Zunächst ein harmlos aussehender Satz, aus dem aber direkt die wichtigsten Ergodensätze folgen:

THEOREM. ***Fundamentaler Ergodensatz:*** *Sei (X, μ) ein Maßraum mit endlichem Gesamtmaß, $T : X \to X$ eine messbare Abbildung, welche T invariant läßt, d.h. für alle messbaren B gilt $\mu(T^{-1}(B)) = \mu(B)$, und sei $B \subset X$ eine (messbare) Teilmenge. Definiere*

$$S_n(x) := \#\{i \in \{0, \dots, n-1\} \mid T^i(x) \in B\},$$

(also die Zahl der „Treffer" in B unter dem Orbitsegment der Länge n mit Startpunkt x), und

$$A_n(x) := \frac{1}{n} S_n(x),$$

(also die relative Häufigkeit, die dieses Orbitsegment in B zubringt).

Dann gilt für μ-fast alle $x \in X$: Der Limes

$$\lim_{n\to\infty} A_n(x)$$

existiert.

BEWEIS. Definiere

$$\overline{A}(x) := \limsup_{n\to\infty} A_n(x),$$

$$\underline{A}(x) := \liminf_{n\to\infty} A_n(x).$$

Die Funktion $\overline{A}$ ist T-invariant, d.h. $\overline{A}(T(x)) = \overline{A}(x)$, und das gilt natürlich auch für $\underline{A}$.

Wähle $\varepsilon > 0$ beliebig (klein) und definiere

$$\tau(x) := \min\{n \in \mathbb{N} \mid A_n(x) \geq \overline{A}(x) - \varepsilon\}.$$

τ hängt natürlich auch von ε ab, aber wir ändern ε bis ganz am Ende von diesem Beweis nicht.

Fall 1: τ ist essentiell beschränkt durch $M < \infty$, d.h. für μ-fast alle $x \in X$ gilt $\tau(x) \leq M$.

Wir betrachten für $n \in \mathbb{N}$ (groß) das Orbitsegment von Länge n ab einem beliebigen $x \in X$, d.h.

$$x, Tx, \dots, T^{n-1}x$$

und stellen fest, dass es $m_1 \leq M$ gibt, so dass $A_{m_1}(x) \geq \overline{A}(x) - \varepsilon$, denn wir können $m_1 = \tau(x)$ wählen. Dann betrachten wir das Orbitsegment von Länge M ab $T^{m_1}x$, d.h.

$$T^{m_1}x, Tx, \dots, T^{m_1+M-1}x$$

und stellen analog fest, dass es $m_2 \leq M$ gibt, so dass $A_{m_2}(T^{m_1}x) \geq \overline{A}(x) - \varepsilon$, denn wir können $m_2 = \tau(T^{m_1}x)$ wählen und es gilt (wie vorhin erwähnt) $\overline{A}(T^{m_1}x) = \overline{A}(x)$. Wir können mit diesem Verfahren fortfahren und das Orbitsegment von Länge n ab x, d.h.

$$x, Tx, \dots, T^{n-1}x$$

mit Teilsegmenten überdecken, auf denen $A_{m_i} \geq \overline{A}(x) - \varepsilon$ ist; das können wir mindestens solange tun, bis das verbleibende Stück

$$T^{m_1+\dots+m_j}x, \dots, T^n x$$

Länge $< M$ hat. Wir haben also vom Orbitsegment der Länge n ein Teil der Länge mindestens $n - M$ abgedeckt. Damit haben wir gezeigt, dass gilt:

$$S_n(x) \geq (n - M)(\overline{A}(x) - \varepsilon)$$

oder äquivalent $A_n(x) \geq \left(1 - \frac{M}{n}\right)(\overline{A}(x) - \varepsilon)$.

Wenn wir das Integral von S_n bezüglich μ über ganz X berechnen, erhalten wir $n\mu(B)$, denn es gilt $S_1 = \chi_B$ (die Indikatorfunktion der Menge B) und somit

$$\int_X S_1(x) d\mu(x) = \mu(B),$$

$S_2 = \chi_B + \chi_{T^{-1}B}$ und wegen T-Invarianz von μ gilt somit $\int_X S_2(x)d\mu(x) = 2\mu(B)$, und analog $\int_X S_n(x)d\mu(x) = n\mu(B)$. Wir erhalten also die Gleichung

$$\mu(B) \geq (1 - \frac{M}{n}) \int_X \overline{A}(x)d\mu(x) - \varepsilon.$$

Hierbei haben wir übrigens verwendet, dass $\mu(X) < \infty$ ist. Durch Grenzübergang $n \to \infty$ folgt

$$\mu(B) \geq \int_X \overline{A}(x)d\mu(x) - \varepsilon.$$

Fall 2: τ ist nicht essentiell beschränkt. Dann nehmen wir die „schlechte Menge" von Punkten x, auf denen $\tau(x) > M$ ist, noch zu B dazu, definieren also eine Menge

$$B' := B \cup \{x \in X \mid \tau(x) > M\}$$

und darauf

$$S'_n(x) := \#\{i \in \{0, \dots, n-1\} \mid T^i(x) \in B'\}.$$

Dann schließen wir: Es gilt analog zu vorhin

$$S'_n(x) \geq (n - M)(\overline{A}(x) - \varepsilon),$$

denn für $x \in X$ gilt: Entweder ist $\tau(x) \leq M$, und dann ist

$$A_{m_1}(x) \geq \overline{A}(x) - \varepsilon$$

für $m_1 = \tau(x) \leq M$. Oder es ist $\tau(x) > M$, und dann ist x schon in der „schlechten Menge", also auch in B', und deshalb ist $S_1(x) = 1$, also sicherlich

$$A_{m_1}(x) = 1 \geq \overline{A}(x) - \varepsilon$$

mit $m_1 = 1$. In beiden Fällen fahren wir dann mit $T^{m_1}x$ anstelle von x fort, wiederholen und erhalten schließlich die gewünschte Abschätzung $S'_n(x) \geq (n-M)(\overline{A}(x) - \varepsilon)$. Daraus schließen wir wieder

$$\mu(B') \geq \int_X \overline{A}(x) d\mu(x) - \varepsilon,$$

und daraus folgt

$$\mu(B) \geq \int_X \overline{A}(x) d\mu(x) - 2\varepsilon.$$

Letztere Gleichung gilt also immer, egal ob Fall 1 oder Fall 2 vorliegt.

Durch Vertauschen von B mit $X \setminus B$ erhalten wir

$$\mu(X \setminus B) \geq \int_X (1 - \underline{A}(x)) d\mu(x) - 2\varepsilon,$$

also

$$\mu(B) \leq \int_X \underline{A}(x) d\mu(x) + 2\varepsilon.$$

Da $\varepsilon > 0$ beliebig war, stellen wir durch Grenzwertübergang $\varepsilon \to 0$ jetzt fest,dass

$$\int_X \overline{A}(x) d\mu(x) \leq \mu(B) \leq \int_X \underline{A}(x) d\mu(x),$$

somit

$$\int_X (\underline{A}(x) - \overline{A}(x)) d\mu(x) \geq 0.$$

Da aber offensichtlich $\underline{A}(x) - \overline{A}(x) \leq 0$ für alle x gilt, muss gelten: $\underline{A}(x) = \overline{A}(x)$ für μ-fast alle $x \in X$. Also existiert der Limes

$$\lim_{n\to\infty} A_n(x)$$

für μ-fast alle $x \in X$, wie gefordert. □

Nun wenden wir uns Sätzen zu, die viel stärker klingen – und auch wirklich aussagekräftig sind – die aber eine direkte Folge aus dem fundamentalen Ergodensatz sind.

DEFINITION. Das System (T, μ) heißt **ergodisch**, wenn für alle (messbaren) und T-invarianten Mengen $B \subset X$ gilt, dass B bezüglich μ Maß 0 hat oder das Komplement von B Maß 0 hat, also $\mu(B) = 0$ oder $\mu(X \setminus B) = 0$ ist. In diesem Fall benutzen wir auch die Sprechweise „T ist ergodisch bezüglich μ".

In der Literatur sieht man oft die Formulierung „T ist ergodisch" ohne Angabe eines Maßes, aber das ist etwas verwirrend und macht nur Sinn, wenn ein ganz bestimmtes Maß dazugedacht ist. Denn Ergodizität hängt von dem Maß ab; wenn T ergodisch ist bezüglich eines Maßes μ_1, dann muss T nicht ergodisch sein bezüglich eines anderen Maßes μ_2. Ebenso findet man häufig die Formulierung „μ ist ergodisch"; offensichtlich macht das auch nur Sinn, wenn ein ganz bestimmtes T dazugedacht ist.

LEMMA. *Folgende Aussagen sind äquivalent:*

(1) (T, μ) *ist ergodisch, d.h. aus* $T^{-1}(B) = B$ *folgt* $\mu(B) = 0$ *oder* $\mu(X \setminus B) = 0$.
(2) *Wenn* $f : X \to \mathbb{R}$ *invariant ist unter* T, *d.h. wenn* $f(T(x)) = f(x)$ *für alle* $x \in X$ *gilt, dann ist* f *essentiell konstant.*
(3) *Der durch* T *gegebene lineare Operator*
$$\hat{T} : L^1(\mu) \to L^1(\mu),$$
der definiert ist durch
$$(\hat{T}(f))(x) := f(T(x)),$$
also
$$\hat{T}(f) := f \circ T,$$
hat keine Eigenfunktionen außer Konstanten.

Die Äquivalenzen sind ganz leicht zu zeigen. Interessant ist hierbei zu bemerken, dass in (2) von „essentiell konstanten" Funktionen und in (3) von „konstanten" Funktionen geredet wird; dies liegt daran, dass eine essentiell konstante Funktion, die in L^1 liegt, in L^1 genau *gleich* einer konstanten Funktion ist, denn L^1 ist ja gerade die Menge von Äquivalenzklassen unter solchen Identifizierungen.