Abstract

We discuss extremal structure in locally convex topological vector spaces, as well as a fundamental result in the theory of topological vector spaces: the Krein–Milman theorem. We also use extremal structure to prove the Stone–Weierstrass Theorem and the Banach–Stone theorem.

Extremal Structure

We need to recall some basic ideas related to convexity and compactness in topological vector spaces.

Theorem: If $A_1, \dots, A_n$ are compact convex sets in a topological vector space X, then $\operatorname{conv}(A_1 \cup \cdots \cup A_n)$ is compact.

Proof. Let $\Delta_n = \operatorname{conv}(e_1, \dots, e_n)$ be the basic simplex in $\mathbb{R}^n$, where elements look like

$$\Delta_n = \left\{ (s_1, \dots, s_n) \;\middle|\; s_i \geq 0, \sum_{i=1}^{n} s_i = 1 \right\}.$$

Define $A = A_1 \times \cdots \times A_n$, and set $f\colon \Delta_n \times A \to X$ to be defined by $f(s,a) = \sum_i s_i a_i$. We set $K = f(S \times A)$.

Note that since f is continuous (as addition and scalar multiplication are continuous), Δ_n is compact, and A is compact, we have that K is compact. Furthermore, $K \subseteq \operatorname{conv}(A_1 \cup \cdots \cup A_n)$. We will now show that the inclusion goes in the opposite direction.

We will do this by showing that K is convex. Let $(s,a), (t,b) \in S \times A$, and let $0 \leq q \leq 1$. Then, defining

$$\begin{aligned} u &= qs + (1-q)t \\ c_i &= \frac{q s_i a_i + (1-q) t_i b_i}{q s_i + (1-q) t_i}, \end{aligned}$$

we have

$$\begin{aligned} qf(s,a) + (1-q)f(t,b) &= f(u,c) \\ &\in K, \end{aligned}$$

meaning K is convex, so $\operatorname{conv}(A_1 \cup \cdots \cup A_n) \subseteq K$. $\square$

Definition: Let K be a subset of a vector space X. A nonempty $S \subseteq K$ is called a *face* for K if the interior of any line in K that is contained in S contains its endpoints. Analytically, this means that if $x, y \in K$ are such that, for all $t \in (0,1)$, $tx + (1-t)y \in S$, then $x, y \in S$.

An *extreme point* of K is an extreme set of K that consists of one point. We write $\operatorname{ext}(K)$ for the extreme points of K.

Example: Let Ω be a LCH space. The extreme points of the regular Borel probability measures on Ω are the Dirac measures. That is,

$$\operatorname{ext}(\mathcal{P}_r(\Omega)) = \{\delta_x \mid x \in \Omega\}.$$

In one direction, we see that if $x \in \Omega$, and $\delta_x = \frac{1}{2}(\mu + \nu)$, then for a Borel set $E \subseteq \Omega$ with $x \in E$, we have $1 = \frac{1}{2}(\mu(E) + \nu(E))$. Therefore, $\mu(E) = \nu(E) = 1$. If $x \notin E$, then $0 = \frac{1}{2}(\mu(E) + \nu(E))$, so $\mu(E) = \nu(E) = 0$. Thus, $\mu = \nu = \delta_x$, so every δ_x is extreme.

In the opposite direction, if $\mu \in \operatorname{ext}(\mathcal{P}_r(\Omega))$, we claim that there is $x_0 \in \Omega$ with $\operatorname{supp}(\mu) = \{x_0\}$. Now, since $\mu(\Omega) = 1$, we know that $\operatorname{supp}(\mu) \neq \emptyset$.

Suppose there exist $x, y \in \operatorname{supp}(\mu)$ with $x \neq y$. Since Ω is Hausdorff, we can separate $x, y \in \operatorname{supp}(\mu)$ with disjoint open sets U and V, where $0 < \mu(U) < 1$ and $0 < \mu(V) < 1$. Set $t = \mu(U)$, and define

$$\begin{aligned} \mu_1(E) &= \frac{\mu(E \cap U)}{\mu(U)} \\ \mu_2(E) &= \frac{\mu(E^c)}{\mu(U^c)}. \end{aligned}$$

Then, μ_1, μ_2 are regular Borel probability measures with $\mu_1 \neq \mu_2$ and $t\mu_1 + (1-t)\mu_2 = \mu$, which contradicts μ being extreme. Therefore, $\operatorname{supp}(\mu) = \{x_0\}$, so $\mu = \delta_{x_0}$.

Example: Let Ω be a LCH space. Then,

$$\operatorname{ext}\bigl(B_{M_r(\Omega)}\bigr) = \{\alpha\delta_x \mid x \in \Omega, \alpha \in \mathbb{T}\}.$$

We start by showing that $\alpha\delta_x$ is extreme. Suppose $\alpha\delta_x = \frac{1}{2}(\mu + \nu)$ for some $\mu, \nu \in B_{M_r(\Omega)}$. Then, if $x \in E$, we have

$$\alpha = \frac{1}{2}(\mu(E) + \nu(E)).$$

Note that

$$\begin{aligned} |\mu(E)| &\leq |\mu|(E) \\ &\leq |\mu|(\Omega) \\ &= \|\mu\| \\ &\leq 1, \end{aligned}$$

and similarly for $|\nu|(E)|$. Thus, $\mu(E) = \nu(E) = \alpha$. In particular,

$$\begin{aligned} 1 &= |\alpha| \\ &= |\mu(\{x\})| \\ &\leq |\mu|(\{x\}) \\ &\leq |\mu|(\Omega) \\ &= \|\mu\| \\ &\leq 1, \end{aligned}$$

so $|\mu|(\Omega) = 1$, and $|\mu|(\{x\}) = 1$, meaning $\mu(\{x\}^c) = 0$. Similarly, we must have $|\nu|(\{x\}^c) = 0$. If E is any Borel set not containing x, we then have

$$\begin{aligned} |\mu(E)| &\leq |\mu|(E) \\ &\leq |\mu|(\{x\}^c) \\ &= 0, \end{aligned}$$

so $\mu(E) = 0$, and similarly $\nu(E) = 0$. Thus, we have $\mu = \nu = \alpha\delta_x$, so $\alpha\delta_x$ is extreme.

Now, we show that if $\mu \in \operatorname{ext}\bigl(B_{M_r(\Omega)}\bigr)$, then $|\mu| \in \operatorname{ext}(\mathcal{P}_r(\Omega))$.

Write $\mu = f\, d|\mu|$ for some $f\colon \Omega \to \mathbb{T}$. Suppose there exist $\nu, \lambda \in \mathcal{P}_r(\Omega)$ such that $|\mu| = \frac{1}{2}(\nu + \lambda)$, Then,

$$\mu = \frac{1}{2}(f\, d\nu + f\, d\lambda).$$

Since ν and λ are positive measures, $|f\, d\nu| = |f|\, d\nu = d\nu$, and $|f\, d\lambda| = |f|\, d\lambda = d\lambda$. Since μ is extreme, we have $f\, d\nu = f\, d\lambda = \mu$, so $|\mu| = |f\, d\nu| = \nu$ and $|\mu| = |f\, d\lambda| = \lambda$.

Since $|\mu| \in \operatorname{ext}(\mathcal{P}_r(\Omega))$, we have $|\mu| = \delta_{x_0}$ for some $x_0 \in \Omega$. Then, for any Borel set E, we have

$$\begin{aligned}
\mu(E) &= \int_E f\, d|\mu| \\
&= \int_\Omega f\mathbb{1}_E\, d\delta_{x_0} \\
&= f(x_0)\mathbb{1}_E(x_0) \\
&= \begin{cases} f(x_0) & x_0 \in E \\ 0 & x_0 \notin E \end{cases} \\
&= f(x_0)\delta_{x_0}(E).
\end{aligned}$$

Thus, $\mu = f(x_0)\delta_{x_0}$. Setting $\alpha = f(x_0)$, we have $|\alpha| = 1$ by definition.

Example: The picture of a face in a convex compact set is relatively simple. If $u\colon X \to \mathbb{R}$ is an $\mathbb{R}$-linear continuous functional, and $P \subseteq X$ is compact and convex, the infimum $\inf_{x\in P} u(x) =: s$ is attained. The subset

$$P_u = \{x \in P \mid u(x) = s\}$$

is a closed face in P.

To start, P_u is nonempty because the infimum is attained. Since u is continuous, P_u is closed. Furthermore, if $t \in [0,1]$ and $x, y \in P_u$, then $(1-t)x + ty \in P_u$, as

$$\begin{aligned}
u((1-t)x + ty) &= (1-t)u(x) + tu(y) \\
&= (1-t)s = ts \\
&= s.
\end{aligned}$$

Now, if $t \in (0,1)$ and $x, y \in P$ with $(1-t)x + ty \in P_u$, then

$$s = (1-t)u(x) + tu(y).$$

Since $u(x) \geq s$ and $u(y) \geq s$, we must have $u(x) = u(y) = s$, meaning $x, y \in P_u$.

The Krein–Milman Theorem

One of the most important results in extremal structure is the fact that every compact convex set of a topological vector space (with some relatively weak conditions) has an extreme point — moreover, there are a lot of extreme points.

Theorem (Krein–Milman)**:** Let X be a topological vector space where X^* separates points. If K is a nonempty compact convex set in X, then

$$K = \overline{\operatorname{conv}}(\operatorname{ext}(K)).$$

Proof. We start with a lemma.

Lemma: If F is a face of K and G is a face of F, then G is a face of K.

Proof. Let $x, y \in K$ be such that for all $t \in (0,1)$, $(1-t)x + ty \in G$. Then, since G is a face of F, we have $(1-t)x + ty \in F$, so since F is a face, $x, y \in F$. However, since G is a face, $x, y \in G$, so G is a face of K. $\square$

We start by showing that $\text{ext}(K) \neq \emptyset$. Let $F \subseteq K$ be a closed face. The family

$$\mathcal{G} = \{G \subseteq F \mid G \text{ is a closed face in } F\}$$

is nonempty, as $F \in \mathcal{G}$. Ordering $\mathcal{G}$ by containment, we will show that $\mathcal{G}$ satisfies the conditions of Zorn's lemma. If $\mathcal{C} \subseteq \mathcal{G}$ is a chain, then we claim that

$$I = \bigcap_{G \in \mathcal{C}} G$$

is an element of $\mathcal{G}$ that is an upper bound for $\mathcal{C}$. First, since I is an arbitrary intersection of convex sets, I is convex.

Furthermore, for any $G_1, \dots, G_n \in \mathcal{C}$, then since $\mathcal{C}$ is a chain, there is j such that $G_i \preceq G_j$ For all $i = 1, \dots, n$, meaning $\bigcap_{i=1}^{n} G_i = G_j \neq \emptyset$. Since K is compact, the finite intersection property gives $I \neq \emptyset$. Finally, let $t \in (0,1)$ with $x, y \in F$ and $(1-t)x + ty \in I$. Then, $(1-t)x + ty \in G$ for all $G \in \mathcal{C}$, so $x, y \in G$ for all $G \in \mathcal{C}$, so $x, y \in I$, meaning I is a face. Notice that for all $G \in \mathcal{C}$, we have $G \preceq I$, so the conditions of Zorn's lemma are satisfied.

By Zorn's lemma, there is a maximal $P \in \mathcal{G}$. We claim that P is a singleton.

Note that P is compact since it is closed. Let $\varphi \in X^*$ and set $u = \text{Re}(\varphi)$. Since P is compact, the set

$$P_u = \left\{ p \in P \;\middle|\; u(p) = \inf_{x \in P} u(x) \right\},$$

and by maximality, we must have $P_u = P$. Since $\varphi(x) = u(x) - iu(ix)$, we must have that φ is constant on P, so $P = \{z\}$ as X^* separates points.

Since F is a face, and $P \subseteq F$ is a face, P is a face, so $z \in \text{ext}(K)$.

Now, note that $C = \overline{\text{conv}}(\text{ext}(K)) \subseteq K$ as K is closed and convex. Suppose that this inclusion is strict. Let $x_0 \in K \setminus C$.

Then, by the Hahn–Banach separation, there is $\varphi \in X^*$ and $t \in \mathbb{R}$ such that for all $y \in C$,

$$u(x_0) < t \leq u(y),$$

where $u = \text{Re}(\varphi)$. Let $s = \inf_{k \in K} u(k)$, so that $K_u = \{x \in K \mid u(x) = s\}$. This is a closed face in K, so it has an extreme point $z \in K$, with $z \in C$. Then, $u(z) \geq t > s$, but $z \in K_u$, so $u(z) = s$. Therefore, the inclusion is not strict. □

Other Uses of Extremal Structure

Extremal structure can often give us a lot of information about the structure of particularly important spaces. We start by proving a particular linear-algebraic lemma.

Lemma: Let X and Y be vector spaces, $T \colon X \to Y$ a linear isomorphism. Let $C \subseteq X$ be nonempty and convex. Then,

$$T(\text{ext}(C)) = \text{ext}(T(C)).$$

In particular, if T is an isometric isomorphism of normed spaces, then $T(\text{ext}(B_X)) = \text{ext}(B_Y)$.

Proof. Let $x \in \text{ext}(C)$. Suppose $T(x) = \frac{1}{2}(y_1 + y_2)$ for some $y_1, y_2 \in T(C)$. We find x_i such that $T(x_i) = y_i$ for each i. Then,

$$T(x) = \frac{1}{2}(T(x_1) + T(x_2))$$

$$= T\left(\frac{1}{2}(x_1 + x_2)\right).$$

Since T is injective, $x = \frac{1}{2}(x_1 + x_2)$, and since x is extreme, $x = x_1 = x_2$, and $T(x) = y_1 = y_2$. Thus, $T(\text{ext}(C)) \subseteq \text{ext}(T(C))$.

Applying the same process on T^{-1}, we have $T^{-1}(\text{ext}(T(C))) \subseteq \text{ext}(C)$. Therefore, $\text{ext}(T(C)) \subseteq T(\text{ext}(C))$, so the sets are equal. □

One of the basic consequences of the Krein–Milman theorem is that it allows us to characterize dual spaces.

Theorem: Let X be a normed vector space. If $\text{ext}(B_X) = \emptyset$, then X is not a dual space.

Proof. If Z is a normed space, then B_{Z^*} in the w^*-topology is a compact and convex set, meaning that $\text{ext}(B_{Z^*}) \neq \emptyset$. The result follows from the contrapositive. □

The Stone–Weierstrass Theorem

Theorem (Stone–Weierstrass)**:** Let Ω be a compact Hausdorff space, and let $A \subseteq C(\Omega)$ be a unital separating $*$-subalgebra. Then,

$$\overline{A}^{\|\cdot\|_u} = C(\Omega).$$

The traditional proof involves showing that if $g \in A$, then $|g| \in A$, which allows for a lattice of functions in A defined over the open cover of Ω to admit a limit point. There is a much more slick proof involving extremal structure. First, we recall some definitions relating to the dual space.

Definition: Let X be a normed space, and let $S \subseteq X$, $T \subseteq X^*$. We define

$$S^{\perp} = \{\varphi \in X^* \mid \varphi(x) = 0 \text{ for all } x \in S\}$$

to be the *annihilator* of S, and the *pre-annihilator* of T to be

$$T_{\perp} = \{x \in X \mid \varphi(x) = 0 \text{ for all } \varphi \in T\}.$$

Note that $S^{\perp} \subseteq X^*$ and $T_{\perp} \subseteq X$ are norm-closed subspaces.

Corollary: Let X be a normed space, and let $S \subseteq X$ be a subset. Then,

$$\left(S^{\perp}\right)_{\perp} = \overline{\text{span}}(S).$$

Proof. Since $S \subseteq \left(S^{\perp}\right)_{\perp}$, we must have $Z \coloneqq \overline{\text{span}}(S) \subseteq \left(S^{\perp}\right)_{\perp}$.

Suppose the inclusion is strict. Then, there exists $x_0 \in \left(S^{\perp}\right)_{\perp} \setminus Z$. By the Hahn–Banach separation for normed spaces, there is $\varphi \in X^*$ such that $\varphi|_Z = 0$ and $\varphi(x_0) = \text{dist}_Z(x_0) \neq 0$, meaning $\varphi \in S^{\perp}$, so $\varphi(x_0) = 0$, a contradiction. □

Proof of the Stone–Weierstrass Theorem. To show the Stone–Weierstrass theorem, we will show that $A^{\perp} = \{0\}$. Note that annihilators are always w^*-closed, so it is enough to show that $B_{A^{\perp}} = A^{\perp} \cap B_{C(\Omega)^*} = \{0\}$. Furthermore, note that $B_{A^{\perp}}$ is w^*-compact, so we will show that $\text{ext}(B_{A^{\perp}}) = \{0\}$.

Suppose $\varphi \in \text{ext}(B_{A^{\perp}})$ with $\|\varphi\| \neq 0$. Then, $\|\varphi\| = 1$, else we would be able to write

$$\varphi = (1 - \|\varphi\|)(0) + \|\varphi\| \frac{\varphi}{\|\varphi\|},$$

and since $0 \neq \varphi$, this would contradict the fact that φ is extreme. Thus, $\|\varphi\| = 1$. By the Riesz–Markov theorem, we know that φ is of the form

$$\varphi(f) = \int_{\Omega} f \, d\mu$$

for some regular Borel complex measure μ with norm 1. We will show now that $\operatorname{supp}(|\mu|) = \{x\}$ for some $x \in B_{A^\perp}$.

Suppose $x \neq y \in \operatorname{supp}(\mu)$. Since A separates points, we may find $g \in A$ such that $g(x) \neq g(y)$. Using the Cartesian decomposition, we write $g = h + ik$, and since A is a $*$-closed subspace, we know that $h, k \in A$. Without loss of generality, we may take $h(x) \neq h(y)$ (else multiply g by $-i$ and replace h with k).

Set $\widetilde{h} = 2\|h\|\mathbb{1}_\Omega + h$, which yet again belongs to A since A is unital, and note that $\widetilde{h}(x) \neq \widetilde{h}(y)$. Finally, set $f = \frac{1}{2\|\widetilde{h}\|}h$. We have that $f\colon \Omega \to (0,1)$ is continuous with $f \in A$ and $f(x) \neq f(y)$. Furthermore, $f \in B_{C(\Omega)}$.

Define the complex measures $\nu = f\, d\mu$ and $\lambda = (1-f)\, d\mu$, where we define

$$\begin{aligned}
\nu(E) &= \int_E f\, d\mu \\
\lambda(E) &= \int_E (1-f)\, d\mu.
\end{aligned}$$

By definition, $\nu, \lambda \in B_{M_r(\Omega)}$, and for all $a \in A$,

$$\begin{aligned}
\int_\Omega a\, d\nu &= \int_\Omega af\, d\mu \\
&= \varphi(af) \\
&= 0,
\end{aligned}$$

as we defined $\varphi \in A^\perp$, and A is a subalgebra. Similarly,

$$\begin{aligned}
\int_\Omega a\, d\lambda &= \int_\Omega a(1-f)\, d\mu \\
&= \varphi(a(1-f)) \\
&= 0.
\end{aligned}$$

Thus, $\nu, \lambda \in A^\perp \cap B_{M_r(\Omega)} = B_{A^\perp}$. Additionally,

$$\begin{aligned}
\|\nu\| + \|\lambda\| &= |\nu|(\Omega) + |\lambda|(\Omega) \\
&= \int_\Omega f\, d|\mu| + \int_\Omega (1-f)\, d|\mu| \\
&= \int_\Omega \mathbb{1}_\Omega\, d|\mu| \\
&= |\mu|(\Omega) \\
&= \|\mu\| \\
&= 1,
\end{aligned}$$

where we use the definition of the total variation norm, $\|\mu\| = |\mu|(\Omega)$.

Thus, we have the convex combination

$$\begin{aligned}
\mu &= \nu + \lambda \\
&= \|\nu\|\left(\frac{\nu}{\|\nu\|}\right) + \|\lambda\|\left(\frac{\lambda}{\|\lambda\|}\right),
\end{aligned}$$

and since μ is extreme, $\mu = \frac{\nu}{\|\nu\|}$, meaning $\nu = \|\nu\|\mu$. Therefore,

$$\int_\Omega f\, d|\mu| = |\nu|(\Omega)$$

$$\begin{aligned}&= \|\nu\||\mu|(\Omega)\\&= \int_{\Omega} \|\nu\| \, d|\mu|,\end{aligned}$$

meaning $f = \|\nu\|$ $|\mu|$-a.e. Furthermore,

$$\operatorname{supp}(|\mu|) \subseteq \{x \mid f(x) = \|\nu\|\},$$

as, taking $E \coloneqq \{x \mid f(x) = \|\nu\|\}$, we must have $E^c \subseteq \ker(|\mu|)$. Since $x, y \in \operatorname{supp}(\mu)$, we have $x, y \in \operatorname{supp}(|\mu|)$, so $f(x) = f(y) = \|\nu\|$, which is a contradiction.

Therefore, we must have $\mu = \alpha\delta_x$ for some $|\alpha| = 1$. Then, for all $a \in A$, since $\varphi \in A^{\perp}$,

$$\begin{aligned}0 &= \varphi(a)\\&= \int_{\Omega} a \, d\mu\\&= \alpha a(x).\end{aligned}$$

In particular, this holds for $\alpha = \alpha \mathbb{1}_{\Omega}(x)$, so $\mu = 0$, which contradicts our assumption that $\|\varphi\| \neq 0$. Thus, we must have $\operatorname{ext}(B_{A^{\perp}}) = \{0\}$.

Applying the Krein–Milman theorem, we have

$$\begin{aligned}B_{A^{\perp}} &= \overline{\operatorname{conv}}(\operatorname{ext}(B_{A^{\perp}}))\\&= \{0\},\end{aligned}$$

or that $\left(A^{\perp}\right)_{\perp} = \overline{A}^{\|\cdot\|_u} = C(\Omega)$. $\square$

The Banach–Stone Theorem

Given two locally compact Hausdorff spaces, X and Y, and a proper[I] map $\tau\colon X \to Y$, there is a natural dual linear map,

$$T_{\tau}\colon C_0(Y) \to C_0(X),$$

given by $T_{\tau}(f) = f \circ \tau$.

Theorem: If $\tau\colon X \to Y$ is a proper map, and $T_{\tau}\colon C_0(Y) \to C_0(X)$ is a proper map, then:

(a) if τ is surjective, then T_{τ} is injective;

(b) if T_{τ} is injective, and $\tau(X) \subseteq Y$ is closed, then τ is surjective;

(c) if T_{τ} is surjective, then τ is injective;

(d) if X, Y are compact, then if τ is injective, T_{τ} is surjective.

Furthermore, T_{τ} is a contractive map; if τ is a homeomorphism, then T_{τ} is an isometric isomorphism.

Proof.

(a) Let τ be surjective. Then, if $T_{\tau}(f) = 0$, we must have $f|_{\operatorname{ran}(\tau)} = 0$; however, since $\operatorname{ran}(\tau) = Y$, we must have $f = 0$.

(b) If T_{τ} is injective, and there is $y \in Y$ such that $y \notin \tau(X)$, Urysohn's lemma gives a compactly supported $f\colon Y \to [0,1]$ such that $f|_{\tau(X)} = 0$ and $f(y) = 1$. However, we would have $T_{\tau}(f) = 0$, but $f \neq 0$, which is a contradiction. Thus, we must have $\tau(X) = Y$.

[I] Preimages of compact sets are compact.

(c) Let T_τ be surjective, and let $x_1 \neq x_2$ in X. By Urysohn's lemma, there is $g \in C_0(X)$ such that $g(x_1) \neq g(x_2)$. We may find $f \in C(Y)$ such that $f \circ \tau = g$, meaning $f(\tau(x_1)) \neq f(\tau(x_2))$, so $\tau(x_1) \neq \tau(x_2)$, and τ is injective.

(d) Let τ be injective. If X is compact, then $\tau(X)$ is compact, hence closed, and $\widetilde{\tau}\colon X \to \tau(X)$ is a homeomorphism. Given $g \in C(X)$, the continuous function $f_0 \coloneqq g \circ \widetilde{\tau}^{-1}$ extends to a continuous $f \in C(Y)$ by Tietze's Extension Theorem. Now,

$$\begin{aligned} T_\tau(f) &= f \circ \tau \\ &= f_0 \circ \widetilde{\tau} \\ &= g \circ \widetilde{\tau}^{-1} \circ \widetilde{\tau} \\ &= g, \end{aligned}$$

so T_τ is surjective.

Computing

$$\begin{aligned} \|T_\tau(f)\|_u &= \sup_{x\in X} |T_\tau(f)(x)| \\ &= \sup_{x \in X} |f(\tau(x))| \\ &\leq \sup_{y\in Y} |f(y)| \\ &\leq \|f\|_u, \end{aligned}$$

so $\|T_\tau\|_{\text{op}} \leq 1$.

Now, if τ is a homeomorphism, then both T_τ and $T_{\tau^{-1}} = T_\tau^{-1}$ are contractions, meaning they must be isometries. Since τ is a bijection, T_τ is also a linear isomorphism, meaning T_τ is an isometric isomorphism. $\square$

Surprisingly, the above statement reverses — i.e., for compact Hausdorff spaces X, Y, if there is an isometric isomorphism $T\colon C(Y) \to C(X)$, there is a corresponding homeomorphism $\tau\colon X \to Y$.

Theorem (Banach–Stone)**:** Suppose $T\colon C(Y) \to C(X)$ is an isometric isomorphism of Banach spaces. Then, there exists a homeomorphism $\tau\colon X \to Y$ and a continuous $\alpha\colon \Omega \to \mathbb{T}$ such that for every $x \in \Omega$ and $g \in C(Y)$,

$$T(g)(x) = \alpha(x) g(\tau(x)).$$

Proof. Let $T\colon C(Y) \to C(X)$ be an isometric isomorphism. Then, by the properties of the transpose map, $T^*\colon C(X)^* \to C(Y)^*$ is an isometric isomorphism and a w^*-w^*-homeomorphism. Since T^* is an isometric isomorphism. $T^*\big(\operatorname{ext}\big(B_{M_r(X)}\big) = \operatorname{ext}\big(B_{M_r(Y)}\big)\big)$.

Fix $x \in X$. Since $\delta_x \in \operatorname{ext}\big(B_{M_r(X)}\big)$, we have $T^*(\delta_x) \in \operatorname{ext}\big(B_{M_r(Y)}\big)$. Thus, there is a $\tau(x) \in Y$ and $\alpha(x) \in \mathbb{T}$ such that $T^*(\delta_x) = \alpha(x)\delta_{\tau(x)}$. This gives maps $\alpha\colon X \to \mathbb{T}$ and $\tau\colon X \to Y$.

We claim that $\alpha\colon X \to \mathbb{T}$ is continuous. If $(x_i)_i$ is a net in X with $(x_i)_i \to x$, then $(\delta_{x_i})_i \xrightarrow{w^*} \delta_x$. Therefore, $(T^*(\delta_{x_0}))_i \xrightarrow{w^*} T^*(\delta_x)$. By definition, we have $\big(\alpha(x_i)\delta_{\tau(x_i)}\big)_i \xrightarrow{w^*} \alpha(x)\delta_{\tau(x)}$. Applying to $\mathbb{1}_Y$, we have

$$\begin{aligned} (\alpha(x_i)_i) &= \big(\alpha(x_i)\delta_{\tau(x_i)}(\mathbb{1}_Y)\big) \\ &\to \alpha(x)\delta_{\tau(x)}(\mathbb{1}_Y) \\ &= \alpha(x), \end{aligned}$$

which proves the claim.

Now, we claim that τ is a homeomorphism. Let $(x_i)_i$ be a net converging to $x \in X$. Then, $(\delta_{x_i})_i \xrightarrow{w^*} \delta_x$ and $(\alpha(x_i))_i \to \alpha(x)$ by the previous claim.

Since scalar multiplication is continuous, we get $\left(\alpha(x_i)\delta_{\tau(x_i)}\right)_i \xrightarrow{w^*} \alpha(x)\delta_{\tau(x)}$. Thus,

$$\begin{aligned}\left(\delta_{\tau(x_i)}\right)_i &= \left(\frac{1}{\alpha(x_i)}\left(\alpha(x_i)\delta_{\tau(x_i)}\right)\right)_i\\ &\xrightarrow{w^*} \frac{1}{\alpha(x)}\alpha(x)\delta_{\tau(x)}\\ &= \delta_{\tau(x)}.\end{aligned}$$

For each $g \in C(Y)$, we have $\left(\delta_{\tau(x_i)}(g)\right)_i \to \delta_{\tau(x)}(g)$, or that $(g(\tau(x_i)))_i \to g(\tau(x))$. Since g is arbitrary, we have that $(\tau(x_i))_i \to \tau(x)$, so τ is continuous.

To see that τ is injective, we let $x_1, x_2 \in X$ with $x_1 \neq x_2$. Then, by Urysohn's lemma, we have $\overline{\alpha(x_1)}\delta_{x_1} \neq \overline{\alpha(x_2)}\delta_{x_2}$, so their images under T^* are not equal as T^* is injective. Therefore, we have $\overline{\alpha(x_1)}\alpha(x_1)\delta_{\tau(x_1)} \neq \overline{\alpha(x_2)}\alpha(x_2)\delta_{\tau(x_2)}$. Since α has modulus 1, we have $\delta_{\tau(x_1)} \neq \delta_{\tau(x_2)}$, so $\tau(x_1) \neq \tau(x_2)$.

Now, we show τ is surjective. For any $y \in Y$, there exists $\mu \in \operatorname{ext}\left(B_{M_r(X)}\right)$ such that such that $T^*(\mu) = \delta_y$. We know that $\mu = \beta\delta_x$ for some $x \in X$ and $\beta \in \mathbb{T}$. Thus,

$$\begin{aligned}\delta_y &= T^*(\mu)\\ &= T^*(\beta\ delta_x)\\ &= \beta T^*(\delta_x)\\ &= \beta\alpha(x)\delta_{\tau(x)}.\end{aligned}$$

By Urysohn's Lemma, we must have $\tau(x) = y$, so τ is surjective.

Since τ is a continuous bijection with X compact and Y Hausdorff, τ is a homeomorphism.

Finally, if $g \in C(Y)$ and $x \in \Omega$,

$$\begin{aligned}T(g)(x) &= \delta_x(T(g))\\ &= T^*(\delta_x)(g)\\ &= \alpha(x)\delta_{\tau(x)}(g)\\ &= \alpha(x)g(\tau(x)).\end{aligned}$$

$\square$