichte 0, so daß fuer jede Folge $(n_k)_k$ in $\mathbb{N}_0 \setminus J$ gilt $a_{n_k} \to 0, k \to \infty$.

- Man kann anschaulich diese Eigenschaften wie folgt deuten:
  - Ergodizitaet: Die Mengen $T^{-k}A$ und B werden asymptotisch im Durchschnitt unabhaengig von einander.
  - Schwaches Mischen: Die Mengen $T^{-k}A$ und B werden asymptotisch bis auf wenige Ausnahmen unabhaengig von einander. (Dazu ist noch etwas zu tun; vgl. vorige Bemerkung und s.u.)
  - Starkes Mischen: Die Mengen $T^{-k}A$ und B werden asymptotisch unabhaengig von einander.

In den Definitionen von Ergodizitaet / Zerlegbarkeit, Schwachem Mischen und starkem Mischen geht es um alle meßbaren Mengen A, B. Tatsaechlich reicht es aber sich jeweils um einen 'Erzeuger zu kuemmern'. Das ist uns auch schon von der Untersuchung der Maßerhaltung einer Transformation bekannt.

Theorem (Charakterisierungen via Erzeuger). *Sei $(X, \mathcal{B}, m)$ ein Wahrscheinlichkeitsraum und $\mathcal{S}$ eine semi-Algebra, die $\mathcal{B}$ erzeugt. Sei $T : X \longrightarrow X$ eine maßerhaltende Transformation. Dann gilt:*
(a) Es ist T ergodisch genau dann wenn fuer alle $A, B \in \mathcal{S}$ gilt

$$\frac{1}{N} \sum_{k=0}^{N-1} m(T^{-k}A \cap B) \to m(A)m(B).$$

$N \to \infty$.
(b) Es ist T schwach mischend genau dann wenn fuer alle $A, B \in \mathcal{S}$ gilt

$$\frac{1}{N} \sum_{k=0}^{N-1} |m(T^{-k}A \cap B) - m(A)m(B)| \to 0,$$

$N \to \infty$.
(c) Es ist T stark mischend genau dann wenn fuer alle $A, B \in \mathcal{S}$ gilt

$$m(T^{-N}A \cap B) \to m(A)m(B),$$

$N \to \infty$.

Beweis. Jedes Element der von $\mathcal{S}$ erzeugten Algebra kann als disjunkte Vereinigung von endlich vielen Elementen aus $\mathcal{S}$ geschrieben werden (wie schon gezeigt wurde). Damit sieht man leicht, daß aus der Gueltigkeit der genannten Konvergenzaussagen fuer Elemente aus $\mathcal{S}$ auch die Gueltigkeit fuer Elemente aus der von $\mathcal{S}$ erzeugten Algebra. Da $\mathcal{S}$ die Algebra $\mathcal{B}$ erzeugt, gibt es weiterhin zu jedem $\varepsilon > 0$ - nach allgemeinen Saetzen aus der Maßtheorie - zu beliebigem $A, B \in \mathcal{B}$ Elemente A_0, B_0 in der von $\mathcal{S}$ erzeugten Algebra mit

$$m(A_0 \triangle A) < \varepsilon, \text{ und } m(B_0 \triangle B) < \varepsilon.$$

Weiterhin gilt offenbar

$$(T^{-k}A \cap B) \triangle (T^{-k}A_0 \cap B_0) \subset (T^{-k}A \triangle T^{-k}A_0) \cup (B \triangle B_0).$$

Da T maßerhaltend ist, folgt dann

$$m(T^{-k}A \triangle T^{-k}A_0) \cup (B \triangle B_0)) \leq m(A \triangle A_0) + m(B \triangle B_0) < 2\varepsilon.$$

Damit folgt schließlich

$$(*)\quad |m(T^{-k}A\cap B)-m(T^{-k}A_0\cap B_0)| \leq m((T^{-k}A\cap B)\triangle(T^{-k}A_0\cap B_0)) < 2\varepsilon.$$

Damit folgen die gewuenschten Aussagen aus der Dreiecksungleichung.

(a) Es gilt nach Dreiecksungleichung und $(*)$

$$\begin{aligned}
& \left|\frac{1}{N}\sum_{k=0}^{N-1} m(T^{-k}A\cap B) - m(A)m(B)\right| \\
= \;& |\frac{1}{N}\sum_{k=0}^{N-1} m(T^{-k}A\cap B) - \frac{1}{N}\sum_{k=0}^{N-1} m(T^{-k}A_0\cap B_0) \\
+ \;& \frac{1}{N}\sum_{k=0}^{N-1} m(T^{-k}A_0\cap B_0) - m(A_0)m(B_0) \\
+ \;& m(A_0)m(B_0) - m(A_0)m(B) \\
+ \;& m(A_0)m(B) - m(A)m(B)| \\
(*)\quad \leq \;& 4\varepsilon + \left|\frac{1}{N}\sum_{k=0}^{N-1} m(T^{-k}A_0\cap B_0) - m(A_0)m(B_0)\right|.
\end{aligned}$$

Damit folgt (a) direkt.

(b) und (c): Es gilt

$$\begin{aligned}
& |m(T^{-k}A\cap B) - m(A)m(B)| \\
= \;& m(T^{-k}A\cap B) - m(T^{-k}A_0\cap B_0) \\
+ \;& m(T^{-k}A_0\cap B_0) - m(A_0)m(B_0) \\
+ \;& m(A_0)m(B_0) - m(A)m(B_0) \\
+ \;& m(A)m(B_0) - m(A)m(B).
\end{aligned}$$

Damit liefert Dreiecksungleichung und $(*)$ nun leicht

$$|m(T^{-k}A\cap B) - m(A)m(B)| \leq 4\varepsilon + |m(T^{-k}A_0\cap B_0) - m(A_0)m(B_0)|.$$

Damit folgen (b) und (c) direkt. □

2. Mischung via Langzeitverhalten von U^n

In diesem Abschnitt betrachten wir ein meßbares dynamisches System (X,T,m) mit $m(X)=1$ und charakterisieren Ergodizitaet und Mischungsverhalten mittels des Langzeitverhaltens der Isometrie

$$U : L^2(X,m) \longrightarrow L^2(X,m), Uf = f\circ T.$$

Dabei wird es um die die Konvergenz von

$$\langle U^n f, g\rangle - \langle f, 1\rangle\langle 1, g\rangle$$

gegen 0 gehen (fuer $f,g\in L^2(X,m)$).

Wir beginnen mit einer Vorueberlegung, die im wesentlichen liefert, daß wir uns auf f und g beschraenken koennen, die senkrecht auf den konstanten

Funktion stehen. Dem liegt dann zugrunde, dass man den Hilbertraum zerlegen kann in

$$L^2(X,m) = Lin\{1\} \oplus (Lin\{1\})^{\perp}$$

und die gewuenschten Konvergenzen auf $Lin\{1\}$ klar sind und man sie dann nur noch auf dem orthogonalen Komplement der konstanten Funktionen ueberpruefen muss.

Proposition (Zerlegung $L^2 = 1 \oplus 1^{\perp}$). *Sei $(X, \mathcal{B}, m)$ ein Wahrscheinlichkeitsraum und $T : X \longrightarrow X$ maßerhaltend und U die assoziierte Isometrie. Seien $f, g \in L^2(X,m)$ beliebig und $f = f_0 + f_{\perp}$ und $g = g_0 + g_{\perp}$ mit $f_{\perp}, g_{\perp} \perp 1$ und $f_0 = \langle 1, f\rangle 1$ und $g_0 = \langle 1, g\rangle 1$. Dann gilt*

$$\langle U^n f, g\rangle - \langle f, 1\rangle\langle 1, g\rangle = \langle U^n f_{\perp}, g_{\perp}\rangle = \langle U^n f_{\perp}, g_{\perp}\rangle - \langle f_{\perp}, 1\rangle\langle 1, g_{\perp}\rangle.$$

Beweis. Die zweite Gleichheit ist klar. Die erste Gleichheit folgt durch direkte Rechnung unter Nutzen von

$$\begin{aligned}
\langle U^n f_{\perp}, g_0\rangle &= \langle U^n f_{\perp}, 1\rangle\langle 1, g\rangle \\
&= \int \overline{f_{\perp} \circ T^n} dm \langle 1, g\rangle \\
&= \int \overline{f_{\perp}} dm \langle 1, g\rangle \\
&= \langle f_{\perp}, 1\rangle\langle 1, g\rangle \\
&= 0.
\end{aligned}$$

und

$$\langle U^n f_0, g_{\perp}\rangle = \langle f, 1\rangle\langle U^n 1, g_{\perp}\rangle = 0.$$

auf folgende Weise

$$\begin{aligned}
\langle U^n(f_{\perp} + f_0), (g_{\perp} + g_0)\rangle &= \langle U^n f_{\perp}, g_{\perp}\rangle \\
&+ \langle U^n f_0, g_{\perp}\rangle + \langle U^n f_{\perp}, g_0\rangle + \langle U^n f_0, g_0\rangle \\
&= \langle U^n f_{\perp}, g_{\perp}\rangle + \langle U^n f_0, g_0\rangle \\
&= \langle U^n f_{\perp}, g_{\perp}\rangle + \langle f, 1\rangle\langle U^n 1, 1\rangle\langle 1, g\rangle. \\
&= \langle U^n f_{\perp}, g_{\perp}\rangle + \langle f, 1\rangle\langle 1, g\rangle.
\end{aligned}$$

Das beendet den Beweis. □

Theorem (Ergodiziaet und Verhalten von U^n). *Sei $(X, \mathcal{B}, m)$ ein Wahrscheinlichkeitsraum und $T : X \longrightarrow X$ maßerhaltend und U die assoziierte Isometrie. Dann sind die folgenden Aussagen aequivalent:*

(i) *Es ist T ergodisch.*

(ii) *Fuer alle $f, g \in L^2(X,m)$ gilt*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} \langle U^k f, g\rangle = \langle f, 1\rangle\langle 1, g\rangle.$$

(ii)' *Fuer alle $f, g \in L^2(X,m)$ mit $f, g \perp 1$ gilt*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} \langle U^k f, g\rangle = 0.$$

(iii) *Fuer alle* $f \in L^2(X,m)$ *gilt*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{k=0}^{N-1} \langle U^k f, f\rangle = \langle f, 1\rangle\langle 1, f\rangle.$$

(iii)' *Fuer alle* $f \in L^2(X,m)$ *mit* $f \perp 1$ *gilt*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{k=0}^{N-1} \langle U^k f, f\rangle = 0.$$

Beweis. Die Aequivalenz von (ii) und (ii)' und von (iii) und (iii)' folgt sofort aus der Proposition zur Zerlegung.

(i) $\Longrightarrow$ (ii): Das folgt fuer $f = 1_A$ und $g = 1_B$ (mit meßbaren Mengen A, B in X) aus einer bekannten Charkterisierung von Ergodizitaet. Nun kann man leicht sehen, daß (bei festem f) die Menge der g fuer die die gewuenschte Konvergenz gilt, stabil ist unter Bilden von endlichen Linearkombinationen und Konvergenz in $L^2(X,m)$. Damit folgt die gewuenschte Aussage fuer $f = 1_A$ mit einer meßbaren Menge $A \subset X$ fuer jedes $g \in L^2$. Nun ist aber bei festem g die Menge der f fuer die die gewuenschte Aussage gilt, stabil unter Bilden von endlichen Linearkombinationen und Konvergenz in L^2. Damit folgt dann (ii) fuer alle f und g.

(ii)$\Longrightarrow$ (iii): Das ist klar.

(iii)$\Longrightarrow$ (i): Aus (iii) folgt sofort, daß jede invariante meßbare Menge das Maß 0 oder 1 haben muß. Das liefert dann (i). $\square$

Bemerkungen.

- Im Zusammehang mit dem von Neumannschen Ergodensatz haben wir gesehen, daß Ergodizitaet die starke Konvergenz der
$$A_N = \frac{1}{N}\sum_{k=0}^{N-1} U^k$$
gegen die Projektion P auf den von der konstanten Funktion 1 erzeugten Unterraum impliziert. (Hier bedeutet starke Konvergenz gerade, daß $A_N f$ gegen Pf konvergiert fuer jedes f aus dem Hilbertraum.) Offenbar ist die Projektion P auf diesen Unterraum gerade gegeben durch
$$Pf = \langle 1, f\rangle 1.$$
Die Aussage (ii) besagt dann gerade die schwache Konvergenz der A_N gegen P. (Hier bedeutet schwache Konvergenz gerade, daß $\langle A_N f, g\rangle$ gegen $\langle Pf, g\rangle$ konvergiert fuer alle f, g aus dem Hilbertraum.)
- Aus den Betrachtungen des vorigen Punktes und dem Theorem erhalten wir dann insgesamt die Aequivalenz Ergodizitaet, starkter Konvergenz der A_N gegen P und schwacher Konvergenz der A_N gegen P.

Ganz aehnlich läßt sich nun folgendes beweisen.

THEOREM (Schwaches Mischen und Verhalten von U^n). *Sei* $(X,\mathcal{B},m)$ *ein Wahrscheinlichkeitsraum und* $T : X \longrightarrow X$ *maßerhaltend und* U *die assoziierte Isometrie. Dann sind die folgenden Aussagen aequivalent:*

(i) *Es ist* T *schwach mischend.*

(ii) *Fuer alle* $f,g \in L^2(X,m)$ *gilt*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} |\langle U^k f, g\rangle - \langle f,1\rangle\langle 1,g\rangle| = 0.$$

(ii)' *Fuer alle* $f,g \in L^2(X,m)$ *mit* $f,g \perp 1$ *gilt*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} |\langle U^k f, g\rangle| = 0.$$

(iii) *Fuer alle* $f \in L^2(X,m)$ *gilt*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} |\langle U^k f, f\rangle - \langle f,1\rangle\langle 1,f\rangle| = 0.$$

(iii)' *Fuer alle* $f \in L^2(X,m)$ *mit* $f \perp 1$ *gilt*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} |\langle U^k f, f\rangle| = 0.$$

Beweis. Die Aequivalenz von (ii) und (ii)' und von (iii) und (iii)' folgt sofort aus der Proposition zur Zerlegung.

(i) $\Longrightarrow$ (ii): Das kann man wie im vorigen Beweis zeigen. (Die Aussage folgt fuer $f = 1_A$ und $g = 1_B$ (mit meßbaren Mengen A,B in X) aus der Definition von schwachem Mischen. Nun kann man leicht sehen, daß (bei festem f) die Menge der g fuer die die gewuenschte Konvergenz gilt, stabil ist unter Bilden von endlichen Linearkombinationen und Konvergenz in $L^2(X,m)$. Damit folgt die gewuenschte Aussage fuer $f = 1_A$ mit einer meßbaren Menge $A \subset X$ fuer jedes $g \in L^2$. Nun ist aber bei festem g die Menge der f fuer die die gewuenschte Aussage gilt, stabil unter Bilden von endlichen Linearkombinationen und Konvergenz in L^2. Damit folgt dann (ii) fuer alle f und g.)

(ii)$\Longrightarrow$ (iii): Das ist klar.

(iii)$\Longrightarrow$ (i): Sei $f \in L^2(X,m)$ fixiert. Definiere

$$\mathcal{U}_f := \{g \in L^2(X,m) : \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} |\langle U^k f, g\rangle - \langle f,1\rangle\langle 1,g\rangle| = 0\}.$$

Dann sieht man leicht, daß $\mathcal{U}_f$ ein abgeschlossener Unterraum von $L^2(X,m)$ ist, der invariant unter U ist. Nach (iii) enthaelt $\mathcal{U}_f$ die Funktion f. Damit enthaelt es dann also alle $U^n f$, $n \in \mathbb{N}$. Offenbar enthaelt $\mathcal{U}_f$ ebenfalls die konstante Funktion 1. Damit ist also der von den $U^n f$, $n \in \mathbb{N}_0$ und der 1 aufgespannte abgeschlossene Unterraum $\mathcal{F}$ in $\mathcal{U}_f$ enthalten. Umgekehrt gilt fuer jedes g mit $g \perp \mathcal{F}$ nach Konstruktion

$$\langle U^n f, g\rangle = \langle 1, g\rangle = 0.$$

Damit sieht man sofort, daß auch jedes solche g zu $\mathcal{U}_f$ gehoert. Insgesamt enthaelt dann $\mathcal{U}_f$ also

$$\mathcal{F} \oplus \mathcal{F}^{\perp} = L^2(X, m).$$

Das liefert (ii) und damit dann nach Wahl von $f = 1_A$ und $g = 1_B$ (mit meßbaren Mengen $A, B \subset X$) sofort auch die Aussage (i). □

Ganz analog laeßt sich nun der folgende Satz beweisen.

Theorem (Starkes Mischen und Verhalten von U^n). *Sei $(X, \mathcal{B}, m)$ ein Wahrscheinlichkeitsraum und $T : X \longrightarrow X$ maßerhaltend und U die assoziierte Isometrie. Dann sind die folgenden Aussagen aequivalent:*

(i) *Es ist T stark mischend.*

(ii) *Fuer alle $f, g \in L^2(X, m)$ gilt*

$$\lim_{N\to\infty} \langle U^N f, g\rangle = \langle f, 1\rangle\langle 1, g\rangle.$$

(ii)' *Fuer alle $f, g \in L^2(X, m)$ mit $f, g \perp 1$ gilt*

$$\lim_{N\to\infty} \langle U^N f, g\rangle = 0.$$

(iii) *Fuer alle $f \in L^2(X, m)$ gilt*

$$\lim_{N\to\infty} \langle U^N f, f\rangle = \langle f, 1\rangle\langle f, 1\rangle.$$

(iii)' *Fuer alle $f \in L^2(X, m)$ mit $f \perp 1$ gilt*

$$\lim_{N\to\infty} \langle U^N f, f\rangle = 0.$$

Beweis. Der Beweis ist eine einfache Modifikation des Beweises des vorangehenden Theorems. □

3. Exkurs: Spektralsatz fuer unitaere Abbildungen

Wir werden die Ergebnisse des vorigen Abschnittes verwenden, um Mischungsverhalten via Spektraltheorie zu studieren. Dazu benoetigen wir noch den Spektralsatz fuer unitaere Abbildungen und eine Folgerung. Wir diskutieren diese Ergebnisse ohne einen Beweis zu geben.

Theorem (Spektralsatz fuer unitaere Abbildungen). *Sei U ein unitaerer Operator im Hilbertraum $\mathcal{H}$. Dann existiert fuer jedes $f \in \mathcal{H}$ ein eindeutiges positives endliches Borelmaß μ_f auf S^1 mit*

$$\langle U^n f, f\rangle = \int_{S^1} z^n d\mu_f(z).$$

Bemerkungen.

- Parametrisiert man S^1 durch $[0, 1)$, so ergibt sich

$$\langle U^n f, f\rangle = \int_{[0,1)} e^{2\pi i t n} d\mu_f(t).$$

Damit wird die Aussage des Spektralsatzes gerade, daß die Funktion

$$\mathbb{Z} \ni n \mapsto \langle U^n f, f\rangle$$

die Fouriertransformation des Maßes μ_f ist. Damit werden Saetze ueber Fouriertransformationen zu nuetzlichen Hilfsmitteln bei der Untersuchung dynamischer Systeme (s.u.).

- Eine Funktion $\Phi : \mathbb{Z} \longrightarrow \mathbb{C}$ heißt *positiv definit*, wenn fuer alle $c : \mathbb{Z} \longrightarrow \mathbb{C}$ mit endlichem Traeger gilt

$$\sum_{n,m\in\mathbb{Z}} \Phi(n-m)c(n)\overline{c(m)} \geq 0.$$

Offenbar ist mit Φ, dann auch $\overline{\Phi}$ positiv definit. Eine direkte Rechnung zeigt, daß fuer jedes positive endliche Maß auf S^1 die (inverse) Fouriertransformation von μ, also die Funktion

$$\Phi = \Phi_\mu : \mathbb{Z} \longrightarrow \mathbb{C},\ \Phi(n) = \int_{S^1} z^n d\mu(z),$$

positiv definit ist:

$$\begin{aligned} \sum_{n,m} \Phi(n-m)c(n)\overline{c(m)} &= \int_{S^1} \sum_{n,m} z^{n-m} c(n)\overline{c(m)} d\mu(z) \\ &= \int_{S^1} \left| \sum_n z^n c(n) \right|^2 d\mu(z) \geq 0. \end{aligned}$$

Es gilt nun auch die Umkehrung (Satz von Bochner), d.h. jede positiv definite Funktion auf $\mathbb{Z}$ ist gerade die (inverse) Fouriertransformation eines Maßes.

Ist f ein Element eines Hilbertraumes und U ein unitaerer Operator auf diesem Hilbertraum, so ist die Funktion $n \mapsto \langle U^n f, f\rangle$ positiv definit. Denn es gilt

$$\begin{aligned} \sum_{n,m} f(n-m)c(n)\overline{c(m)} &= \sum_{n,m} \langle U^n f, U^m f\rangle c(n)\overline{c(m)} \\ &= \left\| \sum_n c(n) U^n f \right\|^2 \geq 0. \end{aligned}$$

Damit folgt also aus dem Satz von Bochner die Existenz eines positiven Maßes μ mit

$$\langle U^n f, f\rangle = \int_{S^1} z^n d\mu(z).$$

Definition. *Sei U ein unitaerer Operator im Hilbertraum $\mathcal{H}$. Dann heißt das $f \in \mathcal{H}$ zugeordnete Maß μ_f aus dem vorigen Theorem das Spektralmaß von f (zu U).*

Ist U eine unitaere Abbildung im Hilbertraum $\mathcal{H}$, so heißt ein $\lambda \in \mathbb{C}$ ein Eigenwert von U, wenn es ein $f \in \mathcal{H}$ mit $f \neq 0$ und $Uf = \lambda f$ gibt. Dann heißt f eine Eigenfunktion zu λ. Es ist nicht schwer zu sehen, daß jeder Eigenwert eines unitaeren Operators zu S^1 gehoeren muß und Eigenfunktionen zu verschiedenen Eigenwerten orthogonal sind.

FOLGERUNG. *Sei U eine unitaere Abbildung im Hilbertraum $\mathcal{H}$. Dann ist das Spektralmaß von $f \in \mathcal{H}$ genau dann ein reines Punktmaß, wenn f im Abschluss der linearen Huelle aller Eigenfunktionen liegt. Genauer gilt*

$$f = \sum_{\lambda} c_\lambda f_\lambda$$

mit normierten paarweise orthogonalen Eigenfunktionen f_λ zu Eigenwerten λ und $c_\lambda \neq 0$, genau dann wenn

$$\mu_f = \sum_{\lambda} |c_\lambda|^2 \delta_\lambda.$$

4. Mischung via Spektralsatz

Ist (X, T, m) ein meßbares dynamisches System und T invertierbar, so ist die zugehoerige Abbildung $U : L^2(X, m) \longrightarrow L^2(X, m)$ unitaer. Gilt $m(X) = 1$, so ist offenbar die konstante Funktion 1 eine Eigenfunktion von U (zum Eigenwert 1). In diesem Sinne ist also die Eigenfunktion 1 unvermeidbar. Uns wird es um Systeme gehen, die keine weiteren Eigenfunktione haben.

DEFINITION. *Sei $(X, \mathcal{B}, m)$ ein Wahrscheinlichkeitsraum und $T : X \longrightarrow X$ maßerhaltend und U die assoziierte Isometrie. Es hat T stetiges Spektrum, wenn* 1 *der einzige Eigenwert ist und die konstanten Funktionen die einzigen Eigenfunktionen sind.*

Bemerkungen.

- Es hat also T stetiges Spektrum genau dann, wenn 1 der einzige Eigenwert ist und T ergodisch ist.
- Ist T invertierbar, so ist U unitaer. Dann bedeutet (nach der Folgerung des vorigen Abschnittes) stetiges Spektrum gerade, daß das Spektralmaß jeder auf 1 senkrechten Funktion stetig ist (d.h. keine Punktanteile besitzt).

Ende der Vorlesung.

Zur Bestimmung der Punktanteile eines Maßes gibt es folgendes fundamentale Resultat von Wiener.

LEMMA (Wiener Lemma). *Sei μ ein positives Borelmaß auf S^1 und*

$$F\mu : \mathbb{Z} \longrightarrow \mathbb{C}, F\mu(n) := \int_{S^1} z^n d\mu(z)$$

seine Fouriertransformation. Dann gilt

$$\sum_{\lambda \in S^1} \mu(\{\lambda\})^2 = \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N} |F\mu(n)|^2.$$

Bemerkungen.

- Fuer die 'meisten' $\lambda \in S^1$ ist $\mu(\{\lambda\}) = 0$. Tatsaechlich werden auf der linken Seite gerade die Quadrate der Punktanteile von μ summiert. Das ist beschraenkt durch das Quadrat des Gesamtpunktmasse von μ. Diese ist endlich (da μ ein endliches Maß ist).

- Es ist Teil der Aussage, daß der Grenzwert auf der rechten Seite existiert (da er ja gleich der endlichen linken Seite ist).
- Es gilt offenbar $F\mu(-n) = \overline{F\mu(n)}$. Damit kann man den Grenzwert auf der rechten Seite auch schreiben als

$$\lim_{N\to\infty} \frac{1}{2N} \sum_{n=-N}^{N} |F\mu(n)|^2.$$

Beweis. Eine direkte Rechnung liefert:

$$\begin{aligned}
\frac{1}{N}\sum_{n=0}^{N} |F\mu(n)|^2 &= \frac{1}{N}\sum_{k=0}^{N-1} \left| \int_{S^1} z^k d\mu(z) \right|^2 \\
&= \frac{1}{N}\sum_{k=0}^{N-1} \left(\int_{S^1} z^k d\mu(k) \int_{S^1} w^{-k} d\mu(w) \right) \\
(Fubini) \quad &= \frac{1}{N}\sum_{k=0}^{N-1} \int_{S^1\times S^1} (z\overline{w})^k d(\mu\times\mu)(z,w) \\
&= \int_{S^1\times S^1} \frac{1}{N}\sum_{k=0}^{N-1} (z\overline{w})^k d(\mu\times\mu)(z,w) \\
(!) \quad &\to \mu\times\mu(\text{Diagonale}) \\
(!!) \quad &= \sum_{\lambda\in S} \mu(\{\lambda\})^2.
\end{aligned}$$

Zu (!): Sei

$$F_N(z,w) := \frac{1}{N}\sum_{k=0}^{N-1} (z\overline{w})^k.$$

Dann gilt offenbar

- $F_N(z,z) = 1$ fuer alle $z \in S^1$.
- $|F_N(z,w)| \le 1$ fuer alle $N \in \mathbb{N}$ und $z,w \in S^1$.
- $F_N(z,w) = \frac{1}{N}\frac{1-(z\overline{w})^N}{1-z\overline{w}} \to 0$, $N\to\infty$ fuer $z \neq w$.

Damit folgt (!) aus dem Satz von der dominierten Konvergenz.

Zu (!!): Das ist klar. □

Theorem (Charakterisierung schwach mischende Systeme). *Sei $(X,\mathcal{B},m)$ ein Wahrscheinlichkeitsraum und $T : X \longrightarrow X$ maßerhaltend und invertierbar. Dann ist T schwach mischend genau dann, wenn es stetiges Spektrum hat.*

Beweis. Nach einer vorangehenden Charakterisierung ist T schwach mischend genau dann, wenn fuer alle $f \in L^2(X,m)$, die senkrecht auf 1 stehen, gilt

$$\frac{1}{N}\sum_{k=0}^{N-1} |\langle U^k f, f\rangle| \to 0, N\to\infty.$$

Hier kann man nun $|\cdot|$ durch $|\cdot|^2$ ersetzen. (Denn: Eine Richtung: Klar, fuer $\|f\| \le 1$ und dann auch nach Skalieren fuer beliebige f. Andere Richtung:

Folgt aus Cauchy-Schwarz angewendet auf $\langle x, y\rangle = \frac{1}{N}\sum_{k=0}^{N-1} x_k y_k$.) Dann folgt die gewuenschte Charakterisierung direkt aus dem Lemma von Wiener. $\square$

Definition (Rein absolut stetiges Spektrum). *Sei $(X, \mathcal{B}, m)$ ein Wahrscheinlichkeitsraum und $T : X \longrightarrow X$ maßerhaltend und invertierbar und U die assoziierte unitaere Abbildung. Es hat T absolut stetiges Spektrum, wenn die Spektralmaße (bzgl U) aller auf 1 senkrechten Funktionen absolut stetig bezueglich des Lebesgumaßes sind.*

Lemma (Lemma von Riemann-Lebesgue). *Sei μ ein positives Borelmaß auf S^1 und*

$$F\mu : \mathbb{Z} \longrightarrow \mathbb{C}, F\mu(n) := \int_{S^1} z^n d\mu(z)$$

seine Fouriertransformation. Dann gilt

$$F\mu(n) \to 0, |n| \to \infty,$$

falls μ absolut stetig bezueglich des Lebesguemaßes ist.

Beweis. Da μ absolut stetig bezueglich des Lebesguemaßes ist, gibt es (nach dem Satz von Radon-Nikodym) eine Funktion $h \in L^1(S^1)$ mit $\mu = h\lambda$, wobei λ das Lebesguemaß auf S^1 ist. Ist h die charakteristische Funktion eines Intervalles, so folgt die Aussage direkt durch Integration. Damit erhaelt man die Aussage dann leicht fuer endliche Linearkombinationen von charakteristischen Funktionen von Intervallen. Offenbar ist die Menge der $h \in L^1(X, m)$ fuer die die Aussage gilt, abgeschlossen unter Konvergenz in $L^1(X, m)$ (Evtl. Details...). Damit folgt dann die Aussage. $\square$

Theorem (Hinreichende Bedingung fuer starkes Mischen). *Sei $(X, \mathcal{B}, m)$ ein Wahrscheinlichkeitsraum und $T : X \longrightarrow X$ maßerhaltend und invertierbar und U die assoziierte unitaere Abbildung. Wenn T absolut stetiges Spektrum hat, dann ist T stark mischend.*

Beweis. Das folgt direkt aus dem Lemma von Riemann-Lebesgue und der vorangegangenen Charakterisierung von starkem Mischen mittels U^n. $\square$

Ende der Vorlesung

Bemerkung - RAGE Theorem fuer Schroedinger Operatoren. Fuer Schroedinger Operatoren in $\mathbb{R}^n$ bzw. $\mathbb{Z}^n$ gibt es - aus den 70 er Jahren - ganz aehnliche Aussagen zum Langzeitverhalten. Diese sind bekannt als RAGE-Theorem: Sei L ein beschraenkter (diese Einschraenkung ist nicht noetig) selbstadjungierter Operator im Hilbertraum $\mathcal{H}$. Dann gibt es zu jedem $f \in \mathcal{H}$ ein eindeutiges Maß μ_f auf $\mathbb{R}$ mit Traeger in $[-\|L\|, \|L\|]$ und

$$\langle L^n f, f\rangle = \int_{\mathbb{R}} t^n d\mu_f(t)$$

fuer alle $n \in \mathbb{N}_0$ (Spektralsatz fuer selbstadjungierte beschraenkte Operatoren). Zu dem selbstadjungierten L definiert dann

$$\mathbb{R} \longrightarrow \text{Unitaere Operatoren auf } \mathcal{H},\ t \mapsto e^{-itL},$$

eine unitaere Darstellung von $\mathbb{R}$. Aus dem Wiener-Lemma bzw. dem Lemma von Riemann-Lebesgue ergibt sich dann ganz analog zu den obigen Betrachtungen:

- μ_f stetig $\Longleftrightarrow \frac{1}{T}\int_0^T |\langle e^{-itL}f, f\rangle|^2 dt \to 0, T \to \infty$.
- μ_f absolut stetig (bzgl. Lebesguemaß) $\Longrightarrow |\langle e^{-itL}f, f\rangle|^2 \to 0, t \to \infty$.

Handelt es sich bei dem Hilbertraum um $L^2(\mathbb{R}^d, \lambda)$ oder $\ell^2(\mathbb{Z}^d)$ (und erfuellt der Operator noch eine schwache Kompaktheitsbedingung), so laeßt sich dann folgern (RAGE Theorem):

- μ_f reines Punktmaß $\Longleftrightarrow$ Zu jedem $\varepsilon > 0$ existiert ein kompaktes K mit
$$\|(1 - 1_K)e^{-itL}f\|^2 \leq \varepsilon$$
fuer alle $t \in \mathbb{T}$.
- μ_f stetig $\Longleftrightarrow \frac{1}{T}\int_0^T \|1_K e^{-itL}f\|^2 dt \to 0, T \to \infty$ fuer jedes kompakte K in $\mathbb{R}^d$ bzw. $\mathbb{Z}^d$.
- μ_f absolut stetig (bzgl. Lebesguemaß) $\Longrightarrow \|1_K e^{-itL}f\|^2 \to 0, t \to \infty$ fuer jedes kompakte K in $\mathbb{R}^d$ bzw. $\mathbb{Z}^d$.

5. Weiteres zu schwach mischenden Systeme

Hier lernen wir einige weitere Charakterisierungen von schwach mischenden Systemen kennen.

Definition. *Eine Teilmenge $A \subset \mathbb{N}$ hat die Dichte 0, wenn*
$$\lim_{N\to\infty} \frac{\sharp A \cap [0, N]}{N} = 0$$
gilt. Eine Teilmenge hat die Dichte 1 (auch volle Dichte), wenn
$$\lim_{N\to\infty} \frac{\sharp A \cap [0, N]}{N} = 0.$$

Bemerkung. Offenbar hat $A \subset \mathbb{N}$ die Dichte 0 genau dann, wenn $\mathbb{N} \setminus A$ die Dichte 1 hat.

Lemma (Cesaro Mittel vs Teilfolgen). *Sei (a_n) eine beschraenkte Folge in $\mathbb{C}$. Dann sind die folgenden beiden Aussagen aequivalent:*

- *Es gilt* $\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} |a_n| = 0$.
- *Es gibt eine Teilmenge J von $\mathbb{N}_0$ mit der Dichte 0 mit*
$$\lim_{n\to\infty, n\notin J} |a_n| = 0.$$

Bemerkung. Es bedeutet $\lim_{n\to\infty, n\notin J} |a_n| = 0$, daß fuer jede Folge (x_n) in $\mathbb{N}_0 \setminus J$ mit $x_n \to \infty$ gilt $a_{x_n} \to 0$.

Beweis. (ii)$\Longrightarrow$ (i): Sei $C \geq 0$ mit $|a_n| \leq C$ fuer alle $n \in \mathbb{N}_0$. Sei $\varepsilon > 0$ beliebig. Sei $N_\varepsilon \in \mathbb{N}$ so gewaehlt, daß
$$|a_n| \leq \varepsilon$$
fuer alle $n \in \mathbb{N}_0 \setminus J$ mit $n \geq N_\varepsilon$. Dann gilt fuer $N \geq N_\varepsilon$ also

$$\frac{1}{N}\sum_{n=0}^{N-1} = \frac{1}{N}\sum_{n\notin J, n\leq N_\varepsilon} |a_n| + \frac{1}{N}\sum_{n\notin J, N_\varepsilon < n \leq N-1} |a_n| + \frac{1}{N}\sum_{n\in J, n\leq N-1} |a_n|.$$

Fuer festes N_ε konvergiert nun der erste Term gegen 0 fuer $N \to \infty$. Der letzte Term konvergiert gegen 0, da er durch

$$C\frac{\sharp\{n \in J : n \leq N-1\}}{N}$$

beschraenkt ist. Der mittlere Term ist jedenfalls durch

$$\varepsilon\frac{\sharp\{n \notin J : N_\varepsilon < n \leq N-1\}}{N} \leq \varepsilon$$

beschraenkt (unabhaengig von N). Damit folgt die gewuenschte Aussage leicht.

(i)$\Longrightarrow$ (ii): Definiere fuer jedes $k \in \mathbb{N}$

$$A_k := \{n \in \mathbb{N}_0 : |a_n| \geq \frac{1}{k}\}.$$

Aus (i) folgt dann

$$\frac{\sharp A_k \cap [0, N-1]}{N} \to 0, N \to \infty.$$

Waehle nun eine - ohne Einschraenkung aufsteigende - Folge N_n in $\mathbb{N}$ mit $N_k \to \infty$ und

$$\frac{\sharp A_k \cap [0, N-1]}{N} \leq \frac{1}{k}$$

fuer alle $N \geq N_k$. Setze

$$J := \bigcup_k A_k \cap [N_k, \infty).$$

Nun ist offenbar $A_k \subset A_{k+1}$. Damit folgt

$$J \cap [0, N_{k+1}) \subset A_k \cap [0, N_{k+1}).$$

(Die 'groessern k' werden sowieso weggeschnitten und die kleineren sind schon enthalten.) Damit folgt fuer $N_k \leq N < N_{k+1}$ leicht

$$\frac{\sharp J \cap [0, N-1]}{N} \leq \frac{1}{k}.$$

Insgesamt ergibt sich also, dass J die Dichte 0 hat. Ausserdem gilt offenbar fuer $n \notin J$ mit $n \geq N_k$ auch $n \notin A_k$ und damit dann $|a_n| \leq \frac{1}{k}$. Es hat also J die gewuenschten Eigenschaften. □

Aus dem vorigen Lemma und den schon bekannten Charakterisierungen schwach mischender Systeme erhalten wir sofort den folgenden Satz

Folgerung (Charakterisierung schwach mischender Systeme). *Sei $(X, \mathcal{B}, m)$ ein Wahrscheinlichkeitsraum und $T : X \longrightarrow X$ maßerhaltend. Dann sind die folgenden Aussagen aequivalent:*

(i) *Es ist T schwach mischend.*
(ii) *Fuer alle meßbaren Mengen A, B in X existiert eine Menge J mit Dichte 0 mit*

$$\lim_{n \notin J, n \to \infty} m(T^{-n}A \cap B) = m(A)m(B).$$

(iii) *Fuer alle* $f, g \in L^2(X, m)$ *existiert eine Menge* J *von Dichte* 0 *mit*

$$\lim_{n \notin J, n \to \infty} \langle U^k f, g\rangle = \langle f, 1\rangle\langle 1, g\rangle.$$

Die Aussagen des vorigen Satzes laesst sich noch verbessern, wenn die zugrundeliegende σ-Algebra eine Abzaehlbarkeitsforderung erfüllt. Diese Abzaehlbarkeitsforderung lernen wir nun kennen.

Lemma. *Sei* $(X, \mathcal{B}, m)$ *ein Maßraum. Dann sind aequivalent:*

(i) *Es hat* $\mathcal{B}$ *eine abzaehlbare Basis d.h. es gibt eine abzaehlbare Menge* $\mathcal{B}'$ *in* $\mathcal{B}$, *so daß fuer jedes* $B \in \mathcal{B}$ *mit* $m(B) < \infty$ *und jedes* $\varepsilon > 0$ *ein* $B' \in \mathcal{B}'$ *existiert mit* $m(B \triangle B') < \varepsilon$.

(ii) *Es ist* $L^2(X, m)$ *separabel, d.h. es gibt eine abzaehlbare dichte Teilmenge* $\mathcal{D}$ *in* $L^2(X, m)$.

Bemerkung. Es ist der Raum $L^2(X, m)$ genau dann separabel, wenn er eine abzaehlbare Orthonormalbasis besitzt (Uebung).

Beweis. (Skizze) (i)$\Longrightarrow$ (ii): Offenbar ist die Menge der endlichen Linearkombinationen von Funktionen der Form 1_B mit $B \in \mathcal{B}]'$ mit Koeffizienten in $\mathbb{Q} + i\mathbb{Q}$ abzaehlbar und dicht in $L^2(X, m)$.

(ii)$\Longrightarrow$ (i): Approximiere die Funktionen aus $\mathcal{D}$ durch endliche Linearkombinationen von charakteristischen Funktionen. Definieren nun $\mathcal{B}'$ als die von den in diesen Approximationen vorkommenden Mengen erzeugte Algebra. $\square$

Theorem. *Sei* $(X, \mathcal{B}, m)$ *ein Wahrscheinlichkeitsraum und* $T : X \longrightarrow X$ *maßerhaltend und* U *die assoziierte Isometrie. Hat die* σ*-Algebra* $\mathcal{B}$ *eine abzaehlbare Basis, so sind die folgenden Aussagen aequivalent:*

(i) *Es ist* T *schwach mischend.*

(ii) *Es existiert eine Menge* J *in* $\mathbb{N}_0$ *mit Dichte* 0 *mit*

$$\lim_{n \notin J, n \to \infty} m(T^{-n}A \cap B) = m(A)m(B)$$

fuer alle meßbaren Mengen A, B *in* X.

(iii) *Es existiert eine Menge* J *von Dichte* 0 *mit*

$$\lim_{n \notin J, n \to \infty} \langle U^k f, g\rangle = \langle f, 1\rangle\langle 1, g\rangle$$

fuer alle $f, g \in L^2(X, m)$.

Beweis. Offenbar gelten die Implikationen (iii)$\Longrightarrow$ (ii) $\Longrightarrow$ (i).

Es bleibt also (i)$\Longrightarrow$ (iii) zu zeigen. Da die σ-Algebra eine abzaehlbare Basis hat, gibt es nach dem vorangegangenen Lemma eine abzaehlbare dichte Teilmenge $\mathcal{D}$ in $L^2(X, m)$. Sei $f_k, k \in \mathbb{N}$, eine Auflistung der Elemente von $\mathcal{D}$. Definiere

$$a_n := \sum_{k,j=1}^{\infty} \frac{1}{2^{k+j}} |\langle U^n f_k, f_j\rangle - \langle f_k, f_j\rangle|.$$

Dann gilt

$$\frac{1}{N}\sum_{n=0}^{N-1} a_n = \sum_{k,j=1}^{\infty} \frac{1}{2^{k+j}} \left(\frac{1}{N}\sum_{n=0}^{N-1} |\langle U^n f_k, f_j\rangle - \langle f_k, f_j\rangle| \right) \to 0, N \to \infty.$$

(Hier nutzen wir im letzten Schritt, daß T schwach mischend ist nach (i) und den Satz von der dominierten Konvergenz.) Nach einem vorausgehenden Lemma gibt es also eine Menge J in $\mathbb{N}_0$ mit Dichte 0 und

$$\lim_{n\to\infty, n\notin J} a_n = 0.$$

Damit folgt dann insbesondere aus der Definition von a_n auch

$$\lim_{n\to\infty, n\notin J} |\langle U^n f_k, f_j\rangle - \langle f_k, f_j\rangle| = 0$$

d.h.

$$\lim_{n\to\infty, n\notin J} \langle U^n f_k, f_j\rangle = \langle f_k, f_j\rangle$$

fuer alle $j, k \in \mathbb{N}$. Das ist gerade die gewueschte Aussage fuer f, g der Form $f = f_k$, $g = f_j$ mit $j, k \in \mathbb{N}$. Da die Menge $\mathcal{D}$ dicht ist und U eine Isometrie, folgt dann die gewuenschte Aussage fuer alle Paare $f, g \in L^2(X, m)$. $\square$

Ende der Vorlesung

Wir kommen nun noch zu einer weiteren komplett anderen Charakterisierung von schwach mischenden Systemen mittels des Produktes

$$T \times T : X \times X \longrightarrow X \times X, (x, y) \mapsto (Tx, Ty).$$

THEOREM. *Sei $(X, \mathcal{B}, m)$ ein Wahrscheinlichkeitsraum und $T : X \longrightarrow X$ maßerhaltend. Dann sind die folgenden Aussagen aequivalent:*

(i) *Es ist T schwach mischend.*
(ii) *Es ist $T \times T$ ergodisch.*
(iii) *Es ist $T \times T$ schwach mischend.*

Bemerkung. Es ist nicht schwer zu sehen, daß T genau dann stark mischend ist, wenn $T \times T$ stark mischend ist.

Beweis. (i)$\Longrightarrow$ (iii): Da die Rechtecke der Form $A \times B$ mit meßbaren $A, B \in X$ eine Semialgebra sind, die die Produkt-σ-Algebra erzeugt, reicht es nach einem schon bekannten Theorem zu zeigen:

$$\frac{1}{N}\sum_{n=0}^{N-1} |(m\times m)(T\times T)^{-n}(A\times B)\cap(C\times D) - (m\times m)(A\times B)(m\times m)(C\times D)| \to 0,\ N \to \infty$$

fuer alle meßbaren A, B, C, D in X. Nach einem vorangehenden Lemma zum Thema Cesaro Mittel vs Teilfolgen reicht es dann eine Menge $J \subset \mathbb{N}$ der Dichte 0 zu finden mit

$$\lim_{n\to\infty, n\notin J} (m\times m)(T\times T)^{-n}(A\times B)\cap(C\times D) = (m\times m)(A\times B)(m\times m)(C\times D).$$

Da T schwach mischend ist, gibt es aber nach einer schon bekannten Charakterisierung Teilmengen J_1, J_2 von $\mathbb{N}$ von Dichte 0 mit

$$\lim_{n\to\infty, n\notin J_1} m(T^{-n}A \cap B) = m(A)m(B)$$

und

$$\lim_{n\to\infty, n\notin J_2} m(T^{-n}C \cap D) = m(C)m(D).$$

Dann rechnet man leicht nach, daß mit $J = J_1 \cup J_2$ gilt

$$\lim_{n\to\infty, n\notin J} (m\times m)(T\times T)^{-n}(A\times B)\cap(C\times D) = (m\times m)(A\times B)(m\times m)(C\times D).$$

(iii)$\Longrightarrow$ (ii): Das ist klar.

(ii)$\Longrightarrow$ (i): Es reicht nach einer vorangegangenen Charakterisierung fuer alle meßbaren A, B in X zu zeigen

$$\frac{1}{N}\sum_{n=0}^{N-1}\big|m(T^{-n}A\cap B)-m(A)m(B)\big|^2\to 0,\ N\to\infty.$$

Es gilt nach (ii) sowohl

$$\begin{aligned}\frac{1}{N}\sum_{n=0}^{N-1}m(T^{-n}A\cap B) &= \frac{1}{N}\sum_{n=0}^{N-1}(m\times m)((T\times T)^{-n}(A\times X)\cap(B\times X))\\ \text{(nach (ii))} &\to (m\times m)(A\times X)(m\times m)(B\times X)\\ &= m(A)m(B)\end{aligned}$$

als auch

$$\begin{aligned}\frac{1}{N}\sum_{n=0}^{N-1}m(T^{-n}A\cap B)^2 &= \frac{1}{N}\sum_{n=0}^{N-1}(m\times m)((T\times T)^{-n}(A\times A)\cap(B\times B))\\ \text{(nach (ii))} &\to (m\times m)(A\times A)(m\times m)(B\times B)\\ &= m(A)^2m(B)^2.\end{aligned}$$

Damit ergibt sich dann

$$\begin{aligned}&\frac{1}{N}\sum_{n=0}^{N-1}\big|m(T^{-n}A\cap B)-m(A)m(B)\big|^2\\ =\ &\frac{1}{N}\sum_{n=0}^{N-1}\big(m(T^{-n}A\cap B)^2-2m(T^{-n}A\cap B)m(A)m(B)+m(A)^2m(B)^2\big)\\ \to\ &m(A)^2m(B)^2-2m(A)m(B)m(A)m(B)+m(A)^2m(B)2\\ =\ &0.\end{aligned}$$

Das beendet den Beweis. □

6. Beispiele

In diesem Abschnitt untersuchen wir die oben gegebenen Beispiele auf ihr Mischungsverhalten.

6.1. Die Identitaet. Sei $(X, \mathcal{B}, m)$ ein beliebiger Maßraum mit $m(X) = 1$. Sei $T : X \longrightarrow X$ die Identitaet. Dann sind aequivalent:

(i) T ist ergodisch.
(ii) T ist schwach mischend.
(iii) T ist stark mischend.
(iv) Jede meßbare Teilmenge von X hat Maß 0 oder Maß 1.

Beweis. Abgesehen von (iv)$\Longrightarrow$ (iii) sind alle Aussagen klar. Dazu: $m(T^{-n}A\cap B) = 1 = m(A)m(B)$ falls $m(A) = m(B) = 1$ und $m(T^{-n}A\cap B) = 0$ falls nicht $m(A) = 1 = m(B)$. □

6.2. Rotation auf dem Einheitskreis. Sei $\mathbb{T} := \mathbb{R}/\mathbb{Z}$ mit dem Lebesguemass λ und sei fuer $\alpha \in RR$ die Transformation T_α definiert durch

$$T_\alpha : \mathbb{T} \longrightarrow \mathbb{T}, T_\alpha(x + \mathbb{Z}) = \alpha + x + \mathbb{Z}.$$

Theorem. *Es ist T_α fuer kein α schwach mischend.*

Beweis. Betrachte die Funktion $f : \mathbb{T} \longrightarrow S^1, f(x) = e^{ix}$. Dann ist f nicht konstant und eine Eigenfunktion zum Eigenwert $e^{i\alpha}$. Damit folgt die Aussage aus der (einfachen Richtung) der Charakterisierung schwach mischender Systeme. □

Das Beispiel kann verallgemeinert werden wie folgt. Sei G eine beliebige kompakte Gruppe zusammen mit der Borel-σ-Algebra und dem Haarmaß. Sei $a \in G$ beliebig. Dann ist $T = T_a$ mit

$$T : G \longrightarrow G, Tx := ax$$

eine meßbare Transformation. Tatsaechlich ist T invertierbar mit inverser Transformation gegeben durch $T_{a^{-1}}$.

Theorem. *Sei G eine kompakte abelsche Gruppe mit Haarmaß λ und $T = T_a$ die Translation um $a \in G$. Dann ist T_a nicht schwach mischend.*

Beweis. Jeder stetige Gruppenhomomorphismus $\gamma : G \longrightarrow S^1$ ist offenbar eine Eigenfunktion. Auf einer kompakten abelschen Gruppe gibt es nun genug solcher Gruppenhomomorphismen, um die Punkte zu trennen. □

Erinnerung. Charaktere auf kompakten abelschen Gruppen (Definition, Orthogonalitaet, Punktetrennung), Fouriertransformation.

Ende der Vorlesung

Bemerkung. Tatsaechlich ist keine Rotation auf einer beliebigen kompakten Gruppe schwach mischend. Denn eine schwach mischende Abbildung ist jedenfalls ergodisch und eine Rotation auf einer Gruppe kann nur ergodisch sein, wenn die Gruppe abelsch ist (Uebung).

6.3. Surjektive Homomorphismen einer kompakten abelschen Gruppe. Wir haben schon fuer fuer $p \in \mathbb{Z}$ die Abbildung

$$T = T_p : \mathbb{T} \longrightarrow \mathbb{T}, T(x + \mathbb{Z}) = px + \mathbb{Z}$$

untersucht. Dann ist T eine meßbare Transformation. Eine alternative Darstellung ist gegeben durch.

$$T : S^1 \longrightarrow S^1, Tz = z^p.$$

Hier geht es um noch allgemeinere Modelle.

Theorem (Rohlin - Halmos). *Sei G eine kompakte abelsche Gruppe ausgestattet mit dem normalisierten Haarmaß. Sei $A : G \longrightarrow G$ ein surjektiver stetiger Gruppenhomomorphismus. Dann sind aequivalent:*

(i) *Es ist A ergodisch.*
(ii) *Der triviale Charakter 1 ist der einzige Charakter γ mit $\gamma \circ A^n = \gamma$ fuer ein $n \in \mathbb{N}$.*

Beweis. (i)$\Longrightarrow$ (ii): Sei A ergodisch und γ ein Charakter mit $\gamma \circ A^n = \gamma$ fuer ein $n \in \mathbb{N}$. Sei n das kleinste solche Zahl. Dann ist

$$f = \gamma + \gamma \circ A + \ldots + \gamma \circ A^{n-1}$$

invariant unter A. Damit ist also nach (i) die Funktion f konstant. Wir zeigen nun $n = 1$: Angenommen $n \geq 2$. Dann sind aber $\gamma, \gamma \circ A, \ldots, \gamma \circ A^{n-1}$ verschieden (da A surjektiv ist und n minimal) und damit (da sie Charaktere sind) orthogonal und damit ist f nicht konstant.

(ii)$\Longrightarrow$ (i): Sei $f \in L^2(G)$ mit $f = f \circ A$. Sei

$$f = \sum_\gamma a_\gamma \gamma$$

mit $\sum |a_\gamma|^2 < \infty$. Wegen $f = f \circ A$ gilt dann

$$\sum a_\gamma \gamma \circ A = \sum a_\gamma \gamma.$$

Sind nun $\gamma, \gamma \circ A, \ldots, \gamma \circ A^n, \ldots$ alle verschieden, so sind ihre Koeffizienten gleich und muessen damit verschwinden. Verschwindet also umgekehrt ein Koeffzient a_γ nicht, so muss es ein $p \in \mathbb{N}$ geben mit

$$\gamma \circ A^p = \gamma.$$

Damit ist also nach (ii) dann $\gamma = 1$ und daher f konstant. □

THEOREM. *Sei G eine kompakte abelsche Gruppe ausgestattet mit dem normalisierten Haarmaß. Sei $A : G \longrightarrow G$ ein surjektiver stetiger Gruppenhomomorphismus. Dann sind aequivalent:*

(i) *Es ist T ergodisch.*
(ii) *Es ist T schwach mischend.*
(iii) *Es ist T stark mischend.*

Beweis. Es gilt (iii)$\Longrightarrow$ (ii)$\Longrightarrow$ (i) nach allgemeinen Erwaegungen (s.o.). Es bleibt also (i)$\Longrightarrow$ (iii) zu zeigen.

Behauptung. Seien γ und ϱ Charaktere auf G und nicht $\gamma = 1 = \varrho$. Dann gilt $\langle U^n\gamma, \varrho\rangle = 0$ fuer alle großen n. Gilt $\gamma = 1 = \varrho$, so gilt$\langle U^n\gamma, \varrho\rangle = 1$ fuer alle n.

Bew. Die zweite Aussage ist klar. Zur ersten Aussage: Ist $\gamma = 1$ und $\varrho \neq 1$, so ist die Aussage klar. Ist $\gamma \neq 1$ so sind aufgrund von (i)$\Longrightarrow$ (ii) im vorigen Theorem $\gamma, \gamma \circ A, \ldots$ paarweise verschieden. Also kann (spaetestens) ab einem gewissen n der Charakter $U^n\gamma$ nicht mehr mit δ uebereinstimmen.

Aufgrund der Behauptung folgt dann

$$\langle U^n\gamma, \varrho\rangle \to \langle\gamma, 1\rangle\langle 1, \varrho\rangle, n \to \infty$$

fuer alle Charaktere γ, ϱ. Damit folgt dann

$$\langle U^n f, g\rangle \to \langle f, 1\rangle\langle 1, g\rangle$$

fuer alle f, g die endliche Linearkombinationen von Charakteren sind. Durch Approximation folgt das dann fuer beliebige $f, g \in L^2(G)$. Das ist dann die gewuenschte Aussage. □

6.4. Der Bernoulli und der $(p_1, \ldots, p_k)$ **Shift.** . Seien $k \geq 2$ und $p_1, \ldots, p_k > 0$ mit $\sum_{j=1}^k p_j = 1$ gegeben. Sei $Y := \{0, \ldots, k-1\}$ mit der σ-Algebra aller Teilmengen ausgestattet und sei μ auf Y das eindeutige Maß mit

$$\mu(\{j\}) = p_{j+1}$$

fuer $j = 0, \ldots, k-1$. Sei $X := Y^{\mathbb{Z}}$ mit der Produkt-σ-Algebra und dem Produktmaß. Sei

$$T : X \longrightarrow X, (Tx)(n) = x(n+1).$$

Dann ist T eine invertierbare maßerhaltende Transformation.

THEOREM. *Der* $(p_1, \ldots, p_k)$*-Shift ist stark mischend*

Beweis. Das folgt aehnlich wie die Ergodizitaet (und eigentlich haben wir es damals schon mitbewiesen). Die Grundidee ist, dass aufgrund der Shiftwirkung die Mengen A und $T^{-n}B$ fuer grosse n 'unabhaengig voneinander' werden. Hier sind die Details:
Nach einem schon bekannten Satz, reicht es fuer meßbare Rechtecke R, S zu zeigen

$$m(T^{-n}R \cap S) \to m(R)m(S), n \to \infty.$$

Wir skizzieren das fuer den Spezialfall $R = [A]_N, S = [B]_L$. (Der allgemeine Fall kann analog behandelt werden.) Es gilt

$$m(T^{-n}R \cap S) = m(T^{-n}[A]_N \cap [B]_L) = m([A]_{n+N} \cap [B]_L) = m([A]_{n+N})m([B]_L)$$

fuer alle grossen n. **Zeichnung.** $\square$

6.5. Der einseitige $(p_1, \ldots, p_k)$ **- Shift.** Der einseitige Shift ist ebenfalls stark mischend. Wir lassen die Details fuer den Leser.

KAPITEL 6

Markov Shifts

In diesem Kapitel untersuchen wir eine weitere wichtige Klasse von Beispielen.

Seien $k \geq 2$ und $Y := \{0, \ldots, k-1\}$ mit der σ-Algebra aller Teilmengen ausgestattet und sei $X := Y^{\mathbb{Z}}$ mit der Produkt-σ-Algebra versehen. Sei

$$T : X \longrightarrow X, (Tx)(n) = x(n+1).$$

Seien nun fuer jedes $n \in \mathbb{N}_0$ und beliebige $a_1, \ldots, a_n \in Y$ reelle Zahlen $p_n(a_0, \ldots, a_n)$ gegeben mit

- $p_n(a_0, \ldots, a_n) \geq 0$,
- $\sum_{a \in Y} p_0(a) = 1$,
- $p_n(a_0, \ldots, a_n) = \sum_{a \in Y} p_{n+1}(a_0, \ldots, a_n, a)$ und $p_n(a_0, \ldots, a_n) = \sum_{a \in Y} p_{n+1}(a, a_0, \ldots, a_n)$

Dann gibt es nach dem Satz von Kolmogorov ein eindeutiges Maß m auf X mit

$$m(\{x : x_q = a_0, \ldots, x_{q+n} = a_n\} = p_n(a_0, \ldots, a_n)$$

fuer all $q \in \mathbb{Z}$, $n \in \mathbb{N}_0$, $a_0, \ldots, a_n \in Y$. Offenbar erhaelt T das Maß von Rechtecken und ist damit dann insgesamt maßerhaltend.

Bemerkung. Man kann (Uebung) den zweiseitigen Shift leicht als Spezialfall der obigen Konstruktion erhalten.

Seien nun ein *Wahrscheinlichkeitsvektor* $p \in \mathbb{R}^k$, d.h. ein $p = (p_0, \ldots, p_{k-1})$ mit

$$p_j > 0 \text{ und } \sum_j p_j = 1,$$

und eine *stochastische Matrix* $P \in \mathbb{R}^{k \times k}$, d.h. ein $P = (p_{ij})$ mit

$$p_{ij} \geq 0 \text{ und } \sum_j p_{ij} = 1$$

fuer alle i, gegeben, so daß gilt

$$pP = p \text{ (d.h. } p_j = \sum_i p_i p_{ij}).$$

Dann erfuellen die Zahlen

$$p_n(a_0, \ldots, a_n) = p_{a_0} p_{a_0 a_1} p_{a_1 a_2} \ldots p_{a_{n-1} a_n}$$

die oben genannte Bedingungen. (Nachrechnen.) Damit gibt es also zugehoerig ein eindeutiges T-invariantes Maß $m = m_{p,P}$. Es heißt dann T auf dem zugehoerigen Maßraum der (p, P)-*Markov Shift*.

Bemerkung. Wir nehmen hier ohne Einschraenkung $p_j > 0$ an fuer alle j. Waere das nicht der Fall, so koennten wir einfach das entsprechende j aus Y entfernen.

Grundlegend fuer die folgenden Untersuchungen ist die naechste Proposition.

PROPOSITION (Verbindung von P und U). *Fuer beliebige $i, j \in Y$ und $n \in \mathbb{N}_0$ gilt:*

$$m(\{x : x_q = i, x_{q+n} = j\} = p_i P^n(i,j).$$

Beweis. Es gilt

$$\begin{aligned}
& m(\{x : x_q = i, x_{q+n} = j\} \\
= \quad & \sum_{j_1, j_2, \ldots, j_{n-1} \in Y} m((\{x : x_q = i, x_{q+1} = j_1, \ldots, x_{q+n-1} = j_{n-1}, x_{q+n} = j\}) \\
\text{(Definition m)} \quad = \quad & \sum_{j_1, j_2, \ldots, j_{n-1} \in Y} p_i p_{ij_1} p_{j_1 j_2} \cdots p_{j_{n-1} j} \\
= \quad & p_i P^n(i,j).
\end{aligned}$$

Das beendet den Beweis. □

Bemerkung. Die Proposition liefert die wesentliche Verbindung zwischen dem durch T auf $L^2(X, m)$ induzierten Operator U und der Matrix P (zusammen mit dem Vektor p) auf folgende Weise. Sei $1_i \in L^2(X, m)$ die charakteristische Funktion der Menge

$$X_i := \{x \in X : x(0) = i\}$$

und sei $\delta_i \in \ell^2(Y) = \mathbb{C}^k$ die charakteristische Funktion von $i \in Y$. Dann gilt, wie man sich leicht ueberlegt,

$$m(\{x : x_0 = i, x_n = j\} = m(X_i \cap T^{-n} X_j) = \langle 1_i, U^n 1_j \rangle.$$

Damit besagt die Proposition also gerade

$$\langle 1_i, U^n 1_j \rangle = p_i \langle \delta_i, P^n \delta_j \rangle.$$

Dabei handelt es sich auf der linken Seite um das Skalarprodukt auf $L^2(X, m)$ und auf der rechten Seite um das Skalarprodukt in $\ell^2(Y)$. Damit werden also die grundlegenden Groessen zur Beschreibung von U mit P und p in Beziehung gesetzt.

Bemerkung / Interpretation. Die Matrix P beschreibt einen Markovprozess. Dabei wird eine Irrfahrt auf Y modelliert, die nach folgender Regel ablaeuft: Wenn man gerade an der Stelle i angekommen ist, so

- waehlt man zufaellig die naechste Station aus, wobei $j \in Y$ mit Wahrscheinlichkeit p_{ij} gewaehlt wird,
- wartet die Zeit 1,
- und springt dann zur Stelle j.

Die Gesamtheit aller moeglichen Bahnen wird dann durch die Menge $X = Y^{\mathbb{Z}}$ beschrieben. Die Wahrscheinlichkeit fuer das 'Abspringen' einer bestimmen Konfiguration ist dann eindeutig durch die Matrix P zusammen mit einer Wahrscheinlichkeitsverteilung fuer den Startpunkt bestimmt. Diese letztere Wahrscheinlichkeitsverteilung wird durch p modelliert. Die Verschiebung T entspricht dann gerade dem Vorruecken um einen Zeitschritt.

Ende der Vorlesung

Uns wird es zunaechst um eine Charakterisierung von Ergodizitaet gehen. Dazu brauchen wir noch etwas Vorarbeit.

LEMMA. *Sei P eine stochastische Matrix mit einem srikt positiven Eigenvektor p mit $pP = p$. Dann existiert*

$$Q = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} P^n.$$

Die Matrix Q ist ebenfalls stochastisch und erfuellt $QP = PQ = Q$ sowie $Q = Q^2$. Jeder (links / rechts) Eigenvektor von P zum Eigenwert 1 ist ebenfalls ein (links / rechts) Eigenvektor von Q zum Eigenwert 1.

Beweis. Wir zeigen zunaechst Existenz des Grenzwertes: Betrachte den (p, P) Markov Shift mit dem zugehoerigen Maß m. Sei 1_j die charakteristische Funktion von $[j]_0$ i.e. von

$$\{x \in Y^{\mathbb{Z}} : x(0) = j\}.$$

Nach dem Birkhoffschen Ergodensatz (gilt ja ohne Ergodizitaetsannahme) konvergiert

$$\frac{1}{N} \sum_{n=0}^{N-1} 1_j(T^n x) \to 1_j^*(x)$$

mit einem geeigneten 1_j^* fuer m-fast alle $x \in X$. Damit folgt nach Multiplikation mit 1_i dann

$$\frac{1}{N} \sum_{n=0}^{N-1} p_i P^n(i,j) \overset{!}{=} \frac{1}{N} \sum_{n=0}^{N-1} \int 1_i(x) 1_j(T^n x) dm(x) \to \int 1_j^*(x) 1_i(x) dm(x)$$

nach dem Satz ueber dominierte Konvergenz. (Hier verwenden wir bei ! die obige grundlegende Proposition zur Verbindung von U und P.) Damit ist dann $Q = (q_{ij})$ mit

$$q_{ij} = \frac{1}{p_i} \int 1_j^*(x) 1_i(x) dm(x)$$

der gewuenschte Grenzwert.

Aus der Definition von Q folgt sofort Stochastizitaet von Q sowie $QP = PQ = Q$. Ebenso sieht man aus der Definition leicht, daß jeder Eigenvektor von P zum Eigenwert 1 auch ein Eigenvektor von Q zum Eigenwert 1 ist.

Zur Gleicheit von Q und Q^2:

$$Q^2 = Q(\lim_N \frac{1}{N} \sum_{n=0}^{N-1} P^n) = \lim_N \frac{1}{N} \sum_{n=0}^{N-1} QP^n = \lim_N \frac{1}{N} \sum_{n=0}^{N-1} Q = Q.$$

Hier verwenden wir in der vorletzten Gleichung das schon bewiesene $QP = Q$. $\square$

THEOREM. *Fuer den (p, P)-Markov Shift mit strikt positivem p und stochastischem P sind aequivalent:*

(i) *Es ist T ergodisch.*
(ii) *Alle Zeilen von Q stimmen ueberein.*
(iii) *Jeder Eintrag von Q ist strikt positiv.*
(iv) *Es ist P irreduzibel (d.h. zu allen $i, j \in Y$ existiert ein N mit $P^N(i,j) > 0$. Deutung im Graph.)*
(v) *Es ist 1 ein einfacher Eigenwert von P.*

Bemerkung. Die Vielfachheit eines Eigenwertes ist unabhaengig davon, ob man Rechts- oder Linkseigenvektoren betrachtet.

Beweis. (i)$\Longrightarrow$ (ii): Aus einer obigen Proposition und dem vorangegangenen Lemma folgt

$$\frac{1}{N}\sum_{n=0}^{N-1} m([i]_0 \cap [j]_n) = \frac{1}{N}\sum_{n=0}^{N-1}\int 1_i(x)1_j(T^n x)dm \overset{Prop}{=} \frac{1}{N}\sum_{n=0} N-1 p_i P^n(i,j) \overset{Lem}{\to} p_i q_{ij}.$$

Aufgrund der Ergodizitaet (i) gilt aber auch

$$\frac{1}{N}\sum_{n=0}^{N-1} m([i]_0 \cap [j]_n) \to m([i]_0)m([j]_0) = p_i p_j.$$

Durch Vergleich folgt

$$q_{ij} = p_j$$

unabhaengig von i. Dast ist die gewuenschte Aussagen.

(ii)$\Longrightarrow$ (iii): Wegen $p = pP$ gilt auch (s.o.) $p = pQ$. Da die Zeilen von Q konstant sind, folgt dann also fuer jedes j die Aussage

$$0 < p_j = \sum_i p_i q_{ij} = q_{1j}\sum_i p_i = q_{1j}.$$

Damit ist die erste (und damit dann jede) Zeile von Q strikt positiv.

(iii)$\Longrightarrow$(iv): Da die Eintraege von Q strikt positiv sind nach (iii) und $Q = \lim_N \frac{1}{N}\sum_{n=0}^{N-1} P^n$ gilt, folgt die gewuenschte Aussage sofort.

(iv)$\Longrightarrow$ (v): Sei v ein (von 1 linear unabhaengiger) Eigenvektor zu P d.h. $Pv = v$. Ohne Einschraenkung gelte $v \perp 1$ (sonst betrachte $v - \frac{1}{\|1\|^2}\langle v, 1\rangle 1$.) Damit hat v also sowohl positive als auch negative Eintraege. Sei $v = v_+ - v_-$ die Zerlegung in nichtnegative und nichtpositive Eintraege. Sei k der Index an dem v den groessten Eintrag hat. Dann gilt fuer jedes n also

$$v_k = (P^n v)_k = (P^n v_+)_k + (P^n v_-)_k.$$

Da P^n stochastisch ist und v_k maximal, gilt abe $P^n v_+(k) \le v_+(k)$. Aufgrund von $v_- \neq 0$ und der Irreduzibilitat, kann man weiterhi n so waehlen, dass $(P^n v_-)_k < 0$. Damit ergibt sich ein Widerspruch.

(v)$\Longrightarrow$ (ii): Wegen $QP = P$ ist jede Zeile von Q ein Eigenvektor von P zum Eigenwert 1. Da 1 ein einfacher Eigenwert ist, muessen dann alle Zeilen Vielfache voneinander sein. Da die Matrix stochastisch ist, haben alle Zeilen die Summe 1 und stimmen damit ueberein.

(ii)$\Longrightarrow$ (i): Nach (ii) gilt (vgl. den Schluss (ii)$\Longrightarrow$ (iii)) $q_{ij} = p_j$ fuer alle i, j. Es reicht zu zeigen

$$\frac{1}{N}\sum_{n=0}^{N-1} m(T^{-n}A \cap B) = m(A)m(B)$$

fuer alle meßbaren Rechtecke. Wir zeige dies (mit einer gewissen Beschraenkung der Allgemeinheit) fuer $A = [j]_0$ und $B = [i]_0$. Es gilt nach der obigen Proposition zur Verbindung von U und (p, P)

$$\begin{aligned} \frac{1}{N}\sum_{n=0}^{N-1} m(T^{-n}A \cap B) &= \frac{1}{N}\sum_{n=0}^{N-1} m([j]_n \cap [i]_0) \\ (Prop) &= \frac{1}{N}\sum_{n=0}^{N-1} p_i P^n(i,j) \\ (Lemma) &\to p_i q_{ij} \\ &= p_i p_j \\ &= m(A)m(B). \end{aligned}$$

Das beendet den Beweis. □

Bemerkung. Es ist eine gute Uebung, direkt weitere Implikationen zu beweisen.

Wir kommen nun noch zu einer Charakterisierung des Mischungsverhalten des Markov Shifts.

THEOREM. *Fuer den (p, P)-Markov Shift mit strikt positivem p und stochastischem P sind aequivalent:*

(i) *Es ist T schwach mischend.*
(ii) *Es ist T stark mischend.*
(iii) *Es gilt $P^n(i,j) \to p_j$ fuer alle $i, j \in Y$.*
(iv) *Die Matrix P ist primitiv (i.e. es existiert ein $N \in \mathbb{N}$, so dass alle Eintraege von P^N strikt positiv sind).*

Beweis. (iv)$\Longrightarrow$ (iii): Das folgt aus dem Perron-Frobenius Theorem (s.u.).

(iii)$\Longrightarrow$ (ii): Es reicht zu zeigen

$$m(T^{-n}A \cap B) = m(A)m(B)$$

fuer alle meßbaren Rechtecke. Wir zeige dies (mit einer gewissen Beschraenkung der Allgemeinheit) fuer $A = [j]_0$ und $B = [i]_0$. Es gilt nach der obigen Proposition zur Verbindung von U und (p, P)

$$\begin{aligned} m(T^{-n}A \cap B) & \quad m([j]_n \cap [i]_0) \\ (Prop) &= p_i P^n(i,j) \\ (iii) &\to p_i p_j \\ &= p_i p_j \\ &= m(A)m(B). \end{aligned}$$

(ii)$\Longrightarrow$ (i): Das ist klar.

(i)$\Longrightarrow$ (iv): Fuer jedes Paar von Elementen $i, j \in Y$ gibt es aufgrund des schwachen Mischungsverhaltens eine Menge J_{ij} von Dichte 0 mit

$$\lim_{n\to\infty, n\notin J_{ij}} m([i]_0 \cap T^{-n}[j]_0) = m([i]_0)m([j]_0),$$

d.h. mit

$$p_i P^n(i,j) \to p_i p_j,$$

d.h. mit

$$P^n(i,j) \to p_j$$

fuer $n \to \infty$ mit $n \notin J_{ij}$. Setzt man

$$J := \bigcup_{i,j} J_{ij},$$

so hat J immer noch die Dichte 0 und es gilt fuer alle $i,j \in Y$

$$\lim_{n\to\infty, n\notin J} P^n(i,j) = p_j > 0.$$

Damit ist dann P primitiv. □

Ende der Vorlesung

THEOREM. *(Perron/Frobenius) Sei M eine quadratische Matrix mit den folgenden beiden Eigenschaften:*

- *Die Matrixelemenete von M sind nichtnegativ.*
- *Es gibt eine natuerliche Zahl N, sodass alle Matrixelemente von M^N strikt positiv sind.*

Dann gelten die folgenden Aussagen.
(a) Es gibt einen positiven Eigenwert θ mit $\theta > |\lambda|$ fuer alle anderen Eigenwerte λ von M.
(b) Es gibt einen strikt positiven Eigenvektor zu θ und θ ist der einzige Eigenwert zu dem ein nichtnegativer Eigenvektor existiert.
(c) Der Eigenwert θ hat die algebraische Vielfachheit eins.

Bemerkung (Uebung). Eine Matrix M mit den beiden Eigenschaften des Theorems heisst primitiv. Eine Matrix M mit nichtnegativen Eintraegen heisst *irreduzibel*, wenn es zu jedem i,j ein $N = N(i,j)$ gibt mit $M^N(i,j) > 0$. Dann heisst der (nicht von i abhaengende) groesste gemeinsame Teiler von

$$\{k : M^k(i,i) > 0\}$$

die Periode p von M. Gilt fuer diese Periode $p = 1$, so heisst M *aperiodisch.* Es gilt:

- M primitiv $\Longleftrightarrow$ M ist irreduzibel und aperiodisch.

Fuer irreduzible Matrizen gibt es einen eindeutigen Eigenwert $\theta > 0$ mit $\theta \geq |\lambda|$ fuer alle Eigenwerte λ von M. Fuer dieses θ gelten (b) und (c). Weiterhin gibt es genau p Eigenwerte von M mit Norm θ. Dieses sind gegeben durch

$$\lambda = \theta e^{2\pi i \frac{k}{p}}, \;\; k = 0, \ldots, p-1.$$

Beweis. In diesem Beweis schreiben wir fuer Vektoren x, y im Euklidischen Raum $x \leq y$ falls jede einzelne Komponente von x kleiner oder gleich der entsprechenden Komponente von y ist und wir schreiben $x < y$, falls jede Komponente von x echt kleiner als die entsprechende Komponente von y ist.

Wir betrachten die Menge (der 'Eigenwerte zu subharmonischen Funktionen')

$$\mathcal{S} := \{s \geq 0 : sf \leq Mf \text{ fuer ein } f \geq 0 \text{ mit } f \neq 0\} \subset [0,\infty).$$

Offenbar enthaelt S die Zahl 0 und ist ein Intervall. Die Menge S ist beschraenkt (z.B. durch $\|M\|$). Die Menge S ist abgeschlossen. (Betrachte s_n mit $s_n \to s$ und $s_n f_n \leq M f_n$; o.E. koennen wir annehmen, dass $\|f_n\| \leq 1$ fuer alle n, also o.E. $f_n \to f$...). Damit ist S also kompakt. Da S also kompakt ist, existiert

$$\theta = maxS.$$

Die Menge S enthaelt ein $s > 0$. (Es hat M keine Nullzeile. Damit kann man $s := \min\{(M1)_j\}$ waehlen, wobei 1 die Konstante Funktion 1 ist.) Damit folgt

$$\theta > 0.$$

Wir zeigen nun, dass θ die gewünschten Eigenschaften hat.

Behauptung. Gilt $\theta|g| \leq M|g|$ fuer ein $g \neq 0$, so folgt $\theta|g| = M|g|$ und alle Eintraege von $|g|$ sind strikt positiv.
Bew. Nach Voraussetzung gilt $M|g| - \theta|g| \geq 0$. Waere $M|g| - \theta|g| \neq 0$, so folgte $0 < M^N(M|g| - \theta|g|) = M(M^N|g|) - \theta M^N|g| = My - \theta y$ mit $y = M^N|g|$. Das ist ein Wiederspruch zur Maximalitaet von θ. Dann gilt aber auch $\theta^N|g| = M^N|g|$. Hier ist die rechte Seite nach Voraussetzung strikt positiv, also auch die linke Seite, also auch $|g|$.

Behauptung. Sei $f \geq 0$ mit $\theta f \leq Mf$ gegeben. Dann gilt $\theta f = Mf$. Insbesondere ist f strikt positiv und θ ist ein positiver Eigenwert.
Bew. Nach Voraussetzung gilt $Mf - \theta f \geq 0$ und $f = |f|$. Nun folgt alles aus der ersten Behauptung.

Behauptung. Ist λ ein anderer Eigenwert, so gilt $|\lambda| < \theta$.
Bew. Ist g eine Eigenfunktion zu λ so gilt $|\lambda||g| = |\lambda g| = |Mg| \leq M|g|$. Damit folgt $|\lambda| \leq \theta$ aufgrund der Maximalitaet von θ. Sei nun $|\lambda| = \theta$. Dann gilt also

$$\theta|g| = |\lambda||g| \leq M|g|.$$

Aus der ersten Behauptung folgt dann

$$\theta|g| = M|g|.$$

Damit ist also $|g|$ eine Eigenfunktion zu θ. Das impliziert weiterhin

$$\theta^N|g| = |\lambda^N g| = |M^N g| \leq M^N|g| = \theta^N|g|.$$

Da alle Eintraege von M^N positiv sind, folgt also, dass g und $|g|$ bis auf einen Faktor uebereinstimmen. Damit ist g (wie auch $|g|$) eine Eigenfunktion zu θ und es folgt $\lambda = \theta$.

Behauptung. Die algebraische Vielfachheit von θ ist eins.
Bew. Die geometrische Vielfachheit ist eins: Sind g, f zwei Eigenfunktionen zu θ so auch $f - cg = h$. Waehle c so dass eine Komponente von h verschwindet. Es gilt

$$\theta|h| = |\theta h| = |Mh| \leq M|h|.$$

Dann muss gelten (vgl erste Behauptung), dass $|h|$ eine Eigenvektor von θ ist. Damit folgt dann aus der ersten Behauptung $h \equiv 0$.

Waere nun die algebraische Vielfachheit groesser als die geometrische Vielfachheit, so gaebe es ein 'Jordankaestchen'. Dann wuerde man also einen strikt positiven Eigenvektor w zu θ finden und ein $v \neq 0$ mit

$$Mv = w + \theta v.$$

Induktiv ergibt sich dann einfach

$$M^n v = n\theta^{n-1} w + \theta^n v$$

fuer alle $n \in \mathbb{N}$. Daraus folgt dann

$$M^n |v| \geq |M^n v| \geq \theta^n n \left| \frac{1}{\theta} w + \frac{1}{n} v \right|.$$

Fuer genuegend große n ist aber

$$n \left| \frac{1}{\theta} w + \frac{1}{n} v \right| \geq 2\theta^n |v|.$$

Damit folgt fuer ein genuegend grosses n also

$$M^n |v| > 2\theta^n |v|.$$

Wendet man nun auf M^n das (fuer M) oben gezeigt an, so schliesst man, dass der (betrags)groesste Eigenwert von M^n mindestens $2\theta^n$ ist. Andererseits ist der (betrags)groesste Eigenwert von M^n natuerlich gerade die n-te Potenz des (betrags)groessten Eigenwertes von M d.h. θ^n. Das ist ein Widerspruch.

Behauptung. Ist λ ein Eigenwert von M mit einem nichtnegativen Eigenvektor v so gilt $\lambda = \theta$.
Bew. Es gilt $\lambda^N v = M^N v$. Daraus folgt dann sofort, dass λ positiv ist und alle Eintraege von v strikt positiv sind (da M^N nur positive Eintraege hat). Sei w ein strikt positiver Eigenvektor zu θ. Dann koennen wir also ein $C > 0$ finden mit

$$w - Cv < 0.$$

Damit gilt dann also

$$0 \geq M^n (w - Cv) = \theta^n (w - \left(\frac{\lambda}{\theta}\right)^n Cv)$$

fuer alle $n \in \mathbb{N}$. Fuer grosse n ist aber die rechte Seite strikt positiv, wenn $\lambda < \theta$ gilt. □

Ende der Vorlesung

FOLGERUNG (Langzeitverhalten primitiver Matrizen). *Die quadratische Matrix M habe nichtnegative Einträge und es habe M^N strikt positive Eintraege fuer ein $N \in \mathcal{N}$. Sei $\theta > 0$ der betragsgroesste Eigenwert von M und von M^T und seien v bzw. w zugehoerige strikt positive Eigenvektoren mit $\langle v, w \rangle = 1$. Sei die Abbildung $P = P_\theta$ gegeben durch*

$$Px = \langle w, x \rangle v$$

und sei

$$N := M - \theta P.$$

Dann gilt:

- $M = \theta P + N$.
- *Es ist P eine Projektion (d.h. es gilt $P^2 = P$).*

- *Die Betraege der Eigenwerte von N sind alle kleiner als θ.*
- *Es gilt $PN = NP = 0$.*

Insbesondere folgt also

$$\left(\frac{1}{\theta}M\right)^n = (P + \frac{1}{\theta}N)^n \to P, n \to \infty.$$

Bemerkung. Es ist P im allgemeinen keine orthogonale Projektion (d.h. es gilt nicht $P^* = P$).

Beweis. Sei U eine invertierbare Matrix, die M diagonalisiert, so dass gilt

$$M = U(\theta, J)U^{-1}$$

und $Ue_1 = v$ und $U^{-1}v = e_1$. Sei $\widetilde{P} := U(10)U^{-1}$ und $\widetilde{N} := U(0J)U^{-1}$. Dann gilt:

- $M = \theta\widetilde{P} + \widetilde{N}$.
- Es ist $\widetilde{P}$ eine Projektion.
- Die Betraege der Eigenwerte von $\widetilde{N}$ sind alle kleiner als θ.
- Es gilt $\widetilde{P}\widetilde{N} = \widetilde{N}\widetilde{P} = 0$.

Es reicht also nun $P = \widetilde{P}$ zu zeigen. (Dann folgt aus $M = \theta\widetilde{P} + \widetilde{N} = \theta P + N$ auch $N = \widetilde{N}$ und die uebrigen Aussagen folgen sofort.)

Nun zum Beweis von $P = \widetilde{P}$: Sei $e_1 = (1, 0, \dots, 0)$.
Es gilt

$$M^t = (U^{-1})^*(\theta, J^t)U^*.$$

Die erste Spalte u von $(U^*)^{-1}$ ist dann also ein Eigenvektor von M^t zum Eigenwert θ ($U^*u = e_1$, $(U^*)^{-1}e_1 = u$...) Weiterhin gilt fuer diesen Eigenvektor

$$1 = \langle e_1, e_1\rangle = \langle U^*e_1, (U^*)^{-1}e_1\rangle = \langle v, u\rangle.$$

Damit folgt

$$u = w$$

(da u und w beide Eigenwerte von M_t zu θ sind mit der gleichen Normierung).

Wir berechnen nun Px zu

$$Px = U(10)U^{-1}x = U((U^{-1}x)_1 e_1) = (U^{-1}x)_1 v.$$

Es bleibt $U^{-1}x$ zu bestimmen. Unter Nutzen von $u = w$ folgt

$$(U^{-1}x)_1 = \langle e_1, U^{-1}x\rangle = \langle (U^{-1})^*e_1, x\rangle = \langle u, x\rangle = \langle w, x\rangle.$$

Nimmt man die letzten beiden Zeilen zusammen so folgt die Aussage. □

FOLGERUNG (Anwendung stochastische Matrix). *Ist P eine primitive stochastische Matrix, so ist $\theta = 1$ und es gilt mit der Projektion P_1 aus der vorigen Folgerung dann*

$$P^n \to P_1, n \to \infty.$$

Beweis. Es ist 1 ein strikt positiver Eigenvektor zum Eigenwert 1. Damit ist nach dem vorangehenden Theorem dann $\theta = 1$. Nun folgt die Aussage aus der vorangehenden Folgerung. $\square$

Bemerkung. Die obigen Betrachtungen lassen sich stark verallgemeinern. Abstrakt findet man sich in folgender Situation: Gegeben ist ein kompaktes Y mit einem endlichen Mass und eine Halbgruppe

$$P : [0, \infty) \longrightarrow \text{Operatoren auf } L^2(Y, Mass)$$

mit folgenden beiden Eigenschaften:

- $P_t f \geq 0$ fuer $f \geq 0$ und $t \geq 0$.
- $P_t 1 = 1$ fuer alle $t \geq 0$.

Die Halbgruppe hat dann den *Grundzustand* 1 (i.e. $P^n 1 = 1$) und man untersucht nun auch noch das Verhalten von $(P^n)^* 1$. Tatsaechlich laeßt sich dies soger noch einmal abstrahieren.

Literaturangaben

Die Notizen verwenden eine ganze Reihe von Quellen. Stark beeinflußt sind sie durch das Buch von Peter Walters 'An Introduction to Ergodic Theory', Graduate Texts in Mathematics 79, Springer 1981. In diesem Zusammenhang sei auch auf das Buch von Karl Pedersen 'Ergodic Theory', Cambridge University Press 1989 und das Buch von I. P. Cornfeld, S. V. Fomin, Y. Sinai 'Ergodic Theory', Grundlehren der mathematischen Wissenschaften 245, Springer, 1982, verwiesen.