FUNCTION SPACES

Analysis of systems of linear algebraic equations leads naturally to the notion of a linear space in which vectors (i.e., the unknowns in the problem) and formation of linear combinations (i.e., the algebraic conditions imposed on the unkowns) are the principle ingredients. Considering linear systems in this setting of a linear space results in efficiency of expression and economy of effort. We are motivated to try to discover similar structures for the analysis of partial differential equations. Here, of course, functions are going to play the role played by vectors in linear algebra and the matrices will be replaced by partial differential operators. An additional difference is that the spaces of functions will be required to be closed under passage to the limit besides being closed under the operation of forming linear combinations.

1. Spaces of Functions

A linear space X is a collection of objects $x, y, z, \ldots$ together with a set of scalars $A, B, C\ldots$ and two operations:

i) Addition $\forall x, y \in X, \quad x + y \in X$

ii) Scalar Multiplication $\forall x \in X, \quad Ax \in X \quad for\ all\ scalars\ A$

Examples of linear spaces are familiar from linear algebra and algebra. These spaces are generally examples of finite dimensional linear spaces. For $1 \leq p < \infty$, let $L_p(U)$ denote the linear space of functions which are defined and whose p-th power is integrable on an open set $U \subset R^n$. That is, the functions which are defined on U and for which

$$\int_I |f(x)|^p dx < \infty. \qquad (1.1)$$

The open set U can be bounded or unbounded and the functions clearly need not be continuous in order for (1.1) to be satisfied. In fact the set of singularities permitted for functions in $L_p(U)$ is quite large and, as a result, the integral used must be more general than the Riemann integral. However, for what we propose to do, we will not require any detailed knowledge of a more general integration theory.

The linear structure on $L_p(U)$ is defined by

$$\forall f, g \in L_p(U),\ \forall A, B \in R, \quad let \quad [Af + Bg](x) = Af(x) + Bg(x) \quad \forall x \in U \qquad (1.2)$$

That $L_p(U)$ is a linear space is the assertion of,

Proposition 1.1 $\forall f, g \in L_p(U), \quad \forall A, B \in R, \qquad Af + Bg \in L_p(U).$

This result follows from the inequality

for each p, $1 \leq p < \infty, \qquad |A + B|^p \leq 2^{p-1}(|A|^p + |B|^p). \quad \forall A, B \in R.$

Note that the scalars here are taken to be the real numbers.

A real valued function, $N(\bullet)$, defined on a linear space X is a ***norm*** if it has the following properties:

1. $N(Ax) = |A| N(x) \qquad \forall A \in R,\ \forall x \in X$

2. $N(x + y) \leq N(x) + N(y) \quad \forall x, y \in X$

3. $N(x) \geq 0 \quad \forall x \in X \quad and \quad N(x) = 0 \quad iff \quad x = 0$

A norm is defined on $L_p(U)$, $1 \le p < \infty$, by

$$\|f\|_p = \left(\int_U |f(x)|^p dx \right)^{1/p}, \qquad for \quad f \in L_p(U).$$

Then $L_p(U)$ becomes a normed linear space. That $N(f) = \|f\|_p$ satisfies 1 and 3 is clear but 2 requires proof. Using the inequality

$$ab \le \frac{a^p}{p} + \frac{b^q}{q} \quad \forall a,b > 0, \quad and\ for \quad \frac{1}{p} + \frac{1}{q} = 1, \qquad (1.3)$$

leads to

Proposition 1.2 (Holder and Minkowski inequalities)

(a) If $\frac{1}{p} + \frac{1}{q} = 1$, then $\|fg\|_1 \le \|f\|_p \|g\|_q \qquad \forall f \in L_p(U), \quad \forall g \in L_q(U)$

(b) If $1 \le p < \infty$, then $\forall f,g \in L_p(U) \qquad \|f+g\|_p \le \|f\|_p + \|g\|_p$.

We should be careful to notice that $\|f-g\|_p = 0$ does not imply that $f(x) = g(x)$ at all $x \in U$. In fact, f and g can differ at infinitely many points in U so long as the set of points contains no positive volume subsets (i.e., f and g can differ on "sets of measure zero"). This means that functions in $L_p(U)$ are defined only up to sets of measure zero and that altering a function in $L_p(U)$ on such a set does not produce a different function in $L_p(U)$. Then functions in $L_p(I)$ are, in fact, equivalence classes of pointwise defined functions where the equivalence relation is equality "almost everywhere". When we say "let f be an arbitrary function in $L_p(U)$", what we really mean is "let f be an arbitrary representative of one of the equivalence classes of functions in $L_p(U)$".

To say a sequence $\{f_n(x)\}$ in $L_p(U)$ converges of to a limit f in $L_p(U)$ means $\|f_n - f\|_p \to 0 \ as \ n \to \infty$. This type of convergence is weaker than uniform convergence on U and neither implies nor is implied by pointwise convergence on U. Every sequence $\{f_n(x)\}$ in $L_p(U)$ that is convergent must be a ***Cauchy sequence***; i.e., it must satisfy $\|f_n - f_m\|_p \to 0 \ as \ m,n \to \infty$. In fact, it is true in any normed linear space that every convergent sequence must be a Cauchy sequence. On the other hand, in an arbitrary normed linear space it is not necessarily the case that every Cauchy sequence converges to a limit that belongs to the space. A space with the property that every Cauchy sequence is convergent is said to be ***complete***. The following result, known as the Riesz-Fischer Theorem in analysis, asserts that $L_p(U)$ is complete. The proof requires techniques of an integration theory more general than Riemann integration.

Proposition 1.3 (Riesz-Fischer) For $1 \le p < \infty$, $L_p(U)$ is complete for the norm $\|f\|_p$.

For each p, $1 \le p < \infty$, $L_p(U)$ is a complete, normed linear space. Such spaces are called Banach spaces. If $1 \le p \le q$, and if $|U|$, the volume of the set U, is finite, then we can show

$$\left(\frac{\int_U |f|^p}{|U|} \right)^{1/p} \le \left(\frac{\int_U |f|^q}{|U|} \right)^{1/q}$$

This implies that under these assumptions on p,q and $U,$ we have the following inclusions

$$L_q(U) \subset L_p(U) \subset L_1(U)$$

In fact, these inclusions are continuous injections. That means that for $1 \le p \le q,$ and $f \in L_q(U),$ we have $\|f\|_p \le \|f\|_q,$ which means that each f in $L_q(U)$ is also an element of $L_p(U)$ and the $L_p(U)$ norm of f is not greater than the $L_q(U)$ norm so the mapping which injects f from $L_p(U)$ into $L_q(U)$ is continuous.

Proposition 1.4 For $1 \le p < \infty,$ the space of test functions $C_c^\infty(U)$ is dense in $L_p(U)$; i.e.,for each f in $L_p(U),$ there exists a sequence of test functions, $\{\phi_n\}$ such that $\|\phi_n - f\|_p \longrightarrow 0$ as $n \to \infty.$

This result is a consequence of the so called "mollifier theorem". We will discuss this theorem in detail later.

2. Inner Product Spaces

An ***inner product space*** is a linear space X, on which there is defined a mapping which associates to every pair of elements $\{x,y\} \in X,$ a scalar value which we denote by $(x,y)_X.$ This mapping must have the following properties

$$\begin{aligned}
&1)\ \ (x,y)_X = (y,x)_X \qquad \forall x,y \in X \\
&2)\ \ (Ax+By,z)_X = A(x,z)_X + B(y,z)_X, \quad \forall x,y,z \in X, \quad \forall A,B \in R, \qquad (2.1)\\
&3)\ \ (x,x)_X \ge 0\ \ \forall x \in X, \quad and \quad (x,x)_X = 0 \quad iff \quad x = 0
\end{aligned}$$

The mapping is called an ***inner product*** on X.

The most familiar example of an inner product space is the space R^n of n-tuples $\vec{x} = [x_1,...,x_n],$ where the inner product is defined as

$$(\vec{x},\vec{y})_{R^n} = \sum_{i=1}^n x_i y_i. \qquad (2.2)$$

An inner product space has a norm, induced by the inner product. That is,

$$\|x\|_X = (x,x)_X^{1/2} \qquad for\ x \in X \qquad (2.3)$$

defines a ***norm*** on the linear space X. Recall that the norm defines a meaning for distance in the linear space X. In any inner product space, the following results are valid.

Proposition 2.1 (Cauchy-Schwartz and Triangle inequalities)

a) $|(x,y)_X| \le \|x\|_X \|y\|_X \qquad \forall x,y \in X$

b) $\|x+y\|_X \le \|x\|_X + \|y\|_X, \qquad \forall x,y \in X$

A ***Cauchy sequence*** in the inner product space X is a sequence of elements $\{x_m\} \subset X$ with the property that $\|x_m - x_k\|_X \to 0$ as $m,k \to \infty.$ A sequence of elements $\{x_m\} \subset X$ is said to be ***convergent*** if there exists an element $x \in X$ such that $\|x_m - x\|_X \to 0$ as $m \to \infty.$ The inner product space X is said to be ***complete*** if every Cauchy sequence in X is convergent. It is well known that the inner product space R^n is complete (this is just a consequence of the fact that the real numbers have been constructed to be complete).

We now consider two additional examples of inner product spaces. The first example is not a function space but will be important later.

The Space of Square Summable Sequences

Let ℓ_2 denote the linear space of infinite sequences of real numbers $\vec{x} = \{x_1, x_2, ...\}$. This space carries the same linear and inner product structure as R^n if we define

$$A\vec{x} + B\vec{y} = \{Ax_1 + By_1, Ax_2 + By_2, ...\} \qquad \forall x,y \in \ell_2 \quad \forall A,B \in R \qquad (2.4)$$

$$(\vec{x},\vec{y})_2 = \sum_{i=1}^{\infty} x_i y_i \qquad and \qquad \|\vec{x}\|_2 = (\vec{x},\vec{x})_2^{1/2} \qquad (2.5)$$

Of course the inner product and norm have to be restricted to those sequences for which the infinite sums are finite. Such sequences are said to be square summable sequences and we use the notation ℓ_2 to indicate the linear space of all such sequences. It is straightforward to show that the dimension of the space ℓ_2 is infinite. In fact, by showing there is a basis for ℓ_2 which is in one to one correspondence with the natural numbers, it follows that the dimension of this space is equal to the cardinality of the natural numbers.

Proposition 2.2 Every Cauchy sequence in ℓ_2 is convergent.
Proof- Let $\{\vec{x}^{(n)}\}$ denote a Cauchy sequence in ℓ_2. Then

$$\|\vec{x}^{(m)} - \vec{x}^{(n)}\|_2^2 = \sum_i \left|x_i^{(m)} - x_i^{(n)}\right|^2 \to 0 \quad as \quad m,n \to \infty.$$

For each fixed i, this implies that for every
$\varepsilon > 0,\ \exists\ N_\varepsilon \quad such\ that \quad \left|x_i^{(m)} - x_i^{(n)}\right|^2 < \varepsilon\ for\ all\ m,n > N_\varepsilon.$
Then for each i, $\left\{x_i^{(m)}\right\}$ is a Cauchy sequence in R, and since R is complete $x_i^{(m)} \to X_i\ as\ m \to \infty,$ for some real number X_i. It remains now to show that $\vec{X} = \{X_1, X_2, ...\}$ belongs to ℓ_2 and that $\{\vec{x}^{(n)}\}$ converges to $\vec{X}$ in ℓ_2.

For each fixed i, and $m,n > N_\varepsilon,$ $\quad \sum_{i=1}^{M}\left|x_i^{(m)} - x_i^{(n)}\right|^2 + \sum_{i=M+1}^{\infty}\left|x_i^{(m)} - x_i^{(n)}\right|^2 < \varepsilon,$ for arbitrary M

and since each sum is nonnegative,

$$\sum_{i=1}^{M}\left|x_i^{(m)} - x_i^{(n)}\right|^2 < \varepsilon.$$

Now fix m and let n tend to infinity. This leads to

$$\sum_{i=1}^{M}\left|x_i^{(m)} - X_i\right|^2 < \varepsilon, \qquad \forall m > N_\varepsilon\ and\ every\ M > 1.$$

Since M is arbitrary, let M tend to infinity to conclude that for every $\varepsilon > 0$ there is an N_ε such that

$$\sum_{i=1}^{\infty}\left|x_i^{(m)} - X_i\right|^2 < \varepsilon, \qquad \forall m > N_\varepsilon.$$

Then it follows that

$$\left|\left|\vec{X} - \vec{X}^{(m)}\right|\right|_2 \to 0 \quad as \quad m \to \infty,$$

and

$$\left|\left|\vec{X}\right|\right|_2 \leq \left|\left|\vec{X}-\vec{X}^{(m)}\right|\right|_2 + \left|\left|\vec{X}^{(m)}\right|\right|_2 \leq \varepsilon + \left|\left|\vec{X}^{(m)}\right|\right|_2 < \infty \quad \text{(so } \vec{X} \in \ell_2\text{).}\blacksquare$$

It should be noted that the set of vectors $\left\{\vec{E}^{(m)}\right\}$ in ℓ_2 where $E_i^{(m)} = \delta_{im}$, forms a basis for ℓ_2 in the sense that for any $\vec{x} \in \ell_2$, the sequence of elements

$$\vec{x}_m = \sum_{i=1}^m \left(\vec{x}, \vec{E}^{(i)}\right)_2 \vec{E}^{(i)}$$

is a Cauchy sequence in ℓ_2 converging to $\vec{x}$ as $m \to \infty$. There is, in general, no finite combination of $\vec{E}'s$ which equals $\vec{x}$.

The Function Space L_2

In the special case, $p = q = 2$, the Holder inequality looks like the Cauchy-Schwartz inequality. In fact, in this special case, $L_p(U) = L_2(U)$ is an inner product space. The inner product on $L_2(U)$ is defined by

$$(f,g)_2 = \int_U f(x)\, g(x)\, dx, \quad and \quad \|f\|_2 = (f,f)_2^{1/2} \tag{2.6}$$

$L_2(U)$ is the only one of the $L_p(U)$ spaces that supports an inner product. Since it is an $L_p(U)$ space, it has all the relevant properties such as completeness.

Two elements in an inner product space are said to be ***orthogonal*** if their inner product is zero. In $L_2(U)$ this has no visualizable significance (e.g. it does not mean the graphs of two orthogonal functions are orthogonal trajectories). However, orthogonality in $L_2(U)$ does imply linear independence. Since it is possible to generate an infinite family of orthogonal functions in $L_2(U)$, it follows that $L_2(U)$ is infinite dimensional. An infinite dimensional inner product space that is complete in the norm induced by the inner product is called a ***Hilbert space***. Note that if U is bounded, so that $\int_U 1\, dx = |U| < \infty$, it follows that $L_2(U)$ is contained in $L_1(U)$ To see this suppose $f \in L_2(U)$ and write

$$\|f\|_1 = \int_U |f|\, dx = \int_U 1\,|f|\, dx \leq \|1\|_2\, \|f\|_2$$

When U is not bounded, then neither space is contained in the other. For example consider the functions

$$f(x) = \left\{ \begin{array}{ll} 1/x & if \quad x > 1 \\ 0 & if \quad x < 1 \end{array} \right\} \quad and \quad g(x) = \left\{ \begin{array}{cc} 1/\sqrt{x} \quad if & 0 < x < 100 \\ 0 & otherwise \end{array} \right\}.$$

Then f belongs to $L_2(R)$ but does not belong to $L_1(R)$, while g belongs to $L_1(R)$ but does not belong to $L_2(R)$. These two spaces have many functions that are in both spaces but each contains some functions that are not in the other.

3. The Fourier Integral Transform

For functions which are defined on all of R^n we can define an alternative representation for the function. This alternative representation is called the ***Fourier transform*** of the function. The Fourier transform of the everywhere defined function $f(\vec{x})$ is defined as follows

$$T_F[f] = F(\vec{\alpha}) = (2\pi)^{-n} \int_{R^n} f(\vec{x})\, e^{-i\vec{x}\bullet\vec{\alpha}}\, dx.$$

Here, we will restrict our attention to the one dimensional case where we have

$$T_F[f] = F(\alpha) = (2\pi)^{-1} \int_R f(x)\, e^{-ix\alpha}\, dx \qquad (3.1)$$

The notations $T_F[f]$ and $F(\alpha)$ will be used interchangeably to denote the Fourier transform of the function f(x). Note that the transform does not exist for any function f(x) for which the improper integral (3.1) fails to converge. If the integral converges, it defines a possibly complex valued function of the real variable α. Evidently, a sufficient condition for the Fourier transform to exist is that f is absolutely integrable, i.e., $f \in L_1(R)$.

***Example* 3.1 Some Fourier Transforms**

(a) Consider

$$f(x) = \left\{ \begin{array}{lll} 1 & if & |x| < 1 \\ 0 & if & |x| > 1 \end{array} \right\}$$

Then $f \in L_1(R)$ and

$$F(\alpha) = \frac{1}{2\pi} \int_R f(x)\, e^{-ix\alpha} dx = \frac{1}{2\pi} \int_{-1}^{1} e^{-ix\alpha} dx = \frac{e^{-ix\alpha}}{-i2\pi\alpha} \Big|_{x=-1}^{x=1} = \frac{\sin\alpha}{\pi\alpha}$$

(b) Consider $f(x) = e^{-|x|}. \in L_1(R)$. Then

$$F(\alpha) = \frac{1}{2\pi} \int_R f(x)\, e^{-ix\alpha} dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-|x|}\, e^{-ix\alpha} dx = \frac{1}{2\pi} \int_{-\infty}^{0} e^{x-ix\alpha} dx + \frac{1}{2\pi} \int_{0}^{\infty} e^{-x-ix\alpha} dx$$

$$= \frac{1}{2\pi} \left[\frac{e^{x(1-i\alpha)}}{1-i\alpha} \Big|_{-\infty}^{0} + \frac{e^{-x(1+i\alpha)}}{1+i\alpha} \Big|_{0}^{\infty} \right] = \frac{1}{2\pi} \left[\frac{1}{1-i\alpha} + \frac{1}{1+i\alpha} \right] = \frac{1}{\pi} \frac{1}{1+\alpha^2}$$

where we used the result

$$\lim_{x\to-\infty} e^{x(1-i\alpha)} = \lim_{x\to\infty} e^{-x(1+i\alpha)} = 0.$$

(c) The Gaussian function $f(x) = e^{-x^2}$ belongs to $L_1(R)$ and

$$F(\alpha) = \frac{1}{2\pi} \int_R f(x)\, e^{-ix\alpha} dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-x^2}\, e^{-ix\alpha} dx$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-(x^2+ix\alpha-\alpha^2/4+\alpha^2/4)}\, dx = e^{-\alpha^2/4} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-(x+i\alpha)^2} dx$$

But

$$\int_{-\infty}^{\infty} e^{-(x+i\alpha)^2} dx = \int_{-\infty}^{\infty} e^{-z^2} dz = \sqrt{\pi}$$

hence

$$F(\alpha) = \frac{1}{\sqrt{4\pi}}\, e^{-\alpha^2/4} = T_F\left[e^{-x^2}\right].$$

Each of the functions in this example belongs to $L_1(R)$ and each transform is a continuous function of the transform variable α. In addition, it appears that the transform tends to zero as $\alpha \to \pm\infty$. In fact, this is true in general. We will use the notation $F \in C_0(R)$ to indicate that F is continuous on R^1 and $F(\alpha)$ tends to zero as $|\alpha|$ tends to infinity. Note that a function in $C_0(R)$ is **not necessarily** absolutely integrable. The function pairs $\{f(x), F(\alpha)\}$ listed above are examples of ***Fourier transform pairs***.

Theorem 3.1 If $f \in L_1(R)$ then $F(\alpha) = T_F[f(x)]$ exists and is a continuous function of $\alpha \in R^1$. Moreover, $F \in C_0(R)$.

The proof of this theorem requires the dominated convergence theorem for Lebesgue integrals and it therefore omitted. We list now several useful properties of the Fourier transform.

Theorem 3.2 If $f(x)$ *and* $g(x)$ have Fourier transforms $F(\alpha),\ G(\alpha),$ respectively, then

1. $\quad T_F[Af(x) + Bg(x)] = AF(\alpha) + BG(\alpha) \qquad \forall A, B \in R$

2. $\quad T_F[f(bx)] = \frac{1}{|b|} F\left(\frac{\alpha}{b}\right) \qquad \forall b \neq 0$

A transformation with property 1 is said to be linear and one with property 2 is said to be homogeneous.

Problem 1 Use the definition of the Fourier transform to prove theorem 3.2

Problem 2 Use theorem 3.2(2) to show that

(a) $T_F[I_A(x)] = \dfrac{\sin(\alpha A)}{\pi\alpha} \quad where \quad I_A(x) = \begin{Bmatrix} 1 & if & |x/A| < 1 \\ 0 & if & |x/A| > 1 \end{Bmatrix} = \begin{Bmatrix} 1 & if & |x| < A \\ 0 & if & |x| > A \end{Bmatrix}$

(b) $T_F[e^{-b|x|}] = \dfrac{b}{\pi}\, \dfrac{1}{b^2 + \alpha^2} \qquad for \quad b > 0$

(c) $T_F\left[e^{-bx^2}\right] = \dfrac{1}{\sqrt{4\pi b}}\, e^{-\alpha^2/4b} \qquad for\ b > 0$

Theorem 3.3 If $f(x)$ has Fourier transform $F(\alpha),$ then for all real values c the following transforms exist and are related to $F(\alpha)$ as indicated

1. $\quad T_F[f(x-c)] = e^{-ic\alpha}F(\alpha)$

2. $\quad T_F[e^{icx} f(x)] = F(\alpha - c)$

Problem 3 Use the definition of the Fourier transform to prove theorem 3.3

Theorem 3.4 If both $f(x)$ and $f'(x)$ have Fourier transforms then the transform of the derivative, $f'(x)$, is given in terms of $F(\alpha) = T_F[f]$ *by*

$$T_F[\,df/dx\,] = i\alpha F(\alpha).$$

More generally, if $f(x)$ and all its derivatives up to order m have Fourier transforms, then

$$T_F[\,d^k f/dx^k\,] = (i\alpha)^k F(\alpha) \quad for \quad k = 1,2,...,m$$

Theorem 3.5 If both $f(x)$ and $xf(x)$ have Fourier transforms then the transform of $xf(x)$, is given in terms of $F(\alpha) = T_F[f]$ *by*

$$T_F[\,xf(x)\,] = i\,d/d\alpha[F(\alpha)].$$

More generally, if $f(x)$ and $x^k f(x)$ have Fourier transforms, for $k = 1,2,...,m$ then

$$T_F[\,x^k f(x)] = (i\,d/d\alpha)^k[F(\alpha)] \qquad for \quad k = 1,2,...,m$$

Problem 4 Use the definition of the Fourier transform to prove theorem 3.4

Problem 5 Use the definition of the Fourier transform to prove theorem 3.5

For functions $f(x)$ *and* $g(x)$ defined on R, we formally define the ***convolution product*** of f and g as

$$f * g(x) = \int_R f(x-y)\,g(y)\,dy \qquad (3.2)$$

Problem 6 Use the definition of the convolution product and the change of variable, $z = x - y$, to show that $f * g(x) = g * f(x)$.

Theorem 3.6 If $f(x)$ *and* $g(x)$ belong to $L_1(R)$, then $f * g \in L_1(R)$ and

$$T_F[f * g] = 2\pi F(\alpha)\,G(\alpha).$$

The $\boldsymbol{L_1(R)}$ Inversion Theorem

We have seen that for each $f \in L_1(R)$, there is a Fourier transform $F \in C_0(R)$. In some sense, knowledge of one of these functions is equivalent to knowledge of the other; i.e., they are two different representations for the same information. To make this assertion more precise, we have the next theorem.

Theorem 3.7 If $f(x)$ belongs to $L_1(R)$, and, in addition, the Fourier transform F is also in $L_1(R)$, then

$$\phi(x) = \int_R F(\alpha)\,e^{ix\alpha} d\alpha \in C_0(R)$$

and

$$\|f - \phi\|_{L_1} = 0.$$

The assertion of this theorem is that if f *and* F both belong to $L_1(R)$, then the **inverse** Fourier transform of F is defined by

$$T_F^{-1}[F(\alpha)] = \int_R F(\alpha)\, e^{ix\alpha} d\alpha \qquad (3.3)$$

and $T_F^{-1}[F(\alpha)] = f(x)$ where the equality is equality in the sense of $L_1(R)$. Then (3.3) is known as the ***Fourier inversion formula***. By comparing (3.3) with (3.1), we arrive at the following result which can be used to increase the number of transform pairs.

Theorem 3.8 If $f(x)$ belongs to $L_1(R)$, and, in addition, the Fourier transform F is also in $L_1(R)$, then

$$T_F[F(\alpha)] = \frac{1}{2\pi} f(-\alpha).$$

Problem 7 Use theorem 3.8 and previous transform pairs to show that:

$$(a)\ T_F\left[\frac{\sin(Ax)}{\pi x}\right] = \frac{1}{2\pi} I_A(-x) = \frac{1}{2\pi} I_A(x)$$

$$(b)\ T_F\left[\frac{2b}{b^2+x^2}\right] = e^{-b|\alpha|} \qquad \textit{for} \quad b > 0$$

$$(c)\ T_F\left[e^{-x^2/4b}\right] = \frac{\sqrt{b}}{\sqrt{\pi}}\, e^{-b\alpha^2} \qquad \textit{for} \quad b > 0$$

4. The Fourier Transform in $L_2(R)$

Note that the functions

$$f_1(x) = I_A(x), \qquad f_2(x) = e^{-b|x|}, \qquad \textit{and} \qquad f_3(x) = e^{-bx^2}$$

all belong to $L_1(R)$ and each has a Fourier transform in $C_0(R)$ by theorem 3.1. Note further that these transforms

$$F_1(\alpha) = \frac{\sin(\alpha A)}{\pi\alpha}, \qquad F_2(\alpha) = \frac{b}{\pi}\frac{1}{b^2+\alpha^2} \qquad \textit{and} \qquad F_3(\alpha) = \frac{1}{\sqrt{4\pi b}}\, e^{-\alpha^2/4b}$$

all belong to $L_2(R)$, but only F_2 *and* F_3 belong to $L_1(R)$. Then only F_2 *and* F_3 can be inverted using Theorem 3.7 to recover the functions f_2 *and* f_3. So in what sense is it possible to say that the inverse transform of F_1 is f_1?

If $f \in L_1(R)$ then its Fourier transform exists and belongs to $C_0(R)$. For an arbitrary function $f \in L_2(R)$, it is not necessarily the case that $f \in L_1(R)$ and it is not clear then that its Fourier transform exists. In order to extend the Fourier transform to $L_2(R)$ we will use an idea that is pervasive in the study of partial differential equations. We first introduce a special subspace of $L_2(R)$ where transforming and inverting works more smoothly. This subspace has the additional property that the results which hold on the subspace can be extended to the whole space, $L_2(R)$, by passing to the limit. We define this subspace as follows.

The Space of Test Functions A function $\phi(x)$ defined on R is a test function if ϕ is infinitely differentiable and vanishes on the complement of some closed bounded interval $[a,b]$; *i.e.*, ϕ has compact support. We denote the linear space of test functions by $C_c^\infty(R)$.

It is obvious that $C_c^\infty(R) \subset L_1(R)$ and $C_c^\infty(R) \subset L_2(R)$ but a more surprising fact is true.

Theorem 4.1 For every $f \in L_p(R),\ 1 \le p < \infty,$ there exists a sequence $\{\phi_n\} \subset C_c^\infty(R)$ such that $\|\phi_n - f\|_p \to 0 \ as\ n \to \infty.$

We describe this by saying that $C_c^\infty(R)$ is a ***dense*** subspace of $L_p(R)$. The proof of this theorem as well as additional information about test functions will be given later when we develop the theory of generalized functions. We now proceed to use the test functions to extend the Fourier transform to $L_2(R)$.

Theorem 4.2 For every $\phi \in C_c^\infty(R)$, the Fourier transform, Φ, exists and, in addition, Φ belongs to $L_2(R)$. Moreover,

$$\|\phi\|_2^2 = 2\pi\|\Phi\|_2^2 \qquad (4.1)$$

It follows now from theorems 4.1 and 4.2 that

1) for every $f \in L_2(R)$, there exists a sequence $\{\phi_n(x)\} \in C_c^\infty(R)$ such that

$$\|\phi_n - f\|_2 \to 0 \ as\ n \to \infty. \quad \text{(hence} \quad \|\phi_n - \phi_m\|_2 \to 0 \ as\ m,n \to \infty\text{)}$$

2) since $\phi_n \in C_c^\infty(R)$, the Fourier transforms, Φ_n, exist and, in addition, $\Phi_n \in L_2(R)$. Moreover,

$$\|\phi_n - \phi_m\|_2^2 = 2\pi\|\Phi_n - \Phi_m\|_2^2 \to 0 \quad as \quad m,n \to \infty$$

3) since $L_2(R)$ is complete, and the sequence $\{\Phi_n\}$ is Cauchy, there exists a unique $\Phi \in L_2(R)$ such that

$$\|\Phi_n - \Phi\|_2 \to 0 \quad as \quad n \to \infty$$

4) since $\phi_n \to f,\ \Phi_n \to \Phi \quad in \quad L_2(R),\ and \quad \Phi_n = T_F[\phi_n]$, we define $\Phi = T_F[f]$

Note that we can also invert the transformation as follows,

1) for every $F \in L_2(R)$, there exists a sequence $\{\Psi_n(x)\} \in C_c^\infty(R)$ such that

$$\|\Psi_n - F\|_2 \to 0 \ as\ n \to \infty. \quad \text{(hence} \quad \|\Psi_n - \Psi_m\|_2 \to 0 \ as\ m,n \to \infty\text{)}$$

2) since $\Psi_n \in C_c^\infty(R)$, the inverse Fourier transforms,

$$\psi_n(x) = \int_R \Psi_n(\alpha)\, e^{ix\alpha} d\alpha$$

exist and, in addition, $\psi_n \in L_2(R)$. Moreover,

$$\|\psi_n - \psi_m\|_2^2 = 2\pi\|\Psi_n - \Psi_m\|_2^2 \to 0 \quad as \quad m,n \to \infty$$

3) since $L_2(R)$ is complete, and the sequence $\{\psi_n\}$ is Cauchy, there exists $\psi \in L_2(R)$ such that

$\|\psi_n - \psi\|_2 \to 0 \quad as \quad n \to \infty$

4) since $\Psi_n \to F, \ \psi_n \to \psi \ \ in \ \ L_2(R), \ and \ \ \psi_n = T_F^{-1}[\Psi_n],$ we define $\psi = T_F^{-1}[F]$

5) since $T_F[f-\psi] = F - F = 0, \ and \ \ \|f-\psi\|_2^2 = 2\pi\|T_F[f-\psi]\|_2^2 = 0, \ we \ have \ f = \psi.$

We summarize these observations in the following theorem.

Theorem 4.3 For every $f \in L_2(R)$ there exists a unique $F \in L_2(R)$ such that $F = T_F[f]$ in the sense that

if $\{\phi_n(x)\} \in C_c^\infty(R)$ is such that $\|\phi_n - f\|_2 \to 0 \ as \ n \to \infty,$

then $\Phi_n = T_F[\phi_n] \in L_2(R)$ is such that $\|\Phi_n - F\|_2 \to 0 \ as \ n \to \infty,$

In addition $T_F^{-1}[F] = f,$ where the equality is in the sense of $L_2(R)$.

Problem 8 Show that the definition of the $L_2(R)$ Fourier transform F does not depend on the choice of the sequence $\{\phi_n\}$

Each of the Fourier transform properties detailed in theorems 3.2 to 3.6 holds for the $L_2(R)$ extension of the Fourier transform. In addition, we have

Theorem 4.4 For $f,g \in L_2(R)$ the Fourier transforms, F,G satisfy

$$(f,g)_2 = \int_R f(x)\, g^*(x)\, dx = 2\pi \int_R F(\alpha)\, G^*(\alpha)\, d\alpha = 2\pi(F,G)_2 \qquad (4.2)$$

Here $g^*, \ G^*$ are used to denote complex conjugates. Even though we are dealing exclusively with real valued functions f,g the Fourier transforms may be complex valued and we have therefore stated the result for the complex form of the inner product on $L_2(R)$. Note that when $f = g$ (4.2) reduces to (4.1). The result (4.1) is known as the Parseval relation and (4.2) is called the Plancherel relation. Together they assert that the Fourier transform is a Hilbert space isometry, meaning that T_F maps $L_2(R)$ onto itself in a one to one norm and inner product preserving fashion.

5. Orthogonal Families and Generalized Fourier Series

Using a basis for representing arbitrary vectors in R^n has numerous advantages in dealing with problems in linear algebra. For most computational purposes, it is convenient if the basis is an orthonormal basis, meaning that the vectors in the basis are mutually orthogonal unit vector. If $\{\vec{u}_{1,\dots,}\vec{u}_n\}$ is such an orthonormal basis then an arbitrary $\vec{v} \in R^n$ can be uniquely expressed as

$$\vec{v} = \sum_{j=1}^n (\vec{v} \bullet \vec{u}_j)\, \vec{u}_j.$$

It is our aim now to develop such representations in infinite dimensional function spaces.

Let $U \subset R^n$ denote a bounded open and connected set in R^n and consider functions $f,g \in L_2(U)$. The functions are said to be ***orthogonal*** if

$$(f,g)_2 = \int_U f(x)\, g(x)\, dx = 0.$$

A countable family $\{g_1(x), g_2(x), ...,\}$ of functions in $L_2(U)$ is said to be an ***orthogonal family*** in $L_2(U)$ if $(g_i, g_j)_2 = 0 \; if \; i \neq j$. The orthogonal family $\{g_j(x)\}$ is said to be an ***orthonormal family*** if, in addition to be mutually orthogonal, the functions satisfy $\|g_j\|_2 = 1$ for every j; e.g., The family of functions $G_n(x) = \sin(n\pi x),\; n = 1,2,...$ is an orthogonal family in $L_2(0,1)$, and the family $g_n(x) = \sqrt{2}\sin(n\pi x),\; n = 1,2,...$ is an orthonormal family in $L_2(0,1)$.

Suppose $\{g_1(x), g_2(x), ...,\}$ is an orthonormal family of functions in $L_2(U)$, and for an arbitrary $f \in L_2(U)$, form the infinite sum

$$\sum_{n=1}^{\infty}(f, g_n)_2\, g_n(x).$$

It is not clear that this sum is convergent in $L_2(U)$, and even if it is convergent, it is not evident that it is convergent to f. In any case, we refer to the sum as the ***generalized Fourier series*** for f and we refer to the coefficients $f_n = (f, g_n)_2$ as the ***generalized Fourier coefficients*** for f.

Proposition 5.1 Suppose $\{g_1(x), g_2(x), ...,\}$ is an orthonormal family of functions in $L_2(U)$, and for an arbitrary $f \in L_2(U)$, let $f_n = (f, g_n)_2$. Then for any integer N, and any choice of constants $a_1, a_2, ...a_N$ we have

i) $\left|\left| f(x) - \sum_{n=1}^{N} a_n g_n(x) \right|\right|_2^2 = \|f\|_2^2 - \sum_{n=1}^{N}(f, g_n)_2^2 + \sum_{n=1}^{N}(f_n - a_n)^2$

ii) $\left|\left| f(x) - \sum_{n=1}^{N} a_n g_n(x) \right|\right|_2 \geq \left|\left| f(x) - \sum_{n=1}^{N} f_n g_n(x) \right|\right|_2$

Proof- For N a fixed positive integer, let

$$S_N(x) = \sum_{n=1}^{N} a_n g_n(x).$$

Then

$$\int_U (f(x) - S_N(x))^2 dx = \int_U f(x)^2 dx - 2\int_U f(x) S_N(x) dx + \int_U S_N(x)^2 dx.$$

But

$$\int_U f(x) S_N(x) dx = \int_U f(x) \sum_{n=1}^{N} a_n g_n(x)\, dx = \sum_{n=1}^{N} a_n \int_U f(x)\, g_n(x) dx = \sum_{n=1}^{N} a_n f_n$$

and

$$\int_U S_N(x)^2 dx = \int_U \sum_{m=1}^{N} a_m g_m(x) \sum_{n=1}^{N} a_n g_n(x)\, dx$$

$$= \sum_{m=1}^{N} a_m \sum_{n=1}^{N} a_n \int_U g_n(x) g_m(x) dx = \sum_{n=1}^{N} a_n^2 \quad \left(\text{since } (g_n, g_m)_2 = \delta_{mn}\right)$$

Then $\int_U (f(x) - S_N(x))^2 dx = \int_U f(x)^2 dx - 2\sum_{n=1}^{N} a_n f_n + \sum_{n=1}^{N} a_n^2$

$$= \int_U f(x)^2 dx - \sum_{n=1}^{N} f_n^2 + \sum_{n=1}^{N} f_n^2 - 2\sum_{n=1}^{N} a_n f_n + \sum_{n=1}^{N} a_n^2$$

$$= \int_U f(x)^2 dx - \sum_{n=1}^{N} f_n^2 + \sum_{n=1}^{N}(f_n - a_n)^2.$$

This is the result (i). Since $\sum_{n=1}^{N}(f_n - a_n)^2 \geq 0$, we get the result (ii). This last result asserts that among all linear combinations of the functions $g_1, ..., g_N$, the one that is closest to f in the $L_2(U) - norm$, is the combination with $a_n = f_n = (f, g_n)_2,\; n = 1, ..., N.$■

Theorem 5.2 Suppose $\{g_1(x), g_2(x), ...,\}$ is an orthonormal family of functions in $L_2(U)$, and for an arbitrary $f \in L_2(U)$, let $f_n = (f, g_n)_2$. Then

(1) (Bessel's inequality) $\sum_{n=1}^{\infty} f_n^2 \leq \|f\|_2^2 = \int_U f(x)^2 dx.$

(2) (Riemann-Lebesgue lemma) $f_n \to 0 \quad as \quad n \to \infty$

Proof-(1) Choosing $a_n = f_n$ in the sum $S_N(x)$, and using the results of the previous proposition,

$$0 \leq \int_U (f - S_N)^2 dx = \int_U f^2 dx - \sum_{n=1}^{N} f_n^2,$$

i.e.,

$$\sum_{n=1}^{N} f_n^2 \leq \int_U f^2 dx = \|f\|_2^2.$$

Since this result holds for all positive integers N, and the right side of the estimate does not depend on N, we are entitled to let N tend to infinity to obtain (1). Then the n-th term test implies the result (2).■

In the special case that the orthogonal family is $\{1, \cos x, \sin x, \cos 2x, \sin 2x, \ldots\}$ in $L_2(0, 2\pi)$ then the (2) takes the form

$$\int_0^{2\pi} f(x) \sin nx\, dx \to 0, \qquad \int_0^{2\pi} f(x) \cos nx\, dx \to 0, \qquad as \qquad n \to \infty.$$

This is what is often referred to as the Riemann-Lebesgue lemma.

The Bessel's inequality implies that for an arbitrary $f \in L_2(U)$, the sequence $f_n = (f, g_n)_2$ of generalized Fourier coefficients belongs to ℓ_2. Then it is also evident that the sequence of partial sums $S_N(x)$, is a Cauchy sequence and therefore converges in $L_2(U)$ to some limit, $S(x)$. However, it is not necessarily the case that $S = f$. An orthonormal family with the property that for every $f \in L_2(U)$, the generalized Fourier series converges to f in $L_2(U)$ is said to be a ***complete*** orthonormal family.

Theorem 5.3 Suppose $\{g_1(x), g_2(x), \ldots,\}$ is an orthonormal family of functions in $L_2(U)$, and for an arbitrary $f \in L_2(U)$, let $f_n = (f, g_n)_2$.
Then the following assertions are all equivalent:

1. $\{g_1(x), g_2(x), \ldots,\}$ is a complete orthonormal family

2. $\sum_{n=1}^{\infty} f_n^2 = \|f\|_2^2$

3. $f_n = 0\ \forall n, \quad if\ and\ only\ if \quad f = 0$

4. $\forall f, g \in L_2(U), \quad (f, g)_2 = \sum_{n=1}^{\infty} f_n g_n$

5. $\|f - S_N\|_2 \to 0 \quad as \quad N \to \infty$

As a result of this theorem, it follows that $L_2(U)$ is isometrically isomorphic to ℓ_2.

Theorem 5.4 Suppose $\{g_1(x), g_2(x), ..., \}$ is a complete orthonormal of functions in $L_2(U)$. Then for an every $f \in L_2(U)$, $f_n = (f, g_n)_2$ belongs to ℓ_2 with $\|\{f_n\}\|_{\ell_2} = \|f\|_2$. Conversely, for every $\{f_n\} \in \ell_2$, the sequence $S_N = \sum_{n=1}^{N} f_n g_n$, converges to a unique limit, f, in $L_2(U)$ and $\|\{f_n\}\|_{\ell_2} = \|f\|_2$.

Just as the Fourier transform provided an alternative but equivalent representation for functions in $L_2(R^n)$, the sequence of generalized Fourier coefficients provides an alternative but equivalent representation for functions in $L_2(U)$. The functions in the complete orthonormal family are the elements of a countable basis for $L_2(U)$, and every function in $L_2(U)$ can be expressed uniquely as a (possibly infinite) linear combination of these elements. While it is relatively easy to find examples of orthonormal families of functions in $L_2(U)$, it is not clear how to determine whether or not a family is complete. We shall see now one source of complete orthonormal families.

Sturm-Liouville Systems

Consider the problem of finding all scalars λ for which there exist nontrivial solutions to the following boundary value problem

$$\begin{aligned} &-d/dx(p(x)\,du/dx) + q(x)u(x) = \lambda r(x)u(x), \qquad a < x < b, \\ &C_1 u(a) + C_2 u'(a) = 0, \qquad C_3 u(b) + C_4 u'(b) = 0, \end{aligned} \tag{5.1}$$

Clearly the trivial solution satisfies (5.1) for all choices of the parameter λ. Any value λ which leads to a nontrivial solution will be called an ***eigenvalue*** of the problem and the corresponding nontrivial solution will be called an ***eigenfunction*** of the problem, corresponding to the eigenvalue λ. Note that if $u = u(x)$ is an eigenfunction corresponding to the eigenvalue λ, the for every nonzero constant k, the function $ku(x)$ is also an eigenfunction for the same eigenvalue. A problem of the form (5.1) is called a ***Sturm-Liouville problem***.

Theorem 5.5 Suppose the coefficients in the Sturm-Liouville problem (5.1) satisfy

$p(x),\ p'(x),\ q(x)\ \mathit{and}\ r(x)$ are all continuous on [a,b]

$$p(x) > 0,\ r(x) > 0 \quad \mathit{for\ all}\ x \in [a,b]$$

$$C_1^2 + C_2^2 > 0,\ \mathit{and}\quad C_3^2 + C_4^2 > 0,$$

Then

i) the Sturm-Liouville problem (5.1) has countably many real eigenvalues

$$\lambda_1 \le \lambda_2 \le ... \le \lambda_n \to +\infty$$

ii) for each eigenvalue there is a single independent eigenfunction, and eigenfunctions corresponding to distinct eigenvalues satisfy

$$\int_a^b u(x,\lambda_j)\, u(x,\lambda_k)\ r(x)\, dx = 0 \qquad j \ne k.$$

iii) The family of normalized eigenfunctions $\{u(x,\lambda_n)\}$ are a complete orthonormal family in $L_2(a,b)$ for the weighted inner product

$$((f,g))_2 = \int_a^b f(x)\,g(x)\;r(x)\,dx.$$

The proof of the Sturm-Liouville theorem is beyond the scope of this course. Instead we will list several examples of S-L problems and the associated eigenvalues and eigenfunctions. In each of the following examples we have $p(x) = r(x) = 1, \quad q(x) = 0, \ on\ [a,b] = [0,1].$ Then since $r(x) = 1,$ the weighted inner product of the theorem reduces to the usual L_2 inner product.

Scales of Hilbert Spaces

When a complete orthonormal family $\{u_n\}$ in $L_2(U)$ is given, we can use it to build a so called scale of Hilbert spaces that are nested in $L_2(U)$ and which can be used to describe the regularity of the functions they contain in the same way that the nested spaces

$$C^m(U) \subset C^{m-1}(U) \subset .... \subset C^1(U) \subset C^0(U)$$

contain continuous functions with more and more continuous derivatives with increasing m. We define the scale of Hilbert spaces as follows. For every $f \in L_2(U),$ the sequence $\{f_n = (f,u_n)_2\}$ belongs to ℓ_2. For $s \geq 0$ we define $H^s(U)$ to consist of those functions in $L_2(U)$ for which

$$\sum_n (1+\lambda_n)^s f_n^2 < \infty;$$

i.e.,

$$H^s(U) = \left\{ f \in L_2(U) \,:\, (1+\lambda_n)^{s/2} f_n \in \ell_2 \right\}.$$

Here, λ_n denotes the eigenvalue associated with the eigenfunction u_n. Since all the eigenfunctions we are going to encounter come from a S-L problem, the eigenvalue will always be known when $\{u_n\}$ in $L_2(U)$ is given. Note that

$$H^s(U) \subset ... \subset H^2(U) \subset H^1(U) \subset H^0(U) = L_2(U)\ .$$

Roughly speaking, $H^s(U)$ consists of those functions in $L_2(U)$ whose derivatives of order m are also in $L_2(U),$ for $m \leq s.$

6. Weak Derivatives

We want to weaken the notion of derivative so that it can be defined on all of $L_p(U)$. In order to do this, it will be convenient to define a new linear space that is slightly larger than $L_p(U)$. We define

$u \in L_p^{loc}(U)$ if, for every compact set K in $U,$ there exists a positive constant C_K such that

$$\int_K |u(x)|^p dx \leq C_K < \infty.$$

The linear space $L_p^{loc}(U)$ is not a normed linear space but we can describe the convergence in this space as follows,

u_n converges to u in $L_p^{loc}(U)$ if and only if u_n converges to u in $L_p(V)$ for all $V \subset\subset U.$

Here, the notation $V \subset\subset U$ means V is an open subset of U. It is clear that the derivative is not defined on all of $L_p^{loc}(U)$ but for a function $u \in C^1(\bar{U}) \subset L_p^{loc}(U)$ we have

$$\int_U \partial_j u(x)\,\phi(x)\,dx = \int_{\partial U} u\,\phi\,\bar{n}_j\,dS - \int_U u\,\partial_j\phi\,dx \qquad \text{for all test functions } \phi$$

$$= 0 - \int_U u\,\partial_j\phi\,dx \qquad \forall\phi \in C_c^\infty(U)$$

More generally, for $u \in C^k(\bar{U}) \subset L_p^{loc}(U)$ we have

$$\int_U D^\alpha u(x)\,\phi(x)\,dx = (-1)^{|\alpha|}\int_U u(x)\,D^\alpha\phi(x)\,dx \qquad \forall\phi \in C_c^\infty(U)$$

Here we are using the following notation,

$\alpha =$ multi-index $(\alpha_1,\ldots,\alpha_n)$ $\qquad \alpha_j = \text{integer}$

$$D^\alpha u = (\partial_1)^{\alpha_1}\cdots(\partial_n)^{\alpha_n}u(x) \qquad where \qquad \partial_j = \frac{\partial}{\partial x_j}$$

$$and \qquad (\partial_j)^{\alpha_j} = \frac{\partial^{\alpha_j}}{\partial x_j^{\alpha_j}}.$$

Motivated by this result, we can define the weak derivative on $L_p^{loc}(U)$,

for $u \in L_p^{loc}(U)$, we say $v = D^\alpha u$ if

$$\int_U v(x)\,\phi(x)\,dx = (-1)^{|\alpha|}\int_U u(x)\,D^\alpha\phi(x)\,dx \qquad \forall\phi \in C_c^\infty(U)$$

Note that the weak derivative is unique since if $v_1 = D^\alpha u = v_2$, for $v_1, v_2 \in L_p^{loc}(U)$, then

$$\int_U v_1(x)\,\phi(x)\,dx = (-1)^{|\alpha|}\int_U u(x)\,D^\alpha\phi(x)\,dx = \int_U v_2(x)\,\phi(x)\,dx \qquad \forall\phi \in C_c^\infty(U)$$

That is,

$$\int_U [v_1(x) - v_2(x)]\phi(x)\,dx = 0 \qquad \forall\phi \in C_c^\infty(U)$$

and this implies that $v_1 = v_2$ almost everywhere in U.

Example 6.1

1. Let $U = (0,2)$ and $\qquad u_1(x) = \begin{cases} x & if \quad 0 < x \le 1 \\ 1 & if \quad 1 < x < 2 \end{cases}$

Then

$$\int_U u_1'(x)\phi(x)dx \overset{def}{=} -\int_U u_1(x)\phi'(x)dx = -\int_0^1 x\,\phi'(x)dx - \int_1^2 \phi'(x)dx$$

$$= -x\phi(x)\big|_{x=0}^{x=1} + \int_0^1 \phi(x)dx - (\phi(2) - \phi(1))$$

$$= -\phi(1) + \phi(1) + \int_0^1 \phi(x)dx;$$

That is,

$$\int_U u_1'(x)\phi(x)dx = -\int_0^2 u_1(x)\phi'(x)dx = \int_0^2 v_1(x)\phi(x)dx = \int_0^1 \phi(x)dx$$

hence

$$v_1(x) = u_1'(x) = \begin{cases} 1 & if \quad 0 < x \le 1 \\ 0 & if \quad 1 < x < 2 \end{cases}.$$

Note that in this case, $v_1 = u_1'$ belongs to $L_p^{loc}(U)$ for $p \ge 1$.

2. Let $U = (0,2)$ and $\quad u_2(x) = \begin{cases} x & if \quad 0 < x \leq 1 \\ a & if \quad 1 < x < 2 \end{cases} \quad$ for $a \neq 1$.

Then

$$\int_U u_2'(x)\phi(x)dx \overset{def}{=} -\int_U u_2(x)\phi'(x)dx = -\int_0^1 x\phi'(x)dx - a\int_1^2 \phi'(x)dx$$

$$= -x\phi(x)|_{x=0}^{x=1} + \int_0^1 \phi(x)dx - a(\phi(2) - \phi(1))$$

$$= (a-1)\phi(1) + \int_0^1 \phi(x)dx;$$

That is,

$$\int_U u_2'(x)\phi(x)dx = -\int_0^2 u_2(x)\phi'(x)dx = \int_0^2 v_2(x)\phi(x)dx = (a-1)\phi(1) + \int_0^1 \phi(x)dx.$$

Now we can show that there is no v_2 in $L_p^{loc}(U)$ such that $v_2 = u_2'$. We choose a sequence of test functions $\{\phi_m(x)\}$ such that

- $0 \leq \phi_m(x) \leq 1$ for each m and all $x \in (0,2)$
- $\phi_m(1) = 1$ for each m
- $\phi_m(x)$ converges pointwise to 0 for all $x \neq 1$.

Then

$$\int_0^2 v_2(x)\phi_m(x)dx = (a-1)\phi_m(1) + \int_0^1 \phi_m(x)dx.$$

By the dominated convergence theorem

$$\int_0^2 v_2(x)\phi_m(x)dx \to 0 \qquad and \qquad \int_0^1 \phi_m(x)dx \to 0$$

and since $\phi_m(1) = 1$ for each m, we get the contradiction, $(a-1) = 0$.

In this case, $u_2'(x) = u_1'(x) + (a-1)\delta(x-1)$, from which we see that while the weak derivative is defined for every function in $L_p^{loc}(U)$, it is not necessarily the case that the weak derivative belongs to $L_p^{loc}(U)$. The space to which all the derivatives of $L_p^{loc}(U)$ functions must belong is the space of distributions which will be discusses later. We can, however, define

$$W^{k,p}(U) = \{u \in L_p(U) : D^\alpha u \in L_p(U) \text{ for all } \alpha,\ |\alpha| \leq k\}$$

Here $D^\alpha u$ denotes the weak derivative.

7. Properties of $W^{k,p}(U)$

For U an open subset of R^n, and $1 \leq p < \infty$, we have defined $W^{k,p}(U)$ as the set of all functions in $L_p(U)$ whose derivatives of order less than or equal to k are again in $L_p(U)$. That this is a linear space follows immediately from the fact that $L_p(U)$ is a linear space. If we define

$$\|u\|_{k,p} = \left(\sum_{|\alpha|\leq k} \|D^\alpha u\|_p^p\right)^{1/p}$$

then we can show this is a norm on $W^{k,p}(U)$ and that $W^{k,p}(U)$ is complete for this norm.

Lemma 7.1 For each k, and $1 \leq p < \infty,$ the function $\|u\|_{k,p}$ defines a norm on $W^{k,p}(U)$

Proof- The only part of the statement that is not easy is the triangle inequality. So write

$$\begin{aligned}\|u+v\|_{k,p} &= \left(\sum_{|\alpha|\leq k}\|D^\alpha u + D^\alpha v\|_p^p\right)^{1/p} \leq \left(\sum_{|\alpha|\leq k}\left(\|D^\alpha u\|_p + \|D^\alpha v\|_p\right)^p\right)^{1/p} \qquad (Minkowski)\\ &\leq \left(\sum_{|\alpha|\leq k}\left(\|D^\alpha u\|_p\right)^p\right)^{1/p} + \left(\sum_{|\alpha|\leq k}\left(\|D^\alpha v\|_p\right)^p\right)^{1/p} \qquad \left(discrete\ Minkowski\right)\\ &= \|u\|_{k,p} + \|v\|_{k,p} \blacksquare\end{aligned}$$

Lemma 7.2 For each k, and $1 \leq p < \infty,$ $W^{k,p}(U)$ is complete for the norm $\|u\|_{k,p}$

Proof- Suppose $\{u_m\}$ is a Cauchy sequence in the norm $\|u\|_{k,p}$. This just means that for each $\alpha,\ |\alpha| \leq k,\ \{D^\alpha u_m\}$ is a Cauchy sequence in $L_p(U)$. Since $L_p(U)$ is complete, it follows that there exist unique functions $U_\alpha \in L_p(U)$ such that $\|D^\alpha u_m - U_\alpha\|_p \to 0$. In particular, $\|u_m - U_0\|_p \to 0$. and $U_0 \in L_p(U)$. Now, for any test function $\phi(x)$ in U, write

$$\int_U U_0\, D^\alpha\phi\, dx = \lim_{m\to\infty}\int_U U_m\, D^\alpha\phi\, dx = \lim_{m\to\infty}(-1)^{|\alpha|}\int_U D^\alpha U_m\,\phi\, dx = (-1)^{|\alpha|}\int_U U_\alpha\,\phi\, dx$$

But this asserts that $U_\alpha = D^\alpha U_0$ and since $U_\alpha \in L_p(U)$ for each $\alpha, |\alpha| \leq k,$ we have shown that U_0 has all its weak derivatives of order $\alpha,\ |\alpha| \leq k,$ in $L_p(U)$. But then $\|u_m - U_0\|_{k,p} \to 0 \ and\ U_0 \in W^{k,p}(U)$.$\blacksquare$

Example 7.1

For $U = \{x \in R^n : \|x\| < 1\}$, let $r = (x_1^2 + \cdots + x_n^2)^{1/2}$ and

$$u(x) = u(r) = \frac{1}{r^\alpha} \text{ for } r > 0.$$

Note that for $r > 0$ we can compute

$$\partial_j u(x) = -\alpha\frac{1}{r^{1+\alpha}}\frac{x_j}{r} = -\alpha\frac{x_j}{r^{\alpha+2}},$$

and

$$|\nabla u(x)| = \left(\sum_{j=1}^n(\partial_j u(x))^2\right)^{1/2} = \frac{|\alpha|}{r^{\alpha+1}}.$$

For $\varepsilon > 0$ fixed, let U_ε denote the open set obtained by removing the ball of radius ε about the origin from U. Then

$$\int_{U_\varepsilon}|\nabla u(x)|\,dx \leq C\int_\varepsilon^1\frac{1}{\varepsilon^{\alpha+1}}r^{n-1}dr \leq C(1-\varepsilon^{n-\alpha-1}).$$

and

$$\int_U|\nabla u(x)|\,dx = \lim_{\varepsilon\to 0}\int_{U_\varepsilon}|\nabla u(x)|\,dx < \infty,$$

provided $n > \alpha + 1; i.e.,\ |\nabla u(x)| \in L_1(U)$ if $0 < \alpha < n-1$. More generally, $|\nabla u(x)| \in L_p(U)$ if $n > (\alpha+1)p$ which is to say, $u \in W^{1,p}(U)$ if $0 < \alpha < \dfrac{n-p}{p}$. Thus, membership in $W^{1,p}(U)$ does not preclude singular behavior, it does impose a limit on how bad the singularity can be.

Lemma 7.3 Suppose $u, v \in W^{k,p}(U)$ and $|\alpha| \leq k$. Then

a) $D^\alpha u \in W^{k-|\alpha|,p}(U)$ and $D^\alpha(D^\beta u) = D^\beta(D^\alpha u) = D^{\alpha+\beta}u$ for $|\alpha| + |\beta| \leq k$

b) $u \in W^{k,p}(V)$ for every open set V contained in U

c) for any test function ϕ in U, $\phi u \in W^{k,p}(U)$ and

$$D^\alpha(\phi u) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} D^\beta \phi D^{\alpha-\beta} u$$

This last result shows that the weak derivative behaves in most respects like the classical derivative.

8. The Hilbert-Sobolev Spaces

The spaces $W^{k,p}(U)$ are Banach spaces such that for p fixed, $W^{k+1,p}(U) \subset W^{k,p}(U)$, and for k fixed, $W^{k,q}(U) \subset W^{k,p}(U)$ if $q > p \geq 1$. In particular, $W^{k,p}(U) \subset W^{k,1}(U)$ for all $p > 1$. When $p = 2$, $W^{k,p}(U)$ is a Hilbert space. In this case we use the notation

$$H^k(U) = W^{k,2}(U) = \left\{u \in L_2(U) : D^\alpha u \in L_2(U),\ |\alpha| \leq k\right\}.$$

Then $H^0(U) = L_2(U)$ is a Hilbert space for the following inner product and norm,

$$(u,v)_0 = \int_U u(x)\,v(x)\,dx \qquad \|u\|_0 = (u,u)_0^{1/2}.$$

Similarly,

$$H^1(U) = \left\{u \in L_2(U) : \partial_j u \in L_2(U)\ 1 \leq j \leq n\right\}$$

is a Hilbert space for the norm and inner product defined by,

$$(u,v)_1 = (u,v)_0 + \int_U \nabla u \bullet \nabla v\,dx = \int_U [uv + \nabla u \bullet \nabla v]\,dx$$
$$\|u\|_1^2 = \|u\|_0^2 + \int_U |\nabla u|^2 dx$$

More generally, $H^k(U)$ is a Hilbert space for the norm and inner product defined by,

$$(u,v)_k = \sum_{|\alpha| \leq k} (D^\alpha u, D^\alpha v)_0 \qquad \textit{and} \qquad \|u\|_k^2 = \sum_{|\alpha| \leq k} \|D^\alpha u\|_0^2$$

Of course these are special cases of $W^{k,p}(U)$ and therefore continue to have all the recently proved properties of these spaces.

Define

$$H_0^k(U) = \text{the completion of } C_c^\infty(U) \text{ in the norm of } H^k(U).$$

Then $H_0^k(U)$ is a closed subspace of $H^k(U)$, and in general, it is a proper closed subspace. We will discuss this in more detail later. Our interest in $H_0^k(U)$ is motivated by the observation that for $u \in H_0^k(U)$, there exists a sequence of test functions $\{\phi_m\}$ such that $\|\phi_m - u\|_k \to 0$, and since each ϕ_m vanishes on the boundary of U, we can think of the functions in $H_0^k(U)$ as the functions in $H^k(U)$ that vanish on the boundary of U in some generalized sense. This too, will be discussed in detail later.

Lemma 8.1 (Poincare Inequality) For $U \subset R^n$ open and bounded, there exists a constant

$C > 0$ depending on U such that

$$\|u\|_0 \le C\|\nabla u\|_0 \qquad \text{for all } u \in H_0^1(U)$$

Proof- We will give the proof in the case $n = 2$. Extension to general n is straightforward. Suppose U is contained in the rectangle $(a,b) \times (c,d)$ and that $\phi \in C_c^\infty(U)$. Then

$$\phi(x,y) = \int_{-\infty}^{x} \partial_x\phi(s,y)\,ds = \int_a^x \partial_x\phi(s,y)\,ds \qquad \text{for } x \le b.$$

Then

$$|\phi(x,y)|^2 = \left|\int_a^x \partial_x\phi(s,y)\,ds\right|^2 \le \int_a^x |\partial_x\phi(s,y)|^2\,ds$$

$$\le \int_U 1^2 dx\ \int_a^x |\partial_x\phi(s,y)|^2\,ds \le (b-a)\int_a^b |\partial_x\phi(s,y)|^2\,ds$$

and
$$\int_a^b |\phi(x,y)|^2 dx \le (b-a)^2 \int_a^b |\partial_x\phi(s,y)|^2\,ds$$

$$\int_c^d\int_a^b |\phi(x,y)|^2 dxdy \le (b-a)^2 \int_c^d\int_a^b |\partial_x\phi(s,y)|^2\,dsdy$$

$$\le (b-a)^2 \int_c^d\int_a^b \left(|\partial_x\phi|^2 + |\partial_y\phi|^2\right) ds\,dy;$$

i.e.,

$$\|\phi\|_0^2 \le (b-a)^2\,\|\nabla\phi\|_0^2.$$

Since this holds for any $\phi \in C_c^\infty(U)$, and the test functions are dense in $H_0^1(U)$, the result extends to all of $H_0^1(U)$.■

It follows from this lemma that for all $u \in H_0^1(U)$,

$$\|u\|_1^2 = \|u\|_0^2 + \int_U |\nabla u|^2 dx \le \left[(b-a)^2 + 1\right]\int_U |\nabla u|^2 dx \le \left[(b-a)^2 + 1\right]\|u\|_1^2$$

which implies that
$$|u|_1^2 = \int_U |\nabla u|^2 dx$$

defines a norm on $H_0^1(U)$ that is equivalent to the norm $\|u\|_1$; i.e., sequences in $H_0^1(U)$ which are convergent in one of these norms must also converge in the other norm as well. By induction we can extend this arguement to show that

$$|u|_k^2 = \sum_{|\alpha|=k} \|D^\alpha u\|_0^2$$

is equivalent on $H_0^1(U)$ to the norm

$$\|u\|_k^2 = \sum_{|\alpha|\le k} \|D^\alpha u\|_0^2$$

Lemma 8.2 For $U \subset R^n$ open and bounded, $H_0^1(U)$ is a closed proper subset of $H^1(U)$.

Proof-For $U \subset R^n$ open and bounded, the function $1(x)$ that is identically equal to 1 is an element of $H^1(U)$ with $\|1\|_1 = \|1\|_0 = |U|$. For any $\phi \in C_c^\infty(U)$,

$$\|1-\phi\|_1 \ge \|1-\phi\|_0 \ge \|1\|_0 - \|\phi\|_0;$$

i.e.,

$$\|1-\phi\|_1 + \|\phi\|_0 \ge \|1\|_0 = |U| > 0$$

Also, by the Poincare inequality,

$$\|\phi\|_0 \leq C\|\nabla\phi\|_0 = C\|\nabla(1-\phi)\|_0 \leq C\|1-\phi\|_1$$

so that

$$\|1-\phi\|_1 \geq \frac{|U|}{1+C} > 0.$$

This implies that there is no test function ϕ that is close to the function 1 in the norm of $H^1(U)$; i.e., 1 does not belong to $H_0^1(U)$.■

Lemma 8.3 For $U = R^n$ $H_0^1(U) = H^1(U)$.

Proof-We begin by defining a "cutoff" function, that is a test function that is identically one inside a bounded set and is zero outside some neighborhood of the set. More precisely, let

$$\phi_r \in C_c^\infty(R^n) \qquad and \qquad \phi_r(x) = \begin{cases} 1 & if \quad |x| \leq r \\ 0 & if \quad |x| > r+1 \end{cases}$$

Then this is a smooth function that decreases smoothly from 1 to 0 in an annular region of fixed thickness so it follows that for each fixed integer, k, and all indices, α, $|\alpha| \leq k$, we have

$$\sup_r \|D^\alpha \phi_r\|_\infty \leq C(\alpha) \leq C(k)$$

Then for every $\psi \in C_c^\infty(R^n)$, $\quad \|\phi_r\psi - \psi\|_0 \to 0$, as $r \to \infty$, since $\phi_r\psi(x) = \psi(x)$ for $r > R$ where the support of $\psi(x)$ is contained in $B_R(0)$. Now it follows that $\|\phi_r u - u\|_0 \to 0$, as $r \to \infty$, for all $u \in H^0(R^n)$.

For $u \in H^1(R^n)$, consider the derivative term

$$\|\partial_x(\phi_r u - u)\|_0 \leq \|\partial_x(\phi_r u) - \phi_r\partial_x u\|_0 + \|\phi_r\partial_x u - \partial_x u\|_0$$

and note that

$$\|\phi_r\partial_x u - \partial_x u\|_0 = 0 \text{ for } r \text{ sufficiently large since } \partial_x u \in H^0(R^n)$$

$$\|\partial_x(\phi_r u) - \phi_r\partial_x u\|_0 = \|u\,\partial_x\phi_r\|_0 \leq \|(u-\psi_\varepsilon)\,\partial_x\phi_r\|_0 + \|\psi_\varepsilon\,\partial_x\phi_r\|_0$$

where, for $\varepsilon > 0$, we choose the test function $\psi_\varepsilon(x)$ so that $\|(u-\psi_\varepsilon)\|_0\|\partial_x\phi_r\|_\infty \leq \varepsilon$. In addition, since the support of ψ_ε is contained in some ball, $B_R(0)$, $R > 0$, we have $\psi_\varepsilon\,\partial_x\phi_r = 0$ for $r > R$. Combining all this, we conclude that $C_c^\infty(R^n)$ is dense in $H^1(R^n)$, which is to say, $H_0^1(R^n) = H^1(R^n)$.■

Evidently the difference between $H^1(R^n)$ and $H^1(U)$ for U bounded is that the norms of the derivatives of the approximating test functions grow with out bound as we approach the boundary of U. When $U = R^n$ there is no boundary and the derivatives of ϕ_r remain bounded.