ENTROPIE UND ERGODIZITÄT

GILLES BELLOT

Forschungsseminar — Ratner's Theoreme
TU Dortmund — Fakultät für Mathematik
Lehrstuhl II: Geometrie und Algebra
A. Maier — L. Pottmeyer — R. Scharlau

Zusammenfassung. Wir beschreiben zuerst zwei, unter dem Gesichtspunkt der Theorie chaotischer Systeme und der Entropie verschiedene, dynamische Systeme. Im zweiten Teil erläutern wir den Zusammenhang zwischen Entropie, Gleichverteilung und dem Punkteweise-Ergoden-Theorem von Birkhoff, indem wir uns unter anderem anschauen, was Schokoladenbrötchen backende Bäcker und ewig segelnde Möwen uns über das Langzeitverhalten dynamischer Systeme verraten.

Wir folgen den Ausführungen in [Wit03] Kapitel 2 und 3. Für eine Einführung in die Wahrscheinlichkeitstheorie siehe [Fel68] und [Fel71].

1. Dynamische Systeme

Definition 1.1. *Sei $(X, \mathcal{B}, \mu)$ ein Maßraum, dann heißt die Abbildung $T : X \to X$ μ-meßbar, falls $\forall A \in \mathcal{B} : T^{-1}A \in \mathcal{B}$. Ein solches T heißt invertierbar, falls $\forall A \in \mathcal{B} : TA \in \mathcal{B}$ und $TX = X$. Ein meßbare Abbildung heißt maßtreu bzgl. μ, wenn*

$$\forall A \in \mathcal{B} : T_*\mu(A) := \mu(T^{-1}A) = \mu(A).$$

Ist T invertierbar, so ist diese Bedingung äquivalent zu

$$\forall A \in \mathcal{B} : T^*\mu(A) := \mu(TA) = \mu(A).$$

Ist T maßtreu, so nennt man $(X, \mathcal{B}, \mu, T)$ ein (diskretes) dynamisches System. Der Orbit von $x \in X$ unter T ist definiert als die Menge $\mathrm{Orb}_T(x) := \{T^n x \mid n \in \mathbb{Z}\}$.

Intuitiv beschreibt die Entropie[1] eines dynamischen Systems den Grad der Unvorhersehbarkeit oder der Unberechenbarkeit des Systems. Bevor wir dies formal genauer definieren, wollen wir uns zunächst zwei verschiedene dynamische Systeme anschauen.

1.1. Das Kreisbillard.

Wir betrachten die Billardabbildung auf einem Kreis als dynamisches System auf $S := \mathbb{R}/2\pi\mathbb{Z}$. Diese ist durch Rotation um $\theta \in \mathbb{R}$ eindeutig definiert: $T_\theta : S \to S$, $T_\theta(x) := x + \theta$, für $\theta \in \mathbb{R}$. Für $n \in \mathbb{N}$ gilt offenbar $x_n := T_\theta^n(x) = x_0 + n\theta \bmod 2\pi$, wobei x_0 der erste Kontaktpunkt der Billardkugel mit dem Rand ist.

Date: 18. November 2014.
2000 *Mathematics Subject Classification.* 28D10, 22E40, 11D09.
Key words and phrases. dynamical systems, entropy, ergodic theory, number theory.

[1] Griechisches Kunstwort: entropía, von en — ein, in und trope — Wendung, Umwandlung

Satz 1.2. *Die eben definierte Billardabbildung T_θ ist maßtreu bzgl. des Lebesgue-Maßes μ.*

Beweis. Betrachte o.B.d.A. $T(x) = x + \theta \bmod 1$. Nach dem Fortsetzungssatz von Kolmogorov (siehe [Fel71] Kapitel IV.6 Theorem 1) reicht es, offene Mengen zu betrachten, da diese unsere σ-Algebra $\mathcal{B}$ erzeugen. Es gilt $T^{-1}(a,b) = \{x \in \mathbb{R}/\mathbb{Z} \mid T(x) \in (a,b)\} = (a-\theta, b-\theta)$, also erhalten wir

$$\begin{aligned} T_*\mu(a,b) &= \mu T^{-1}(a,b) \\ &= \mu(a-\theta, b-\theta) \\ &= (b-\theta) - (a-\theta) \\ &= b-a \\ &= \mu(a,b), \text{ wie gewünscht.} \end{aligned}$$

□

Ist θ 2π-rational, so ist die Billardbahn offensichtlich periodisch, genauer sogar: gilt $\theta = 2\pi\frac{p}{q}$ mit $p,q \in \mathbb{N}$, so ist die Bahn $q-$ periodisch (mit Windungszahl) p, d. h. $T_\theta^q(x) = x + 2\frac{qp}{q}\pi x \equiv_{2\pi} x + 2p\pi \equiv_{2\pi} x$, für alle $x \in S$, also $x_{n+q} \equiv_{2\pi} x_n$.

Theorem 1.3 (Jacobi)**.** *Ist θ irrational, dann liegt der Orbit eines jeden Startpunktes dicht in S, d. h. jedes Intervall beinhaltet Punkte aus dem Orbit.*

Beweis. Sei x unser Startpunkt. Wir laufen also auf dem Kreis entlang und machen dabei Schritte der Länge θ. Nach einer gewissen Anzahl von Schritten, sagen wir n, sind wir wieder beim Startpunkt und gehen über x hinweg. Es ist klar, dass wir nicht wieder genau beim Startpunkt ankommen, sonst wäre $\theta = \frac{2\pi}{n}$. Seien nun $y = T_\theta^n(x) \equiv_{2\pi} x + n\theta$ der Punkt vor x und $z = T_\theta^{n+1}(x) = T_\theta(y) \equiv_{2\pi} y + \theta$ der nächste Punkt, dann ist klar, dass eines der Kreisabschnitte yx oder zx eine Länge kleiner als $\frac{\theta}{2}$ hat. Wir definieren θ_1 als diese Länge (immer noch 2π-irrational) und bemerken, dass $T_{\theta_1} = T_\theta^n$, somit haben wir also eine neue Rotation definiert. Diese Konstruktion führen wir induktiv fort und erhalten so eine Folge von Rotationen um Winkel $\theta_k \to 0$ und jede dieser Rotationen ist eine Iteration von T_θ.

Sei nun I ein beliebiges Intervall auf dem Kreis, dann können wir k so wählen, dass $\theta_k < \|I\|$, also enthält I einen Punkt aus $\mathrm{Orb}_{T_{\theta_k}}(x)$, wie gewünscht. □

Ist $\theta = 2\pi\frac{p}{q}$, so besteht die Folge $x_n \equiv_{2\pi} x + n\theta$ aus q Elementen, welche sehr regulär auf dem Kreis verteilt sind. Wir wollen nun zeigen, dass dies auch für 2π-irrationale Winkel gilt und sprechen von Gleichverteilung. Seien I ein Intervall und $\mathrm{k}(n)$ die Anzahl der Folgenglieder $x_0, \dots, x_{n-1} \in I$, dann heißt unsere Folge gleichverteilt auf S, falls

$$\lim_{n\to\infty} \frac{\mathrm{k}(n)}{n} = \frac{|I|}{2\pi}. \tag{1}$$

Theorem 1.4 (Kronecker-Weyl)**.** *Ist θ irrational, so ist die Folge $x_n \equiv_{2\pi} x + n\theta$ auf dem Kreis gleichverteilt.*

Beweisskizze. Kronecker und Weyl haben dies mit klassischen Methoden der Theorie der diophantischen Approximation gezeigt, wir wollen aber etwas mehr zeigen und den Ergoden-Satz von Birkhoff schon mit ins Spiel bringen. Sei f eine integrierbare

Abbildung auf dem Kreis, dann gilt

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(x_j) = \frac{1}{2\pi} \int_0^{2\pi} f(x)dx. \tag{2}$$

Wir erhalten also sozusagen eine Gleichheit zwischen dem zeitlichen Mittelwert und dem Scharmittelwert, siehe Theorem 2.7. Um aus (1) (2) zu folgern, definiere f als die charakteristische Funktion von I (siehe auch ($Korollar$ 2.8), also

$$\chi_I : I \to I, x \mapsto \begin{cases} 0\ , & x \notin I \\ 1\ , & x \in I \end{cases},$$

dann gilt

$$\lim_{n\to\infty} \frac{1}{n} k(n) = \frac{1}{2\pi}|I|,$$

wie gewünscht.

Die Abbildung f kann durch trigonometrische Polynome approximiert werden, wir werden den Satz für harmonische Abbildungen beweisen, oder genauer, für $f(x) = e^{ikx}$, also für eine komplexe Abbildung deren Real- und Imaginärteile harmonisch sind (als Real- und Imaginärteil einer holomorphen Abbildung).

Ist $k = 0$, also $f \equiv 1$, dann sind beide Seiten von (2) gleich 1. Sei also nun $k \geq 1$, dann ist die linke Seite der Gleichung (2) eine geometrische Reihe:

$$\frac{1}{n} \sum_{j=0}^{n-1} e^{ikj\theta} = \frac{1}{n} \frac{e^{ikn\theta} - 1}{e^{ik\theta} - 1} \to 0.$$

Offensichtlich ist $\int_0^{2\pi} e^{ikx}dx = 0$, wie gewünscht. □

Bemerkung 1.5. *Wir werden im zweiten Teil sehen, dass die Kreisrotation um einen irrationalen Winkel ergodisch* (*Theorem* 2.9) *ist. Der Satz von Birkhoff* (*Theorem* 2.7) *liefert uns dann das gewünschte Ergebnis* automatisch *mit. Genauer ist die Billardabbildung sogar eindeutig ergodisch* (*Theorem* 2.14)*, dies ist aber eine* zu starke *und nicht gebrauchte Voraussetzung (vgl. Kapitel V des ersten Vortrags).*

1.2. **Die Shift-Abbildung.** Wir wollen nun ein zweites System betrachten, welches sich *chaotisch* verhält, wir müssen also die Shift-Abbildung ins Spiel bringen. Betrachte dazu zunächst was passiert, wenn eine Münze unendlich oft geworfen wird. Der nun betrachtete Wahrscheinlichkeitsraum ist also ein unendliches kartesisches Produkt

$$C^\infty := \{f : \mathbb{Z} \to \mathbb{F}_2\},$$

mit dem Produktmaß, d. h. für beliebige disjunkte endliche Teilmengen $\mathcal{K}$ und $\mathcal{Z}$ ist die Wahrscheinlichkeit, dass

- $f(n) = 0$ für alle $n \in \mathcal{K}$, und
- $f(n) = 1$ für alle $n \in \mathcal{Z}$

genau gleich $\frac{1}{2}^{|\mathcal{K}|+|\mathcal{Z}|}$.

Nun beschreibt f die Geschichte eines Mannes welcher, ohne Müdigkeit, tapfer, jeden Tag, bis in alle Ewigkeit, seine *faire* Münze einmal wirft. Dabei ist $f(0)$ das Ergebnis seines *heutigen* Wurfes, $f(n)$ das Ergebnis, welches er in n Tagen erhalten wird und $f(-n)$ das Ergebnis, welches er vor n Tagen erzielt hat. *Morgen* wird seine

Geschichte von $g \in C^\infty$ erzählt, wobei $g(n) = f(n+1)$ für alle $n \in \mathbb{Z}$. Bekannt ist, dass fast jede Folge von Münzwürfen (im Limes) halb aus 0en (Kopf) und halb aus 1en (Zahl) besteht.

Wir möchten das System noch etwas genauer betrachten. Dies führt uns zu einem symbolischen dynamischen System, dem sogenannten Bernoulli-Shift.

Betrachte $\Sigma := \{0,1\}^{\mathbb{Z}}$ die Menge aller 0-1-Folgen, dann heißt

$$\sigma : \Sigma \to \Sigma$$
$$(s_i)_{i\in\mathbb{Z}} \mapsto (s_{i+1})_{i\in\mathbb{Z}}$$

shift-Abbildung auf Σ.

Mit der Bernoulli-Verteilung (und den in der Literatur üblichen Bezeichnungen) erhalten wir ein Maß auf $\{0,1\}$, also $p_1, p_2 \geq 0$ mit $p_1 + p_2 = 1$. Betrachte nun die Zylinder $[a_1, \dots, a_n]_t := \{x \in \Sigma \mid \forall 0 \leq i \leq n : x_{n+i} = a_i\}$ (Zylinder eines Subshifts), dies ist die Menge aller Folgen $(x_i)_{i\in\mathbb{Z}} \in \Sigma$, die ab der n-ten Koordinate die vorher gegebene Symbolfolge $a_0, a_1, \dots a_n$ haben. Diese Zylinder sind sowohl offen als auch abgeschlossen und bilden eine abzählbare Basis der Produkttopologie, jede offene Menge lässt sich als Vereinigung von Zylindermengen darstellen.

Genauer definieren wir nun μ auf diesen Zylindermengen wie folgt:

$$\mu([a_1, \dots, a_t]_t) := \prod_{i=1}^{n} p_{a_i} \in [0,1]\,,$$
$$\mu(\emptyset) := 0$$

und erhalten so

$$\Sigma = [0]_1 \cup [1]_1 \Rightarrow \mu(\Sigma) = p_1 + p_2 = 1.$$

Damit haben wir also ein Wahrscheinlichkeitsmaß, wie gewünscht, auf der durch die Zylinder erzeugten Algebra definiert. Die gleiche Argumentation/Konstruktion definiert ein Wahrscheinlichkeitsmaß auf *bi-unendlichen* Folgen über einem beliebigen Alphabet.

Satz 1.6. *Die Shift-Abbildung ist μ-invariant.*

Beweis.

$$\sigma_*\mu([a_1, a_2, \dots, a_n]_t) = \mu([a_1, a_2, \dots, a_n]_{t+1}) = \prod_{i=1}^{n} p_{a_i} = \mu([a_1, a_2, \dots, a_n]_t).$$

□

Satz 1.7. *Durch $d(s,t) := \max\{2^{-|i|} \mid s_i \neq t_i\}$ wird eine Metrik auf Σ definiert.*

Beweis. Seien $s, t, u \in \Sigma$ beliebig, aber fest.

Ist $s = t$, so gilt $s_i = t_i$ für alle $i \in \mathbb{Z}$, also ist dann $d(s,t) = 0$. Umgekehrt folgt aus $d(s,t) = 0$ sofort, dass $s_i = t_i$ für alle i, also $s = t$. Da d als Maximum einer Menge positiver Zahlen definiert ist, gilt $d(s,t) > 0$. Die Symmetrie ist offensichtlich.

Zeige nun noch $d(s,u) \leq d(s,t) + d(t,u)$. Ist $s = u$, so ist nichts zu zeigen. Sei also o.B.d.A. $s \neq u$ und q die erste Stelle mit $s_q \neq u_q$, also $d(s,u) = 2^{-q}$. Betrachte nun folgende Fälle:

(1) $t_i = s_i$ für alle $i \Rightarrow d(s,t) = 0 \wedge d(t,u) = 2^{-q}$
(2) $t_i = u_i$ für alle $i \Rightarrow d(t,u) = 0 \wedge d(s,t) = 2^{-q}$
(3) $t_p \neq s_p$ für $p < q \Rightarrow d(s,t) = 2^{-p} > 2^{-q}$

(4) $t_p \neq s_p$ für $p > q \Rightarrow d(s,t) = 2^{-p} < 2^{-q}$ aber $d(t,u) = 2^{-q}$.

Alle weiteren Fälle ergeben sich hieraus. □

Satz 1.8. *Die Shift-Abbildung $\sigma : \Sigma \to \Sigma$, $\sigma((s_i)) = (s_{i+1})$ ist ein Homöomorphismus auf Σ.*

Beweis. Es ist klar, dass σ bijektiv ist, die Umkehrabbildung ist *der Rechtsshift*. Um zu zeigen, dass σ stetig ist, wähle $\delta := \frac{\varepsilon}{2}$, dann gilt $\sigma(y) \in B_\varepsilon(\sigma(x))$ für alle $y \in B_\delta(x)$. Für den Beweis der Stetigkeit von σ^{-1} setze natürlicherweise $\delta := 2\varepsilon$. □

Definition 1.9. *Ein diskretes dynamisches System welches durch einen metrischen Raum (M,d) und einen Homöomorphismus $f : M \to M$ gegeben ist, heißt sensitiv bzgl. Anfangsbedingungen, falls es ein $c > 0$ gibt, so dass*

$$\forall \varepsilon > 0, x \in M \exists y \in M, n \in \mathbb{Z} : d(x,y) < \varepsilon \wedge d(f^n(x), f^n(y)) > c.$$

Ein solches System heißt chaotisch, falls es (Σ, σ) als Untersystem besitzt, d.h. falls es einen Homöomorphismus h und eine f-invariante Teilmenge Λ gibt, mit $h : \Lambda \to \Sigma$ und $h \circ f = \sigma \circ h$ auf Λ.

Satz 1.10. *Für das diskrete dynamische System (Σ, σ) gilt:*

(1) Die periodischen Bahnen liegen dicht in Σ.
(2) σ ist sensitiv bzgl. Anfangsbedingungen
(3) σ besitzt eine dichte Bahn, d.h. $\exists x \in \Sigma$ mit $\overline{\{\sigma^k(s) \mid k \in \mathbb{Z}\}} = \Sigma$.

Beweis. Triviale *Spielerei* mit der obigen Metrik. □

Wir wollen den Bernoulli-Shift nun noch einmal etwas konkreter (geometrischer) darstellen und betrachten dazu die sogenannte Bäcker-Abbildung

$$T : [0,1]^2 \to [0,1]^2$$
$$(x,y) \mapsto \begin{cases} (2x, \frac{y}{2}) & \text{für } x \leq \frac{1}{2} \\ (2x-1, \frac{y+1}{2} & \text{für } x > \frac{1}{2}. \end{cases}$$

Wir dehnen das Einheitsquadrat (den Teig) also in horizontaler Richtung zu einem $2 \times \frac{1}{2}$-Rechteck, *schneiden* den Teig vertikal in zwei Hälften ($x = \frac{1}{2}$) und platzieren die rechte Hälfte dann genau über der linken. Die Bäcker-Abbildung ist nicht stetig in $x = \frac{1}{2}$, aber überall sonst, d.h. T ist stückweise stetig, sogar stückweise glatt und linear. Die Umkehrabbildung T^{-1} hat die gleichen Eigenschaften (für $y = \frac{1}{2}$).

Satz 1.11. *Die Bäcker-Abbildung ist μ-maßtreu bzgl. dem Lebesgue-Maß auf* $[0,1]^2$.

Beweis. Nach dem Fortsetzungssatz von Kolmogorov können wir uns auf *kleine* Rechtecke, d.h. Rechtecke mit Durchmesser $0 < d < \varepsilon$ beschränken. Betrachten wir den Limes $\varepsilon \to 0$, dann reduziert sich die Bedingung $T_*\mu(A) = \mu(A)$, für alle $A \in \mathcal{I}$, wobei $\mathcal{I}$ ein Erzeugendensystem von $\mathcal{B}$ ist, zu $|\det J_T| = 1$, wobei J_T die Jakobi-Matrix von T ist. In unserem Fall gilt

$$J_T = \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix},$$

also $|\det J_T| = 1$, wie gewünscht. □

Um nun zu sehen, was die Bäcker-Abbildung mit dem Bernoulli-Shift zu tun hat, betrachten wir binäre Brüche. Wir können jede Zahl $x \in [0,1]$ als unendlichen binären Bruch $(a_1a_2a_3\dots)$, mit $a_i \in \mathbb{F}_2$, darstellen. Es gilt also $x = \sum_{i=1}^{\infty} a_i 2^{-i}$.

Betrachten wir nun zwei Punkte $x = (a_1a_2a_3\dots)$ und $y = (b_1b_2b_3\dots)$, so gilt $T(x,y) = (X,Y)$ mit $X = (a2a3\dots)$ und $Y = (a_1b_1b_2\dots)$. Codieren wir (x,y) wie folgt als bi-infinite 0-1-Folge $(\dots b_2b_1.a_1a_2\dots)$, dann ist T nichts anderes als die Shift-Abbildung σ, d.h. $T(x,y) = \sigma(x,y) = (\dots b2b1a1.a2a3\dots)$.

1.3. **Unvorhersehbarkeit.** Beide Systeme, die Billardabbildung und die Bäckerabbildung, sind deterministische dynamische Systeme, in diesem Sinne gibt es also keine *Unvorhersehbarkeiten*, jedoch gibt es einen fundamentalen Unterschied zwischen den beiden Systemen. Da T_θ eine Isometrie des Kreises ist, gilt für $d(x,x_0) < c$, $c \in \mathbb{R}$, auch $d(T_\theta^n(x), T_\theta^n(x_0)) < c$, also können wir die Position von x auch dann noch ziemlich genau bestimmen, wenn wir ein wenig am Anfangspunkt *wackeln.* Die Kreisrotation ist nicht chaotisch.

Die Bäcker-Abbildung ist chaotisch in dem Sinne, dass es unmöglich ist, das Verhalten des dynamischen Systems bei unvollkommenem Wissen vorherzusagen. Hier herrscht das pure Chaos - die Schmetterlinge lassen Grüßen. Angenommen wir möchten Schokoladenbrötchen backen und verteilen deswegen kleine Schokoladenstückchen im Teig und angenommen die beiden Stückchen hatten einen Abstand von etwa $2^{-33} = 10^{-10}$m (Atomdurchmesser). Offensichtlich verdoppelt sich deren Abstand in x-Richtung mit jeder Iteration, d.h. ist $x_0 = x + \varepsilon_x(0)$, dann ist der Abstand nach n Iterationen $\varepsilon_x(n) = 2^n\varepsilon_x(0) = \varepsilon_x(0)e^{n\ln(2)}$ (Liapunuv-Exponent). Unsere Schokoladenstückchen sind also nach nur 33 Iterationen 1m voneinander entfernt. Damit ist klar, dass selbst die einfache Aussage, in welcher der beiden Hälften des Teiges ein bestimmtes Schokoladenstückchen nach 33 Iterationen liegen wird, unmöglich ist. In welcher Hälfte das Stückchen x nach n Iterationen liegen wird, hängt von der n-ten Binärstelle des Anfangswertes ab, d.h. wir müssen die erste Koordinate von x mit einer Präzision von $\frac{1}{2^{n+1}}$ kennen. Kleine Störungen bei den Anfgangswerten, führen also zu unberechenbaren, oder unvorhersehbaren, Folgen.

Um nun zu verstehen was die Entropie eines diskreten dynamischen Systems ist, müssen wir dieses *Chaos* aus einem anderen Blickwinkel betrachten.

Angenommen wir haben nun eine faire Münze 1000 Tage lang je einmal pro Tag geworfen und die Ergebnisse notiert. Wir betrachten die Abbildung $\chi : C^\infty \to \mathbb{F}_2$, $\chi(f) = f(0)$ und obwohl wir also $\chi(T^{-1000})(f)), \chi(T^{-999}(f)), \dots, \chi(T^{-1}(f))$ kennen, haben wir keine Ahnung, was $\chi(f)$ sein wird. Die Vergangenheit liefert uns keine Informationen über das Jetzt oder über die Zukunft des Systems. Intuitiv bedeutet das, dass die Entropie der Shift-Abbildung nicht Null ist.

Anders verhält sich wieder die Billardabbildung. Setzen wir $\theta = \frac{\sqrt{3}}{100} \cong 0.01732\dots$ und betrachten wir die charakteristische Funktion $\chi : [0,1) \to [0,1)$ von $[\frac{1}{2}, 1)$, so können wir, wenn wir $\chi(T_\theta^{-1}(x)) = 0$ und $\chi(x) = 1$ kennen, voraussagen, dass $\chi(T_\theta^k(x)) = 1$ für $1 \leq k \leq 25$. Es ist offensichtlich, dass je mehr Informationen wir über die Vergangenheit haben, desto weiter können wir in die Zukunft des Systems sehen, d.h. im Limes geht die Unvorhersehbarkeit des Systems gegen Null, also ist die Entropie von T_θ gleich Null. Diese Ideen führen uns zu folgender

Definition 1.12. *Die Entropie eines dynamischen Systems ist genau dann Null, wenn die Vergangenheit fast immer die Zukunft entscheidet. Genauer gesagt ist die*

Entropie eines Systems genau dann Null, wenn für jede Partition $\mathcal{A}$ von X (in endlich viele meßbare Teilmengen) eine conull *Teilmenge $X^{'}$ existert, so dass für alle $x, y \in X^{'}$ gilt*

$$\forall k < 0 : T^k(x) \sim \forall k : T^k(y) \Rightarrow T^k(x) \sim T^k(y),$$

mit $x \sim y$, falls es ein $A \in \mathcal{A}$ gibt mit $\{x, y\} \in A$.

Die Entropie eines dynamischen Systems sagt uns also, wie viele neue Informationen wir erhalten, wenn wir wissen, in welchem Teil der Partition $x \in X$ nun liegt.

Zurück zu unseren Beispielen. Es ist nun noch klarer, dass die Entropie der Shift-Abbildung ungleich Null ist. Die Vergangenheit zu kennen, bringt uns keine weiteren Informationen für die Zukunft des Systems. Das gleiche gilt dann natürlich auch für die Bäcker-Abbildung. Betrachte dazu die Partition $\mathcal{A} = \{[0, \frac{1}{2}) \times [0, 1], [\frac{1}{2}, 1] \times [0, 1]\}$. Das Urbild eines horizontalen Streifens liegt immer komplett in einer dieser beiden *Hälften*, d.h. gilt $x \sim y$, so ist $T^k(x) \sim T^k(y)$ für alle $k < 0$. Aber obwohl wir diese Informationen haben, können wir nichts über die nächsten Iterationen voraussagen, es ist sogar $x \not\sim y$, z. B. für $x = (\frac{1}{4}, \frac{1}{2})$ und $y = (\frac{3}{4}, \frac{1}{2})$. Also ist die Entropie der Bäcker-Abbildung nicht Null.

Sei nun $\mathcal{A} = \{I, S \setminus I\}$ für einen beliebigen Kreisbogen I. Wir betrachten wieder T_θ, mit θ irrational, dann wissen wir, dass $\{T_\theta^k(x) \mid k < 0\}$ dicht in S liegt, d.h. wenn $T_\theta^k(x) \sim T_\theta^k(y)$ für alle $k < 0$, so ist $x = y$. Die Vergangenheit definiert also die Zukunft und die Entropie des Kreisbilliard ist Null.

2. Ergodizität

Die Ergodentheorie[2] studiert das Langzeitverhalten von dynamischen Systemen mit Maß bzw. Wahrscheinlichkeitstheoretischen Methoden. Die Ursprünge der Ergodentheorie liegen in der Himmelsmechanik in den Arbeiten von Ludwig Boltzmann und Henri Poincaré. Ziel dieses Kapitels ist es, den Punkteweise-Ergoden-Satz von Birkhoff einzuführen und an Beispielen zu erläutern. Intuitiv sagt der Satz, dass wenn T ergodisch ist, das Mittel über die Zeit eines festen Punktes $x \in X$, d.h. die Häufigkeit wie oft x in einem bestimmten Intervall landet, gleich dem Mittel des Raumes, d.h. dem Maß des Raumes ist. So gilt für das Lebesgue-Maß zum Beispiel

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \chi_[a, b](T^k(x)) = b - a,$$

für fast alle $x \in X$.

Wenn wir uns eine Möwe vorstellen, welche für alle Zeiten *ergodisch* über die Erdoberfläche segelt, dann wissen wir also, dass nach einer *Ewigkeit* die Zeit die die Möwe über einem Gewässer segelt gleich dem Anteil der Wasserfläche[3] auf der Erdoberfläche ist.

Im ersten Vortrag haben wir folgende Definition gesehen.

Definition 2.1. *Sei X eine Mannigfaltigkeit und φ_t ein Fluss auf X. Ein Wahrscheinlichkeitsmaß μ auf X heißt ergodisch, genau dann wenn jede meßbare Abbildung $f : X \mapsto \mathbb{R}$ mit $f \circ \varphi_t = t$, für alle t fast überall konstant ist. In der Sprache*

---

[2] Griech. ergon = Arbeit und odon = Weg

[3] ca. 361,2 Millionen km^2 - etwa 71%

unserer dynamischen Systeme bedeutet das, dass T ergodisch ist, falls für alle meßbaren (oder L^1, L^2, ...) Abbildungen f mit $f \circ T = f$ μ-fast überall gilt, dass f μ-fast überall konstant ist.

Theorem 2.2. *Sei T eine maßtreu Abbildung auf $(X, \mathcal{B}, \mu)$, dann sind folgende zwei Aussagen äquivalent:*

- *T ist ergodisch*
- *Für $A \in \mathcal{B}$: $T^{-1}A = A \Rightarrow \mu(A) = 0 \vee \mu(A) = 1$.*

Ist ein System ergodisch, bedeutet das also, dass jede meßbare T-invariante Menge entweder eine Nullmenge ist oder volles Maß hat. Ergodizität ist also eine Bedingung an die Unzerlegbarkeit von T, ist nämlich das System nicht ergodisch und gilt $T^{-1}A = A$ mit $0 < \mu(a) < 1$, so kann man $T : X \to X$ in $T : A \to A$ und $T : X \setminus A \to X \setminus A$ aufspalten.

Beweis. Die Rückrichtung geht schnell. Sei $A \in \mathcal{B}$ mit $T^{-1}A = A$, dann ist $\chi_A \in L^1(X, \mathcal{B}, \mu)$ und $\chi_A \circ T(x) = \chi_A(x)$ für alle $x \in X$, also ist χ_a μ-fast überall konstant. Da aber χ_A nur die Werte 0 oder 1 annimmt, ist $\chi_A = 0$ μ-fast überall oder $\chi_A = 1$ μ-fast überall. Insgesamt gilt

$$\mu(A) = \int_X \chi_A d\mu = 0 \vee \mu(A) = 1$$

und T ist also ergodisch.

Für die Hinrichtung siehe [Ein11] Proposition 2.14. □

Wir wollen nun den Satz von Birkhoff formulieren und uns dann einige Beispiele anschauen. Für einen Beweis des Satzes und weitere Details, siehe [Ein11] Kapitel 2.6.4.

Theorem 2.3. *Seien μ ein Wahrscheinlichkeitsmaß auf einem lokal-kompakten, separablen metrischen Raum X, φ_t ein ergodischer, maßtreuer Fluss auf X und $f \in L^1(X, \mu)$, dann gilt*

$$\frac{1}{T}\int_0^T f(\varphi_t(x))dt \to \int_X f d\mu, \tag{3}$$

für μ-fast alle $x \in X$.

Definition 2.4. *Ein Punkt $x \in X$ heißt generisch für μ, wenn* (3) *für jede gleichmäßig stetige, beschränkte Abbildung auf X gilt, d. h. $x \in X$ ist generisch, wenn sein Orbit gleichverteilt in X liegt.*

Beispiel 2.5. $0^{\mathbb{Z}}$ *ist kein generischer Punkt für σ. Jeder Punkt $x \in S$ ist ein generischer Punkt für T_θ, für irrationale θ.*

Das folgende Korollar zeigt wieder die Verbindung zwischen ergodisch und gleichverteilt.

Korollar 2.6. *Ist φ_t ein ergodischer Fluss auf (X, μ), so ist fast jeder Punkt $x \in X$ generisch für μ. Ist φ_t nicht ergodisch, dann ist fast kein Punkt $x \in X$ generisch für μ.*

In der Sprache unserer diskreten dynamischen Systeme liest sich Theorem 2.3 wie folgt

Theorem 2.7. *Seien* $(X, \mathcal{B}, \mu)$ *ein Wahrscheinlichkeitsraum,* $T : X \to X$ *eine ergodische maßtreu Abbildung und* $f \in L^1(X, \mathcal{A}, \mu)$*, dann gilt*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) \to \int_X f d\mu,$$

für μ*-fast alle* $x \in X$*.*

Den Zusammenhang mit der bereits in Theorem 1.4 *bewiesenen* Gleichverteilung liefert folgendes

Korollar 2.8. *Seien* $A \in \mathcal{B}$ *und* T *ergodisch, dann gilt für* μ*-fast alle* $x \in X$*, dass die Häufigkeit wie oft der Orbit von* x *in* A *liegt durch* $\mu(A)$ *gebeben wird, d.h.*

$$\lim_{n\to\infty} \frac{k(n)}{n} = \mu(A).$$

Insbesondere sind also μ*-fast alle* $x \in X$ *generische Punkte.*

Betrachten wir wieder unsere beiden Beispiele. Betrachte zunächst die Kreisrotation, o.B.d.A. $T_\theta : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$, $T_\theta(x) = x + \theta \bmod 1$.

Satz 2.9. *Die Kreisrotation* T_θ *ist genau dann ergodisch, wenn* θ *irrational ist.*

Beweis. Sei zuerst $\theta \in \mathbb{Q}$, also $\theta = \frac{p}{q}$, mit $p \in \mathbb{Z}, q \in \mathbb{N}$. Definiere nun

$$f(x) = e^{2i\pi qx} \in L^2(S, \mathcal{A}, \mu),$$

dann ist f nicht konstant, aber es gilt

$$f(T_\theta(x)) = e^{2i\pi q(x+\frac{p}{q})} = e^{2i\pi(qx+p)} = e^{2i\pi qx} = f(x).$$

Also ist T nicht ergodisch.

Seien nun θ irrational und $f \in L^2(S, \mathcal{B}, \mu)$ mit $f \circ T = f$ μ-fast überall und Fourier-Reihe $\sum c_n e^{2i\pi nx}$. Dann hat $f \circ T$ die Fourier-Reihe $\sum c_n e^{2i\pi n\theta} e^{2i\pi nx}$ und Koeffizientenvergleich liefert $c_n = c_n e^{2i\pi n\theta}$, für alle $n \in \mathbb{Z}$. Da aber $\theta \notin \mathbb{Q}$, ist $e^{2i\pi n\theta} \neq 1$, ausser für $n = 0$. Also ist $c_n = 0$ für $n \neq 0$ und somit hat f die Fourier-Reihe c_0, ist also μ-fast überall konstant.

Siehe auch [Ein11] Proposition 2.16 für einen anderen Beweis. □

Satz 2.10. *Die Shift-Abbildung ist ergodisch.*

Beweis. Siehe [Ein11] Proposition 2.15. □

Beide dynamischen Systeme sind also ergodisch, aber wir wissen bereits, dass die Systeme fundamental verschieden sind. Wir wollen uns noch kurz anschauen, wie wir diesen Unterschied in der Sprache der Ergodentheorie erklären können. Wir betrachten dazu zu einem Maßraum $(X, \mathcal{B})$ und einer meßbaren Abbildung $T : X \to X$ den Banachraum der zugehörigen Borelmaße $M(X)$.

Definition 2.11. *Seien* $(X, \mathcal{B})$ *ein Maßraum und* $T : X \to X$ *meßbar. Gibt es ein eindeutiges* T*-invariantes Wahrscheinlichkeitsmaß auf* X*, so heißt* T *eindeutig ergodisch.*

Die eindeutige Ergodizät impliziert eine *strenge Konvergenz* Aussage für stetige Abbildungen:

Theorem 2.12. *Sei* X *ein kompakter metrischer Raum und* $T : X \to X$ *stetig, dann sind folgende zwei Aussagen äquivalent:*

(i) T ist eindeutig ergodisch, d. h. $M(X,T) = \{\mu\}$.
(ii) $\forall f \in C(X)\ \exists c_f \in X$:

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f(T^k x) = c_f, \tag{4}$$

gleichmäßig für $x \in X$.

Beweis. $(i) \Rightarrow (ii)$: Sei also $M(X,T) = \{\mu\}$. Ist (ii) erfüllt, so gilt nach dem Satz von der majorisierten Konvergenz notwendigerweise $c_f = \int f d\mu$. Angenommen (ii) sei falsch, dann gibt es ein $f \in C(x)$ und Folgen $(x_k)_{k\in\mathbb{N}}$, $(n_k)_{k\in\mathbb{N}}$ so, dass

$$\lim_{k\to\infty} \frac{1}{n_k}\sum_{i=0}^{n_k-1} f(T^i x_k) \neq \int f d\mu.$$

Definiere nun für jedes $k \geq 1$ ein Maß $\nu_k \in M(X,T)$ durch

$$\nu_k = \frac{1}{n_k}\sum_{i=0}^{n_k-1} T_*^i \delta_{x_k}$$

so, dass

$$\int f d\nu_k = \frac{1}{n_k}\sum_{i=0}^{n_k-1} f(T^i x_j).$$

Nach dem Satz von Krylov-Bogoloubov (siehe [KB]), gibt es eine Teilfolge von ν_k, welche schwach-* gegen ein Maß $\nu \in M(X,T)$ konvergiert. Genauer gilt

$$\int f d\nu = \lim_{k\to\infty} \int f d\nu_k \neq \int f d\mu.$$

Also ist $\nu \neq \mu$, im Widerspruch zur eindeutigen Ergodizität.

$(ii) \Rightarrow (i)$: Seien μ, ν zwei T-invariante Wahrscheinlichkeitsmaße. Wir wollen zeigen, dass dann $\mu = \nu$ gilt. Integrieren wir (ii), so erhalten wir

$$\begin{aligned}\int f d\mu &= \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \int f \circ T^i d\mu \\ &= \int \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f \circ T^i d\mu \\ &= \int c_f d\mu = c_f\end{aligned}$$

und analog gilt

$$\int f d\nu = c_f.$$

Insgesamt gilt also

$$\int f d\mu = \int f d\nu\ \forall f \in C(x)$$

und mit dem Darstellungssatz von Riesz folgt dann die Behauptung $\mu = \nu$. □

Den Zusammenhang von eindeutiger Ergodizität und generischen Punkten, also der Gleichverteilung, liefert folgender

Satz 2.13. *$T : X \to X$ ist genau dann eindeutig ergodisch, wenn es ein $\mu \in M(X,T)$ gibt, so dass jeder Punkt $x \in X$ μ-generisch ist, d. h. wenn der Orbit jedes Punktes gleichverteilt in X liegt.*

Beweis. Die Rückrichtung folgt wieder sofort aus dem Satz von der majorisierten Konvergenz, denn seien $\nu \in M(X,T)$ und $f \in C(X)$ beliebig, dann ist die linke Seite von (4) durch $\|f\|_\mu$ gleichmäßig beschränkt und konvergiert für jedes $x \in X$ gegen $\int f d\mu$. Da ν T-invariant ist erhalten wir also wie im Beweis oben

$$\int f d\nu = \int \frac{1}{n} \sum_{k=0}^{n_1} f(T^k x) \to \int \int f d\mu d\nu = \int f d\mu,$$

wie gewünscht.

Für die Hinrichtung benutzen wir, dass die linke Seite von (4) gegen eine Konstante c_f konvergiert, denn dann gilt für ein beliebiges $\nu \in M(X,T)$

$$\int f d\nu = \int \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) d\nu \to \int c_f d\nu = c_f.$$

Dies bedeutet aber, dass $M(X,T) = \{\mu\}$ und $c_f = \int f d\mu$. $\square$

Bewaffnet mit diesem Satz, können wir die beiden Systeme wieder unterscheiden. Schon nach Konstruktion ist klar, dass die Shift-Abbildung nicht eindeutig ergodisch sein kann.

Theorem 2.14. *Die Kreisrotation T_θ ist für irrationale θ eindeutig ergodisch und das Lebesgue-Maß μ ist das eindeutige T-invariante Wahrscheinlichkeitsmaß.*

Beweis. Sei ν ein weiteres, beliebiges T-invariantes Wahrscheinlichkeitsmaß, dann gilt

$$\begin{aligned}\int e^{2i\pi kx} d\nu &= \int e^{2i\pi k T(x)} d\nu \\ &= \int e^{2i\pi k(x+\theta)} d\nu \\ &= 2^{2i\pi k\theta} \int e^{2i\pi x} d\nu.\end{aligned}$$

Da θ irrational ist, ist $e^{2i\pi k\theta} \neq 1$ für $k \neq 0$, also gilt

$$e^{2i\pi k} d\nu = 0. \tag{5}$$

Sei nun $f \in C(X)$ beliebig, mit Fourier-Reihe $\sum a_k e^{2i\pi k}$ so, dass $a_0 = \int f d\nu$, dann konvergiert die Folge der Mittel der Partialsummen (linke Seite von (4)) gleichmässig gegen f und wir erhalten

$$\lim_{n\to\infty} \int \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) d\nu = \int f d\nu.$$

Aus (5) folgt damit aber

$$\int \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) d\nu = a_0 = \int f d\mu,$$

also $\int f d\nu = \int f d\mu$, für alle $f \in C(X)$. Ingesamt gilt also $\nu = \mu$ wie gewünscht. $\square$

Literatur

[Ein11] Manfred Einsiedler, *Ergodic theory : With a view towards number theory*, Springer, London, 2011.

[Fel68] William Feller, *An introduction to probability theory and its applications. Vol. I.*, Third edition, Wiley, 1968.

[Fel71] ———, *An introduction to probability theory and its applications. Vol. II.*, Second edition, Wiley, 1971.

[KB] Nicolas Kryloff and Nicolas Bogoliouboff, *La théorie générale de la mesure dans son application à l'étude des systémes dynamiques de la mécanique non linéaire*, Annals of Mathematics **38**.

[Wit03] D. Witte Morris, *Ratner's Theorems on Unipotent Flows*, ArXiv Mathematics e-prints (2003).

TU Dortmund, Fakultät für Mathematik, Institut für Geometrie
E-mail address: gilles.bellot@uni-dortmund.de