THE SCHWARTZ SPACE: A BACKGROUND TO WHITE NOISE ANALYSIS

JEREMY J. BECNEL AND AMBAR N. SENGUPTA

Abstract. An account of the Schwartz space of rapidly decreasing functions as a nuclear space is presented, along with a description of the Gaussian measure on the dual space.

1. Introduction

In this expository paper we present an account of

- the Schwartz space $\mathcal{S}(\mathbf{R}^d)$ of rapidly decreasing functions as a nuclear space,

in a largely self-contained way, concluding with a construction of the standard Gaussian measure on the dual space $\mathcal{S}'(\mathbf{R}^d)$, directed primarily to those who plan to delve further into white noise analysis. We work out the properties of the useful operator

$$T = -\frac{d^2}{dx^2} + \frac{x^2}{4} + \frac{1}{2} \tag{1.1}$$

on $\mathcal{S}(\mathbf{R})$, in terms of creation and annihilation operators, and describe in brief the origins of these notions in quantum mechanics. The operator T arose in quantum mechanics as the Hamiltonian for a harmonic oscillator and, in that context as well as in white noise analysis, the operator $N = T - 1$ is called the *number operator*.

Our exposition of the properties of T and of $\mathcal{S}(\mathbf{R})$ follows Barry Simon's [4], but we provide more detail (and our notational conventions are slightly different). For the Gaussian measure, we describe a direct construction using the Kolmogorov theorem, instead of the more general (and more difficult) Minlos theorem.

2. Objectives

In this section we summarize the basic notions, notation, and results that we discuss in more detail in later sections. Here, and later in this paper, we will work mainly with the case of functions of one variable and then describe the generalization to the multi-dimensional case.

We will use the letter W to denote the set of all non-negative integers:

$$W = \{0, 1, 2, 3, ...\} \tag{2.1}$$

Date: August 2004.
Research supported by US NSF grant DMS-0201683.

2.1. **The Schwartz Space.** The *Schwartz space* $\mathcal{S}(\mathbf{R})$ is the linear space of all functions $f : \mathbf{R} \to \mathbf{C}$ which have derivatives of all orders and which satisfy the condition

$$p_{a,b}(f) \stackrel{\text{def}}{=} \sup_{x \in \mathbf{R}} |x^a f^b(x)| < \infty$$

for all $a, b \in W = \{0, 1, 2, ...\}$. The finiteness condition for all $a \geq 1$ and $b \in W$, implies that $x^a f^b(x)$ actually goes to 0 as $|x| \to \infty$, for all $a, b \in W$, and so functions of this type are said to be *rapidly decreasing*.

2.2. **The Schwartz Topology.** The functions $p_{a,b}$ are *semi-norms* on the vector space $\mathcal{S}(\mathbf{R})$, in the sense that

$$p_{a,b}(f + g) \leq p_{a,b}(f) + p_{a,b}(g)$$

and

$$p_{a,b}(zf) = zp_{a,b}(f)$$

for all $f, g \in \mathcal{S}(\mathbf{R})$, and $z \in \mathbf{C}$. For this semi-norm, an open ball of radius r centered at some $f \in \mathcal{S}(\mathbf{R})$ is given by

$$B_{p_{a,b}}(f; r) = \{g \in \mathcal{S}(\mathbf{R}) : p_{a,b}(g - f) < r\} \tag{2.2}$$

Thus each $p_{a,b}$ specifies a topology $\tau_{a,b}$ on $\mathcal{S}(\mathbf{R})$. A set is open according to $\tau_{a,b}$ if it is a union of open balls.

The topologies $\tau_{a,b}$ put all together, generate the standard *Schwartz topology* τ on $\mathcal{S}(\mathbf{R})$. This is the smallest topology containing all the sets of $\tau_{a,b}$ for all $a, b \in W$.

There is a different approach to the topology on $\mathcal{S}(\mathbf{R})$ that is very useful for analysis, and much of our effort in this paper will go into demonstrating the equivalence of the two ways of understanding the topology on $\mathcal{S}(\mathbf{R})$.

2.3. **The operator T.** The operator

$$T = -\frac{d^2}{dx^2} + \frac{x^2}{4} + \frac{1}{2} \tag{2.3}$$

plays a very useful role in working with the Schwartz space. As we shall see, there is an orthonormal basis $\{\phi_n\}_{n \in W}$ of $L^2(\mathbf{R}, dx)$, where

$$W = \{0, 1, 2, ...\},$$

consisting of eigenfunctions ϕ_n of T:

$$T\phi_n = (n + 1)\phi_n \tag{2.4}$$

The functions ϕ_n, called the *Hermite functions* are actually in the Schwartz space $\mathcal{S}(\mathbf{R})$. Let B be the bounded linear operator on $L^2(\mathbf{R})$ given on each $f \in L^2(\mathbf{R})$ by

$$Bf = \sum_{n \in W} (n + 1)^{-1} \langle f, \phi_n \rangle \phi_n \tag{2.5}$$

It is readily checked that the right side does converge and, in fact,

$$\|Bf\|^2_{L^2} = \sum_{n \in W} (n + 1)^{-2} |\langle f, \phi_n \rangle|^2_{L^2} \leq \sum_{n \in W} |\langle f, \phi_n \rangle|^2_{L^2} = \|f\|^2_{L^2} \tag{2.6}$$

Note that B and T are inverses of each other on the linear span of the vectors ϕ_n:

$$TBf = f \text{ and } BTf = f \text{ for all } f \in \mathcal{L}, \tag{2.7}$$

where

$$\mathcal{L} = \text{linear span of the vectors } \phi_n, \text{ for } n \in W \tag{2.8}$$

2.4. **The L^2 approach.** For any $p \geq 0$, the image of B^p consists of all $f \in L^2(\mathbf{R})$ for which

$$\sum_{n\in W}(n+1)^{2p}|\langle f,\phi_n\rangle|^2 < \infty.$$

Let

$$\mathcal{S}_p(\mathbf{R}) = B^p\left(L^2(\mathbf{R})\right) \tag{2.9}$$

This is a subspace of $L^2(\mathbf{R})$, and on $\mathcal{S}_p(\mathbf{R})$ there is an inner-product $\langle\cdot,\cdot\rangle_p$ given by

$$\langle f,g\rangle_p \stackrel{\text{def}}{=} \sum_{n\in W}(n+1)^{2p}\langle f,\phi_n\rangle\langle\phi_n,g\rangle = \langle B^{-p}f, B^{-p}g\rangle_{L^2} \tag{2.10}$$

which makes it a Hilbert space, with $\mathcal{L}$, and hence also $\mathcal{S}(\mathbf{R})$, a dense subspace. We will see later that functions in $\mathcal{S}_p(\mathbf{R})$ are p-times differentiable.

We will prove that *the intersection $\cap_{p\in W}\mathcal{S}_p(\mathbf{R})$ is exactly equal to* $\mathcal{S}(\mathbf{R})$. In fact,

$$\mathcal{S}(\mathbf{R}) = \cap_{p\in W}\mathcal{S}_p(\mathbf{R}) \subset \cdots \mathcal{S}_2(\mathbf{R}) \subset \mathcal{S}_1(\mathbf{R}) \subset \mathcal{S}_0(\mathbf{R}) = L^2(\mathbf{R}) \tag{2.11}$$

We will also prove that *the topology on $\mathcal{S}(\mathbf{R})$ generated by the norms $\|\cdot\|_p$ coincides with the standard topology.* Furthermore, the elements $(n+1)^{-p}\phi_n \in \mathcal{S}_p(\mathbf{R})$ form an orthonormal basis of $\mathcal{S}_p(\mathbf{R})$, and

$$\sum_{n\in W}\|(n+1)^{-(p+1)}\phi_n\|_p^2 = \sum_{n\geq 1}\frac{n^{2p}}{n^{2(p+1)}} < \infty,$$

showing that

the inclusion map $\mathcal{S}_{p+1}(\mathbf{R}) \to \mathcal{S}_p(\mathbf{R})$ is *Hilbert-Schmidt.*

The topological vector space $\mathcal{S}(\mathbf{R})$ has topology generated by a *complete metric*, and *has a countable dense subset* given by all finite linear combinations of the vectors ϕ_n with rational coefficients.

2.5. **Coordinatization as a Sequence Space.** All of the results described above follow readily from the identification of $\mathcal{S}(\mathbf{R})$ with a space of sequences. Let $\{\phi_n\}_{n\in W}$ be the orthonormal basis of $L^2(\mathbf{R})$ mentioned above, where

$$W = \{0,1,2,...\}$$

Then we have the set $\mathbf{C}^W$. An element $a \in \mathbf{C}^W$ is a map $W \to \mathbf{C} : n \mapsto a_n$. So we shall often write such an element a as $(a_n)_{n\in W}$.

We have then the coordinatizing map

$$I : L^2(\mathbf{R}) \to \mathbf{C}^W : f \mapsto (\langle f,\phi_n\rangle)_{n\in W} \tag{2.12}$$

For each $p \in W$ let E_p be the subset of $\mathbf{C}^W$ consisting of all $(a_n)_{n\in W}$ such that

$$\sum_{n\in W}(n+1)^{2p}|a_n|^2 < \infty$$

On E_p define the inner-product $\langle\cdot,\cdot\rangle_p$ by

$$\langle a,b\rangle_p = \sum_{n\in W}(n+1)^{2p}a_n\overline{b}_n \tag{2.13}$$

This makes E_p a Hilbert space, essentially the Hilbert space $L^2(W, \mu_p)$ where μ_p is the measure on W given by $\mu_p(\{n\}) = (n+1)^{2p}$ for all $n \in W$.

The definition (2.9) for $\mathcal{S}_p(\mathbf{R})$ shows that it is the set of all $f \in L^2(\mathbf{R})$ for which $I(f)$ belongs to E_p.

We will prove in Theorem 11.1 that I *maps* $\mathcal{S}(\mathbf{R})$ *exactly onto*

$$E \stackrel{\text{def}}{=} \cap_{p\in W} E_p \tag{2.14}$$

This will establish essentially all of the facts mentioned above concerning the spaces $\mathcal{S}_p(\mathbf{R})$.

Note the chain of inclusions:

$$E = \cap_{p\in W} E_p \subset \cdots \subset E_2 \subset E_1 \subset E_0 = L^2(W, \mu_0) \tag{2.15}$$

2.6. **The multi-dimensional setting.** In the multidimensional setting, the Schwartz space $\mathcal{S}(\mathbf{R}^d)$ consists of all infinitely differentiable functions f on $\mathbf{R}^d$ for which

$$\sup_{x\in\mathbf{R}^d} \left| x_1^{k_1} \ldots x_d^{k_d} \frac{\partial^{m_1+\cdots+m_k} f(x)}{\partial x_1^{m_1} \ldots \partial x_d^{m_d}} \right| < \infty,$$

for all $(k_1, ..., k_d) \in W^d$ and $m = (m_1, ..., m_d) \in W^d$. For this setting, it is best to use some standard notation:

$$|k| = k_1 + \cdots + k_d \qquad \text{for } k = (k_1, ..., k_d) \in W^d \tag{2.16}$$

$$x^k = x^{k_1} \ldots x^{k_d} \tag{2.17}$$

$$D^m = \frac{\partial^{|m|}}{\partial x_1^{m_1} \ldots \partial x_d^{m_d}} \tag{2.18}$$

For the multi-dimensional case, we use the indexing set W^d whose elements are d–tuples $j = (j_1, ..., j_d)$, with $j_1, ..., j_d \in W$, and counting measure μ_0 on W^d. The sequence space is replaced by $\mathbf{C}^{W^d}$; a typical element $a \in \mathbf{C}^{W^d}$, is a map

$$a : W^d \to \mathbf{C} : j = (j_1, ..., j_d) \mapsto a_j = a_{j_1 \ldots j_d} \tag{2.19}$$

The orthonormal basis $(\phi_n)_{n\in W}$ of $L^2(\mathbf{R})$ yields an orthonormal basis of $L^2(\mathbf{R}^d)$ consisting of the vectors

$$\phi_j = \phi_{j_1} \otimes \cdots \otimes \phi_{j_d} : (x_1, ..., x_d) \mapsto \phi_{j_1}(x_1) \ldots \phi_{j_d}(x_d) \tag{2.20}$$

The coordinatizing map I is replaced by the map

$$I_d : L^2(\mathbf{R}^d) \to \mathbf{C}^{W^d} \tag{2.21}$$

where

$$I_d(f)_j = \langle f, \phi_j \rangle_{L^2(\mathbf{R}^d)} \tag{2.22}$$

Replace the operator T by

$$T_d \stackrel{\text{def}}{=} T^{\otimes d} = \left(-\frac{\partial^2}{\partial x_d^2} + \frac{x_d^2}{4} + 1 \right) \cdots \left(-\frac{\partial^2}{\partial x_1^2} + \frac{x_1^2}{4} + 1 \right) \tag{2.23}$$

Then

$$T_d \phi_j = (j_1 + 1) \ldots (j_d + 1) \phi_j$$

for all $j \in W^d$.

In place of B, we now have the bounded operator B_d on $L^2(\mathbf{R}^d)$ given by

$$B_d f = \sum_{j\in W^d} [(j_1 + 1) \ldots (j_d + 1)]^{-1} \langle f, \phi_j \rangle \phi_j \tag{2.24}$$

Again, T_d and B_d are inverses of each other on the linear subspace $\mathcal{L}_d$ of $L^2(\mathbf{R}^d)$ spanned by the vectors ϕ_j.

The space E_p is now the subset of $\mathbf{C}^{W^d}$ consisting of all $a \in \mathbf{C}^{W^d}$ for which

$$\sum_{j \in W^d} [(j_1+1)...(j_d+1)]^{2p}|a_j|^2 < \infty$$

and this is a Hilbert space with inner-product

$$\langle a, b\rangle_b = \sum_{j \in W^d} [(j_1+1)...(j_d+1)]^{2p} a_j \overline{b}_j$$

Again we have the chain of spaces

$$E \stackrel{\text{def}}{=} \cap_{p \in W} E_p \subset \cdots E_2 \subset E_1 \subset E_0 = L^2(W^d, \mu_0),$$

with the inclusion $E_{p+1} \to E_p$ being Hilbert-Schmidt.

To go back to functions on $\mathbf{R}^d$, define $\mathcal{S}_p(\mathbf{R^d})$ to be the range of B_d. Thus $\mathcal{S}_p(\mathbf{R^d})$ is the set of all $f \in L^2(\mathbf{R}^d)$ for which

$$\sum_{j \in W^d} [(j_1+1)...(j_d+1)]^{2p}|\langle f, \phi_j\rangle|^2 < \infty$$

The inner-product $\langle \cdot, \cdot\rangle_p$ comes back to an inner-product, also denoted $\langle \cdot, \cdot\rangle_p$, on $\mathcal{S}_p(\mathbf{R^d})$ and is given by

$$\langle f, g\rangle_p = \langle B_d^{-p} f, B_d^{-p} g\rangle_{L^2(\mathbf{R}^d)} \tag{2.25}$$

The intersection $\cap_{p \in W}\mathcal{S}_p(\mathbf{R^d})$ equals $\mathcal{S}(\mathbf{R^d})$. Moreover, the topology on $\mathcal{S}(\mathbf{R^d})$ is the smallest one generated by the inner-products obtained from $\langle \cdot, \cdot\rangle_p$, with p running over W.

3. Topological Vector Spaces

The Schwartz space is a *topological vector space*, i.e. it is a vector space equipped with a Hausdorff topology with respect to which the vector space operations (addition, and multiplication by scalar) are continuous. In this section we shall go through a few of the basic notions and results for topological vector spaces.

Let V be a real vector space.

A *vector topology* τ on V is a topology such that addition $V \times V \to V : (x, y) \mapsto x + y$ and scalar multiplication $\mathbf{R} \times V \to V : (t, x) \mapsto tx$ are continuous. If V is a complex vector space we require that $\mathbf{C} \times V \to V : (\alpha, x) \mapsto \alpha x$ be continuous.

It is useful to observe that when V is equipped with a vector topology, the translation maps

$$t_x : V \to V : y \mapsto y + x$$

are continuous, for every $x \in V$, and are hence also homeomorphisms since $t_x^{-1} = t_{-x}$.

A *topological vector space* is a vector space equipped with a Hausdorff vector topology.

A *local base* of a vector topology τ is a family of open sets $\{U_\alpha\}_{\alpha \in I}$ containing 0 such that if W is any open set containing 0 then W contains some U_α. If U is any open set and x any point in U then $U - x$ is an open neighborhood of 0 and hence contains some U_α, and so U itself contains a neighborhood $x + U_\alpha$ of x:

(3.1) *If U is open and $x \in U$ then $x + U_\alpha \subset U$, for some $\alpha \in I$*

Doing this for each point x of U, we see that each open set is the union of translates of the local base sets U_α.

3.1. **Local Convexity and the Minkowski Functional.** A vector topology τ on V is *locally convex* if for any neighborhood W of 0 there is a convex open set B with $0 \in B \subset W$. Thus, local convexity means that there is a local base of the topology τ consisting of convex sets. The principal consequence of having a convex local base is the Hahn-Banach theorem which guarantees that continuous linear functionals on subspaces of V extend to continuous linear functionals on all of V. In particular, if $V \neq \{0\}$ is locally convex then there exist non-zero continuous linear functionals on V.

Let B be a convex open neighborhood of 0. Continuity of $\mathbf{R} \times V \to V : (s, x) \mapsto sx$ at $s = 0$ shows that for each x the multiple sx lies in B if s is small enough, and so $t^{-1}x$ lies in B if t is large enough. The smallest value of t for which $t^{-1}x$ is just outside B is clearly a measure of how large x is relative to B. The *Minkowski functional* μ_B is the function on V given by

$$\mu_B(x) = \inf\{t > 0 : t^{-1}x \in B\}$$

Note that $0 \leq \mu_B(x) < \infty$. The definition of μ_B shows that $\mu_B(kx) = k\mu_B(x)$ for any $k \geq 0$. Convexity of B can be used to show that

$$\mu_B(x + y) \leq \mu_B(x) + \mu_B(y)$$

If B is symmetric, i.e. $B = -B$, then $\mu_B(kx) = |k|\mu_B(x)$ for all real k. If V is a complex vector space and B is *balanced* in the sense that $\alpha B = B$ for all complex numbers α with $|\alpha| = 1$, then $\mu_B(kx) = |k|\mu_B(x)$ for all complex k. Note that in general it could be possible that $\mu_B(x)$ is 0 without x being 0; this would happen if B contains the entire ray $\{tx : t \geq 0\}$.

3.2. **Semi-norms.** A typical vector topology on V is specified by a *semi-norm* on V, i.e. a function $\mu : V \to \mathbf{R}$ such that

$$\mu(x + y) \leq \mu(x) + \mu(y), \qquad \mu(tx) = |t|\mu(x) \tag{3.2}$$

for all $x, y \in V$ and $t \in \mathbf{R}$ (complex t if V is a complex vector space). Note that then, using $t = 0$, we have $\mu(0) = 0$ and, using $-x$ for y, we have $\mu(x) \geq 0$. For such a semi-norm, an open ball around x is the set

$$B_\mu(x; r) = \{y \in V : \mu(y - x) < r\}, \tag{3.3}$$

and the topology τ_μ consists of all sets which can be expressed as unions of open balls. These balls are convex and so the topology τ_μ is locally convex. If μ is actually a norm, i.e. $\mu(x)$ is 0 only if x is 0, then τ_μ is Hausdorff.

A consequence of the triangle inequality (3.2) is that a semi-norm μ is uniformly continuous with respect to the topology it generates. To see this, consider any $x, y \in V$. From

$$\mu(x) = \mu(x - y + y) \leq \mu(x - y) + \mu(y)$$

we have

$$\mu(x) - \mu(y) \leq \mu(x - y)$$

Interchanging x and y, we conclude that

$$|\mu(x) - \mu(y)| \leq \mu(x - y), \tag{3.4}$$

which implies that μ, as a function on V, is continuous with respect to the topology τ_μ it generates. Now suppose μ is continuous with respect to a vector topology τ. Then the open balls $\{y \in V : \mu(y-x) < r\}$ are open in the topology τ and so $\tau_\mu \subset \tau$.

3.3. **Topologies generated by families of topologies.** Let $\{\tau_\alpha\}_{\alpha\in I}$ be a collection of topologies on a space. It is natural and useful to consider the the least upper bound topology τ, i.e. the smallest topology containing all sets of $\cup_{\alpha\in I}\tau_\alpha$. In our setting, we work with each τ_α a vector topology on a vector space V.

Theorem 3.1. *The least upper bound topology τ of a collection $\{\tau_\alpha\}_{\alpha\in I}$ of vector topologies is again a vector topology. If $\{W_{\alpha,i}\}_{i\in I_\alpha}$ is a local base for τ_α then a local base for τ is obtained by taking all finite intersections of the form $W_{\alpha_1,i_1}\cap\cdots\cap W_{\alpha_n,i_n}$.*

Proof. Let $\mathcal{B}$ be the collection of all sets which are of the form $W_{\alpha_1,i_1}\cap\cdots\cap W_{\alpha_n,i_n}$.

Let τ' be the collection of all sets which are unions of translates of sets in $\mathcal{B}$ (including the empty union). Our first objective is to show that τ' is a topology on V. It is clear that τ' is closed under unions and contains the empty set. We have to show that the intersection of two sets in τ' is in τ'. To this end, it will suffice to prove the following:

(3.5) If C_1 and C_2 are sets in $\mathcal{B}$, and x is a point in the intersection of the translates $a+C_1$ and $b+C_2$, then $x+C\subset(a+C_1)\cap(b+C_2)$ for some C in $\mathcal{B}$.

Clearly, it suffices to consider finitely many topologies τ_α. Thus, consider vector topologies $\tau_1,...,\tau_n$ on V.

Let $\mathcal{B}_n$ be the collection of all sets of the form $B_1\cap\cdots\cap B_n$ with B_i in a local base for τ_i, for each $i\in\{1,...,n\}$. We can check that if $D, D'\in\mathcal{B}_n$ then there is an $E\in\mathcal{B}_n$ with $E\subset D\cap D'$.

Working with B_i drawn from a given local base for τ_i, let z be a point in the intersection $B_1\cap\cdots\cap B_n$. Then there exist sets B_i', with each B_i' being in the local base for τ_i, such that $z+B_i'\subset B_i$ (this follows from our earlier observation (3.1)). Consequently,

$$z+\cap_{i=1}^n B_i'\subset\cap_{i=1}^n B_i$$

Now consider sets C_1 an C_2, both in $\mathcal{B}_n$. Consider $a,b\in V$ and suppose $x\in(a+C_1)\cap(b+C_2)$. Then since $x-a\in C_1$ there is a set $C_1'\in\mathcal{B}_n$ with $x-a+C_1'\subset C_1$; similarly, there is a $C_2'\in\mathcal{B}_n$ with $x-b+C_2'\subset C_2$. So $x+C_1'\subset a+C_1$ and $x+C_2'\subset b+C_2$. So

$$x+C\subset(a+C_1)\cap(b+C_2),$$

where $C\in\mathcal{B}_n$ satisfies $C\subset C_1\cap C_2$.

This establishes (3.5), and shows that the intersection of two sets in τ' is in τ'.

Thus τ' is a topology. The definition of τ' makes it clear that τ' contains each τ_α. Furthermore, if any topology σ contains each τ_α then all the sets of τ' are also open relative to σ. Thus

$$\tau'=\tau,$$

the topology generated by the topologies τ_α.

Observe that we have shown that if $W\in\tau$ contains 0 then $W\supset B$ for some $B\in\mathcal{B}$.

Next we have to show that τ is a vector topology. The definition of τ shows that τ is translation invariant, i.e. translations are homeomorphisms. So, for addition, it will suffice to show that addition $V \times V \times V : (x,y) \mapsto x+y$ is continuous at $(0,0)$. Let $W \in \tau$ contain 0. Then there is a $B \in \mathcal{B}$ with $0 \in B \subset W$. Suppose $B = B_1 \cap \cdots \cap B_n$, where each B_i is in the given local base for τ_i. Since τ_i is a vector topology, there are open sets $D_i, D_i' \in \tau_i$, both containing 0, with

$$D_i + D_i' \subset B_i$$

Then choose C_i, C_i' in the local base for τ_i with $C_i \subset D_i$ and $C_i' \subset D_i'$. Then

$$C_i + C_i' \subset B_i$$

Now let $C = C_1 \cap \cdots \cap C_n$, and $C' = C_1' \cap \cdots \cap C'n$. Then $C, C' \in \mathcal{B}$ and $C + C' \subset B$. Thus, addition is continuous at $(0,0)$.

Now consider the multiplication map $\mathbf{R} \times V \to V : (t,x) \mapsto tx$. Let $(s,y),(t,x) \in \mathbf{R} \times V$. Then

$$sy - tx = (s-t)x + t(y-x) + (s-t)(y-x)$$

Suppose $F \in \tau$ contains tx. Then

$$F \supset tx + W',$$

for some $W' \in \mathcal{B}$. Using continuity of the addition map

$$V \times V \times V \to V : (a,b,c) \mapsto a+b+c$$

at $(0,0,0)$, we can choose $W_1, W_2, W_3 \in \mathcal{B}$ with $W_1 + W_2 + W_3 \subset W'$. Then we can choose $W \in \mathcal{B}$, such that

$$W \subset W_1 \cap W_2 \cap W_3$$

Then $W \in \mathcal{B}$ and

$$W + W + W \subset W'$$

Suppose $W = B_1 \cap \cdots \cap B_n$, where each B_i is in the given local base for the vector topology τ_i. Then for s close enough to t, we have $(s-t)x \in B_i$ for each i, and hence $(s-t)x \in W$. Similarly, if y is τ–close enough to x then $t(y-x) \in W$. Lastly, if $s-t$ is close enough to 0 and y is close enough to x then $(s-t)(y-x) \in W$. So $sy - tx \in W'$, and so $sy \in F$, when s is close enough to t and y is τ–close enough to x. □

The above result makes it clear that if each τ_α has a convex local base then so is τ. Note also that if at least one τ_α is Hausdorff then so is τ.

A family of topologies $\{\tau_\alpha\}_{\alpha \in I}$ is *directed* if for any $\alpha, \beta \in I$ there is a $\gamma \in I$ such that

$$\tau_\alpha \cup \tau_\beta \subset \tau_\gamma$$

In this case every open neighborhood of 0 in the generated topology contains an open neighborhood in one of the topologies τ_γ.

3.4. **Topologies generated by families of semi-norms.** We are concerned mainly with the topology τ generated by a family of semi-norms $\{\mu_\alpha\}_{\alpha\in I}$; this is the smallest topology containing all sets of $\cup_{\alpha\in I}\tau_{\mu_\alpha}$. An open set in this topology is a union of translates of finite intersections of balls of the form $B_{\mu_i}(0;r_i)$. Thus any open neighborhood of f contains a set of the form

$$B_{\mu_1}(f;r_1)\cap\cdots\cap B_{\mu_n}(f;r_n)$$

This topology is Hausdorff if for any non-zero $x\in V$ there is some norm μ_α for which $\mu_\alpha(x)$ is not zero.

The description of the neighborhoods in the topology τ shows that a sequence f_n converges to f with respect to τ if and only if $\mu_\alpha(f_n-f)\to 0$, as $n\to\infty$, for all $\alpha\in I$.

We will need to examine when two families of semi-norms give rise to the same topology:

Theorem 3.2. *Let τ be the topology on V generated by a family of semi-norms $\mathcal{M}=\{\mu_i\}_{i\in I}$, and τ' the topology generated by a family of semi-norms $\mathcal{M}'=\{\mu'_j\}_{j\in J}$. Suppose each μ_i is bounded above by a linear combination of the μ'_j. Then $\tau\subset\tau'$.*

Proof. Let $\mu\in\mathcal{M}$. Then there exist $\mu'_1,...,\mu'_n\in\mathcal{M}'$, and real numbers $c_1,...,c_n>0$, such that

$$\mu\le c_1\mu'_1+\cdots+c_n\mu'_n$$

Now consider any $x,y\in V$. Then

$$|\mu(x)-\mu(y)|\le\mu(x-y)\le\sum_{i=1}^{n}|c_i|\mu'_i(x-y)$$

So μ is continuous with respect to the topology generated by $\mu'_1,...,\mu'_n$. Thus, $\tau_\mu\subset\tau'$. Since this is true for all $\mu\in\mathcal{M}$, we have $\tau\subset\tau'$.

□

3.5. **Completeness.** A sequence $(x)_{n\in\mathbf{N}}$ in a topological vector space V is *Cauchy* if for any neighborhood U of 0 in V, the difference x_n-x_m lies in U when n and m are large enough.

The topological vector space V is *complete* if every Cauchy sequence converges.

Theorem 3.3. *Let $\{\tau_\alpha\}_{\alpha\in I}$ be a directed family of Hausdorff vector topologies on V, and τ the generated topology. If each τ_α is complete then so is τ.*

Proof. Let $(x_n)_{n\ge1}$ be a sequence in V, which is Cauchy with respect to τ. Then clearly it is Cauchy with respect to each τ_α. Let $x_\alpha=\lim_{n\to\infty}x_n$, relative to τ_α. If $\tau_\alpha\subset\tau_\gamma$ then the sequence $(x_n)_{n\ge1}$ also converges to x_γ relative to the topology τ_α, and so $x_\gamma=x_\alpha$. Consider $\alpha,\beta\in I$, and choose $\gamma\in I$ such that $\tau_\alpha\cup\tau_\beta\subset\tau_\gamma$. This shows that $x_\alpha=x_\gamma=x_\beta$, i.e. all the limits are equal to each other. Let x denote the common value of this limit. We have to show that $x_n\to x$ in the topology τ. Let $W\in\tau$ contain x. Since the family $\{\tau_\alpha\}_{\alpha\in I}$ which generates τ is directed, it follows that there is a $\beta\in I$ and a $B_\beta\in\tau_\beta$ with $x\in B_\beta\subset W$. Since $(x_n)_{n\ge1}$ converges to x with respect to τ_β, it follows $x_n\in B_\beta$ for large n. So $x_n\to x$ with respect to τ. □

3.6. **Metrizability.** Suppose the topology τ on the topological vector space V is generated by a countable family of semi-norms $\mu_1, \mu_2, \ldots$. For any $x, y \in V$ define

$$d(x,y) = \sum_{n\geq 1} 2^{-n} d_n(x,y) \tag{3.6}$$

where

$$d_n(x,y) = \min\{1, \mu_n(x-y)\}$$

Then d is a metric, it is translation invariant, and generates the topology τ.

4. The Schwartz Space $\mathcal{S}(\mathbf{R})$

Our objective in this section is to show that the Schwartz space is *complete*, in the sense that every Cauchy sequence converges.

Recall that $\mathcal{S}(\mathbf{R})$ is the set of all C^∞ functions f on $\mathbf{R}$ for which

$$p_{a,b}(f) \stackrel{\text{def}}{=} \|f\|_{a,b} \stackrel{\text{def}}{=} \sup_{x\in\mathbf{R}} |x^a D^b f(x| < \infty \tag{4.1}$$

for all $a, b \in W = \{0, 1, 2, \ldots\}$. The functions $p_{a,b}$ are semi-norms, with $\|\cdot\|_{0,0}$, being just the sup-norm, is actually a norm. Thus the family of semi-norms given above specify a Hausdorff vector topology on $\mathcal{S}(\mathbf{R})$. We will call this the *Schwartz topology* on $\mathcal{S}(\mathbf{R})$.

Theorem 4.1. *The topology on $\mathcal{S}(\mathbf{R})$ generated by the family of semi-norms $\|\cdot\|_{a,b}$ for all $a, b \in \{0, 1, 2, \ldots\}$, is complete.*

Proof. . Let $(f_n)_{n\geq 1}$ be a Cauchy sequence on $\mathcal{S}(\mathbf{R})$. Then this sequence is Cauchy in each of the semi-norms $\|\cdot\|_{a,b}$, and so each sequence of functions $x^a D^b f_n(x)$ is uniformly convergent. Let

$$g_b(x) = \lim_{n\to\infty} D^b f_n(x) \tag{4.2}$$

Let $f = g_0$. Using a Taylor theorem argument it follows that g_b is $D^b f$. For instance, for $b = 1$, observe first that

$$f_n(y) = f_n(x) + \int_0^1 \frac{df_n\big((1-t)x + ty\big)}{dt}\,dt = f_n(x) + \int_0^1 f_n'\big((1-t)x + ty\big)(y-x)\,dt,$$

and so, letting $n \to \infty$, we have

$$f(y) = f(x) + \int_0^1 g_1\big((1-t)x + ty\big)(y-x)\,dt,$$

which implies that $f'(x)$ exists and equals $g_1(x)$.

In this way, we have

$$x^a D^b f_n(x) \to x^a D^b f(x) \text{ pointwise.} \tag{4.3}$$

Note that our Cauchy hypothesis implies that the sequence of functions $x^a D^b f_n(x)$ is Cauchy in sup-norm, and so the convergence

$$x^a D^b f_n(x) \to x^a D^b f(x)$$

is uniform. In particular, the sup-norm of $x^a D^b f(x)$ is finite, since it is the limit of a uniformly convergent sequence of bounded functions. Thus $f \in \mathcal{S}(\mathbf{R})$.

Finally, we have to check that f_n converges to f in the topology of $\mathcal{S}(\mathbf{R})$. We have noted above that $x^a D^b f_n(x) \to x^a D^b f(x)$ uniformly. Thus $f_n \to f$ relative to the semi-norm $\|\cdot\|_{a,b}$. Since this holds for every $a, b \in \{0, 1, 2, 3, ...\}$, we have $f_n \to f$ in the topology of $\mathcal{S}(\mathbf{R})$. □

Now lets take a quick look at the Schwartz space $\mathcal{S}(\mathbf{R}^d)$. First some notation. A *multi-index* a is an element of $\{0, 1, 2, ...\}^d$, i.e. it is a mapping

$$a : \{1, ..., d\} \to \{0, 1, 2, ...\} : j \mapsto a_j$$

If a is a multi-index, we write $|a|$ to mean the sum $a_1 + \cdots + a_d$, x^a to mean the product $x_1^{a_1}...x_d^{a_d}$, and D^b to mean the differential operator $D_{x_1}^{a_1}...D_{x_d}^{a_d}$. The space $\mathcal{S}(\mathbf{R}^d)$ consists of all C^∞ functions f on R^d such that each function $x^a D^b f(x)$ is bounded. On $\mathcal{S}(\mathbf{R^d})$ we have the semi-norms

$$\|f\|_{a,b} = \sup_{x \in \mathbf{R}^d} |x^a D^b f(x)|$$

for each pair of multi-indices a and b. The Schwartz topology on $\mathcal{S}(\mathbf{R}^d)$ is the smallest topology making each semi-norm $\|\cdot\|_{a,b}$ continuous. This makes $\mathcal{S}(\mathbf{R}^d)$ a topological vector space.

The argument for the proof of the preceding theorem goes through with minor alterations and shows that

Theorem 4.2. *The topology on $\mathcal{S}(\mathbf{R}^d)$ generated by the family of semi-norms $\|\cdot\|_{a,b}$ for all $a, b \in \{0, 1, 2, ...\}^d$, is complete.*

5. Hermite Polynomials, Creation and Annihilation Operators

We shall summarize the definition and basic properties of Hermite polynomials (our approach is essentially that of Hermite's original [1]).

A central role is played by the *Gaussian kernel*

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \tag{5.1}$$

Properties of translates of p are obtained from

$$e^{xy - \frac{y^2}{2}} = \frac{p(x-y)}{p(x)} \tag{5.2}$$

Expanding the *right side* in a Taylor series we have
(5.3)

$$\boxed{e^{xy - \frac{y^2}{2}} = \frac{p(x-y)}{p(x)} = \sum_{n=0}^{\infty} \frac{1}{n!} H_n(x) y^n,}$$

where the Taylor coefficients, denoted $H_n(x)$, are

$$\boxed{H_n(x) = \frac{1}{p(x)} \left(-\frac{d}{dx}\right)^n p(x)} \tag{5.4}$$

This is the n–th *Hermite polynomial* and is indeed an n–th degree polynomial in which x^n has coefficient 1, facts which may be checked by induction.

Observe the following

$$\begin{aligned}
\int_{\mathbf{R}} \frac{p(x-y)}{p(x)} \frac{p(x-z)}{p(x)} p(x) dx &= e^{-\frac{y^2+z^2}{2}} \int_{\mathbf{R}} e^{-x(y+z)} p(x)\, dx \\
&= e^{-\frac{y^2+z^2}{2} + \frac{(y+z)^2}{2}} \\
&= e^{yz}
\end{aligned}$$

Going over to the Taylor series and assuming integrals and series can be interchanged at will, we have

$$\sum_{n,m\geq 0}\langle H_n, H_m\rangle_{L^2(p(x)dx)}\frac{y^n z^m}{n!m!} = \sum_{n=0}^{\infty}\frac{y^n z^n}{n!}$$

Thus

$$\boxed{\langle H_n, H_m\rangle_{L^2(p(x)dx)} = n!\delta_{nm}} \tag{5.5}$$

Thus an orthonormal set of functions is given by

$$\boxed{h_n(x) = \tfrac{1}{\sqrt{n!}}H_n(x)} \tag{5.6}$$

Since these are orthogonal polynomials, the n–th one being exactly of degree n, their span contains all polynomials. It can be shown that the span is in fact dense in $L^2(p(x)dx)$. Thus the polynomials above constitute an orthonormal basis of $L^2(p(x)dx)$.

Next, consider the derivative of H_n:

$$\begin{aligned} H_n'(x) &= (-1)^n p(x)^{-1}p^{(n+1)}(x) - (-1)^n p(x)^{-1}p'(x)p(x)^{-1}p^{(n+1)}(x) \\ &= -H_{n+1}(x) + xH_n(x) \end{aligned}$$

So

$$\boxed{\left(-\tfrac{d}{dx}+x\right)h_n(x) = \sqrt{n+1}\,h_{n+1}(x)} \tag{5.7}$$

The operator

$$\left(-\frac{d}{dx}+x\right)$$

is the *creation operator* in $L^2\big(\mathbf{R}; p(x)dx\big)$.

More officially, we can take the creation operator to have domain consisting of all functions f which can be expanded in $L^2(p(x)dx)$ as $\sum_{n\geq 0} a_n h_n$, with each a_n a complex number, and satisfying the condition $\sum_{n\geq 0}(n+1)|a_n|^2 < \infty$; the action of the operator on f yields the function $\sum_{n\geq 0}\sqrt{n+1}\,a_n h_{n+1}$. This makes the creation operator unitarily equivalent to a multiplication operator (in the sense discussed below in subsection 12.5) and hence a *closed* operator (see 12.1 for definition). For the type of smooth functions f we will mostly work with, the effect of the operator on f will in fact be given by application of $\left(-\frac{d}{dx}+x\right)$ to f.

Next, differentiating the fundamental generating relation (5.3) with respect to x we have

$$\begin{aligned} \sum_{n\geq 0}\frac{1}{n!}H_n'(x)y^n &= ye^{xy-y^2/2} \\ &= \sum_{n\geq 0}\frac{1}{n!}y^{n+1}H_n(x) \\ &= \sum_{n\geq 1}\frac{1}{(n-1)!}H_{n-1}(x)y^n \end{aligned}$$

From this we see that

$$H_n'(x) = nH_{n-1}(x) \tag{5.8}$$

where $H_{-1} = 0$. Thus:

$$\frac{d}{dx}h_n(x) = \sqrt{n}\, h_{n-1}(x) \tag{5.9}$$

The operator

$$\frac{d}{dx}$$

is the *annihilation operator* in $L^2\big(\mathbf{R}; p(x)dx\big)$. As with the creation operator, we may define it in a more specific way, as a closed operator.

6. Hermite Functions, Creation and Annihilation Operators

In the preceding section we studied Hermite polynomials in the setting of the Gaussian space $L^2\big(\mathbf{R}; p(x)dx\big)$. Let us translate the concepts are results back to the usual space $L^2\big(\mathbf{R}; dx\big)$.

To this end, consider the unitary isomorphism:

$$\mathcal{U} : L^2(\mathbf{R}, p(x)dx) \to L^2(\mathbf{R}, dx) : f \mapsto \sqrt{p}f \tag{6.1}$$

Then the orthonormal basis polynomials h_n go over to the functions ϕ_n given by

$$\phi_n(x) = (-1)^n \tfrac{1}{\sqrt{n!}}(2\pi)^{-1/4} e^{x^2/4} \tfrac{d^n e^{-x^2/2}}{dx^n} \tag{6.2}$$

The family $\{\phi_m\}_{n\geq 0}$ forms an orthonormal basis for $L^2(\mathbf{R}, dx)$.

We will now determine the annihilation and creation operators on $L^2(\mathbf{R}, dx)$. If $f \in L^2(\mathbf{R}, dx)$ is differentiable and has derivative f' also in $L^2(\mathbf{R}, dx)$, we have:

$$\begin{aligned}
\left(\mathcal{U}\frac{d}{dx}\mathcal{U}^{-1}\right) f(x) &= \sqrt{p(x)}\,\frac{d}{dx}\left[p(x)^{-1/2}f(x)\right] \\
&= f'(x) + p(x)^{1/2}\big(-1/2\big)p(x)^{-3/2}p'(x)f(x) \\
&= f'(x) + \frac{1}{2}xf(x)
\end{aligned}$$

So, on $L^2(\mathbf{R}, dx)$, the annihilator operator is

$$A = \tfrac{d}{dx} + \tfrac{1}{2}x \tag{6.3}$$

which will satisfy

$$A\phi_n = \sqrt{n}\,\phi_{n-1} \tag{6.4}$$

where $\phi_{-1} = 0$. For the moment, we proceed by taking the domain of A to be the Schwartz space $\mathcal{S}(\mathbf{R})$.

Next,

$$\begin{aligned}
\left(\mathcal{U}\left(-\frac{d}{dx} + x\right)\mathcal{U}^{-1}\right) f(x) &= -f'(x) + xf(x) - \frac{1}{2}xf(x) \\
&= \left(-\frac{d}{dx} + \frac{1}{2}x\right) f(x)
\end{aligned}$$

Thus the *creation operator* is

$$C = A^* = -\tfrac{d}{dx} + \tfrac{1}{2}x \tag{6.5}$$

The reason we have written A^* is that, as is readily checked, we have the adjoint relation

$$\langle Af, g\rangle = \left\langle f, \left(-\frac{d}{dx} + \frac{1}{2}x\right) g\right\rangle \tag{6.6}$$

with the inner-product being the usual one on $L^2(\mathbf{R}, dx)$. Again, for the moment, we take the domain of C to be the Schwartz space $\mathcal{S}(\mathbf{R})$ (though, technically, in that case we should not write C as A^*, since the latter, if viewed as the L^2–adjoint operator, has a larger domain).

For this we have

$$C\phi_n = \sqrt{n+1}\,\phi_{n+1} \tag{6.7}$$

Observe also that

$$AC = \frac{1}{4}x^2 - \frac{d^2}{dx^2} + \frac{1}{2}I \quad \text{and} \quad CA = \frac{1}{4}x^2 - \frac{d^2}{dx^2} - \frac{1}{2}I \tag{6.8}$$

which imply:

$$[A, C] = AC - CA = I, \quad \text{the identity} \tag{6.9}$$

Next observe that

$$CA\phi_n = \sqrt{n}\sqrt{n}\phi_n = n\phi_n \tag{6.10}$$

and so CA is called the *number operator* N:

$$N = A^*A = CA = -\tfrac{d^2}{dx^2} + \tfrac{x^2}{4} - \tfrac{1}{2} \qquad \text{the } \textit{number operator} \tag{6.11}$$

As noted above in (6.10), the number operator N has the eigenfunctions ϕ_n:

$$N\phi_n = n\phi_n \tag{6.12}$$

Integration by parts shows that

$$\langle f, g'\rangle = -\langle f', g\rangle$$

for every $f, g \in \mathcal{S}(\mathbf{R})$, and so also

$$\langle f, g''\rangle = \langle f'', g\rangle$$

It follows that the operator N satisfies

$$\langle Nf, g\rangle = \langle f, Ng\rangle \tag{6.13}$$

for every $f, g \in \mathcal{S}(\mathbf{R})$.

Now consider the case of $\mathbf{R}^d$. Then for each $j \in \{1, ..., d\}$, there are creation, annihilation, and number operators:

$$A_j = \tfrac{\partial}{\partial x_j} + \tfrac{1}{2}x_j, \quad C_j = -\tfrac{\partial}{\partial x_j} + \tfrac{1}{2}x_j, \quad N_j = C_jA_j \tag{6.14}$$

These map $\mathcal{S}(\mathbf{R}^d)$ into itself and satisfy the commutation relations

$$[A_j, C_k] = \delta_{jk} I, \quad [N_j, A_k] = -\delta_{jk} A_j, \quad [N_j, C_k] = \delta_{jk} C_j \tag{6.15}$$

Now let us be more specific about the precise definition of the creation and annihilation operators. Given its effect on the orthonormal basis $\{\phi_m\}_{m \in W^d}$, the operator C_k has the form:

$$\phi_m \mapsto \sqrt{m_k + 1}\phi_{m'},$$

where $m'_i = m_i$ for all $i \in \{1, ..., d\}$ except when $i = k$, in which case $m'_k = m_k + 1$. The domain of C_k is the set $\mathcal{D}(C_k)$ given by

$$\mathcal{D}(C_k) = \left\{ f \in L^2(\mathbf{R}) \;\middle|\; \sum_{m \in W^d} (m_k + 1)|a_m|^2 < \infty \text{ where } a_m = \langle f, \phi_m \rangle \right\}$$

for The operator C_k is then officially defined by specifying its action on a typical element of its domain:

$$C_k\Big(\sum_{m \in W^d} a_m \phi_m \Big) = \sum_{m \in W^d} a_m \sqrt{m_k + 1}\phi_{m'}, \tag{6.16}$$

where m' is as before. The operator C_k is essentially the composite of a multiplication operator and a bounded linear map taking $\phi_m \to \phi_{m'}$ where m' is as defined above. (See subsection 12.5 for precise formulation of a multiplication operator.) Noting this it can be readily checked that C_k is a closed operator using the following argument: Let T be a bounded linear operator and M_h a multiplication operator (any closed operator will do); we show that the composite $M_h T$ is a closed operator. Suppose $x_n \to x$. Since T is a bounded linear operator, $Tx_n \to Tx$. Now suppose also that $M_h(Tx_n) \to y$. Since M_h is closed, it follows then that $Tx \in \mathcal{D}(M_h)$ and $y = M_h T x$.

The operators A_k and N_k are defined analogously.

Proposition 6.1. *Let $\mathcal{L}_0$ be the vector subspace of $L^2(\mathbf{R}^d)$ spanned by the basis vectors $\{\phi_m\}_{m \in W^d}$. Then for $k \in \{1, 2, \ldots, d\}$, $C_k|\mathcal{L}_0$ and $A_k|\mathcal{L}_0$ have closures given by C_k and A_k, respectively* (see subsection 12.4 for the notion of closure).

Proof. We need to show that the graph of C_k, $Gr(C_k)$, is equal to the closure of the graph of $C_k|\mathcal{L}_0$, $\overline{Gr(C_k|\mathcal{L}_0)}$ (refer to subsection 12.1 for the notion of graph). It is clear that $Gr(C_k|\mathcal{L}_0) \subseteq Gr(C_k)$. Using this and the fact that C_k is a closed operator we have that

$$\overline{Gr(C_k|\mathcal{L}_0)} \subseteq \overline{Gr(C_k)} = Gr(C_k)$$

Going in the other direction, take $(f, C_k f) \in Gr(C_k)$. Now $f = \sum_{m \in W^d} a_m \phi_m$ where $a_m = \langle f, \phi_m \rangle$. Let f_N be given by

$$f_N = \sum_{m \in W_N^d} a_m \phi_m \text{ where } W_N^d = \{\, m \in W^d \mid 0 \le m_1 \le N, \ldots, 0 \le m_d \le N\}.$$

Observe $f_N \in \mathcal{L}_0$. Moreover

$$\lim_{N \to \infty} f_N = f \text{ and } \lim_{N \to \infty} C_k f_N = \lim_{N \to \infty} \sum_{m \in W_N^d} (m_k + 1) a_m \phi_m = C_k f$$

in $L^2(\mathbf{R}^d)$. Thus $(f, C_k f) \in \overline{Gr(C_k|\mathcal{L}_0)}$ and we have that $Gr(C_k) \subseteq \overline{Gr(C_k|\mathcal{L}_0)}$.

The proof for A_k follows similarly. □

Linking this new definition for C_k with our earlier formulas (6.14) we have:

Proposition 6.2. *If* $f \in \mathcal{S}(\mathbf{R}^d)$ *then*

$$C_k f = -\frac{\partial f}{\partial x_k} + \frac{x_k}{2} f, \qquad \textit{and} \qquad A_k f = \frac{\partial f}{\partial x_k} + \frac{x_k}{2} f$$

Proof. Let $g = -\frac{\partial f}{\partial x_k} + \frac{x_k}{2} f$. Since $f \in \mathcal{S}(\mathbf{R}^d)$, we have that $g \in L^2(\mathbf{R}^d)$. So we can write g as $g = \sum_{j \in W^d} a_j \phi_j$ where $a_j = \langle g, \phi_j \rangle$. Let us examine these a_j's more closely. Observe

$$\begin{aligned} a_j = \langle g, \phi_j \rangle &= \left\langle -\frac{\partial f}{\partial x_k} + \frac{x_k}{2} f, \phi_j \right\rangle \\ &= \left\langle f, \frac{\partial \phi_j}{\partial x_k} + \frac{x_k}{2} \phi_j \right\rangle \\ &= \langle f, \sqrt{j_k} \phi_{j'} \rangle \end{aligned}$$

where $j'_i = j_i$ for all $i \in \{1, ..., d\}$ except when $i = k$, in which case $j'_k = j_k - 1$.

Bringing this information back to our expression for g we see that

$$\begin{aligned} g &= \sum_{j \in W^d} \sqrt{j_k} \langle f, \phi_{j'} \rangle \phi_j \\ &= \sum_{m \in W^d} \sqrt{m_k + 1} \langle f, \phi_m \rangle \phi_{m'} && \text{where } m' \text{ is as defined above} \\ &= C_k f && \text{by (6.16)} \end{aligned}$$

The second equality is obtained by letting $m = j'$ and noting that $\phi_{j'} = 0$ when j'_k is -1. The proof follows similarly for A_k. □

7. Properties of $\mathcal{S}_p(\mathbf{R})$ functions

Our aim here is to demonstrate that functions in $\mathcal{S}_p(\mathbf{R})$ are p-times differentiable. The main tool we will use is the Fourier transform:

$$\hat{f}(p) = \mathcal{F}f(p) = (2\pi)^{-1/2} \int_{\mathbf{R}} e^{-ipx} f(x)\, dx \tag{7.1}$$

This is meaningful whenever f is in $L^1(\mathbf{R})$, but we will work mainly with f in $\mathcal{S}(\mathbf{R})$. We will use the following facts:

- $\mathcal{F}$ maps $\mathcal{S}(\mathbf{R})$ *onto* itself and satisfies the Plancherel identity:

$$\int_{\mathbf{R}} |f(x)|^2\, dx = \int_{\mathbf{R}} |\hat{f}(p)|^2\, dp \tag{7.2}$$

- for any $f \in \mathcal{S}(\mathbf{R})$,

$$f(x) = (2\pi)^{-1/2} \int_{\mathbf{R}} e^{ipx} \hat{f}(p)\, dp \tag{7.3}$$

- if $f \in \mathcal{S}(\mathbf{R})$ then

$$p\hat{f}(p) = -i\mathcal{F}(f')(p) \tag{7.4}$$

Consequently, we have

$$
\begin{aligned}
\|f\|_{\sup} &\le (2\pi)^{-1/2}\int_{\mathbf{R}} |\hat{f}(p)|\,dp \\
&= (2\pi)^{-1/2}\int_{\mathbf{R}} (1+p^2)^{1/2}|\hat{f}(p)|(1+p^2)^{-1/2}\,dp \\
&= (2\pi)^{-1/2}\left[\int_{\mathbf{R}} (1+p^2)|\hat{f}(p)|^2\,dp\right]^{1/2}\pi^{1/2} \quad \text{by Cauchy-Schwartz} \\
&\le 2^{-1/2}\left[\|\hat{f}\|_{L^2} + \|p\hat{f}\|_{L^2(\mathbf{R},dp)}\right] \\
&\le 2^{-1/2}\left[\|f\|_{L^2} + \|f'\|_{L^2}\right] \quad \text{by Plancherel and (7.4)}
\end{aligned} \tag{7.5}
$$

Recall that for f in $\mathcal{S}_p(\mathbf{R}) \subset L^2(\mathbf{R})$ we have that

$$f = \sum_{n\ge 0} a_n\phi_n$$

where $a_n = \langle f, \phi_n\rangle$ for every $n \ge 0$. Let

$$f_N = \sum_{n=0}^{N} a_n\phi_n$$

and observe the following:

Lemma 7.1. *If $f \in \mathcal{S}_p(\mathbf{R})$ for $p \ge 1$, then $\{f_N'\}$ is Cauchy in $L^2(\mathbf{R})$.*

Proof. Note that

$$f_N' = \left(\frac{A-C}{2}\right) f_N$$

So $\|f_N' - f_M'\|_{L^2} \le \frac{1}{2}\|Af_N - Af_M\|_{L^2} + \frac{1}{2}\|Cf_N - Cf_M\|_{L^2}$.

Now for $M < N$ we have

$$
\begin{aligned}
\|Af_N - Af_M\|_{L^2}^2 &= \|\sum_{n=M+1}^{N} a_n\sqrt{n}\phi_{n-1}\|_{L^2}^2 \\
&= \sum_{n=M+1}^{N} |a_n|^2 n \\
&\le \sum_{n=M+1}^{N} |a_n|^2(n+1)^2
\end{aligned}
$$

Likewise,

$$
\begin{aligned}
\|Cf_N - Cf_M\|_{L^2}^2 &= \|\sum_{n=M+1}^{N} a_n\sqrt{n+1}\phi_{n+1}\|_{L^2}^2 \\
&= \sum_{n=M+1}^{N} |a_n|^2(n+1) \\
&\le \sum_{n=M+1}^{N} |a_n|^2(n+1)^2
\end{aligned}
$$

Since $f \in \mathcal{S}_p(\mathbf{R})$, we know $\sum_{M+1}^{N} |a_n|^2(n+1)^2$ tends to 0 as M goes to infinity. Thus $\{f_N'\}$ is Cauchy in $L^2(\mathbf{R})$.

□

Lemma 7.2. *If $f \in \mathcal{S}_p(\mathbf{R})$ for $p \geq 1$, then $\{f_N\}$ converges uniformly to f, i.e. $\|f - f_N\|_{\sup} \to 0$ as $N \to \infty$.*

Proof. It is enough to show that $\|f_M - f_N\|_{\sup} \to 0$ as $M, N \to \infty$. Note that

$$\|f_M - f_N\|_{\sup} \leq 2^{-1/2}\left[\|f_N - f_M\|_{L^2} + \|f_N' - f_M'\|_{L^2}\right]$$

by (7.5). Since $f \in L^2(\mathbf{R})$ we have $\|f_N - f_M\|_{L^2} \to 0$ as $M, N \to \infty$ and by Lemma 7.1 we have that $\|f_N' - f_M'\|_{L^2} \to 0$ as $M, N \to \infty$. Therefore $\{f_N\}$ converges uniformly to f.

□

The Lemma above immediately gives us that $f \in \mathcal{S}_p(\mathbf{R})$ is continuous, being the uniform limit of the continuous functions f_N. Next we address the differentiablity of functions in $\mathcal{S}_p(\mathbf{R})$. We show that functions in $\mathcal{S}_p(\mathbf{R})$ are differentiable and their derivatives lie in $\mathcal{S}_{p-1}(\mathbf{R})$.

Theorem 7.3. *If $f \in \mathcal{S}_p(\mathbf{R})$ for $p \geq 1$, then f is differentiable and $f' = \lim_{n\to\infty} f_N'$ in $L^2(\mathbf{R})$. Moveover f' is in $\mathcal{S}_{p-1}(\mathbf{R})$.*

Proof. By Lemma 7.1, the sequence of derivatives f_N' is Cauchy in $L^2(\mathbf{R})$. Let $g = \lim_{N\to\infty} f_N'$ in $L^2(\mathbf{R})$. Observe that

$$f_N(y) = f_N(x) + \int_0^1 f_N'\big((x + t(y - x)\big)(y - x)\,dt \tag{7.6}$$

Now $\int_0^1 |f_N'(x+t(y-x))(y-x) - g(x+t(y-x))(y-x)|\,dt \leq \sqrt{|y-x|}\|f_N' - g\|_{L^2}$ by the Cauchy-Schwartz inequality. Since $\|f_N' - g\|_{L^2} \to 0$ as $N \to \infty$, we have

$$\int_0^1 f_N'(x + t(y - x))(y - x)\,dt \to \int_0^1 g(x + t(y - x))(y - x)\,dt$$

Since f_N converges to f uniformly by Lemma 7.2, taking the limit as $N \to \infty$ in (7.6) we obtain

$$f(y) = f(x) + \int_0^1 g\big((x + t(y - x)\big)(y - x)\,dt$$

Therefore $f' = g \in L^2(\mathbf{R})$.

Now we have the L^2 limits:

$$f' = \lim_{N\to\infty} f_N' = \lim_{N\to\infty}\left(\frac{A - C}{2}\right) f_N = \lim_{N\to\infty} \frac{1}{2}\sum_{n=0}^{N}(\sqrt{n+1}a_{n+1} - \sqrt{n}a_{n-1})\phi_n$$

Observe that

$$\begin{aligned}
&\frac{1}{2}\sum_{n\geq 0}(n+1)^{2(p-1)}|\sqrt{n+1}a_{n+1} - \sqrt{n}a_{n-1}|^2\\
&\quad\leq \sum_{n\geq 0}(n+1)^{2(p-1)}(n+1)|a_{n+1}|^2 + \sum_{n\geq 1}(n+1)^{2(p-1)}n|a_{n-1}|^2\\
&\qquad\qquad\text{since } |a-b|^2 \leq 2(a^2+b^2) \text{ for } a, b \in \mathbb{R}\\
&\quad\leq \sum_{n\geq 0} n^{(2p-1)}|a_n|^2 + \sum_{n\geq 0}(n+2)^{2(p-1)}(n+1)|a_n|^2
\end{aligned}$$

This sum is finite since $f \in \mathcal{S}_p(\mathbf{R})$ and since, for n large enough, $n^{(2p-1)} \leq (n+1)^{2p}$ and $(n+2)^{2(p-1)}(n+1) \leq (n+1)^{2p}$. Thus we have f' is in $\mathcal{S}_{p-1}(\mathbf{R})$. □

Corollary 7.4. *If $f \in \mathcal{S}_p(\mathbf{R})$, then $f^{(k)}$ exists for $k \in \{0, 1, 2, \ldots, p\}$ and $f^{(k)} \in \mathcal{S}_{p-k}(\mathbf{R})$.*

Proof. Recursively apply Theorem 7.3 to f and it's derivatives. □

8. Inner-products on $\mathcal{S}(\mathbf{R})$ from N

For $f \in L^2(\mathbf{R})$, define

$$\|f\|_t = \left\{ \sum_{n\geq 0} (n+1)^t |\langle f, \phi_n\rangle|^2 \right\}^{1/2} \tag{8.1}$$

for every $t > 0$. More generally, define

$$\langle f, g\rangle_t = \sum_{n\geq 0} (n+1)^t \langle f, \phi_n\rangle_{L^2} \langle \phi_n, g\rangle_{L^2}, \tag{8.2}$$

for all f, g in the subspace of $L^2(\mathbf{R})$ consisting of functions F for which $\|F\|_t < \infty$.

Theorem 8.1. *Let $f \in \mathcal{S}(\mathbf{R})$. Then every $t > 0$ we have $\|f\|_t < \infty$. Moreover, for every integer $m \geq 0$, we also have*

$$N^m f = \sum_{n\geq 0} n^m \langle f, \phi_n\rangle \phi_n, \tag{8.3}$$

where on the left N^m is the differential operator $-\frac{d^2}{dx^2} + \frac{x^2}{4} - \frac{1}{2}$ applied n times, and on the right the series is taken in the sense of $L^2(\mathbf{R}, dx)$. Furthermore,

$$\|f\|_m^2 = \langle f, (N+1)^m f\rangle \tag{8.4}$$

This result will be strengthened and a converse proved later.

Proof. Let $m \geq 0$ be an integer. Since $f \in \mathcal{S}(\mathbf{R})$, it is readily seen that Nf is also in $\mathcal{S}(\mathbf{R})$, and thus, inductively, so is $N^m f$. Then we have

$$\begin{aligned}
\langle f, N^m f\rangle &= \sum_{n\geq 0} \langle f, \phi_n\rangle \langle \phi_n, N^m f\rangle \\
&= \sum_{n\geq 0} \langle f, \phi_n\rangle \langle N^m \phi_n, f\rangle \quad \text{by (6.13)} \\
&= \sum_{n\geq 0} \langle f, \phi_n\rangle \, \langle n^m \phi_n, f\rangle \\
&= \sum_{n\geq 0} n^m |\langle f, \phi_n\rangle|^2
\end{aligned}$$

Thus we have proven the relation

$$\langle f, N^m f\rangle = \sum_{n\geq 0} n^m |\langle f, \phi_n\rangle|^2 \tag{8.5}$$

An exactly similar argument shows

$$\langle f, (N+1)^m f\rangle = \sum_{n\geq 0} (n+1)^m |\langle f, \phi_n\rangle|^2 = \|f\|_m^2 \tag{8.6}$$

So if $t > 0$, choosing any integer $m \geq t$ we have

$$\|f\|_t^2 \leq \|f\|_m^2 = \langle f, (N+1)^m f\rangle < \infty$$

Observe that the series

$$\sum_{n\geq 0} n^m \langle f, \phi_n\rangle \phi_n \tag{8.7}$$

is convergent in $L^2(\mathbf{R}, dx)$ since

$$\sum_{n\geq 0} n^{2m} |\langle f, \phi_n\rangle|^2 = \langle N^{2m} f, f\rangle < \infty.$$

So for any $g \in L^2(\mathbf{R}, dx)$ we have, by an argument similar to the calculations done above:

$$\begin{aligned} \langle N^m f, g\rangle &= \sum_{n\geq 0} n^m \langle f, \phi_n\rangle \langle \phi_n, g\rangle \\ &= \sum_{n\geq 0} \langle n^m \langle f, \phi_n\rangle \phi_n, g\rangle \\ &= \left\langle \sum_{n\geq 0} n^m \langle f, \phi_n\rangle \phi_n, g \right\rangle \end{aligned}$$

This proves the statement about $N^m f$.

□

We have similar observations concerning $C^m f$ and $A^m f$. First observe that since C and A are operators involving d/dx and x, they map $\mathcal{S}(\mathbf{R})$ into itself. Also,

$$\langle Af, g\rangle = \langle f, Cg\rangle,$$

for all $f, g \in \mathcal{S}(\mathbf{R})$, as already noted. Using this we have for $f \in \mathcal{S}(\mathbf{R})$, we have

$$\begin{aligned} \langle \phi_{n+m}, C^m f\rangle &= \langle A^m \phi_{n+m}, f\rangle \\ &= \sqrt{(n+m)(n+m-1)\cdots(n+1)}\, \langle \phi_n, f\rangle. \end{aligned}$$

Therefore,

$$C^m f = \sum_{n\geq 0} \left[\frac{(n+m)!}{n!}\right]^{1/2} \langle f, \phi_n\rangle \phi_{n+m} \tag{8.8}$$

Similarly,

$$A^m f = \sum_{n\geq 0} \left[\frac{n!}{(n-m)!}\right]^{1/2} \langle f, \phi_n\rangle \phi_{n-m} \tag{8.9}$$

More generally, if $B_1, ..., B_k$ are such that each B_i is either A or C then

$$B_1 ... B_k f = \sum_{n\geq 0} \theta_{n,k} \langle f, \phi_n\rangle \phi_{n+r}, \tag{8.10}$$

where the integer r is the excess number of C's over the A's in the sequence $B_1, ..., B_k$, and $\theta_{n,k}$ is a real number determined by n and k. We do have the upper bound

$$\theta_{n,k}^2 \leq (n+k)^k \leq [(n+1)k]^k = (n+1)^k k^k \tag{8.11}$$

Note also that

$$\|B_1...B_kf\|^2 = \langle (B_1...B_k)^* B_1...B_k f, f\rangle = \sum_{n\geq 0} \theta_{n,k}^2 |\langle f, \phi_n\rangle|^2 \tag{8.12}$$

Actually, it seems that $(B_1...B_k)^*B_1...B_k$ is a polynomial in N of degree k, and so $\theta_{n,k}^2$ would be a polynomial in n of degree k.

Lets look at the case of $\mathbf{R}^d$. The functions ϕ_n generate an orthonormal basis by tensor products. In more detail, if $a \in W^d$ is a multi-index, define $\phi_a \in L^2(\mathbf{R}^d)$ by

$$\phi_a(x) = \phi_{a_1}(x_1)...\phi_{a_d}(x_d)$$

Now, for each $t > 0$, and $f \in L^2(\mathbf{R}^d)$ define

$$\|f\|_t \stackrel{\text{def}}{=} \left\{ \sum_{a\in W^d} [(a_1+1)...(a_d+1)]^t |\langle f, \phi_a\rangle|^2 \right\}^{1/2}, \tag{8.13}$$

and then define

$$\langle f, g\rangle_t = \sum_{a\in W^d} [(a_1+1)...(a_d+1)]^t \langle f, \phi_a\rangle_{L^2(\mathbf{R}^d)} \langle \phi_a, g\rangle_{L^2(\mathbf{R}^d)}, \tag{8.14}$$

for all f, g in the subspace of $L^2(\mathbf{R}^d)$ consisting of functions F for which $\|F\|_t < \infty$.

Let T_d be the operator on $\mathcal{S}(\mathbf{R}^d)$ given by

$$T_d = (N_d + 1)...(N_1 + 1)$$

Then, for every non-negative integer m, we have

$$\|f\|_m^2 = \langle f, T_d^m f\rangle$$

The other results of this section also extend in a natural way to $\mathbf{R}^d$.

9. L^2–type norms on $\mathcal{S}(\mathbf{R})$

For integers $a, b \geq 0$, and $f \in \mathcal{S}(\mathbf{R})$, define

$$\|f\|_{a,b,2} = \|x^a D^b f(x)\|_{L^2(\mathbf{R},dx)} \tag{9.1}$$

Recall the operators

$$A = \frac{d}{dx} + \frac{1}{2}x, \quad C = -\frac{d}{dx} + \frac{1}{2}x, \quad N = CA$$

and the norms

$$\|f\|_m = \langle f, (N+1)^m f\rangle$$

Our objective for this section is summarized as:

Theorem 9.1. *The system of semi-norms given by $\|f\|_{a,b,2}$ and the system given by the norms $\|f\|_m$ generate the same topology on $\mathcal{S}(\mathbf{R})$.*

Proof. Let a, b be non-negative integers. Then

$$\begin{aligned} \|f\|_{a,b,2} &= \|(A+C)^a 2^{-b}(A-C)^b f\|_{L^2} \\ &\leq \text{a linear combination of terms} \\ &\quad\ \text{of the form } \|B_1...B_kf\|_{L^2}, \end{aligned}$$

where each B_i is either A or C, and $k = a + b$. Writing $c_n = \langle f, \phi_n\rangle$, we have

$$\|B_1...B_kf\|_{L^2}^2 = \|\sum_{n\geq 0} c_n \theta_{n,k} \phi_{n+r}\|$$

$$= \sum_{n\geq 0} |c_n|^2 \theta_{n,k}^2,$$

where

$$r = \#\{j : B_j = C\} - \#\{j : B_j = A\},$$

and, as noted earlier in (8.11),

$$\theta_{n,k}^2 \leq (n+k)^k \leq [(n+1)k]^k = (n+1)^k k^k$$

So

$$\|B_1...B_k f\|_{L^2}^2 \leq \sum_{n\geq 0} |c_n|^2 (n+1)^k k^k = k^k \|f\|_k^2$$

Thus $\|f\|_{a,b,2}$ is bounded above by a multiple of the norm $\|f\|_{a+b}$.

It follows, that the topology generated by the semi-norms $\|\cdot\|_{a,b,2}$ is contained in the topology generated by the norms $\|\cdot\|_k$.

Now we show the converse inclusion. From

$$\|f\|_k^2 = \langle f, (N+1)^k f\rangle \leq \|f\|_{L^2}^2 + \|(N+1)^k f\|_{L^2}^2$$

and the expression of N as a differential operator we see that $\|f\|_k^2$ is bounded above by a linear combination of $\|f\|_{a,b,2}^2$ for appropriate a and b. It follows then that the topology generated by the norms $\|\cdot\|_k$ is contained in the topology generated by the semi-norms $\|\cdot\|_{a,b,2}$.

□

Now consider $\mathbf{R}^d$. Let $a, b \in W^d$ be multi-indices, where $W = \{0, 1, 2, ...\}$. Then for $f \in \mathcal{S}(\mathbf{R}^d)$ define

$$\|f\|_{a,b,2} = \left\{\int_{\mathbf{R}^d} |x^a D^b f(x)|^2\, dx\right\}^{1/2}$$

These specify semi-norms and they generate the same topology as the one generated by the norms $\|\cdot\|_m$, with $m \in W$. The argument is a straightforward modification of the one used above.

10. Equivalence of the three topologies

We will demonstrate that the topology generated by the family of norms $\|\cdot\|_k$, or, equivalently, by the semi-norms $\|\cdot\|_{a,b,2}$, is the same as the Schwartz topology on $\mathcal{S}(\mathbf{R})$.

Recall from (7.5) we have that for $f \in \mathcal{S}(\mathbf{R})$

$$\|f\|_{\sup} \leq 2^{-1/2}\left[\|f\|_{L^2} + \|f'\|_{L^2}\right]$$

Putting in $x^a D^b f(x)$ in place of $f(x)$ we then have

$$\|f\|_{a,b} \leq \text{a linear combination of } \|f\|_{a,b,2},\ \|f\|_{a-1,b,2}, \text{ and } \|f\|_{a,b+1,2} \tag{10.1}$$

Next we bound the semi-norms $\|f\|_{a,b,2}$ by the semi-norms $\|f\|_{a,b}$. To this end, observe first

$$\begin{aligned} \|f\|_{L^2}^2 &= \int_{\mathbf{R}} (1+x^2)^{-1}(1+x^2)|f(x)|^2\, dx \\ &\leq \|(1+x^2)|f(x)|^2\|_{\sup}\pi \\ &\leq \pi\left(\|f\|_{\sup}^2 + \|xf(x)\|_{\sup}^2\right) \\ &\leq \pi\left(\|f\|_{\sup} + \|xf(x)\|_{\sup}\right)^2 \end{aligned}$$

So for any integers $a, b \geq 0$, we have

$$\|f\|_{a,b,2} = \|x^a D^b f\|_{L^2} \leq \pi^{1/2}(\|f\|_{a,b} + \|f\|_{a+1,b}) \tag{10.2}$$

Thus, the topology generated by the semi-norms $\|\cdot\|_{a,b,2}$ coincides with the Schwartz topology.

Now lets look at the situation for $\mathbf{R}^d$. The same result holds in this case and the arguments are similar. The appropriate Sobolev inequalities require using $(1+|p|^2)^d$ instead of $1 + p^2$. For $f \in \mathcal{S}(\mathbf{R}^d)$, we have the Fourier transform given by

$$\mathcal{F}(f)(p) = \hat{f}(p) = (2\pi)^{-d/2} \int_{\mathbf{R}^d} e^{-i\langle p,x\rangle} f(x)\, dx \tag{10.3}$$

Again, this preserves the L^2 norm, and transforms derivatives into multiplications:

$$p_j \hat{f}(p) = -i\mathcal{F}\left(\frac{\partial f}{\partial x_j}\right)(p).$$

Repeated application of this shows that

$$|p|^2 \hat{f}(p) = -\mathcal{F}(\Delta f)(p), \tag{10.4}$$

where

$$\Delta = \sum_{j=1}^{n} \frac{\partial^2}{\partial x_j^2}$$

is the Laplacian. Iterating this gives, for each $r \in \{0, 1, 2, ...\}$ and $f \in \mathcal{S}(\mathbf{R}^d)$,

$$|p|^{2r} \hat{f}(p) = (-1)^r \mathcal{F}(\Delta^r f)(p), \tag{10.5}$$

which in turn implies, by the Plancherel formula (7.2), the identity:

$$\int_{\mathbf{R}^d} \left||p|^{2r}|\hat{f}(p)\right|^2 dp = \int_{\mathbf{R}^d} |\Delta^r f(x)|^2\, dx \tag{10.6}$$

Then we have, for any $m > d/4$,

$$\begin{aligned} \|f\|_{\sup} &\leq (2\pi)^{-d/2} \int_{\mathbf{R}^d} |\hat{f}(p)|\, dp \\ &= (2\pi)^{-d/2} \int_{\mathbf{R}^d} (1 + |p|^2)^m |\hat{f}(p)| (1 + |p|^2)^{-m}\, dp \\ &= \left[\int_{\mathbf{R}^d} (1 + |p|^2)^{2m} |\hat{f}(p)|^2\, dp\right]^{1/2} K \quad \text{by Cauchy-Schwartz} \end{aligned}$$

$$\text{where } K = (2\pi)^{-d/2} \left[\int_{\mathbf{R}^d} \frac{dp}{(1+|p|^2)^{2m}}\right]^{1/2} < \infty \tag{10.7}$$

The function $(1 + s)^n/(1 + s^n)$, for $s \geq 0$, attains a maximum value of 2^{n-1}, and so the inequality $(1 + s)^{2m} \leq 2^{2m-1}(1 + s^{2m})$, which leads to

$$(1 + |p|^2)^{2m} \leq 2^{2m-1}(1 + |p|^{4m})$$

Then, from (10.7), we have

$$\|f\|^2_{\sup} \leq K^2 2^{4m-2} \left(\|f\|^2_{L^2} + \|\Delta^m f\|^2_{L^2}\right) \tag{10.8}$$

This last quantity is clearly bounded above by a linear combination of $\|f\|_{0,b,2}$ for certain multi-indices b. Thus $\|f\|_{\sup}$ is bounded above by a linear combination of $\|f\|_{0,b,2}$ for certain multi-indices b. It follows that $\|x^a D^b f\|_{\sup}$ is bounded above by a linear combination of $\|f\|_{a',b',2}$ for certain multi-indices a', b'.

For the inequality going the other way, the reasoning used above for (10.2) generalizes readily, again with $(1 + x^2)$ replaced by $(1 + |x|^2)^d$.

Thus, on $\mathcal{S}(\mathbf{R}^d)$ the topology generated by the family of semi-norms $\|\cdot\|_{a,b,2}$ coincides with the Schwartz topology.

Now we return to (10.8) for some further observations. First note that

$$\Delta = \frac{1}{4}\sum_{j=1}^{d}(C_j - A_j)^2$$

and so Δ^m consists of a sum of multiples of $(3d)^m$ terms each a product of $2m$ elements drawn from the set $\{A_1, C_1, ..., A_d, C_d\}$. Consequently,

$$\|\Delta^m f\|_{L^2}^2 \le c_{d,m}^2 \|f\|_m^2, \tag{10.9}$$

for some positive constant $c_{d,m}$. Combining this with (10.8), we see that for $m > d/4$, there is a constant $k_{d,m}$ such that

$$\|f\|_{\sup} \le k_{d,m}\|f\|_m \tag{10.10}$$

holds for all $f \in \mathcal{S}(\mathbf{R}^d)$.

Now consider $f \in \mathcal{S}_p(\mathbf{R}^d)$, with $p > d/4$. Let

$$f_N = \sum_{j\in W^d, |j|\le N} \langle f, \phi_j\rangle\phi_j$$

Then $f_N \to f$ in L^2 and so a subsequence $\{f_{N_k}\}_{k\ge1}$ converges pointwise almost everywhere to f. It follows then that the essential supremum $\|f\|_\infty$ is bounded above as follows:

$$\|f\|_\infty \le \limsup_{N\to\infty} |f_N|_{\sup}$$

Note that $f_N \to f$ also in the $\|\cdot\|_p$–norm. It follows then from (10.10) that

$$\|f\|_\infty \le k_{d,p}\|f\|_p \tag{10.11}$$

holding for all $f \in \mathcal{S}_p(\mathbf{R}^d)$ with $p > d/4$. Replacing f by the difference $f - f_N$ in (10.11), we see that f is the L^∞–limit of a sequence of continuous functions which, being Cauchy in the sup-norm, has a continuous limit; thus f is a.e. equal to a continuous function, and may thus be redefined to be continuous.

11. Identification of $\mathcal{S}(\mathbf{R})$ with a sequence space

Suppose $a_0, a_1, ...$ are a sequence of complex numbers such that

$$\sum_{n\ge0}(n+1)^m|a_n|^2 < \infty, \qquad \text{for every integer } m \ge 0. \tag{11.1}$$

We will show that the sequence of functions given by

$$s_n = \sum_{j=0}^{n} a_j\phi_j$$

converges in the topology of $\mathcal{S}(\mathbf{R})$ to a function $f \in \mathcal{S}(\mathbf{R})$ for which $a_n = \langle f, \phi_n\rangle$ for every $n \ge 0$.

All the hard work has already been done. From (11.1) we see that $(s_n)_{n\ge0}$ is Cauchy in each norm $\|\cdot\|_m$. So it is Cauchy in the Schwartz topology of $\mathcal{S}(\mathbf{R})$, and hence convergent to some $f \in \mathcal{S}(\mathbf{R})$. In particular, $s_n \to f$ in L^2. Taking inner-products with ϕ_j we see that $a_j = \langle f, \phi_j\rangle$.

Thus we have

Theorem 11.1. *Let* $W = \{0, 1, 2, ...\}$*, and define*

$$F : L^2(\mathbf{R}) \to \mathbf{C}^W$$

by requiring that

$$F(f)_n = \langle f, \phi_n \rangle_{L^2}$$

for all $n \in W$*. Then the image of* $\mathcal{S}(\mathbf{R})$ *under* F *is the set of all* $a \in \mathbf{C}^W$ *for which* $\|a\|_m^2 \stackrel{\text{def}}{=} \sum_{n \geq 0}(n+1)^m |a_n|^2 < \infty$ *for every integer* $m \geq 0$*. Moreover, if* $F\big(\mathcal{S}(\mathbf{R})\big)$ *is equipped with the topology generated by the norms* $\|\cdot\|_m$ *then* F *is a homeomorphism.*

12. Spectral Theory in Brief

Let H be a complex Hilbert space. A *linear operator* on H is a linear map

$$A : D_A \to H,$$

where D_A is a subspace of H. Usually, we work with densely defined operators, i.e. operators A for which D_A is dense.

12.1. **Graph and Closed Operators.** The *graph* of the operator A is

$$Gr(A) = \{(x, Ax) : x \in D_A\} \tag{12.1}$$

Thus $Gr(A)$ is A viewed as a set of ordered pairs, and is thus A itself taken as a mapping in the set-theoretic sense. The operator A is said to be *closed* if its graph is a closed subset of $H \oplus H$; put another way, this means that if $(x_n)_{n \geq 1}$ is any sequence in H which converges to a limit x and if $\lim_{n \to \infty} Ax_n = y$ also exists then x is in the domain of A and $y = Ax$.

12.2. **The adjoint** A^***.** If A is a densely defined operator on H then there is an adjoint operator A^* defined as follows. Let D_{A^*} be the set of all $y \in H$ for which the map

$$f_y : D_A \to \mathbf{C} : x \mapsto \langle Ax, y \rangle$$

is bounded linear. Clearly, D_{A^*} is a subspace of H. The bounded linear functional f_y extends to a bounded linear functional f_y on H. So there exists a vector $z \in H$ such that $f_y(x) = \langle z, x \rangle$ for all $x \in H$. Since D_A is *dense* in H, the element z is uniquely determined by x and A. Denote z by A^*y. Thus, A^*y is the unique vector in H for which

$$\langle x, A^*y \rangle = \langle Ax, y \rangle \tag{12.2}$$

holds for all $x \in D_A$. Using the definition of A^* for a densely-defined operator A it is readily seen that A^* is a closed operator.

12.3. **Self-adjoint Operators.** The operator A is *self-adjoint* if it is densely defined and $A = A^*$. Thus, if A is self-adjoint then $D_A = D_{A^*}$ and

$$\langle x, Ay \rangle = \langle Ax, y \rangle \tag{12.3}$$

for all $x, y \in D_A$. Note that a self-adjoint operator A, being equal to its adjoint A^*, is automatically a closed operator.

12.4. **Closure, and Essentially Self-adjoint Operators.** Consider a densely-defined linear operator S on H. Assume that the closure of the graph of S is the graph of some operator $\overline{S}$. Then $\overline{S}$ is called the *closure* of S. We say that S is *essentially self-adjoint* if its closure is a self-adjoint operator. In particular, S must then be a *symmetric operator*, i.e. it satisfies

$$\langle Sx, y\rangle = \langle x, Sy\rangle \tag{12.4}$$

for all $x, y \in H$. A symmetric operator may not, in general, be essentially self-adjoint.

12.5. **The Multiplication Operator.** Let us turn to a canonical example. Let $(X, \mathcal{F}, \mu)$ be a sigma-finite measure space. Consider the Hilbert space $L^2(\mu)$. Let $f : X \to \mathbf{C}$ be a measurable function. Define the operator M_f on $L^2(\mu)$ by setting

$$M_f g = fg, \tag{12.5}$$

with the domain of M_f given by

$$D(M_f) = \{g \in L^2(\mu) : fg \in L^2(\mu)\} \tag{12.6}$$

Let us check that $D(M_f)$ is dense in $L^2(\mu)$. By sigma-finiteness of μ, there is an increasing sequence of measurable sets X_n such that $\cup_{n\geq 1} X_n = X$ and $\mu(X_n) < \infty$. For any $h \in L^2(\mu)$ let $h_n = 1_{X_n \cap \{|f|\leq n\}} h$. Then

$$\int |fh_n|\, d\mu \leq n \int |h| 1_{X_n}\, d\mu \leq n\mu(X_n)^{1/2} \|h\|_{L^2} < \infty$$

and so $h_n \in D(M_f)$. On the other hand,

$$\|h_n - h\|_{L^2}^2 \to 0$$

by dominated convergence. So $D(M_f)$ is dense in H.

It may be shown that

$$M_f^* = M_{\overline{f}} \tag{12.7}$$

Thus M_f is self-adjoint if f is real-valued.

A very special case of the preceding example is obtained by taking X to be a finite set, say $X = \{1, 2, ..., d\}$, and μ as counting measure on the set of all subsets of X. In this case, $L^2(\mu) = \mathbf{C}^d$, and the operator M_f, viewed as a linear map

$$M_f : \mathbf{C}^d \to \mathbf{C}^d$$

is given by the diagonal matrix

$$\begin{pmatrix} f_1 & 0 & 0 & \cdots & 0 \\ 0 & f_2 & 0 & \cdots & 0 \\ 0 & 0 & f_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & f_d \end{pmatrix} \tag{12.8}$$

Now take the case where μ is counting measure on the sigma-algebra of all subsets of a countable set X. Let f be any real-valued function on X. Let D_f^0 be the subspace of $L^2(\mu)$ consisting of all functions g for which $\{g \neq 0\}$ is a finite set, and let M_f^0 be the restriction of M_f to D_f^0. Then it is readily checked that M_f^0 is essentially self-adjoint. Consequently, the restriction of M_f to any subspace of D_f larger than D_f^0 is also essentially self-adjoint.

12.6. **The Spectral Theorem.** The *spectral theorem* for a self-adjoint operator A on a separable complex Hilbert space H says that there is a sigma-finite measure space $(X, \mathcal{F}, \mu)$, a unitary isomorphism

$$U : H \to L^2(\mu)$$

and a measurable real-valued function f on X such that

$$A = U^{-1} M_f U \tag{12.9}$$

Expressing A in this way is called a *diagonalization* of A (the terminology being motivated by (12.8)).

12.7. **The Functional Calculus.** If g is any measurable function on $\mathbf{R}$ we can then form the operator

$$g(A) \stackrel{\text{def}}{=} U^{-1} M_{g \circ f} U \tag{12.10}$$

If g is a polynomial then $g(A)$ works out to be what it should be, a polynomial in A. Another example, is the function $g(x) = e^{ikx}$, where k is any constant; this gives the operator e^{ikA}.

12.8. **The Spectrum.** The *essential range* of f is the smallest closed subset of $\mathbf{R}$ whose complement U satisfies $\mu\big(f^{-1}(U)\big) = 0$. It consists of all $\lambda \in \mathbf{R}$ for which the operator $M_f - \lambda I = M_{f-\lambda}$ has a bounded inverse (which is $M_{(f-\lambda)^{-1}}$). This essential range forms the *spectrum* $\sigma(A)$ of the operator A. Thus $\sigma(A)$ is the set of all real numbers λ for which the operator $A - \lambda I$ has a bounded linear operator as inverse.

12.9. **The Spectral Measure.** Associate to each Borel set $E \subset \mathbf{R}$ the operator

$$P'_E = M_{1_{f^{-1}(E)}}$$

on $L^2(X, \mu)$. This is readily checked to be an orthogonal *projection operator*. Hence, so is the operator

$$P^A(E) = U^{-1} P'_E U$$

Moreover, it can be checked that the association $E \mapsto P^A(E)$ is a *projection-valued measure*, i.e. $P^A(\emptyset) = 0$, $P^A(\mathbf{R}) = I$, $P^A(E \cap F) = P^A(E)P^A(F)$, and for any disjoint Borel sets $E_1, E_2, ...$ and any vector $x \in H$ we have

$$P^A(\cup_{n\geq 1} E_n)x = \sum_{n\geq 1} P^A(E_n)x \tag{12.11}$$

This is called the *spectral measure* for the operator A, and is uniquely determined by the operator A.

12.10. **The Number Operator.** Let us examine an example. Let $W = \{0, 1, 2, ...\}$, and let μ be counting measure on W. On W we have the function

$$N' : W \to \mathbf{R} : n \mapsto n$$

Correspondingly we have the multiplication operator $M_{N'}$ on the Hilbert space $L^2(W, \mu)$.

Now consider the Hilbert space $L^2(\mathbf{R})$. We have the unitary isomorphism

$$U : L^2(\mathbf{R}) \to L^2(W, \mu) : f \mapsto (\langle f, \phi_n \rangle)_{n\geq 0}$$

Consider the operator N on $L^2(\mathbf{R})$ given by

$$N = U^{-1}M_{N'}U$$

Then

$$Nf = \sum_{n\in W} n\langle f, \phi_n\rangle\phi_n$$

and the domain of N is

$$D_N = \{f \in L^2(\mathbf{R}) : \sum_{n\in W} n^2|\langle f, \phi_n\rangle|^2 < \infty\}$$

Comparing with (8.3) we see that

$$(Nf)(x) = \left(-\frac{d^2}{dx^2} + \frac{x^2}{4} - \frac{1}{2}\right) f(x)$$

for every $f \in \mathcal{S}(\mathbf{R})$.

Thus the self-adjoint operator N extends the differential operator $-\frac{d^2}{dx^2} + \frac{x^2}{4} - \frac{1}{2}$, and, notationally, we will often not make a distinction. In view of the observation made at the end of subsection 12.5, the differential operator $-\frac{d^2}{dx^2} + \frac{x^2}{4} - \frac{1}{2}$ on the domain $\mathcal{S}(\mathbf{R})$ is essentially self-adjoint, with closure equal to the operator N.

The operator U above helps realize the operator N as the multiplication operator $M_{N'}$, and is thus an explicit realization of the fact guaranteed by the spectral theorem.

13. Explanation of physics terminology

In quantum theory, one associates to each physical system a complex Hilbert space $\mathcal{H}$. Each state of the system is represented by a bounded self-adjoint operator $\rho \geq 0$ for which $\mathrm{tr}(\rho) = 1$. An observable is represented by a self-adjoint operator A on $\mathcal{H}$. The relationship of the mathematical formalism with physics is obtained by declaring that

$$\mathrm{tr}(P^A(E)\rho)$$

is the probability that in state ρ the observable A has value in the Borel set $E \subset \mathbf{R}$. Here, P^A is the spectral measure for the self-adjoint operator A.

The states form a convex set, any convex linear combination of any two states being also clearly a state. There are certain states which cannot be expressed as a convex linear combination of distinct states. These are called *pure states.* A pure state is always given by the orthogonal projection onto a ray (1–dimensional subspace of $\mathcal{H}$). If ϕ is any unit vector on such a ray then the orthogonal projection onto the ray is given by: $P_\phi: \psi \mapsto \langle\psi, \phi\rangle\phi$ and then the probability of the observable A having value in a Borel set E in the state P_ϕ then works out to be

$$\langle P^A(E)\phi, \phi\rangle$$

Suppose, for instance, the spectrum of A consists of eigenvalues $\lambda_1, \lambda_2, ...,$ with $Au_n = \lambda_n u_n$ for an orthonormal basis $\{u_n\}_{n\geq 1}$ of $\mathcal{H}$. Then the probability that the observable represented by A has value in E in state P_ϕ is

$$\sum_{\{n:\lambda_n\in E\}} |\langle u_n, \phi\rangle|^2$$

Thus the spectrum $\sigma(A)$ here consists of all the possible values of A which could be realized.

To every system there is a special observable H called the *Hamiltonian*. The physical significance of this observable is that it describes the energy of the system. There is a second significance to this observable: if ρ is the state of the system at a given time then time t later the system evolves to the state

$$\rho_t = e^{-i\frac{t}{\hbar}H}\rho e^{i\frac{t}{\hbar}H},$$

where $\hbar$ is Planck's constant.

A basic system considered in quantum mechanics is the *harmonic oscillator*. One may think of this crudely as a ball attached to a spring, but the model is used widely, for instance also for the quantum theory of fields. The Hilbert space for the harmonic oscillator is $L^2(\mathbf{R})$. The Hamiltonian operator, up to scaling and addition of the constant $-\frac{1}{2}$, is

$$H = -\frac{d^2}{dx^2} + \frac{x^2}{4} - \frac{1}{2}.$$

The energy levels are then the spectrum of this operator. In this case the spectrum consists of all the eigenvalues $0, 1, 2, \ldots$. The creation operator bumps an eigenstate of energy n up to a state of energy $n+1$; an annihilation operator lowers the energy by 1 unit.

In many applications, the eigenstates represent quanta, i.e. particles. Thus raising the energy by one unit corresponds to the creation of a particle, while lowering the energy by one unit corresponds to annihilating a particle.

14. The abstract formulation

As before, we use the notation $W = \{0, 1, 2, \ldots\}$. We work with a real separable Hilbert space H_0, and a positive Hilbert-Schmidt operator B on H_0. Thus H_0 has an orthonormal basis $\{u_n\}_{n\in W}$ of eigenvectors of B, with

$$Bu_n = \lambda_n u_n$$

and

$$\textstyle\sum_{n\geq 0} |\lambda_n|^2 < \infty, \text{ with each } \lambda_n > 0.$$

The example to keep in mind is $H_0 = L^2(\mathbf{R})$, and $B = (-\frac{d^2}{dx^2} + \frac{x^2}{4} + \frac{1}{2})^{-1}$.

We have the coordinate map

$$I : H_0 \mapsto \mathbf{R}^W : f \mapsto \big(\langle f, u_n\rangle\big)_{n\in W}$$

Let

$$F_0 = I(H_0) = \{(x_n)_{n\in W} : \sum_{n\in W} x_n^2 < \infty\} \tag{14.1}$$

Now, for each $p \in W$, let

$$F_p = \{(x_n)_{n\in W} : \sum_{n\in W} \lambda_n^{-2p} x_n^2 < \infty\} \tag{14.2}$$

On F_p we have the inner-product $\langle\cdot,\cdot\rangle_p$ given by

$$\langle a, b\rangle_p = \sum_{n\in W} \lambda_n^{-2p} a_n b_n$$

This makes F_p a real Hilbert space, unitarily isomorphic to $L^2(W, \mu_p)$ where μ_p is the measure on W specified by $\mu_p(\{n\}) = \lambda_n^{-2p}$. Moreover, we have

$$F \stackrel{\text{def}}{=} \cap_{p\in W} F_p \subset \cdots F_2 \subset F_1 \subset F_0 = L^2(W, \mu_0) \tag{14.3}$$

Each inclusion $F_{p+1} \to F_p$ is Hilbert-Schmidt.

Now we pull all this back to $L^2(\mathbf{R})$. First set

$$H_p = I^{-1}(F_p) = \{x \in H_0 : \sum_{n\geq 0} \lambda_n^{-2p} |\langle x, u_n\rangle|^2 < \infty\} \tag{14.4}$$

It is readily checked that

$$H_p = B^p(H_0) \tag{14.5}$$

On H_p we have the pull back inner-product $\langle \cdot, \cdot \rangle_p$, which works out to

$$\langle f, g\rangle_p = \langle B^{-p} f, B^{-p} g\rangle \tag{14.6}$$

Then we have the chain

$$H \stackrel{\text{def}}{=} \cap_{p\in W} H_p \subset \cdots H_2 \subset H_1 \subset H_0, \tag{14.7}$$

with each inclusion $H_{p+1} \to H_p$ being Hilbert-Schmidt.

Equip H with the topology generated by the norms $\|\cdot\|_p$ (i.e. the smallest topology making all inclusions $H \to H_p$ continuous). Then H is, more or less by definition, a nuclear space.

The vectors u_n all lie in H and the set of all rational-linear combinations of these vectors produces a countable dense subspace of H.

Consider a linear functional on H which is continuous. Then it must be continuous with respect to some norm $\|\cdot\|_p$. Thus the topological dual H' is the union of the duals H_p'. In fact, we have:

$$H' = \cup_{p\in W} H_p \supset \cdots H_2' \supset H_1' \supset H_0' \simeq H_0, \tag{14.8}$$

where in the last step we used the usual Hilbert space isomorphism between H_0 and its dual H_0'.

Going over to the sequence space, H_p' corresponds to

$$F_{-p} \stackrel{\text{def}}{=} \{(x_n)_{n\in W} : \sum_{n\in W} \lambda_n^{2p} x_n^2 < \infty\} \tag{14.9}$$

The element $y \in F_{-p}$ corresponds to the linear functional on F_p given by

$$x \mapsto \sum_{n\in W} x_n y_n$$

which, by Cauchy-Schwartz, is well-defined and does define an element of the dual F_p' with norm equals to $\|y\|_{-p}$.

Consider now the product space $\mathbf{R}^W$, along with the coordinate projection maps

$$\hat{X}_j : \mathbf{R}^W \to \mathbf{R} : x \mapsto x_j$$

for each $j \in W$. Equip $\mathbf{R}^W$ with the product σ–algebra, i.e. the smallest sigma-algebra with respect to which each projection map $\hat{X}_j$ is measurable. A fundamental result in probability measure theory (a special case of Kolmogorov's theorem, for instance) says that there is a unique probability measure ν on the product σ–algebra such that each function $\hat{X}_j$, viewed as a random variable, has standard Gaussian distribution. Thus,

$$\int_{\mathbf{R}^W} e^{it\hat{X}_j}\,d\nu = e^{-t^2/2}$$

for $t \in \mathbf{R}$, and every $j \in W$. The measure ν is the product of the standard Gaussian measure $e^{-x^2/2}(2\pi)^{-1/2}dx$ on each component $\mathbf{R}$ of the product space $\mathbf{R}^W$.

Since, for any $p \geq 1$, we have

$$\int_{\mathbf{R}^W} \sum_{j\in W} \lambda_j^{2p} x_j^2\,d\nu(x) = \sum_{j\in W} \lambda_j^{2p} < \infty,$$

it follows that

$$\nu(F_{-p}) = 1$$

for all $p \geq 1$. Thus $\nu(F') = 1$.

We can, therefore, transfer the measure ν back to H', obtaining a probability measure μ on the sigma–algebra of subsets of H' generated by the maps

$$\hat{u}_j : H' \to \mathbf{R} : f \mapsto f(u_j),$$

where $\{u_j\}_{j\in W}$ is the orthonormal basis of H_0 we started with (note that each u_j lies in $H = \cap_{p\geq 0} H_p$). This is clearly the sigma–algebra generated by the weak topology on H' (which happens to be equal also to the sigma–algebras generated by the strong/inductive-limit topology).

Specialized to the example $H_0 = L^2(\mathbf{R})$, and $B = (-\frac{d^2}{dx^2} + \frac{x^2}{4} + \frac{1}{2})^{-1}$, we have the standard Gaussian measure on the distribution space $\mathcal{S}'(\mathbf{R})$.

The above discussion gives a simple direct description of the measure μ. Its existence is also obtainable by applying the well–known Minlos theorem.

To summarize, we can state the starting point of much of infinite-dimensional distribution theory (white noise analysis): Given a real, separable Hilbert space H_0 and a positive Hilbert-Schmidt operator B on H_0, we have constructed a nuclear space H and a unique probability measure μ on the Borel sigma–algebra of the dual H' such that there is a linear map

$$H_0 \to L^2(H', \mu) : x \mapsto \hat{x},$$

satisfying

$$\int_{H'} e^{it\hat{x}}\,d\nu = e^{-t^2\|x\|_0^2/2},$$

for every real t and $x \in H_0$. This Gaussian measure μ is often called the *white noise measure* and forms the background measure for white-noise analysis.

References

[1] Ch. Hermite, *Sur un Nouveau Développement en Série des Fonctions*, Comptes rendus de l'Academie des Sciences**14**, 93-266 (1864); *in* Oeuvres de Charles Hermite, Tome II, Gauthier-Villars (1908)

[2] T. Hida, H. -H. Kuo, J. Potthoff, L. Streit, *White Noise : An Infinite Dimensional Calculus*, Kluwer Academic Publishers (1993)

[3] H.-H. Kuo, *White Noise Distribution Theory*, CRC Press (1996).

[4] B. Simon, *Distributions and Their Hermite Expansions*, J. Math. Phys. **12**(1) 140 (1971).

Department of Mathematics, Louisiana State University, Baton Rouge, LA 70803, USA, e-mail: *sengupta@math.lsu.edu*