The mean ergodic theorem of von Neumann:
a very elementary proof

Let \( T : H \to H \) be any map on a Hilbert space and let

\[ \text{Fix}(T) = \{ x \in H : Tx = x \} \]

be the fixed point set of \( T \). If \( x \in \text{Fix}(T) \), then the iterates \( T^k x \) are all equal to \( x \), hence \( T^k x \to x \). On the other hand, if \( T \) is continuous and \( (T^n x) \) converges to some \( y \), then \( Ty = T(\lim_n T^n x) = \lim_n T(T^n x) = \lim_n T^{n+1} x = y \), so \( y \in \text{Fix}(T) \).

But, even for a unitary operator \( T \), it can happen that the sequence \( (T^n x) \) converges only in the trivial case \( x = 0 \). Example: the bilateral shift.

The situation is much better if we take averages:

**Theorem 1** Let \( T \in B(H) \) be a contraction. If

\[ S_n = \frac{1}{n} \sum_{k=0}^{n-1} T^k \quad (n = 0, 1, \ldots) \]

are the averages of the iterates \( T^k \) of \( T \), then

(i) \( (S_n) \) converges strongly (i.e. pointwise) and
(ii) its limit is the orthogonal projection \( F \) onto the fixed point set

\[ \text{Fix}(T) = \ker(I - T) = \{ x \in H : Tx = x \} \].

**Proof.** (a) Suppose first that \( x = (I - T)(H) \), hence there exists \( y \in H \) with \( x = (I - T)y \). Then for each \( k \in \mathbb{Z}_+ \) we have \( T^k x = T^k y - T^{k+1} y \), therefore

\[ S_n x = \frac{1}{n} (y - T^n y) \]

hence

\[ \| S_n x \| \leq \frac{1}{n} \| y - T^n y \| \leq \frac{2 \| y \|}{n} \to 0. \]

Thus \( S^n x \to 0 \) for all \( x = (I - T)(H) \).

(b) It follows that for all \( x \in (I - T)(H) \) we have \( S^n x \to 0 \). Indeed given \( \varepsilon > 0 \) choose \( z = (I - T)y \in (I - T)(H) \) so that \( \| x - z \| < \varepsilon \), and then choose \( n_0 \in \mathbb{N} \) such that \( \| S_n z \| < \varepsilon \) for all \( n \geq n_0 \).

If \( n \geq n_0 \) then, since each \( S_n \) is a contraction,

\[ \| S_n x \| \leq \| S_n (x - z) \| + \| S_n z \| \leq \| x - z \| + \| S_n z \| < 2 \varepsilon. \]

(c) It remains to consider the case \( x \in (I - T)(H) \) \( \bot = \ker(I - T^*) \), i.e. \( x = T^* x \).

But then \( x = Tx \): indeed

\[ \| x - Tx \|^2 = \| x \|^2 + \| Tx \|^2 - 2 \text{Re} \langle x, Tx \rangle = \| x \|^2 + \| Tx \|^2 - 2 \text{Re} \langle T^* x, x \rangle = \| x \|^2 + \| Tx \|^2 - 2 \| x \|^2 \leq 0 \]

because \( T \) is a contraction; hence \( \| x - Tx \|^2 = 0 \).

Thus \( x \in \text{Fix}(T) \) and so, as noted above, \( S_n x = x \) for all \( n \), hence \( \lim_n S_n x = x \). Therefore for all \( x \in H \),

\[ \lim_n S_n x = \lim_n S_n F x + \lim_n S_n F^* x = F x + 0. \quad \square \]

See also the very interesting blog, [Terry Tao: The mean Ergodic Theorem](http://terrytao.wordpress.com/).