1 | Locally Convex Spaces & Duality

1.1 | Locally Convex Spaces

The linear structures of vector spaces is closely related to the notion of lines. A set K is said to be convex (respectively absolutely convex) if it contains the smallest line segments (respectively the smallest complex disc) joining its points, that is. We will use this line segment[1] description of convex sets for heuristics and intuition, and formalise this intuition in the abstract framework using conjugate spaces. Since Banach spaces are topological vector spaces with respect to the norm. The linear operations, such as vector addition, and scaling are continuous with respect to the topology induced by the norm. Every line passing through origin in $\mathcal{X}$ corresponds to a real linear map from $\mathbb{R}$ to $\mathcal{X}$,

$$L : \mathbb{R} \to \mathcal{X}.$$

Since norms are necessarily non-vanishing, the composition,

$$\mathbb{R} \xrightarrow{L} \mathcal{X} \xrightarrow{\|\cdot\|_{\mathcal{X}}} \mathbb{R},$$

defines a non-trivial function from every line. Linear maps are transformation of lines passing through origin to lines passing through origin, and we need to generalise them to be not dependent on the origin. Affine maps are compositions of a translation map after a linear map, and take lines to lines. Affine maps act as the correct starting point for convex analysis.

Heuristically, the norm on a Banach space sees every line on the space, and hence we may expect to characterise convex sets in terms of the properties of the norm.

1.1.1 | The Hahn-Banach Theorem

For any complex number λ, we denote its real part by $\Re(\lambda)$ and the imaginary part by $\Im(\lambda)$. We are interested in the two cases corresponding to convexity and absolute convexity. We will now prove the Hahn-Banach theorem for the real linearity case, and extend to the complex case as a corollary. The Hahn-Banach separation theorems make precise the heuristic relation between the notion of convexity and norm on the Banach space.

[1]Note that in the vector space context, convexity is closely related to the notion of lines. This notion of convexity is intuitive since it is very geometric. If needed this intuitive notion of convexity can be replaced with more general heuristics based on the needs. It might be possible to obtain an abstract categorical notion of convexity where we replace the order relation on reals by morphisms between objects of a category. A full subcategory of this category could then correspond to a *convex thing*. We will however not explore such abstract territory here since we do not see the need for it.

The proof of Hahn-Banach theorem also works for sublinear maps on $\mathcal{X}$. A sublinear map is a real valued map f on $\mathcal{X}$ such that,

$$f(x+y) \leqq f(x) + f(y).$$

When $f(\lambda x) = |\lambda| f(x)$, it is called a semi-norm, and if $f(x)$ is zero only when x is the origin of $\mathcal{X}$ they correspond to norms. Hence Hahn-Banach theorem would still be useful if the *accuracy* function for instruments does not correspond to a norm.

Note that we can write any complex number λ as $\Re(\lambda) + i\Im(\lambda)$. Hence we can write $\omega(x) = \Re(\omega(x)) + i\Im(\omega(x))$. Since ω is a complex functional, we also have $\omega(ix) = i\omega(x)$, and hence we have $\Re(i\omega(x)) = \Re(\omega(ix)) = -\Im(\omega(x))$. Therefore we must have,

$$\omega(x) = \Re(\omega(x)) - i\Re(\omega(ix)).$$

Hence the real part $\Re(\omega)$ already determines the functional ω. For absolute convexity we replace $\Re(\varphi)$ by the modulus function.

Theorem 1.1.1. (Hahn-Banach) *Let $\mathcal{X}$ be a vector space with a sublinear map f. Let $\mathcal{Y}$ be a subspace of $\mathcal{X}$. If φ be a linear functional on $\mathcal{Y}$ such that for all y in $\mathcal{Y}$*

$$\Re\big(\varphi(y)\big) \leqq f(y)$$

then there exists a linear functional $\widehat{\varphi}$ on $\mathcal{X}$ with

$$\Re\big(\widehat{\varphi}(x)\big) \leqq f(x).$$

Proof

Let $\mathcal{E}_\varphi$ be the collection of all extensions of φ, such that whenever ψ belongs to $\mathcal{E}_\varphi$, it satisfies $\psi(x) \leqq \|x\|_{\mathcal{X}}$ on the domain $\mathcal{X}_\psi$ of ψ. This is a partially ordered collection. By Zorn's lemma there must exists a maximal element $\widehat{\varphi}$ such that

$$\Re\big(\widehat{\varphi}(x)\big) \leqq f(x).$$

for all x in its domain. Note that the domain of $\widehat{\varphi}$ consists of all of $\mathcal{X}$, because otherwise we can construct an extension of $\widehat{\varphi}$, and hence contradicting the maximality. We can construct such an extension as follows;

Let $\mathcal{X}_{\widehat{\varphi}}$ be the domain of $\widehat{\varphi}$, suppose there exists some x in $\mathcal{X} \backslash \mathcal{X}_{\widehat{\varphi}}$, and let $\mathbb{R}x$ denote the line containing both the origin of $\mathcal{X}$ and x. We can obtain an extension of the linear functional φ to the subspace $\mathcal{Y} \oplus \mathbb{R}x$ by fixing its value at x to be say λ_x and extending linearly to all of $\mathbb{R}x$, as

$$\varphi_x(y + \lambda x) = \varphi(y) + \lambda\lambda_x.$$

By adjusting λ_x appropriately we can make it to satisfy the majorisation requirement. For this, we need to have $\Re\big(\varphi_x(y + \lambda x)\big) \leqq f(y + \lambda x)$, and it follows that any λ_x with

$$\sup_{y \in \mathcal{Y}} \Big[\Re\big(\varphi(y)\big) - f(y - x)\Big] \leqq \big|\lambda_x\big| \leqq \inf_{y \in \mathcal{Y}} \Big[f(x+y) - \Re\big(\varphi(y)\big)\Big]$$

can be used for the construction of such an extension. Such a choice of λ_x keeps the extension well-behaved on $\mathbb{R}x$ with respect to the sublinear function f. This completes the proof. $\square$

To extend to the complex case, we use the fact $\varphi(x) = \Re(\varphi(x)) - i\Re(\varphi(ix))$ whenever φ is complex linear, φ is determined by its real part $\Re(\varphi)$. Since $\Re(\lambda) \leqq |\lambda|$ it follows that if φ is complex linear functional with

$$\big|\varphi(y)\big| \leqq f(y)$$

for all y in $\mathcal{Y}$ then there exists an extension $\widehat{\varphi}$ of φ with

$$\big|\widehat{\varphi}(x)\big| \leqq f(x)$$

for all $x \in \mathcal{X}$. The Hahn-Banach theorem can now be used to study convexity properties of $\mathcal{X}$ since the sub-linear map relates the convexity structure on $\mathcal{X}$ with the convexity structure on $\mathbb{R}$. Given two disjoint convex sets, we can consider the collection of all lines joining the points of the two convex sets. We can describe these lines in terms of affine linear functionals, and describe when a line leaves a convex set using the values of the functional. This allows us to separate disjoint convex subsets using linear functionals. We now formalise this heuristics to obtain a characterisation of disjointness of convex sets in terms of the conjugate space.

A locally convex space $\mathcal{X}$ is a topological vector space (vector space operations are continuous with respect to its topology) such that its topology is generated by a family of convex neighborhoods of origin together with the vector space operations. Let $\mathcal{N}_{\mathcal{X}}$ be such a neighborhood basis of convex sets of origin. Then for each $\mathcal{N}$ in $\mathcal{N}_{\mathcal{X}}$, define a function $\mu_{\mathcal{N}}$ which maps

$$\mu_{\mathcal{N}}(x) = \inf_{x \in \lambda\mathcal{N}} \big[|\lambda|\big].$$

Here $\mu_{\mathcal{N}}(x)$ is the smallest amount descaling of x needed to bring it inside the neighborhood $\mathcal{N}$. By convention we assume infimum of the empty set to be $-\infty$. It immediately follows from definition that $\mu_{\mathcal{N}}(x) < 1$ only when x belongs to $\mathcal{N}$, and $\mu_{\mathcal{N}}(y) > 1$ if y is outside of $\mathcal{N}$. $\mu_{\mathcal{N}}$ is called the Minkowski functional for $\mathcal{N}$.

It follows that each $\mu_{\mathcal{N}}$ is a semi-norm on $\mathcal{X}$ and the collection $\{\mu_{\mathcal{N}}\}_{\mathcal{N} \in \mathcal{N}_{\mathcal{X}}}$ of semi-norms on $\mathcal{X}$ generate the topology on $\mathcal{X}$. The Minkowski functional $\mu_{\mathcal{N}}$ can distinguish points inside $\mathcal{N}$ from the points outside. For a set $\mathcal{M}$ in $\mathcal{X}$, we denote by $\langle\!\langle \mathcal{M} \rangle\!\rangle$ and $\langle\!\langle\!\langle \mathcal{M} \rangle\!\rangle\!\rangle$ the convex and absolutely convex closures respectively of $\mathcal{M}$.

Theorem 1.1.2. (Separation Theorem) *Let $\mathcal{X}$ be a locally convex space, and let $\mathcal{M}$ be a compact set with non-empty interior. Let $\mathcal{N}$ be such that $\langle\!\langle \mathcal{M} \rangle\!\rangle \cap \langle\!\langle \mathcal{N} \rangle\!\rangle$ is empty. Then there exists a functional φ in $\mathcal{X}^*$ such that*

$$\sup_{x \in \mathcal{M}} \Big[\Re\big(\varphi(x)\big)\Big] < \lambda_{[\mathcal{M}:\mathcal{N}]} < \inf_{y \in \mathcal{N}} \Big[\Re\big(\varphi(y)\big)\Big],$$

for some $\lambda_{[\mathcal{M}:\mathcal{N}]}$ in $\mathbb{R}$. In such a case, we say φ separates $\mathcal{M}$ and $\mathcal{N}$.

Proof

Recall that for any complex linear functional ω we must have, $\omega(x) = \Re(\omega(x)) - i\Re(\omega(ix))$. Hence the real part $\Re(\omega)$ already determines the functional ω. Hence we can focus on real vector spaces. We now separate the two convex set by constructing a linear functional using a line joining the two convex sets. If $\mathcal{M}$ has non-empty interior, then we can assume by translation that the origin is an internal point of $\mathcal{M}$.

Let y be in $\mathcal{N}$. The Minkowski function can distinguish points inside a convex set and those outside it, the set $\mathcal{K} \equiv \langle\!\langle \mathcal{M} \rangle\!\rangle - \langle\!\langle \mathcal{N} \rangle\!\rangle + y$ is a convex set. By construction $\mathcal{K}$ does not contain y, and if $\mu_{\mathcal{K}}$ is the Minkowski functional associated with $\mathcal{K}$, then we must have

$\mu_{\mathcal{K}}(y) \geqq 1$. The line joining origin which is contained in $\langle\!\langle \mathcal{M} \rangle\!\rangle$ and y is give by $\mathbb{R}y$, and we can construct a functional using this line by,

$$\begin{aligned} \varphi_y : \mathbb{R}y &\to \mathbb{R} \\ \lambda y &\mapsto \lambda \mu_{\mathcal{K}}(y). \end{aligned}$$

Clearly φ_y is a linear functional majorised by $\mu_{\mathcal{K}}$, and by Hahn-Banach theorem it must be a restriction of a $\mu_{\mathcal{K}}$-majorised functional on $\mathcal{X}$, say φ. Then we have, $\varphi(y) = \mu_{\mathcal{K}}(y) \geqq 1$, and $\varphi(x) \leqq \mu_{\mathcal{K}}(x) \leqq 1$ for all x in $\mathcal{K}$. Hence we must have,

$$\sup_{x \in \mathcal{M}} \Big[\Re\big(\varphi(x)\big) \Big] \leqq \inf_{y \in \mathcal{N}} \Big[\Re\big(\varphi(y)\big) \Big].$$

Assume by translation again that the origin is the interior of $\mathcal{M}$. Since $\mathcal{X}$ is a locally convex space there exists a absolutely convex neighborhood $U_{\mathcal{M}}$ centered at origin entirely contained inside $\mathcal{M}$. By construction above we must have,

$$\Re\big(\varphi(U_{\mathcal{M}})\big) \leqq \sup_{x \in \mathcal{M}} \Big[\Re\big(\varphi(x)\big) \Big] \equiv S_{\varphi}.$$

For any ϵ in $\mathbb{R}^+$, let $\mathcal{M}_\epsilon \equiv (\epsilon / S_\varphi) U_{\mathcal{M}}$, then by linearity it follows that $\Re\big(\varphi(\mathcal{M}_\epsilon)\big) \leqq \epsilon$. Hence $\Re(\varphi)$ is continuous at origin. Since $\mathcal{X}$ is a topological vector space, continuity of a linear map at one point implies its continuity everywhere. Hence it follows that $\Re(\varphi)$ as constructed above must be continuous, whenever $\langle\!\langle \mathcal{M} \rangle\!\rangle$ has non-empty interior.

Since $\mathcal{M}$ is compact $\Re(\varphi)$ attains its the supremum on $\mathcal{M}$, and since $\langle\!\langle \mathcal{M} \rangle\!\rangle \cap \langle\!\langle \mathcal{N} \rangle\!\rangle$ is empty we must have $\sup_{\mathcal{M}} \Re(\varphi) < \inf_{\mathcal{N}} \Re(\varphi)$. In particular there must exist some $\lambda_{[\mathcal{M}:\mathcal{N}]}$ in $\mathbb{R}$ such that,

$$\sup_{x \in \mathcal{M}} \Big[\Re\big(\varphi(x)\big) \Big] < \lambda_{[\mathcal{M}:\mathcal{N}]} < \inf_{y \in \mathcal{N}} \Big[\Re\big(\varphi(y)\big) \Big].$$

This completes the proof. □

The set of all x in $\mathcal{X}$ with $\Re\big(\varphi(x)\big) = \lambda_{[\mathcal{M}:\mathcal{N}]}$ defines a hyperplane in $\mathcal{X}$ such that $\langle\!\langle \mathcal{M} \rangle\!\rangle$ and $\langle\!\langle \mathcal{N} \rangle\!\rangle$ lie on its opposite sides. We say that the hyperplane $\Re\big(\varphi(x)\big) = \lambda_{[\mathcal{M},\mathcal{N}]}$ separates $\langle\!\langle \mathcal{M} \rangle\!\rangle$ and $\langle\!\langle \mathcal{N} \rangle\!\rangle$. If $\langle\!\langle\!\langle \mathcal{M} \rangle\!\rangle\!\rangle \cap \langle\!\langle\!\langle \mathcal{N} \rangle\!\rangle\!\rangle$ is non-empty, then there exist φ in $\mathcal{X}$ and

$$\sup_{x \in \mathcal{M}} \Big[\big|\varphi(x)\big| \Big] < \lambda_{[\mathcal{M}:\mathcal{N}]} < \inf_{y \in \mathcal{N}} \Big[\big|\varphi(y)\big| \Big].$$

The above proof can be used for this case by using the complex version of Hahn-Banach theorem instead of the real version.

Weak & Weak* Topologies

Continuous functions relate the topologies of their domain with the topology of their range. By relating the topologies, continuous functions measure the nearness of points in the domain, via the nearness of points in the range. Let $\mathcal{X}$ and $\mathcal{Y}$ be two sets. If there exists a non-degenerate bilinear pairing

$$\big\langle \cdot \,|\, \cdot \big\rangle : \mathcal{X} \times \mathcal{Y} \to \mathbb{C}$$

then we say $\langle \mathcal{X} | \mathcal{Y} \rangle$ is a dual pair. The elements of $\mathcal{Y}$ define a topology on $\mathcal{X}$, and dually elements of $\mathcal{X}$ define a topology on $\mathcal{Y}$. Such topologies are called weak topologies, and are the universal topologies on the spaces making the pairing continuous.

A locally convex topology τ on $\mathcal{X}$ is said to be $\langle \mathcal{X}|\mathcal{Y}\rangle$-compatible if the space of all continuous functionals on $\mathcal{X}$ with respect to the topology τ is $\mathcal{Y}$. That is, $(\mathcal{X}, \tau)^* \equiv \mathcal{Y}$. If $\mathcal{X}$ already comes equipped with a locally convex topology, then we denote by $\mathcal{X}^*$ the space of all continuous linear functionals on $\mathcal{X}$. The evaluation pairing is called the conjugate pairing and is denoted by,

$$\big\langle x|\varphi\big\rangle_{\mathcal{X}} \equiv \varphi(x).$$

It follows that

$$\big\langle \mathcal{X}|\mathcal{X}^*\big\rangle_{\mathcal{X}}$$

is a dual pair.

The weak topology on $\mathcal{X}^*$ is such that $\{\varphi \mapsto \varphi(x)\}_{x\in\mathcal{X}}$ is a continuous family of functions. The universal topology on $\mathcal{X}^*$ with this property is the one induced by the collection of seminorms,

$$\sigma(\mathcal{X}^*, \mathcal{X}) \equiv \Big\{\varphi \mapsto \big|\varphi(x)\big|\Big\}_{x\in\mathcal{X}}$$

Since semi-norms are convex functions, the conjugate space $\mathcal{X}^*$ equipped with this topology is a locally convex space. Dually, the smallest topology on $\mathcal{X}$ making $x \mapsto \varphi(x)$ continuous for all φ in $\mathcal{X}^*$ is called the weak topology on $\mathcal{X}$. The topology on $\mathcal{X}^*$ generated by $\sigma(\mathcal{X}^*, \mathcal{X})$ is called the weak*-topology.

The second conjugate space $\mathcal{X}^{**}$ of $\mathcal{X}$, is the space of all continuous linear functionals on $\mathcal{X}^*$ equipped with the weak*-topology. In such a case, we have a canonical embedding of $\mathcal{X}$ inside $\mathcal{X}^{**}$,

$$\begin{aligned} \iota_{\mathcal{X}} \,:\, \mathcal{X} &\to \mathcal{X}^{**} \\ x &\mapsto \big\langle x \,|\, \cdot\,\big\rangle_{\mathcal{X}}. \end{aligned}$$

where $\big\langle x|\,\cdot\,\big\rangle_{\mathcal{X}} : \varphi \mapsto \langle x|\varphi\rangle_{\mathcal{X}}$ is a continuous function on $\mathcal{X}^*$. We will denote $\langle\cdot|\cdot\rangle_{\mathcal{X}}$ by $\langle\cdot|\cdot\rangle$ whenever there is no confusion.

1.1.2 | Conjugate Duality

Let $\mathcal{X}$ be a locally convex space. A function f from $\mathcal{X}$ to $\mathbb{R}$ is convex if it maps line segments in $\mathcal{X}$ to line segments in $\mathbb{R}$, equivalently characterised by the Jensen inequality; $f(\lambda x + (1-\lambda)y) \leqq \lambda f(x) + (1-\lambda)f(y)$. Such functions describe convex sets $\mathcal{M}_{\lambda_f}$ of all x in $\mathcal{X}$ with $f(x) \leqq \lambda_f$ for each real number $\lambda_f \in \mathbb{R}$. Convex functions can compress or stretch the line segments in $\mathcal{X}$. The interesting case is when the compression is not continuous, that is parts of line segments are collapsed completely. In such cases $\mathcal{M}_{\lambda_f}$ can contain full affine subspaces. The convexity properties contained in such subspaces cannot be seen by the convex function. Conjugate duality allows us to quantify the convexity unseen by a given convex function. One may study how much more convexity the set of affine functionals see compared to a given convex function. The total convexity seen by the linear functionals on $\mathcal{X}$ can be quantified by the taking supremum over all points of $\mathcal{X}$. We now make this precise by defining the conjugate function of a function.

Given a function $f : \mathcal{X} \to \overline{\mathbb{R}}$, the conjugate function or the Legendre-Fenchel transform of f is the map $\mathcal{F}f : \mathcal{X}^* \to \overline{\mathbb{R}}$ defined by

$$\mathcal{F}f(\varphi) \equiv (f^*)(\varphi) \coloneqq \sup_{x\in\mathcal{X}} \Big[\Re\big\langle x|\varphi\big\rangle - f(x)\Big].$$

$f^*(\varphi)$ quantifies how much less convexity the function f sees when compared to the linear functional χ. If we take f to be $\Re\langle\cdot|\varphi\rangle$, then $f^*(\varphi)$ vanishes and f is just as convex as φ. If

$g : \mathcal{X}^* \to \overline{\mathbb{R}}$ is a function then we define the coconjugate of g by,

$$\mathcal{G}g(x) \equiv \big(g_*\big)(x) = \sup_{\varphi \in \mathcal{X}^*} \Big[\Re\big\langle x|\varphi\big\rangle - g(\varphi)\Big].$$

As the supremum of affine linear functionals the Legendre transform is a convex function. Since f^* is the pointwise supremum of weak*-continuous affine functions, the conjugate function f^* is a weak*-continuous function on $\mathcal{X}^*$. The coconjugate g_* is a weakly continuous function on $\mathcal{X}$.

We immediately obtain the following properties of the Legendre transform;

$$\begin{aligned}
&\mathcal{F}(f+c)(\chi) = \mathcal{F}f(\chi) - c\\
&\mathcal{F}f(\chi - x) = \mathcal{F}f(\chi) + \langle x, \chi\rangle\\
&\mathcal{F}f(t\chi) = \mathcal{F}f(\chi/t)\\
&\mathcal{F}(tf)(\chi) = t \cdot \mathcal{F}f(\chi/t).
\end{aligned}$$

Similarly for $\mathcal{G}$. By definition of the Legendre-Fenchel transform, we must have

$$f(x) + f^*(\varphi) \geqq \Re\big\langle x|\varphi\big\rangle.$$

This inequality is called the Fenchel-Young inequality.

We can also define absolute Legendre transform to focus on absolute convexity instead of convexity. In such a case, the Legendre transform of interest can be defined to be;

$$f^*(\varphi) \equiv \sup_{x \in \mathcal{X}} \Big[\big|\big\langle x|\varphi\big\rangle\big| - f(x)\Big].$$

We will use either of these Legendre transforms depending on whether the context is about convexity or absolute convexity, and only explicitly indicate which one we are using if there would be confusion.

1.1.2.1 | Properness & Lower Semi-Continuity

Topology and the vector space structure on the real line are closely related to each other. If a convex function compresses certain line segments in the space, then it should be possible to notice it from the topological properties of the real line. Hence the compression issues of convex functions can be studied by their continuity properties. This is possible in locally convex spaces since the topology is generated by neighborhood bases of convex sets, and hence the convexity and topological properties of locally convex spaces can be related with convexity and topological properties on the real line.

We must also ensure that f does not stretch line segments to infinity. A function f is said to be proper if it is not always ∞ and is never $-\infty$. The set of all x with $f(x) < \infty$ is called the domain of f and denoted by $\mathcal{D}_f$. For a proper function f on $\mathcal{X}$ we define the conjugate function as,

$$\mathcal{F}f(\varphi) \equiv f^*(\varphi) \coloneqq \sup_{x \in \mathcal{D}_f} \Big[\Re\big\langle x|\varphi\big\rangle_{\mathcal{X}} - f(x)\Big].$$

Similarly if g is a proper function on $\mathcal{X}^*$ then we define the coconjugate of g as

$$\mathcal{G}g(x) \equiv g_*(x) = \sup_{\varphi \in \mathcal{D}_g} \Big[\Re\big\langle x|\varphi\big\rangle - g(\varphi)\Big].$$

A convex functions f on $\mathcal{X}$ is said to be lower semi-continuous if for every x with $\lambda < f(x)$ there exists an open neighborhood N_x of x with $\lambda < f(N_x)$, which means that for every y in N_x we have $\lambda < f(y)$.

As a supremum over affine functions, for any function f on $\mathcal{X}$, the conjugate function f^* is lower semi-continuous function on $\mathcal{X}^*$ with respect to the weak*-topology. We denote the collection of all lower semi-continuous functions on $\mathcal{X}$ by $\Gamma(\mathcal{X})$. We will denote the collection of all lower semi-continuous proper convex functions on $\mathcal{X}$ denoted by $\overline{\Gamma}(\mathcal{X})$.

Lemma 1.1.3. *If f is in $\overline{\Gamma}(\mathcal{X})$ then f^* is in $\overline{\Gamma}(\mathcal{X}^*)$ and the Legendre-Fenchel conjugate transform is an injective mapping;*

$$\mathcal{F} : \overline{\Gamma}(\mathcal{X}) \rightarrowtail \overline{\Gamma}(\mathcal{X}^*)$$

In particular, then there exists φ_f in $\mathcal{X}^*$ such that

$$\begin{aligned} f(x) - f(y) &\geqq \Re\big[\big\langle x - y|\varphi_f\big\rangle\big]; \\ f^*(\varphi_f) &\leqq \Re\big\langle y|\varphi_f\big\rangle - f(y). \end{aligned}$$

Proof

Since f is lower semi-continuous, then the set of all (x, λ) with $f(x) \leqq \lambda$ is a closed convex set of $\mathcal{X} \oplus \mathbb{R}$ denoted by $\mathcal{M}_f$. Any (y, μ) with $\mu < f(y)$ lies outside, and hence by Hahn-Banach separation theorem, there exists a continuous linear functional φ on $\mathcal{X} \oplus \mathbb{R}$ such that,

$$\sup_{\mathcal{M}_f} \Big[\Re\big\langle (x, \lambda)|\varphi\big\rangle_{\mathcal{X}\oplus\mathbb{R}}\Big] < \Big[\Re\big\langle (y, \mu)|\varphi\big\rangle_{\mathcal{X}\oplus\mathbb{R}}\Big].$$

The functional φ can be decomposed as an extension of the linear functional φ on $\mathcal{X}$ with, $\varphi(x, \lambda) = \varphi(x) + \lambda\lambda_\varphi$ where $\lambda_\varphi \equiv \varphi(0, 1)$, for simplicity we denote $\Re(\lambda_\varphi) = \lambda_\varphi$. Hence we must have

$$\sup_{f(x)\leqq\lambda} \big[\Re\big\langle x|\varphi\big\rangle + \lambda\lambda_\varphi\big] < \Re\big\langle y|\varphi\big\rangle + \mu\lambda_\varphi.$$

If f is proper it must be bounded below. By taking $-\mu$ large enough, we infer that λ_φ is negative. Hence by taking $\lambda = f(x)$ and $\mu = f(y) - \epsilon$ we must have; $\Re[\langle x|\varphi\rangle] + \lambda_\varphi f(x) \leqq \Re[\langle y|\varphi\rangle] + \lambda_\varphi f(y)$. By rescaling φ, and considering the functional $\varphi_f \equiv -\lambda_\varphi^{-1}\varphi$, it follows that

$$f(x) - f(y) \geqq \Re\big[\big\langle x - y|\varphi_f\big\rangle\big]$$

The evaluation of the conjugate function f^* at φ_f is

$$f^*(\varphi_f) = \sup_{x\in\mathcal{D}_f} \Big[\Re\big\langle x|\varphi_f\big\rangle - f(x)\Big] \leqq \Re\big\langle y|\varphi_f\big\rangle - f(y).$$

Hence f^* is a proper function. This completes the proof of the lemma. □

The lemma says that is every f in $\overline{\Gamma}(\mathcal{X})$ is bounded below by an affine linear functional. In the absence of lower semi-continuity we would have $f(x) \geqq \Re\langle x - y|\varphi_f\rangle + \mu_f$, and here the shortest line segment joining the point y and $\mathcal{M}_f$ gets mapped by φ_μ to the line segment joining $f(y)$ and μ in $\mathbb{R}$.

We now show that it is indeed a bijection by showing $\mathcal{G} \circ \mathcal{F}$ is the identity map on $\overline{\Gamma}(\mathcal{X})$. $\mathcal{G} \circ \mathcal{F}f$ is called the biconjugate of f and is given by

$$(\mathcal{G} \circ \mathcal{F}f)(x) \equiv (f)^*_*(x) \equiv \sup_{\varphi\in\mathcal{D}_{\mathcal{F}f}} \Big[\Re\big\langle x|\varphi\big\rangle - f^*(\varphi)\Big].$$

Clearly, $(f)^*_*$ is a proper lower semi-continuous convex function that is, $(\mathcal{G} \circ \mathcal{F}f)$ is in $\overline{\Gamma}(\mathcal{X})$. We obtain the Fenchel-Young inequality for biconjugate from the above expression;

$$(f)^*_*(x) + f^*(\chi) \geqq \Re\langle x|\chi\rangle.$$

We now prove the duality theorem due to Fenchel-Moreau-Rockafellar;

Theorem 1.1.4. (Conjugate Duality)

$$\overline{\Gamma}(\mathcal{X}) \xrightarrow[\mathcal{F}]{\cong} \overline{\Gamma}(\mathcal{X}^*).$$

Proof

We show that $\mathcal{G}$ is the inversion of $\mathcal{F}$. That is, for every f in $\overline{\Gamma}(\mathcal{X})$, we now show that $(f)^*_*(x) = f(x)$ for all x in $\mathcal{X}$. By definition of the Legendre-Fenchel transform, we must have

$$f(x) + f^*(\varphi) \geqq \Re\langle x|\varphi\rangle.$$

Combining this with the definition of the biconjugate we obtain,

$$f(x) \geqq \sup_{\varphi \in \mathcal{D}_{\mathcal{F}f}} \Big[\Re\langle x|\varphi\rangle - f^*(\varphi)\Big] \equiv (f)^*_*(x).$$

From the discussion before the previous lemma, we have that whenever f is a lower semi-continuous proper convex function, there exists some φ_f such that $f(x) - f(y) \geqq \Re\langle x|\varphi_f\rangle - \Re\langle y|\varphi_f\rangle$ and $f(y) \leqq \Re\big\langle y|\varphi_f\big\rangle - f^*(\varphi_f)$. By definition of the biconjugate as supremum over all such linear functionals, we must have, $\Re\langle x|\varphi_f\rangle - f^*(\varphi_f) \leqq (f)^*_*(x)$ for all x in its domain. Hence we have,

$$f(x) \geqq (f)^*_*(x).$$

Combining with the above inequality we have $(f)^*_*(x) = f(x)$ for all x in $\mathcal{D}_f$.

We now show that if $f(x)$ is infinite, then $(f)^*_*(x)$ must also be infinite. If φ_f is as above, it must belong to the domain of f^*. For any other functional φ in its domain consider the functional $\chi = \varphi + \lambda\varphi_f$. By definition of the conjugate function we have,

$$\begin{aligned} f^*(\chi) &= \sup_{y \in \mathcal{D}_f} \Big[\Re\big\langle y|\varphi\big\rangle + \lambda\Re\big\langle y|\varphi_f\big\rangle - f(y)\Big] \\ &\leqq \underbrace{\sup_{y \in \mathcal{D}_f} \Big[\Re\big\langle y|\varphi\big\rangle - f(y)\Big]}_{f^*(\varphi)} + \lambda \sup_{y \in \mathcal{D}_f} \Big[\Re\big\langle y|\varphi_f\big\rangle\Big]. \end{aligned}$$

Hence we have;

$$f^*(\chi) - f^*(\varphi) \leqq \sup_{y \in \mathcal{D}_f} \Big[\Re\big[\big\langle y|\chi - \varphi\big\rangle\big]\Big] = \lambda \sup_{y \in \mathcal{D}_f} \Big[\Re\big\langle y|\varphi_f\big\rangle\Big].$$

That is, $\chi = \varphi + \lambda\varphi_f$ is in the domain of f^*. By the Fenchel-Young inequality for biconjugate function we have, $(f)^*_*(x) + f^*(\chi) \geqq \Re\big\langle x|\chi\big\rangle$. Substituting for $f^*(\chi)$ we obtain,

$$(f)^*_*(x) \geqq \Big[\Re\langle x|\varphi\rangle_{\mathcal{X}} - f^*(\varphi)\Big] + \lambda\Big[\inf_{y \in \mathcal{D}_f} \big[\Re\langle x - y|\varphi_f\rangle\big]\Big].$$

Since $f(x) - f(y) \geqq \Re\langle x - y|\varphi_f\rangle$ it follows that

$$(f)^*_*(x) \geqq \lambda_{x,\varphi} + \lambda\big[f(x) - f(y)\big]$$

where $\lambda_{x,\varphi} = \Re\langle x|\varphi\rangle - f^*(\varphi)$ is a finite. Since f is a proper function, there must exist some y for which $f(y)$ is finite. Since λ can be arbitrarily large it follows that $(f)^*_*(x) \geqq \infty$, whenever $f(x) = \infty$. This completes the proof. □

Since there exists a natural embedding $\iota_{\mathcal{X}}$ of $\mathcal{X}$ into $\mathcal{X}^{**}$, for every function h on $\mathcal{X}^{**}$, we obtain a function on $\mathcal{X}$ given by the composition of the mappings; $h \circ \iota_{\mathcal{X}} : \mathcal{X} \to \mathbb{R}$. Let $\mathbb{R}^{\mathcal{X}}$ and $\mathbb{R}^{\mathcal{X}^{**}}$ denote the space functions from $\mathcal{X}$ to $\mathbb{R}$ and the space of functions from $\mathcal{X}^{**}$ to $\mathbb{R}$ respectively. The natural embedding $\iota_{\mathcal{X}} : \mathcal{X} \rightarrowtail \mathcal{X}^{**}$ gives us a mapping;

$$\begin{aligned} |_{\mathcal{X}} : \mathbb{R}^{\mathcal{X}^{**}} &\to \mathbb{R}^{\mathcal{X}} \\ h &\mapsto h \circ \iota_{\mathcal{X}}. \end{aligned}$$

We can take iterated compositions of Legendre-Fenchel transforms,

$$\cdots \leftrightharpoons \overline{\Gamma}(\mathcal{X}) \underset{\mathcal{G}}{\overset{\mathcal{F}}{\leftrightharpoons}} \overline{\Gamma}(\mathcal{X}^*) \underset{\mathcal{G}}{\overset{\mathcal{F}}{\leftrightharpoons}} \overline{\Gamma}(\mathcal{X}^{**}) \leftrightharpoons \cdots$$

where each arrow is an isomorphism. Since $\mathcal{G} \circ \mathcal{F}$ is an isomorphism on $\overline{\Gamma}(\mathcal{X})$ it follows that for any f in $\overline{\Gamma}(\mathcal{X})$, we must have, $f(x) = (f)^*_*(x) = (f^*)^*(\iota_{\mathcal{X}}(x))$.

$$|_{\mathcal{X}} : \overline{\Gamma}(\mathcal{X}^{**}) \to \overline{\Gamma}(\mathcal{X}).$$

Hence it follows that $|_{\mathcal{X}} \equiv \mathcal{G} \circ \mathcal{G}$. Hence, the conjugate transform relates the weak topology on $\mathcal{X}$ and $\mathcal{X}^{**}$. Hence we have proved the following lemma;

Corollary 1.1.5. $|_{\mathcal{X}} : \overline{\Gamma}(\mathcal{X}^{**}) \to \overline{\Gamma}(\mathcal{X})$ *is an isomorphism with inverse* $\mathcal{F} \circ \mathcal{F}$.

The conjugate function f^* of f is lower semi-continuous regardless of the continuity properties of f. Hence $(f)^*_*$ acts as weakly lower semi-continuous approximate of f. If the second conjugate $(f)^*_*$ is identical to f, then it immediately follows that the function is a lower semi-continuous function.

1.2 | Locally Convex Topologies

Since every locally convex space has an absolutely convex neighborhoods of origin which generates the topology, the Minkowski functionals associated to such neighborhoods are very closely related to the topology and convexity properties of the locally convex space. Hence we can expect to study the topological and convexity properties in terms of the Minkowski functionals. We now make this duality precise by studying how to transfer the topological and convexity properties from the locally convex space to the functions on the space via Minkowski functionals, and back.

A function is said to be affine functional on a set $\mathcal{K}$ if it maps line segments joining points in $\mathcal{K}$ to line segments in the target space. The restriction of an affine functional on $\mathcal{X}$ to $\mathcal{K}$ is an affine functional on $\mathcal{K}$, and conversely, every affine functional, can be extended by linearity, and using the Hahn-Banach theorem to an affine functional on $\mathcal{X}$. The space of affine functionals on $\mathcal{K}$ is denoted by $\mathcal{A}(\mathcal{K})$. We have,

$$\mathcal{A}(\mathcal{X}) \cong \mathcal{X}^* \oplus \mathbb{R}.$$

Any affine functional α in $\mathcal{A}(\mathcal{K})$ is a restriction of a functional $\Re(\varphi)$ to the set $\mathcal{K} \times \{0\}$.

If f is in $\overline{\Gamma}(\mathcal{X})$, then by conjugate duality,

$$(f)^*_*(x) = \sup_{\varphi \in \mathcal{D}_{f^*}} \Big[\Re\langle x|\varphi\rangle - f^*(\varphi) \Big] = f(x).$$

Hence f can be written as a supremum of affine continuous functions. Since each affine continuous functional is proper lower semi-continuous conex function, the supremum of a collection of affine continuous functionals must also be proper lower semi-continuous. Hence we have proved the following lemma;

Lemma 1.2.1. *A function f is in $\overline{\Gamma}(\mathcal{X})$ if and only if*

$$f(x) = \sup_{\alpha \leq f} \big[\alpha(x)\big],$$

where α are continuous affine functionals on the domain of f.

Note that $(f)^*_*$ is weakly lower semi-continuous. $(f)^*_*$ must always be supremum of affine functionals, and hence must possess nice convexity properties. We also know by conjugate duality that the $(f)^*_*$ is identical to f whenever the proper function f is lower semi-continuous convex function. Hence the conjugate duality provides a duality between convexity properties and topological properties.

1.2.0.1 | Supporting Functions & Indicators

Since convex functions on $\mathcal{X}$ are closely related to convex subsets of $\mathcal{X}$, the conjugate duality also establishes a duality between convex subsets of $\mathcal{X}$ and $\mathcal{X}^*$ with appropriate closure properties. To make the relation between convex sets and convex functions precise, we define the notion of supporting functions. We will assume $\mathcal{X}$ is a locally convex space, and $\mathcal{N}_{\mathcal{X}}$ is an convex neighborhood basis of origin of $\mathcal{X}$.

Convex functions are closely related to convex sets in the space. For any convex function f on a locally convex space $\mathcal{X}^*$, the set of all χ such that $f(\chi) \leqq \lambda$ is a convex set. A lower semi-continuous proper convex function f is said to be positively homogeneous if for every t in $\mathbb{R}_+$ the function satisfies,

$$f(\lambda\varphi) = \lambda f(\varphi).$$

Positively homogeneous convex functions are well-behaved around the origin.

Since locally convex topologies are closely related to the convex neighborhoods of origin, we may expect to describe the semi-norms in terms of positively homogeneous functions in $\overline{\Gamma}(\mathcal{X})$. Suppose a function f in $\overline{\Gamma}(\mathcal{X})$ is positively homogeneous, then $f(\lambda x) = \lambda f(x)$ for all λ in $\mathbb{R}^+$ and we must have,

$$\begin{aligned} f^*(\varphi) &= \sup_{x \in \mathcal{X}} \Big[\big|\big\langle x|\varphi\big\rangle_{\mathcal{X}}\big| - f(x) \Big] \\ &= \sup_{x \in \mathcal{X}} \Big[\big|\big\langle \lambda x|\varphi\big\rangle_{\mathcal{X}}\big| - f(\lambda x) \Big] = \lambda f^*(\varphi). \end{aligned}$$

Hence f^* can only take the values 0 or ∞. Since f^* is a convex function its domain $\mathcal{D}_{f^*}$ is a closed convex set which contains origin. Hence we must have,

$$(f)^*_*(x) = \sup_{\varphi \in \mathcal{D}_{f^*}} \Big[\big|\big\langle x|\varphi\big\rangle\big| \Big] = f(x).$$

We denote the positively homogeneous functions in $\overline{\Gamma}(\mathcal{X})$ by $\overline{\Gamma}_+(\mathcal{X})$. Every function in $\overline{\Gamma}_+(\mathcal{X})$ is the supremum of linear functionals in the domain of f^*. Hence we have proved the following lemma;

Lemma 1.2.2. *If f is in $\overline{\Gamma}_+(\mathcal{X})$ if and only if*

$$f(x) = \sup_{\varphi \in \mathcal{K}} \Big[\big| \langle x|\varphi \rangle_{\mathcal{X}} \big| \Big],$$

for a convex neighborhood $\mathcal{K}$ of origin in $\mathcal{X}^*$.

The semi-norms of a locally convex topological space $\mathcal{X}$ correspond to positively homogeneous continuous proper convex functions with respect to the topology. By the above lemma, every semi-norm must be a supremum over convex neighborhood of origin in $\mathcal{X}^*$. Hence the conjugate space allows us to describe the topology in terms of its neighborhoods of origin. For any subset $\mathcal{N}$ of $\mathcal{X}$ the supporting function of $\mathcal{N}$ is defined by,

$$H_{\mathcal{N}}(\varphi) = \sup_{x \in \mathcal{N}} \Big[\big| \langle x|\varphi \rangle \big| \Big].$$

By linearity of each φ, it follows immediately that

$$H_{\mathcal{N}} \equiv H_{\langle\!\langle\!\langle \mathcal{N} \rangle\!\rangle\!\rangle}.$$

Hence the supporting functions can see sets upto, (absolutely) convex and topological closures. Clearly, supporting functions are positively homogeneous. As supremum over continuous functionals, supporting functions of non-empty sets are lower semi-continuous proper convex functions. Hence we have, $H_{\mathcal{N}} = (H_{\mathcal{N}})^*_*$. The Legendre transform of $H_{\mathcal{N}}$ is called the indicator of $\mathcal{N}$ and is denoted by $I_{\mathcal{N}}$. We have,

$$I_{\mathcal{N}} = (H_{\mathcal{N}})^* \,;\, H_{\mathcal{N}} = (I_{\mathcal{N}})_*.$$

Upon expanding the definition of Legendre-Fenchel transform for $H_{\mathcal{N}}$ we obtain,

$$I_{\mathcal{N}}(x) \equiv \sup_{\varphi \in \mathcal{X}^*} \inf_{y \in \mathcal{N}} \Big[\big| \langle x|\varphi \rangle \big| - \big| \langle y|\varphi \rangle \big| \Big].$$

Hence the indicator $I_{\mathcal{N}}$ must vanish inside $\mathcal{N}$, and be infinite outside.

1.2.0.2 | Minkowski Functionals & Duality

The Minkowski functional associated with absolutely convex neighborhoods basis of origin of a locally convex space generate its topology, and they are convenient to study since we can apply conjugate duality to Minkowski functionals. However we usually do not have the neighborhood basis to work with. Hence it is useful to have tools that allow us to construct such absolutely convex neighborhoods of origin starting from functions on the space.

We can use this property to construct absolutely convex neighborhoods around origin of $\mathcal{X}^*$, using the supporting functions. Let $\mathcal{N}$ be any set in $\mathcal{X}$, and let $H_{\mathcal{N}}$ be its supporting function. Then for any λ, let $\mathcal{N}^\lambda$ denote the absolutely convex set of all χ in $\mathcal{X}^*$ such that $H_{\mathcal{N}}(\chi) \leqq \lambda$. We can represent $\mathcal{N}^\lambda$ more conveniently in terms of its indicator as;

$$I_{\mathcal{N}^\lambda} \equiv I_{[H_{\mathcal{N}} \leqq \lambda]},$$

where $[H_{\mathcal{N}} \leqq \lambda]$ is the set of all x such that $H_{\mathcal{N}}(x) \leqq \lambda$. Since $H_{\mathcal{N}}$ is positively homogeneous it follows that K^1 is a closed absolutely convex neighborhood of origin. $\mathcal{N}^1$ is called the polar of $\mathcal{N}$ and is denoted by $\mathcal{N}^\circ$. We have,

$$I_{\mathcal{N}^\circ} \equiv I_{[H_{\mathcal{N}} \leqq 1]}.$$

Similarly, for any subset $\mathcal{M}$ of $\mathcal{X}^*$, we define the polar $\mathcal{M}_\circ$ of $\mathcal{M}$ in $\mathcal{X}$ to be the absolutely convex neighborhood of origin given by;

$$I_{\mathcal{M}_\circ} \equiv I_{[H_{\mathcal{M}} \leqq 1]}.$$

Hence we can obtain absolutely convex neighborhoods of origin using the subsets of the conjugate pair of the space, and using the conjugate function of the supporting function of the subset.

Relation to Supporting Functions

The idea is to use the subsets of the conjugate space and use the supporting functions on the conjugate space as an alternative way to construct the Minkowski functionals. We denote by $\mathcal{N}_{\mathcal{X}}$ and $\mathcal{N}_{\mathcal{X}^*}$ the absolutely convex neighborhood basis of origin which generate the locally convex topologies $\mathcal{X}$ and $\mathcal{X}^*$ respectively. By taking the interior of closed sets, or by taking closures of open sets, we can study the topologies by either by their closed sets or open sets.

The Minkowski functional $\mu_{\mathcal{N}}$ corresponding to an absolutely convex neighborhood $\mathcal{N}$ of origin is corresponds to be the function which assigns to each x, the minimal descaling of x needed to bring it inside $\mathcal{N}$. It is hence defined by,

$$\mu_{\mathcal{N}}(x) = \inf_{x \in \lambda\mathcal{N}} \big[|\lambda|\big].$$

The Minkowski functional is a semi-norm, and for every λ in $\mathbb{R}^+$,

$$\mu_{\mathcal{N}}(\lambda x) = \lambda \mu_{\mathcal{N}}(x).$$

The Minkowski functional is always greater than 1 outside of $\mathcal{N}$ and less than 1 inside of $\mathcal{N}$. Hence, the Minkowski functional can also be expressed as an infimum over all indicators for neighborhoods $\lambda\mathcal{N}$, and for every λ and hence we have,

$$\mu_{\mathcal{N}}(x) = \inf_{x \in \lambda\mathcal{N}} \big[|\lambda|\big] = \inf_{\lambda \in \mathbb{R}^+} \big[I_{\lambda\mathcal{N}}(x) + \lambda\big].$$

This is justified because if x belongs $\lambda\mathcal{N}$ then $I_{\lambda\mathcal{N}}$ is zero, and if λ is too small and x lies outside of $\lambda\mathcal{N}$ then $I_{\lambda\mathcal{N}}$ is infinite. Hence the infimum is attained for the minimal λ for which x lies inside $\lambda\mathcal{N}$. By the definition of the conjugate function of $\mu_{\mathcal{N}}$ is then given by,

$$\begin{aligned}(\mu_{\mathcal{N}})^*(\chi) &= \sup_{\lambda \in \mathbb{R}^+} \big[H_{\lambda\mathcal{N}}(\chi) - \lambda\big] \\ &= \sup_{\lambda \in \mathbb{R}^+} \big[\lambda(H_{\mathcal{N}}(\chi) - 1)\big].\end{aligned}$$

For the last step; $\lambda(H_{\mathcal{N}}(\chi) - 1)$ is positive whenever χ is outside of $\mathcal{N}^\circ$, in which case the above supremum will be infinite. When χ is inside $\mathcal{N}^\circ$, $H_{\mathcal{N}}(\chi) - 1$ is negative, and hence the supremum must be zero. Hence we must have,

$$(\mu_{\mathcal{N}})^* = I_{\mathcal{N}^\circ}.$$

Similarly, if $\mathcal{M}$ is in $\mathcal{N}_{\mathcal{X}^*}$, then we must have

$$\left(\mu_{\mathcal{M}}\right)_* = I_{\mathcal{M}_\circ}.$$

Since the polar sets are absolutely convex neighborhoods of origin, we can define their Minkowski functionals, and by the above discussion for any subset $\mathcal{K}$ in $\mathcal{X}^*$ we must have,

$$\mu_{\mathcal{K}_\circ} = \left(I_{\mathcal{K}}\right)_*.$$

Similarly, for any subset $\mathcal{L}$ of $\mathcal{X}$ we must have,

$$\mu_{\mathcal{L}^\circ} \equiv \left(I_{\mathcal{L}}\right)^*.$$

We have, $I_{(\mathcal{N}^\circ)_\circ} = (\mu_{\mathcal{N}^\circ})_* = (I_{\mathcal{N}})^*_* = I_{\mathcal{N}}$, and since $H_{\mathcal{N}} = H_{\langle\!\langle\!\langle \mathcal{N} \rangle\!\rangle\!\rangle}$ whenever $\mathcal{N}$ contains the origin, we must have $I_{(\mathcal{N}^\circ)_\circ} = I_{\langle\!\langle\!\langle \mathcal{N} \rangle\!\rangle\!\rangle}$. If $\langle\!\langle\!\langle \mathcal{M} \rangle\!\rangle\!\rangle^*$ is the absolutely convex weak*-closure of $\mathcal{M}$. Then we should have $I_{(\mathcal{M}_\circ)^\circ} = I_{\langle\!\langle\!\langle \mathcal{M} \rangle\!\rangle\!\rangle^*}$ whenever $\mathcal{M}$ contains the origin of $\mathcal{X}^*$. Hence we have proved the bipolar theorem;

Theorem 1.2.3. (Bipolar Theorem) *If $\mathcal{N} \subseteq \mathcal{X}, \mathcal{M} \subseteq \mathcal{X}^*$ contain origin, then*

$$(\mathcal{N}^\circ)_\circ = \langle\!\langle\!\langle \mathcal{N} \rangle\!\rangle\!\rangle,$$
$$(\mathcal{M}_\circ)^\circ = \langle\!\langle\!\langle \mathcal{M} \rangle\!\rangle\!\rangle^*$$

For any non-empty absolutely convex neighborhood of origin $\mathcal{N}$ the polar $\mathcal{N}^\circ$ is an absolutely convex neighborhood of origin for the conjugate space $\mathcal{X}^*$. For any x by definition of the Minkowski functional we have $x \in \mu_{\mathcal{N}}(x)\mathcal{N}$, and for any small ϵ, $\lambda_\epsilon = [\mu_{\mathcal{N}}(x) + \epsilon]$ is such that $x \in \lambda_\epsilon \mathcal{N}$. Similarly, by taking $\lambda_{\epsilon^*} = [\mu_{\mathcal{N}^\circ}(\chi) + \epsilon]$ we must have, $\chi \in \lambda_{\epsilon^*}\mathcal{N}^\circ$. Hence by the definition of polar sets we must have,

$$\left\langle \lambda_\epsilon^{-1} x | \lambda_{\epsilon^*}^{-1} \chi \right\rangle \leqq 1.$$

Since $\mu_{\mathcal{N}} = H_{\mathcal{N}^\circ}$, and $\mu_{\mathcal{N}^\circ} = H_{(\mathcal{N}^\circ)_\circ} = H_{\mathcal{N}}$. Hence we must have,

$$\left\langle x | \chi \right\rangle \leqq \left[H_{\mathcal{N}}(\chi) + \epsilon\right]\left[H_{\mathcal{N}^\circ}(x) + \epsilon\right].$$

Here ϵ can be arbitrarily small. This proves a Cauchy-Schwarz type inequality for the conjugate pairing in terms of the neighborhood bases of origin and their polars.

Lemma 1.2.4. (Cauchy-Schwarz)

$$\left\langle x | \chi \right\rangle \leqq H_{\mathcal{N}}(\chi) H_{\mathcal{N}^\circ}(x),$$

whenever $\mathcal{N}$ is an absolutely convex neighborhood of origin.

The space of all continuous functions from $\mathcal{X}$ to $\mathbb{K}$ coincides with the product space $\mathbb{K}^{\mathcal{X}} = \prod_{x \in \mathcal{X}} \mathbb{K}$, equipped with the product topology. The space $\mathcal{X}^*$ of continuous linear functionals on $\mathcal{X}$ is a subset of $\mathbb{K}^{\mathcal{X}}$. Given a subset $\mathcal{U}$ of $\mathcal{X}^*$, we denote by $\langle \cdot \,|\, \mathcal{U} \rangle$ the image of $\mathcal{U}$ under the map,

$$\varphi \mapsto \left\langle \cdot \,|\, \varphi \right\rangle \in \mathbb{K}^{\mathcal{X}}.$$

Since linearity, and topology are related to each other on $\mathbb{K}$, and since the product topology on $\mathbb{K}^{\mathcal{X}}$ corresponds to pointwise evaluations, it follows that $\mathcal{X}^*$ is a closed subset of $\mathbb{K}^{\mathcal{X}}$.

The product topology on $\mathbb{K}^{\mathcal{X}}$ restricted to continuous functionals coincides with the weak* topology on $\mathcal{X}^*$.

Let $\mathcal{X}$ be a locally convex space and $\mathcal{N}_{\mathcal{X}}$ be a neighborhood basis of origin of absolutely convex subsets. For each element $\mathcal{N}$ of $\mathcal{N}_{\mathcal{X}}$ the associated Minkowski functional $\mu_{\mathcal{N}}$ possesses useful convexity property, and continuity properties. We now exploit these properties, to study topological properties of subsets of the conjugate space $\mathcal{X}^*$.

Theorem 1.2.5. (Alaoglu-Bourbaki) *If $\mathcal{N}$ is in $\mathcal{N}_{\mathcal{X}}$ then $\mathcal{N}^\circ$ is weak*-compact.*

Proof

Let K_x be the disc around origin in $\mathbb{K}$ of radius $\mu_{\mathcal{N}}(x)$. Then each K_x is a compact subset of $\mathbb{K}$. If $\mathbb{K}^{\mathcal{X}}$ is the collection of all pointwise continouus functions from $\mathcal{X}$ to $\mathbb{K}$, then $\mathbb{K}^{\mathcal{X}}$ comes equipped with the product topology. Then the product space

$$\Big[\prod_{x\in\mathcal{X}} K_x\Big] \Subset \mathbb{K}^{\mathcal{X}}.$$

is compact by Tychonoff's theorem.

The Minkowski functional $\mu_{\mathcal{N}}$ of $\mathcal{N}$ is a continuous function on $\mathcal{X}$ and $x \mapsto \mu_{\mathcal{N}}(x)$, and for every φ in $\mathcal{N}^\circ$ we have $|\langle x|\varphi\rangle| \leqq \mu_{\mathcal{N}}(x)$. Hence we have an embedding of $\mathcal{N}^\circ$ inside $\prod_{x\in\mathcal{X}} K_x$,

$$\begin{aligned} \big\langle \cdot \,|\, \cdot \big\rangle : \mathcal{N}^\circ &\rightarrowtail \prod_{x\in\mathcal{X}} K_x; \\ \varphi &\mapsto \big(\langle x|\varphi\rangle\big)_{x\in\mathcal{X}}. \end{aligned}$$

where $\langle \cdot \mid \mathcal{N}^\circ\rangle \equiv \big[\prod_{x\in\mathcal{X}} K_x\big] \cap \langle \cdot \mid \mathcal{X}^*\rangle_{\mathcal{X}}$. Hence it is a closed subset of a compact set, and must be compact in the product topology. Since $\mathcal{N}^\circ \rightarrowtail \langle \cdot \mid \mathcal{N}^\circ\rangle$ is continuous when $\mathcal{X}^*$ is equipped with the weak* topology on $\mathcal{X}^*$, it follows that $\mathcal{N}^\circ$ is weak*-compact. □

The supporting functions can see sets only upto their absolutely convex closures. Hence, we can isolate the relation between convexity and topology by constructing topologies using supporting functions. Since supporting function are proper, lower semi-continuous with respect to the weak topologies we can make use of conjugate duality and related theorems for studying the properties. The topology on $\mathcal{X}$ generated by the collection of semi-norms,

$$\tau(\mathcal{X}, \mathcal{X}^*) \equiv \Big\{ x \mapsto \underbrace{\sup_{\varphi\in\mathcal{K}} \Big[\big|\langle x|\varphi\rangle\big|\Big]}_{H_{\mathcal{K}}(x)} \Big\},$$

where each $\mathcal{K}$ is weak*-compact subset of $\mathcal{X}^*$, is called the Mackey topology on $\mathcal{X}$. The topology on $\mathcal{X}^*$ generated by the collection of semi-norms,

$$\tau(\mathcal{X}^*, \mathcal{X}) \equiv \Big\{ \varphi \mapsto \underbrace{\sup_{x\in\mathcal{N}} \Big[\big|\langle x|\varphi\rangle\big|\Big]}_{H_{\mathcal{N}}(\varphi)} \Big\},$$

where each $\mathcal{N}$ is weakly compact in $\mathcal{X}$, is called the Mackey topology on $\mathcal{X}^*$.

The Mackey topologies isolate the convexity aspects of a space with respect to the pairing space, and hence we may expect the Mackey topologies to have some universal property for locally convex topologies.

Theorem 1.2.6. (Mackey-Arens)

$$(\mathcal{X},\tau)^* = \mathcal{X}^* \Leftrightarrow \sigma(\mathcal{X},\mathcal{X}^*) \subseteq \tau \subseteq \tau(\mathcal{X},\mathcal{X}^*).$$

Proof

Since $\sigma(\mathcal{X},\mathcal{X}^*)$ is the weakest topology such that the maps $x \mapsto \langle x|\varphi\rangle_{\mathcal{X}}$ are continuous for all φ in $\mathcal{X}^*$, and since $(\mathcal{X},\tau)^* = \mathcal{X}^*$ we must have, $\sigma(\mathcal{X},\mathcal{X}^*) \subseteq \tau$.

We must now prove that for any τ-continuous linear functional φ the map $|\varphi|$ is a semi-norm generating the topology $\tau(\mathcal{X},\mathcal{X}^*)$, that is, there exists some weak*-compact subset $\mathcal{K}_\varphi$ of $\mathcal{X}^*$ such that $|\varphi| = H_{K_\varphi}$. Since $(\mathcal{X},\tau)^* = \mathcal{X}^*$, the functional φ must be in $\mathcal{X}^*$. The polar $\{\varphi\}_\circ$ is a weakly-closed absolutely convex neighborhood of origin in $\mathcal{X}$. By Alaoglu-Bourbaki theorem $(\{\varphi\}_\circ)^\circ$ must be weak*-compact. Let

$$\mathcal{K}_\varphi \equiv \big(\{\varphi\}_\circ\big)^\circ.$$

By the bipolar theorem it is the closed absolutely convex hull generated by φ and the origin in $\mathcal{X}^*$. Hence the map $\chi \mapsto |\langle x|\chi\rangle|$ attains suprema at φ in the set $\mathcal{K}_\varphi$. Hence we have,

$$\big|\varphi(x)\big| = \sup_{\chi\in\mathcal{K}_\varphi}\Big[\big|\big\langle x|\chi\big\rangle\big|\Big] \equiv H_{\mathcal{K}_\varphi}(x).$$

Hence φ must be $\tau(\mathcal{X},\mathcal{X}^*)$-continuous, since it is one of the semi-norms generating $\tau(\mathcal{X},\mathcal{X}^*)$. Hence we have $\tau \subseteq \tau(\mathcal{X},\mathcal{X}^*)$. This completes the proof. □

A locally convex space, whose topology is induced by a translation invariant metric[2] and is complete with respect to the metric is called a Fréchet space. By the translation invariance we can describe the metric in terms of the convex function $\|\cdot\|$.

By the density properties of rational numbers inside $\mathbb{R}$, only a countable collection of semi-norms consisting of Minkowski functionals of the neighborhoods of the form $\mathcal{N}_i$ consisting of all x with $\|x\| \leqq 1/2^i$ is sufficient to describe the topology. The Minkowski functionals of $\mathcal{N}_i$ are given by,

$$\mu_{\mathcal{N}_i}(x) = H_{\mathcal{N}_i^\circ}(x) = 2^i\big\|x\big\|.$$

By the Cauchy-Schwarz inequality we must have, $\big|\big\langle x|\chi\big\rangle_{\mathcal{X}}\big| \leqq H_{\mathcal{N}_i}(\chi)H_{\mathcal{N}_i^\circ}(x)$, for all x in $\mathcal{X}$, and φ in $\mathcal{X}^*$. By substituting the above formula for the supporting function we have $2^{-i}|\langle x|\chi\rangle_{\mathcal{X}}| \leqq H_{\mathcal{N}_i}(\chi)\|x\|$, and since $\sum_{i\in\mathbb{N}} 2^{-i} = 1$, we have,

$$\big|\big\langle x|\chi\big\rangle_{\mathcal{X}}\big| \leqq \|x\|\Big[\sum_{i\in\mathbb{N}} H_{\mathcal{N}_i}\Big](\chi).$$

If a linear functional ω is weak*-continuous on $\mathcal{N}^\circ$, then there exists a finite collection of F of $\mathcal{X}$ such that for all χ in $\mathcal{N}^\circ$ we must have

$$\kappa(\chi) \leqq \lambda_\kappa \sup_{x\in F}\Big[\big|\big\langle x|\chi\big\rangle_{\mathcal{X}}\big|\Big] \leqq \lambda_\kappa \sup_{x\in F}\Big[\|x\|\big[\textstyle\sum_{\mathbb{N}} H_{\mathcal{N}_i}\big](\chi)\Big].$$

[2] A metric g on $\mathcal{X}$ is a symmetric non-degenerate real valued pairing of a space with itself, which satisfies triangle inequality, that is, $g(x|z) \leqq g(x|y) + g(y|z)$, for all x, y and z in $\mathcal{X}$. A metric space is a topological space whose topology is induced by a metric. If $\mathcal{X}$ is a vector space, the metric g is said to be translation invariant if

$$g(x|y) = g(x-y|0),\ \ \forall x, y \in \mathcal{X}.$$

In such a case $\|\cdot\| \equiv g(\cdot|0)$. The function $\|\cdot\|$ is convex and positively homogeneous on $\mathcal{X}$.

Since $\mathcal{X}$ is a Fréchet space, we can exhaust it by a countable collection $\{K_i\}_{i\in\mathbb{N}}$ of compact subsets of $\mathcal{X}$ and since κ is weak*-continuous in each K_i, we obtain a countable collection F_i of finite subsets of $\mathcal{X}$. If y belongs to $\mathcal{X}$ then consider the ball $\mathcal{N}_{\epsilon,y}$ of radius $\|y\|+\epsilon$. Then we should have $\mu_{\mathcal{N}_{\epsilon,y}}(x) = \|x\|$.

[[Incomplete]]

If I cant figure it out, I will only state the theorem

From the above expression, it follows that a linear functional is weak*-continuous if and only if it is weak*-continuous on $\|\cdot\|$-bounded subsets of $\mathcal{X}$.

Theorem 1.2.7. (Krein-Smulian) *Let $\mathcal{X}$ be a Fréchet space and $\|\cdot\|$ be as above. A linear functional on $\mathcal{X}^*$ is $\sigma(\mathcal{X}^*,\mathcal{X})$-continuous iff its restriction to $\mathcal{N}^\circ$ is $\sigma(\mathcal{X}^*,\mathcal{X})$-continuous, for every $\mathcal{N}\in\mathcal{N}_{\mathcal{X}}$.*

1.2.1 | Conjugate Duality for Banach Space

A localy convex space $\mathcal{X}$ whose topology is generated by a norm $\|\cdot\|_{\mathcal{X}}$ is called a normed space. If $\mathcal{X}$ is complete with respect to the norm topology, then it is called a Banach space. The non-degeneracy of the norm $\|\cdot\|_{\mathcal{X}}$ allows it to see all the lines passing through origin.

By linearity, the unit ball $B_{\mathcal{X}}$ in $\mathcal{X}$ consisting of all x in $\mathcal{X}$ with $\|x\|_{\mathcal{X}} \leqq 1$ contains the data required to generate all lines of $\mathcal{X}$. The supporting function of $B_{\mathcal{X}}$ is denoted by,

$$\|\varphi\|_{\mathcal{X}^*} \coloneqq \sup_{\|x\|_{\mathcal{X}}\leqq 1}\Big[\big|\langle x|\varphi\rangle_{\mathcal{X}}\big|\Big] \equiv H_{B_{\mathcal{X}}}(\varphi).$$

As a supremum of weak* continuous functionals $\|\cdot\|_{\mathcal{X}^*}$ is in $\overline{\Gamma}_+(\mathcal{X}^*)$. $\|\cdot\|_{\mathcal{X}^*}$ defines a norm on $\mathcal{X}^*$, and $\mathcal{X}^*$ equipped with the norm $\|\cdot\|_{\mathcal{X}^*}$ is a Banach space.

Any subspace $\mathcal{Y}$ of $\mathcal{X}$ inherits the norm $\|\cdot\|_{\mathcal{X}}$ by restriction, and hence we can describe the conjugate space $\mathcal{Y}^*$ as $\|\cdot\|_{\mathcal{X}}$-bounded functionals φ with the norm, $\|\varphi\|_{\mathcal{Y}^*} \equiv H_{B_{\mathcal{Y}}}(\varphi)$, where $B_{\mathcal{Y}}$ is the collection of all y in $\mathcal{Y}$ with $\|y\|_{\mathcal{X}} \leqq 1$. The map

$$f^{\varphi}_{\mathcal{X}}(x) = \|\varphi\|_{\mathcal{Y}^*}\|x\|_{\mathcal{X}}$$

is a sub-linear map since the norm $\|\cdot\|_{\mathcal{X}}$ on $\mathcal{X}$ is sub-additive. Hence by the Hahn-Banach theorem, there exists an extension $\widehat{\varphi}$ of φ such that $|\langle x|\widehat{\varphi}\rangle_{\mathcal{X}}| \leqq f^{\varphi}_{\mathcal{X}}(x) = \|\varphi\|_{\mathcal{Y}^*}\|x\|_{\mathcal{X}}$. Since $\|\cdot\|_{\mathcal{X}^*}$ is the supremum over all such functionals, it follows that $\|\widehat{\varphi}\|_{\mathcal{X}^*}$ is less than $\|\varphi\|_{\mathcal{Y}^*}$. Since $\mathcal{X}$ already contains $\mathcal{Y}$ we have

$$\|\widehat{\varphi}\|_{\mathcal{X}^*} = \|\varphi\|_{\mathcal{Y}^*}.$$

We now prove an lemma which will simplify many of the proceding theorems;

For any x in $\mathcal{X}$, we can consider the subspace $\mathbb{K}x$ generated by x, and define a linear functional on $\mathbb{K}x$ by, $\lambda x \mapsto \lambda\|x\|^2_{\mathcal{X}}$. This is clearly a bounded linear functional on $\mathbb{K}x$, and by Hanh-Banach theorem, can be thought of as a restriction of the linear functional φ_x with the same norm. The norm of the linear functional φ_x is given by,

$$\|\varphi_x\|_{\mathcal{X}^*} = \sup_{\|\lambda x\|\leqq 1}\Big[\big|\langle \lambda x|\varphi_x\rangle_{\mathcal{X}}\big|\Big] = \|x\|_{\mathcal{X}},$$

with

$$\langle x|\varphi_x\rangle_{\mathcal{X}} = \|x\|^2_{\mathcal{X}}.$$

Since $\|\cdot\|_{\mathcal{X}^*} = H_{B_{\mathcal{X}}}$, by the Caucy-Schwarz inequality we must have $|\langle x|\varphi\rangle_{\mathcal{X}}| \leqq \|x\|_{\mathcal{X}}\|\varphi\|_{\mathcal{X}^*}$. The set $B_{\mathcal{X}^*}$ consists of all φ which are majorised by $\|\cdot\|_{\mathcal{X}}$. Hence it follows that

$$\|x\|_{\mathcal{X}} \equiv \sup_{\varphi \leq \|\cdot\|_{\mathcal{X}}} \Big[\big|\langle x|\varphi\rangle_{\mathcal{X}}\big| \Big].$$

As a supremum of weak*-continuous functionals, $\|\cdot\|_{\mathcal{X}}$ must be weak*-continuous and positively homoegneous. Hence it must belong to $\overline{\Gamma}_+(\mathcal{X})$. Hence we have proved the following lemma;

Theorem 1.2.8. *If $\|\cdot\|_{\mathcal{X}}$ is a norm on $\mathcal{X}$, then $\|\cdot\|_{\mathcal{X}}$ is in $\overline{\Gamma}_+(\mathcal{X})$, with*

$$\|x\|_{\mathcal{X}} = \sup_{\varphi \leq \|\cdot\|_{\mathcal{X}}} \Big[\big|\langle x|\varphi\rangle_{\mathcal{X}}\big| \Big].$$

Heuristically, the norm already contains data about all line segments on the space, and relates each line segment to a line-segment on the real line. Since the topology and vector space structure on $\mathbb{R}$ are related to each other, we may expect the norm to introduce such a relation on the vector space.

1.2.1.1 | The Second Conjugate

Let $\mathcal{X}$ be a locally convex space, which is also a Banach space with respect to a norm $\|\cdot\|_{\mathcal{X}}$. We may expect topological closure properties to be closely related to closure with respect to convexity structure, as is the case for real numbers. The second conjugate space of $\mathcal{X}$ also inherits the sup-norm from the norm $\|\cdot\|_{\mathcal{X}^*}$, defined by $\|\cdot\|_{\mathcal{X}^{**}} := H_{B_{\mathcal{X}^*}}$ and hence

$$\|X\|_{\mathcal{X}^{**}} = \sup_{\varphi \in B_{\mathcal{X}^*}} \Big[\big|\langle \varphi|X\rangle_{\mathcal{X}^*}\big| \Big].$$

The second conjugate $\mathcal{X}^{**}$ inherits the weak topology from the conjugate space $\mathcal{X}^*$. We call this the weak** topology on $\mathcal{X}^{**}$.

The conjugate transform of $\|\cdot\|_{\mathcal{X}}$ is $(\|\cdot\|_{\mathcal{X}})^*(\varphi) = \sup_{\lambda \in \mathbb{R}^+} \lambda \sup_{x \in B_{\mathcal{X}}} \big[|\langle x|\varphi\rangle_{\mathcal{X}}| - \|x\|_{\mathcal{X}} \big]$. Whenever x is φ is not in $B_{\mathcal{X}^*}$, then we have $\|\varphi\|_{\mathcal{X}^*} > 1$, then there exists some x in $B_{\mathcal{X}}$ such that $\langle x|\varphi\rangle > 1$. Hence $(\|\cdot\|_{\mathcal{X}})^*(\varphi)$ must be infinite. When φ is in $B_{\mathcal{X}^*}$, by the Cauchy-Schwarz inequality we must have, $\sup_{x \in B_{\mathcal{X}}} \big[|\langle x|\varphi\rangle_{\mathcal{X}}| - \|x\|_{\mathcal{X}} \big] \leqq \sup_{x \in B_{\mathcal{X}}} \big[\|x\|_{\mathcal{X}} [\|\varphi\|_{\mathcal{X}^*} - 1] \big] \leqq 0$. Hence must be zero. Hence it follows that

$$(\|\cdot\|_{\mathcal{X}})^* = I_{B_{\mathcal{X}^*}}.$$

For every f in $\overline{\Gamma}(\mathcal{X})$ we must have $(\mathcal{F} \circ \mathcal{F} f)|_{\mathcal{X}}(x) = f(x)$. Hence we must have $\|x\|_{\mathcal{X}} = \|\iota_{\mathcal{X}}(x)\|_{\mathcal{X}^{**}}$. Hence the canonical embedding of $\mathcal{X}$ into $\mathcal{X}^{**}$ is an isometric embedding.

In terms of the polar sets, we have $(B_{\mathcal{X}})^\circ = B_{\mathcal{X}^*}$. Hence we have

$$({B_{\mathcal{X}}}^\circ)^\circ = (B_{\mathcal{X}^*})^\circ = B_{\mathcal{X}^{**}}.$$

From the relation between indicators and Minkowski functionals we must have,

$$I_{B_{\mathcal{X}^{**}}} = I_{((B_{\mathcal{X}})^\circ)^\circ} = (\mu_{(B_{\mathcal{X}})^\circ})^* = (I_{B_{\mathcal{X}}})^{**}.$$

Since $f^{**}(\iota_{\mathcal{X}}(x)) = f(x)$ for every function f in $\overline{\Gamma}(\mathcal{X})$, it follows that $B_{\mathcal{X}^{**}}$ is the weak** closure of the set $\iota_{\mathcal{X}}(B_{\mathcal{X}})$. Hence we have proved the following theorem;

Theorem 1.2.9. (Goldstine) *$\iota_{\mathcal{X}}(B_{\mathcal{X}})$ is weak* dense in $B_{\mathcal{X}^{**}}$.*

1.2.1.2 | Duality Mapping & Adjoints

We now consider the collection of all such functionals which possess $\|x\|_{\mathcal{X}}^2 = |\langle x|\varphi\rangle_{\mathcal{X}}| = \|\varphi\|_{\mathcal{X}^*}^2$. The duality mapping for the norm $f = \|\cdot\|_{\mathcal{X}}$ is the mapping,

$$\begin{aligned} J : \mathcal{X} &\to \mathbf{2}^{\mathcal{X}^*}. \\ x &\mapsto J(x) \end{aligned}$$

where $J(x) \subseteq \mathcal{X}^*$ consists of all φ such that

$$\|\varphi\|_{\mathcal{X}^*}^2 = \langle x|\varphi\rangle = \|x\|_{\mathcal{X}}^2.$$

By the above lemma, $J(x)$ is non-empty for every x in $\mathcal{X}$, and by the bilinearity of the pairing, we must have $J(\lambda x) = \overline{\lambda} J(x)$. A point x in the unit sphere $B_{\mathcal{X}}$ is said to be smooth point if $J(x)$ is a singleton.

A selector of the duality mapping is a choice of an element of $J(x)$, and the composition gives us a map, $j : \mathcal{X} \to \mathcal{X}^*$. A selector of the duality mapping is called a supporting map if

$$j(\lambda x) = \lambda j(x), \quad \forall \lambda \in \mathbb{R}^+.$$

A supporting map i is called strong supporting map if

$$i(\lambda x) = \overline{\lambda} i(x), \quad \forall \lambda \in \mathbb{C}.$$

Let A be a linear map from the locally convex space $\mathcal{X}$ to a locally convex space $\mathcal{Y}$. Then the adjoint mapping of A is the continuous linear mapping A^* from the conjugate space $\mathcal{Y}^*$ to the conjugate space $\mathcal{X}^*$ such that,

$$\langle x|A^*\chi\rangle_{\mathcal{X}} = \langle Ax|\chi\rangle_{\mathcal{Y}}$$

for all x in $\mathcal{X}$ and χ in $\mathcal{Y}^*$. The second adjoint A^{**} is defined to be the adjoint of A^*, from $\mathcal{X}^*$ to $\mathcal{Y}^*$ such that

$$\langle \varphi|A^{**}X\rangle_{\mathcal{Y}^*} = \langle A^*\varphi|X\rangle_{\mathcal{X}^*}.$$