ON THE MEAN ERGODIC THEOREM OF SINE

STUART P. LLOYD

Abstract. Robert Sine has shown that \((1/n)(I + T + \cdots + T^{n-1})\), the ergodic averages, converge in the strong operator topology iff the invariant vectors of \(T\) separate the invariant vectors of the adjoint operator \(T^*\), \(T\) being any Banach space contraction. We prove a generalization in which \((\text{spectral radius of } T) \leq 1\) replaces \(||T|| \leq 1\), and any bounded averaging sequence converging uniformly to invariance replaces the ergodic averages; it is necessary to assume that such sequences exist.

1. Introduction. Let \(T: \mathcal{X} \to \mathcal{X}\) be a bounded linear operator of spectral radius \(r(T) \leq 1\) on real or complex Banach space \(\mathcal{X}\), let \(\mathcal{N} = \{x \in \mathcal{X}: Tx = x\}\) be the invariant vectors of \(T\), and let \(\mathcal{N}^* = \{\xi \in \mathcal{X}^*: T^*\xi = \xi\}\) be the invariant vectors of the adjoint operator \(T^*: \mathcal{X}^* \to \mathcal{X}^*\). Robert Sine has shown [4] that in the case \(||T|| \leq 1\), the condition \(\langle \mathcal{N} \text{ separates } \mathcal{N}^* \rangle\) is necessary and sufficient for strong convergence of the ergodic averages \((1/n)(I + T + \cdots + T^{n-1})x, x \in \mathcal{X}\). We show here that \(\langle \mathcal{N} \text{ separates } \mathcal{N}^* \rangle\) is necessary and sufficient for strong convergence of any bounded regular invariant summability method, and that the limiting operator, a projection onto \(\mathcal{N}\), is independent of the method when it exists.

2. Mean ergodic theorem. Let \(\mathcal{A}_1\) be the set of power series \(p(z) = \sum_{i=0}^{\infty} p_i z^i\) which have radius of convergence greater than 1 and are such that \(\sum_{i=0}^{\infty} p_i = 1\). With \(\mathcal{B}(\mathcal{X})\) the bounded linear operators on \(\mathcal{X}\), each \(p \in \mathcal{A}_1\) determines \(p(T) = \sum_{i=0}^{\infty} p_i T^i \in \mathcal{B}(\mathcal{X})\), the series converging in the uniform operator topology. We put \(\mathcal{P}_1 = \{p(T): p \in \mathcal{A}_1\}\), noting that if \(P \in \mathcal{P}_1\) then \(Px = x, x \in \mathcal{N}\).

We introduce further \(\mathcal{P}_1(M) = \{P \in \mathcal{P}_1: ||P|| \leq M\}\) for \(1 \leq M < \infty\), and also \(\mathcal{P}_1(M, \epsilon) = \{P \in \mathcal{P}_1(M): ||(I - T)P|| \leq \epsilon, 1 \leq M < \infty, \epsilon > 0\}\). Our basic assumption will be:

(UI) There exists \(1 \leq M_0 < \infty\) such that
\[\mathcal{P}_1(M_0, \epsilon) \neq \emptyset\text{ for each } \epsilon > 0.\]

In the terminology of Day [1], this is the assertion that \(\mathcal{P}_1\) contains bounded sequences \(\{P_n\}\) converging uniformly to invariance, i.e., \(\lim_n ||P_n - TP_n|| = 0\). In §3 we give various sufficient conditions for (UI).

Let \(\overline{\mathcal{P}_1}\) be the closure of \(\mathcal{P}_1 \subset \mathcal{B}(\mathcal{X})\) in the strong operator topology (SOT), and let \(\overline{\mathcal{P}_1}(M, \epsilon)\) be the SOT closure of \(\mathcal{P}_1(M, \epsilon)\).

Theorem 1. If (UI) holds then (i) is equivalent to (ii):

Presented to the Society, January 26, 1975; received by the editors March 1, 1975.
Key words and phrases. Mean ergodic theorem.

© American Mathematical Society 1976
(1) \(\mathcal{H}\) separates \(\mathfrak{K}\);
(ii) There exists a projection \(Q \in \mathcal{P}_1\) onto \(\mathfrak{K}\), necessarily unique, and
\[\{Q\} = \mathcal{P}_1(M, \varepsilon)\] for any \(M \geq M_0\).

**Proof.** We show first that \(\mathcal{P}_1\) contains at most one projection \(Q\) onto \(\mathfrak{K}\).
Suppose \(\lim_{n \to \infty} p_n = Q\) in SOT with \(\{p_n\} \subset \mathcal{P}_1\) a generalized sequence and \(Q\) a projection onto \(\mathfrak{K}\). Such a projection commutes with \(T\), from
\[Q = TQ = T \lim_{n \to \infty} p_n = \lim_{n \to \infty} TP_n = \lim_{n \to \infty} p_n T = QT,
the limits being SOT. If \(\lim_{n \to \infty} p_n' = Q'\) and \(\lim_{n \to \infty} p_n'' = Q''\), then \(Q' = Q'p_n'' = Q''\), whence \(Q\) is unique in \(\mathcal{P}_1\) if it exists.

(i) \(\Rightarrow\) (ii) Let us prove that if projection \(Q\) exists then the adjoint projection \(Q^* : \mathfrak{K}^* \to \mathfrak{K}^*\) is onto \(\mathfrak{K}\). From \(Q = TQ = QT\) follows \(Q^* = Q^* T^* = T^* Q^*\), so that \(Q^*\) has range in \(\mathfrak{K}\). If \(Q = \lim_{n \to \infty} p_n\) with \(\{p_n\} \subset \mathcal{P}_1\) and if \(x \in \mathfrak{K}, \xi \in \mathfrak{K}\), then
\[(x, \xi) = (x, p_n^*\xi) = (p_n x, \xi) = (Q x, \xi) = (x, Q^* \xi),
giving \(Q^* \xi = \xi, \xi \in \mathfrak{K}\), whence \(Q^*\) is onto \(\mathfrak{K}\).

Suppose \(0 \neq \xi \in \mathfrak{K}\), and let \(x \in \mathfrak{K}\) be such that \((x, \xi) \neq 0\). Then \((Q x, \xi) = (x, Q^* \xi) = (x, \xi) \neq 0\), so that \(Q x \in \mathfrak{K}\) separates \(\xi\) and \(0\). That is, \(\mathfrak{K}\) separates \(\mathfrak{K}\).

(ii) \(\Rightarrow\) (i) Let functional \(\Phi(x), x \in \mathfrak{K}\), be defined by
\[\Phi(x) = \lim_{\varepsilon \to 0} \sup_{P', P'' \in \mathcal{P}_1(M, \varepsilon)} \|P'x - P''x\|,\]
for any fixed \(M \geq M_0\). The properties
(a) \(\Phi(x + y) \leq \Phi(x) + \Phi(y), x, y \in \mathfrak{K},\)
(b) \(\Phi(cx) = |c|\Phi(x), x \in \mathfrak{K},\) scalar \(c,\)
(c) \(0 \leq \Phi(x) \leq 2M\|x\|, x \in \mathfrak{K},\)
(d) \(\Phi(x) = 0, x \in \mathfrak{K},\)
(e) \(\Phi(x - Tx) = 0, x \in \mathfrak{K},\)
are straightforward or obvious; e.g. for (e) we use
\[\|P'(x - Tx) - P''(x - Tx)\| \leq \|(I - T)P'|| + \|(I - T)P''\|| \leq 2\varepsilon\|x\|, \quad x \in \mathfrak{K}\] and \(P', P'' \in \mathcal{P}_1(M, \varepsilon).\)

Let \(\Phi^0\) denote the set of all \(\xi \in \mathfrak{K}^*\) satisfying \((x, \xi) \leq \Phi(x), x \in \mathfrak{K};\) by the Hahn-Banach theorem, \(\Phi(x) = \max_{\xi \in \Phi^0} (x, \xi), x \in \mathfrak{K}.\) Properties (d), (e) of \(\Phi\) yield properties: (d*) \(\xi \in \mathfrak{K}^*\); (e*) \(\xi \in \mathfrak{K},\) for the \(\xi \in \Phi^0,\) where \(\mathfrak{M}^\perp \subset \mathfrak{K}^*\) is the annihilator of \(\mathfrak{K}.\) Now, condition (1)(i) is the assertion \(\mathfrak{M}^\perp \cap \mathfrak{K} = 0;\) from \(\Phi^0 \subset \mathfrak{M}^\perp \cap \mathfrak{K}\) just shown follows then \(\Phi^0 = 0\) and hence \(\Phi = 0.\)

Let \(\{\eta_n\}\) be any sequence of positive numbers such that \(\lim_{n \to \infty} \eta_n = 0,\) let \(P_n\) for each \(n\) be an arbitrarily chosen member of \(\mathcal{P}_1(M, \eta_n),\) and put \(\varepsilon_n = \max_{r \geq n} \eta_r.\) Any such sequence \(\{P_n\}\) is SOT Cauchy when \(\Phi = 0,\) from
\begin{align*}
\lim_{n \to \infty} \sup_{n \geq r < s} \|P_n x - P_s x\| &\leq \lim_{n \to \infty} \sup_{P', P'' \in \mathcal{P}_1(M, \varepsilon)} \|P' x - P'' x\| \\
&= \lim_{\varepsilon \to 0} \sup_{P', P'' \in \mathcal{P}_1(M, \varepsilon)} \|P' x - P'' x\| \\
&= \Phi(x) = 0, \quad x \in \mathcal{X};
\end{align*}

we have used the nested property of the \(\{\mathcal{P}_1(M, \varepsilon) : \varepsilon > 0\}\). The SOT limit \(Q = \lim_n P_n\) is the projection sought. \(\square\)

We remark that when \(\mathcal{H} = 0\) the result takes the form: given (UI), \(0 \in \mathcal{F}_1\) iff \(\mathcal{H} = 0\).

3. Condition (UI). Consider the familiar ergodic averages \(A_n = (1/n) \sum_{i=0}^{n-1} T^i\), \(n \geq 1\), for which \((I - T)A_n = (I - T^n)/n\). If these satisfy \(\|A_n\| \leq M_0 < \infty\), \(n \geq 1\), and if \(\lim inf_n \|T^n\|/n = 0\), then (UI) holds, clearly. A fortiori, (UI) holds when \(\|T^n\| \leq M_0 < \infty\), \(n \geq 1\); this case is covered by the arguments of [4], although Sine assumes \(\|T\| \leq 1\).

The resolvent of \(T\) is given by \(R_\lambda = \sum_{i=0}^{\infty} T^i/\lambda^{i+1}\) when \(|\lambda| > 1 \geq r(T)\), the series converging in the uniform operator topology. If we introduce \(P_\lambda = (\lambda - 1)R_\lambda\) for \(|\lambda| > 1\), then \(P_\lambda \in \mathcal{P}_1\), and it is easily verified that \((I - T)P_\lambda = (\lambda - 1)(I - P_\lambda)\). Thus if the \(\{P_\lambda\}\) satisfy

\[
\lim_{\lambda \to 1; |\lambda| > 1} \inf \|P_\lambda\| < M_0 < \infty,
\]

then (UI) holds for such \(M_0\).

In the other direction, suppose \(T\) has index \(1 < \mu < \infty\) at \(\lambda = 1\). (The index of \(T\) at \(\lambda = 1\) is the least integer \(\mu \geq 0\) with the property: all vectors \(x \in \mathcal{X}\) satisfying \((I - T)\mu X = 0\) satisfy also \((I - T)x = 0\) [2, p. 556].)

**THEOREM 2.** If \(T\) has index \(1 < \mu < \infty\) at \(\lambda = 1\) then (UI) cannot hold and \(\mathcal{P}_1\) contains no projection onto \(\mathcal{H}\).

**PROOF.** If a generalized sequence \(\{P\} \subset \mathcal{P}_1\) is SOT convergent to a projection \(Q \in \mathcal{F}_1\) onto \(\mathcal{H}\), then, necessarily,

\[
\lim_{\nu} \|(I - T)P_{\nu} x\| = \|(I - T)Q x\| = 0, \quad x \in \mathcal{X}.
\]

If \(T\) has index \(1 < \mu < \infty\) at \(\lambda = 1\), then unit vectors \(x_1, x_2 \in \mathcal{X}\) exist such that \(T x_1 = x_1, T x_2 = x_2 + c x_1\) for some \(c > 0\). From

\[
(I - T)T^i x_2 = T^i(I - T)x_2 = T^i(-c x_1) = -c x_1, \quad i \geq 0,
\]

follows \((I - T)P x_2 = -c x_1, P \in \mathcal{P}_1\). We have then \(\|(I - T)P x_2\| = c > 0\), \(P \in \mathcal{P}_1\), showing that no \(Q\) exists, and \(\|(I - T)P\| \geq c > 0\), \(P \in \mathcal{P}_1\), showing that (UI) fails. \(\square\)

For the same \(x_1, x_2\) an easy calculation gives \(P_\lambda x_2 = x_2 + c x_1/(\lambda - 1)\), whence

\[
\|P_\lambda\| \geq |c/|\lambda - 1|-1|, \quad |\lambda| > 1,
\]

whence \(\Phi(x) = 0, \quad x \in \mathcal{X}\);
when \( T \) has index \( 1 < \mu < \infty \) at \( \lambda = 1 \). Thus if some condition on \( \|P_\lambda\|, |\lambda| > 1, \lambda \to 1 \), is necessary and sufficient for (UI) then it lies between (2) and

\[
\|P_\lambda\| = o(1/|\lambda - 1|), \quad |\lambda| > 1, \lambda \to 1.
\]

Apart from changes in variable, the following result is the \((C, \alpha)\) generalization of Theorem 1' of [3]. For \( \alpha \neq -1, -2, \ldots \) the \((C, \alpha)\) averages \( A_n^{(\alpha)}(T) \) of \( \{T^n\} \) have as generating function

\[
\left(1 - \frac{1}{\lambda}\right)^{-\alpha} R_\lambda(T) = \sum_{n=0}^{\infty} \frac{\alpha + n!}{\alpha! n!} \frac{A_n^{(\alpha)}(T)}{\lambda^{\alpha+1}}, \quad |\lambda| > 1,
\]

so that

\[
A_n^{(\alpha)}(T) = \frac{\alpha! n!}{\alpha + n!} \frac{1}{2\pi i} \int \lambda^{\alpha} \left(1 - \frac{1}{\lambda}\right)^{-\alpha} R_\lambda(T) d\lambda, \quad n \geq 0,
\]

the contour being a large circle around the origin, say.

For \( \lambda \) in the resolvent set of \( T \) let \( T_\lambda \in \mathbb{B}(\mathcal{H}) \) be defined by

\[
T_\lambda = P_\lambda T = (\lambda - 1)(\lambda I - T)^{-1} T = \lambda P_\lambda - (\lambda - 1)I.
\]

The spectrum of \( T_\lambda \) is \( \sigma(T_\lambda) = \{(\lambda - 1)\lambda^'/(\lambda - \lambda^'): \lambda^' \in \sigma(T)\} \), and we find \( 1 - \lambda \notin \sigma(T_\lambda) \) provided \( \lambda \neq 0, 1 \). Thus the inversion formula

\[
T = \lambda T_\lambda[(\lambda - 1)I + T_\lambda]^{-1}
\]

is valid for \( \lambda \neq 0, 1 \) in the resolvent set of \( T \); moreover, \( r(T_\lambda/(1 - \lambda)) < 1 \) if \( |\lambda| > 2 \). Note that the relations between \( T_\lambda \) and \( T \) admit the involution \( T \leftrightarrow T_\lambda, \lambda \leftrightarrow 1 - \lambda \).

**Theorem 3.** For any fixed \( \alpha \geq 0, 1 \leq M < \infty \), the conditions

(i) \[
||A_n^{(\alpha)}(T)|| \leq M, \quad n \geq 0,
\]

(ii) \[
||A_n^{(\alpha)}(T_\lambda)|| \leq M \left|\frac{\lambda - 1}{|\lambda| - 1}\right|^{\alpha+n}, \quad n \geq 0, |\lambda| > 1,
\]

are equivalent. When they are satisfied, (UI) holds and

\[
||P_\lambda(T)|| \leq M(|\lambda - 1|/(|\lambda| - 1)^{\alpha+1}, \quad |\lambda| > 1.
\]

**Proof.** In

\[
A_n^{(\alpha)}(T_\lambda) = \frac{\alpha! n!}{\alpha + n!} \frac{1}{2\pi i} \int \mu^{\alpha} \left(1 - \frac{1}{\mu}\right)^{-\alpha} R_\mu(T_\lambda) d\mu, \quad n \geq 0,
\]

we have

\[
R_\mu(T_\lambda) = \frac{1}{\mu I - T_\lambda} = \frac{1}{\lambda + \mu - 1} \left(I + \frac{\lambda(\lambda - 1)}{\lambda \mu I - (\lambda + \mu - 1)T}\right);
\]

the change of variable \( \xi = \lambda \mu/(\lambda + \mu - 1) \) gives
\[ A_n^{(a)}(T) \]
\[ = \frac{\alpha! n!}{\alpha + n!} \left( 1 - \frac{1}{\lambda} \right)^{\alpha + n} \frac{1}{2\pi i} \int \left[ \frac{\xi}{1 - (\xi/\lambda)} \right]^n \left( 1 - \frac{1}{\lambda} \right)^{-\alpha} \left[ \frac{I}{\xi - \lambda} + R_\xi(T) \right] d\xi \]
\[ = \left( 1 - \frac{1}{\lambda} \right)^{\alpha + n} \sum_{j=0}^{\infty} \frac{\alpha + n - 1 + j!}{\alpha + n!} \left\{ \frac{\alpha j}{(\alpha + n)(j + n)} \right\} \left[ 1 - \frac{\alpha j}{(\alpha + n)(j + n)} \right] A_n^{(a)}(T) \}
\[ n \geq 0, |\lambda| > 1, \]

with \( 1 < |\xi| < |\lambda| \) on the contour. If \( \alpha \geq 0 \) then \( 0 \leq \alpha/(\alpha + n) \cdot j/(j + n) \)
\( \leq 1 \) in the last expression, so if (3)(i) holds, then

\[ \|A_n^{(a)}(T_\lambda)\| \leq \left| 1 - \frac{1}{\lambda} \right|^{\alpha + n} \left( 1 - \frac{1}{|\lambda|} \right)^{-\alpha - n} M \]
\[ = \left| \frac{\lambda - 1}{|\lambda| - 1} \right|^{\alpha + n} M, \quad |\lambda| > 1, \]

which is (3)(ii).

In the same way, using the bound (3)(ii) in the inversion formula

\[ A_n^{(a)}(T) = \left( 1 - \frac{1}{\lambda} \right)^{-\alpha - n} \]
\[ = \sum_{j=0}^{\infty} \frac{\alpha + n - 1 + j!}{\alpha + n!} \left\{ \frac{\alpha j}{(\alpha + n)(j + n)} \right\} \left[ 1 - \frac{\alpha j}{(\alpha + n)(j + n)} \right] A_n^{(a)}(T) \}
\[ n \geq 0, |\lambda| > 2, \]

gives

\[ \|A_n^{(a)}(T)\| \leq \left( |\lambda|/(|\lambda| - 2) \right)^{\alpha + n} M, \quad n \geq 0, |\lambda| > 2; \]

we let \( |\lambda| \to \infty \) to obtain (3)(i).

If conditions (3) are satisfied then

\[ \|R_\lambda(T)\| \leq \left[ \frac{|\lambda - 1|}{|\lambda|} \right]^a \left[ 1 - \frac{1}{|\lambda|} \right]^{-a - 1} M \]
\[ = \frac{|\lambda - 1|^a M}{(|\lambda| - 1)^{a+1}}, \quad |\lambda| > 1, \]

which is (4). For a (UI) sequence we may take \( \{P_n\} \) for some \( \{\lambda_n \downarrow 1\} \), or \( \{A_n^{(a+1)}(T)\} \), since

\[ A_n^{(a+1)}(T) = \sum_{j=0}^{n} \frac{\alpha + j!}{\alpha! j!} A_j^{(a)}(T) / \sum_{j=0}^{n} \frac{\alpha + j!}{\alpha! j!}, \]
\[ (I - T)A_n^{(a+1)}(T) = (\alpha + 1)/(n + 1)[I - A_n^{(a)}(T)], \quad n \geq 0. \]
Note the slight strengthening of the result of the first paragraph: if \( \|A_n^{(1)}(T)\| \leq M < \infty, n \geq 0 \), then (UI) holds, with no condition on \( \|T^n\|/n \). All that is involved in Theorem 3 is \( \|T - T_\lambda\| = O(1/|\lambda|) \) at \( |\lambda| \to \infty \), of course; the interesting part at \( \lambda \to 0 \) has no force.

4. The adjoint projections. Recall that a generalized sequence \( \{\Gamma_x\} \subset \mathfrak{B}(\mathfrak{X}^*) \) is W* OT convergent to \( \Gamma \in \mathfrak{B}(\mathfrak{X}^*) \) iff \( \lim_{\nu}(x, \Gamma_x, \xi) = (x, \Gamma \xi) \) for each \( x \in \mathfrak{X} \), \( \xi \in \mathfrak{X}^* \); and further, that bounded W* OT closed sets are W* OT compact. Let \( \mathfrak{P}_1(M, \varepsilon) \subset \mathfrak{B}(\mathfrak{X}^*) \) be the set of adjoints of members of \( \mathfrak{P}_1(M, \varepsilon) \), and let \( \overline{\mathfrak{P}_1}(M, \varepsilon) \) denote the W* OT closure of \( \mathfrak{P}_1(M, \varepsilon) \). If (UI) holds and \( M \supseteq M_0 \), then \( \overline{\mathfrak{P}_1}(M, \varepsilon) \) is a nested family of nonempty convex W* OT compact sets; the intersection \( \mathfrak{S}(M) = \cap_{\varepsilon > 0} \overline{\mathfrak{P}_1}(M, \varepsilon) \), necessarily nonempty, consists of projections onto \( \mathfrak{N} \) (with \( \mathfrak{S}(M) = \{0\} \) if \( \mathfrak{N} = 0 \)). The members of \( \mathfrak{S}(M) \) commute with \( T^* \) and satisfy \( \Gamma \Gamma'' = \Gamma'' \Gamma, \Gamma'' \in \mathfrak{S}(M) \), clearly. If the projection \( Q \in \mathfrak{P}_1 \) onto \( \mathfrak{N} \) exists, then it satisfies \( Q^* \Gamma = Q^* \), \( \Gamma \in \mathfrak{S}(M) \). For, suppose \( \lim_{\nu} P_\nu^* = \Gamma \in \mathfrak{S}(M) \) in W* OT; then

\[
(x, Q^* \Gamma \xi) = (Qx, \Gamma \xi) = \lim_{\nu} (Qx, P_\nu^* \xi) = \lim_{\nu} (P_\nu Qx, \xi)
\]

\[
= (Qx, \xi) = (x, Q^* \xi), \quad x \in \mathfrak{X}, \xi \in \mathfrak{X}^*.
\]

This and \( Q^* \Gamma = \Gamma \) give \( \Gamma = Q^* \), which is to say, \( \mathfrak{S}(M) \) is the singleton \( \mathfrak{S}(M) = \{Q^*\} \) when \( Q \) exists. The author is not able to prove the following plausible converse: If \( \mathfrak{S}(M) = \{\Gamma\} \) is a singleton, then \( \Gamma = Q^* \) for \( Q \in \mathfrak{P}_1 \) a projection onto \( \mathfrak{N} \).

References


Bell Laboratories, Murray Hill, New Jersey 07974