Proof. We can assume f real. Consider any $K \Subset \Omega$ and $\Phi = \operatorname{sign} f$ on K and zero elsewhere. There is a sequence of $\phi_n \in C_0^\infty(\Omega)$ with $|\phi_n(x)| \leq 1$ and $\phi_n \to \Phi$ a.e. on K. Thus, by Lebesgue dominated convergence theorem

$$\int\limits_K |f| d\mathbf{x} = \int\limits_K f \lim \phi_n d\mathbf{x} = \lim \int\limits_K f\phi_n d\mathbf{x} = 0.$$

From the considerations above it is clear that $\partial_{\mathbf{x}}^\beta$ is a closed operator extending the classical differentiation operator (from $C^{|\beta|}(\Omega)$). One can also prove that $\partial_{\mathbf{x}}^\beta$ is the closure of the classical differentiation operator. If $\Omega = \mathbb{R}^n$, then this statement follows directly from (1.7) and (1.8). Indeed, let $f \in L_p(\mathbb{R}^n)$ and $g = D^\alpha f \in L_p(\mathbb{R}^n)$. We consider $f_\epsilon = J_\epsilon * f \to f$ in L_p. By Fubini theorem, we prove

$$\begin{aligned}\int\limits_{\mathbb{R}^n} (J_\epsilon * f)(\mathbf{x}) D^\alpha \phi(\mathbf{x}) d\mathbf{x} &= \int\limits_{\mathbb{R}^n} \omega_\epsilon(y) \int\limits_{\mathbb{R}^n} f(\mathbf{x}-\mathbf{y}) D^\alpha \phi(\mathbf{x}) d\mathbf{x} d\mathbf{y} \\ &= (-1)^{|\alpha|} \int\limits_{\mathbb{R}^n} \omega_\epsilon(y) \int\limits_{\mathbb{R}^n} g(\mathbf{x}-\mathbf{y}) \phi(\mathbf{x}) d\mathbf{x} d\mathbf{y} \\ &= (-1)^{|\alpha|} \int\limits_{\mathbb{R}^n} (J_\epsilon * g) \phi(\mathbf{x}) d\mathbf{x}\end{aligned}$$

so that $D^\alpha f_\epsilon = J_\epsilon * D^\alpha f = J_\epsilon * g \to g$ as $\epsilon \to 0$ in L_p. This shows that action of the distributional derivative can be obtained as the closure of the classical derivation.

Otherwise the proof is more complicated (see, e.g., [4, Theorem 3.16]) since we do not know whether we can extend f outside Ω in such a way that the extension still will have the generalized derivative.

In one-dimensional spaces the concept of the generalised derivative is closely related to a classical notion of absolutely continuous function. Let $I = [a, b] \subset \mathbb{R}^1$ be a bounded interval. We say that $f : I \to \mathbb{C}$ is *absolutely continuous* if, for any $\epsilon > 0$, there is $\delta > 0$ such that for any finite collection $\{(a_i, b_i)\}_i$ of disjoint intervals in $[a, b]$ satisfying $\sum_i (b_i - a_i) < \delta$, we have $\sum_i |f(b_i) - f(a_i)| < \epsilon$. The fundamental theorem of calculus, [150, Theorem 8.18], states that any absolutely continuous function f is differentiable almost everywhere, its derivative f' is Lebesgue integrable on $[a, b]$, and $f(t) - f(a) = \int_a^t f'(s) ds$. It can be proved (e.g., [61, Theorem VIII.2]) that absolutely continuous functions on $[a, b]$ are exactly integrable functions having integrable generalised derivatives and the generalised derivative of f coincides with the classical derivative of f almost everywhere.

Let us briefly explore this connection. First, let us observe that if

$$F(x) = \int\limits_a^x f(y) dy$$

where $f \in L_{p,loc}(\mathbb{R})$, then f is differentiable almost everywhere (it is absolutely continuous). Consider

$$A_h f(x) = \frac{1}{h} \int\limits_x^{x+h} \cancel{g}f(y)dy.$$

Clearly it a jointly continuous function on $\mathbb{R}_+ \times \mathbb{R}$. Further, denote

$$Hf(x) = \sup_{h>0} A_h|f|(x).$$

We restrict considerations to some bounded open interval I. Then $A_h f(x) \to f(x)$ if there is no n such that $x \in S_n = \{x;\ \limsup_{h\to 0} |A_h f(x) - f(x)| \geq 1/n\}$. Thus, we have to prove $\mu(S_n) = 0$ for any n.

Then we can assume that f is of bounded support and therefore, by Luzin theorem, for any ϵ there is a continuous function g with bounded support with $\mu(\{x \in I, f(x) \neq g(x)\}) \leq \epsilon$. Fix any ϵ and corresponding g. Then

$$\limsup_{h\to 0} |A_h f(x) - f(x)| \leq \sup_{h>0} |A_h(f(x) - g(x))| + \lim_{h\to 0} |A_h g(x) - g(x)| + |f(x) - g(x)|$$

The second term is zero, the third is 0 outside a set of measure ϵ. We need to estimate the first term. For a given ϕ consider an open set $E_\alpha = \{x \in I; H\phi(x) > \alpha\}$. For any $x \in E_\alpha$ we find $I_{x,r_x} = (x - r_x, x + r_x)$ such that $A_h f(x) > \alpha$. Thus, E_α is covered by these intervals. From the theory of Lebesque measure, the measure of any measurable set S is supremum over measures of compact sets $K \subset S$. Thus, for any $c < \mu(E_\alpha)$ we can find compact set $K \subset E_\alpha$ with $c < \mu(K) \subset \mu(E_\alpha)$ and a finite cover of K by $I_{x_i, r_{x_i}}$, $i = 1, \ldots, i_K$. Let us modify this cover in the following way. Let I_1 be the element of maximum length $2r_1$, I_2 be the largest of the remaining which are disjoint with I_1 and so on, until the collection is exhausted with $j = J$. According to the construction, if some $I_{x_i, r_{x_i}}$ is not in the selected list, then there is j such that $I_{x_i, r_{x_i}} \cap I_j \neq \emptyset$. Let as take the smallest such j, that is, the largest I_j. Then $2r_{x_i}$ is at most equal to the length of I_j, $2r_j$, and thus $I_{x_i, r_{x_i}} \subset I_j^*$ where the latter is the interval with the same centre as I_j but with length $6r_j$. The collection of I_j^* also covers K and we have

$$c \leq 6\sum_{j=1}^{J} r_j = 3\sum_{j=1}^{J} \mu(I_j) \leq \frac{3}{\alpha} \sum_{j=1}^{J} \int\limits_{I_j} |\phi(y)|dy \leq \frac{3}{\alpha} \int\limits_{I} |\phi(y)|dy.$$

Passing with $c \to \mu(E_\alpha)$ we get

$$\mu(E_\alpha) = \mu(\{x \in I; H\phi(x) > \alpha\}) \leq \frac{3}{\alpha} \int\limits_{I} |\phi(y)|dy.$$

Using this for $\phi = f - g$ we see that for any $\epsilon > 0$ we have

$$\mu(S_n) \leq 3n\epsilon + \epsilon$$

and, since ϵ is arbitrary, $\mu(S_n) = 0$ for any n. So, we have differentiability of $x \to \int_{x_0}^{x} f(y)dy$ almost everywhere.

Now, we observe that if $f \in L_{1,loc}(\mathbb{R})$ satisfies

$$\int_{\mathbb{R}} f\phi' dx = 0$$

for any $\phi \in C_0^\infty(\mathbb{R})$, then $f = 0$ almost everywhere. To prove this, we observe that if $\phi \in C_0^\infty(\mathbb{R})$ satisfies $\int_{\mathbb{R}} \psi dx = 1$, the for any $\omega \in C_0^\infty(\mathbb{R})$ there is $\phi \in C_0^\infty(\mathbb{R})$ satisfying

$$\phi' = \omega - \psi \int_{\mathbb{R}} \omega dx.$$

Indeed, $h = \omega - \psi \int_{\mathbb{R}} \omega dx$ is continuous compactly supported with $\int_{\mathbb{R}} h dx = 0$ and thus it has a unique compactly supported primitive. Hence

$$\int_{\mathbb{R}} f\phi' dx = \int_{R} f(\omega - \psi \int_{\mathbb{R}} \omega dy) dx = 0$$

or

$$\int_{\mathbb{R}} (f - \int_{\mathbb{R}} f\psi dy)\omega dx = 0$$

for any $\omega \in C_0^\infty(\mathbb{R})$ and thus $f = const$ almost everywhere.

Next, if $v(x) = \int_{x_0}^{x} f(y)dy$ for $f \in L_{1,loc}(\mathbb{R})$, then v is continuous and the generalized derivative of v, Dv, equals f. We can put $x_0 = 0$. Then

$$\begin{aligned}
\int_{\mathbb{R}} v\phi' dx &= \int_0^\infty (\int_0^x f(y)\phi'(x)dy)dx - \int_{-\infty}^0 (\int_x^0 f(y)\phi'(x)dy)dx \\
&= \int_0^\infty f(y)(\int_y^\infty \phi'(x)dx)dy - \int_{-\infty}^0 f(y)(\int_{-\infty}^0 \phi'(x)dx)dy \\
&= -\int_{\mathbb{R}} f(y)\phi(y)dy.
\end{aligned}$$

With these results, let $u \in L_{1,loc}(\mathbb{R})$ the distributional derivative $Du \in L_{1,loc}(\mathbb{R})$ and set $\bar{u}(x) = \int_0^x Du(t)dt$ Then $D\bar{u} = Du$ almost everywhere and hence $\bar{u} + C = u$ almost everywhere. Defining $\widetilde{u} = \bar{u} + C$ we see that $\widetilde{u}$ is continuous and has integral representation and thus it is differentiable almost everywhere.

1.2 Fundamental Theorems of Functional Analysis

The foundation of classical functional analysis are the four theorems which we formulate and discuss below.

1.2.1 Hahn–Banach Theorem

Theorem 1.6. *(Hahn–Banach) Let X be a normed space, X_0 a linear subspace of X, and x_1^* a continuous linear functional defined on X_0. Then there exists a continuous linear functional x^* defined on X such that $x^*(x) = x_1^*(x)$ for $x \in X_0$ and $\|x^*\| = \|x_1^*\|$.*

The Hahn–Banach theorem has a multitude of applications. For us, the most important one is in the theory of the dual space to X. The space $\mathcal{L}(X, \mathbb{R})$ (or $\mathcal{L}(X, \mathbb{C})$) of all continuous functionals is denoted by X^* and referred to as the *dual space*. The Hahn–Banach theorem implies that X^* is nonempty (as one can easily construct a continuous linear functional on a one-dimensional space) and, moreover, there are sufficiently many bounded functionals to separate points of x; that is, for any two points $x_1, x_2 \in X$ there is $x^* \in X^*$ such that $x^*(x_1) = 0$ and $x^*(x_2) = 1$. The Banach space $X^{**} = (X^*)^*$ is called the *second dual*. Every element $x \in X$ can be identified with an element of X^{**} by the evaluation formula

$$x(x^*) = x^*(x); \tag{1.19}$$

that is, X can be viewed as a subspace of X^{**}. To indicate that there is some symmetry between X and its dual and second dual we shall often write

$$x^*(x) = <x^*, x>_{X^* \times X},$$

where the subscript $X^* \times X$ is suppressed if no ambiguity is possible.

In general $X \neq X^{**}$. Spaces for which $X = X^{**}$ are called *reflexive*. Examples of reflexive spaces are rendered by Hilbert and L_p spaces with $1 < p < \infty$. However, the spaces L_1 and L_∞, as well as nontrivial spaces of continuous functions, fail to be reflexive.

Example 1.7. If $1 < p < \infty$, then the dual to $L_p(\Omega)$ can be identified with $L_q(\Omega)$ where $1/p + 1/q = 1$, and the duality pairing is given by

$$<f, g> = \int\limits_{\Omega} f(\mathbf{x}) g(\mathbf{x}) d\mathbf{x}, \quad f \in L_p(\Omega), \; g \in L_q(\Omega). \tag{1.20}$$

This shows, in particular, that $L_2(\Omega)$ is a Hilbert space and the above duality pairing gives the scalar product in the real case. If $L_2(\Omega)$ is considered over the complex field, then in order to get a scalar product, (1.20) should be modified by taking the complex adjoint of g.

Moreover, as mentioned above, the spaces $L_p(\Omega)$ with $1 < p < \infty$ are reflexive. On the other hand, if $p = 1$, then $(L_1(\Omega))^* = L_\infty(\Omega)$ with duality pairing given again by (1.20). However, the dual to L_∞ is much larger than $L_1(\Omega)$ and thus $L_1(\Omega)$ is not a reflexive space.

Another important corollary of the Hahn–Banach theorem is that for each $0 \neq x \in X$ there is an element $\bar{x}^* \in X^*$ that satisfies $\|\bar{x}^*\| = 1$ and $<\bar{x}^*, x> = \|x\|$. In general, the correspondence $x \to \bar{x}^*$ is multi-valued: this is the case in L_1-spaces and spaces of continuous functions it becomes, however, single-valued if the unit ball in X is strictly convex (e.g., in Hilbert spaces or L^p-spaces with $1 < p < \infty$; see [82]).

1.2.2 Spanning theorem and its application

A workhorse of analysis is the spanning criterion.

Theorem 1.8. *Let X be a normed space and $\{y_j\} \subset X$. Then $z \in Y := \overline{\mathcal{L}in}\{y_j\}$ if and only if*

$$\forall_{x^* \in X^*} < x^*, y_j >= 0 \quad \text{implies} \quad < x^*, z >= 0.$$

Proof. In one direction it follows easily from linearity and continuity.

Conversely, assume $< x^*, z >= 0$ for all x^* annihilating Y and $z \neq Y$. Thus, $\inf_{y \in Y} \|z - y\| = d > 0$ (from closedness). Define $Z = \mathcal{L}in\{Y, z\}$ and define a functional y^* on Z by $< y^*, \xi >=< y^*, y + az >= a$. We have

$$\|y + az\| = |a| \|\frac{y}{a} + z\| \geq |a| d$$

hence

$$| < y^*, \xi >| = |a| \leq \frac{\|y + az\|}{d} = d^{-1} \|\xi\|$$

and y^* is bounded. By H.-B. theorem, we extend it to $\widetilde{y}^*$ on X with $< \widetilde{y}^*, x >= 0$ on Y and $< \widetilde{y}^*, z >= 1 \neq 0$.

Next we consider the Müntz theorem.

Theorem 1.9. *Let $(\lambda_j)_{j \in \mathbb{N}}$ be a sequence of positive numbers tending to ∞. The functions $\{t^{\lambda_j}\}_{j \in \mathbb{N}}$ span the space of all continuous functions on $[0, 1]$ that vanish at $t = 0$ if and only if*

$$\sum_{j=1}^{\infty} \frac{1}{\lambda_j} = \infty.$$

Proof. We prove the 'sufficient' part. Let x^* be a bounded linear functional that vanishes on all t^{λ_j}:

$z \to f(z) \in \mathbb{C}$

$z \to f(z) \in X$

$$< x^*, t^{\lambda_j} >= 0, \qquad j \in \mathbb{N}.$$

For $\zeta \in \mathbb{C}$ such that $\Re\zeta > 0$, the functions $\zeta \to t^{\zeta}$ are analytic functions with values in $C([0, 1])$ This can be proved by showing that

$$\lim_{\mathbb{C} \ni h \to 0} \frac{t^{\zeta + h} - t^{\zeta}}{h} = (\ln t) t^{\zeta}$$

uniformly in $t \in [0,1]$. Then

$$f(\zeta) = < x^*, t^\zeta >$$

is a scalar analytic function of ζ with $\Re\zeta > 0$. We can assume that $\|x^*\| \leq 1$. Then

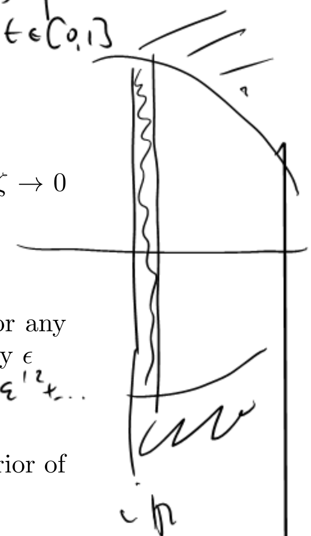
$$|f(\zeta)| \leq 1$$

$|f(\zeta)| \leq \|x^*\| \cdot \sup_{t \in [0,1]} |t^\zeta|$

for $\Re\zeta > 0$ and $f(\lambda_j) = 0$ for any $j \in \mathbb{N}$.

Next, for a given N, we define a Blaschke product by

$$B_N(\zeta) = \prod_{j=1}^{N} \frac{\zeta - \lambda_j}{\zeta + \lambda_j}.$$



We see that $B_N(\zeta) = 0$ if and only if $\zeta = \lambda_j$, $|B_N(\zeta)| \to 1$ both as $\Re\zeta \to 0$ and $|\zeta| \to \infty$. Hence

$$g_N(\zeta) = \frac{f(\zeta)}{B_N(\zeta)}$$

is analytic in $\Re\zeta > 0$. Moreover, for any ϵ' there is $\delta_0 > 0$ such that for any $\delta > \delta_0$ we have $|B_N(\zeta)| \geq 1 - \epsilon'$ on $\Re\zeta = \delta$ and $|\zeta| = \delta^{-1}$. Hence for any ϵ

$$|g_N(\zeta)| \leq 1 + \epsilon$$

$\frac{1}{1-\epsilon'} = 1 + \epsilon' + \epsilon'^2 + \ldots$

there and by the maximum principle the inequality extends to the interior of the domain. Taking $\epsilon \to 0$ we obtain $|g_N(\zeta)| \leq 1$ on $\Re\zeta > 0$.

Assume now there is $k > 0$ for which $f(k) \neq 0$. Then we have

$\langle x^*, t^k \rangle$

$$\prod_{j=1}^{N} \left| \frac{\lambda_j + k}{\lambda_j - k} \right| \leq \frac{1}{f(k)}.$$

$1 \geq |B_N(k)| = \frac{|f(k)|}{\prod \left|\frac{\lambda_j - k}{\lambda_j + k}\right|}$

Note, that this estimate is uniform in N. If we write

$$\frac{\lambda_j + k}{\lambda_j - k} = 1 + \frac{2k}{\lambda_j - k}$$

$\prod (1+\alpha_k)$

then, by $\lambda_j \to \infty$ almost all terms bigger then 1. Remembering that boundedness of the product is equivalent to the boundedness of the sum

$\ln \prod (1+\alpha_k)$
$=$
$\sum \ln(1+\alpha_k)$

$$\sum_{j=1}^{N} \frac{1}{\lambda_j - k}$$

$\alpha_k \to 0$

$\ln(1+\alpha_k) \leq \alpha_k$

we see that we arrived at contradiction with the assumption. Hence, we must have $f(k) = 0$ for any $k > 0$. This means, however, that any functional that vanishes on $\{t^{\lambda_j}\}$ vanishes also on t^k for any k. But, by the Stone- Weierstrass theorem, it must vanish on any continuous function (taking value 0 at zero). Hence, by the spanning criterion, any such continuous function belongs to the closed linear span of $\{t^{\lambda_j}\}$.

Example 1.10. The existence of an element $\bar{x}^*$ satisfying $<\bar{x}^*, x>= \|x\|$ has an important consequence for the relation between X and X^{**} in a nonreflexive case. Let B, B^*, B^{**} denote the unit balls in X, X^*, X^{**}, respectively. Because $x^* \in X^*$ is an operator over X, the definition of the operator norm gives

$$\|x^*\|_{X^*} = \sup_{x \in B} | <x^*, x> | = \sup_{x \in B} <x^*, x>, \tag{1.21}$$

and similarly, for $x \in X$ considered as an element of X^{**} according to (1.19), we have

$$\|x\|_{X^{**}} = \sup_{x^* \in B^*} | <x^*, x> | = \sup_{x^* \in B^*} <x^*, x> \leq \underbrace{\sup_{x^* \in B^*} \|x^*\|}_{=1} \cdot \|x\| \tag{1.22}$$

Thus, $\|x\|_{X^{**}} \leq \|x\|_X$. On the other hand,

$$\|x\|_X =<\bar{x}^*, x>\leq \sup_{x^* \in B^*} <x^*, x>= \|x\|_{X^{**}}$$

and

$$\|x\|_{X^{**}} = \|x\|_X. \tag{1.23}$$

Hence, in particular, the identification given by (1.19) is an isometry and X is a closed subspace of X^{**}.

The existence of a large number of functionals over X allows us to introduce new types of convergence. Apart from the standard *norm (or strong) convergence* where $(x_n)_{n \in \mathbb{N}} \subset X$ converges to x if

$$\lim_{n \to \infty} \|x_n - x\| = 0,$$

we define *weak convergence* by saying that $(x_n)_{n \in \mathbb{N}}$ weakly converges to x, if for any $x^* \in X^*$,

$$\lim_{n \to \infty} <x^*, x_n>=<x^*, x> .$$

In a similar manner, we say that $(x_n^*)_{n \in \mathbb{N}} \subset X^*$ converges $*$-*weakly* to x^* if, for any $x \in X$,

$$\lim_{n \to \infty} <x_n^*, x>=<x^*, x> .$$

Remark 1.11. It is worthwhile to note that we have a concept of a *weakly convergent* or *weakly Cauchy* sequence if the finite limit $\lim_{n \to \infty} <x^*, x_n>$ exists for any $x^* \in X^*$. In general, in this case we do not have a limit element. If every weakly convergent sequence converges weakly to an element of X, the Banach space is said to be *weakly sequentially complete*. It can be proved that reflexive spaces and L_1 spaces are weakly sequentially complete. On the other hand, no space containing a subspace isomorphic to the space c_0 (of sequences that converge to 0) is weakly sequentially complete (see, e.g., [6]).

Remark 1.12. Weak convergence is indeed weaker than the convergence in norm. However, we point out that a theorem proved by Mazur (e.g., see [172], p. 120) says that if $x_n \to x$ weakly, then there is a sequence of convex combinations of elements of $(x_n)_{n \in \mathbb{N}}$ that converges to x in norm. Thus, in particular, the norm and the weak closure of a convex sets coincide.