Spectral theorem for self-adjoint continuous operators on Hilbert spaces

Paul Garrett  garrett@math.umn.edu  http://www.math.umn.edu/~garrett/

[This document is http://www.math.umn.edu/~garrett/m/fun/notes_2016-17/spectral_theorem_bdd.pdf]

1. Spectral theorem, part one
2. Schur’s lemma and other corollaries
3. Spectral theorem, part two: projectors
4. Appendix: goofy lemma on polynomials
5. Appendix: Tietze-Urysohn-Brouwer extension theorem
6. Appendix: Urysohn’s lemma
7. Appendix: notes

1. Spectral theorem, part one

The isomorphism of the theorem is a special, concrete case of the Gelfand isomorphism.

Let T be a continuous linear map V → V for a (separable) Hilbert space V. Its spectrum σ(T) is a compact subset of ℝ, so is certainly contained in some finite interval [a,b]. As usual, for a self-adjoint continuous operator S on V, write S ≥ 0 when ⟨Sv,v⟩ ≥ 0 for all v ∈ V. For self-adjoint S,T, write S ≤ T when T − S ≥ 0. At the outset, with a ≤ −|T|op and b ≥ |T|op, we have, ⟨a · v,v⟩ ≤ ⟨Tv,v⟩ ≤ ⟨b · v,v⟩. That is, a ≤ T ≤ b, where the scalars refer to scalar operators on V. As a corollary of part one of the spectral theorem, we will see that, in fact, inf_{λ∈σ(T)} λ ≤ T ≤ sup_{λ∈σ(T)}, but it seems difficult (and un-necessary) to prove this inequality directly.

In the following all functions are real-valued, so C₀[a,b] refers to real-valued continuous functions on [a,b].

[1.1] Theorem: The map ℝ[x] → ℝ[T] on polynomials given by f → f(T) is continuous for ℝ[x] with the sup-norm on [a,b], and ℝ[T] with the operator norm. Thus, by Weierstraß approximation, this map extends to a continuous map C₀[a,b] → ℝ[T], the latter being the operator-norm completion of ℝ[T]. This map factors through C₀(σ(T)):

\[ C₀[a,b] \rightarrow C₀(σ(T)) \rightarrow ℝ[T] \]

and the map C₀(σ(T)) → ℝ[T] is an isometric isomorphism, where C₀(σ(T)) has sup-norm.

Proof: We claim that for f ∈ ℝ[x] with f(x) ≥ 0 on [a,b], then f(T) ≥ 0. From the goofy lemma on polynomials, f is expressible as a finite sum of the form

\[ f = \sum_i P_i^2 + (x-a) \sum_j Q_j^2 + (b-x) \sum_k R_k^2 \]

for polynomials P_i, Q_j, R_k in ℝ[x]. We have

[1.2] Lemma: For commuting self-adjoint S,T with T ≥ 0, also S²T ≥ 0.

Proof: ⟨S²Tv,v⟩ = ⟨TSv,S⁻¹v⟩ = ⟨S(Sv),v⟩ ≥ 0. ///

Thus, since a ≤ T ≤ b, and all these operators commute (being polynomials in T), each P_i²(T) ≥ 0, each (T − a)Q_j²(T) ≥ 0, and (b − T)R_k²(T) ≥ 0. Thus, f(T) ≥ 0, proving the claim.

Since g(x) = sup_{[a,b]} |f| ± f(x) ≥ 0 on [a,b], sup_{[a,b]} |f| ± f(T) ≥ 0. That is, −sup_{[a,b]} |f| ≤ f(T) ≤ sup_{[a,b]} |f|, which gives

\[ |f(T)|_{op} = \sup_{|v| ≤ 1} |f(T)v| ≤ \sup_{|v| ≤ 1} \sup_{[a,b]} |f| · |v| = \sup_{[a,b]} |f| \]
which is the desired inequality. Thus, we can extend by continuity to the sup-norm closure of \( \mathbb{R}[x] \) in \( \mathcal{C}^o[a, b] \), which by Weierstraß is the whole \( \mathcal{C}^o[a, b] \), giving \( \mathcal{C}^o[a, b] \to \mathbb{R}[T] \), the latter being the operator-norm closure of \( \mathbb{R}[T] \), with \( |f(T)|_{\text{op}} \leq |f|_{\mathcal{C}^o[a, b]} \). Since \( \mathbb{R}[x] \to \mathbb{R}[T] \) is a ring homomorphism, the extension by continuity is also a ring homomorphism.

We capture some useful partial results:

**[1.3 Corollary]**: (*Existence of square roots of positive operators*) For \( T \geq 0 \), there is \( S \in \mathbb{R}[T] \) such that \( S \geq 0 \) and \( S^2 = T \).

**Proof**: Since \( T \geq 0 \), without yet claiming anything about the spectrum of \( T \), we can take \([a, b] = [0, b]\) in the previous discussion. The function \( f(x) = \sqrt{x} \in \mathcal{C}^o[0, b] \) is non-negative on \([0, b]\), and \( f(T)^2 = f^2(T) = T \). Take \( S = f(T) \).

**[1.4 Corollary]**: (*Positivity of products of commuting positive operators*) For \( S \geq 0 \) and \( T \geq 0 \) with \( ST = TS \), also \( ST \geq 0 \).

**Proof**: From the previous corollary, there is \( R \in \mathbb{R}[S] \) such that \( R \geq 0 \) and \( R^2 = S \). Also, \( R \) commutes with \( T \), by continuity. Thus,

\[
\langle STv, v \rangle = \langle R^2Tv, v \rangle = \langle RTTv, v \rangle = \langle TTv, v \rangle \geq 0
\]

because \( T \geq 0 \).

Next, let \( I \) be the kernel of \( \mathcal{C}^o[a, b] \to \mathbb{R}[T] \). It is an *ideal* in \( \mathcal{C}^o[a, b] \), and is (topologically) *closed* because \( \mathcal{C}^o[a, b] \to \mathbb{R}[T] \) is continuous. Let \( \tau(T) \subset [a, b] \) be the simultaneous zero-set of all the functions in \( I \). Shortly, we will see that \( \tau(T) = \sigma(T) \), but we cannot use this yet.

The following is a sort of *Nullstellensatz* for the present situation:

**[1.5 Claim]**: The restriction map \( \mathcal{C}^o[a, b] \to \mathcal{C}^o(\tau(T)) \) has kernel \( I \). That is, if \( f|_{\tau(T)} = 0 \), then \( f(T) = 0 \). More precisely, \( f \geq 0 \) on \( \tau(T) \) if and only if \( f(T) \geq 0 \).

**Proof**: It suffices to show that \( f(T) \geq 0 \) implies \( f \geq 0 \) on \( \tau(T) \). If \( f \) is not non-negative on \( \tau(T) \), then there is \( x_0 \in \tau(T) \) where \( f(x_0) < 0 \). Using the continuity of \( f \), take a small neighborhood \( N \) of \( x_0 \) in \([a, b] \) such that \( f(x) < 0 \) on \( N \). Let \( g \in \mathcal{C}^o[a, b] \) be supported inside \( N \), non-negative, and strictly positive at \( x_0 \). Then \( fg \leq 0 \), and \( fg(x_0) < 0 \), so \( -fg(T) \leq 0 \). But \( f(T) \geq 0 \) and \( g(T) \geq 0 \), so by the corollary on positivity of commuting positive operators, \( fg(T) \geq 0 \). Thus, \( fg(T) = 0 \), so \( fg \in I \), and \( fg|_{\tau(T)} = 0 \), contradiction. Thus, \( f \geq 0 \) on \( \tau(T) \).

Thus, if \( f = 0 \) on \( \tau(T) \), both \( f \geq 0 \) and \( -f \geq 0 \) on \( \tau(T) \), so both \( f(T) \geq 0 \) and \( -f(T) \geq 0 \), so \( f(T) = 0 \), and \( f \in I \).

**[1.6 Corollary]**: \( \mathcal{C}^o[a, b] \to \mathbb{R}[T] \) factors through \( \mathcal{C}^o(\tau(T)) \), giving a commutative diagram

\[
\begin{array}{ccc}
\mathcal{C}^o[a, b] & \longrightarrow & \mathcal{C}^o(\tau(T)) \\
\longrightarrow & & \longrightarrow \\
& & \mathbb{R}[T]
\end{array}
\]

The induced map \( \mathcal{C}^o(\tau(T)) \to \mathbb{R}[T] \) is a *bijection*, and \( |f(T)|_{\text{op}} \geq |f|_{\mathcal{C}^o(\tau(T))} \).

**Proof**: By the Tietze-Urysohn-Brouwer extension theorem (see appendix), every continuous function on \( \tau(T) \) has an extension to a continuous function on \([a, b]\), with the same sup-norm. This gives the surjectivity of \( \mathcal{C}^o[a, b] \to \mathcal{C}^o(\tau(T)) \). By the claim, \( \mathcal{C}^o(\tau(T)) \approx \mathcal{C}^o[a, b]/I \), giving the injectivity to \( \mathbb{R}[T] \).

Given the positivity, since \( |f(T)|_{\text{op}} \pm f(T) \geq 0 \), from the previous claim \( |f(T)|_{\text{op}} \pm f(x) \geq 0 \) for \( x \in \tau(T) \). Thus, \( \sup_{x \in \tau(T)} |f(x)| \leq |f(T)|_{\text{op}} \).
Now we can refine the earlier argument to give the other inequality on norms:

[1.7] Corollary: The induced map \( C^o(\tau(T)) \) \( \to \mathbb{R}/T \) is an isometric isomorphism. That is, the map is a bijection, and \( |f(T)|_{op} = |f|_{C^o(\tau(T))} \).

Proof: For \( f \geq 0 \) on \( \tau(T) \), again by Tietze-Urysohn-Brouwer, there is an extension \( g \geq 0 \) of \( f \) to \([a, b]\) with the same sup norm. The first claim of the proof showed that \( |f(T)|_{op} \leq |g|_{C^o(a, b)} \), so

\[
|f|_{C^o(\tau(T))} \leq |f(T)|_{op} \leq |g|_{C^o(a, b)} = |f|_{C^o(\tau(T))}
\]

giving the isometry. In particular, for \( f_n(T) \) a Cauchy sequence in the operator norm (for \( f_n \in C^o(\tau(T)) \)), the sequence \( f_n \) is Cauchy in \( C^o(\tau(T)) \), so converges to some \( f \in C^o(\tau(T)) \). By the isometry, \( f_n(T) \to f(T) \), giving the surjection to the closure.

What remains is to show that \( \tau(T) = \sigma(T) \).

First, we reprove the fact that \( \sigma(T) \subset \mathbb{R} \). For \( \lambda \in \mathbb{C} \) such that there is no \( (T - \lambda)^{-1} \), the polynomial \( g(x) = (x - \lambda)(x - \bar{\lambda}) \) is non-zero on \( \mathbb{R} \), so certainly on every interval, so has an inverse \( h(x) = 1/g(x) \in C^o(\tau(T)) \). Then \( h(T)(T - \lambda) \) would be an inverse for \( T - \lambda \), contradiction. Thus, \( \sigma(T) \subset \mathbb{R} \).

For \( \lambda \) real and not in \( \tau(T) \), \( x - \lambda \) is invertible on \( \tau(T) \) with inverse \( h \in C^o(\tau(T)) \), so

\[
h(T) \circ (T - \lambda) = (h \cdot (x - \lambda))(T) = 1(T) = 1
\]

and similarly \( (T - \lambda) \circ h(T) = 1 \), so \( T - \lambda \) is invertible. For \( \lambda \in \tau(T) \), for \( n > 0 \), let \( f_n(x) \in C^o[a, b] \) be

\[
f_n(x) = \begin{cases} 
\frac{n}{|x - \lambda|} & \text{for } |x - \lambda| \leq \frac{1}{n} \\
\frac{1}{|x - \lambda|} & \text{for } |x - \lambda| \geq \frac{1}{n}
\end{cases}
\]

Thus, \( |(x - \lambda) \cdot f_n|_{C^o(\tau(T))} \leq 1 \), and \( (T - \lambda)f_n(T)|_{op} \leq 1 \). If \( T - \lambda \) had an inverse \( S \), then for all \( n \)

\[
n \leq |f_n|_{C^o(\tau(T))} = |f_n(T)|_{op} = |1 \cdot f_n(T)|_{op} = |S \cdot (T - \lambda) \cdot f_n(T)|_{op} \leq |S|_{op} \cdot |(T - \lambda) \cdot f_n(T)|_{op} \leq |S|_{op}
\]

This is impossible, so there is no inverse. This proves that \( \tau(T) = \sigma(T) \). ///

2. Schur’s lemma and other corollaries

[2.1] Corollary: For \( \inf \sigma(T) \leq T \leq \sup \sigma(T) \).

Proof: Let \( a = \inf \sigma(T) \) and \( b = \sup \sigma(T) \). Since \( x - a \geq 0 \) on \( \sigma(T) \), \( (x - a)(T) = T - a \geq 0 \). Since \( b - x \geq 0 \) on \( \sigma(T) \), \( (b - x)(T) = b - T \geq 0 \). ///

[2.2] Corollary: If \( \sigma(T) = \{\lambda\} \), then \( T \) is the scalar operator \( \lambda \).

Proof: Because the function \( f(x) = x \) restricted to \( \{\lambda\} \) is equal to the restriction of the constant function \( g(x) = \lambda \),

\[
T = f(T) = g(T) = \lambda
\]

meaning the scalar operator. ///

[2.3] Remark: Certainly the converse is not true: there easily can be eigenvalues imbedded in continuous spectrum.
Suppose that $\sigma(T)$ contains at least two distinct points $x_1, x_2$, and show that $V$ is not $R$-irreducible. Let $f, g$ be continuous functions with disjoint supports, such that $f(x_1) = 1$ and $g(x_2) = 1$. Thus, $fg = 0$, and $f(T)g(T) = g(T)f(T) = 0$. The image $f(T)(V)$ is not 0, because $f(T) \neq 0$. Also, $f(T)(V)$ is inside the kernel of $g(T)$, because $g(T)f(T) = (gf)(T) = 0$. By continuity of $g(T)$, the closure $W$ of $f(T)(V)$ is also inside the kernel of $g(T)$. Since $g(T) \neq 0$, necessarily $W \neq V$.

Since $T$ commutes with all operators in $R$, $\mathbb{R}[T]$ commutes with $R$, and by continuity of operators in $R$, $\mathbb{R}[T]$ commutes with $R$. Thus, $R$ commutes with $f(T)$ and $g(T)$, so for $S \in R$,

$$S(f(T)(V)) = f(T)(SV) \subset f(T)(V)$$

That is, $R$ stabilizes $f(T)(V)$. By continuity of operators in $R$, $R$ stabilizes the closure $W$ of $f(T)(V)$. But $W$ is a proper closed subspace of $V$, so $V$ is not $R$-irreducible. Since $\sigma(T) \neq \phi$, it is a singleton $\{\lambda\}$. By the previous corollary, $T$ is the scalar operator $\lambda$. ///

Recall that Liouville’s theorem on bounded entire functions implies that the spectrum of a continuous linear operator on a Hilbert space is not empty, as follows. If a continuous $R_\lambda = (T - \lambda)^{-1}$ exists for every complex $\lambda$, then for $0 \neq v \in V$, $R_\lambda v \in V$ is never $0 \in V$. Take $w \in V$ such that $\langle R_\lambda v, w \rangle \neq 0$ for some $\lambda_v \in \mathbb{C}$. Then $f(\lambda) = \langle R_\lambda v, w \rangle$ is a not-identically 0 entire function. At the same time, for large $|\lambda|$, the operator norm of $R_\lambda$ is small. Thus, $f(\lambda)$ is small for large $|\lambda|$, and must be identically 0, by Liouville, contradiction. ///

### 3. Spectral theorem, part two: projectors

The earlier spectral theorem gives an isomorphism of the operator-norm-topology closure of $\mathbb{R}[T]$ to $C^\sigma(\sigma(T))$. In the (larger) closure in the strong operator topology, given by seminorms $\mu_v(S) = |Sv|$ for all $v \in V$, we will exhibit a family of projectors that includes projectors to eigenspaces.

For each $t \in \mathbb{R}$ and $\varepsilon > 0$, consider a family of continuous approximations to step functions:

$$h_{t,\varepsilon}(x) = \begin{cases} 1 & \text{for } x \leq t \\ 1 - \frac{x - t}{\varepsilon} & \text{for } t \leq x \leq t + \varepsilon \\ 0 & \text{for } x \geq t + \varepsilon \end{cases}$$

The proof of the following accomplishes more along the way than in its assertion. The operators $P_t$ in the theorem are projectors:

**[3.1] Theorem:** The strong operator topology limit $P_t = \lim_{\varepsilon \to 0^+} h_{t,\varepsilon}(T)$ exists. For $s \leq t$, we have $P_s \leq P_t$. For real scalars $a \leq b$ such that $a \leq T \leq b$, $P_t = 0$ for $t < a$, and $P_t = 1$ for $t > b$. For all $t \in \mathbb{R}$, $P_t^2 = P_t$. There is one-sided continuity: $\lim_{\varepsilon \to 0^+} P_{t+\varepsilon} = P_t$.

**Proof:** First we need

**[3.2] Claim:** Monotone-decreasing, bounded-from-below limits of self-adjoint operators converge in the strong operator topology. That is, for self-adjoint operators $T_1 \geq T_2 \geq \ldots$ with $T_n \geq c$ for some $c$, for all $n$, the $\lim_n T_n$ exists in the strong operator topology, namely, $\lim_n T_nv$ exists (in the topology on $V$) for all $v \in V$. 

4
Proof: The inequality gives $\langle T_n v, v \rangle \geq \langle T_{n+1} v, v \rangle \geq \ldots \geq c \cdot \langle v, v \rangle$. That is, $\langle T_n v, v \rangle$ is a monotone-decreasing, bounded-from-below sequence of reals. Thus, it has a limit. By polarization, $\lim_n (T_n v, w)$ exists for all $v, w \in V$. That is, $\lim_n T_n$ exists in the weak operator topology. We must improve the result to obtain the better result that it converges in the strong operator topology. To this end, let $\lambda_t(w) = \lim_n (T_n v, w)$. This is a conjugate-linear functional on $V$, and is continuous:

$$|\lambda_t(w)| \leq \|T_1 v, w\| \leq \|T\|_{op} \cdot |v| \cdot |w|$$

Thus, $\lambda_t$ is bounded, hence continuous. By Riesz–Fréchet, for all $v \in V$ there is $Tv \in V$ such that $\langle Tv, v \rangle = \lim_n (T_n v, w)$, for all $w \in V$. One can check that $T$ is linear, continuous, and self-adjoint. ///

[3.3] Claim: For a bounded-from-below function $f$ on $\sigma(T)$ expressible as a monotone-decreasing limit $f$ of $f_n \in C^0(\sigma(T))$ (bounded from below on $\sigma(T)$), the strong operator limit $\lim_n f_n(T)$ is independent of the sequence $\{f_n\}$ having limit $f$. Thus, there is an unambiguous operator $f(T) = \lim_n f_n(T)$.

Proof: Let $g_n$ also decrease monotonically to $f$. For all $\varepsilon > 0$, for all $m$, for sufficiently large $n$, max$g_n(x), f_m(x)) \leq f_m(x) + \varepsilon$ for all $x$, so $g_n(x) \leq f_m(x) + \varepsilon$. Thus, $g_n(T) \leq f_m(T) + \varepsilon$. Thus, $\lim_n g_n(T) \leq f_m(T) + \varepsilon$, and then $\lim_n g_n(T) = \lim_n f_m(T) + \varepsilon$. Since this holds for all $\varepsilon > 0$, and the roles of $g_n$ and $f_m$ can be reversed, we have equality. ///

[3.4] Corollary: Thus, the map $C^0(\sigma(T)) \to \mathbb{R}[T]$ (operator-norm-closure) extends to a map defined on monotone-decreasing, bounded-below limits of functions in $C^0(\sigma(T))$, mapping continuously to the strong operator topology closure of $\mathbb{R}[T]$. The extension is still additive, inequality-preserving, and multiplicative (in the sense that $fg(T) = f(T) \cdot g(T)$).

Proof: The first assertion is a special case of the previous claim. Since $\chi_{(-\infty, t]} \leq \chi_{(-\infty, s]}$ for $s \leq t$, the second assertion follows. The third assertion follows from $\chi_{(-\infty, t]} = \chi_{(-\infty, t]}$ and $(1 - \chi_{(-\infty, t]}) \cdot \chi_{(-\infty, t]} = 0$. Since $P_t$ is in the strong. Since $P_t$ is in the strong operator topology closure of a set of self-adjoint operators, it is self-adjoint. The computation

$$\langle P_t v, (1 - P_t) w \rangle = \langle (1 - P_t) P_t v, w \rangle = \langle 0, w \rangle = 0$$

shows that $(1 - P_t)V \subset (P_t)V$. Equality follows from $(1 - P_t) + P_t = 1$. ///

[3.6] Claim: $a \leq T|_{(P_b - P_a)V} \leq b$ for $a \leq b$.

Proof: Let

$$f(x) = \begin{cases} 0 & \text{for } x \leq t \\ x - t & \text{for } x \geq t \end{cases}$$

and

$$g(x) = \begin{cases} |x - t| & \text{for } x \leq t \\ 0 & \text{for } x \geq t \end{cases}$$

Thus, $f(x) + g(x) = |x - t|$. Since $(x - t) \cdot (1 - \chi_{(-\infty, s]} = f(x)$, we have $(T - t)(1 - P_t) = f(T)$. Thus, $T - t = f(T)$ on $(P_t V)^\perp$. Since $\varepsilon \geq 0$, $T - t \geq 0$ on $(P_t V)^\perp$. Take $t = a$.$$

Since $(x - t)\chi_{(-\infty, s]} = -g(x)$, we have $(T - t)P_t = -g(T)$. Thus, $T - b = -g(T)$ on $P_t V$. Since $-g \leq 0$, $T \leq t$ on $P_t V$. Take $t = b$. ///

[3.7] Claim: $t \to P_t$ is strong operator topology continuous on the right.
Conversely, we show that $Q$ is the identity map on the $t$-eigenspace of $T$.

Proof: Certainly $\text{ch}_{(-\infty, t]} - \text{ch}_{(-\infty, t-\varepsilon]}$ is monotone decreasing (and bounded below) as $\varepsilon \to 0^+$, so $P_t - P_{t-\varepsilon}$ converges in the strong operator topology to a continuous operator $Q$. From $a \leq T|_{(P_h - P_0)V} \leq b$ for $a \leq b$,

$$(t-a)(P_t - P_{t-\varepsilon}) \leq T(P_t - P_{t-\varepsilon}) \leq t(P_t - P_{t-\varepsilon})$$

so $|T-t|(P_t - P_{t-\varepsilon})_{\text{op}} \leq \varepsilon$. Let $w = Qv$. Then $|(T-t)w| \leq \varepsilon$ for all $\varepsilon > 0$, so $(T-t)w = 0$. Thus, $Q$ maps to the $t$-eigenspace.

Conversely, we show that $Q$ is the identity map on the $t$-eigenspace $V_t$. For continuous, real-valued $f$, and for a $T$-stable subspace $W$ of $V$, $f(T|_W) = f(T)|_W$, so we may assume without loss of generality that $T$ is a scalar $t$ on $V$. For all $\varepsilon > 0$, $h_{t,\varepsilon} = 1$ on $\sigma(T) = \{t\}$, so $h_{t,\varepsilon}(T) = 1$, and the strong operator topology limit is 1. For $s < t$, for sufficiently small $\varepsilon > 0$, $h_{s,\varepsilon} = 0$ on $\sigma(T) = \{t\}$, so $h_{s,\varepsilon}(T) = 0$. Thus, $Q = 1$. //

[3.9] Corollary: An isolated point $\lambda$ of $\sigma(T)$ is an eigenvalue of $T$.

Proof: $\text{ch}_{(-\infty, \lambda]} - \text{ch}_{(-\infty, \lambda-\varepsilon]}$ is non-zero for $\varepsilon > 0$, because $\lambda \in \sigma(T)$. Thus, $P_\lambda - P_{\lambda-\varepsilon}$ is non-zero for $\varepsilon > 0$. For $\lambda$ isolated, $P_{\lambda-\varepsilon}$ is constant for $\varepsilon > 0$ sufficiently small. Thus, the limit $\lim_{\varepsilon \to 0^+} P_\lambda - P_{\lambda-\varepsilon}$ is non-zero, and by the previous theorem this is the projector to the eigenspace. //

4. Appendix: goofy lemma on polynomials

The following peculiar lemma is not surprising, is essentially elementary, and facilitates a usefully gradual approach to the spectral theorem and its corollaries.

[4.1] Lemma: Let $f \in \mathbb{R}[x]$ be non-negative-valued on a finite interval $[a, b]$. Then $f$ is expressible as a finite sum of the form

$$f = \sum_i P_i^2 + (x-a) \sum_j Q_j^2 + (b-x) \sum_k R_k^2$$

for polynomials $P_i, Q_j, R_k$ in $\mathbb{R}[x]$.

Proof: It suffices to consider monic $f$, since positive constants can be absorbed. Factor $f$ into irreducibles over $\mathbb{R}$, show that each of the linear and quadratic factors can be expressed in the given form, and then show that a product of such expressions can be re-written in the same form.

For quadratic irreducibles with complex-conjugate roots $z, \overline{z}$, by completing the square,

$$(x-z)(x-\overline{z}) = x^2 - (z + \overline{z})x + z\overline{z} = x - \left(\frac{z + \overline{z}}{2}\right)^2 + (z\overline{z} - \left(\frac{z + \overline{z}}{2}\right)^2)$$
Proof: Without loss of generality, the image of that function must occur to an even power, since otherwise $f(x)$ would take opposite signs on the two sides of $\alpha$, contradicting the positivity of $f$ on $[a, b]$.

A linear factor $x - \alpha$ with $a < \alpha < b$ must occur to an even power, since otherwise $f(x)$ would take opposite signs on the two sides of $\alpha$, contradicting the positivity of $f$ on $[a, b]$.

A linear factor $x - \alpha$ with $a \leq \alpha$ can be rewritten as

$$x - \alpha = (x - a) + (a - \alpha) = (x - a) \cdot 1 + (a - \alpha)$$

Since $a - \alpha \geq 0$, it is a square of an element of $\mathbb{R}$, and this gives the desired expression. Similarly, a linear factor $\alpha - x$ with $\alpha \geq b$ can be rewritten as

$$\alpha - x = (b - x) + (\alpha - b)$$

Thus, all the factors of $f$ can be written in the desired form. As for products, we can inductively rewrite them by

$$P^2 \cdot Q^2 = (PQ)^2 \quad (x - a)P^2 \cdot Q^2 = (x - a) \cdot (PQ)^2 \quad (x - a)P^2 \cdot (x - a)Q^2 = ((x - a)PQ)^2 \quad (b - x)P^2 \cdot Q^2 = (b - x) \cdot (PQ)^2 \quad (b - x)P^2 \cdot (b - x)Q^2 = ((b - x)PQ)^2$$

The only possible issue is the form $(x - a)P^2 \cdot (b - x)Q^2$. By luck,

$$(x - a)(b - x) = (x - a)(b - x) \cdot \left(\frac{b - x + (x - a)}{b - x + (x - a)}\right) = \frac{(x - a) \cdot (b - x)^2 + (b - x) \cdot (x - a)^2}{b - a}$$

which is of the desired form. Iterating these rewritings gives the lemma.

\\

5. Appendix: Tietze-Urysohn-Brouwer extension theorem

Granting Urysohn’s lemma, this result is not difficult.

[5.1] Theorem: For $X$ a normal space (meaning that any two disjoint closed sets have disjoint open neighborhoods), closed subset $E \subset X$, every continuous, bounded, real-valued $f$ on $E$ extends to $F$ on $X$ such that $\sup_X |F| = \sup_E |f|$.

Proof: Without loss of generality, the image of $f$ is contained in $[0, 1]$. Urysohn’s lemma will be repeatedly invoked: given disjoint, closed $B_n, C_n$ in $X$, there is continuous $g_n$ on $X$ taking values in $[0, \frac{1}{2}(2/3)^n]$ such that $g_n = 0$ on $B_n$ and $g_n = \frac{1}{2}(2/3)^n$ on $C_n$. Specify the subsets $B_n, C_n (n = 1, 2, \ldots)$ of $E$ inductively by

$$B_1 = \{x \in E : f(x) \leq \frac{1}{3}\} \quad C_1 = \{x \in E : f(x) \geq \frac{2}{3}\}$$

and

$$B_n = \{x \in E : f(x) - \sum_{i=1}^{n-1} g_i(x) \leq \frac{2^{n-1}}{3^n}\} \quad C_n = \{x \in E : f(x) - \sum_{i=1}^{n-1} g_i(x) \geq \frac{2^n}{3^n}\}$$

These are disjoint closed subsets of $E$, so are closed in $X$. The sum $F = \sum_{i=1}^{\infty} g_i$ converges uniformly, so is continuous. On $E$, $0 \leq f - F \leq (2/3)^n$ for all $n$, so $F = f$ on $E$. 

///
6. Appendix: Urysohn's lemma

[6.1] Theorem: (Urysohn) In a locally compact Hausdorff topological space $X$, given a compact subset $K$ contained in an open set $U$, there is a continuous function $0 \leq f \leq 1$ which is 1 on $K$ and 0 off $U$.

Proof: First, we prove that there is an open set $V$ such that

$$K \subset V \subset \overline{V} \subset U$$

For each $x \in K$ let $V_x$ be an open neighborhood of $x$ with compact closure. By compactness of $K$, some finite subcollection $V_{x_1}, \ldots, V_{x_n}$ of these $V_x$ cover $K$, so $K$ is contained in the open set $W = \bigcup_i V_{x_i}$ which has compact closure $\bigcup_i \overline{V}_{x_i}$ since the union is finite.

Using the compactness again in a similar fashion, for each $x$ in the closed set $X - U$ there is an open $W_x$ containing $K$ and a neighborhood $U_x$ of $x$ such that $W_x \cap U_x = \emptyset$.

Then

$$\bigcap_{x \in X-U} (X-U) \cap \overline{W} \cap \overline{W}_x = \emptyset$$

These are compact subsets in a Hausdorff space, so (again from compactness) some finite subcollection has empty intersection, say

$$(X-U) \cap (\overline{W} \cap \overline{W}_{x_1} \cap \ldots \cap \overline{W}_{x_n}) = \emptyset$$

That is,

$$\overline{W} \cap \overline{W}_{x_1} \cap \ldots \cap \overline{W}_{x_n} \subset U$$

Thus, the open set

$$V = W \cap W_{x_1} \cap \ldots \cap W_{x_n}$$

meets the requirements.

Using the possibility of inserting an open subset and its closure between any $K \subset U$ with $K$ compact and $U$ open, we inductively create opens $V_r$ (with compact closures) indexed by rational numbers $r$ in the interval $0 \leq r \leq 1$ such that, for $r > s$, $K \subset V_r \subset \overline{V}_r \subset V_s \subset \overline{V}_s \subset U$.

From any such configuration of opens we construct the desired continuous function $f$ by

$$f(x) = \sup\{r \text{ rational in } [0,1] : x \in V_r, \} = \inf\{r \text{ rational in } [0,1] : x \in \overline{V}_r, \}$$

It is not immediate that this sup and inf are the same, but if we grant their equality then we can prove the continuity of this function $f(x)$. Indeed, the sup description expresses $f$ as the supremum of characteristic functions of open sets, so $f$ is at least lower semi-continuous. The inf description expresses $f$ as an infimum of characteristic functions of closed sets so is upper semi-continuous. Thus, $f$ would be continuous.

To finish the argument, we must construct the sets $V_r$ and prove equality of the inf and sup descriptions of the function $f$.

To construct the sets $V_r$, start by finding $V_0$ and $V_1$ such that

$$K \subset V_1 \subset \overline{V}_1 \subset V_0 \subset \overline{V}_0 \subset U$$

[1] A (real-valued) function $f$ is lower semi-continuous when for all bounds $B$ the set $\{x : f(x) > B\}$ is open. The function $f$ is upper semi-continuous when for all bounds $B$ the set $\{x : f(x) < B\}$ is open. It is easy to show that a sup of lower semi-continuous functions is lower semi-continuous, and an inf of upper semi-continuous functions is upper semi-continuous. As expected, a function both upper and lower semi-continuous is continuous.
Fix a well-ordering $r_1, r_2, \ldots$ of the rationals in the open interval $(0, 1)$. Supposing that $V_{r_1}, \ldots, V_{r_n}$ have been chosen, let $i, j$ be indices in the range $1, \ldots, n$ such that

$$r_j > r_{n+1} > r_i$$

and $r_j$ is the smallest among $r_1, \ldots, r_n$ above $r_{n+1}$, while $r_i$ is the largest among $r_1, \ldots, r_n$ below $r_{n+1}$. Using the first observation of this argument, find $V_{r_{n+1}}$ such that

$$V_{r_j} \subset \overline{V}_{r_j} \subset V_{r_{n+1}} \subset \overline{V}_{r_{n+1}} \subset V_{r_i} \subset \overline{V}_{r_i}$$

This constructs the nested family of opens.

Let $f(x)$ be the sup and $g(x)$ the inf of the characteristic functions above. If $f(x) > g(x)$ then there are $r > s$ such that $x \in V_r$ and $x \notin \overline{V}_s$. But $r > s$ implies that $V_r \subset \overline{V}_s$, so this cannot happen. If $g(x) > f(x)$, then there are rationals $r > s$ such that

$$g(x) > r > s > f(x)$$

Then $s > f(x)$ implies that $x \notin V_s$, and $r < g(x)$ implies $x \in V_r$. But $V_r \subset \overline{V}_s$, contradiction. Thus, $f(x) = g(x)$. ///

### 7. Appendix: historical notes

To a considerable degree, our first section follows the outline in Appendix A1 of S. Lang’s $SL_2(\mathbb{R})$. There, on page 362, it is noted that F. Riesz first used a positivity argument to get a form of the spectral theorem, but the sequel of that argument aimed toward an integral form of a spectral decomposition, while, in contrast, a 1950 seminar of J. von Neumann followed the outline of that appendix and section one here.