Notes on the Reisz-Markoff Theorem and Related Topics

Eric A. Carlen[1]
Rutgers University

May 1, 2017

0.1 Locally compact Hausdorff spaces

0.1 DEFINITION. A topological space $(X,\mathcal{U})$ is *locally compact* in case every point of X has a neighborhood with compact closure.

When $(X,\mathcal{U})$ is locally compact, every neighborhood U of x contain another neighborhood of x that has compact closure: Let V be any neighborhood of x that has compact closure, and then $V \cap U \subset U$ and $\overline{V \cap U} \subset \overline{V}$ which is compact.

For example, $\mathbb{R}^n$ is compact since for each x, $\overline{B_1(x)}$ is compact. Also, it is evident that any compact space is locally compact. However, an infinite dimensional Hilbert space with its norm topology is not locally compact: As we have seen, the closed unit ball – and thus any closed ball – in such a space fails to be compact, and every closed neighborhood must contain a closed ball.

In a locally compact Hausdorff space, one can separate compact sets K and points $y \in K^c$ as follows: For each $x \in K$, let V_x be a neighborhood of x with compact closure, and let U_x be a neighborhood of y such that $V_x \cap U_x = \emptyset$, which is possible since X is Hausdorff. Then $\{V_x \ : \ x \in K\}$ is an open cover of K so that there exist $\{x_1,\dots,x_n\} \subset K$ such that $K \subset \cup_{j=1}^n V_{x_j} =: V$. let $U = \cup_{j=1}^n U_{x_j}$. Since $\overline{V} \subset \cup_{j=1}^n \overline{V_{x_j}}$ which is compact, V has compact closure, and since $y \notin \overline{V_{x_j}}$ for any j, $y \notin \overline{V}$. In summary:

0.2 LEMMA. *Let $(X,\mathcal{U})$ be a locally compact Hausdorff space. Suppose $K \subset X$ is compact and $y \notin K$. Then there exists disjoint open sets V and U such that $K \subset V$, $y \in U$, and $\overline{V}$ is compact.*

The slightly more elaborate result in the next lemma is fundamental.

0.3 LEMMA. *Let $(X,\mathcal{U})$ be a locally compact Hausdorff space. Suppose $K \subset U \subset X$ with K compact and U open. Then there exists an open set V with compact closure such that*

$$K \subset V \subset \overline{V} \subset U \ . \tag{0.1}$$

Proof of Lemma 0.3. If $y \in U^c$, then $y \in K^c$, and using Lemma 0.2 we may choose for each $y \in U^c$ disjoint open sets V_y and W_y such that $K \subset V_y$, $y \in W_y$ and $\overline{V_y}$ is compact. Since $y \notin \overline{V_y}$,

$$\bigcap_{y \in U^c} U^c \cap \overline{V_y} = \emptyset \ .$$

and each set in the intersection is compact. Therefore there exist $\{y_1, \ldots, y_n\} \subset U^c$ such that

$$\bigcap_{j=1,\ldots,n} U^c \cap \overline{V_{v_j}} = \emptyset \ .$$

Define $V = \cap_{j=1}^n V_{y_j}$ which is open and contains K. Since $\overline{V} \subset \cap_{j=1}^n \overline{V_{y_j}}$, which is compact and disjoint from U^c, $\overline{V}$ is compact and $\overline{V} \cap U^c = \emptyset$, which is the same as $\overline{V} \subset U$. □

The main feature of locally compact Hausdorff spaces that makes it possible to develop a rich theory linking topology and integration for them is that locally compact Hausdorff spaces are rich in continuous functions in the sense that we now explain. We shall write

$$K \prec f$$

to mean that K is a compact subset of X, and that f is a continuous and compactly supported function on X with values in $[0,1]$ and with $f(x) = 1$ for all $x \in K$.

We shall write

$$f \prec U$$

to mean that U is a compact subset of X, and that f is a continuous and compactly supported function on X with values in $[0,1]$ and that the support of f is contained in U.

0.4 THEOREM (Urysohn's Lemma)**.** *Let $(X,\mathcal{U})$ be a locally compact Hausdorff space. Suppose $K \subset U \subset X$ with K compact and U open. Then there exists a continuous function f such that*

$$K \prec f \prec U \ . \tag{0.2}$$

Proof. First, pick an open G with compact closure such that $K \subset G \subset \overline{G} \subset U$, which we may do by Lemma 0.3. In the next step we construct a sequence of open sets $\{V_s\}$ indexed by the dyadic rational numbers s in $(0,1)$ such that for each $t > s$,

$$K \subset V_t \subset \overline{V_t} \subset V_s \subset G \ .$$

We proceed inductively. By what we have shown in the first step, there exists an open set $V_{1/2}$ such that

$$K \subset V_{1/2} \subset \overline{V_{1/2}} \subset G \ .$$

For the same reason, there exist open sets $V_{1/4}$ and $V_{3/4}$ such that

$$K \subset V_{3/4} \subset \overline{V_{3/4}} \subset V_{1/2} \qquad \text{and} \qquad \overline{V_{1/2}} \subset V_{1/4} \subset \overline{V_{1/4}} \subset G \ .$$

Now an obvious induction argument provides the construction.

Having constructed our sequence $\{V_s\}$, s ranging over the dyadic rationals in $(0,1)$, we are ready to define f: First, set $V_0 := X$. Next, for $x \in X$, define

$$f(x) = \sup\{s \ : \ x \in V_s\} \ .$$

If $x \in K$, then $x \in V_s$ for all s, and hence $f(x) = 1$, and if $x \in G^c$, then $f(x) = 0$. Hence the support of f is contained in $\overline{G}$, which is compact. Finally, we claim that f is continuous. It suffices to show that for all λ, $f^{-1}((\lambda,\infty))$ is open and that and $f^{-1}([\lambda,\infty))$ is closed.

First, we claim that

$$f^{-1}((\lambda,\infty)) = \bigcup_{s>\lambda} V_s \ ,$$

which is open. To see this, note that if $f(x) > \lambda$ then for some $s > \lambda$, $x \in V_s$, and so x belongs to the union on the right. On the other hand, if x belongs to the union on the right, then $x \in V_s$ for some $s > \lambda$, and then $f(x) > \lambda$

Second, we claim that

$$f^{-1}([\lambda,\infty)) = \bigcap_{s<\lambda} \overline{V_s} \ ,$$

To see this, note that if $f(x) \geq \lambda$, then for all $s < \lambda$, $x \in V_s$, and hence $x \in \overline{V_s}$, and so x belongs to the intersection on the right. On the other hand, if x belongs to the intersection on the right, then for all $r < \lambda$, there is an $s > r$ so that $x \in \overline{V_s}$. Since $\overline{V_s} \subset V_r$, $x \in V_r$ and $f(x) \geq r$. Since $r < \lambda$ is arbitrary, $f(x) \geq \lambda$. □

0.5 THEOREM (Continuous partitions of unity)**.** *Let $(X,\mathcal{U})$ be a locally compact Hausdorff space. Let K be a compact subset of X, and let $\{U_1,\dots,U_n\}$ be a finite open cover of K. Then there exist functions g_j, $j = 1,\dots,n$ such that $g_j \prec U_j$ for each j and such that $K \prec \sum_{j=1}^n g_j$.*

Proof. Each $x \in K$ belongs to some U_j, and applying Lemma 0.3 to $\{x\} \subset U_j$, we can choose an open set W_x with compact closure such that $x \in W_x \subset \overline{W_x} \subset U_j$. Since K is compact, we may cover K by finitely many such sets $\{W_{x_1},\dots,W_{x_m}\}$. Let C_ℓ be the union of those $\overline{W_{x_k}}$ that were constructed choosing $W_x \subset U_\ell$. Then C_ℓ is compact, and $C_\ell \subset U_\ell$, and $K \subset \cup_{\ell=1}^n C_\ell$.

By Urysohn's Lemma, there exists f_ℓ such that $C_\ell \prec f_\ell \prec U_\ell$. How consider the function

$$h := 1 - \prod_{j=1}^n (1-f_j) \ .$$

Clearly, h is continuous and $0 \leq h \leq 1$. If $x \in K$, $x \in C_j$ for some j, and then $f_j(x) = 1$ so that $h(x) = 1$.

We next claim that

$$h = f_1 + \sum_{j=1}^{n-1} f_j \prod_{k=1}^{j-1}(1-f_k) \ .$$

To see this note that

$$1 - \prod_{j=1}^n (1-f_j) = f_1 + (1-f_1) - \prod_{j=1}^n (1-f_j) = f_1 + (1-f_1)\left[1 - \prod_{j=2}^n (1-f_j)\right] \ ,$$

and then apply the obvious induction argument.

Finally, we define

$$g_1 = f_1 \qquad \text{and} \qquad g_j = f_j \prod_{k=1}^{j-1}(1-f_k) \quad \text{for} \quad j = 1,\dots,n-1 \ .$$

Thus, $h = \sum_{j=1}^n g_j$ and $g_j \prec U_j$ for each j. □

0.2 Some spaces of continuous functions

Given a topological space $(X, \mathcal{U})$, there are three natural normed vector spaces of continuous functions:

0.6 DEFINITION (Spaces of continuous functions)**.** Let $(X, \mathcal{U})$ be a topological space. We define:
(i) $\mathcal{C}_c(X)$ is the normed vector space of continuous, compactly supported functions f on X with values in $\mathbb{C}$ on which the norm, $\|\cdot\|_\infty$, is given by $\|f\|_\infty = \sup\{\ |f(x)|\ :\ x \in X\ \}$.
(ii) $\mathcal{C}_0(X)$ is the normed vector space of continuous functions f on X with values in $\mathbb{C}$ such that for each $\epsilon > 0$, there exists a compact set $K_{\epsilon,f}$ such that $|f(x)| < \epsilon$ for all x outside of $K_{\epsilon,f}$. The norm is once again $\|\cdot\|_\infty$, given by $\|f\|_\infty = \sup\{\ |f(x)|\ :\ x \in X\ \}$.
(iii) $\mathcal{C}_b(X)$ is the normed vector space of continuous functions f on X with values in $\mathbb{C}$ such that $\|f\|_\infty = \sup\{\ |f(x)|\ :\ x \in X\ \} < \infty$. The norm is once again $\|\cdot\|_\infty$.

When X is not compact, $\mathcal{C}_c(X)$ is not complete, but if $(X, \mathcal{U})$ is a locally compact Hausdorff space, then it is dense in $\mathcal{C}_0(X)$, and both $\mathcal{C}_0(X)$ and $\mathcal{C}_b(X)$ are Banach spaces.

0.7 THEOREM. *$\mathcal{C}_0(X)$ equipped with the sup norm is a Banach space. If X is a locally compact Hausdorff space, then the subspace $\mathcal{C}_c(X)$ is dense.*

Proof. If $\{f_n\}$ is a Cauchy sequence in $\mathcal{C}_0(X)$, then $\{f_n(x)\}$ is a Cauchy sequence in $\mathbb{C}$. Hence the limit $\lim_{n\to\infty} f_n(x)$ exists for each x, and we define a function f by

$$f(x) = \lim_{n\to\infty} f_n(x)\ .$$

Given the uniform convergence, it is easy to check, using an $\epsilon/3$ argument, that $f \in \mathcal{C}_0(X)$ so that $\mathcal{C}_0(X)$ is complete.

To see that $\mathcal{C}_c(X)$ is dense, pick $f \in \mathcal{C}_0(X)$ and $\epsilon > 0$. Let K_ϵ be a compact set such that $|f(x)| \le \epsilon$ for all $x \notin K_\epsilon$. Because X is locally compact, it is possible to find an open set U containing K_ϵ such that U has compact closure. Then by Urysohn's Lemma, there exists a continuous function g with $K \prec g \prec U$. Then $fg = f$ on K and and $|fg - f| = |f||g-1| \le |f|$ everywhere, so that $|fg - f| \le \epsilon$ on K_ϵ^c. In particular, $\|fg - f\|_\infty \le \epsilon$, and $fg \in \mathcal{C}_c(X)$. □

0.3 Radon measures

Radon measures, are, roughly speaking, the class of Borel measures on a locally compact Hausdorff space for which the measure theory and the topology are "nicely compatible".

0.8 DEFINITION (Inner and outer regularity)**.** Let $(X, \mathcal{U})$ be a topological space. A Borel measure μ on X is *outer regular* in case for each Borel set E

$$\mu(E) = \inf\{\ \mu(U)\ :\ E \subset U\ ,\ U \text{ open}\ \}\ . \tag{0.3}$$

A Borel measure μ is *inner regular* in case for each Borel set E

$$\mu(E) = \sup\{\ \mu(K)\ :\ K \subset E\ ,\ K \text{ compact}\ \}\ . \tag{0.4}$$

A Borel measure μ is *inner regular for open sets* in case (0.4) holds for all *open* E. A Borel measure is *regular* if it is both inner and outer regular.

0.9 DEFINITION (Radon measure). A *Radon measure* on a topological space $(X,\mathcal{U})$ is a positive Borel measure μ on X such that $\mu(K) < \infty$ for all compact sets $K \subset X$ that is outer regular and inner regular for open sets.

The next results demonstrate the compatibility of the topology and the measure theory for Radon measures.

0.10 THEOREM. *Let $(X,\mathcal{U})$ be a locally compact Hausdorff space, and let μ be a Radon measure on X. Then $\mathcal{C}_c(X)$ is dense in $L^p(X,\mathcal{B},\mu)$ for all $p \in [1,\infty)$.*

Proof. Since $L^1 \cap L^\infty$ is dense in Lp for all $p \in [1,\infty)$, it suffices to deal with $p = 1$. We know that integrable simple functions are dense in $L^1(\mu)$. Therefore, it suffices to show that whenever E is a Borel set with $\mu(E) < \infty$, for all $\epsilon > 0$, there exists $f \in \mathcal{C}_c(X)$ such that

$$\int_X |f - 1_E| \mathrm{d}\mu \leq \epsilon \ .$$

Since μ is outer regular, there exists U open with $E \subset U$ such that $\mu(U) \leq \mu(E) + \epsilon$. Since μ is inner regular for open sets, there is a compact set $K \subset U$ such that $\mu(K) \geq \mu(U) - \epsilon/$.

By Urysohn's Lemma, there exists $f \in \mathcal{C}_c(X)$ such that $K \prec f \prec U$. But then $|f - 1_U| \leq 1_{U \cap K^c}$, and so

$$\|f - 1_E\|_1 \leq \|f - 1_U\|_1 + \|1_U - 1_E\|_1 \leq \mu(U \backslash E) + \mu(U \cap K^c) \leq 2\epsilon \ .$$

□

0.11 THEOREM. *Let $(X,\mathcal{U})$ be a locally compact Hausdorff space. Let μ and ν be a Radon measures on X such that*

$$\int_X f \mathrm{d}\mu = \int_X f \mathrm{d}\nu$$

for all $f \in C_c(X)$. Then $\mu = \nu$.

Proof. Let U be open, and let $f \prec U$. For all compact $K \subset U$, Urysohn's lemma provides g with $K \prec g \prec U$, and hence

$$\mu(K) \leq \int_X g \mathrm{d}\mu = \int_X g \mathrm{d}\nu = \nu(U) \ .$$

By the inner regularity of μ on open sets, $\mu(U) = \sup\{\mu(K) \ : \ K \subset U, K \text{ compact}\}$. Hence $\mu(U) \leq \nu(U)$ By symmetry, $\nu(U) \leq \mu(U)$ so that $\mu(U) = \nu(U)$ for all open U, and then by the outer regularity of μ and ν, $\mu(E) = \nu(E)$ for all Borel sets E. □

0.12 THEOREM. *Let $(X,\mathcal{U})$ be a locally compact Hausdorff space. Every σ-finite Radon measure μ on X is regular.*

Proof. Let E be a Borel set with $\mu(E) < \infty$. Pick $\epsilon > 0$. By the outer regularity of μ, there is an open set U such that $E \subset U$ and $\mu(U) \leq \mu(E) + \epsilon$. For the same reason, there is an open set V such that $U \backslash E \subset V$ and $\mu(V) \leq \mu(U \backslash E) + \epsilon \leq 2\epsilon$.

By the inner regularity of μ for open sets, there is a compact set $C \subset U$ such that $\mu(U) \leq \mu(C) + \epsilon$.

Define $K := C \cap V^c$ which is compact. By construction, every point $x \in U$ that is not in E lies in V, so that $C \cap V^c \subset E$. Therefore,

$$\mu(K) \geq \mu(C) - \mu(V) \geq \mu(U) - 2\epsilon \ .$$

Hence $\mu(K) \leq \mu(E) - 2\epsilon$, and since $\epsilon > 0$ is arbitrary, $\mu(E) = \sup\{\ \mu(K)\ :\ K \subset E\ ,\ K \text{ compact }\}$.

Finally, if $\mu(E) = \infty$, there is an increasing sequence of sets E_n with $\mu(E_n) < \infty$ whose union is E and such that $\lim_{n\to\infty} \mu(E_n) = \infty$. By what was proved above, within each E_n there exists a compact K_n with $\mu(K_n) \geq \mu(E_n) - 1$. Each K_n is contained in E and $\lim_{n\to\infty} \mu(K_n) = \infty$ □

0.4 The Riesz-Markov Theorem for locally compact Hausdorff spaces

Throughout this section, let $(X,\mathcal{U})$ be a locally compact Hausdorff space. The spaces $\mathcal{C}_c(X)$, $\mathcal{C}_0(X)$ and $\mathcal{C}_b(X)$ are more than topological spaces: They contain a distinguished *cone* of non-negative elements: We say that *a function f on X is non-negative* in case for each x, $f(x) \in \mathbb{R}$, and $f(x) \geq 0$. In this case we write $f \geq 0$.

Let L be a linear functional on $\mathcal{C}_c(X)$. We say that L *is a positive linear functional* on $\mathcal{C}_c(X)$ in case

$$f \geq 0 \quad \Rightarrow \quad L(f) \geq 0 \ .$$

Evidently. if f is real, $L(f)$ is real, and $|L(f)| \leq L(|f|)$. If $f = g + ih$ where g and h are real, $|L(f)| = |L(g) + iL(h)| \leq L(|f|)$.

There is a close connection between the topology on $\mathcal{C}_c(X)$ and the partial order structure on $\mathcal{C}_c(X)$ induced by its cone of positive elements.

0.13 THEOREM. *Let L be a positive linear functional on $\mathcal{C}_c(X)$. Then for each compact $K \subset X$, there exists a finite constant C_K such that*

$$|f| \prec K \quad \Rightarrow \quad |L(f)| \leq C_K \|f\|_\infty \ .$$

Proof. The uniqueness is immediate from Theorem 0.11, and we turn to existence. By Lemma 0.3, there exists an open set U with compact closure $\overline{U}$ such that $K \subset U \subset \overline{U}$, and then by Urysohn's Lemma, there exists a continuous function φ on X such that $K \prec \varphi \prec U$.

Then evidently $\|f\|_\infty\varphi - |f| \geq 0$, and hence $L(\|f\|_\infty\varphi - |f|) \geq 0$. Thus,

$$|L(f)| \leq L(|f|) \leq L(\|f\|_\infty\varphi) = L(\varphi)\|f\|_\infty \ .$$

Thus,

$$\sup\{|L(f)|\ :\ |f| \prec K\ , \|f\|_\infty \leq 1\ \} \leq L(\varphi)\|f\|_\infty \ .$$

We may take $C_K = L(f)$, or, better yet, $C_K = \inf\{L(f)\ :\ K \prec f\ \}$. □

To construct an example of a positive linear functional on $\mathcal{C}_c(X)$, let μ be a Borel measure on X that is finite on every compact set $K \subset X$. If $f \in \mathcal{C}_c(X)$, let K denote the compact support of f. Then $0 \leq |f| \leq \|f\|_\infty 1_K$, and since $\mu(K) < \infty$, f is integrable. Thus, we may define

$$L_\mu(f) = \int_X f \mathrm{d}\mu \ ,$$

which is therefore a linear functional on $\mathcal{C}_c(X)$, and evidently it is positive.

The Reisz-Markov Theorem asserts that *every* example is of this type. Moreover, we shall show that one consequence of the Reisz-Markov Theorem is that every Borel measure on X such that $\mu(K) < \infty$ for all compact K automatically has certain regularity properties:

0.14 THEOREM (Riesz-Markov Theorem)**.** *Let L be any positive linear functional on $\mathcal{C}_c(X)$. Then there exists a unique Radon measure μ such that*

$$L(f) = \int_X f \mathrm{d}\mu \quad \text{for all} \quad f \in \mathcal{C}_c(X) \ . \tag{0.5}$$

Moreover,

$$\mu(U) = \sup\{ \ L(f) \ : \ f \prec U \ , f \in \mathcal{C}_c(X) \ \} \tag{0.6}$$

for all open sets U, and

$$\mu(K) = \inf\{ \ L(f) \ : \ K \prec f \ , f \in \mathcal{C}_c(X) \ \} \tag{0.7}$$

for all compact sets K.

Proof. Step 1: Use L to construct an outer measure μ^*. We define a set function μ^* on open subsets of X by

$$\mu^*(U) = \sup\{ \ L(f) \ : \ f \prec U \ , f \in \mathcal{C}_c(X) \ \}$$

for open sets U, and then on arbitrary subsets E of X by

$$\mu^*(E) = \inf\{ \ \mu^*(U) \ : \ E \subset U \ , \ U \text{ open } \} \ .$$

It is clear that $\mu^*(\emptyset) = 0$, and that if $A \subset B$, then $\mu^*(A) \le \mu^*(B)$. Therefore, to show that μ^* is an outer measure, we must show that for any sequence $\{E_n\}_{n\in\mathbb{N}}$ of subsets of X,

$$\mu^*\left(\bigcup_{n=1}^\infty E_n \right) \le \sum_{n=1}^\infty \mu^*(E_n) \ .$$

Let E denote $\cup_{n=1}^\infty E_n$. It suffices to consider the case in which $\mu^*(E_n) < \infty$ for all n.

Pick any $\epsilon > 0$. Then by construction, there exists an open set U_n with $E_n \subset U_n$, and $\mu^*(U_n) \le \mu^*(E_n) + 2^{-n}\epsilon$. But then

$$E \subset U := \bigcup_{n=1}^\infty U_n \qquad \text{and} \qquad \sum_{n=1}^\infty \mu^*(U_n) \le \sum_{n=1}^\infty \mu^*(E_n) + \epsilon \ .$$

It therefore suffices to prove that

$$\mu^*\left(U\right) \le \sum_{n=1}^\infty \mu^*(U_n) \ . \tag{0.8}$$

To do this, consider any $f \in \mathcal{C}_c$ such that $f \prec U$. Let K denote the support of f. Then $\{U_n\}_{n\in\mathbb{N}}$ is an open cover of K, and so there exists a finite sub cover, which we may take to be $\{U_1, \dots, U_N\}$.

Let $\{h_1, \ldots, h_N\}$ be a partition of unity on K, subordinate to the open cover $\{U_1, \ldots, U_N\}$. Then $f = \sum_{n=1}^{N} fh_n$ and $fh_n \prec U_n \quad n = 1, \ldots, N$. It follows that

$$L(f) = \sum_{n=1}^{N} L(fh_n) \leq \sum_{n=1}^{N} \mu^*(U_n) \leq \sum_{n=1}^{\infty} \mu^*(U_n) \ .$$

Since, $f \in \mathcal{C}_c$ such that $f \prec U$ is arbitrary, we obtain (0.8).

Step 2: Caratheodory σ-algebra contains all open sets, and hence all Borel sets.

Let U be open, and let $E \subset X$ be arbitrary. We must show that

$$\mu^*(E) \geq \mu^*(E \cap U) + \mu^*(E \cap U^c) \ . \tag{0.9}$$

Since $\mu^*(E)$ is the infimum of $\mu^*(V)$, V open with $E \subset V$, it suffices to prove (0.9) when $E = V$, where V is open. We may also suppose that $\mu^*(V) < \infty$. Note that $U \cap V$ is open. Hence for any $\epsilon > 0$, there is an $f \prec V \cap U$ so that $L(f) \geq \mu^*(V \cap U) - \epsilon$. Let K denote the support of f. Since $K \subset U$, $U^c \subset K^c$, and so $V \cap U^c \subset V \cap K^c$, which is open. Choose $g \prec V \cap K^c$ so that $L(g) \geq \mu^*(V \cap K^c) - \epsilon$. Then $f + g$ has compact support contained in V and since the supports of f and g are disjoint, $f + g \prec V$. Therefore,

$$\mu^*(V) \geq L(f+g) = L(f) + L(g) \geq \mu^*(V \cap U) + \mu^*(V \cap U^c) - 2\epsilon \ .$$

At this point, we know that the Caratheodory σ-algebra contains the Borel σ-algebra $\mathcal{B}(X)$, and that the restriction of μ^* to $\mathcal{B}(X)$ is countably additive. We define μ to be this restriction.

Step 3: Compact sets have finite measure

Let U be any open set containing K such that U has compact closure. Such sets exist by Lemma 0.3. By Urysohn's Lemma, there exists an $f \in \mathcal{C}_c(X)$ such that $\overline{U} \prec f$. Thus, if $g \prec U$, $g \leq f$, and so $L(g) \leq L(f)$, and hence $\mu(U) \leq L(f)$ since $g \prec U$ is arbitrary. Therefore, $\mu(K) \leq \mu(U) < \infty$.

Step 4: μ is inner regular for open sets

Let K be compact. We claim that for K compact,

$$\mu(K) = \inf\{L(f) \ : \ K \prec f \ \} \tag{0.10}$$

Fix $\epsilon > 0$. Then there exists an open set U so that $K \subset U$ and $\mu(K) \geq \mu(U) - \epsilon$. By Urysohn's Lemma, there exists a function f with $K \prec f \prec U$. Then $L(f) \leq \mu(U) \leq \mu(K) + \epsilon$. Since $\epsilon > 0$ is arbitrary,

$$\mu(K) \geq \inf\{L(f) \ : \ K \prec f \ \} \ .$$

Next, suppose that $K \prec f$. Then for any $0 < \epsilon < 1$, Let $U_\epsilon = \{x \ : \ f(x) > 1 = \epsilon\}$. Then U_ϵ is open and has compact support, and $K \subset U_\epsilon$.

There exists a function g such that $g \prec U_\epsilon$ and $L(g) \geq \mu(U_\epsilon) - \epsilon$. Note that $f \geq fg \geq (1-\epsilon)g$. Therefore

$$L(f) \geq (1-\epsilon)L(g) \geq (1-\epsilon)[\mu(U_\epsilon) - \epsilon] \geq (1-\epsilon)[\mu(K) - \epsilon] \ .$$

Since $0 < \epsilon < 1$ is arbitrary, $\mu(K) \le L(f)$ whenever $K \prec f$. This completes the proof of (0.10).

Now let V be open. Suppose $\mu(V) < \infty$. Then for every $\epsilon > 0$, there exists a function $f \prec V$ such that $\mu(V) - \epsilon \le L(f)$. Let K be the support of f. By (0.10), there is a function g with $K \prec g$ such that $L(g) \le \mu(K) + \epsilon$. But since $f \le 1_K \le g$,

$$\mu(V) - \epsilon \le L(f) \le L(g) \le \mu(K) + \epsilon\ .$$

The same sort of reasoning shows that if $\mu(V)$ is infinite, we can find compact sets in V of arbitrarily large measure.

Step 5: For $f \in \mathcal{C}_c(X)$, $\int_X f\mathrm{d}\mu = L(F)$.

It suffices to treat the case in which f takes values in $[0,1]$. For each $t \in (0,1]$, define $K_t = \{x\ :\ f(x) \ge t\}$. Since $f \in \mathcal{C}_c(X)$, each K_t is compact Let K_0 denote the support of f.

For any $0 \le a < b \le 1$, define $f_{[a,b]} = (f-a)_+ \wedge b$. Then for all $0 < t < a$,

$$K_b \prec \frac{1}{b-a} f_{[a,b]} \le 1_{K_a}\ .$$

Therefore, for all open U with $K_a \subset U$, $\mu(K_b) \le \dfrac{1}{b-a} L(f_{[a,b]})\mu(U)$. By outer regularity,

$$\mu(K_b) \le \frac{1}{b-a} L(f_{[a,b]}) \le \mu(K_a)\ .$$

Then for any $n \in \mathbb{N}$, $f = \displaystyle\sum_{j=1}^n f_{[(j-1)/n, j/n]}$, and so $L(f) = \displaystyle\sum_{j=1}^n L\left(f_{[(j-1)/n, j/n]}\right)$. Therefore,

$$\sum_{j=1}^n \frac{1}{n}\mu(K_{j/n}) \le L(f) \le \sum_{j=1}^n \frac{1}{n}\mu(K_{(j-1)/n})\ .$$

Since $\lim_{n\to\infty}\left(\sum_{j=1}^n \frac{1}{n}\mu(K_{j/n})\right) = \lim_{n\to\infty}\left(\sum_{j=1}^n \frac{1}{n}\mu(K_{j/n})\right) = \int_X f\mathrm{d}\mu$, the proof is complete. □

0.15 DEFINITION (Radon space)**.** A topological space $(X, \mathcal{U})$ is a *Radon space* in case every Borel measure on X that is finite on all compact sets is regular, and hence a Radon measure.

0.16 DEFINITION (Polish space)**.** A topological space $(X, \mathcal{U})$ is a *Polish space* in case it is separable and homeomorphic to a metric space.

The term “Polish space” recognizes the work of a group of Polish mathematicians including Kuratowski, Sierpinski and Tarski, who proved a number results pertaining to the concept. Note that a Polish space is necessarily Hausdorff. Therefore, a locally compact Polish space is, in particular, a locally compact Hausdorff space. For example, $\mathbb{R}^n$ is a locally compact Polish space, as is any Riemannain manifold.

0.17 LEMMA. *Let $(X, \mathcal{U})$ be a locally compact Polish space. Then every open set U in X is the countable union of compact sets in X. In particular, X is the countable union of compact sets.*

Proof. Let ρ be a metric that induces the topology on $(X,\mathcal{U})$, and let $B(r,x)$ denote the open ball of radius $r > 0$ about $x \in X$. Since $(X,\mathcal{U})$ is locally compact, for each $x \in X$, there exists some $r > 0$ $\overline{B(r,x)}$ is compact . (It is not excluded that $\overline{B(r,x)}$ is compact for *all* $r > 0$; this is the case in $\mathbb{R}^n$, for example.) Define $r_x = \sup\{r > 0 \ : \ \overline{B(r,x)} \text{ is compact}\}$. If $0 < \delta < r_x$ and $\rho(y,x) < \delta$, then for all r with $\delta < r < r_x$, $B(r-\delta,y) \subset B(r,x)$ and $\overline{B(r,x)}$ is compact. Hence $r_y \geq r_x - \delta$.

Likewise, given the open set U and $x \in U$ define $d_x = \sup\{s > 0 \ : \ B(s,x) \subset U\}$. That is, d_x is the distance from x to U^c. Just as above, one shows that if $\rho(y,x) < \delta < d_x$, then $d_y \geq d_x - \delta$.

Let $\{x_n\}_{n\in\mathbb{N}}$ be a dense sequence in the open set U. For $n \in \mathbb{N}$, define

$$K_n := \overline{B(\min\{d_{x_n} + r_{x_n}\}/2, x_n)} \ .$$

Then K_n is compact and contained in U. For any $x \in U$, there is some n such that $\rho(x_n,x) < \frac{1}{4}\min\{d_x, r_x\}$. Then $r_{x_n} \geq \frac{3}{4}r_x$ and $d_{x_n} \geq \frac{3}{4}d_x$, so that $\min\{d_{x_n} + r_{x_n}\}/2 > \rho(x_n,x)$. It follows that $x \in K_n$, Since $x \in U$ is arbitrary, $U = \cup_{n=1}^\infty K_n$. □

0.18 THEOREM. *Let $(X,\mathcal{U})$ be a locally compact Polish space. Then $(X,\mathcal{U})$ is a Radon space.*

Proof. Let μ be any Borel measure on X that is finite on every compact set. Then each $f \in C_c(X)$ is integrable with respect to μ, and hence $f \mapsto \int_X f\mathrm{d}\mu =: L(f)$ is a well-defined positive linear functional on $C_c(X)$. By the Riesz-Markov Theorem, there exists a Radon measure ν such that for all $f \in C_c(X)$,

$$\int_X f\mathrm{d}\mu = \int_X f\mathrm{d}\nu \ . \tag{0.11}$$

Since by Lemma 0.17, $(X,\mathcal{B},\nu)$ is σ-finite, Theorem 0.12 then says that ν is regular. We next show that μ and ν agree on all open sets U.

Let U be open and write $U = \cup_{n=1}^\infty K_n$ where K_n is compact, which is possible by Lemma 0.17. By Uryson's lemma, for each $n \in \mathbb{N}$ there exists $f_n \in C_c(X)$ such that $\cup_{j=1}^n K_n \prec f_n \prec U$, Then define $g_n := \max\{f_1,\dots,f_n\}$. Then $g_n \uparrow 1_U$.

We have seen that there is a monotone increasing sequence $\{g_n\}_{n\in\mathbb{N}}$ with $g_n \uparrow 1_U$. By the Lebesgue Monotone Convergence Theorem and (0.11),

$$\mu(U) = \lim_{n\to\infty}\int_X g_n\mathrm{d}\mu = \lim_{n\to\infty}\int_X g_n\mathrm{d}\nu = \nu(U) \ .$$

Hence μ and ν agree on all open sets.

Now let K be compact. By Lemma 0.3, there is an open set V such that $K \subset V$ and $\overline{V}$ is compact. It follows that $\nu(V) = \mu(V) < \infty$. Then since $K = V\backslash(V\backslash K)$, and since V and $V\backslash K$ are both open

$$\mu(K) = \mu(V) - \mu(V\backslash K) = \nu(V) - \nu(V\backslash K) = \nu(K) \ .$$

Hence μ and ν agree on all compact sets. Let E be any Borel set. Since ν is regular,

$$\nu(E) = \sup\{\nu(K) \ : \ K \subset E \ , K \text{ compact } \} = \sup\{\mu(K) \ : \ K \subset E \ , K \text{ compact } \} \leq \mu(E) \ .$$

Hence if $\nu(E)$ is infinite, so is $\mu(E)$. Suppose that $\nu(E) < \infty$, and pick $\epsilon > 0$. By the regularity of ν, there exist K compact and U open such that $K \subset E \subset U$ and $\nu(U) - \nu(K) < \epsilon$. Then because $\mu(K) = \nu(K)$ and $\mu(U) = \nu(U)$, both $\mu(E)$ and $\nu(E)$ lie in the interval $[\nu(K), \nu(K)+\epsilon]$, and hence $|\mu(E) - \nu(E)| \leq \epsilon$. Since $\epsilon > 0$, $\mu(E) = \nu(E)$. Hence μ and ν agree on all Borel sets. □

0.5 The Hahn-Saks Theorem

Throughout this section, $(X,\mathcal{U})$ is a locally compact Hausdorff space, $\mathcal{C}_c(X)$ denotes the *real* normed space of continuous compactly supported *real valued* functions on X, and $\mathcal{C}_0(X)$ denotes the *real* Banach space consisting of all uniform limits of functions in $\mathcal{C}_c(X)$.

The real vector spaces $\mathcal{C}_c(X)$ and $\mathcal{C}_0(X)$ are *ordered* vector spaces: We say $f \geq g$ in case $f(x) - g(x) \geq 0$ for all x. Let $\mathcal{C}_c^+(X)$ and $\mathcal{C}_0^+(X)$ denote the sets of point-wise non-negative functions in $\mathcal{C}_c(X)$ and $\mathcal{C}_0(X)$ respectively. Then $f \geq g$ in $\mathcal{C}_c(X)$ if and only if $f - g \in \mathcal{C}_c^+(X)$, and likewise for the order in $\mathcal{C}_0(X)$.

Our goal is to concretely identify the elements of the dual space $(\mathcal{C}_0(X))^*$ in terms of measures on X. An element of $(\mathcal{C}_0(X))^*$ is *positive* in case $L(f) \geq 0$ wherever f is a non-negative function in $\mathbb{C}_0(X)$. Restricting such a functional L to $C_c(X)$, yields a positive linear functional on $C_c(X)$, and then by the Riesz-Markoff Theorem, there is a Radon measure μ_L on X and such that

$$L(f) = \int_X f \mathrm{d}\mu_L \tag{0.12}$$

for all $f \in \mathcal{C}_c(X)$.

0.19 LEMMA. *Let L be a positive linear functional on $\mathbb{C}_0(X)$, and let μ_L be the Radon measure on X such that (0.12) is valid for all $f \in \mathbb{C}_c(X)$. Then $\mu_L(X) \leq \|L\|$, and hence (0.12) extends by continuity to all of $C_0(X)$.*

Proof. Since μ_L is inner regular on open sets, there exists an increasing sequence $\{K_n\}_{n\in\mathbb{N}}$ such that $\mu_L(K_n) \uparrow \mu_L(X)$. By Urysohn's Lemma, for each n there exists $f_n \in V_c(X)$ such that $K_n \prec f_n$. Then $\mu_L(K_n) \leq \displaystyle\int_X f_n \mathrm{d}\mu_L = L(f_n) \leq \|L\|$. □

Suppose that $L \in (\mathcal{C}_0(X))^*$ can be written as the difference of two *positive* linear functionals,

$$L = L_1 - L_2 \ . \tag{0.13}$$

Then by Lemma 0.19, there are finite Radon measures μ_1 and μ_2 such that for all $f \in C_0(X)$,

$$L(f) = \int_X f \mathrm{d}\mu_1 - \int_X f \mathrm{d}\mu_2 \ . \tag{0.14}$$

Define $\nu = \mu_1 + \mu_2$, which is also a finite Radon measure. Evidently μ_1 and μ_2 are absolutely continuous with respect to ν, and hence there exist non-negative functions $h_1, h_2 \in L^1(X, \mathcal{B}, \nu)$ such that $\mathrm{d}\mu_1 = h_1 \mathrm{d}\nu$ and $\mathrm{d}\mu_2 = h_2 \mathrm{d}\nu$. Define $\widetilde{h}_1 := h_1 - h_1 \wedge h_2$ and $\widetilde{h}_2 := h_2 - h_1 \wedge h_2$. Let $A = \{x \ : \ h_1 \wedge h_2(x) = h_1(x)\}$ and observe that $\widetilde{h}_1(x) = 0$ for all $x \in A$, while $\widetilde{h}_2(x) = 0$ for all $x \in A^c$. Then since $h_1 - h_2 = \widetilde{h}_1 - \widetilde{h}_2$, (0.14) becomes

$$L(f) = \int_X f h_1 \mathrm{d}\nu - \int_X f h_2 \mathrm{d}\nu = \int_X f(\widetilde{h}_1 - \widetilde{h}_2) \mathrm{d}\nu \ . \tag{0.15}$$

Define measures $\mathrm{d}\widetilde{\mu}_1 := \widetilde{h}_1 \mathrm{d}\nu$ and $\widetilde{\mu}_2 := \widetilde{h}_2 \mathrm{d}\nu$. Then $\widetilde{\mu}_1(A) = 0$ and $\widetilde{\mu}_2(A^c) = 0$ so that $\widetilde{\mu}_1$ and $\widetilde{\mu}_2$ are mutually singular. Since any Borel measure that is absolutely continuous with respect to a Radon measure is itself as Radon measure, $\widetilde{\mu}_1$ and $\widetilde{\mu}_2$ are also Radon measures.

In summary, if there exists a decomposition of $L \in (\mathcal{C}_0(X))^*$ as the difference of two positive elements of $(\mathcal{C}_0(X))^*$, then there exists a pair of *mutually singular* finite Radon measures μ_+ and μ_- such that for all $f \in C_0(X)$,

$$L(f) = \int_X f\mathrm{d}\mu_+ - \int_X f\mathrm{d}\mu_- \ . \tag{0.16}$$

Such a decomposition is necessarily unique: Let A be a Borel set such that $\mu_+(A) = 0$ and $\mu_-(A^c) = 0$. Pick $\epsilon > 0$, and then $h \in C_c(X)$ such that such that $\|h - 1_A\|_{L^1(X,\mathcal{B},\mu_++\mu_-)} < \epsilon$. Without loss of generality, we may assume that $0 \le h \le 1$. Then for any $0 \le g \le f$ in $C_0(X)$, by the choice of h,

$$\begin{aligned} L(g) &= L((1-h)g) + L(hg) \\ &= \int_X (1-h)g\mathrm{d}\mu_+ - \int_X (1-h)g\mathrm{d}\mu_- + \int_X hg\mathrm{d}\mu_+ - \int_X hg\mathrm{d}\mu_- \\ &\le \int_X (1-h)g\mathrm{d}\mu_+ + \|f\|_\infty\epsilon \\ &\le \int_X (1-h)f\mathrm{d}\mu_+ + \|f\|_\infty\epsilon \le \int_X f\mathrm{d}\mu_+ + 2\|f\|_\infty\epsilon \end{aligned}$$

This shows that

$$\sup\{\ L(g)\ :\ 0 \le g \le f\ \} \le \int_X f\mathrm{d}\mu_+ \ . \tag{0.17}$$

On the other hand,

$$L((1-h)f) = \int_X (1-h)f\mathrm{d}\mu_+ - \int_X (1-h)f\mathrm{d}\mu_- \ge \int_X f\mathrm{d}\mu_+ - 2\|f\|_\infty\epsilon \ .$$

Since $0 \le (1-h)f \le f$, equality holds in (0.17) and therefore, the linear functional $f \mapsto \int_X f\mathrm{d}\mu_+$ is uniquely determined by L under the assumption that (0.16) with μ_+ and μ_- mutually singular. Then by Theorem 0.11, μ_+ is then uniquely determined by L. We have proved:

0.20 LEMMA. *Let $L \in (\mathcal{C}_0(X))^*$ have a decomposition as the difference of two* positive *linear functionals as in (0.14). Then there is a unique pair of mutually singular Radon measures μ_+, μ_- such that (0.16) is valid for all $f \in C_0(X)$, and moreover, μ_+ is determined through L by (0.17).*

It is now a simple matter to show that, in fact, *every* $L \in (\mathcal{C}_0(X))^*$ can be written as the difference of two *positive* linear functionals as in (0.14). The identity (0.17) gives us a candidate for the components of the decomposition:

0.21 LEMMA. *Let $L \in (\mathcal{C}_0(X))^*$. For $f \in \mathcal{C}_0^+(X)$, define*

$$L_+(f) = \sup\{\ L(g)\ :\ 0 \le g \le f\ \} \ . \tag{0.18}$$

For general $f \in \mathcal{C}_0(X)$, define

$$L_+(f) = L_+(f_+) - L_+(f_-) \ . \tag{0.19}$$

where f_+ and f_- are, respectively, the positive and negative parts of f. Then $L_+(f) \in (\mathcal{C}_0(X))^*$, and both L_+ and $L_- := L - L_+$ are positive, so that $L = L_+ - L_-$ is a decomposition of L into the difference between two positive linear functionals.

Proof. We first show that for $f_1, f_2 \in \mathcal{C}_0^+(X)$, $L_+(f_1+f_2) = L_+(f_1)+L_+(f_2)$. First, let $0 \le g_1 \le f_1$ and $0 \le g_2 \le f_2$. Then $0 \le g_1 + g_2 \le f_1 + f_2$ so that $L_+(f_1 + f_2) \ge L(g_1 + g_2) = L(g_1) + L(g_2)$ Taking the supremum over $g_1 \le f_1$ and $g_2 \le f_2$ yields $L_+(f_1 + f_2) \ge L_+(f_1) + L_+(f_2)$.

Fix $\epsilon > 0$, and choose $g \in \mathcal{C}_0^+(X)$ with $0 \le g \le f_1 + f_2$ so that $L_+(f_1 + f_2) \le L(g) + \epsilon$. Define $g_1 = f_1 \wedge g$ and $g_2 = g - g_1$. Then $0 \le g_1 \le f_1$, $0 \le g_2 \le f_2$ and $g_1 + g_2 = g$. Therefore

$$L_+(f_1 + f_2) \le L(g) + \epsilon \le L(g_1) + L(g_2) + \epsilon \le L_+(f_1) + L_+(f_2) + \epsilon \ .$$

Since $\epsilon > 0$ is arbitrary, $L_+(f_1+f_2) \le L_+(f_1)+L_+(f_2)$. Together with what we proved above, this shows $L_+(f_1 + f_2) = L_+(f_1) + L_+(f_2)$.

To see that L_+ as extended to all of $C_0(X)$ by (0.19), is additive, we observe that if $f = g_1 - g_2$ and $f = h_1 - h_2$ are two ways of writing $f \in C_0(X)$ as a difference of two elements in $C_0^+(X)$, then $g_1 + h_2 = h_1 + g_2$. By what was proved above,

$$L_+(g_1) + L_+(h_2) = L_+(g_1 + h_2) = L_+(h_1 + g_2) = L_+(h_1) + L_+(g_2) \ ,$$

and hence

$$L_+(g_1) - L_+(g_2) = L_+(h_1) - L_+(h_2) \ . \tag{0.20}$$

Therefore, one can replace the specific decomposition $f = f_+ - f_-$ in (0.19) by any other decomposition of f into the difference between elements of $C_0^+(X)$, and the result is the same. Then for $f, g \in \mathcal{C}_0(X)$, $f + g = (f_+ + g_+) - (f_- + g_-)$ is one way of $f + g$ as the difference of elements of $\mathcal{C}_0^+(X)$,

$$L_+(f+g) = L_+(f_++g_+)-L_+(f_-+g_-) = (L_+(f_+)-L_+(f_-))-(L_+(g_+)-L_+(g_-)) = L_+(f)+L_+(g) \ .$$

This together with the evident fact that $L_+(\alpha f) = \alpha L_+(f)$ for all $\alpha \in \mathbb{R}$ and $f \in \mathcal{C}_0(X)$ shows that L_+ is a linear functional.

By (0.19) and then (0.18) that

$$\|L_+\| = \sup\{L_+(f) \ : \ 0 \le f \le 1\} \le \sup\{L(g) \ : \ 0 \le g \le 1\} \le \|L\| \ .$$

Thus, L_+ is a bounded linear functional and It is evident from (0.18) that $L_+(f) \ge 0$ for all $f \in \mathcal{C}_0(X)$, so that it is also positive.

Finally, defining $L_- = L_+ - L$, we have that for all $f \in \mathcal{C}_0^+(X)$, $L_-(f) = L_+(f) - L(f) \ge 0$ since $L_+(f) \ge L(f)$. Hence L_- is a positive linear functional. Since $L = L_+ - L_-$, this shows that every element L of $(\mathcal{C}_0(X))^*$ can be written as the difference of two positive linear functionals on $\mathcal{C}_0(X)$. □

0.22 LEMMA. *Let $L \in (C_0(X))^*$, and let μ_+, μ_- be the unique pair of mutually singular Radon measures on X such that (0.16) is valid for all $f \in C_0(X)$. Then*

$$\|L\| := \mu_+(X) + \mu_-(X) \ . \tag{0.21}$$

Proof. For all $f \in C_0(X)$ with $\|f\|_\infty \le 1$, $|L(f)| \le \int_X |f| \mathrm{d}\mu_+ + \int_X |f| \mathrm{d}\mu_- \le \mu_+(X) + \mu_-(X)$. On the other hand, let A be a Borel set such that $\mu_+(A) = 0$ and $\mu_-(A^c) = 0$. Pick $\epsilon > 0$, and then

$h \in C_c(X)$ such that such that $\|h - (1_{A^c} - 1_A)\|_{L^1(X,\mathcal{B},\mu_+ + \mu_-)} < \epsilon$. Without loss of generality, we may assume that $-1 \leq h \leq 1$. Then

$$\begin{aligned} \|L\| \geq L(h) &= \int_X h \mathrm{d}\mu_+ - \int_X h \mathrm{d}\mu_- \\ &\geq \int_X (1_{A^c} - 1_A)\mathrm{d}\mu_+ - \int_X (1_{A^c} - 1_A)\mathrm{d}\mu_- - 2\epsilon = \mu_+(X) + \mu_-(X) - 2\epsilon \ . \end{aligned}$$

□

Collecting results from the lemmas, we have proved:

0.23 THEOREM (Riesz Representation Theorem for $((C_0(X))^*)$**.** *Let $(X, \mathcal{U})$ be a locally compact Hausdorff space. Then for each $L \in ((C_0(X))^*$ there exists a unique pair of mutually singular finite Radon measures μ_+ and μ_- such that (0.16) is valid for all $f \in C_0(X)$ and such that the norm $\|L\|$ of L is given by (0.21).*

A number of consequences of this theorem deserve further discussion. We begin with a definition:

0.24 DEFINITION. A *signed measure* on X is a real valued function μ on the Borel σ–algebra of X such that there exist two positive finite Borel measures μ_1 and μ_2 such that for all Borel sets E, $\mu(E) = \mu_1(E) - \mu_2(E)$.

The set of signed measures is evidently a real vector space. (Complex measures are defined in the analogous way, and would constitute a complex measure space.) We denote the real vector space of signed measures on X by $\mathcal{M}(X)$.

For a bounded Borel function f, define the integral $\int_X f \mathrm{d}\mu$ by $\displaystyle\int_X f \mathrm{d}\mu = \int_X f \mathrm{d}\mu_1 - \int_X f \mathrm{d}\mu_2$. This gives us a continuous linear functional L on $\mathcal{C}_0(X)$ where $L(f) = \displaystyle\int_X f \mathrm{d}\mu$. Theorem 0.23 then gives us the existence of uniquely determined positive Borel measures μ_+ and μ_- that are mutually singular – i.e., supported on disjoint sets – and such that

$$\mu(E) = \mu_+(E) - \mu_-(E)$$

for all Borel sets E in X.

0.25 DEFINITION. For any signed measure μ, the positive measure $|\mu|$ given by

$$|\mu| = \mu_+ + \mu_-$$

is called the *total variation measure* of μ, and and the function $\mu \mapsto \|\mu\|_{\mathrm{TV}}$ where

$$\|\mu\|_{\mathrm{TV}} = \mu_+(X) + \mu_-(X)$$

is called the *total variation norm* of μ.

It is easy to see from our analysis above that

$$\|\mu\|_{\mathrm{TV}} = \sup \left\{ \int_X f \mathrm{d}\mu \ \middle| \ f \in C_c(X) \ , \ -1 \leq f \leq 1 \right\} \ ,$$

and from this the Minkowski inequality is easily seen to hold, so that $\|\cdot\|_{\rm TV}$ is actually a norm, as the name indicates.

We know that the dual of a Banach space is complete in the dual norm, and so $\mathcal{M}(X)$ is complete in the total variation norm. Moreover, the map $L \mapsto \mu_+ - \mu_-$ where μ_+ and μ_- are related to L as in Theorem 0.23 is evidently an isometric isomorphism of $(\mathcal{C}_0(X))^*$ onto $\mathcal{M}(X)$.

The Banach space $C_0(X)$ is not reflexive expect when X is very simple. As long as there exists a single Borel set E that is not both open and closed, we may define a linear functional on $\mathcal{M}$ by

$$\Lambda(\mu) = \mu(E) = \int_X 1_E {\rm d}\mu\ .$$

Clearly Λ is linear and

$$|\Lambda(E)| \le |\mu|(E) \le |\mu|(X) = \|\mu\|_{\rm TV}\ ,$$

so Λ is indeed bounded, and so is an element of $(\mathcal{M}(X))^*$. But if there were a function $g \in \mathcal{C}_0(X)$ for which $\Lambda(\mu) = \int_X g{\rm d}\mu$ for all $\mu \in \mathcal{M}$, we would have $g(x) = 1_E(x)$ for all x since the point mass δ_x that concentrates unit mass at x belongs to $\mathcal{M}$ for each $x \in X$. But since E is not both open and closed, 1_E is not continuous.

0.6 Wiener measure

In this section, Ω denotes the set of continuous functions $\omega : [0,\infty) \to \mathbb{R}$ such that $\omega(0) = 0$. We equip Ω with the topology of uniform convergence on compact subsets of $[0,\infty)$. This makes it a Frechét space, and in particular, a Hausdorff space. However, Ω is not locally compact. For each $t \in [0,\infty)$, let X_t denote the *evaluation functional* $X_t(\omega) = \omega(t)$. Let $\mathcal{F}$ be the σ-algebra on Ω generated by the X_t, $t \ge 0$. *Wiener measure* is a probability measure ν_W on $(\Omega,\mathcal{F})$ such that if $E_1,\dots,E_n$ are Borel sets in $\mathbb{R}$, and $0 \le t_1 < \cdots < t_n$, a particle performing "Browninan motion" starting from x at time $t = 0$ is in E_j at time t_j for each $j = 1,\dots,n$ with probability

$$\nu_W(\{\omega \in \Omega\ :\ \omega(t_j) \in E_j\ ,\quad j = 1,\dots,n\ \}) = \int_{E_1\times\cdot\times E_n} \prod_{j=1}^n \gamma_{t_j - t_j - 1}(x_{j-1} - x_j){\rm d}x_1\cdots{\rm d}x_n \qquad (0.22)$$

where $\gamma_t(x)$ is the Gaussian probability density used to define the heat semigroup, $t_0 := 0$ and $x_0 := 0$. (We restrict ourselves to one dimension only to keep the notation simple. Everything we say in this section about "Brownian motion" in $\mathbb{R}$ extends readily to Brownian motion in $\mathbb{R}^n$ for any $n \in \mathbb{N}$.)

The formula on the right side of (0.22) for the probabilities of such events follows from Einstein's work on Browninan motion in 1905, in which he related Brownian motion to diffusion and the heat equation. Einstein's precise explanation for Brownian motion in terms of molecular collisions – at a time when the very existence of atoms and molecules was still a matter of dispute – made it possible to determine Avogadro's number, the number of atoms of hydrogen in one gram of hydrogen, by making observations through a microscope of pollen-sized particles undergoing Brownian motion. This was actually done in 1908 by Jean Baptiste Perrin, and he was awarded the Nobel prize in 1926 for his experimental work. Einstein received the prize in 1921 for his theoretical work. A simplified version of Einstein's formulae, leaving out the constants that make it possible to determine Avogadro's number by looking though a microscope, says that the probability that a Brownian particle starting at 0 at time $t = 0$ is in E_j at time t_j is given in terms of the heat kernel $\gamma_t(x - y)$ by

$$\int_{E_1 \times \cdot \times E_n} \prod_{j=1}^{n} \gamma_{t_j - t_j - 1}(x_{j-1} - x_j) \mathrm{d}x_1 \cdots \mathrm{d}x_n \tag{0.23}$$

where $t_0 := 0$ and $x_0 := 0$, which is the formula on the right side of (0.22). For example if $n = 2$, this reduces to, using the Fubini-Tonelli Theorem,

$$\int_{E_1} \gamma_{t_1}(x) \left(\int_{E_2} \gamma_{t_2 - t_1}(x - y) \mathrm{d}y \right) \mathrm{d}x \ .$$

The inner integral gives the probability of the Brownian particle making a transition from x to E_2 in time $t_2 - t_1$, and $\gamma_{t_1}(x)$ is the probability density for the Brownian particle making a transition from 0 to x in time t_1. The composite formula is justified on account of the increments of the motion being "statistically independent".

We shall not go further into the physical origins of the formula on the right side of (0.22), but turn to the mathematical question of whether or not there exists a probability measure ν_W on Ω, the infinite dimensional "path space" for the Brownian particle, such that (0.22) is valid.

The sets of the form $\{\omega \in \Omega \ : \ \omega(t_j) \in E_j \ , \quad j = 1, \dots, n \ \}$ are very special in $\mathcal{F}$, and the existence of a countably additive probability measure $(\Omega, \mathcal{F})$ that assigns the specified probabilities to these sets is far from trivial. It is therefore somewhat amazing that Norbert Wiener constructed the measure ν_W on $(\Omega, \mathcal{F})$ in 1923, well before Kolmogorov had even given his measure-theoretic formulation of probability theory.

In this section we give a proof of Wiener's Theorem due to Edward Nelson that makes use of the Riesz-Markov Theorem, and the regularity of the measures that it provides. There are two main parts to the proof: In the first part, we embed Ω into a *compact* Hausdorff space, and use the formula on the right hand side of (0.22) to define a positive linear functional on $\mathcal{C}(X) = \mathcal{C}_c(X)$. The Riesz-Markoff Theorem then provides a regular Borel probability measure μ_X on X. In the second second part, we show that Ω is a Borel set, and that $\mu_W(\Omega) = 1$. We then obtain ν_W by restricting μ_W to Ω. It is in the course of the proof that $\mu_W(\Omega) = 1$ that we make essential use of the inner regularity of μ_W to deal with the fact that continuity of a function ω from $[0, \infty)$ into $\mathbb{R}$ depends on the behavior of ω at uncountably many points.

This is not the only way to construct Wiener measure, and indeed Wiener's paper predates the work of Markoff on which this approach depends. However, it is flexible and powerful, and can be used to construct many other measures on "path spaces". It will be clear that very little specific information about the heat kernel is used in the proof.

Our first task, which is essentially notational, is to recast the formula (0.23) in terms of a probability measure on $\mathbb{R}^n$. Let S be an arbitrary finite subset of distinct elements t_j of $(0, \infty)$ arrange in increasing order: $S = \{t_1, \dots, t_n\}$ with $t_j \le t_{j+1}$ for $j = 1, \dots, n - 1$.

Given such a set S, define a measure $\mu_{x,S}$ on $\mathbb{R}^n$ by

$$\mathrm{d}\mu_{x,S} = \prod_{j=1}^{n} \gamma_{t_j - t_j - 1}(x_{j-1} - x_j) \mathrm{d}x_1 \cdots \mathrm{d}x_n \ . \tag{0.24}$$

where $t_0 := 0$ and $x_0 := x$. (We need the more general formula in which x_0 is arbitrary and not set equal to 0 for reasons that are explained below.) For example, if $S = \{t_1, t_2\}$, then

$$\mathrm{d}\mu_{x,\{t_1,t_2\}} = \gamma_{t_1}(x - x_1))\gamma(x_1 - x_2)\mathrm{d}x_1\mathrm{d}x_2 \ . \tag{0.25}$$

Integrating first in x_2, and then in x_1 and using the fact that $\int_{\mathbb{R}} \gamma_t(x)\mathrm{d}x = 1$ for all t, one readily sees that $\mu_{x,\{t_1,t_2\}}$ is a probability measure. The same reasoning shows that for all S, $\mu_{x,S}$ is a probability measure. For Borel set E_1 and E_2, $\mu_{x,\{t_1,t_2\}}(E_1 \times E_2)$ is the probability that the particle, initially at x, is in E_1 at time t_1, and then in E_2 at time t_2. Likewise,

$$\mu_{x,\{t_1,\dots,t_n\}}(E_1 \times \cdots \times E_n)$$

is to be thought of as the probability the particle, initially at x, is in E_j at time t_j for $j = 1, \dots, n$. Notice that if some $E_j = \mathbb{R}$, the condition that the particle is in E_j is vacuous given that the particle must be *somewhere* in $\mathbb{R}$. This corresponds to an important consistence condition of the family of measures $\{\mu_{x,S}\}$ indexed by the finite ordered sets S Let $S = \{t_1, \dots, t_n\}$ and for some $j = 1, \dots, n$, let $S' := S\backslash\{t_j\}$ ordered as above.

Let $E_1, \dots, E_n$ be n Borel sets in $\mathbb{R}$. Let $E = E_1 \times \cdot \times E_n$ and let E' be the Cartesian product of $\{E_1, \dots, E_n\}\backslash\{E_j\}$ taken in the order induced by the subscripts. Then since

$$\int_{\mathbb{R}} \gamma_{s-t_{j-1}}(x_{j-1} - y)\gamma_{t_j - s}(y - x_j)\mathrm{d}y = \gamma_{t_j - t_{j-1}}(x_{j-1} - x_j) \ ,$$

doing the integration over the x_j first on the left we find that in case $E_j = \mathbb{R}$,

$$\mu_{x,S}\left(E\right) = \mu_{x,S'}\left(E'\right) \ . \tag{0.26}$$

There is another important relation satisfied by these measures, this times with x and S both being variable. Let $S = \{t_1, \dots, t_n\}$ be given, $n \geq 2$. Let $1 \leq k \leq n-1$ be given and let f be a bounded Borel function on $\mathbb{R}^k$ and let g be a bounded Borel function on $\mathbb{R}^{n-k}$. Let $S' = \{t_1, \dots, t_k\}$ and let $S'' := \{t_{k+1}, \dots, t_n\}$. Then doing the integral over $x_{k+1}, \dots, x_n$ first, we find that

$$\int_{\mathbb{R}^n} f(x_1, \dots, x_k)g(x_{k+1}, \dots, x_n)\mathrm{d}\mu_{x,S} = \int_{\mathbb{R}^k} f(x_1, \dots, x_k)\left(\int_{R^{n-k}} g(x_{k+1}, \dots, x_n)\mathrm{d}\mu_{x_k,S''}\right)\mathrm{d}\mu_{x,S'} \ . \tag{0.27}$$

In probabilistic terms, one may regard the functions $X_j : (x_1, \dots, x_n) \mapsto x_j$, $j = 1, \dots, n$ as "random variables" on the probability space $(\mathbb{R}^n, \mathcal{B}, \mu_{x,S})$. The product structure in (0.27) can then be interpreted as saying that "given X_k, the future variables $X_{k+1}, \dots, X_n$ are statistically independent of the past variables'$X_1, \dots, X_{k-1}$." This is the *Markov property.* In what follows we do not need to know precisely what "given X_k" means; it involves the notion of conditional probability. We shall only make direct analytic use of the factorization identity (0.27). However, it would be a grave injustice not to at least mention mention in passing the Markov property at this point. For our purposes, the Markov property of the measure $\mu_{x,S}$ is precisely the factorization formula (0.27) relating it to the measures $\mu_{x,S'}$ and $\mu_{x_k,S''}$.

0.26 THEOREM. *There exists a unique probability measure ν_W on $(\Omega, \mathcal{F})$ such that if φ is any function on $\mathbb{R}^n$ and $S = \{t_1, \dots, t_n\}$, $0 < t_0 < \cdots < t_n$, then*

$$\int_{\Omega} \varphi(\omega(t_1), \dots, \omega(t_n))\mathrm{d}\nu_W(\omega) = \int_{\mathbb{R}^n} \varphi(x_1, \dots, x_n)\mathrm{d}\mu_{0,\{t_1,\dots,t_n\}} \ .$$

Proof of Theorem 0.26. Let $\dot{\mathbb{R}}$ denote the one-point compactification of $\mathbb{R}$, and let X denote the Cartesian product $(\dot{\mathbb{R}})^{[0,\infty)}$ with the product topology. Then by Tychonov's Theorem, X is a compact Hausdorff space. The general element ω of X is an arbitrary function from $[0,\infty)$ into $\dot{\mathbb{R}}$.

Let $\mathcal{A}$ denote the set of all functions f on X of the form

$$f(\omega) = \phi(\omega(t_1), \dots, \omega(t_n)) \tag{0.28}$$

for some $n \in \mathbb{N}$, some $S = \{t_1, \dots, t_n\}$, and some bounded continuous function $\varphi : \mathbb{R}^n \to \mathbb{R}$. Then $\mathcal{A}$ is an algebra consisting of continuous functions on X which contains the constant function 1 and separates points. By the Stone Wierstrass Theorem, it is dense in $\mathcal{C}(X)$.

We now define a linear functional L on $\mathcal{A}$ by

$$L(f) := \int_{\mathbb{R}}^{n} \phi(x_1, \dots, x_n) \mathrm{d}\mu_{0,\{t_1,\dots,t_n\}} \tag{0.29}$$

where f and ϕ are related by (0.28). Any given function in $\mathcal{A}$ has many different representations of the form (0.28) since one can always enlarge the set $\{t_1, \dots, t_n\}$ but then have ϕ depends trivially on the inserted coordinates. Because of the consistency relation (0.26), the right hand side of (0.29) is independent of the choice of the representative. Hence $f \mapsto L(f)$ is a well-defined linear functional on $\mathcal{A}$, and evidently when $f \in \mathcal{A}$ is given by (0.28), $f(\omega) \geq 0$ for all ω if and only if $\varphi(x_1, \dots, x_n) \geq 0$ for all $x_1, \dots, x_n$. Therefore, since each $\mu_{x,S}$ is a probability measure, L is positive on $\mathcal{A}$, and for all $f \in \mathcal{A}$, $|L(f)| \leq \|f\|_\infty$. Thus L has a unique extension by continuity from the dense sub algebra $\mathcal{A}$ to all of $\mathcal{C}(X)$, and evidently this extension, still denotes by L, is positive with $\|L\| = 1$.

We may now invoke the Riesz-Markov Theorem to assert the existence of a unique regular Borel measure μ_W on X such that $L(f) = \displaystyle\int_X f(\omega)\mathrm{d}\mu_W(\omega)$ for all $f \in \mathcal{C}(X)$.

The proof will be completed by showing that Ω is a Borel set in X, and that $\mu_W(\Omega) = 1$. Whether $\omega \in X$ is continuous or not depends on the behavior of ω at uncountably many values of t. The fact that μ_W is regular, specifically inner regular, will be crucial for estimating the probabilities of sets depending on the behavior of ω at all of the uncountably many points in an interval $[a,b]$ in terms of sets depending on the behavior of ω at only finitely many points in $[a,b]$.

For $\epsilon, \delta > 0$, define

$$\rho(\epsilon, \delta) := \sup_{0<t<\delta} \int_{|x|>\epsilon} \gamma_t(y) dy = \int_{|x|>\epsilon} \gamma_\delta(y) dy \tag{0.30}$$

It is easy to see that as $\delta \downarrow 0$, $\rho(\epsilon, \delta) = o(\delta)$. In fact, $\rho(\epsilon, \delta) = o(\delta^n)$ for any $n \in \mathbb{N}$, and even that is not all. However, all we shall use is that $\rho(\epsilon, \delta) = o(\delta)$. The rest of the proof is broken into several steps.

Step 1. Let $\epsilon, \delta > 0$. Let $0 < t_1 < \cdots < t_n$ with $t_n - t_1 < \delta$. Define

$$A = \{\omega \ : \ |\omega(t_1) - \omega(t_j)| > \epsilon \quad \text{for some } j = 2, \dots, n \ \}$$

We show in this step that $\mu_W(A) \leq 2\rho(\epsilon, \delta/2)$, independent of n. This is based on the Markov property. Define

$$B := \{\omega \ : \ |\omega(t_1) - \omega(t_n)| > \epsilon/2\}$$

and then for $j = 2, \dots n$, define

$$C_j := \{\omega \ : \ |\omega(t_1) - \omega(t_j)| > \epsilon \quad \text{and} \quad |\omega(t_1) - \omega(t_i)| \leq \epsilon \quad \text{for} \quad i \leq j-1\}$$

and

$$D_j := \{\omega \ : \ |\omega(t_j) - \omega(t_n)| > \epsilon/2\} \ .$$

Note that if $\omega \in A$, there is some least value of j such that $|\omega(t_1) - \omega(t_j)| > \epsilon$, and then, if $\omega \notin B$, $j \leq n-1$, and $|\omega(t_j) - \omega(t_n)| > \epsilon/2$. That is,

$$A = B \bigcup_{j=2}^{n-1} (C_j \cap D_j) \ .$$

Note that 1_{C_j} can be written in the form $1_{C_j}(\omega) = \varphi(\omega(t_1), \dots \omega(t_{j-1}))$ where φ is the characteristic function of an open set on $\mathbb{R}^{j-1}$, and that 1_{C_j} can be written in the form $1_{C_j}(\omega) = \psi(\omega(t_j), \omega(t_n))$ where φ is the characteristic function of an open set on $\mathbb{R}^2$. Then by the Markov property (0.27) and the definition of μ_W,

$$\mu_W(C_j \cap D_j) = \int_{\mathbb{R}^{j-1}} \varphi(x_1, \cdots, x_{j-1}) \left(\int_{\mathbb{R}^2} \psi(x_j, x_n) \mathrm{d}\mu_{x_j, \{t_n - t_j\}} \right) \mathrm{d}\mu_{0, \{t_1, \dots, t_j\}} \tag{0.31}$$

Since

$$\int_{\mathbb{R}^2} \psi(x_j, x_n) \mathrm{d}\mu_{x_j, \{t_n - t_j\}} = \int_{|x| > \epsilon/2} \gamma_{t_n - t_j}(x) \mathrm{d}x \leq \rho(\epsilon/2, \delta) \ ,$$

(0.31) yields $\mu_W(C_j \cap D_j) \leq \rho(\epsilon/2, \delta) \mu_W(C_j)$. Then since the sets C_j are mutually disjoint, $\mu_W(\cup_{j=2}^n C_j) \leq 1$. Thus, $\mu_W(A) \leq \mu_W(B) + \rho(\epsilon/2, \delta) \leq 2\rho(\epsilon/2, \delta)$.

Step 2. Fix $\epsilon, \delta > 0$. Fix $a < b \in [0, \infty)$ with $b - a < \delta$. Define the set

$$E(a, b, \epsilon) := \{\omega \ : |\omega(s) - \omega(t)| \geq 2\epsilon \quad \text{for some} \quad s, t \in [a, b] \ \} \ .$$

In this step we show that $\mu_W(E(a, b, \epsilon)) \leq 2\rho(\epsilon/2, \delta)$.

To do this, let S be a finite subset of $[a, b]$, and define

$$E(a, b, \epsilon, S) := \{\omega \ : |\omega(s) - \omega(t)| \geq 2\epsilon \quad \text{for some} \quad s, t \in S \ \} \ .$$

Note that each $E(a, b, \epsilon, S)$ is open, and $E(a, b, \epsilon)$ is the union of all of the $E(a, b, \epsilon, S)$ as S ranges over all finite subsets S of $[a, b]$. Since μ_W is regular, for all $\eta > 0$, there is a compact set $K \subset E(a, b, \epsilon)$ such that $\mu_W(K) \geq \subset E(a, b, \epsilon) - \eta$. Since the sets of the form $E(a, b, \epsilon, S)$, $S \subset [a, b]$ finite, are an open cover of K, there exists a finite sub-cover. But any finite union of sets of the form $E(a, b, \epsilon, S)$ is again of this form. Hence there exists a finite $S_\eta \subset [a, b]$ such that $\mu_W(E(a, b, \epsilon, S_\eta) \geq \mu_W(E(a, b, \epsilon)$. It follows that

$$\mu_W(E(a, b, \epsilon) = \sup\{\mu_W(E(a, b, \epsilon, S) \ : \ S \subset [a, b] \text{ finite } \} \ . \tag{0.32}$$

Now for any finite set $S = \{t_1, \dots, t_n\}$ if for some $1 \leq i < j \leq n$, $|\omega(t_i) - \omega(t_j)| > 2\epsilon$, then either $|\omega(t_1) - \omega(t_i)| > \epsilon$ or $|\omega(t_1) - \omega(t_j)| > \epsilon$. Hence the bound proved in Step 1 implies that $\mu_W(E(a, b, \epsilon, S) \leq 2\rho(\epsilon/2, \delta)$.

Step 3. Fix $\epsilon, \delta > 0$ with $1/\delta \in \mathbb{N}$. Let $k \in \mathbb{N}$. Define

$$F(k, \epsilon, \delta) := \{\omega \ : \ |\omega(t) - \omega(s)| > 4\epsilon \quad \text{for some} \quad t, s \in [0, k] \ \} \ .$$

In this step we show that $\mu_W(F(k, \epsilon, \delta)) \leq 2k\rho(\epsilon/2\delta)/\delta$.

To do this, write $[0, k]$ as the union of k/δ subintervals of the form $[(j-1)\delta, j\delta]$, $j = 1, \dots, k/\delta$. If $\omega \in F(k, \epsilon, \delta)$, $|\omega(t) - \omega(s)| > 4\epsilon$ for some s, t in the same or adjacent intervals. But that means that $|\omega(u) - \omega(v)| > 2\epsilon$ for some u, v belonging to one of these intervals. Thus, ω belongs to $E((j-1)\delta, j\delta, \epsilon)$, as defined in Step 2, for some $j = 1, \dots, k/\delta$. This proves the desired bound.

Step 4. We complete the proof. A function $\omega : [0, \infty] \to \dot{\mathbb{R}}$ is continuous with $\omega(0) \in \mathbb{R}$ if and only if its restriction to each $[0, k]$, $k \in \mathbb{N}$ is uniformly continuous, meaning that for for all $\epsilon > 0$, there is a $\delta > 0$ so that $|\omega(s) - \omega(t)| < 4\epsilon$ whenever $|s - t| \leq \delta$. That is, ω is continuous if and only if for each $k \in \mathbb{N}$, ω belongs to

$$\bigcap_{\epsilon > 0} \bigcup_{\delta > 0} F^c(k, \epsilon, \delta) \ ,$$

and $F(k, \epsilon, \delta)$ is defined as in the previous step. Moreover, we can restrict ϵ and δ to be the reciprocals on positive integers so that the intersection and union are countable. Since each $F(k, \epsilon, \delta)$ is open, in X,

$$\Omega = \bigcap_{k \in \mathbb{N}} \bigcap_{\epsilon > 0} \bigcup_{\delta > 0} F^c(k, \epsilon, \delta) \ ,$$

is a Borel set. It follows that Ω^c is a countable union of sets of the form $\bigcap\limits_{\delta > 0} F(k, \epsilon, \delta)$ and

$$\mu_W\left(\bigcap_{\delta > 0} F(k, \epsilon, \delta)\right) = \lim_{\delta \downarrow 0} \mu_W(F(k, \epsilon, \delta)) = 0$$

by the estimate of Step 3 since $\rho(\epsilon, \delta) = o(\delta)$. Thus Ω is a Borel subset of X, and $\mu_W(\Omega) = 1$. We define ν_W to be the restriction of μ_W to Ω. □

0.7 Functions of Bounded Variation

The class of *functions of bounded variation* on an interval $I \subset \mathbb{R}$ was introduced by Camile Jordan in 1881 in an investigation of the point-wise convergence of Fourier Series. The notion turns out to be quite useful in many contexts, and we shall give a modern development of the theory that readily admits generalization to functions on open sets in $\mathbb{R}^n$, which is crucial for many modern applications. However, first we recall the classical definition:

0.27 DEFINITION (Bounded variation)**.** Let I be an interval in $\mathbb{R}$. Let $\mathcal{P}_I$ denote the set of all ordered sets $\{x_0, x_1, \dots, x_n\} \subset I$, $n \in \mathbb{N}$, with $x_{j-1} < x_j$ for all $j = 1, \dots, n$. For any real valued function h defined on I, define

$$TV(h; I) = \sup\left\{\sum_{j=1}^{n} |h(x_j) - h(x_{j-1})| \ : \ \{x_0, x_1, \dots, x_n\} \in \mathcal{P}_I \right\} \ . \qquad (0.33)$$

A function $h : I \to \mathbb{R}$ is of *bounded variation on* I in case $TV(h; I) < \infty$.

0.28 EXAMPLE. Let h be continuously differentiable and such that the derivative h' is integrable on an interval (a,b). Then for any $\{x_0, x_1, \dots, x_n\} \in \mathcal{P}_{(a,b)}$,

$$\sum_{j=1}^{n} |h(x_j) - h(x_{j-1})| = \sum_{j=1}^{n} \left| \int_{x_{j-1}}^{x_j} h'(y)\mathrm{d}y \right| \le \int_a^b |h'(x)|\mathrm{d}x \ ,$$

and hence $TV(h;(a,b)) \le \int_a^b |h'(x)|\mathrm{d}x$. In fact, it is not hard to show that actually equality holds in this inequality.

However, functions of bounded variation need not be continuous. Consider for example the function f on $\mathbb{R}$ defined by

$$h(x) = \begin{cases} 0 & x < 0 \\ 1 & x \ge 0 \ . \end{cases}$$

Evidently $TV(h,\mathbb{R}) = 1$. More generally, any monotone non-decreasing function has bounded variation on any interval on which it is bounded.

We shall soon we that every bounded variation function on any interval $[a,b]$ is the difference of two monotone non-decreasing functions; this is the *Jordan decomposition.* The Jordan decomposition is closely related to the Hahn-Saks decomposition of a signed Borel measure into its positive and negative parts. This means that any function of bounded variation is measurable, and in fact, has left and right limits (which need not be equal) at every point. In particular, functions of bounded variation are measurable. In the modern development presented here, we shall deduce the Jordan Decomposition from the Hahn-Saks decomposition, but we need to know from the outset that functions of bounded variation are at least measurable, which we show after proving a simple lemma that will be useful again.

0.29 LEMMA. *Let $\epsilon > 0$. Let h be a real valued function defined on the interval $[c,d]$ such that $TV(h,[c,d]) < \epsilon$. Let $\widetilde{h}$ be the linear interpolation of h between c and d, That is, for any $\lambda \in [0,1]$ define $x_\lambda = (1-\lambda)c + \lambda d$, and $\widetilde{h}(x_\lambda) = (1-\lambda)h(c) + \lambda h(d)$. Then $|\widetilde{h}(x) - h(x)| < \epsilon$ for all $x \in [c,d]$.*

Proof. By hypothesis, for any $x \in (c,d)$ $|h(x) - h(c)| + |h(x) - h(d)| \le TV(h,[c,d]) < \epsilon$. By the convexity of $t \mapsto |t|$,

$$|h(x_\lambda) - (1-\lambda)h(c) - \lambda h(d)| \le (1-\lambda)|h(x_\lambda) - h(c)| + \lambda |h(x_\lambda) - h(d)| \le \epsilon \ .$$

□

Then next simple observation to make is that if $b < c < d$, $TV(h;[b,d]) = TV(h;[b,c]) + TV(h;[c,d])$. This may be iterated in the obvious way.

0.30 LEMMA. *Let m denote Lebesgue measure on $\mathbb{R}$. Let h be a real valued function defined on the interval $[c,d]$ such that $TV(h,[c,d]) < \infty$. For each $\epsilon > 0$ there is a piecewise linear function h_ϵ and a Borel set $E \subset [c,d]$ with $m(E) < \epsilon$ such that for all $x \in [c,d] \cap E^c$, $|h(x) - h_\epsilon(x)| < \epsilon$ and $|h(x) - h_\epsilon(x)| \le TV(h,[c,d])$ for* all *$x \in [a,d]$. Moreover, $TV(h_\epsilon,[c,d]) \le TV(h,[c,d])$*

Proof. Choose $k \in \mathbb{N}$ such that $\frac{1}{k}TV(h;[c,d]) < \epsilon$. Then choose n so large that $k(d-c)/n < \epsilon$. Divide $[c,d]$ into n closed intervals $\{I_j \ : \ j = 1,\dots,n\}$ of equal length, with consecutive intervals overlapping at endpoints only. Since $\displaystyle\sum_{j=1}^{n} TV(h; I_j) = TV(h; [c,d])$, $TV(h; I_j) \geq \frac{1}{k} TV(h; I_j)$ for at most k values of j. Let E be the union of any such intervals, and then by the choice of n, $m(E) < \epsilon$. On each I_j not included in E, h is uniformly within ϵ of its linear interpolation between the endpoints $\{x_0, \dots, x_n\}$ of the intervals of I_j according to Lemma 0.29. Thus if we define h_ϵ be the linear interpolation of h at the endpoints of the intervals I_j, $j = 1, \dots, n$, $|h(x) - h_\epsilon(x)| < \epsilon$ for $x \notin E$. Finally, it is clear that since h_ϵ is linear between x_{j-1} and x_j for each $j = 1, \dots, n$,

$$TV(h_\epsilon) = \sum_{j=1}^{n} |h(x_j - h(x_{j-1})| \leq TV(h, [c,d]) \ .$$

□

0.31 Remark. Given any real valued function h on $[c,d]$ such that $TV(h,[c,d]) < \infty$ Lemma 0.30 provides a sequence of continuous functions converging to h in measure on $[c,d]$, and hence h is measurable on $[c,d]$.

In what follows, (a,b) denotes any open interval in $\mathbb{R}$. In particular, it is not excluded that either $a = -\infty$, $b = \infty$, or both. Let $\mathcal{C}_c^1((a,b))$ denote the set of continuously differentiable real-valued functions with compact support in (a,b). Then $\mathcal{C}_c^1((a,b))$ is dense in $\mathcal{C}_c((a,b))$. This is easily shown by convolving $f \in \mathcal{C}_c((a,b))$ by a smooth probability density supported in some sufficiently small interval $[-\delta,\delta]$. In fact, the same argument shows that $\mathcal{C}_c^\infty((a,b))$, the set of infinitely differentiable real-valued functions with compact support in (a,b), is dense in $\mathcal{C}_c((a,b))$. For $f \in \mathcal{C}_c^1((a,b))$, let f' denote the derivative of f. Throughout the rest of this section, $\mathcal{B}$ denotes the Borel σ-algebra of (a,b), and m denote Lebesgue measure on (a,b), though we shall often write integrals with respect to Lebesgue measure using dx in place of dm.

Now let h be a real-valued function on (a,b) with $TV(h;(a,b)) < \infty$. In particular, h is uniformly bounded on (a,b), and so if both a and b are finite, h is integrable with respect to Lebesgue measure m since it is also measurable by the remark following Lemma 0.30. If either $a = -\infty$ or $b = \infty$ this need not be the case, and then we further suppose that h is integrable. The next lemma leads to the modern characterization of functions of bounded variation.

0.32 LEMMA. *Let h be an integrable real-valued function on (a,b) with $TV(h;(a,b)) < \infty$. Let μ denote the finite Borel measure on (a,b) that is absolutely continuous with respect to Lebesgue measure on (a,b) with Radon-Nikodym derivative h. That is, $\mathrm{d}\mu = h\mathrm{d}x$. Define $L \in (\mathcal{C}_0((a,b)))^*$ by*

$$L(f) = \int_a^b f \mathrm{d}\mu \ . \tag{0.34}$$

Then for all $f \in \mathcal{C}_c^1((a,b))$,

$$|L(f')| \leq TV(h;[a,b]) \|f\|_\infty \ . \tag{0.35}$$

Proof. Let $f \in \mathcal{C}_c^1((a,b))$, and let $[c,d]$ contain the support of f. Let $\epsilon > 0$. By Lemma 0.30, there exists a piecewise linear function h_ϵ on $[c,d]$ such that $|h_\epsilon(x) - h(x)| \leq TV(h,[c,d])$ for all $x \in [c,d]$ and such that $|h_\epsilon(x) - h(x)| \leq \epsilon$ outside a set E with $m(E) < \epsilon$. Therefore,

$$L(f') = \int_a^b f'(x) h_\epsilon(x) \mathrm{d}x + \int_a^b f'(x)(h - h_\epsilon((x)\mathrm{d}x \ ,$$

and hence

$$\begin{aligned}\left|L(f') - \int_a^b f'(x)h_\epsilon(x)\mathrm{d}x\right| &\leq \int_c^d |f'(x)||h - h_\epsilon((x)|\mathrm{d}x \\ &\leq \|f'\|_\infty \int_{[c,d]\setminus E} |h - h_\epsilon((x)|\mathrm{d}x + \|f'\|_\infty \int_E |h - h_\epsilon((x)|\mathrm{d}x \\ &\leq \epsilon\|f'\|_\infty(d - c + TV(h, [c, d])) \ .\end{aligned}$$

Then if $\{x_0, \dots, x_n\}$ denotes set in $\mathcal{P}_{[c,d]}$ such that h_ϵ is the linear interpolation of h through $\{x_0, \dots, x_n\}$, integration by parts yields

$$\left|\int_a^b f'(x)h_\epsilon(x)\mathrm{d}x\right| = \left|\sum_{j=1}^n \frac{h(x_j) - h(x_{j-1})}{x_j - x_{j-1}} \int_{x_{j-1}}^{x_j} f(y)\mathrm{d}y\right| \leq \|f\|_\infty \sum_{j=1}^n |h(x_j) - h(x_{j-1})| \ .$$

Combining the last two estimates, $|L(f')| \leq TV(h; [a, b])\|f\|_\infty(1+\epsilon(d-c)+TV(h, [c, d]))$. Since $\epsilon > 0$ is arbitrary, (0.35) is proved. □

The construction in the previous lemma show how integrable functions of bounded variation give rise to a special class of linear functionals L on $\mathcal{C}_0((a, b))$: Those for which there exists $C < \infty$ such that for all $f \in \mathcal{C}_c^1((a, b))$

$$|L(f')| \leq C\|f\|_\infty \ . \tag{0.36}$$

We shall soon see that every such functional L $\mathcal{C}_0((a, b))$ arises in this way: There is a unique integrable function h of bounded variation on (a, b) such that $L(f) = \int_{(a,b)} fh\mathrm{d}x$ for all $f \in \mathcal{C}_0((a, b))$. In the course of proving this, we shall obtain further information about the class of functions of bounded variation. We begin with a lemma that provides a crucial "generalized integration by parts" formula.

0.33 LEMMA. *$L \in (\mathcal{C}_0((a, b)))^*$ be such that for some $C < \infty$ and all all $f \in \mathcal{C}_c^1((a, b))$, (0.36) is valid. Then there exists a unique pair of signed Borel measures μ and ν on (a, b) such that for all $f \in \mathcal{C}_0((a, b))$*

$$L(f) = \int_{(a,b)} f\mathrm{d}\mu \tag{0.37}$$

and for all $f \in \mathcal{C}_c^1((a, b))$

$$L(f') = -\int_{(a,b)} f\mathrm{d}\nu \ . \tag{0.38}$$

Moreover, $\|\nu\|_{\mathrm{TV}} \leq C$, and for all $f \in \mathcal{C}_c^1((a, b))$,

$$\int_{(a,b)} f'\mathrm{d}\mu = -\int_{(a,b)} f\mathrm{d}\nu \ . \tag{0.39}$$

Proof. Direct application of the Riesz Representation Theorem for $(\mathcal{C}_0((a, b)))^*$, provides the signed measure μ such that (0.37) is valid. Next, note that $\mathcal{C}_c^1((a, b))$ is dense in $\mathcal{C}_0((a, b))$ in the uniform norm, and the functional $f \mapsto L(f')$ is linear since differentiation and L are both linear. It is bounded on the dense subspace $\mathcal{C}_c^1((a, b))$ by (0.36), and hence it extends by continuity to a linear functional M on all of $\mathcal{C}_0((a, b))$ and $\|M\|_* \leq C$. Now a second application of the Riesz Representation Theorem for $(\mathcal{C}_0((a, b)))^*$ provides the signed measure ν such that (0.38) is valid. Finally, (0.39) follows directly from (0.37) and (0.38). □

0.34 EXAMPLE. Let h be a continuously differentiable function on (a,b) such that $\int_a^b |h'(x)|\mathrm{d}x = C < \infty$. Define

$$L(f) = \int_a^b h(x)f(x)\mathrm{d}x \tag{0.40}$$

for all $f \in \mathcal{C}_c((a,b))$. Then, integrating by parts, for $f \in \mathcal{C}_c^1((a,b))$,

$$L(f') = \int_a^b h(x)f'(x)\mathrm{d}x = -\int_a^b f(x)h'(x)\mathrm{d}x \tag{0.41}$$

and hence $|L(f')| \le C\|f\|_\infty$ so that (0.36) is satisfied with $C = \displaystyle\int_a^b |h'(x)|\mathrm{d}x$. Moreover, in this case the measure ν is evidently given by

$$\mathrm{d}\nu = -h'(x)\mathrm{d}x \ .$$

We may regard $-\nu$ as the "generalized derivative" of the bounded variation function h even when h is not differentiable.

0.35 LEMMA. *Let L, C, μ and ν be as in Lemma 0.33. If $b < \infty$, then the limits*

$$\lim_{c\uparrow b} \frac{1}{b-c}\mu(c,b) =: h(b) \quad \text{and} \quad \lim_{c\downarrow a} \frac{1}{c-a}\mu(c,b) =: h(a) \tag{0.42}$$

exit and

$$|h(a)|, |h(b)| \le \frac{1}{b-a}\|\mu\|_{\mathrm{TV}} + C \ . \tag{0.43}$$

Proof. Let f be a C^1 function that monotonically increases from 0 to 1 such that f' has support contained in $[a',b'] \subset (a,b)$. For each $c \in (b',b)$ define

$$f_c(x) = \begin{cases} f(x) & x \le c \\ 1 - x/(b-c) & c < x < b \ . \end{cases}$$

While f_c does not belong to $C_c^1((a,b))$, it is easy to approximate f_c by a sequence $\{g_{c,k}\}_{k\in\mathbb{N}}$ of such functions that converge uniformly to f_c and whose derivatives converge at each x to

$$f'(x) - \frac{1}{b-c}1_{(c,b)}(x)$$

and are uniformly bounded in absolute value by $\max\{\|f'\|_\infty, (b-c)^{-1}\}$. Applying (0.39) to $g_{c,k}$ and taking the limit $k \to \infty$, we obtain

$$\int_{(a,b)} f'\mathrm{d}\mu - \frac{\mu((c,b))}{b-c} = -\int_{(a,b)} f_c\mathrm{d}\nu \ . \tag{0.44}$$

By the Lebesgue Dominated Convergence Theorem,

$$\lim_{c\uparrow b}\int_{(a,b)} f_c\mathrm{d}\nu = \int_{(a,b)} f\mathrm{d}\nu \ , \tag{0.45}$$

and since the integral on the left hand side of (0.44) is independent of c, the first limit in (0.42) exists. Moreover, for each c and k, $|L(g'_{n,k})| \le C\|g_{n,k}\|_\infty = C$, and hence for all c,

$$\left| \int_{(a,b)} f' \mathrm{d}\mu - \frac{\mu((c,b))}{b-c} \right| \le C \ . \tag{0.46}$$

Moreover, f can be chosen so that $\|f'\|_\infty$ is arbitrarily close to $(b'-a)^{-1}$, and we may take b' arbitrarily close to c. In particular, if $a = -\infty$, we can choose f with $\|f'\|_\infty$ is arbitrarily close to 0. Then (0.46) yields

$$\frac{|\mu((c,b))|}{b-c} \le (c-a)^{-1}\|\mu\|_{\mathrm{TV}} + C \ ,$$

and this proves (0.43) for $h(b)$. The statements involving $h(a)$ are valid by symmetry. □

0.36 LEMMA. *Let L, C, μ and ν be as in Lemma 0.33. If $b < \infty$ let $h(b)$ be defined as in (0.42), and if $b = \infty$, define $h(b) = 0$. Let f be any $C^1((a,b))$ function with $f'(x) \ge 0$ for all x and $\lim_{x\downarrow a} f(x) = 0$ and $\|f\|_\infty = \lim_{x\uparrow b} f(x) < \infty$. Then*

$$\int_{(a,b)} f' \mathrm{d}\mu = -\int_{(a,b)} f \mathrm{d}\nu + h(b)\|f\|_\infty \ . \tag{0.47}$$

Proof. Note that $\|f'\|_{L^1(m)} = \|f\|_{L^\infty(m)}$. Since $f(x) = \int_a^x f'(y)\mathrm{d}y$, making a change in f' that has a small $L^1(m)$ norm results in a change in f that is correspondingly small in the uniform norm. We may therefore suppose without loss of generality that f' has compact support in $[a',b'] \subset (a,b)$. By homogeneity, we may assume that $\|f\|_\infty = 1$. In case $b < \infty$, we may then proceed as in the last lemma, and then (0.42), (0.44) and (0.45) yield (0.47).

In case $b = \infty$, pick $\epsilon > 0$. Since ν is a finite Radon measure, by increasing b' as necessary, we may suppose that $\nu((b',b)) < \epsilon$. Let $g \in C^1_c((a,b))$ be such that $g'(x) = f'(x)$ for $x \le b'$, and such that $|g'(x)| < \epsilon$ for $x > b'$, which is easy to do since $(b',b) = (b',\infty)$. Define $g(x) = \int_a^x g'(y)\mathrm{d}y$, and note that for $x \le b'$, $g(x) = f(x)$, and we may arrange that $\|g - f\|_\infty = \|f\|_\infty$. Then by (0.39),

$$\begin{aligned}
\int_{(a,b)} f' \mathrm{d}\mu &= \int_{(a,b)} g' \mathrm{d}\mu - \int_{(b',b)} g' \mathrm{d}\mu \\
&= -\int_{(a,b)} g \mathrm{d}\nu - \int_{(b',b)} g' \mathrm{d}\mu \\
&= -\int_{(a,b)} f \mathrm{d}\nu + \int_{(b',b)} (f-g) \mathrm{d}\nu - \int_{(b',b)} g' \mathrm{d}\mu
\end{aligned}$$

Now note that $\left| \int_{(b',b)} (f-g)\mathrm{d}\nu \right| \le \|f\|_\infty \nu(b',b)$ and $\left| \int_{(b',b)} g' \mathrm{d}\nu \right| \le \epsilon\|\nu\|_{\mathrm{TV}}$. Thus, recalling that $\|\nu\|_{\mathrm{TV}} \le C$,

$$\left| \int_{(a,b)} f' \mathrm{d}\mu + \int_{(a,b)} f \mathrm{d}\nu \right| \le \epsilon(\|f\|_\infty + C) \ .$$

Since ϵ is arbitrary, the identity (0.47) is proved. □

0.37 THEOREM. *$L \in (\mathcal{C}_0((a,b)))^*$ be such that for some $C < \infty$ and all all $g \in \mathcal{C}^1_c((a,b))$, (0.36) is valid. Let μ and ν be the unique signed Borel measures on (a,b) such that (0.37) and (0.38) are valid for all $f \in \mathcal{C}_0((a,b)$ and all $f \in \mathcal{C}^1_c((a,b)$ respectively.*

Then μ is absolutely continuous with respect to Lebesgue measure. Let h denote the Radon-Nikodymn derivative of μ with respect to Lebesgue measure so that by the Lebesgue Differentiation Theorem, at almost every x,

$$h(x) = \lim_{y \downarrow x} \frac{\mu((x,y)}{y-x} \ . \tag{0.48}$$

Then there is a preferred representative of the a.e. equivalence class of h such that (0.48) is valid for every *$x \in (a,b)$, and with this version of h,*

$$TV(h;(a,b)) \leq C \ , \tag{0.49}$$

and for all $a \leq x < y < b$,

$$h(x) - h(y) = \nu((x,y]) \ . \tag{0.50}$$

In particular, h is right continuous and has a left limit at each point $x \in (a,b)$.

Before giving the proof, we recall a basic measure theoretic result that we shall use. Let $\mathcal{A}$ denote the half-open interval algebra which consists of all finite disjoint unions of sets of the form $(c,d]$ with $a \leq c < d < b$ or of the form (c,b) where $a < c < b$. By a standard application of the Monotone Class Theorem, for every positive Radon measure λ on $\mathbb{R}$ and every Borel set E, for each $\epsilon > 0$, there is a set $A \in \mathcal{A}$ such that $\lambda(E\Delta A) < \epsilon$.

Proof. Let $\mu = \mu_+ - \mu_-$ be the unique decomposition of μ into a different of mutually singular finite (positive) Borel measures, and let $|\mu| = \mu_+ + \mu_-$, and likewise define ν_+ and ν_- in terms of ν. Let m denote Lebesgue measure and note that $|\mu| + m$ is a Radon measure.

Let E be a Borel set such that $\mu_-(E) = 0$ and $\mu_+(E^c) = 0$. Then for any $\epsilon > 0$, there exists $A_\epsilon \in \mathcal{A}$ such that

$$(|\mu| + m)(E\Delta A_\epsilon) < \epsilon \ .$$

Now let F be any Borel subset of E and suppose that $\mu_+(F) > 0$. For $\epsilon > 0$, let $B \in \mathcal{A}$ be such that $\mu(F\Delta B) < \epsilon$. Then since $F \subset E$, $1_B 1_{A_\epsilon} = (1_B - 1_F)1_{A_\epsilon} + 1_F(1_{A_\epsilon} - 1_E) + 1_F$. It follows that

$$(|\mu| + m)(B \cap A_\epsilon - F) \leq 2\epsilon \ . \tag{0.51}$$

Since $B \cap A_\epsilon \in \mathcal{A}$, it can be written in as $B \cap A_\epsilon = \displaystyle\sum_{j-1}^n (c_j, d_j]$, where $c_1 < d_1 < c_2 < \cdots < d_n$. At the cost of another ϵ, we may assume that $c_1 > -\infty$ and $d_n < \infty$, which is automatic in case a and b are finite.

Now for $j = 1, \ldots, n$ choose disjoint open intervals U_j such that $[c_j, d_j] \subset U_j$ and such that

$$\sum_{j=1}^n (|\mu| + m)(U_j) \leq (|\mu| + m)(B \cap A_\epsilon) + \epsilon \ , \tag{0.52}$$

which is possible by the outer regularity of $|\mu| + m$. For each j, choose $[c_j, d_j] \prec g_j \prec U_j$, and define $g := \sum_{j=1}^n g_j$. Define $f(x) = \displaystyle\int_a^x \sum_{j=1}^n g_j(y)\mathrm{d}y$.

Then

$$\|f\|_\infty \leq \sum_{j=1}^{n} m(U_j) \leq \epsilon + m(B \cap A_\epsilon) \leq 3\epsilon + m(F) \ . \tag{0.53}$$

Also, with $U := \cap_{j=1}^n U_j$,

$$\int_{(a,b)} f' \mathrm{d}\mu \geq \mu_+(B \cap A_\epsilon) - \mu_-(U) \ .$$

By (0.52), $\mu_-(U) \leq \mu_-(A_\epsilon) + \epsilon \leq 2\epsilon$. By (0.51), $\mu_+(B \cap A_\epsilon) \geq \mu_+(F) - \epsilon$. Altogether, using Lemma 0.36 together with the positivity of f,

$$\begin{aligned}
\mu_+(F) &\leq 3\epsilon + \int_{(a,b)} f' \mathrm{d}\mu \\
&\leq 3\epsilon + \int_{(a,b)} f \mathrm{d}\nu_- + h(b)\|f\|_\infty \\
&\leq 3\epsilon + (C + h(b))\|f\|_\infty \ .
\end{aligned}$$

Combining this with (0.43) and (0.53), we have

$$\mu_+(F) \leq 3\epsilon + (2C + (b-a)^{-1})(3\epsilon + m(F)) \ .$$

Since $\epsilon > 0$ is arbitrary, $\mu_+(F) \leq (2C + (b-a)^{-1}\|\mu\|_{\mathrm{TV}})m(F)$. and then since F is arbitrary, this proves that μ_+ is absolutely continuous with respect to Lebesgue measure, and then using the Lebesgue Differentiation Theorem that the Radon-Nikodymn derivative h satisfies $\|h\|_\infty \leq 2C + (b-a)^{-1}\|\mu\|_{\mathrm{TV}}$.

For the final part, given any $a \leq x < y < b$, for n sufficiently large that and $y - 1/n > x$, define the "ramp function" f_n approximation of $1_{(x,y]}$ by

$$f_n(t) = \begin{cases} n(t-x) & x < t \leq x + 1/n \\ 1 & x + 1/n \leq t \leq y \\ 1 - n(t-y) & y < t \leq y + 1/n \end{cases}$$

and $f_n(t) = 0$ for all other t. Notice that $\lim_{n\to\infty} f_n(t) = 1_{(x,y]}(t)$ for all t. Though f is not in $C^1_c((a,b))$, a simple approximation argument such as we made in the proof of Lemma 0.35 shows that (0.39) is nonetheless valid and therefore

$$n\mu((x - 1/n, x)) - n\mu((y - 1/n, y)) = -\int f_n \mathrm{d}\nu \ .$$

Taking the limit $n \to \infty$, we obtain $h(y) - h(x) = \nu((x,y])$. It now follows that when $a < x_0 < x_1 \cdots x_n < b$, $\displaystyle\sum_{j=1}^{n} |h(x_j) - h(x_{j-1})| = \sum_{j=1}^{n} |\nu((x_{j-1}, x_j]) \leq \|\nu\|_{\mathrm{TV}}$. □

0.38 COROLLARY. *Let h be integrable on (a,b) and suppose that $TV(h;(a,b)) < \infty$. Then in the a.e. equivalence class of h there is a preferred representative, also denoted by h, such that h is right continuous and has a left limit at each point $x \in (a,b)$. Moreover, h is the difference of two monotone non-decreasing right continuous functions $h = h_+ - h_-$.*

Proof. The first part is a direct consequence of Lemma 0.32 and Theorem 0.37. Let ν be the signed measure associated to h as in Theorem 0.37. Let $\nu = nu_+ - \nu_-$ be the Hahn-Saks decomposition of ν into its positive and negative parts. Then since for all $a < x < b$, $h(x) - h(a) = \nu_-((a,x]) - \nu_+((a,x])$. Define $h_+(x) = h(a) + \nu_-((a,x])$ and define $h_-(x) = \nu_+((a,x])$. □

0.39 DEFINITION (BV((a,b)))**.** The real normed space $BV((a,b))$, called the space of BV functions on (a,b) is the vector space of real integrable functions h on (a,b) such that $TV(h;(a,b)) < \infty$. For $h \in BV((a,b))$, let $\mu = h\mathrm{d}x$, so that by Lemma 0.32 and Lemma **??**, there is a signed Borel measure ν with $\|\nu\|_{\mathrm{TV}} = TV(h;(a,b)) < \infty$, and such that (0.39) is valid for all $f \in \mathcal{C}^1_c((a,b))$. Then we define the BV norm of h as

$$\|h\|_{\mathrm{BV}} = \|h\|_{L^1(m)} + \|\nu\|_{\mathrm{TV}} = \|\mu\|_{\mathrm{TV}} + \|\nu\|_{\mathrm{TV}} \tag{0.54}$$

By the uniqueness in Lemma 0.33 and the linearity of (0.39), the map $h \mapsto \nu$ is linear, and so it is evident that $\|\cdot\|_{\mathrm{BV}}$ is a norm on $BV((a,b))$. To see that $BV((a,b))$ is complete in this norm, Let $\{h_n\}_{n\in\mathbb{N}}$ be a Cauchy sequence in $BV((a,b))$. Then $\{h_n\}_{n\in\mathbb{N}}$ is also a Cauchy sequence in $L^1((a,b),\mathcal{B},m)$ and so there exist $h \in L^1((a,b),\mathcal{B},m)$ such that $\lim_{n\to\infty}\|h_n - h\|_{L^1(m)} = 0$. For each $n \in \mathbb{N}$, let μ_n be the signed Borel measures such that

$$\int_{(a,b)} f' h_n \mathrm{d}x = -\int_{(a,b)} f \mathrm{d}\nu_n$$

for all $f \in \mathcal{C}_0((a,b))$.

Since the space of signed Borel measures on $\mathcal{C}_0((a,b))$ is, like every dual space, complete, and since by the definition of the BV norm $\{\nu_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in the total variation norm, there exists a signed measure ν such that $\lim_{n\to\infty}\|\nu_n - \nu\|_{\mathrm{TV}} = 0$. It follows that

$$\int_{(a,b)} f' h \mathrm{d}x = -\int_{(a,b)} f \mathrm{d}\nu$$

for all $f \in \mathcal{C}_0((a,b))$, and hence $h \in BV((a,b))$ and $\lim_{n\to\infty}\|h_n - h\|_{\mathrm{BV}} = 0$.

There is in fact a better way to see that $BV((a,b))$ is a Banach space: It turns out the $BV((a,b))$ is the dual of another Banach space. Let Y be the Banach space of of all pairs (f_1,f_2) with $f_1, f_2 \in \mathcal{C}_0((a,b))$ and with the norm

$$\|(f_1,f_2)\|_Y = \|f_1\|_\infty + \|f_2\|_\infty\ .$$

The dual space Y^* is the space of all pairs (μ_1,μ_2) of signed Borel measures on (a,b) with the norm

$$\|(\mu_1,\mu_2)\|_Y = \|\mu_1\|_{\mathrm{TV}} + \|\mu_2\|_{\mathrm{TV}}\ .$$

Now let Z be closure of the subspace of Y consisting on (f_1,f_2) with $f_1, f_2 \in \mathcal{C}^1_c((a,b))$ and $f_1 = -f_2'$. The annihilator of Z is the subspace of Y^* consisting of pairs (μ,ν) such that

$$\int_{(a,b)} f' \mathrm{d}\mu = \int_{(a,b)} f' \mathrm{d}\nu \tag{0.55}$$

for all $f \in \mathcal{C}^1_c((a,b))$. By Theorem **??**, (μ,ν) belongs to the annihilator of Z precisely when $\mu = h\mathrm{d}x$ where $h \in BV((a,b))$, and then μ is the unique signed measure associated to h such (0.55) is valid for all $f \in \mathcal{C}^1_c((a,b))$, and in this case

$$\|h\|_{\mathrm{BV}} = \|\mu\|_{\mathrm{TV}} + \|\nu\|_{\mathrm{TV}} = \|(\mu,\nu)\|_{Y^*}\ .$$

Since Z is a closed subspace of a Banach space Y, the annihilator of Z in Y^* is the dual of Y/Z by a general result in the theory of Banach spaces. Therefore, $BV((a,b))$ is the dual of Y/Z. By the Banach-Aloglu Theorem, the unit closed ball in $BV((a,b))$ is compact in the weak-$*$ topology induced on $BV((a,b))$ by Y/Z.

The space $BV((a,b))$ is not separable: For each $y \in (a,b)$, let h_y be the step function with $h_y(x) = 1$ for $y \geq x$ and $h_y(x) = 0$ otherwise. Then for $y \neq z$, $\|h_y - h_z\|_{\mathrm{BV}} \geq 2$.