The Abdus Salam
International Centre for Theoretical Physics

United Nations
Educational, Scientific and
Cultural Organization

IAEA
International Atomic Energy Agency

1950-2

School and Workshop on Dynamical Systems

30 June - 18 July, 2008

Ergodic Theorems and Amenable Groups

V.Bergelson
Ohio State University, USA

A. Gorodnik
University of Bristol, UK

LECTURE 1: ERGODIC THEOREMS AND AMENABLE GROUPS

VITALY BERGELSON AND ALEXANDER GORODNIK

Preliminary version

The standard set-up of ergodic theory consists of a measurable space $(X, \mathcal{B})$ equipped with a probability measure μ and a measurable transformation $T : X \to X$. Usually we assume that the transformation is measure-preserving, that is, $\mu(T^{-1}B) = \mu(B)$ for all $B \in \mathcal{B}$. In ergodic theory, one is interested in statistical properties of the orbits $\{T^n x\}_{n\geq 0}$. Origins of this subject can be traced back to the works of Boltzmann, Gibbs and Poincare, and the word "ergodic" comes from so-called ergodic hypothesis in statistical physics. A function $f : X \to \mathbb{R}$ is considered as an observable, which we can sample along the trajectory $\{T^n x\}_{n\geq 0}$ to compute its time average $\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x)$. The ergodic hypothesis of Boltzmann roughly stated that if the space does not split into invariant pieces, then the time average should converge to the space average $\int_X f\, d\mu$. A rigorous version of the ergodic hypothesis is the mean ergodic theorem, which was proved by von Neumann in 40s. We formulate this result in a more general setting of operators on a Hilbert space:

Theorem 0.1 (von Neumann)**.** *Let U be a unitary operator on a Hilbert space $\mathcal{H}$ and P_U denotes the orthogonal projection on the space U-invariant vectors. Then for every $v \in \mathcal{H}$*

$$\frac{1}{N}\sum_{n=0}^{N-1} U^n v \xrightarrow{\|\cdot\|} P_U v \quad \text{as } n \to \infty, \tag{0.1}$$

where P denotes the orthogonal projection on the space of U-invariant vectors.

Proof. We will use the following decomposition:

$$\mathcal{H} = \ker(U^* - I) \perp \overline{\mathrm{Im}(U - I)}, \tag{0.2}$$

which is valid for an operator which is not necessarily unitary. To prove this, we observe that for $v, w \in \mathcal{H}$,

$$\langle v, (U - I)w\rangle = \langle (U^* - I)v, w\rangle .$$

So if v is orthogonal to $\mathrm{Im}(U-I)$, it follows that $(U^*-I)v=0$. Conversely, if $v\in\ker(U^*-I)$, then v is orthogonal to $\mathrm{Im}(U-I)$.

Since U is unitary, $\ker(U-I)=\ker(U^*-I)$. For $v\in\ker(U-I)$, (0.1) is obvious, and for $v=Uw-w$, we have

$$\left\|\frac{1}{N}\sum_{n=0}^{N-1}U^nv\right\|=\left\|\frac{1}{N}(U^Nw-w)\right\|\le\frac{2\|w\|}{N}\to 0\quad\text{as } N\to\infty.$$

Hence, it follows from (0.2) that (0.1) holds for a dense family of vectors. The convergence for general vectors is easy to deduce with a help of the triangle inequality. □

To apply Theorem 0.1 to the study of the dynamical systems, we observe that a measure-preserving transformation $T:X\to X$ defines a unitary operator U_T on the Hilbert space $L^2(X)$:

$$(U_Tf)(x)=f(Tx),\quad f\in L^2(X). \tag{0.3}$$

This simple observation provides a fundamental connection between ergodic theory and harmonic analysis, and it will be crucial for our purposes.

The transformation T is called *ergodic* if $L^2(X)$ contains no nonconstant U_T-invariant functions. Note that in this case, the projection map P_{U_T} is given by $P_{U_T}f=\int_X f\,d\mu$, and we have the following corollary:

Corollary 0.2 (von Neumann mean ergodic theorem). *Let T be an ergodic measure-preserving transformation of a probability space (X,μ). Then for every $f\in L^2(X)$,*

$$\frac{1}{N}\sum_{n=0}^{N-1}f(T^nx)\xrightarrow{L^2}\int_X f\,d\mu\quad \text{as } n\to\infty$$

in L^2-norm.

Here we consider some generalizations of this classical mean ergodic theorem and applications in number theory, Diophantine approximation, and combinatorics.

In these lectures G denotes a topological group. We always assume that

$$\boxed{G \text{ is locally compact and compactly generated.}}$$

A reader might think about some concrete examples: $\mathbb{R}^d$, the group of upper triangular matrices, $\mathrm{SL}_d(\mathbb{R})$, etc.

The group G supports a *right invariant* (regular) Borel measure m, which is called *Haar measure*:

$$m(Bg)=m(B)\text{ for every } g\in G \text{ and Borel } B\subset G.$$

Moreover, this measure is unique up to a scalar multiple. It follows from uniqueness that there exists a function $\Delta : G \to \mathbb{R}^+$, which is called the *modular function*, such that

$$m(gB) = \Delta(g)m(B) \text{ for every Borel } B \subset G.$$

This measure satisfies

$$m(U) > 0 \quad \text{for open } U \subset G,$$
$$m(K) < \infty \quad \text{for compact } K \subset G.$$

We note that in most examples such measure can be given explicitly. It is easy to construct for Lie groups by integrating a nonzero invariant differential form of top degree.

Example 0.3. The Lebesgue measure on $\mathbb{R}^d$ is a Haar measure.

Exercise 0.4.
(1) Show that left/right Haar measure on $\mathrm{GL}_d(\mathbb{R})$ is given by
$$\det(g)^{-d} \prod_{i,j} dg_{ij}$$
(2) Show that left and right Haar measure on
$$\mathrm{SL}_2(\mathbb{R}) = \left\{ \begin{pmatrix} x & y \\ z & t \end{pmatrix} : xt - yz = 1 \right\}$$
is given by $\frac{1}{x} dx dy dz$.
(3) Compute left and right Haar measures for the affine group
$$\left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} : a \in \mathbb{R}^+, b \in \mathbb{R} \right\}. \tag{0.4}$$
Compute the function Δ.
(4) Prove that Δ is a continuous homomorphism.
(5) Prove that if the Haar measure is finite, then G is compact.

A group G is called *unimodular* if Haar measure is both left and right invariant

Exercise 0.5.
(1) Every compact group is unimodular.
(2) $\mathrm{SL}_d(\mathbb{R})$ is unimodular.

We consider a measure-preserving action of a group G on a probability space (X, μ) and aim to prove an ergodic theorem for such actions. As a natural generalization of the von Neumann ergodic theorem, we consider the averages

$$A_n f(x) := \frac{1}{m(B_n)} \int_{B_n} f(g^{-1}x)\, dm(g) \tag{0.5}$$

defined for $f \in L^2(X)$ and a sequence of measurable sets B_n such that $0 < m(B_n) < \infty$. Similarly to (0.3), we can introduce a unitary representation π_X of G on the space $L^2(X)$:

$$\pi_X(g)f(x) = f(g^{-1}x), \quad g \in G,\ f \in L^2(X).$$

The averaging operators (0.5) can be defined for a general unitary representation of G on a Hilbert space $\mathcal{H}$. The operators $A_n : \mathcal{H} \to \mathcal{H}$ are defined by

$$\langle A_n v, w\rangle := \frac{1}{m(B_n)} \int_{B_n} \langle \pi(g)v, w\rangle \; dm(g), \quad v, w \in \mathcal{H}. \tag{0.6}$$

After a short contemplation on the proof of Theorem 0.1, one realizes that the key ingredient was the asymptotic invariance of the intervals $[0, N-1]$ under translations. This leads to the notion of a Følner sequence. We say that a sequence of measurable sets B_n such that $0 < m(B_n) < \infty$ is a *Følner sequence* if for a compact generating set Q of G, we have

$$\sup_{g\in Q} \frac{m(B_n \triangle B_n g)}{m(B_n)} \to 0 \quad \text{as } n \to \infty.$$

A group G is called *amenable* if such a sequence exists.

Exercise 0.6. Prove that for every Følner sequence B_n,

$$\frac{m(B_n \triangle B_n g)}{m(B_n)} \to 0 \quad \text{as } n \to \infty$$

uniformly on g in compact sets. (Hint: use the inclusion $A\triangle C \subset (A\triangle B) \cup (B\triangle C)$ and invariance of the measure.)

Exercise 0.7.
(1) Show that a sequence of boxes $[a_n^{(1)}, b_n^{(1)}] \times \cdots \times [a_n^{(d)}, b_n^{(d)}]$ is a Følner sequence iff $\min_i(b_n^{(i)} - a_n^{(i)}) \to \infty$.
(2) Construct a Følner sequence for the affine group (0.4).

The original argument of von Neumann easily generalizes to prove

Theorem 0.8. *Let B_n be a Følner sequence in a group G and π a unitary representation of G on a Hilbert space $\mathcal{H}$. Then for the sequence of operators A_n defined in (0.6) and every $v \in \mathcal{H}$,*

$$A_n v \xrightarrow{\|\cdot\|} P_G v \quad \textit{as } n \to \infty, \tag{0.7}$$

where P_G is the orthogonal projection on the space of G-invariant functions.

In particular, we immediately obtain the following

Corollary 0.9. *Let B_n be a Følner sequence in a group G that acts ergodically on a probability space (X, μ). Then for every $f \in L^2(X)$,*

$$\frac{1}{m(B_n)} \int_{B_n} f(g^{-1}x)\, dm(g) \xrightarrow{L^2} \int_X f\, d\mu \quad \text{as } n \to \infty. \tag{0.8}$$

Proof of Theorem 0.9. As in the proof of Theorem 0.1, one checks the orthogonal decomposition

$$\mathcal{H} = \mathcal{H}_{inv} \perp \mathcal{H}_{erg}$$

where

$$\begin{aligned} \mathcal{H}_{inv} &= \{v \in \mathcal{H} : \pi(g)v = v \text{ for all } g \in G\}, \\ \mathcal{H}_{erg} &= \overline{\text{span}\{\pi(g)w - w : \, g \in G,\, w \in \mathcal{H}\}}. \end{aligned}$$

For $v \in \mathcal{H}_{inv}$, (0.8) is obvious, and for $v = \pi(g_0)w - w$,

$$\begin{aligned} \|A_n v\| &= \left| \frac{1}{m(B_n)} \int_{B_n} \pi(gg_0)w\, dm(g) - \frac{1}{m(B_n)} \int_{B_n} \pi(g)w\, dm(g) \right| \\ &= \frac{1}{m(B_n)} \left| \int_{B_n g_0} \pi(g)w\, dm(g) - \int_{B_n} \pi(g)w\, dm(g) \right| \\ &\le \frac{1}{m(B_n)} m(B_n g_0 \triangle B_n)\|w\| \to 0 \quad \text{as } n \to \infty. \end{aligned}$$

Taking linear combinations, we obtain a dense family of vectors for which (0.8) holds. Since $\|A_n\| \le 1$ and $\|P_G\| \le 1$, the general case follows by triangle inequality. □

Amenable groups were introduced by von Neumann in relation to the Banach–Tarski paradox. Since then the notion of amenability has found an amazing number of applications. We mention only several characterizations of amenability which appear naturally in our context:

- Existence of invariant means,
- Existence of almost invariant vectors (this will lead to the notion of Kazhdan groups),
- Bounds on the spectrum of the averaging operators (this will lead to the notion of spectral gap).

The important property of amenable groups (which, in fact, characterizes amenability) is existence of invariant measures.

Theorem 0.10 (Bogolubov, Krylov)**.** *Consider a continuous action of an amenable group G on a compact space X. Then X support a G-invariant probability measure.*

Proof. By the Riesz representation theorem, the space $\mathcal{M}(X)$ of Borel measures on X can be identified with the dual of the space $C(X)$ of continuous functions. We equip $\mathcal{M}(X)$ with the weak* topology. Namely, $\mu_n \to \mu$ if

$$\int_X f\, d\mu_n \to \int_X f\, d\mu \quad \text{for every } f \in C(X).$$

The key ingredient of the proof is the compactness of the space $\mathcal{M}_1(X)$ of probability measures, which is a special case of the Banach–Alaoglu theorem. Consider the sequence of the probability measures μ_n defined by

$$\int_X f\, d\mu_n = \frac{1}{m(B_n)} \int_{B_n} f(g^{-1}x)\, dm(g), \quad f \in C(X),$$

where B_n is a Følner sequence. By compactness, we have convergence $\mu_{n_i} \to \mu$ along a subsequence to a probability measure μ. We claim that μ is G-invariant. As in the proof of Theorem 0.9, we have, for $g_0 \in G$ and $f \in C(X)$,

$$\begin{aligned}
&\left| \int_X f(g_0^{-1}x)\, d\mu_n(x) - \int_X f(x)\, d\mu_n(x) \right| \\
=&\frac{1}{m(B_n)} \left| \int_{B_n} f(g_0^{-1}g^{-1}x)\, dm(g) - \int_{B_n} f(g^{-1}x)\, dm(g) \right| \\
\le&\frac{1}{m(B_n)} m(B_n g_0 \triangle B_n) \|f\|_\infty \to 0 \quad \text{as } n \to \infty.
\end{aligned}$$

Hence,

$$\int_X f(g_0^{-1}x)\, d\mu(x) = \int_X f(x)\, d\mu(x)$$

for every $g_0 \in G$ and $f \in C(X)$, as claimed. □

Exercise 0.11.
(1) Consider the action of $\mathrm{SL}_2(\mathbb{R})$ on the space $X = \mathbb{R} \cup \{\infty\}$ defined by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot x = \frac{ax+b}{cx+d}.$$

Show that there are no finite $\mathrm{SL}_2(\mathbb{R})$-invariant measure X. (Hint: the north-south pole dynamics of the diagonal subgroup.)

Note that it follows from Theorem 0.10 that $\mathrm{SL}_2(\mathbb{R})$ is not amenable.

(2) Show that a nonabelean free group, equipped with the discrete topology, is not amenable.

Let π be a unitary representation of a group G on a Hilbert space $\mathcal{H}$. Given subset $Q \subset G$ and $\varepsilon > 0$, we call a vector v a (Q, ϵ)-*invariant* if

$$\|\pi(g)v - v\| \le \varepsilon\|v\| \quad \text{for every } g \in Q.$$

We say that π has *almost invariant vectors* if for every compact subset Q and $\varepsilon > 0$, there exists a nonzero (Q, ϵ)-invariant vector.

Another characterization of amenability is in terms of almost invariant vectors for the regular representation. The *regular representation* λ of a group G is defined by

$$\lambda(g)f(x) = f(xg), \quad g \in G,\ f \in L^2(G).$$

Theorem 0.12 (Hulanicki, Reiter)**.** *A group G is amenable if the regular representation of G has almost invariant vectors.*

Note that $L^2(G)$ contains no invariant vectors when G is not compact.

Lemma 0.13. *Let ϕ, ψ be nonnegative functions in $L^1(G)$. Then for $E_t = \{g \in G : \phi(g) \ge t\}$ and $F_t = \{g \in G : \psi(g) \ge t\}$,*

$$\|\phi - \psi\|_1 = \int_0^\infty m(E_t \triangle F_t)dt.$$

Exercise 0.14. Prove this lemma. (Hint: Observe that by the Fubini theorem, $\phi(x) = \int_0^\infty \chi_{E_t}(x)dt$.)

Proof of Theorem 0.12. Let B_n be a Følner sequence in G and $f_n = \frac{\chi_{B_n}}{m(B_n)^{1/2}}$. Then $\|f_n\|_2 = 1$ and

$$\|\lambda(g)f_n - f_n\|_2 = \frac{m(B_n g^{-1} \triangle B_n)^{1/2}}{m(B_n)^{1/2}} \to 0,$$

uniformly on g in compact sets. Hence, $L^2(G)$ contains almost invariant vectors.

Conversely, let Q be a compact generating set of G and f_n be a sequence of vectors in $L^2(G)$ such that $\|f_n\| = 1$ and

$$\|\lambda(g)f_n - f_n\|_2 \to 0$$

uniformly on $g \in Q$. Consider $\phi_n = f_n^2 \in L^1(G)$. By the Cauchy-Schwartz inequality,

$$\begin{aligned}\|\lambda(g)\phi_n - \phi_n\|_1 &\le \|\lambda(g)f_n + f_n\|_2 \|\lambda(g)f_n - f_n\|_2 \\ &\le 2\|\lambda(g)f_n - f_n\|_2 \to 0.\end{aligned}$$

Let $B_{n,t} = \{g \in G : \phi_n(g) \ge t\}$. Then by Lemma 0.13, we have

$$\|\lambda(g)\phi_n - \phi_n\|_1 = \int_0^\infty m(B_{n,t} g^{-1} \triangle B_{n,t})dt$$

and for every compact $K \subset G$,

$$\alpha_n := \int_0^\infty m(B_{n,t}) \left(\int_K \frac{m(B_{n,t}g^{-1} \triangle B_{n,t})}{m(B_{n,t})} dm(g) \right) dt \to 0.$$

Since $\int_0^\infty m(B_{n,t})dt = \|\phi_n\|_1 = 1$, the measure $d\nu_n(t) = m(B_{n,t})dt$ is a probability measure. For the set

$$\Omega_n = \left\{ t > 0 : \int_K \frac{m(B_{n,t}g^{-1} \triangle B_{n,t})}{m(B_{n,t})} dm(g) \geq \alpha_n^{1/2} \right\},$$

we have $\nu_n(\Omega_n) \leq \alpha_n^{1/2} \to 0$. Hence, there exists t_n such that for $B_n = B_{n,t_n}$, we have $0 < m(B_n) < \infty$ and

$$\int_K \frac{m(B_n g^{-1} \triangle B_n)}{m(B_n)} dm(g) \to 0.$$

This already implies convergence in measure, but we need to prove the uniform convergence!

Let $\varepsilon > 0$, and

$$K_{n,\varepsilon} = \{k \in K : \frac{m(B_n k^{-1} \triangle B_n)}{m(B_n)} \leq \epsilon\}.$$

Note that for $g \in G$, we have

$$m(K_{n,\varepsilon} g \cap K_{n,\varepsilon}) \to m(Kg \cap K) \quad \text{as } n \to \infty, \tag{0.9}$$

uniformly on $g \in G$, and

$$K_{n,\varepsilon}^{-1} \cdot K_{n,\varepsilon} \subset K_{n,2\varepsilon}.$$

Now we assume that Q is a compact symmetric generating set that contains identity, $m(Q) > 0$, and $K = Q^2$. Then for $g \in Q$, we have $m(Kg \cap K) \geq m(Q)$. Then it follows from (0.9) that for $n > n_0(\epsilon)$, we have $K_{n,\varepsilon} g \cap K_{n,\varepsilon} \neq \emptyset$. Hence, $g \in K_{n,2\varepsilon}$. This proves that

$$\frac{m(B_n g^{-1} \triangle B_n)}{m(B_n)} \leq 2\epsilon$$

for every $g \in Q$ and $n \geq n_0(\varepsilon)$ as required. □

It would be convenient to generalize definition (0.6). Given a probability Borel measure μ on G and a unitary representation π of G on a Hilbert space $\mathcal{H}$, we define an operator $\pi(\mu) : \mathcal{H} \to \mathcal{H}$:

$$\langle \pi(\mu)v, w \rangle := \int_G \langle \pi(g)v, w \rangle \, d\mu(g), \quad v, w \in \mathcal{H}.$$

Note that we always have $\|\pi(\mu)\| \leq 1$, and nontrivial upper bounds are very useful.

Theorem 0.15 (Kesten). *Let μ be a probability measure on G which is absolutely continuous with respect to Haar measure, and* $\mathrm{supp}(\mu)$ *generates a dense subgroup of G. Then G is amenable iff the spectrum of the operator $\lambda(\mu)$ contains* 1 *(i.e., the operator $\lambda(\mu) - I$ does not have a bounded inverse).*

Proof. We will prove that the condition on the spectrum is equivalent to the condition from Theorem 0.12. Assume that the regular representation λ contains almost invariant vectors. Let Q_n be a compact subset of G such that $\mu(Q_n) \geq 1 - 1/n$. There exists a sequence of unit vectors $v_n \in L^2(G)$ such that $\|\lambda(g)v_n - v_n\| \leq 1/n$ for all $g \in Q_n$. Then it follows that

$$\begin{aligned}\|\lambda(\mu)v_n - v_n\| &\leq \int_G \|\lambda(g)v_n - v_n\| d\mu(g)\\ &\leq \int_{Q_n} \|\lambda(g)v_n - v_n\| d\mu(g) + \frac{2}{n} \leq \frac{3}{n} \to 0.\end{aligned}$$

This implies that $\lambda(\mu) - I$ does not have a bounded inverse.

Now we assume that $1 \in \mathrm{spec}(\pi(\mu))$. We first claim that there exists a sequence of unit vectors $v_n \in L^2(G)$ such that

$$\|\lambda(\mu)v_n - v_n\| \to 0. \tag{0.10}$$

It suffices to consider the case when $\mathrm{Ker}(\lambda(\mu) - I) = 0$. If $\mathrm{Im}(\lambda(\mu) - I)$ is dense in $L^2(G)$, then the inverse of $\lambda(\mu) - I$ should be unbounded, and this implies (0.10). Otherwise, it follows from (0.2) that there exists $v \in \mathrm{Ker}(\lambda(\mu)^* - I)$, $v \neq 0$, such that

$$\langle \lambda(\mu)v, v\rangle = \langle v, \lambda(\mu)^* v\rangle = \|v\|^2,$$

i.e., we have equality in the Cauchy-Schwartz inequality. Hence $\lambda(\mu)v = v$. This proves (0.10).

Since μ is absolutely continuous, $d\mu(g) = \phi(g)dm(g)$ for some $\phi \in L^1(G)$, so

$$\lambda(\mu)v_n(x) = \int_G v_n(xy)\phi(y)dm(y)$$

and

$$\lambda(g)\lambda(\mu)v_n(x) = \int_G v_n(xy)\phi(g^{-1}y)\Delta(g)^{-1}dm(y).$$

This implies that for every $\varepsilon > 0$, there exists a neighborhood U of identity in G such that

$$\|\lambda(g)\lambda(\mu)v_n - \lambda(\mu)v_n\| < \varepsilon \quad \text{for all } g \in U \text{ and } n \geq 0. \tag{0.11}$$

We set $w_n = \lambda(\mu)v_n/\|\lambda(\mu)v_n\|$. Since $\|\lambda(\mu)\| \le 1$, it follows from (0.10) that

$$\|\lambda(\mu)w_n - w_n\| \to 0,$$

and

$$|\langle\lambda(\mu)w_n, w_n\rangle - 1| = |\langle\lambda(\mu)w_n - w_n, w_n\rangle| \le \|\lambda(\mu)w_n - w_n\| \to 0.$$

Hence,

$$\int_G (1 - \langle\lambda(g)w_n, w_n\rangle)dm(g) \to 0,$$

i.e., $\langle\lambda(g)w_n, w_n\rangle \to 1$ in measure. Passing to a subsequence, we may assume that convergence holds almost everywhere. Since

$$\|\lambda(g)w_n - w_n\|^2 = 2 - 2\mathrm{Re}\,\langle\lambda(g)w_n, w_n\rangle\,,$$

the set

$$H = \{g \in G : \|\lambda(g)w_n - w_n\| \to 0\}$$

has full measure. Then $\mathrm{supp}(\mu) \subset \bar{H}$, and by the assumption on μ, $\bar{H} = G$.

Let $\varepsilon > 0$ and Q be a compact generating set of G. It follows from (0.12) that for a suitable neighborhood U of identity in G,

$$\|\lambda(u)w_n - w_n\| < \varepsilon \quad \text{for all } u \in U \text{ and } n \ge n_0. \tag{0.12}$$

Since H is dense, there exist $h_1, \dots, h_k \in H$ such that $Q \subset \cup_{i=1}^k h_i U$. Note that for all sufficiently large n,

$$\|\lambda(h_i)w_n - w_n\| < \varepsilon \quad \text{for all } i = 1, \dots, k.$$

Writing $g \in Q$ as $g = h_i u$ for some $i = 1, \dots, k$ and $u \in U$, we get

$$\begin{aligned}\|\lambda(g)w_n - w_n\| &\le \|\lambda(h_i u)w_n - \lambda(h_i)w_n\| + \|\lambda(h_i)w_n - w_n\| \\ &\le \|\lambda(u)w_n - w_n\| + \|\lambda(h_i)w_n - w_n\| < 2\varepsilon.\end{aligned}$$

This shows that $L^2(G)$ contains almost invariant vectors, and completes the proof of the theorem. $\square$

Theorem 0.15 implies in particular that for amenable groups, we always have

$$\|\lambda(\mu)\| = 1.$$

This hints that there is no rate of convergence in the mean ergodic theorem. The situation is quite different for semisimple Kazhdan groups such as $\mathrm{SL}_d(\mathbb{R})$, $d \ge 3$, as we will see later.

We say that the unitary representation π of a group G has a *spectral gap* if $1 \notin \mathrm{spec}(\pi(\mu))$ for some absolutely continuous probability measure μ on G whose support generates G topologically. The action of G on a probability space has a *spectral gap* if the representation π_X^0 on the space $L_0^2(X)$, the subspace of functions orthogonal to the constant, has spectral gap.

Exercise 0.16. Show that the notion of spectral gap does not depend on a choice of the measure μ.

Exercise 0.17.
(1) Consider the rotation $T : x \mapsto x + \alpha \mod 1$ on the circle $X = \mathbb{R}/\mathbb{Z}$ and the measures $\mu_n = \frac{1}{N}\sum_{n=0}^{N-1} \delta_{T^i}$ where δ_x denotes the Dirac measure. Show that the spectrum of $\pi_X(\mu_n)$ contains one for all n.
(2*) Let T be an invertible measure-preserving transformation of a general probability measure space (X, μ) and μ_n is defined as above. Show that $\|\pi_X(\mu_n)\| = 1$. (Hint: One can use the notion of Rohklin tower (see, for instance, Halmos' book on ergodic theory).

Exercise 0.18. Let G be a free group with generators a, b and $\mu = \frac{1}{2}\delta_a + \frac{1}{2}\delta_b$. Prove that $\|\lambda(\mu)\| = 1$, but $1 \notin \mathrm{spec}(\pi(\mu))$.

Although the regular representation of an amenable group never has a spectral gap, there are many natural examples of actions of amenable groups with spectral gap.

Exercise 0.19. Consider the action of $G = \mathbb{R}^2$ on $X = \mathbb{R}^2/\mathbb{Z}^2$ by translations. Show that this action has a spectral gap.