(*November 28, 2016*)

Vector-valued integrals

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/vector-valued_integrals.pdf]

Quasi-complete, locally convex topological vector spaces V have the useful property that continuous compactly-supported V-valued functions have *integrals* with respect to finite, regular Borel measures. Rather than being *constructed* as limits, these vector-valued integrals are *characterized*. Uniqueness follows from the Hahn-Banach theorem, and existence follows from a construction.

An immediate application is justification of differentiation with respect to a parameter inside an integral, under mild, easily understood hypotheses, a special case of the general assertion that Gelfand-Pettis integrals commute with continuous operators, as in the first section. A subtler application is to passage of compactly-supported distributions inside the integrals expressing Fourier inversion, as in [14.3]. Uniqueness of group-invariant measures, distributions, and other functionals is another corollary. Other applications are to holomorphic vector-valued functions, to holomorphically parametrized families of *generalized functions* (distributions), as in chapter 14. Many distributions which are not classical functions appear naturally as residues or analytic continuations of meromorphic families of classical functions.

---

1. *Characterization*

For a topological vectorspace V over $\mathbb{C}$ and for f a continuous V-valued function on a topological space X with a regular Borel measure, a *Gelfand-Pettis integral* of f is a vector $I_f \in V$ so that

$$\lambda(I_f) = \int_X \lambda \circ f \qquad \text{(for all } \lambda \in V^*)$$

If it exists and is unique, this vector I_f is reasonably denoted

$$I_f \;=\; \int_X f$$

In contrast to *construction* of integrals as limits, this *characterization* surely should apply to any reasonable notion of integral, without asking how the property comes to be. Since the property of allowing continuous linear functionals to pass inside the integral is an irreducible minimum, the Gelfand-Pettis integral is sometimes called a *weak integral*.

We only consider *locally convex* vectorspaces, so *uniqueness* of the integral is immediate, since V^* *separates points* on V, by Hahn-Banach. Similarly, for such V, *linearity* of $f \to I_f$ follows by Hahn-Banach. The issue is *existence*. [1] We only consider V-valued functions that are *continuous* on *compact* measure spaces with

[1] We want the integral to be in V itself, rather than in a larger space containing V, such as a double dual V^{**}, for

regular Borel measures. Under these assumptions, all the $\mathbb{C}$-valued integrals

$$f \longrightarrow f \circ \lambda \longrightarrow \int_X \lambda \circ f \qquad\qquad (\text{for } \lambda \in V^*)$$

exist for elementary reasons, being integrals of compactly-supported $\mathbb{C}$-valued continuous functions on compact sets with respect to a regular Borel measure.

For *existence* of Gelfand-Pettis integrals of compactly-supported, continuous V-valued functions, the literal requirement on V turns out to be that *the closure of the convex hull of a compact set is compact.* We show below that *local convexity* and *quasi-completeness* suffice. For the following, a *probability measure* is a positive, regular, Borel measure with total measure 1.

[1.1] Theorem: Let X be a compact Hausdorff topological space with a probability measure. Let V be a quasi-complete, locally convex vectorspace. Then *continuous* V-valued functions f on X have Gelfand-Pettis integrals. The *basic estimate* holds:

$$\int_X f \;\in\; \Big(\text{closure of convex hull of } f(X)\Big)$$

substituting for the estimate of a $\mathbb{C}$-valued integral by the integral of its absolute value. *(Proof below.)*

[1.2] Corollary: In the situation of the theorem, but when the total measure of X is *finite* but not necessarily 1, the *basic estimate* becomes

$$\int_X f \;\in\; \Big(\text{closure of convex hull of } f(X)\Big) \cdot \int_X 1$$

(Replace the measure by a constant multiple.) ///

[1.3] Corollary: For a continuous linear map of locally convex, quasi-complete topological vectorspaces $T : V \to W$, and f a continuous, compactly-supported V-valued function on a finite, regular, positive Borel measure space X. Then

$$T\Big(\int_X f\Big) \;=\; \int_X T \circ f$$

Proof: To verify that the left-hand side of the asserted equality is a Gelfand-Pettis integral of $T \circ f$, show that

$$\mu\Big(\text{left-hand side}\Big) \;=\; \int_X \mu \circ (T \circ f) \qquad\qquad (\text{for all } \mu \in W^*)$$

Starting with the left-hand side,

$$\begin{aligned} \mu\Big(T\Big(\textstyle\int_X f\Big)\Big) &= (\mu \circ T)\Big(\textstyle\int_X f\Big) && (\text{associativity}) \\ &= \textstyle\int_X (\mu \circ T) \circ f && (\mu \circ T \in V^* \text{ and } \textstyle\int_X f \text{ is a weak integral}) \\ &= \textstyle\int_X \mu \circ (T \circ f) && (\text{associativity}) \end{aligned}$$

proving that $T\big(\int_X f\big)$ is a weak integral of $T \circ f$. ///

---

example, to make *existence* trivial, but then leaving technical issues. Some discussions of vector-valued integration do allow integrals to exist in larger spaces, but this only delays certain issues, rather than resolving them directly.

2. *Differentiation of parametrized integrals*

Differentiation under the integral is an immediate corollary, in many useful situations.

[2.1] **Claim:** A $\mathbb{C}$-valued C^k function F on $[a,b] \times [c,d]$ gives a *continuous* $C^k[c,d]$-valued function $f(x) = F(x,-)$ of $x \in [a,b]$.

Proof: For each $0 \le i \le k$, the function $(x,y) \to \frac{\partial^i}{\partial y^i} F(x,y)$ is continuous as a function of two variables. For each $\varepsilon > 0$ and each $x_o \in [a,b]$, we want $\delta > 0$ such that

$$|x - x_o| < \delta \implies \sup_y \Big| \frac{\partial^i}{\partial y^i} F(x,y) - \frac{\partial^i}{\partial y^i} F(x_o,y) \Big| < \varepsilon$$

The continuous function $(x,y) \to \frac{\partial^k}{\partial y^i} F(x,y)$ is *uniformly* continuous on the compact $[a,b] \times [c,d]$, so there is $\delta > 0$ such that

$$\Big| \frac{\partial^k}{\partial y^i} F(x_1,y_1) - \frac{\partial^i}{\partial y^i} F(x_2,y_2) \Big| < \varepsilon \qquad \text{(for all } (x_1,y_1),(x_2,y_2) \text{ with } |x_1 - x_2| < \delta \text{ and } |y_1 - y_2| < \delta)$$

In particular, this holds for all $y_1 = y_2$, and $x_1 = x$, and $x_2 = x_o$. ///

[2.2] **Corollary:** For a $\mathbb{C}$-valued C^k function F on $[a,b] \times [c,d]$,

$$\frac{\partial}{\partial y} \int_a^b F(x,y)\, dx \;=\; \int_a^b F \frac{\partial}{\partial y}(x,y)\, dx$$

Proof: The function-valued function $x \to (y \to F(x,y))$ is a continuous, $C^k[c,d]$-valued function, and $\frac{\partial}{\partial y}$ is a continuous linear map $C^k[c,d] \to C^{k-1}[c,d]$, so the Gelfand-Pettis property allows interchange of the operator and the integral. ///

3. *Fourier transforms*

Certainly an integral expressing Fourier inversion

$$f(x) \;=\; \frac{1}{2\pi} \int_{\mathbb{R}} e^{i\xi x}\; \widehat{f}(\xi)\; d\xi$$

for Schwartz function f cannot converge as a Schwartz-function-valued integral, because $x \to e^{i\xi x}$ is in $C^\infty(\mathbb{R})$, but not Schwartz. Multiplying by $\widehat{f}$ does not affect decay in x, so does not alter the situation. Examination of the situation is complicated by the fact that the integrand is not compactly supported, but we can follow Schwartz' device of suitably *compactifying* $\mathbb{R}^n$ to a sphere S^n, and then invoke the Gelfand-Pettis property for compactly-supported functions. Then we will see that the integral *does* converge as a $C^\infty(\mathbb{R})$-valued Gelfand-Pettis integral.

First,

[3.1] **Claim:** For any $\Phi \in C^\infty(\mathbb{R}^2)$, the $C^\infty(\mathbb{R})$-valued function $\xi \to \Phi(-,\xi)$ that is, $\xi \to (x \to \Phi(x,\xi))$ is a *continuous*, $C^\infty(\mathbb{R})$-valued function on $\mathbb{R}$. (Similarly, it is a *smooth* $C^\infty(\mathbb{R})$-valued function, but we do not need this.)

Proof: The function $(x,\xi) \to \Phi(x,\xi)$ is C^∞ as a function of two variables. In particular, $(x,\xi) \to \frac{\partial^k}{\partial x^k}\Phi(x,\xi)$ is continuous as a function of two variables. For each k, compact $C \subset \mathbb{R}$, $\varepsilon > 0$ and each $\xi_o \in \mathbb{R}$, we want $\delta > 0$ such that

$$|\xi - \xi_o| < \delta \implies \sup_{x\in C}\Big|\frac{\partial^k}{\partial x^k}\Phi(x,\xi) - \frac{\partial^k}{\partial x^k}\Phi(x,\xi_o)\Big| < \varepsilon$$

Let I be the interval $[\xi_1, \xi_o + 1]$. The continuous function $(x,\xi) \to \frac{\partial^k}{\partial x^k}\Phi(x,\xi)$ is *uniformly* continuous on the compact $C \times I$ that is, there is $\delta > 0$ such that

$$\Big|\frac{\partial^k}{\partial x^k}\Phi(x_1,\xi_1) - \frac{\partial^k}{\partial x^k}\Phi(x_2,\xi_2)\Big| < \varepsilon \qquad \text{(for all } (x_1,\xi_1),(x_2,\xi_2) \in I \text{ with } |x_1 - x_2| < \delta \text{ and } |\xi_1 - \xi_2| < \delta)$$

In particular, this holds for all $x_1 = x_2$, and $\xi = \xi_o \in I$, and $\xi_2 = \xi_o$, giving the desired continuity. This previous applies to $\Phi(x,\xi) = e^{i\xi x}$. Since $\xi \to F(\xi)$ is a continuous $\mathbb{C}$-valued function, the product $x \to F(\xi)\cdot e^{i\xi x}$ is a continuous $C^\infty(\mathbb{R})$-valued function of $\xi \in \mathbb{R}$.

Compactify $\mathbb{R}$ to the circle $\mathbb{T} \subset \mathbb{R}^2$ via by stereographic projection

$$\sigma : x \longrightarrow \Big(\frac{x}{\sqrt{1+x^2}}, \frac{1}{\sqrt{1+x^2}}\Big)$$

and adding the point $\infty = (0,1)$.

[3.2] Claim: $\xi \to F(\xi)\cdot\psi_\xi$ extends (by $0 \in C^\infty(\mathbb{R})$) to a continuous, $C^\infty(\mathbb{R})$-valued function on the compactification $\mathbb{T}$ of $\mathbb{R}$.

Proof: We must check continuity in ξ near ∞. That is, for each k, compact $C \subset \mathbb{R}$, and $\varepsilon > 0$, we want (large) B such that

$$|\xi| > B \implies \sup_{x\in C}\Big|F(\xi)\cdot\frac{\partial^k}{\partial x^k}e^{i\xi x} \;-\; 0\Big| < \varepsilon$$

The exponential function is easy to estimate: for example, with M a bound so that $|(1+\xi^2)^k \cdot F(\xi)| \le M$,

$$\sup_{x\in C}\Big|F(\xi)\cdot\frac{\partial^k}{\partial x^k}e^{i\xi x}\Big| \;=\; \Big|F(\xi)\cdot(i\xi)^k\Big|\cdot 1 \;\le\; \frac{M\cdot|\xi|^k}{(1+\xi^2)^k}$$

Take B large enough so that $M\cdot B^k/(1+B^2)^k < \varepsilon$. For any continuous linear functional, $\xi \to \lambda\circ(\psi_\xi\cdot F(\xi))$ is a continuous scalar-valued function on the compact set $\mathbb{T}$, so is *bounded.* The same is true of any $\xi \to \lambda\circ(\psi_\xi\cdot(1+\xi^2)^N F(\xi))$, so $\xi \to \lambda\circ(\psi_\xi\cdot F(\xi))$ is *rapid decreasing.* Adjust the measure on $\mathbb{R}$ to give total measure 1:

$$\int_{\mathbb{R}} \psi_\xi\cdot F(\xi)\;d\xi \;=\; \int_{\mathbb{R}} \psi_\xi\cdot\pi(1+\xi)^2 F(\xi)\;\frac{d\xi}{\pi(1+\xi^2)}$$

The function $\pi(1+\xi)^2F(\xi)$ is still continuous and of rapid decay. Being continuous and compactly supported on a measure space with total measure 1, with values in a quasi-complete, locally convex topological vector space, $\xi \to \pi(1+\xi^2)F(\xi)\cdot\psi_\xi$ has a *Gelfand-Pettis integral* J with respect to the measure $d\xi/\pi(1+\xi^2)$, lying inside the closed convex hull of the image. That is,

$$\lambda(J) \;=\; \int_{\mathbb{R}} \lambda(\psi_\xi)\cdot\pi(1+\xi)^2F(\xi)\;\frac{d\xi}{\pi(1+\xi^2)}$$

for every continuous linear functional λ. In the the latter scalar-valued integral the adjustment factors cancel:

$$\int_{\mathbb{R}} \lambda\Big(\psi_\xi\cdot\pi(1+\xi)^2F(\xi)\Big)\;\frac{d\xi}{\pi(1+\xi^2)} \;=\; \int_{\mathbb{R}} \lambda(\psi_\xi)\cdot\pi(1+\xi)^2F(\xi)\;\frac{d\xi}{\pi(1+\xi^2)} \;=\; \int_{\mathbb{R}} \lambda(\psi_\xi)\cdot F(\xi)\;d\xi$$

That is, $\lambda(J) = \int_{\mathbb{R}} \lambda(\psi_\xi) \cdot F(\xi)\, d\xi$, and the Gelfand-Pettis integral J of the mutually adjusted function and measure is the Gelfand-Pettis integral of the original. ///

[3.3] Corollary: For rapidly decreasing $F \in C^o(\mathbb{R})$, for any continuous linear $T : C^\infty \to V$ for another topological vector space V,

$$T\Big(\int_{\mathbb{R}} \psi_\xi \cdot F(\xi)\, d\xi\Big) \;=\; \int_{\mathbb{R}} T\big(\psi_\xi \cdot F(\xi)\big)\, d\xi \;=\; \int_{\mathbb{R}} T(\psi_\xi) \cdot F(\xi)\, d\xi$$

as V-valued Gelfand-Pettis integral. ///

[3.4] Corollary: For rapidly decreasing $F \in C^o(\mathbb{R})$, for any continuous, for any compactly-supported distribution u,

$$u\Big(\int_{\mathbb{R}} \psi_\xi \cdot F(\xi)\, d\xi\Big) \;=\; \int_{\mathbb{R}} u\big(\psi_\xi \cdot F(\xi)\big)\, d\xi \;=\; \int_{\mathbb{R}} u(\psi_\xi) \cdot F(\xi)\, d\xi$$

with absolutely convergent integral. ///

[3.5] Corollary: For rapidly decreasing $F \in C^o(\mathbb{R})$, the Fourier transform is a C^∞ function on $\mathbb{R}$, and its derivative is computed by the expected expression

$$\frac{\partial}{\partial x}\Big(\int_{\mathbb{R}} \psi_\xi \cdot F(\xi)\, d\xi\Big) \;=\; \int_{\mathbb{R}} \frac{\partial \psi_\xi}{\partial x} \cdot F(\xi)\, d\xi \;=\; i\int_{\mathbb{R}} \psi_\xi \cdot \xi F(\xi)\, d\xi$$

since $\partial/\partial x$ is a continuous map of $C^\infty(\mathbb{R})$ to itself. ///

---

4. *Uniqueness of invariant distributions*

We prove uniqueness of invariant *functionals* on suitable function spaces V on topological spaces X on which a topological group acts transitively. This includes uniqueness of invariant (Haar) *measures*, and uniqueness of invariant *distributions*, as special cases.

A translation-invariant function f on the real line, that is, a function with $f(x + y) = f(x)$ for all $x, y \in \mathbb{R}$, is *constant*, by a point-wise argument:

$$f(x) \;=\; (T_x f)(0) \;=\; f(0)$$

where $T_x f(y) = f(x + y)$ is translation. The same conclusion holds for translation-invariant *distributions*, but we cannot argue in terms of point-wise values.

Let G be a *topological group*,[2] with right translation-invariant measure dg, meaning that

$$\int_G f(g \cdot h)\, dg \;=\; \int_G f(g)\, dg \qquad\qquad \text{(for all } h \in G)$$

We assume only *existence* of a right translation-invariant measure. The theorem proves uniqueness:

For present purposes, a continuous *approximate identity* on a topological group G is a sequence $\{\varphi_i\}$ of non-negative, continuous, real-valued functions such that $\int_G \varphi_i = 1$ for all i, and such that the supports shrink

[2] A *topological group* is usually understood to be locally compact and Hausdorff, and multiplication and inversion are continuous. To avoid measure-theoretic pathologies, a *countable basis* is often assumed. Perhaps oddly, the local compactness excludes most topological vector spaces.

to $\{1\}$, in the sense that for every neighborhood N of 1 in G, there is an index i_o so that the support of φ_i is inside N for all $i \geq i_o$. From Urysohn's Lemma, there always exists a *continuous* approximate identity. Not all classes of functions contain an approximate identity in this strict sense: (real-) analytic functions on a non-compact group, such as $\mathbb{R}$ cannot be compactly supported, so a compromise notion would be needed. The following theorem refers to the strict sense that supports shrink to $\{1\}$:

[4.1] Theorem: Let $V \subset C_c^o(G)$ be a quasi-complete, locally convex topological vector space of complex-valued functions on G stable under left and right translations, so that $G \times V \to V$ is continuous, and containing an *approximate identity* $\{\varphi_i\}$. Then there is a unique *right* G-invariant element of the dual space V^* (up to constant multiples), and it is

$$f \to \int_G f(g)\, dg \qquad \text{(with right translation-invariant measure)}$$

Proof: Let $T_g f(y) = f(yg)$ be right translation. With right translation-invariant measure dg on G, since $\int_G \varphi_i(g)\, dg = 1$ and φ_i is non-negative, $\varphi_i(g)\, dg$ is a probability measure (total mass 1) on the (compact) support of φ_i. Thus, for any $f \in V$, we have a V-valued Gelfand-Pettis integral

$$T_{\varphi_i} f \;=\; \int_G \varphi_i(g)\, T_g f \; dg \;\in\; \text{closure of convex hull of } \{\varphi_i(g) f : g \in G\} \;\subset\; V$$

By continuity, given a neighborhood N of 0 in V, we have $T_{\varphi_i} f \in f + N$ for all sufficiently large i. That is, $T_{\varphi_i} f \to f$. For a right-invariant (continuous) functional $u \in V^*$,

$$u(f) \;=\; \lim_i u\left(g \to \int_G \varphi_i(h)\, f(gh)\, dh\right)$$

This is

$$u\left(g \to \int_G f(hg)\, \varphi_i(h^{-1})\, dh\right) \;=\; u\left(g \to \int_G f(h)\, \varphi_i(gh^{-1})\, dh\right)$$

by replacing h by hg^{-1}. By properties of Gelfand-Pettis integrals, and since f is guaranteed to be a compactly-supported continuous function, we can move the functional u inside the integral: the above becomes

$$\int_G f(h)\, u\left(g \to \varphi_i(gh^{-1})\right)\, dh$$

Using the *right* G-invariance of u the evaluation of u with right translation by h^{-1} gives

$$\int_G f(h)\, u(g \to \varphi_i(g))\, dh \;=\; u(\varphi_i) \cdot \int_G f(h)\, dh$$

By assumption the latter expressions approach $u(f)$ as $i \to \infty$. For f so that the latter integral is non-zero, we see that the limit of the $u(\varphi_i)$ exists, and then we conclude that $u(f)$ is a constant multiple of the indicated integral with right Haar measure. ///

---

5. *Smoothing of distributions*

Every locally integrable [3] function f on $\mathbb{R}^n$, for example, gives a *distribution* u_f by *integrating against* it:

$$u_f(\varphi) \;=\; \int_{\mathbb{R}^n} \varphi \cdot f \qquad (\text{for } \varphi \in \mathcal{D}(\mathbb{R}^n))$$

[3] Again, *locally integrable* means that $|f|$ is in $L^1(K)$ for every compact K. This makes best sense for positive regular Borel measures, so that the measures of compact sets are finite.

Conversely, we prove here that the distributions u_f from $f \in \mathcal{D} = C_c^\infty(\mathbb{R}^n)$ are *dense* in the whole space $\mathcal{D}^*$ of distributions, with the weak dual topology. Further, a *sequence* of such smooth functions approaching a given distribution can be expressed in terms of *smoothing* or *mollifying* u.

Let $g \to T_g$ be the *regular representation* of $\mathbb{R}^n$ on test functions $f \in \mathcal{D} = \mathcal{D}(\mathbb{R}^n)$ by $(T_g f)(x) = f(x+g)$, for $x, g \in \mathbb{R}^n$. As usual, $x \times f \to T_x f$ gives a continuous map $\mathbb{R}^n \times \mathcal{D} \longrightarrow \mathcal{D}$. The corresponding *adjoint* action of $\mathbb{R}^n$ on distributions u is

$$(T_g^* u)(f) \;=\; u(T_g^{-1} f)$$

For the usual reasons, this gives a continuous map $x \times u \;\longrightarrow\; x \cdot u \;=\; T_x^* u$ with the *weak dual* topology: for $f \in \mathcal{D}$, let ν_f be the semi-norm $\nu_f(u) = |u(f)|$ on $\mathcal{D}^*$, and then

$$\nu(T_g^* u - T_h^* v) \;=\; |u(T_g^{-1} f) - v(T_h^{-1} f)| \;\le\; |u(T_g^{-1} f) - u(T_h^{-1} f)| + |u(T_h^{-1} f) - v(T_h^{-1} f)|$$

$$= \; \big|u\big(T_g^{-1} f - T_h^{-1} f\big)\big| + \nu_{T_h^{-1} f}(u - v)$$

For g close to h, since the translation action of $\mathbb{R}^n$ on $\mathcal{D}$ is continuous and u is a continuous functional, $|u(T_g^{-1} f) - u(T_h^{-1} f)|$ is small. And for u close to v in the weak dual topology, the second term is small. This proves the continuity.

As earlier and throughout, the action of a function $\varphi \in C_c^o(\mathbb{R}^n)$ on distributions u is by *integrating* the group action

$$T_\varphi^* u \;=\; \int_{\mathbb{R}^n} \varphi(x)\, T_x^* u \, dx \;\in\; \mathcal{D}^*$$

Suppressing the T^*, this is

$$\varphi \cdot u \;=\; \int_{\mathbb{R}^n} \varphi(x)\, x \cdot u \, dx \;\in\; \mathcal{D}^*$$

A *smooth* approximate identity on $\mathbb{R}^n$ is a sequence $\{\psi_i\} \subset \mathcal{D}$ which are non-negative, real-valued, have $\int_{\mathbb{R}^n} \psi_i = 1$, and supports shrink to $\{0\} \subset \mathbb{R}^n$.

[5.1] Theorem: *For a smooth approximate identity $\{\psi_i\}$ and distribution u, the distributions $T_{\psi_i}^* u$ go to u in the weak dual topology on $\mathcal{D}^*$, and are (integration against) the functions $x \to u(T_x^{-1}\psi_i)$, which are smooth functions.*

Proof: That $T_{\psi_i}^* u \to u$ as distributions is an instance of an important general property of such Gelfand-Pettis integrals, which we review, as follows. Let U be any *convex* open neighborhood of 0 in $\mathcal{D}^*$. Let N be a sufficiently small neighborhood of 0 in $\mathbb{R}^n$ such that under $\mathbb{R}^n \times \mathcal{D}^* \;\longrightarrow\; \mathcal{D}^*$ we have

$$N \times u \;\longrightarrow\; N \cdot u \;\subset\; u + \frac{1}{2} U$$

For i sufficiently large so that the support of ψ_i is inside N, the measure $\psi_i(x)\, dx$ is a probability measure supported in N, so by Gelfand-Pettis

$$T_{\psi_i}^* u \;\in\; \text{closure of convex hull of image of } \psi_i(x)\, x \cdot u$$

Since $u + U' + \frac{1}{2}U$ contains the closure of the convex set $u + \frac{1}{2}U$ for any open U' containing 0 (in $\mathcal{D}^*$), this shows that $T_{\psi_i}^* u \in u + U$. This holds for all open neighborhoods U of 0, so $T_{\psi_i}^* u \to u$.

To prove that every $T_f u$ for $f \in \mathcal{D}$ is (integration against) a *continuous* or *smooth* function, we first guess what that continuous function is, by determining its point-wise values. Indeed, if $u = u_\varphi$ were known to be integration against a continuous function φ, then with an approximate identity $\{\psi_i\}$

$$\lim_i u_\varphi(\psi_i) \;=\; \lim_i \int_{\mathbb{R}^n} \varphi(x)\, \psi_i(x)\, dx \;=\; \varphi(0)$$

Thus, we anticipate determining values of the alleged continuous function $f \cdot u$ by computing

$$\text{alleged value } (f\cdot u)(0) \;=\; \lim_i (f\cdot u)(\psi_i)$$

For a continuous function F on $\mathbb{R}^n$, let $F^\vee(x) = F(-x)$. For for f and ψ in $\mathcal{D}$, since that Gelfand-Pettis integrals commute with continuous linear maps,

$$(T_f^* u)(\psi) \;=\; \Big(\int_{\mathbb{R}^n} f(x)\, T_x^* u\, dx\Big)(\psi) \;=\; \int_{\mathbb{R}^n} f(x)\,(T_x^* u)(\psi)\, dx$$

$$= \int_{\mathbb{R}^n} f(x)\, u(T_x^{-1}\cdot\psi)\, dx \;=\; u\Big(\int_{\mathbb{R}^n} f(x)\,(T_x^{-1}\cdot\psi)\, dx\Big) \;=\; u\Big(\int_{\mathbb{R}^n} f(-x)\,(T_x\psi)\, dx\Big) \;=\; u(T_{f^\vee}\psi)$$

The function $T_{f^\vee}\psi$ admits a rewriting that reverses the roles of f and ψ, namely

$$(T_{f^\vee}\psi)(y) \;=\; \int_{\mathbb{R}^n} f(-x)\,\psi(y+x)\, dx \;=\; \int_{\mathbb{R}^n} f(y-x)\,\psi(x)\, dx$$

$$= \int_{\mathbb{R}^n} f(y+x)\,\psi(-x)\, dx \;=\; \int_{\mathbb{R}^n} f(y+x)\,\psi^\vee(x)\, dx \;=\; (T_{\psi^\vee} f)(y)$$

Thus,

$$(T_f^*\cdot u)(\psi) \;=\; u(T_{f^\vee}\psi) \;=\; u(T_{\psi^\vee} f) \;=\; (T_\psi^* u)(f)$$

We already know that $T_{\psi_i}^* u \to u$ for an approximate identity ψ_i, so the limit exists, and has an understandable value:

$$(T_f^* u)(\psi_i) \;=\; (T_{\psi_i}^* u)(f) \;\longrightarrow\; u(f) \;=\; \text{supposed value of } f\cdot u \text{ at } 0$$

Thus, we would guess that $T_f^* u$ should be a function with value $u(f)$ at 0. More generally, for the distribution u_φ given by integration against φ, we have

$$(T_z^* u_\varphi)(\psi_i) \;=\; u_\varphi(T_z^{-1}\psi_i) \;=\; \int_{\mathbb{R}^n} \varphi(x)\,\psi_i(x-z)\, dx \;=\; \int_{\mathbb{R}^n} \varphi(x+z)\,\psi_i(x)\, dx \;\to\; \varphi(z)$$

The analogous computation suggests the values of the function $T_f^* u$ at z. First, a more elaborate version of the identity reverses the roles of test functions f and φ, namely

$$(T_{f^\vee} T_z^{-1}\psi)(y) \;=\; \int_{\mathbb{R}^n} f(-x)\,\psi(y+x-z)\, dx \;=\; \int_{\mathbb{R}^n} f(y-x-z)\,\psi(x)\, dx$$

$$= \int_{\mathbb{R}^n} f(y+x-z)\,\psi(-x)\, dx \;=\; \int_{\mathbb{R}^n} (T_z^{-1} f)(y+x)\,\psi^\vee(x)\, dx \;=\; (T_{\psi^\vee} T_z^{-1} f)(y)$$

The same sort of computation gives

$$(T_y^*(T_f^* u))(\psi_i) \;=\; (T_f^* u))(T_y^{-1}\psi_i) \;=\; u(T_{f^\vee} T_y^{-1}\psi_i) \;=\; u(T_{\psi_i^\vee} T_y^{-1} f)$$

$$= \; (T_y^*(T_{\psi_i}^* u))(f) \to (T_y^* u)(f) \;=\; u(T_y^{-1} f) \;=\; \text{supposed value of } f\cdot u \text{ at } y$$

Since $\mathbb{R}^n \times \mathcal{D} \to \mathcal{D}$ is continuous, and u is continuous, the composition

$$y \times f \;\longrightarrow\; T_y^{-1} f \;\longrightarrow\; u(T_y^{-1} f)$$

is indeed *continuous* as a function of $y \in \mathbb{R}^n$.

Now we check that the distribution $f\cdot u$ is truly given by integration against the continuous function

$$\varphi(y) \;=\; u(T_y^{-1} f)$$

that apparently gives the pointwise values of $T_f^* u$. Letting $h \in \mathcal{D}$,

$$\int_{\mathbb{R}^n} \varphi(x)\, h(x)\, dx \;=\; \int_{\mathbb{R}^n} u(T_x^{-1} f)\, h(x)\, dx \;=\; \Big(\int_{\mathbb{R}^n} h(x)\, x\cdot u\, dx\Big)(f) \;=\; (T_h^* u)(f)$$

We already computed directly that

$$(T_h^* u)(f) \;=\; u(T_{h^\vee} f) \;=\; u(T_{f^\vee} h) \;=\; (T_f^* u)(h)$$

which shows that integration against the continuous function $\varphi(y) = u(T_y^{-1} f)$ gives the distribution $T_f^* u$.

Smoothness of $\varphi(y) = u(T_y^{-1} f)$ would follow from the assertion that $y \to T_y^{-1} f$ is a smooth, $\mathcal{D}$-valued function. The latter assertion is existence of the limit

$$\lim_{t\to 0} \frac{T_{y+tX}^{-1} f - T_y^{-1} f}{t} \qquad\qquad (\text{for } X \in \mathbb{R}^n \text{ and } y \in \mathbb{R}^n)$$

in $\mathcal{D}$ for each $X \in \mathbb{R}^n$. It suffices to consider $y = 0$. By design, differentiation is a continuous map of $\mathcal{D}$ to itself, giving the requisite *smoothness*. ///

That is, given the idea that $f \cdot u$ has been smoothed, *determination* of it as a classical function is straightforward. The proof that $T_{\psi_i}^* u \to u$ did not use the specifics of the situation: the same argument applies to representations of *Lie groups*.

6. *Density of smooth vectors*

Let G be a Lie group, so that the notion of C^∞ function on G makes sense. A *representation* of G on a locally convex, quasi-complete topological vectorspace V is a continuous map $G \times V \to V$ that is linear in V, and has the associativity $(gh)\cdot v = g\cdot(h\cdot v)$ for $g, h \in G$. The subspace V^∞ of *smooth vectors* is

$$V^\infty \;=\; \{v \in V : g \to g\cdot v \text{ is a } C^\infty\ V\text{-valued function on } G\}$$

It suffces to consider derivatives associated to the Lie algebra $\mathfrak{g}$ of G:

$$(x\cdot f)(g) \;=\; \frac{\partial}{\partial t}\Big|_{t=0} \Big((ge^{tx})\cdot v\Big) \qquad\qquad (\text{for } x \in \mathfrak{g})$$

where $x \to e^x$ is the exponential map $\mathfrak{g} \to G$.

Note that in the representation of $\mathbb{R}^n$ on distributions $\mathcal{D}^*$ *every* distribution is a smooth vector, since every distribution is infinitely differentiable as a distribution. Thus, smooth *vectors* are not necessarily smooth *functions*. Nevertheless, as in the previous section, distributions are approximable by smooth functions.

7. *Quasi-completeness and convex hulls of compacts*

A subset E of a *complete metric space* X is *totally bounded* if, for every $\varepsilon > 0$ there is a covering of E by *finitely-many* open balls of radius ε. The property of *total boundedness* in a metric space is generally stronger than mere *boundedness*. It is immediate that any subset of a totally bounded set is totally bounded. Recall:

[7.1] Proposition: A subset of a complete metric space has compact closure if and only if it is *totally bounded*.

Proof: Certainly if a set has compact closure then it admits a finite covering by open balls of arbitrarily small (positive) radius. On the other hand, suppose that a set E is totally bounded in a complete metric space X. To show that E has compact closure it suffices to show that any sequence $\{x_i\}$ in E has a Cauchy subsequence.

Choose such a subsequence as follows. Cover E by finitely-many open balls of radius 1. In at least one of these balls there are infinitely-many elements from the sequence. Pick such a ball B_1, and let i_1 be the smallest index so that x_{i_1} lies in this ball.

The set $E \cap B_1$ is still totally bounded, and contains infinitely-many elements from the sequence. Cover it by finitely-many open balls of radius $1/2$, and choose a ball B_2 with infinitely-many elements of the sequence lying in $E \cap B_1 \cap B_2$. Choose the index i_2 to be the smallest one so that both $i_2 > i_1$ and so that x_{i_2} lies inside $E \cap B_1 \cap B_2$.

Inductively, suppose that indices $i_1 < \ldots < i_n$ have been chosen, and balls B_i of radius $1/i$, so that

$$x_i \in E \cap B_1 \cap B_2 \cap \ldots \cap B_i$$

Cover $E \cap B_1 \cap \ldots \cap B_n$ by finitely-many balls of radius $1/(n+1)$ and choose one, call it B_{n+1}, containing infinitely-many elements of the sequence. Let i_{n+1} be the first index so that $i_{n+1} > i_n$ and so that

$$x_{n+1} \;\in\; E \cap B_1 \cap \ldots \cap B_{n+1}$$

For $m < n$ we have $d(x_{i_m}, x_{i_n}) \leq \frac{1}{m}$ so this subsequence is Cauchy. ///

In a not-necessarily metric *topological vectorspace* V, a subset E is *totally bounded* if, for every neighborhood U of 0 there is a finite subset F of V so that $E \subset F + U$, where

$$F + U \;=\; \bigcup_{v \in F} v + U \;=\; \{v + u : v \in F, u \in U\}$$

[7.2] Proposition: A totally bounded subset E of a *locally convex* topological vectorspace V has totally bounded *convex hull*.

Proof: First, recall that the convex hull of a *finite* set $F = \{x_1, \ldots, x_n\}$ in a topological vectorspace is *compact*, since it is the continuous image of the compact set $\{(c_1, \ldots, c_n) \in \mathbb{R}^n : \sum_i c_i = 1,\ 0 \leq c_i \leq 1,\ \text{ for all } i\} \subset \mathbb{R}^n$ under $(c_1, \ldots, c_n) \to \sum_i c_i x_i$.

Given a neighborhood U of 0 in V, let U_1 be a *convex* neighborhood of 0 so that $U_1 + U_1 \subset U$. For some finite subset F we have $E \subset F + U_1$, by total boundedness. The convex hull K of F is *compact*. Then $E \subset K + U_1$, and the latter is *convex*. Therefore, the convex hull H of E lies inside $K + U_1$. Since K is compact, it is totally bounded, so can be covered by a finite union $\Phi + U_1$ of translates of U_1. Thus, since $U_1 + U_1 \subset U$, $H \subset (\Phi + U_1) + U_1 \subset \Phi + U$. Thus, H lies inside this finite union of translates of U. This holds for any open U containing 0, so H is totally bounded. ///

[7.3] Corollary: In a Fréchet space, the closure of the convex hull of a compact set is compact.

Proof: A compact set in a Fréchet space (or in any complete metric space) is totally bounded, as recalled above. By the previous, the convex hull of a totally bounded set in a Fréchet space is totally bounded. Thus, this convex hull has compact closure, since totally bounded sets in complete metric spaces have compact closure. ///

The general case reduces to the case of Fréchet spaces.

[7.4] Proposition: In a *quasi-complete*, locally convex topological vectorspace X, the closure of the convex hull of a compact set is *compact*.

Proof: Since X is locally convex, its topology is given by a collection of seminorms v. For each seminorm v, let X_v be the completion of the quotient $X/\{x \in X : v(x) = 0\}$ with respect to the *metric* that v induces on the latter quotient. Thus, X_v is a Banach space. Consider $Z = \prod_v X_v$ with product topology, with the natural injection $j : X \to Z$, and with projection p_v to the v^{th} factor. By construction, and by definition of the topology given by the seminorms, j is a (linear) homeomorphism to its image. That is, X is homeomorphic to the subset jX of Z, given the subspace topology.

Let $K \subset X$ be compact, with convex hull H, and C the closure of H. The continuous image $p_v jK$ of compact K is compact. Since X_v is Fréchet, the convex hull H_v of $p_v jK$ has compact closure C_v. The convex hull jH of jK is contained in the product $\prod_v H_v$ of the convex hulls H_v of the projections $p_v jK$. By Tychonoff's theorem, the product $\prod_v C_v$ is *compact*.

Since jC is contained in the compact set $\prod_v C_v$, to prove that the closure jC of jH in jX is compact, it suffices to prove that jC is closed in Z. Since jC is a subset of the compact set $\prod_v C_v$, it is totally bounded and so is certainly *bounded* (in Z, hence in $X \approx jX$). By the quasi-completeness, a Cauchy net in jC is necessarily *bounded* and converges to a point in jC. Since any point in the closure of jC in Z has a Cauchy net in jC converging to it, jC is closed in Z. ///

---

8. *Existence proof*

To simplify, divide by a constant to make X have total measure 1. The closure H of the convex hull of $f(X)$ in V is *compact* by hypothesis. We will show that there is an integral of f inside H.

For *finite* $L \subset V^*$, let

$$V_L \;=\; \{v \in V : \lambda v = \int_X \lambda \circ f,\ \forall \lambda \in L\} \qquad \text{and} \qquad I_L \;=\; H \cap V_L$$

Since H is compact and V_L is closed, I_L is *compact*. Certainly $I_L \cap I_{L'} = I_{L\cup L'}$ for two finite subsets L, L' of V^*. If all the I_L are *non-empty*, then the intersection of *all* these compact sets I_L is non-empty, by the *finite intersection property*, giving *existence*.

To prove that each I_L is non-empty for finite subsets L of V^*, choose an ordering $\lambda_1, \ldots, \lambda_n$ of the elements of L. Make a continuous linear mapping $\Lambda = \Lambda_L$ from V to $\mathbb{R}^n$ by $\Lambda(v) = (\lambda_1 v, \ldots, \lambda_n v)$. Since this map is continuous, the image $\Lambda(f(X))$ is compact in $\mathbb{R}^n$.

For a finite set L of functionals, the integral $y = y_L = \int_X \Lambda f(x)\, dx$ is readily defined by component-wise integration. Take y in the convex hull of $\Lambda(f(X))$. Since Λ_L is linear, $y = \Lambda_L v$ for some v in the convex hull of $f(X)$. Then

$$\Lambda_L v \;=\; y \;=\; (\ldots, \int \lambda_i f(x)\, dx, \ldots)$$

Thus, $v \in I_L$ as desired. It remains to show that y lies in the convex hull of $\Lambda_L(f(x))$.

Suppose *not*. From the lemma below, in a *finite-dimensional* space the convex hull of a compact set is still compact, *without* taking closure. By the finite-dimensional case of the Hahn-Banach theorem, there would be a linear functional η on $\mathbb{R}^n$ so that $\eta y > \eta z$ for all z in this convex hull. That is, letting $y = (y_1, \ldots, y_n)$, there would be real $c_1, \ldots, c_n$ so that for all $(z_1, \ldots, z_n)$ in the convex hull $\sum_i c_i z_i < \sum c_i y_i$. In particular, for all $x \in X$

$$\sum_i c_i \lambda_i(f(x)) \;<\; \sum_i c_i y_i$$

Integration of both sides of this over X *preserves ordering*, giving the impossible $\sum_i c_i y_i < \sum_i c_i y_i$. Thus, y *does* lie in this convex hull. ///

[8.1] Lemma: The convex hull of a compact set K in $\mathbb{R}^n$ is compact.

Proof: First claim that, for $E \subset \mathbb{R}^n$ and for any x a point in the convex hull of E, there are $n+1$ points $x_0, x_1, \ldots, x_n$ in E of which x is a convex combination.

By induction, it suffices to consider a convex combination $v = c_1v_1 + \ldots + c_Nv_N$ of vectors v_i with $N > n+1$ and show that v is actually a convex combination of $N-1$ of the v_i. Further, without loss of generality that all the coefficients c_i are non-zero. Define a linear map

$$L : \mathbb{R}^N \longrightarrow \mathbb{R}^n \times \mathbb{R} \qquad \text{by} \qquad L(x_1, \ldots, x_N) \longrightarrow (\sum_i x_iv_i, \sum_i x_i)$$

By dimension-counting, since $N > n+1$ the kernel of L is non-trivial. Let $(x_1, \ldots, x_N)$ be a non-zero vector in the kernel. Since $c_i > 0$ for every index, and since there are only finitely-many indices altogether, there is a constant c so that $|cx_i| \le c_i$ for every index i, and so that $cx_{i_o} = c_{i_o}$ for at least one index i_o. Then

$$v = v - 0 = \sum_i c_iv_i - c \cdot \sum x_iv_i = \sum_i (c_i - cx_i)v_i$$

Since $\sum_i x_i = 0$ this is still a convex combination, and since $cx_{i_o} = c_{i_o}$ at least one coefficient has become zero. This is the induction proving the claim.

By this claim, a point v in the convex hull of K is a convex combination $c_ov_o + \ldots + c_nv_n$ of $n+1$ points $v_o, \ldots, v_n$ of K. Let σ be the compact set $(c_o, \ldots, c_n)$ with $0 \le c_i \le 1$ and $\sum_i c_i = 1$. The convex hull of K is the image of the compact set $\sigma \times K^{n+1}$ under the continuous map

$$L : (c_o, \ldots, c_n) \times (v_o, v_1, \ldots, v_n) \longrightarrow \sum_i c_iv_i$$

so is compact. This proves the lemma, finishing the proof of the theorem. ///

---

9. *Appendix: Hahn-Banach theorems*

For a locally convex vectorspace V, functionals $\lambda \in V^*$ separate points, and convex sets can be separated by linear functionals. Continuous linear functionals on arbitrary subspaces have continuous extensions to the whole space. In contrast, in general, linear maps from subspaces W to not-finite-dimensional topological vectorspaces need not extend to V. Indeed, if the identity map $T : W \to W$ extended to $T' : V \to W$, then $\ker T'$ would be a *complementary subspace*, which need not exist even for *closed* subspaces W.

Let k be either $\mathbb{R}$ or $\mathbb{C}$, and let V be a k-vectorspace, without any assumptions about topologies for the moment. A k-linear k-valued function on V is a *linear functional.* A linear functional λ on V is *bounded* when there is a neighborhood U of 0 in V and constant c so that $|\lambda x| \le c$ for $x \in U$, where $|\ |$ is the usual absolute value on k. The following proposition is the general analogue of the corresponding assertion for Banach spaces, in which *boundedness* has a different sense.

[9.1] Proposition: The following conditions on a linear functional λ on a topological vectorspace V over k are equivalent: (i) λ is continuous, (ii) λ is continuous at 0, (iii) λ is bounded.

Proof: The first assertion certainly implies the second. Assume the second. Then, given $\varepsilon > 0$, there is a neighborhood U of 0 so that $|\lambda|$ is bounded by ε on U. This proves boundedness. Finally, suppose that $|\lambda(x)| \le c$ on a neighborhood U of 0. Then given $x \in V$ and given $\varepsilon > 0$, we *claim* that for $y \in x + \frac{\varepsilon}{2c}U$ we have $|\lambda(x) - \lambda(y)| < \varepsilon$. Indeed, letting $x - y = \frac{\varepsilon}{2c}u$ with $u \in U$, we have

$$|\lambda(x) - \lambda(y)| = \frac{\varepsilon}{2c}|\lambda(u)| \le \frac{\varepsilon}{2c} \cdot c = \frac{\varepsilon}{2} < \varepsilon$$

This proves the proposition. ///

The immediate goal is to *extend* a linear functional while preserving a comparison to another function (denoted p below). For this, we need *not* suppose that the vectorspaces involved are *topological* vectorspaces. Let V be a *real* vectorspace, without any assumption about topologies. Let $p : V \to \mathbb{R}$ be a *non-negative* real-valued function on V so that

$$p(tv) \;=\; t \cdot p(v) \qquad (\text{for } t \geq 0) \qquad \textit{(positive-homogeneity)}$$

$$p(v+w) \leq p(v) + p(w) \qquad \textit{(triangle inequality)}$$

Lacking a description of $p(tv)$ for $t < 0$, p is not quite a *semi-norm.*

[9.2] Theorem: Let λ be a real-linear function on a real vector subspace W of V, so that $\lambda(w) \leq p(w)$ for all $w \in W$. There is an extension of λ to a real-linear function Λ on all of V, so that $-p(-v) \leq \Lambda(v) \leq p(v)$ for all $v \in V$.

Proof: The key issue is extending the functional *one step*. That is, for $v_o \in V$, attempt to extend λ' of λ to $W + \mathbb{R}v_o$ by $\lambda'(w + tv_o) = \lambda(w) + ct$ and examine the resulting conditions on c.

For all $w, w' \in W$

$$\lambda(w) - p(w - v_o) \;=\; \lambda(w + w') - \lambda(w') - p(w - v_o)$$
$$\leq\; p(w + w') - \lambda(w') - p(w - v_o) \;=\; p(w - v_o + w' + v_o) - \lambda(w') - p(w - v_o)$$
$$\leq\; p(w - v_o) + p(w' + v_o) - \lambda(w') - p(w - v_o) \;=\; p(w' + v_o) - \lambda(w')$$

That is,

$$\lambda(w) - p(w - v_o) \;\leq\; p(w' + v_o) - \lambda(w') \qquad\qquad (\text{for all } w, w' \in W)$$

Let σ be the sup of all the left-hand sides as w ranges over W. Since the right-hand side is finite, this sup is finite. With μ the inf of the right-hand side as w' ranges over W,

$$\lambda(w) - p(w - v_o) \;\leq\; \sigma \;\leq\; \mu \;\leq\; p(w' + v_o) - \lambda(w')$$

Take any real number c so that $\sigma \leq c \leq \mu$ and define $\lambda'(w + tv_o) = \lambda(w) + tc$.

To compare with p is easy: in the inequality $\lambda(w) - p(w - v_o) \leq \sigma$ replace w by w/t with $t > 0$, multiply by t and invoke the positive-homogeneity to obtain $\lambda(w) - p(w - tv_o) \leq t\sigma$ from which

$$\lambda'(w - tv_o) \;=\; \lambda(w) - tc \;\leq\; \lambda(w) - t\sigma \;\leq\; p(w - tv_o)$$

Likewise, from $\mu \leq p(w + v_o) - \lambda(w)$ a similar trick produces

$$\lambda'(w + tv_o) \;=\; \lambda(w) + tc \;\leq\; \lambda(w) + t\mu \;\leq\; p(w + tv_o)$$

for $t > 0$, the other half of the desired inequality. Thus, for all $v \in W + Rv_o$ we have $\lambda'(v) \leq p(v)$. Using the linearity of λ', $\lambda'(v) = -\lambda'(-v) \geq -p(-v)$ giving the bottom half of the comparison of λ' and p.

Extend to a functional on the *whole* space dominated by p by transfinite induction, as follows. Let $\mathcal{X}$ be the collection of all pairs (X, μ), where X is a subspace of V (containing W), and where μ is real-linear real-valued function on X so that μ restricted to W is λ, and so that $-p(-x) \leq \mu(x) \leq p(x)$ for all $x \in X$. *Order* these by writing $(X, \mu) \leq (Y, \nu)$ when $X \subset Y$ and $\nu|_X = \mu$. By the Hausdorff Maximality Principle, there is a *maximal* totally ordered subset $\mathcal{Y}$ of $\mathcal{X}$. Let

$$V' \;=\; \bigcup_{(X,\mu) \in \mathcal{Y}} X$$

be the ascending union of all the subspaces in $\mathcal{Y}$. Define a linear functional λ' on this union as follows: for $v \in V'$, take any X so that $(X, \mu) \in \mathcal{Y}$ and $v \in X$ and define $\lambda'(v) = \mu(v)$. The total ordering on $\mathcal{Y}$ makes the choice of (X, μ) not affect the definition of λ'. If V' were not the whole space V the first part of the proof would create an extension to a properly larger subspace, contradicting the maximality. ///

[9.3] Theorem: For a non-empty convex open subset X of a *locally convex* topological vectorspace V, and a non-empty convex set Y in V with $X \cap Y = \phi$, there is a *continuous* real-linear real-valued functional λ on V and a constant c so that $\lambda(x) < c \le \lambda(y)$ for all $x \in X$ and $y \in Y$.

Proof: Fix $x_o \in X$ and $y_o \in Y$. Since X is open, $X - x_o$ is open, and thus

$$U = (X - x_o) - (Y - y_o) = \{(x - x_o) - (y - y_o) : x \in X, \ \ y \in Y\}$$

is open. Further, since $x_o \in X$ and $y_o \in Y$, U contains 0. Since X, Y are convex, U is convex. The *Minkowski functional* $p = p_U$ attached to U is $p(v) = \inf\{t > 0 : v \in tU\}$. The convexity assures that this function p has the *positive-homogeneity* and *triangle-inequality* properties of the auxiliary functional p above.

Let $z_o = -x_o + y_o$. Since $X \cap Y = \phi$, $z_o \notin U$, so $p(z_o) \ge 1$. Define a linear functional λ on $\mathbb{R}z_o$ by $\lambda(tz_o) = t$. Check that λ is dominated by p in the sense of the previous section:

$$\lambda(tz_o) = t \le t \cdot p(z_o) = p(tz_o) \qquad \text{(for } t \ge 0\text{)}$$

while

$$\lambda(tz_o) = t < 0 \le p(tz_o) \qquad \text{(for } t < 0\text{)}$$

Thus, $\lambda(tz_o) \le p(tz_o)$ for all real t, and λ extends to a real-linear real-valued functional Λ on V, still so that $-p(-v) \le \Lambda(v) \le p(v)$ for all $v \in V$. From the definition of p, $|\Lambda| \le 1$ on U. Thus, on $\frac{\varepsilon}{2}U$ we have $|\Lambda| < \varepsilon$. That is, the linear functional Λ is *bounded*, so is *continuous* at 0, so is *continuous* on V.

For arbitrary $x \in X$ and $y \in Y$,

$$\Lambda x - \Lambda y + 1 = \Lambda(x - y + z_o) \le p(x - y + z_o) < 1$$

since $x - y + z_o \in U$. Thus, $\Lambda x - \Lambda y < 0$ for all such x, y. Therefore, $\Lambda(X)$ and $\Lambda(Y)$ are *disjoint* convex subsets of $\mathbb{R}$. Since Λ is not the zero functional, it is *surjective* to $\mathbb{R}$, and so is an *open* map. Thus, $\Lambda(X)$ is open, and $\Lambda(X) < \sup \Lambda(X) \le \Lambda(Y)$ as desired. ///

The analogous results for complex scalars are corollaries of the real-scalar cases, as follows. Let V be a complex vectorspace. Given a complex-linear complex-valued functional λ on V, let its real part be

$$u(v) = \mathrm{Re}\,\lambda(v) = \frac{\lambda(v) + \overline{\lambda(v)}}{2}$$

where the overbar denotes complex conjugation. On the other hand, given a *real*-linear *real*-valued functional u on V, its *complexification* is $Cu(x) = u(x) - iu(ix)$ where $i = \sqrt{-1}$.

[9.4] Lemma: For a real-linear functional u on the complex vectorspace V, the complexification Cu is a complex-linear functional so that $\mathrm{Re}\,(Cu) = u$ and for a complex-linear functional λ $C(\mathrm{Re}\lambda) = \lambda$. *(Straightforward computation).* ///

[9.5] Corollary: Let p be a *seminorm* on the complex vectorspace V. Let λ be a complex-linear function on a complex vector subspace W of V, so that $|\lambda(w)| \le p(w)$ for all $w \in W$. Then there is an extension of λ to a complex-linear function Λ on all of V, so that $|\Lambda(v)| \le p(v)$ for all $v \in V$.

Proof: Certainly if $|\lambda| \le p$ then $|\mathrm{Re}\lambda| \le p$. By the theorem for *real*-linear functionals, there is an extension u of $\mathrm{Re}\lambda$ to a *real*-linear functional so that still $|u| \le p$. Let $\Lambda = Cu$. In light of the lemma, it remains to show that $|\Lambda| \le p$. To this end, given $v \in V$, let μ be a complex number of absolute value 1 so that $|\Lambda(v)| = \mu\Lambda(v)$. Then

$$|\Lambda(v)| \;=\; \mu\Lambda(v) \;=\; \Lambda(\mu v) \;=\; \mathrm{Re}\,\Lambda(\mu v) \;\le\; p(\mu v) \;=\; p(v)$$

using the seminorm property of p. Thus, the complex-linear functional made by complexifying the *real*-linear extension of the real part of λ satisfies the desired bound. ///

[9.6] Corollary: Let X be a non-empty convex open subset of a locally convex topological vectorspace V, and let Y be an arbitrary non-empty convex set in V so that $X \cap Y = \phi$. Then there is a *continuous* complex-linear complex-valued functional λ on V and a constant c so that

$$\mathrm{Re}\,\lambda(x) \;<\; c \;\le\; \mathrm{Re}\,\lambda(y) \qquad \text{(for all } x \in X \text{ and } y \in Y\text{)}$$

Proof: Invoke the real-linear version of the theorem to make a *real*-linear functional u so that $u(x) < c \le u(y)$ for all $x \in X$ and $y \in Y$. By the lemma, u is the real part of its own complexification. ///

[9.7] Corollary: Let V be a locally convex topological vectorspace. Let K and C be *disjoint* sets, where K is a *compact* convex non-empty subset of V, and C is a *closed* convex subset of V. Then there is a continuous linear functional λ on V and there are real constants $c_1 < c_2$ so that

$$\mathrm{Re}\,\lambda(x) \;\le\; c_1 \;<\; c_2 \;\le\; \mathrm{Re}\,\lambda(y) \qquad \text{(for all } x \in K \text{ and } y \in C\text{)}$$

Proof: Take a small-enough convex neighborhood U of 0 in V so that $(K+U) \cap C = \phi$. Apply the separation theorem to $X = K + U$ and $Y = C$. The constant c_2 can be taken to be $c_2 = \sup \mathrm{Re}\,\lambda(K+U)$. Since $\mathrm{Re}\,\lambda(K)$ is a compact subset of $\mathrm{Re}\,\lambda(K+U)$, its sup c_1 is strictly less than c_2. ///

[9.8] Corollary: Let V be a locally convex topological vectorspace, W a subspace, and $v_o \in V$. Let $\overline{W}$ be the topological closure of W. Then $v_o \notin \overline{W}$ if and only if there is a *continuous* linear functional λ on V so that $\lambda(W) = 0$ while $\lambda(v) = 1$.

Proof: On one hand, if v_o lies in the closure of W, then any continuous function which is 0 on W must be 0 on v_o, as well. On the other hand, suppose that v_o does *not* lie in the closure of W. Then apply the previous corollary with $K = \{v_o\}$ and $C = \overline{W}$. We find that

$$\mathrm{Re}\,\lambda(\{v_o\}) \cap \mathrm{Re}\,\lambda(\overline{W}) \;=\; \phi$$

Since $\mathrm{Re}\,\lambda(\overline{W})$ is a vector subspace of the real line, and is not the whole real line, it is just $\{0\}$, and $\mathrm{Re}\,\lambda(v_o) \ne 0$. Divide λ by the constant $\mathrm{Re}\,\lambda(v_o)$ to obtain a continuous linear functional zero on W but 1 on v_o. ///

[9.9] Corollary: Let V be a locally convex topological (real) vectorspace. Let λ be a continuous linear functional on a subspace W of V. Then there is a continuous linear functional Λ on V extending λ.

Proof: Without loss of generality, take $\lambda \ne 0$. Let W_o be the kernel of λ (on W), and pick $w_1 \in W$ so that $\lambda w_1 = 1$. Evidently w_1 is not in the closure of W_o, so there is Λ on the whole space V so that $\Lambda|_{W_o} = 0$ and $\Lambda w_1 = 1$. It is easy to check that this Λ is an extension of λ. ///

[9.10] Corollary: Let V be a locally convex topological vectorspace. Given two distinct vectors $x \ne y$ in V, there is a continuous linear functional λ on V so that $\lambda(x) \ne \lambda(y)$

Proof: The set $\{x\}$ is compact convex non-empty, and the set $\{y\}$ is closed convex non-empty, so we can apply a corollary just above. ///

10. *Historical notes and references*

Most investigation and use of integration of vector-valued functions is in the context of *Banach-space*-valued functions. Nevertheless, the idea of [Gelfand 1936] extended and developed by [Pettis 1938] immediately suggests a viewpoint not confined to the Banach-space case. A hint appears in [Rudin 1991].

This is in contrast to many of the more detailed studies and comparisons of varying notions of integral specific to the Banach-space case, such as [Bochner 1935]. A variety of developmental episodes and results in the Banach-space-valued case is surveyed in [Hildebrandt 1953]. Proofs and application of many of these results are given in [Hille-Phillips 1957]. (The first edition, authored by Hille alone, is sparser in this regard.) See also [Brooks 1969] to understand the viewpoint of those times.

One of the few exceptions to the apparent limitation to the Banach-space case is [Phillips 1940]. However, it seems that in the United States after the Second World War consideration of anything fancier than Banach spaces was not popular.

The present pursuit of the issue of quasi-completeness (and compactness of the closure of the convex hull of a compact set) was motivated originally by the discussion in [Rudin 1991], although the latter does not make clear that this condition is fulfilled in more than Fréchet spaces, and does not mention quasi-completeness. Imagining that these ideas must be applicable to distributions, one might cast about for means to prove the compactness condition, eventually hitting upon the hypothesis of quasi-completeness in conjunction with ideas from the proof of the Banach-Alaoglu theorem. Indeed, in [Bourbaki 1987] it is shown (by apparently different methods) that quasi-completeness implies this compactness condition, although there the application to vector-valued integrals is not mentioned. [Schaeffer-Wolff 1999] is a very readable account of further important ideas in topological vector spaces.

The fact that a bounded subset of a countable strict inductive limit of closed subspaces must actually be a bounded subset of one of the subspaces, easy to prove once conceived, is attributed to Dieudonne and Schwartz in [Horvath 1966]. See also [Bourbaki 1987], III.5 for this result. Pathological behavior of uncountable colimits was evidently first exposed in [Douady 1963].

Evidently *quotients* of quasi-complete spaces (by closed subspaces, of course) *may fail to be quasi-complete:* see [Bourbaki 1987], IV.63 exercise 10 for a construction.

[Birkhoff 1935] G. Birkhoff, *Integration of functions with values in a Banach space*, Trans. AMS **38** (1935), 357-378.

[Bochner 1935] S. Bochner, *Integration von Funktionen deren Werte die Elemente eines Vektorraumes sind*, Fund. Math. **20** (1935), 262-276.

[Bourbaki 1987] N. Bourbaki, *Topological Vector Spaces, ch. 1-5*, Springer-Verlag, 1987.

[Brooks 1969] J.K. Brooks, *Representations of weak and strong integrals in Banach spaces*, Proc. Nat. Acad. Sci. U.S., 1969, 266-270.

[Douady 1963] A. Douady, *Parties compactes d'un espace de fonctions continues a support compact*, C. R. Acad. Sci. Paris **257** (1963), 2788-2791.

[Dunford 1936] N. Dunford, *Integration of abstract functions*, abstract, Bull. AMS **44** (1936), 178.

[Dunford 1938] N. Dunford, *Uniformity in linear spaces*, Trans. AMS **44** (1938), 305-356.

[Gelfand 1936] I. M. Gelfand, *Sur un lemme de la theorie des espaces lineaires*, Comm. Inst. Sci. Math. de Kharkoff, no. 4, vol. 13, 1936, 35-40.

[Gelfand-Shilov 1964] I. M. Gelfand, G. E. Shilov *Generalized Functions, vol. 1*, translated by E. Saletan, Academic Press, 1964.

[Hildebrandt 1953] T. H. Hildebrandt, *Integration in abstract spaces*, Bull. AMS, **59** (1953), 111-139.

[Hille-Phillips 1957] E. Hille with R. Phillips, *Functional Analysis and Semigroups*, AMS Colloquium Publication, second edition, 1957.

[Horvath 1966] J. Horvath, *Topological Vector Spaces and Distributions*, Addison-Wesley, 1966.

[Pettis 1938] B. J. Pettis, *On integration in vector spaces*, Trans. AMS **44** (1938), 277-304.

[Phillips 1940] R. S. Phillips, *Integration in a convex linear topological space*, Trans. AMS, **47** (1940), 114-145.

[Rudin 1991] W. Rudin, *Functional Analysis*, second edition, McGraw-Hill, 1991.

[Schaeffer-Wolff 1999] H.H. Schaeffer, with M.P. Wolff, *Topological Vector Spaces*, second edition, Springer, 1999.

[Taylor 1938] A. E. Taylor, *The resolvent of a closed transformation*, Bull. AMS **44** (1938), 70-74.

[Taylor 1971] A. E. Taylor, *Notes on the history of the uses of analyticity in operator theory*, Amer. Math. Monthly **78** (1971), 331-342.