Theory of Radon measures in LCH spaces

Yuxuan Song

May 2024

Prologue

This notes consists of my learnings in the structure of Radon measures on an abstract measure space, in particular, on a locally compact Hausdorff (LCH) topological space. I have tried to link this rather under-appreciated part of measure theory with some aspects in functional analysis, including weak-* topology, positive linear functionals and their representations. Some aspects arising from geometric measure theory and the calculus of variations are discussed after having introduced the general framework with Radon measures. This note should be more or less self-contained in the sense that the readers are only expected to know some elementary measure theory. The material here is largely inspired by the content in the excellent books by Folland [Fol13] and by Ambrosio-Fusco-Pallara [AFP00]. All mistakes or typos are almost certainly my own and any comment or corrections will be appreciated.

1 Locally compact Hausdorff topological spaces

Recall a locally compact Hausdorff(LCH) topological space admits a compact neighbourhood around each point and the defining topology is Hausdorff. For sufficiently regular measures, we can approximate any measurable sets with compact subsets from the inside. Moreover, the space of Radon measures naturally gives rise to a characterisation of the continuous dual space for the space of continuous functions. This fact constitutes of an important link between integration theory and functional analysis, as well as the study of rough geometric objects. Throughout this text, we will refer to countable unions/intersections of closed/open sets F_σ or G_δ sets.

Definition 1.1 (Separation axioms)**.** *Let X be a topological space. We say X is a T_j-space if:*

- *T_0 (Kolmogorov), if for $x \neq y \in X$, there exists $U \in \mathscr{T}_X$ such that $x \in U$ and $y \notin U$ or an open set $V \in \mathscr{T}_X$ such that $x \notin V$ and $y \in V$;*
- *T_1 (Fréchet), if for $x \neq y \in X$, there exists $U \in \mathscr{T}_X$ such that $x \in U$ and $y \notin U$;*
- *T_2 (Hausdorff), if for $x \neq y \in X$, we can find disjoint open sets $U, V \in \mathscr{T}_X$ with $x \in U$ and $y \in V$;*
- *T_3 (regular), if X is T_1-space and for any $F \subset X$ closed, $x \in X \setminus F$, there are disjoint open sets $U, V \in \mathscr{T}_X$ with $x \in U$ and $F \subset V$;*
- *T_4 (normal), if X is T_1-space and for any disjoint closed sets $F, K \subset X$, we can find disjoint open neighbourhoods $U, V \in \mathscr{T}_X$ with $F \subset U$ and $K \subset V$.*

We will establish another topological separation condition. A prototype on normal spaces is discussed below.

Lemma 1.1. *Let X be T_4-space and $A, B \subset X$ disjoint closed subsets. Denote the set of dyadic rationals in $(0,1)$ by $\mathscr{D} = \{k2^{-i} \mid i \in \mathbb{N},\, k \in (0, 2^i)\}$. Then there exists a family $\{U_d \mid \delta \in \mathscr{D}\}$ of open subsets in X such that $A \subset U_\delta \subset X \setminus B$ for any $\delta \in \mathscr{D}$ and $\overline{U}_\delta \subset U_s$ for all $\delta < s$.*

Proof. By normality, we find disjoint $V, W \in \mathscr{T}_X$ with $A \subset V$ and $B \subset W$. Then since $X \setminus W$ is closed, setting $U_{\frac{1}{2}} = V$ gives us $A \subset U_{\frac{1}{2}} \subset \overline{U}_{\frac{1}{2}} \subset X \setminus W \subset X \setminus B$.
We now proceed by induction on i to find $U_{k2^{-i}}$. Suppose for induction we have found $U_{k2^{-i}}$ for any $k \in (0, 2^i)$ and $i \leqslant N$. Set for convenience $\overline{U}_0 = A$ and $X \setminus U_1 = B$. Note $\overline{U}_{j2^{-N}}$ and $X \setminus U_{(j+1)2^{-N}}$ are disjoint closed sets for any $j \in [0, 2^N)$. Thus by the base case, we can find $U_{(2j+1)2^{-(N+1)}} \in \mathscr{T}_X$ such that:

$$A \subset \overline{U}_{j2^{-N}} \subset U_{(2j+1)2^{-(N+1)}} \subset \overline{U}_{(2j+1)2^{-(N+1)}} \subset U_{(j+1)2^{-N}} \subset X \setminus B.$$

The collection $\{U_\delta \mid \delta \in \mathscr{D}\}$ clearly satisfies all the desired conditions. $\square$

A geometric intuition for the above lemma comes from simply taking X to be the plane $\mathbb{R}^2$. Then the sets U_δ can be thought of as regions bounded by plane curves ∂U_δ, forming a topographical map for f.

Proposition 1.1 (Urysohn's lemma on normal spaces)**.** *Let X be a normal space and $A, B \subset X$ disjoint closed subsets. Then there exists $f \in C^0(X; [0,1])$ such that $f \equiv 0$ on A and $f \equiv 1$ on B.*

Proof. Let $\{U_\delta \mid \delta \in \mathscr{D}\}$ be the collection from lemma 1.1 and set for convenience $U_1 = X$. We can define:

$$f\colon X \to [0,1], \quad x \longmapsto \inf\{\delta \in \mathscr{D} \cup \{1\} \mid x \in U_\delta\},$$

which is now well-defined. We clearly have the required separation since $A \subset U_\delta \subset X \setminus B$ for all $\delta \in (0,1)$:

$$f(x) = 0 \,\forall x \in A, \quad f(x) = 1 \,\forall x \in B, \quad f(x) \in [0,1] \,\forall x \in X.$$

Now it suffices to show that the pre-images of the half-lines under f are open for continuity of f. Indeed, note $f(x) < \lambda$ if and only if $x \in U_\delta$ for some $\delta < \lambda$, which is then equivalent to $x \in \bigcup_{\delta<\lambda} U_\delta$ by construction. Thus, $f^{-1}((-\infty, \lambda)) = \bigcup_{\delta<\lambda} U_\delta$ is open. Similarly, $f(x) > \lambda$ if and only if $x \notin \bigcap_{\delta>\lambda} \overline{U}_\delta = \bigcup_{\delta>\lambda} X \setminus \overline{U}_\delta$, whence $f^{-1}((\lambda, \infty)) = \bigcup_{\delta>\lambda} X \setminus \overline{U}_\delta$ is open. Then it suffices to recall the half-lines generates the topology on $\mathbb{R}$. $\square$

Lemma 1.2. *If X is a locally compact Hausdorff topological space and $U \in \mathscr{T}_X$, then for any $x \in U$, there exists a compact neighbourhood K of x such that $K \subset U$.*

Proof. By shrinking the neighbourhood(e.g. replacing U with $U \cap V$ with $V \ni x$ open precompact), we can assume without loss of generality that U is precompact. Since X is Hausdorff and in particular regular, we can separate the singleton x and ∂U by relatively open neighbourhoods $V, W \subset \overline{U}$ of x and ∂U respectively. Then $\overline{V}$ is a compact neighbourhood of x contained in U since it is disjoint from the boundary. $\square$

Lemma 1.3. *Let X be locally compact Hausdorff topological space and $K \subset U \subset X$, where K is compact and U is open. Then there exists a precompact open set $V \in \mathscr{T}_X$ such that $K \subset V \subset \overline{V} \subset U$.*

Proof. Choose for any $x \in K$, a compact neighbourhood $F(x) \subset U$. Then $\{\operatorname{int}(F(x)) \mid x \in K\}$ forms an open cover for compact K. The finite subcover $\{\operatorname{int}(F(x_i)) \mid i = 1, \dots, N\}$ gives rise to an open neighbourhood:

$$V = \bigcup_{i=1}^{N} \operatorname{int}(F(x_i)) \text{ with } K \subset V \subset \overline{V} = \bigcup_{i=1}^{N} F(x) \subset U,$$

where $\overline{V}$ is a finite union of compact sets and thus compact. $\square$

We can now prove the more general version of Urysohn's lemma on LCH spaces.

Theorem 1.1 (Urysohn's lemma on LCH spaces)**.** *Let X be a locally compact Hausdorff topological vector space and $K \subset U \subset X$ with $U \in \mathscr{T}_X$ and K compact. Then there exists $f \in C^0(X; [0,1])$ such that $f \equiv 1$ on K and $f \equiv 0$ outside a compact subset of U.*

Proof. Choose a neighbourhood V of K with $K \subset V \subset \overline{V} \subset U$ as in lemma 1.3. Then $\overline{V}$ is a compact Hausdorff space and thus normal. Indeed, for given disjoint closed subsets $E, F \subset \overline{V}$, we can separate each point in E from F by some open neighbourhoods $N(x), Z(x) \subset X$. This gives rise to an open cover $\{N(x) \mid x \in E\}$ of the compact subset E. Thus for some finite subcover, $E' = \bigcup_{j=1}^{k} N(x_j)$ and $F' = \bigcap_{j=1}^{k} Z(x_j)$ are disjoint open neighbourhoods separating E and F.
Now we can apply Urysohn's lemma for normal spaces on $\overline{V}$, which separates K and ∂V via $f \in C^0(\overline{V}; [0,1])$, i.e. $f \equiv 1$ on K and $f \equiv 0$ on ∂V. Now extend f by zero to the whole space. $\square$

Remark 1.1. I like to think of the proof as saying LCH spaces can be locally modelled as normal spaces, which then reduces to the topographical intuition for normal spaces. This result should be compared with the existence of a distance functional on Banach spaces(as a consequence of the analytic form of the Hahn-Banach theorem).

Urysohn's lemma can be used to introduce another separation axiom, which lies somewhere between T_2 and T_3:

Definition 1.2 (Completely regular ($T_{2\frac{1}{2}}$) spaces)**.** *A topological space X is completely regular if X is T_1-space and for any $A \subset X$ closed and $x \notin A$, we can find $f \in C^0(X; [0,1])$ such that $f(x) = 1$ and $f \equiv 0$ on A.*

Given apriori a continuous function defined on a compact subset in a LCH space, we can define an extension similar to the bump functions on Euclidean spaces, or more generally on smooth manifolds. The construction of which uses the Urysohn's lemma and structure of continuous functions on a general topological space.

Theorem 1.2 (Tietze extension theorem on LCH spaces)**.** *Let X be a locally compact Hausdorff topological space and $K \subset X$ compact. If $f_0 \in C^0(K)$, then we can find $f \in C^0(X)$ extending f_0, i.e. $f \equiv f_0$ on K, where we can choose f such that $f \equiv 0$ outside some compact set containing K.*

Proof. Note any continuous function is necessarily bounded on compact sets. Thus, we may as well assume $f_0(x) \in [0, 1]$ for every $x \in K$ by scaling and translation. Note K is compact in a Hausdorff space, thus closed.
Claim: There exists $g_i \in C^0(X; [0, \frac{2^{i-1}}{3^i}])$ for any $j \in \mathbb{N}$ such that $0 \leqslant f_0 - \sum_{j \leqslant i} g_j \leqslant \frac{2^i}{3^i}$ on K.
Indeed, consider the sets $C_{11} = f_0^{-1}([0, \frac{1}{3}])$ and $C_{12} = f_0^{-1}([\frac{2}{3}, 1])$, which are closed subsets in compact K thus closed in X. Using the Urysohn's lemma, we find $g_1 \in C^0(X; [0, \frac{1}{3}])$(by scaling) such that $g_1 \equiv 0$ on C_{11} and $g_1 \equiv \frac{1}{3}$ on C_{12}. An inductive argument allows us to further separate the pre-image of the Cantor splitting. Suppose $(g_i)_{i=1}^N$ have been found, we can use the same argument to find $g_{N+1} \in C^0(X; [0, \frac{2^N}{3^{N+1}}])$ such that:

$$g_{N+1} \equiv 0 \text{ on } \left\{ 0 \leqslant f_0 - \sum_{i=1}^N g_i \leqslant \frac{2^N}{3^{N+1}} \right\} \quad \text{and} \quad g_{N+1} \equiv \frac{2^N}{3^{N+1}} \text{ on } \left\{ f_0 - \sum_{i=1}^N g_i \geqslant \frac{2^N}{3^{N+1}} \right\}.$$

Now notice that $\|g_i\|_{C^0(X)} \leqslant \frac{2^i}{3^i}$ for any $i \in \mathbb{N}$, where $\sum_{i \in \mathbb{N}} \frac{2^i}{3^i} < \infty$ and in particular the tail sum is infinitesimal. Thus the partial sums $f_N = \sum_{i \leqslant N} g_i$ defines a uniformly Cauchy sequence. By completeness of the space of continuous functions, $f_N \to f = \sum_{i \in \mathbb{N}} g_i \in C^0(X; [0, 1])$ with $0 \leqslant f_0 - f \leqslant \frac{2^i}{3^i}$ for all i. $\square$

The above results give hint towards the following versatile construction. Let X be a topological space. A partition of unity on $E \subset X$ is a family of functions $\{\varphi_i \mid i \in \mathcal{I}\} \subset C^0(X; [0, 1])$ such that:

- every $x \in X$ admits a neighbourhood $U(x) \subset X$ such that $\varphi_i = 0$ on $U(x)$ for all but finitely many $i \in \mathcal{I}$;
- $\sum_{i \in \mathcal{I}} \varphi_i = 1$ pointwise everywhere on E.

We say that a partition of unity $\{\varphi_i \mid i \in \mathcal{I}\}$ of E is subordinate to an open cover $\mathscr{U}$ of E if for any $U \in \mathscr{U}$, there exists $i_U \in \mathcal{I}$ such that $\operatorname{supp} \varphi_{i_U} \subset U$. If X is a LCH space and $E \subset X$ is compact, then we can choose partition of unity consisting of only compactly supported continuous functions.

Proposition 1.2 (Compact exhaustion)**.** *Any second-countable, locally compact Hausdorff space X admits a compact exhaustion, i.e. there exists sequence of compact sets $(K_i \mid i \in \mathbb{N})$ such that $X = \bigcup_{i=1}^\infty K_i$ and $K_i \subset\subset K_{i+1}$ for all $i \in \mathbb{N}$.*

Proof. We will first show that every $x \in X$ admits a neighbourhood basis of precompact sets. Indeed, pick by local compactness of X, a compact set $K \subset X$ containing a neighbourhood $U \subset X$ of x. Now the collection $\mathcal{C}(U, x)$ of all neighbourhoods of x contained in U forms a neighbourhood basis for x since for any neighbourhood V of x, $U \cap V \in \mathcal{C}(U, x)$ is contained in V. Now since K is closed, we have any $V \in \mathcal{C}(U, x)$ is a precompact set contain in K, which thus shows that $\mathcal{C}(U, x)$ defines a neighbourhood basis of precompact sets. In particular, $\mathscr{T}_X$ admits a basis of precompact subsets. By second-coutability, we can find a countable open covering of X by precompact subsets $(U_i \mid i \in \mathbb{N})$. Now construct the compact exhaustion inductively by first setting $K_1 := \overline{U_1}$ and $k_1 = 1$. For each $i \in \mathbb{N}$, pick $k_i \in \mathbb{N}$ by compactness such that:

$$K_i \subset U_{k_1} \cup \dots \cup U_{k_i} \quad \text{since } X \subset \bigcup_{i=1}^\infty U_i.$$

Then we can set $K_{i+1} := \bigcup_{j=1}^i \overline{U_{k_j}}$ such that K_{i+1} is compact and $K_i \subset\subset K_{i+1}$. This gives us the desired exhausion by compact sets. $\square$

If X is a non-compact LCH space, we can modify X into a compact topological space by the means of Alexandroff 1-point compactification. Indeed, recall the space of continuous functions vanishing at infinity:

$$C_0(X) := \{f \in C^0(X) \mid \forall \varepsilon > 0, \{|f| \geqslant \varepsilon\} \text{ compact}\} = \overline{C_c^0(X)} \subset \mathscr{C}^0(X).$$

We can adjoin the space X with a single point ∞ at infinity such that $C_0(X)$ consists of precisely the functions such that $f(x) \to 0$ as $x \to \infty$. The compactification of X is denoted by $X_\infty = X \cup \{\infty\}$. The topology on X_∞ is defined as the collection $\mathscr{T}_\infty$ of subsets $U \subset X_\infty$ such that:

(i) either $U \subset X$ is an open subset of X, i.e. $U \in \mathscr{T}$;

(ii) or $\infty \in U$ and $X_\infty \setminus U \subset X$ is open in X.

Then the inclusion map $\iota\colon X \hookrightarrow X_\infty$ is in fact an topological embedding and $(X_\infty, \mathscr{T}_\infty)$ defines a compact topological space. Furthermore, $f \in C^0(X)$ extends continuously to X_∞ if and only if $f = f_0 + \lambda$ for some $f_0 \in C_0(X)$ and a constant $\lambda \in \mathbb{R}$, where the extension is given by setting $f(\infty) = \lambda$.

2 Positive linear functionals on $C_c^0(X)$

For the remainder of this section, X will always denote a locally compact Hausdorff topological space.

Definition 2.1 (Positive linear functionals)**.** *A linear functional $I\colon C_c^0(X) \to \mathbb{R}$ is called positive if for any $f \in C_c^0(X)$ with $f \geqslant 0$ everywhere on X, we have $I(f) \geqslant 0$.*

Positivity of linear functionals on $C_c^0(X)$ in fact implies a rather strong continuity property.

Proposition 2.1 (Local continuity)**.** *Let $I\colon C_c^0(X) \to \mathbb{R}$ be a positive linear functional on $C_c(X)$. Then for any compact subset $K \subset X$, there exists $C(K) = const.$ such that:*

$$|\langle f, I\rangle| = |I(f)| \leqslant C(K) \cdot \|f\|_{C^0(X)} \quad \textit{for any } f \in C_c^0(X) \textit{ with } \operatorname{supp} f \subset K.$$

Proof. Without loss of generality f is $\mathbb{R}$-valued. Let $K \subset X$ be compact and by Urysohn's lemma, we can choose some $\phi_K \in C_c^0(X; [0,1])$ such that $\phi_K \equiv 1$ on K. Then we have for any $f \in C_c^0(X)$ with $\operatorname{supp} f \subset K$:

$$|f| = |f\phi_K| \leqslant \phi_K \cdot \|f\| \Leftrightarrow \|f\|\phi_K \pm f \geqslant 0 \Rightarrow I(\|f\|\phi_K \pm f) = \|f\| I(\phi_K) - I(f) \geqslant 0 \Rightarrow |I(f)| \leqslant C(K)\cdot\|f\|,$$

where $C(K) = I(\phi_K)$ is a fixed constant depending only on K by Urysohn's lemma. □

If μ is a Borel measure on X with $\mu K < \infty$ for any compact $K \subset X$, then since $C_c^0(X) \subset L^1(\mu)$:

$$I_\mu = \mu\colon C_c^0(X) \to \mathbb{R}, \quad f \longmapsto \int f d\mu,$$

defines a positive linear functional on $C_c(X)$ and thus locally continuous. We will prove the Riesz-Markov-Kakutani theorem on $C_c^0(X)$, which shows that by imposing further regularity assumptions on such Borel measures, any positive linear functional on $C_c^0(X)$ arises as integration against such Borel measure.

Definition 2.2 (Regularity of Borel measure)**.** *Let μ be a Borel measure on X and $E \in \mathcal{B}(X)$. We say μ is:*

(i) outer regular on E if $\mu E = \inf\{\mu U \mid U \subset X$ open neighbourhood of $E\}$;

(ii) inner regular on E if $\mu E = \sup\{\mu K \mid K \subset E$ compact$\}$.

Definition 2.3 (Radon measure)**.** *A Radon measure on X is a Borel measure μ that is inner regular on all open sets and outer regular on all Borel sets such that $\mu K < \infty$ for any $K \subset X$ compact.*

Theorem 2.1 (Riesz-Markov-Kakutani)**.** *If $I\colon C_c^0(X) \to \mathbb{R}$ is a positive linear functional, then there exists unique Radon measure $\mu \in \mathcal{M}^+(X)$ such that:*

$$I(f) = \langle f, \mu\rangle = \int f d\mu \quad \textit{for any } f \in C_c^0(X).$$

Furthermore, we have for any open $U \subset X$ and compact $K \subset X$:

$$\mu U = \sup\{I(f) \mid f \in C_c^0(X; [0,1]), \operatorname{supp} f \subset U\} \quad \textit{and} \quad \mu K = \inf\{I(f) \mid f \in C_c^0(X), f \geqslant \chi_K\}.$$

Proof. Consider first the uniqueness statement. Suppose $\mu \in \mathcal{M}^+(X)$ is a Radon measure such that we have $I(f) = \langle f, \mu\rangle$ for any $f \in C_c^0(X)$. Then for any $U \in \mathscr{T}_X$, we obtain the following estimate:

$$f \leqslant \chi_U \text{ everywhere} \Rightarrow I(f) = \langle f, \mu\rangle \leqslant \langle \chi_U, \mu\rangle = \mu U \quad \text{for any } f \in C_c^0(X; [0,1]) \text{ with } \operatorname{supp} f \subset U.$$

For any compact subset $K \subset U$, using Urysohn's lemma, we find $f \in C_c^0(U; [0,1])$ with $f \equiv 1$ on K. Thus:

$$\chi_K \leqslant f \text{ everywhere} \Rightarrow \mu K = \langle \chi_K . \mu \rangle \leqslant \langle f, \mu \rangle = I(f).$$

Inner regularity of μ on open sets thus gives $\mu U = \sup\{\mu K \mid K \subset U \text{ compact}\} \leqslant I(f)$, which then gives the representation formula for μU in terms of the functional I. In particular, μ is determined by I on open sets, thus on all Borel sets by the Carathéodory-Hahn criterion. This gives uniqueness of the representation measure.
For proving the existence of such Radon measure, we define the candidate measure on open sets via:

$$\mu U := \sup\{I(f) \mid f \in C_c^0(X; [0,1]), \operatorname{supp} f \subset U\} \quad \text{for all } U \in \mathscr{T}_X,$$

and define a set function via the outer regularity condition for arbitrary subsets $E \subset X$:

$$\mu^* E := \inf\{\mu U \mid U \supset E,\, U \in \mathscr{T}_X\}.$$

By definition, we obtain monotonicity $\mu U \leqslant \mu V$(thus also for μ^*) whenever $U \subset V$ and $\mu = \mu*$ on open sets. We will now establish that μ^* is an outer measure and the μ^*-measurability of open sets in X. The non-negativity and monotonicity of μ^* are clear, it thus remains to show σ-subadditivity. It suffices to construct an outer approximation of μ^* by μ-measure of open subsets, i.e. $\mu^* E = \inf\{\sum_i \mu U_i \mid E \subset \bigcup_i U_i, U_i \text{ open}\}$, which defines an outer measure if we can show σ-subadditivity of μ^* on open sets.
Let $(U_i \mid i \in \mathbb{N}) \in \mathscr{T}_X$ and choose some $f \in C_c^0(X; [0,1])$ with $\operatorname{supp} f \subset U := \bigcup_{i \in \mathbb{N}} U_i$. Since $K = \operatorname{supp} f$ is compact, we can find finite subcover $(U_j)_{j=1}^N$ for K and a partition of unity $(\varphi_j)_{j=1}^N \in C_c^0(X; [0,1])$ subordinate to this subcover. Then $f = \sum_{j=1}^N f\varphi_j$ and for all $j = 1, \dots, N$, $f\varphi_j \in C_c^0(U_j; [0,1])$, which gives us:

$$I(f) = \sum_{j=1}^N I(f\varphi_j) \leqslant \sum_{j=1}^N \mu U_j \leqslant \sum_{i \in \mathbb{N}} \mu U_i \Rightarrow \mu U = \sup\{I(f) \mid f \in C_c^0(U; [0,1])\} \leqslant \sum_{i \in \mathbb{N}} \mu U_i.$$

Let $U \in \mathscr{T}_X$ and $E \subset X$ an arbitrary subset. By σ-subadditivity, we automatically have $\mu^* E \leqslant \mu^*(E \cap U) + \mu^*(E \setminus U)$, thus it suffices to show the reverse inequality for μ^*-measurability of open sets. Without loss of generality, we can assume $\mu^* E < \infty$. If E open, for any $\varepsilon > 0$, there exists $f \in C_c^0(E \cap U; [0,1])$ such that $I(f) > \mu(E \cap U) - \varepsilon$. For oen subset $V = E \setminus \operatorname{supp} f$, we can also find $h \in C_c^0(V; [0,1])$ such that $I(h) > \mu(E \setminus \operatorname{supp} f) - \varepsilon$. Since we have $f + h \in C_c^0(E; [0,1])$, it yields:

$$\mu E \geqslant I(f+g) \geqslant \mu(E \cap U) + \mu(E \setminus \operatorname{supp} f) - 2\varepsilon \geqslant \mu(E \cap U) + \mu(E \setminus U) - 2\varepsilon,$$

where since $\varepsilon > 0$ is arbitrarily chosen, we have $\mu E \geqslant \mu(E \cap U) + \mu(E \setminus U)$. For a general subset $E \subset X$, we can find an open set $V \supset E$ approximating the measure of E from outside. Thus any open set is μ^*-measurable. In particular, by the means of the Carathéodory-Hahn theorem, any Borel set is μ^*-measurable and $\mu := \mu^*|_{\mathcal{B}(X)}$ defines a Borel measure on X and satisfies the exterior approximation by I.
Let $K \subset X$ be a compact subset and $f \in C_c^0(X; [0,1])$ with $f \geqslant \chi_K$ everywhere. Set for any $\varepsilon > 0$, $U_\varepsilon := \{f > 1 - \varepsilon\} \in \mathscr{T}_X$. Then for any $h \in C_c^0(U_\varepsilon; [0,1])$, we would have:

$$(1-\varepsilon)^{-1} f - h \geqslant 0 \text{ everywhere } \Rightarrow I((1-\varepsilon)^{-1} f - h) \geqslant 0 \Rightarrow \mu U_\varepsilon = \sup_{h \in C_c^0(U_\varepsilon; [0,1])} I(h) \leqslant (1-\varepsilon)^{-1} I(f).$$

In particular, notice that $\mu K \leqslant \mu U_\varepsilon$ by monotonicity for any $\varepsilon > 0$, thus sending $\varepsilon \searrow 0$ gives $\mu K \leqslant I(f)$. On the contrary, using Urysohn's lemma, for any $U \supset K$ open neighbourhood of K, we find $f \in C_c^0(U; [0,1])$ with $f \equiv 1$ on K. In particular, we have $I(f) \leqslant \mu U$ by exterior approximation of μ with I. Thus, we can conclude, by outer regularity of μ, $\mu K \geqslant I(f)$ and $\mu K = \inf\{I(f) \mid f \in C_c^0(X), f \geqslant \chi_K\}$.
Finally, it remains to check $I(f) = \langle f, \mu \rangle$ for any $f \in C_c^0(X)$, whereas by dilation it suffices to assume f is $[0,1]$-valued. We use a monotone class argument for approximation of the integral. Let $N \in \mathbb{N}$ be fixed and set $K_j := \{f \geqslant jN^{-1}\}$ for each $j = 1, \dots, N$. Denote by $K_0 = \operatorname{supp} f$. We define $f_1, \dots f_N \in C_c^0(X)$ via:

$$f_j = \min\{\max\{f - \frac{j-1}{N}, 0\}, \frac{1}{N}\} \colon x \longmapsto \begin{cases} 0, & x \notin K_{j-1} \\ f(x) - \frac{j-1}{N}, & x \in K_{j-1} \setminus K_j \\ \frac{1}{N}, & x \in K_j. \end{cases}$$

Then we obtain $N^{-1}\chi_{K_j} \leqslant f_j \leqslant N^{-1}\chi_{K_{j-1}}$ everywhere, which leads to $N^{-1}\mu K_j \leqslant \langle f_j, \mu \rangle \leqslant N^{-1}\mu K_{j-1}$. Note also for any open $U \supset K_{j-1}$, we have $Nf_j \in C_c^0(U; [0,1])$, which then gives us $I(f_j) \leqslant N^{-1}\mu U$. We can thus use outer regularity to obtain $N^{-1}\mu K_j \leqslant I(f_j) \leqslant N^{-1}\mu K_{j-1}$. Summing over $j = 1, \dots, N$ gives:

$$\frac{1}{N}\sum_{j=1}^{N}\mu K_j \leqslant \sum_{j=1}^{N}\langle f_j, \mu\rangle = \langle f, \mu\rangle \leqslant \frac{1}{N}\sum_{j=0}^{N-1}\mu K_j \quad \text{and} \quad \frac{1}{N}\sum_{j=1}^{N}\mu K_j \leqslant \sum_{j=1}^{N} I(f_j) = I(f) \leqslant \frac{1}{N}\sum_{j=0}^{N-1}\mu K_j.$$

In particular, we can deduce the following estimate from alternating sum:

$$|I(f) - \langle f, \mu\rangle| \leqslant \frac{\mu K_0 - \mu K_N}{N} \leqslant \frac{\mu(\operatorname{supp} f)}{N} \overset{N\to\infty}{\to} 0,$$

which tells us $I(f) = \langle f, \mu\rangle$ for any $f \in C_c^0(X)$. $\square$

Remark 2.1. The Radon measure μ obtained from the Riesz-Markov-Kakutani theorem in fact comes from a complete measure μ^* defined on the space of all μ^*-measurable subsets, which is in fact an outer measure inducing the desired Radon measure μ and the completion of μ.

3 Regularity and approximation of Radon measures

Proposition 3.1 (σ-finite inner regularity)**.** *Every Radon measure is inner regular on the σ-finite subsets.*

Proof. Let μ be a Radon measure and $E \subset X$ σ-finite. Suppose first $\mu E < \infty$. Then for any $\varepsilon > 0$, we can pick open neighbourhood $U \supset E$ and compact subset $K \subset E$ such that $\mu E > \mu U - \varepsilon$ and $\mu K > \mu E - \varepsilon$. By exterior approximation, since $\mu(U \setminus E) < \varepsilon$, we can choose another open set $V \supset U \setminus E$ such that $\mu V < \varepsilon$. Set $F = K \setminus V$, which defines a compact subset of E satisfying:

$$\mu F = \mu K - \mu(K \cap V) > \mu E - \varepsilon - \mu V > \mu E - 2\varepsilon.$$

This gives inner regularity of μ on subsets of finite measures. If $\mu E = \infty$, then we can write $E = \bigcup_{i\in\mathbb{N}} E_i$ for some increasing sequence of finite measure subsets (E_i) with $\mu E_i < \infty$ and $\mu E_i \to \infty$ as $i \to \infty$. In particular, for any $N \in \mathbb{N}$, we find E_i such that $\mu E_i > N$. Inner regularity on finite measure sets gives us the existence of a compact subset $K \subset E_i \subset E$ with $\mu K > N$. In particular, μ inner regular on any σ-finite subsets. $\square$

Corollary 3.1. *Every σ-finite Radon measure is regular. In particular, every Radon measure on a σ-finite measure space is regular.*

Proposition 3.2. *Let $\mu \in \mathcal{M}^+(X)$ be a σ-finite Radon measure and $E \in \mathcal{B}(X)$. Then:*

(i) for all $\varepsilon > 0$, there exists $U \in \mathscr{T}_X$ and $F \subset X$ closed with $F \subset E \subset U$ and $\mu(U \setminus F) < \varepsilon$;

(ii) there exists F_σ-set $A = \bigcup_i A_i$ and G_δ-set $B = \bigcap_i B_i$ such that $A \subset E \subset B$ and $\mu(B \setminus A) = 0$.

Proof. (i) We have $E = \bigcup_{i=1}^{\infty} E_i$ for some disjoint sequence (E_i) with $\mu E_i < \infty$. Pick for each $i \in \mathbb{N}$, an open neighbourhood $U_i \supset E_i$ with $\mu U_i < \mu E_i + \varepsilon 2^{-(i+1)}$. Similarly, there exists disjoint sequence (G_i) with $X \setminus E = \bigcup_{i=1}^{\infty} G_i$ and open neighbourhoods $V_i \supset G_i$ with $\mu V_i < \mu G_i + \varepsilon 2^{-(i+1)}$. Summing over i gives:

$$\mu U \leqslant \sum_{i\in\mathbb{N}}\mu U_i < \sum_{i\in\mathbb{N}}\mu E_i + \varepsilon 2^{-(i+1)} = \mu E + \frac{\varepsilon}{2} \quad \text{and} \quad \mu V \leqslant \sum_{i\in\mathbb{N}}\mu V_i < \sum_{i\in\mathbb{N}}\mu G_i + \varepsilon 2^{-(i+1)} = \mu(X \setminus E) + \frac{\varepsilon}{2},$$

where we have denoted $U = \bigcup_i U_i$ and $V = \bigcup_i V_i$. Setting $F = X \setminus V \subset E$ gives a closed subset such that:

$$\mu(U \setminus F) = \mu(U \setminus E) + \mu(E \setminus F) = \mu(U \setminus E) + \mu(V \cap E) < \varepsilon.$$

(ii) follows from applying (i) with a choice of infinitesimal sequence (ε_i). $\square$

Theorem 3.1 (Borel regularity on LCH spaces)**.** *Let X be a LCH space where any open $U \subset X$ is σ-compact. Then every Borel measure on X, finite on compact subsets, defines a regular measure and thus a Radon measure.*

Proof. Let μ be a Borel measure, finite on compact sets. Thus any compactly supported continuous function is μ-integrable, i.e. $C_c^0(X) \subset L^1(\mu)$. In particular, $I(f) := \langle f, \mu \rangle$ defines a positive linear functional by the Riesz-Markov-Kakutani theorem, which is then induced from some Radon measure $\nu \in \mathcal{M}^+(X)$ such that $\langle f, \mu \rangle = \langle f, \nu \rangle$(this does not imply $\mu \equiv \nu$ yet since we do not know whether μ is a Radon measure).
If $U \in \mathscr{T}_X$, we can write $U = \bigcup_{i=1}^\infty K_i$ for some sequence of compact sets (K_i). Pick $f_1 \in C_c^0(U; [0,1])$ such that $f_1 \equiv 1$ on K_1 by using Urysohn's lemma. Choose inductively $(f_i) \in C_c^0(X; [0,1])$ with $\operatorname{supp} f_i \subset U$ for all i and $f_i \equiv \bigcup_{k=1}^i K_k \cup \bigcup_{k=1}^{i-1} \operatorname{supp} f_k$. Then it holds $f_i \nearrow \chi_U$ pointwise and thus:

$$\mu U = \lim_{i \to \infty} \langle f_i, \mu \rangle = \lim_{i \to \infty} \langle f_i, \nu \rangle = \nu U,$$

by monotone convergence theorem. For any $E \in \mathcal{B}(X)$ and $\varepsilon > 0$, we can find open neighbourhood $V \supset E$ and closed subset $F \subset E$ such that $\mu(V \setminus F) = \nu(V \setminus F) < \varepsilon$. Thus $\mu V \leqslant \mu F + \varepsilon \leqslant \mu E + \varepsilon$, giving us outer regularity. Note F is σ-compact and $\mu F \geqslant \mu E - \varepsilon$ by outer regularity, which allows us to find a sequence of compact sets (K_i) such that $\mu K_i \to \mu F$. Thus μ regular and $\mu \equiv \nu$ by uniqueness from Riesz's theorem. □

Proposition 3.3 (Density theorem)**.** *Let $\mu \in \mathcal{M}^+(X)$ be a Radon measure. Then $C_c^0(X) \subset L^p(\mu)$ dense subspace for any $p \in [1, \infty)$. In particular, $L^p(X, \mathcal{B}(X), \mu)$ is separable for any $p \in [1, \infty)$.*

Proof. Let $E \in \mathcal{B}(X)$ such that $\mu E < \infty$. Then for any $\varepsilon > 0$, we can find open neighbourhood $U \supset E$ and compact subset $K \subset E$ with $\mu(U \setminus K) < \varepsilon^p$. By Urysohn's lemma, we find $f \in C_c^0(X)$ such that $\chi_K \leqslant f \leqslant \chi_U$. Integrating this gives us $\|\chi_E - f\|_{L^p(\mu)} \leqslant \mu(U \setminus K)^{\frac{1}{p}} < \varepsilon$. Then the result follows from the L^p-density of integrable simple functions. □

Theorem 3.2 (Egoroff)**.** *Let $\mu \in \mathcal{M}^+(X)$ be a finite Borel measure and $(f_i), f \colon X \to \mathbb{C}$ Borel functions such that $f_i \to f$ pointwise a.e. Then f_i converges to f μ-almost uniformly on X, i.e. for any $\varepsilon > 0$, there exists $E \subset X$ with $\mu E < \varepsilon$ such that $f_i \to f$ uniformly on $X \setminus E$.*

Proof. Assume, upto correction on a μ-negligible set, $f_i \to f$ pointwise everywhere on X. Set for $m, k \in \mathbb{N}$:

$$E_{m,k} := \bigcup_{i \geqslant m} \{x \in X \mid |f_i(x) - f(x)| \geqslant \frac{1}{k}\}.$$

Fix $k \in \mathbb{N}$, then $(E_{m,k} \mid m \in \mathbb{N})$ forms a sequence of decreasing sets such that by pointwise convergence, $\bigcap_{m \in \mathbb{N}} E_{m,k} = \emptyset$. In particular, $\mu X < \infty$, we can conclude $\mu E_{m,k} \to 0$ as $m \to \infty$. Now for each $k \in \mathbb{N}$, pick $m(k)$ sufficiently large such that $\mu E_{m(k),k} < \varepsilon 2^{-k}$ and define $E := \bigcup_{k \in \mathbb{N}} E_{m(k),k}$. Then $\mu E \leqslant \sum_{k \in \mathbb{N}} \varepsilon 2^{-k} = \varepsilon$ and $\sup_{x \in X \setminus E} |f_i(x) - f(x)| < k^{-1}$ whenever $i > m(k)$ by construction, giving us uniform convergence. □

Theorem 3.3 (Lusin)**.** *Let $\mu \in \mathcal{M}^+(X)$ be a Radon measure and $f \colon X \to \mathbb{C}$ measurable such that we have $\operatorname{ess\,supp} f \subset E$ with $\mu E < \infty$. Then for all $\varepsilon > 0$, there exists $\varphi \in C_c^0(X)$ such that $\varphi \equiv f$ away from a set of measure $< \varepsilon$. If $f \in L^\infty(\mu)$, we can find such φ with $\|\varphi\|_{C^0(X)} \leqslant \|f\|_{L^\infty(\mu)}$.*

Proof. If $f \in L^\infty(\mu)$ supported on a set of finite measure, then f is μ-integrable and by density we find $(f_i) \in C_c^0(X)$ such that, upto passing to subsequence, $f_i \to f$ pointwise a.e. By Egoroff's theorem, we can find $A \subset E = \{f \neq 0\}$ with $\mu(E \setminus A) < \frac{\varepsilon}{3}$ and $f_i \to f$ uniformly on A. Now we can choose compact subset $K \subset A$ and open neighbourhood $U \supset E$ with $\mu(A \setminus K) < \frac{\varepsilon}{3}$ and $\mu(U \setminus E) < \frac{\varepsilon}{3}$. The uniform convergence on $K \subset A$ gives us $f|_K \in C^0(X)$. Then we can find a compactly supported continuous function $\psi \in C_c^0(X)$ where $\psi \equiv f$ on K and $\operatorname{supp} \psi \subset U$. In particular, $\{f \neq \psi\} \subset U \setminus K$, giving $\mu\{f \neq \psi\} \leqslant \mu(U \setminus K) < \varepsilon$. Post-composing h with a smooth cut-off gives us the case for bounded measurable functions.
If f unbounded, then $E_i = \{|f| \in (0, i]\} \nearrow E$ as $i \to \infty$. Thus for all i large, we have $\mu(E \setminus E_i) < \frac{\varepsilon}{2}$. Fix one such $i \in \mathbb{N}$ and by the case for L^∞-functions, we can find $\varphi_i \in C_c^0(X)$ with $\varphi_i \equiv f \chi_{A_i}$ away from a set of measure $< \frac{\varepsilon}{2}$. In particular, $\varphi_i \equiv f$ outside a set of arbitrarily small measure. □

Proposition 3.4 (Structure of l.s.c. functions)**.** *Let X be a topological space. Then the following holds:*

(i) $U \in \mathscr{T}_X \Leftrightarrow \chi_U$ *is lower-semicontinuous;*

(ii) $F \subset X$ *closed* $\Leftrightarrow \chi_F$ *upper-semicontinuous;*

(iii) $f \colon X \to \mathbb{R} \cup \{\infty\}$ *l.s.c.(or u.s.c.)* $\Rightarrow \lambda f$ *l.s.c.(or u.s.c.) for all* $\lambda \in [0, \infty)$*;*

(iv) $\mathcal{G}_>$ *family of l.s.c. functions* $\Rightarrow f := \sup \mathcal{G}_>$ *l.s.c.;*

(v) $\mathcal{G}_<$ *family of u.s.c. functions* $\Rightarrow f := \inf \mathcal{G}_<$ *u.s.c.;*

(vi) $f_1, f_2 \colon X \to \mathbb{R}$ *l.s.c.(or u.s.c.)* $\Rightarrow f_1 + f_2$ *l.s.c.(or u.s.c.).*

If we assume further that X *is a locally compact Hausdorff topological space, then:*

(I) $f \geqslant 0$ *l.s.c.* $\Rightarrow f(x) = \sup\{h(x) \mid h \in C_c^0(X),\, 0 \leqslant h \leqslant f\}$*;*

(II) $f \leqslant 0$ *u.s.c.* $\Rightarrow f(x) = \inf\{h(x) \mid h \in C_c^0(X),\, 0 \geqslant h \geqslant f\}$*.*

Proof. The usual l.s.c. and u.s.c. properties on a general topological space are routine and hence omitted here. The statements for LCH spaces are effectively consequences of Urysohn's lemma.
Indeed, for simplicity we only consider the statement for u.s.c. function $f < 0$(equality case is trivial). Let $x \in X$ and $\lambda \in (0, -f(x))$ be arbitrary. Then by upper-semicontinuity, $U(\lambda) = \{f < -\lambda\}$ is an open neighbourhood of x. Using the Urysohn's lemma, we find $h \in C_c^0(X)$ with $h(x) = -\lambda$ and $0 \geqslant h \geqslant -\lambda\chi_{U(\lambda)} \geqslant f$ everywhere. Letting $\lambda \nearrow -f(x)$ gives us statement (II). Treatment for (I) is similar. $\square$

Theorem 3.4 (l.s.c. monotone convergence theorem)**.** *Let* X *be a LCH space and* $\mathcal{G}$ *a family of non-negative l.s.c. functions as a directed set with the pre-order* $\leqslant$*, i.e. for any* $h_0, h_1 \in \mathcal{G}$*, there exists* $h \in \mathcal{G}$ *with* $g_0, g_1 \leqslant g$ *everywhere on* X*. Then for any* $\mu \in \mathcal{M}^+(X)$ *Radon measure:*

$$\int f d\mu = \langle f, \mu\rangle = \sup_{h\in\mathcal{G}} \langle h, \mu\rangle = \sup_{h\in\mathcal{G}} \int h d\mu, \quad \textit{where } f := \sup_{h\in\mathcal{G}} h.$$

Proof. By Proposition 3.4 (I), f is necessarily l.s.c. and thus Borel measurable, which then gives the well-definedness of the integral. By monotonicity of integral, it suffices to show $\langle f, \mu\rangle \leqslant \sup_{h\in\mathcal{G}}\langle h, \mu\rangle$. We construct the approximating sequence of simple functions as follows: for all $k, m \in \mathbb{N}$, we set

$$\varphi_m := 2^{-m} \sum_{k=1}^{2^{2m}} \chi_{U_{k,m}}, \quad \text{where } U_{k,m} := \{x \in X \mid f(x) > k2^{-m}\}.$$

Clearly, we have $\varphi_m \nearrow f$ pointwise a.e. as $m \to \infty$. Then by the usual monotone convergence theorem, for any $\lambda < \langle f, \mu\rangle$, we can find m sufficiently large with $2^{-m}\sum_{k=1}^{2^{2m}} \mu U_{k,m} = \langle \varphi_m, \mu\rangle \geqslant \lambda$. By the inner regularity of μ on open sets, we can find compact subsets $K_k \subset U_{k,m}$ such that $\sum_{k=1}^{2^{2m}} \mu K_k > 2^m\lambda$. We set for each m:

$$\psi_m := 2^{-m} \sum_{k=1}^{2^{2m}} \chi_{K_k} \Rightarrow f > \varphi_m \geqslant \psi_m \text{ on } K := \bigcup_{k=1}^{2^{2m}} K_k,$$

whence for each $x \in K$, we can pick $h_x \in \mathcal{G}$ with $g_x(x) > \psi_m(x)$ by definition of f. In particular, $\{V_x := \{\psi_m < h_x\} \mid x \in K\}$ forms an open cover for the compact set K by lower-semicontinuity, and thus admits a finite subcover $\{V_{x_i} \mid i = 1, \dots, N\}$. Now since $\mathcal{G}$ is a directed set, we can find an upper bound $h \in \mathcal{G}$ such that $h \geqslant \max_{k\leqslant N} h_{x_i} \geqslant \psi_m$. Then monotonicity of integral again gives $\langle h, \mu\rangle > \lambda$ for all $\lambda < \langle f, \mu\rangle$. $\square$

Corollary 3.2. $\mu \in \mathcal{M}^+(X)$ *Radon measure,* $f \colon X \to [0, \infty]$ *l.s.c. Then the following holds:*

$$\int f d\mu = \sup\left\{\int h d\mu \mid h \in C_c^0(X),\, 0 \leqslant h \leqslant f\right\}$$

Proposition 3.5 (l.s.c. approximation)**.** *Let* $\mu \in \mathcal{M}^+(X)$ *be a Radon measure and* $f \colon X \to [0, \infty]$ *Borel function. If* $\{x \mid f(x) > 0\}$ *is* σ*-finite, then:*

$$\int f d\mu = \inf\left\{\int h d\mu \mid h \geqslant f,\, h\colon X \to [0,\infty] \textit{ l.s.c.}\right\} = \sup\left\{\int h d\mu \mid 0 \leqslant h \leqslant f,\, h\colon X \to [0,\infty] \textit{ u.s.c.}\right\}$$

Proof. Let (φ_i) be sequence of non-negative simple functions such that $\varphi_i \nearrow f$ pointwise (upto correction on μ-negligible sets). Writing $\varphi_0 \equiv 0$ for convenience, we thus have $f = \sum_{i\in\mathbb{N}} \varphi_i - \varphi_{i-1}$. In particular, we can write $f = \sum_{i\in\mathbb{N}} \lambda_i \chi_{E_i}$ by reordering, where $\lambda_i > 0$ for all $i \in \mathbb{N}$. Now for all $\varepsilon > 0$, pick open neighbourhoods $U_i \supset E_i$ by outer regularity such that $\mu U_i \leqslant \mu E_i + \varepsilon\lambda_i^{-1}2^{-i}$. By the structure of l.s.c. functions, $h := \sum_{i\in\mathbb{N}} \lambda_i \chi_{U_i}$ is l.s.c. with $h \geqslant f$ everywhere and:

$$\int h d\mu = \sum_{i\in\mathbb{N}} \lambda_i \mu U_i \leqslant \liminf_{\varepsilon\downarrow 0} \sum_{i\in\mathbb{N}} \lambda_i(\mu E_i + \varepsilon\lambda_i^{-1}2^{-i}) = \sum_{i\in\mathbb{N}} \lambda_i \mu E_i = \int f d\mu,$$

whereas the reverse inequality is trivial since f is l.s.c. This gives us the first claim.
Let $\lambda < \langle f, \mu\rangle$ be arbitrary and choose N large such that $\sum_{i\leqslant N} \lambda_i \mu E_i > \lambda$. By inner regularity of μ on σ-finite sets, we can choose compact subsets $K_i \subset E_i$ approximating the measure of E_i such that $\sum_{i\leqslant N} \lambda_i \mu K_i > \lambda$. Then $h := \sum_{i\leqslant N} \lambda_i \chi_{K_i}$ is an u.s.c. function with $0 \leqslant h \leqslant f$ and $\langle h, \mu\rangle > \lambda$. □

4 Dual of continuous functions vanishing at infinity

Let X be a locally compact Hausdorff topological space. Any positive bounded linear functional $I\colon C_c^0(X) \to \mathbb{R}$ has a natural continuous extension to the uniform closure $C_0(X)$ via a density argument. In particular, recall by the Riesz-Markov-Kakutani theorem, I is given by integration against a Radon measure $\mu \in \mathcal{M}^+(X)$. Since the constant 1 function is clearly continuous and thus lower-semicontinuous, we can use 3.2 to approximate the total measure of X and the boundedness of our linear functional I gives:

$$\mu X = \sup_{\varphi\in C_c^0(X;[0,1])} \int \varphi d\mu \leqslant \|I\|_{(C_c^0(X))^*} \sup_{\varphi\in\mathbb{B}_{C_c(X)}} \|\varphi\|_{C^0(X)} = \|I\|_{(C_c^0(X))^*} < \infty.$$

Thus we can deduce that any finite Radon measure on X induces a positive bounded linear functional on $C_0(X)$ via extension through the dense subset $C_c^0(X) \subset C_0(X)$. We will see that such bounded linear functionals in fact give hint to a complete characterisation of the dual space $C_0(X)^*$, which is the content of the Riesz representation theorem. The result crucially relies on a Jordan decomposition of the elements in $C_0(X)^*$.

Lemma 4.1 (Jordan decomposition of continuous linear functionals)**.** *If $I \in C_0(X)^*$, then there exists positive linear functionals $I^\pm \in C_0(X)^*$ such that $I = I^+ - I^-$.*

Proof. Define for $f \in C_0(X;[0,\infty))$, $I^+(f) := \sup\{I(\psi) \mid \psi \in C_0(X), 0 \leqslant \psi \leqslant f\}$. For any continuous compactly supported $0 \leqslant h \leqslant f$, we can deduce from linearity:

$$|I(h)| \leqslant \|I\|_{C_0(X)^*}\|h\|_{C^0(X)} \leqslant \|I\|_{C_0(X)^*}\|f\|_{C^0(X)} \Rightarrow 0 \leqslant I^+(f) \leqslant \|I\|_{C_0(X)^*}\|f\|_{C^0(X)}.$$

Inspired from the proof for linearity of integral, we will show that I^+ is the restriction of some linear functional to $C_0(X;[0,\infty))$. Indeed, for any $f \in C_c^0(X;[0,\infty))$ and $\lambda \geqslant 0$, we clearly have $I^+(\lambda f) = \lambda I^+(f)$ from linearity of I. If $0 \leqslant \psi_0 \leqslant f_0$ and $0 \leqslant \psi_1 \leqslant f_1$, then we obtain by definition:

$$0 \leqslant \psi_0 + \psi_1 \leqslant f_0 + f_1 \Rightarrow I^+(f_0 + f_1) \geqslant \sup_{0\leqslant\psi_i\leqslant f_i|i=0,1} I(\psi_0) + I(\psi_1) = I^+(f_0) + I^+(f_1).$$

For any choice of $0 \leqslant \psi \leqslant f_0+f_1$, we can set $\psi_0 := \max\{\psi, f_0\}$ and $\psi_1 := \psi-\psi_0$. Then we have $0 \leqslant \psi_i \leqslant f_i$ for $i = 0, 1$, which allows us to conclude $I(\psi) = I(\psi_0)+I(\psi_1) \leqslant I^+(f_0)+I^+(f_1)$ by definition. In particular, since ψ is chosen arbitrarily, $I^+(f_0 + f_1) = I^+(f_1) + I^+(f_2)$. Thus, we can define a linear extension by decomposing every $f \in C_0(X)$ into positive and negative parts, i.e. $I^+(f) = I^+(f_+) - I^+(f_-)$. In particular:

$$|I^+(f)| = |I^+(f_+) - I^+(f_-)| \overset{I^+\geqslant 0}{\leqslant} \max\{I^+(f_+), I^+(f_-)\} \leqslant \|I\|_{C_0(X)^*}\|f\|_{C^0(X)} \quad \text{for all } f \in C_0(X).$$

Hence, $I^+ \in C_0(X)^*$ with $\|I^+\|_{C_0(X)^*} \leqslant \|I\|_{C_0(X)^*}$, whereas we can set, by the linear structure of the dual space, $I^- := I - I^+ \in C_0(X)^*$. It is easy to see that I is necessarily positive. □

Remark 4.1. By considering the real and imaginary parts separately, we can carry out the analogous discussion in the proof above for $I \in C_0(X;\mathbb{C})^*$. In particular, in the view of the Riesz-Markov-Kakutani theorem, for any continuous $\mathbb{C}$-linear functionals $I \in C_0(X;\mathbb{C})$, there are finite Radon measures $(\mu_i)_{i=1}^4 \in \mathcal{M}^+(X)$ such that $I(f) = \langle f, \mu\rangle$ for any $f \in C_0(X;\mathbb{C})$, where we have denoted the $\mathbb{C}$-valued representation Radon measure by $\mu = (\mu_1 - \mu_2) + i(\mu_3 - \mu_4) \in \mathcal{M}(X;\mathbb{C})$.

Definition 4.1 (Vector-valued Radon measures)**.** *Let $(V, \|\cdot\|)$ be a finite-dimensional normed space. A set function $\mu\colon \mathcal{B}(X) \to V$ is called a vector-valued Radon measure if* *(i)* $\mu\emptyset = 0$*;*

(ii) $\mu(\bigcup_{i\in\mathbb{N}} E_i) = \sum_{i\in\mathbb{N}} \mu E_i$ *for any* $(E_i) \in \mathcal{B}(X)$ *disjoint;*

(iii) each μ^i *is a signed Radon measure on* X*.*

We denote the space of vector-valued Radon measures by $\mathcal{M}(X;V)$*. In particular, since complex measures are bounded, every complex Borel measure is Radon on a second-countable LCH space. The total variation measure induce a norm on* $\mathcal{M}(X;V)$ *by setting for all* $\mu \in \mathcal{M}(X;V)$*,* $\|\mu\|_{\mathcal{M}} := |\mu|(X)$*. Furthermore,* $(\mathcal{M}(X;V), \|\cdot\|_{\mathcal{M}})$ *forms a normed vector space, which is in particular closed under linear operations.*

Theorem 4.1 (Riesz representation theorem)**.** *Let* X *be a locally compact Hausdorff topological space. Define:*

$$T\colon \mathcal{M}(X;\mathbb{C}) \longrightarrow C_0(X;\mathbb{C})^*, \quad \mu \longmapsto I_\mu := \langle\cdot,\mu\rangle.$$

Then $T\colon \mu \mapsto I_\mu$ *is an isometric isomorphism between* $(\mathcal{M}(X;\mathbb{C}), \|\cdot\|_{\mathcal{M}})$ *and* $(C_0(X;\mathbb{C})^*, \|\cdot\|_{op})$*.*

Proof. From the lemma and remark above, we know that any $I \in C_0(X;\mathbb{C})^*$ is induced by some complex Radon measure. On the contrary, for $\mu \in \mathcal{M}(X;\mathbb{C})$, we can see for any $E \in \mathcal{B}(X)$:

$$\mu E = \int \chi_E \left(\frac{d\mu}{d|\mu|}\right) d|\mu| \leqslant \int \left|\chi_E \frac{d\mu}{d|\mu|}\right| d|\mu| = |\mu|(E),$$

where we have used the polar decompostion of μ and $|\frac{d\mu}{d|\mu|}| = 1$ $|\mu|$-a.e. Then via a monotone class argument:

$$|\int f d\mu| \leqslant \int |f| d|\mu| \leqslant \|f\|_{C^0(X)}\|\mu\|_{\mathcal{M}} \quad \text{for any } f \in C_0(X;\mathbb{C}).$$

In particular, in combination with the linearity of integral, we have $I_\mu \in C_0(X;\mathbb{C})^*$ where $\|I_\mu\|_{op} \leqslant \|\mu\|_{\mathcal{M}}$. It now remains to show the reverse inequality. Using the Lusin's theorem, for any $\varepsilon > 0$, we can find some $\varphi \in C_c^0(X) \subset C_0(X)$ and measurable subset $E \subset X$ such that $\overline{\frac{d\mu}{d|\mu|}} \equiv \varphi$ on E and $\mu(X \setminus E) \leqslant \varepsilon$. Thus:

$$\|\mu\|_{\mathcal{M}} = |\mu|(X) = \int \left|\frac{d\mu}{d|\mu|}\right|^2 d|\mu| = \int \overline{\frac{d\mu}{d|\mu|}} d\mu \leqslant |\int \varphi d\mu| + \int \left|\varphi - \overline{\frac{d\mu}{d|\mu|}}\right| d\mu \leqslant \|I_\mu\|_{op} + 2\varepsilon,$$

where we could normalise $\varphi \in \mathbb{B}_{C_c^0(X)}$ since the polar of μ has unit length. Hence $\mu \mapsto I_\mu$ is an isometry. □

Recall that the continuous dual space of $L^\infty(\mu)$ consists of elements which are finitely additive complex measures absolutely continuous with respect to μ. This in turn characterises the second dual space of $L^1(\mu)$. Clearly, since $C_0(X) \hookrightarrow L^\infty(\mu)$, the embedding implies the canonical inclusion of dual spaces $(L^\infty(\mu))^* \hookrightarrow C_0(X)^*$ via restriction. In particular, we can construct the canonical embedding into the second dual space:

$$\Lambda_{L^1(\mu)}\colon L^1(X,\mathcal{B}(X),\mu) \hookrightarrow C_0(X;\mathbb{C})^* = \mathcal{M}(X;\mathbb{C}), \quad f \longmapsto \nu_f := f d\mu.$$

5 Lower-semicontinuity of functionals on Radon measures and generalised products

We discuss the lower-semicontinuity and continuity with respect to the weak-* topology of functional defined on Radon measures. Some preliminary facts about the interactions between continuous functions and Radon measures on a given measure space need to be recalled. A useful compactness criterion is given below, which becomes useful in the discussion of varifold geometry.

Definition 5.1 (Total variation measure)**.** *Let* X *be a locally compact Hausdorff topological space and* $\nu \in \mathcal{M}(X;V)$ *a vector-valued measure for some normed vector space* $(V, \|\cdot\|_V)$*. We define the corresponding total variation measure* $|\nu|$ *on the Borel subsets of* X*:*

$$|\nu|(E) := \sup\left\{\sum_{i=1}^{\infty} \|\nu E_i\| \mid E = \bigcup_{i=1}^{\infty} E_i \text{ for some pairwise disjoint } (E_i) \in \mathcal{B}(X)\right\} \quad \text{for any } E \in \mathcal{B}(X).$$

Proposition 5.1 (Equivalent formulation of total variation measure)**.** *Let X be a locally compact separable metric space and $\mu \in \mathcal{M}(X;\mathbb{R}^m)$ a finite Radon measure. Then for any open subset $U \subset \mathbb{R}^n$:*

$$|\mu|(U) := \sup\left\{ \int_X \varphi^i d\mu_i \,\middle|\, \varphi = (\varphi^1,\ldots,\varphi^m) \in C^0_c(U;\mathbb{R}^m) \textit{ with } \|\varphi\|_{C^0(X)} \leqslant 1 \right\}.$$

Proof. Note $\mu << |\mu|$ and by the Radon-Nikodym theorem, we obtain the polar decomposition $d\mu = fd|\mu|$ for some μ-integrable function $f\colon X \to \mathbb{S}^{m-1} \subset \mathbb{R}^m$. Now fix an arbitrary open set $U \subset X$:

$$\left|\int_U \varphi \cdot d\mu\right| \leqslant \int_U \langle \varphi, f\rangle_{Euc} d|\mu| \leqslant \|\varphi\|_{C^0(X)} \cdot |\mu|(U), \quad \text{for all } \varphi \in C^0_c(U),$$

by the decomposition of measure and Cauchy-Schwarz inequality. Choosing $\varphi \in \mathbb{B}_{C^0_c(X;\mathbb{R}^m)}$ gives us the backward inequality. On the contrary, recall the dense injection $C^0_c(U;\mathbb{R}^m) \subset L^1(A,\mathcal{B}(A),\mu)$. We can thus find an approximating sequence $(\varphi_i) \in \mathbb{B}_{C^0_c(U;\mathbb{R}^m)}$ such that $\varphi_i \to f\chi_U$ in L^1. In particular, we have:

$$\lim_{i\to\infty} \int_U \varphi_i^j d\mu_j = \lim_{i\to\infty}\int_U \langle \varphi_i, f\rangle_{Euc} d|\mu| = |\mu|(U);$$

by dominated convergence theorem, since f is $\mathbb{S}^{m-1}$-valued and L^1-convergence gives us $\varphi_i \to f\chi_U$ μ-a.e. $\square$

Remark 5.1. The above characterisation of the total variation measure allows us to give an explicit Radon-Nikodym decomposition. Suppose $\nu << \mu$, whence by the Lebesgue-Radon-Nikodym theorem, we obtain:

$$|\nu|(U) = \sup_{\varphi\in C^0_c(U;\mathbb{R}^m);\|\varphi\|_\infty\leqslant 1} \int_X \langle \varphi, d\nu\rangle_{Euc} = \sup_{\varphi\in C^0_c(U;\mathbb{R}^m);\|\varphi\|_\infty\leqslant 1} \int_X \langle \varphi, \frac{d\nu}{d\mu}\rangle_{Euc} d\mu = \int_X \chi_U |\frac{d\nu}{d\mu}| d\mu,$$

for any open subset $U \subset X$ via an argument using the dense injection $C^0_c(U;\mathbb{R}^m) \hookrightarrow L^1(U;\mathbb{R}^m)$.

Proposition 5.2 (De La Vallée Poussin compactness criterion)**.** *Let $(\mu_i) \in \mathcal{M}(X;\mathbb{R}^m)$ be a sequence of finite Radon measures such that $\sup_{i\in\mathbb{N}} |\mu_i|(X) < \infty$. Then:*

(i) (μ_i) admits a weak- convergent subsequence;*

(ii) $\mu \longmapsto |\mu|(X)$ is lowe-semicontinuous with respect to the weak- topology.*

Proof. Suppose without loss generality, e.g. via scaling, $\sup_{i\in\mathbb{N}} |\mu_i|(X) \leqslant 1$. Recall that the space of continuous functions on a LCH topological space is separable. Thus we can find a sequence $(u_i) \in C_0(X;\mathbb{R}^m)$ (argue componentwise if necessary) such that $\mathrm{clin}_{\mathbb{R}}(\{u_i \mid i \in \mathbb{N}\}) = C_0(X;\mathbb{R}^m)$.
After normalising each u_i, use Cantor's diagonal argument to extract a subsequence (μ_{i_k}) such that:

$$\langle u_m, \mu_{i_k}\rangle = \int_X u_m \, d\mu_{i_k} \to \lambda_m \text{ as } k \to \infty, \quad \text{where } \sup_{m\in\mathbb{N}} |\lambda_m| \leqslant 1.$$

By the linearity of integral and limits, we can can use the density assumption to define via duality:

$$\langle u, \mu\rangle = \lim_{k\to\infty} \langle u, \mu_{i_k}\rangle, \quad \text{for all } u \in C_0(X;\mathbb{R}^m),$$

where the limit exists via a $\varepsilon/3$-argument. In particular, we necessarily have $\mu \in (C_0(X;\mathbb{R}^m))^* \equiv \mathcal{M}(X;\mathbb{R}^m)$ and $\mu_{i_k} \overset{*}{\rightharpoonup} \mu$ by construction. The lower-semicontinuity statement follows from the equivalent definition of the total variation measure as the supremum of a family of weak-* continuous functionals. As a consequence, we necessarily have $|\mu|(X) \leqslant \liminf_{k\to\infty} |\mu_{i_k}|(X) \leqslant 1$. $\square$

Definition 5.2 (Weak-* convergence)**.** *Let $(\mu_i) \in \mathcal{M}_{\mathrm{loc}}(X;\mathbb{R}^m)$. We say that $\mu_i \overset{*}{\rightharpoonup} \mu$ locally on X if*

$$\int_X u d\mu_i \to \int_X u d\mu \quad \textit{for any } u \in C_c(X).$$

For finite Radon measures $\mu, (\mu_i) \in \mathcal{M}_{\mathrm{loc}}(X;\mathbb{R}^m)$, we say $\mu_i \overset{*}{\rightharpoonup} \mu$ if

$$\int_X u d\mu_i \to \int_X u d\mu \quad \textit{for any } u \in C_0(X) := \overline{C_c(X)} \subset \mathscr{C}^0(X).$$

Remark 5.2. We can see that, on locally compact Hausdorff spaces, for $\mu, (\mu_i) \in \mathcal{M}_{\mathrm{loc}}(X;\mathbb{R}^m)$ finite Radon measures, $\mu_i \xrightharpoonup{*} \mu$ if and only if $\mu_i \xrightharpoonup{*} \mu$ locally and $\sup_{i\in\mathbb{N}} |\mu_i|(X) < \infty$ in addition.
Indeed, we have that, in a locally compact Hausdorff topological space, $C_c(X) \subset L^p(\mu)$ is $\|\cdot\|_{L^p}$-dense for any $p \in [1,\infty)$. Thus for any Borel partition, we can approximate the constant-1 function on each partition set by compactly supported continuous functions in the L^1-sense, which then gives us $\sup_{i\in\mathbb{N}} |\mu_i|(X) < |\mu|(X)$ by convergence of the measure of the whole space. The converse direction follows since uniform density implies L^1-density, which then allows us to use a $\varepsilon/3$-argument to conclude. For simplicity, we present a prototypical criterion for weak-* convergence in $\mathcal{M}(\mathbb{R};\mathbb{C})$ below.

Lemma 5.1 (Integration by parts formula)**.** *Let $\mu \in \mathcal{M}(\mathbb{R};\mathbb{C})$. Then the following formula holds:*

$$\int \psi d\mu = \int \psi'(t)\mu(-\infty,t]dt = \int \psi'(t)F(t)dt \quad \text{for all } \psi \in C_c^1(\mathbb{R};\mathbb{C}).$$

Proof. Indeed, since ψ is compactly supported, we necessarily have $\lim_{|x|\to\infty} \psi(x) = 0$. Thus using the FTC:

$$\psi(s) = \int_{-\infty}^{s} \psi'(t)dt \Rightarrow \int \psi(s)d\mu(s) = \iint \psi'(t)\chi_{(-\infty,s]}(t)\,dt\,d\mu(s) = \int \psi'(t)\mu(-\infty,t]dt,$$

where we have used Fubini's theorem to justify changing order of integration. □

Proposition 5.3 (Convergence in distribution)**.** *Suppose $(\mu_i), \mu \in \mathcal{M}(\mathbb{R})$ and consider their respective distribution functions $F_i(x) := \mu_i(-\infty,x]$ and $F(x) := \mu(-\infty,x]$ on $\mathbb{R}$. Then:*

(i) $\sup_{i\in\mathbb{N}} \|\mu_i\|_{\mathcal{M}} = \sup_{i\in\mathbb{N}} |\mu|(X) < \infty$ *and $F_i \to F$ at all continuity points of $F \Rightarrow \mu_i \xrightharpoonup{*} \mu$;*

(ii) $\mu_i \xrightharpoonup{*} \mu$ *and $(\mu_i) \in \mathcal{M}^+(X) \Rightarrow F_i \to F$ at continuity points of F.*

Proof. (i) Denote by $F(x_+) := \lim_{y\uparrow x} F(y)$ and $F(x_-) := \lim_{y\downarrow x} F(y)$. Suppose first that μ is a $\mathbb{R}$-valued signed measure and note F is clearly increasing by the monotonicity of integral. Thus $\{(F(x_-), F(x_+)) \mid x \in \mathbb{R}\}$ forms a disjoint collection of intervals such that $F(-N) \leqslant F(x_-) \leqslant F(x_+) \leqslant F(N)$ if $|x| < N$. Then:

$$\mu\{x \in (-N,N) \mid F(x_+) \neq F(x_-)\} = \sum_{|x|<N} |F(x_+) - F(x_-)| \leqslant F(N) - F(-N) < \infty \quad \text{for any } N \in \mathbb{N}.$$

Thus $|\{x \mid F(x_+) \neq F(x_-)\}|$ is at most countable. In particular, any distribution function is necessarily a function of bounded variation and thus F is continuous away from at most countably many points. In particular, the assumption implies $F_i \to F$ $\mathcal{L}^1$-a.e. and $\sup_{i\in\mathbb{N}} \|F_i\|_{L^\infty(\mathbb{R})} \leqslant \sup_{i\in\mathbb{N}} |\mu_i|(\mathbb{R}) < \infty$. Recall $C_c^1(\mathbb{R}) \subset C_0(\mathbb{R})$ is a dense subspace and the integration by parts formula gives for any $\varphi \in C_c^1(\mathbb{R})$:

$$\int \varphi d\mu_i = \int \varphi'(t)F_i(t)dt \longrightarrow \int \varphi'(t)F_i(t)dt = \int \varphi d\mu,$$

by Lebesgue's dominated convergence theorem. A 3ε-argument and density allow us to conclude $\mu_i \xrightharpoonup{*} \mu$.
(ii) By weak-* convergence, $(\langle f,\mu_i\rangle)_{i\in\mathbb{N}}$ defines a Cauchy and thus bounded sequence for any $f \in C_0(X)$. By the principle of uniform boundedness, $\sup_{i\in\mathbb{N}} \|\mu_i\| < \infty$. Let $x \in \mathbb{R}$ be a continuity point for the distribution function F of μ. For any $\varepsilon > 0$, we can define $f_x, f^x \in C_c^0(\mathbb{R})$ via the following:

$$f_x(y) := \begin{cases} 1, & y \in [-N+\varepsilon, x]; \\ 1 - \frac{y-x}{\varepsilon} & y \in (x, x+\varepsilon); \\ 0 & \text{otherwise.} \end{cases} \quad \text{and} \quad f^x(y) := \begin{cases} 1, & y \in [-N+\varepsilon, x-\varepsilon]; \\ 1 - \frac{y-(x-\varepsilon)}{\varepsilon} & y \in (x-\varepsilon, x); \\ 0 & \text{otherwise.} \end{cases}$$

We obtain the following from the weak-* convergence $\mu_i \xrightharpoonup{*} \mu$:

$$\limsup_{i\to\infty} F_i(x) - F_i(-N+\varepsilon) = \limsup_{i\to\infty} \mu_i(-N+\varepsilon, x] \leqslant \int f_x d\mu \leqslant F(x+\varepsilon) - F(-N+\varepsilon)$$
$$\liminf_{i\to\infty} F_i(x) - F_i(-N) = \limsup_{i\to\infty} \mu_i(-N, x] \geqslant \int f^x d\mu \geqslant F(x-\varepsilon) - F(-N+\varepsilon)$$

whereas if we send $N \to \infty$, since $F_i(-N), F(-N \pm \varepsilon) \to 0$, it yields for $\varepsilon > 0$ arbitrarily small:

$$F(x-\varepsilon) \leqslant \liminf_{i\to\infty} F_i(x) \leqslant \limsup_{i\to\infty} F_i(x) \leqslant F(x+\varepsilon).$$

Since x is chosen to be a continuity point, we can send $\varepsilon \searrow 0$ to obtain $\lim_{i\to\infty} F_i(x) = F(x)$. □

Another useful result describes a generalised notion of product measures, allowing us to decompose measures on product spaces and giving hints towards the method of disintegration. We briefly describe the approach below.

Definition 5.3 (Measure-valued maps)**.** *Let $E \subset \mathbb{R}^n$, $F \subset \mathbb{R}^m$ be open sets, $\mu \in \mathcal{M}^+(E)$ Radon measure. We say a function $x \mapsto \nu_x$, assigning $x \in E$ to a finite Radon measure $\nu_x \in \mathcal{M}(F;\mathbb{R}^k)$, is μ-measurable if $x \mapsto \nu_x B$ is μ-measurable map in the usual sense for any $B \in \mathcal{B}(F)$.*

Proposition 5.4 (Measurability criterion)**.** *If $x \mapsto \nu_x A$ is μ-measurable for any open set $A \subset F$, then $x \mapsto \nu_x$ is a μ-measurable measure-valued map. Furthermore, $x \mapsto \int_F h(x,y) d\nu_x(y)$ is μ-measurable for any bounded $(\mathcal{B}_\mu(E) \otimes \mathcal{B}(F))$-measurable function $h\colon E \times F \to \mathbb{R}$.*

Proof. The first part is an exercise using the Caratheodory-Hahn extension. The second part follows from a monotone class argument and checking μ-measurablility of the rectangle sets, which is clear from definition. $\square$

Remark 5.3. In particular, if $x \mapsto \nu_x$ is μ-measurable and $A \subset F$ open, we obtain:

$$|\nu_x|(A) = \sup\left\{ \int_F \sum_{i=1}^k u^i d(\nu_x)^i \mid u \in D \subset \mathbb{B}_{C_c(A;\mathbb{R}^k)} \right\},$$

where D is a dense subset in the unit ball. This gives μ-measurability of $x \mapsto |\nu_x|$.

Definition 5.4 (Generalised product measure)**.** *Assuming the following holds true:*

$$\int_{E'} |\nu_x|(F) d\mu(x) < \infty \text{ for any } E' \subset\subset E \text{ open.}$$

The product Radon measure $\mu \otimes \nu_x \in \mathcal{M}(E \times F;\mathbb{R}^k)$ is defined via the following: for any $B \in \mathcal{B}(K \times F)$:

$$\mu \otimes \nu_x(B) := \int_E \left(\int_F \chi_B(x,y) d\nu_x(y) \right) d\mu(x),$$

where $K \subset E$ is an arbitrary compact set. In particular, we have for any Borel function $f\colon E \times F \to \mathbb{R}$ with $\operatorname{supp} f \subset E' \times F$ *with $E' \subset\subset E$, the following integration formula holds:*

$$\int_{E \times F} f(x,y) d(\mu \otimes \nu_x)(x,y) = \int_E \left(\int_F f(x,y) d\nu_x(y) \right) d\mu(x).$$

Theorem 5.1 (Disintegration of measures on product spaces)**.** *Let $k \geqslant 1$, $E \subset \mathbb{R}^n$ and $F \subset \mathbb{R}^m$ open subsets, $\nu \in \mathcal{M}(E \times F;\mathbb{R}^k)$ Radon measure, $\pi\colon E \times F \to E$ projection onto E. Denote $\mu = \pi_*|\nu|$, which is assumed to be a Radon measure, i.e. $|\nu|(K \times F) < \infty$ for any $K \subset E$ compact. Then there exist μ-measurable map:*

$$E \to \mathcal{M}(F;\mathbb{R}^m) \quad x \longmapsto \nu_x,$$

mapping onto the space of $\mathbb{R}^m$-valued finite Radon measures such that:

(i) $|\nu_x| \in \mathcal{P}(F)$, *i.e.* $|\nu_x|(F) = 1$ *for μ-a.e. $x \in E$;*

(ii) $f(x,\cdot) \in L^1(F,|\nu_x|)$ *for μ-a.e. $x \in E$;*

(iii) $x \longmapsto \int_F f(x,y) d\nu_x(y) \in L^1(E,\mu)$ *for any $f \in L^1(E \times F, |\nu|)$;*

(iv) $\int_{E \times F} f(x,y) d\nu(x,y) = \int_E \left(\int_F f(x,y) d\nu_x(y) \right) d\mu(x)$ *for any $f \in L^1(E \times F, |\nu|)$.*

Moreover, any other μ-measurable map $x \mapsto \nu'_x$ satisfying the above for any bounded, compactly supported Borel function and $\nu'_x F \in L^1(E,\mu)$ satisfies $\nu_x = \nu'_x$ for μ-a.e. $x \in E$.

Corollary 5.1. *The total variation of the given Radon measure satisfies $|\nu| = \mu \otimes |\nu_x|$.*

Recall the definition of a convex function defined on a vector space X: $f\colon X \to \mathbb{R}_\infty = \mathbb{R} \cup \{\pm\infty\}$ is convex if

$$f(\lambda x + (1-\lambda)y) \leqslant \lambda f(x) + (1-\lambda) f(y) \quad \text{for any } x,y \in X, \lambda \in [0,1].$$

We say that f is lower-semicontinuous if for any $x_0 \in X$ and $t \in (-\infty, f(x_0))$, there exists open neighbourhood $U(x_0) \subset X$ of x_0 such that $f(x) < t$ for any $x \in U(x_0)$. Equivalently, f lower-semicontinuous if:

$$f(x_0) \leqslant \liminf_{x \to x_0} f(x) \text{ for any } x_0 \in X.$$

Proposition 5.5 (Equivalent formulations of lower-semicontinuity)**.** *Let X be a separable Banach space and $f\colon X^* \to \mathbb{R} \cup \{\infty\}$ convex. Then:*

(i) f weak- lower-semicontinuous $\Leftrightarrow$ f sequentially weak-* lower-semicontinuous;*

(ii) f weak- lower-semicontinuous $\Leftrightarrow$ there exists $(x_i) \in X$ and $(\lambda_i) \in \mathbb{R}$ such that*

$$f(\phi) = \sup_{i \in \mathbb{N}} \phi(x_i) + \lambda_i \text{ for any } \phi \in X^*.$$

In particular, f is positively 1-homogeneous if and only if $\lambda_i = 0$ for all $i \in \mathbb{N}$.

Proof. (i) Note by definition, lower-semicontinuity always implies sequential lower-semicontinuity with respect to any topology. Now let f be sequentially weak-* lower-semicontinuous and convex. Consider the sets $C_t = \{f \leqslant t\}$, which are sequentially weak-* closed and convex by our assumption on f. Note the intersection of C_t with any closed balls in X is necessarily weak-* closed. Then by the Krein-Šmulian theorem, any C_t is weak-* closed, which implies the weak-* lower-semicontinuity of f.
(ii) Assume without loss of generality $f \not\equiv \infty$. Now consider the set:

$$\mathcal{R} := \{(x, \lambda) \in X \times \mathbb{R} \mid f(\phi) \geqslant \phi(x) + \lambda \,\forall \phi \in X^*\},$$

which reduces to proving for any $\phi \in X^*$, $f(\phi) = \sup\{\phi(x) + \lambda \mid (x, \lambda) \in \mathcal{R}\}$, which is equivalent to:

$$\mathrm{Epi}(f) := \{(\phi, t) \in X^* \times \mathbb{R} \mid t \geqslant f(\phi)\} \subset \bigcup_{(x,\lambda) \in \mathcal{R}} \{(\phi, t) \in X^* \times \mathbb{R} \mid t < \phi(x) + \lambda\}.$$

Indeed, since $X^* \times \mathbb{R}$ is σ-compact with respect to the weak-* topology, the open cover above can be refined into a countable cover, whence $f(\phi) = \sup_{i \in \mathbb{N}} \phi(x_i) + t_i$ for any $\phi \in X^*$.
Suppose for now f is a positive functional. Choose $\varphi \in X^*$ and $s < f(\varphi)$, thus it follows $(\varphi, s) \notin \mathrm{Epi}(f)$. Using the functional separation theorem, we can find $z \in X$ and $\alpha, \beta \in \mathbb{R}$ such that

$$\langle \phi, z \rangle + \beta t > \alpha > \langle \varphi, z \rangle + \beta s \text{ for any } (\phi, t) \in \mathrm{Epi}(f) \Rightarrow \langle \psi, z \rangle + \beta f(\psi) \geqslant \alpha > \langle \varphi, z \rangle + \beta s \,\forall \psi \in \{f < \infty\}.$$

Now choose some $\psi \in \{f < \infty\}$ and $t > f(\psi)$, sending $t \to \infty$ gives $\beta \geqslant 0$. If $\beta = 0$, we would have $f(\varphi) = \infty$ otherwise we obtain $\varphi(z) > \alpha > \varphi(z)$ from above which is a contradiction. Moreover, for any $\varepsilon > 0$, $\psi \in \{f < \infty\}$, we have $\psi(\frac{-z}{\varepsilon}) + \frac{a}{\varepsilon} < 0 \leqslant f(\psi)$, which then implies $(\frac{-z}{\varepsilon}, \frac{\alpha}{\varepsilon}) \in \mathcal{R}$. Thus:

$$\varphi(\frac{-z}{\varepsilon}) + \frac{\alpha}{\varepsilon} \to \infty \text{ as } \varepsilon \downarrow 0 \Rightarrow f(\phi) = \sup_{(x,\alpha) \in \mathcal{R}} \phi(x) + \alpha \text{ for all } \phi \in X^*.$$

If $\beta > 0$, then using the inequality from the functional separation theorem, we obtain for any $\psi \in X^*$:

$$\psi(z)\beta f(\psi) > \varphi(z) + \beta s \Leftrightarrow f(\psi) > \psi(\frac{-z}{\beta}) + \frac{\beta s + \varphi(z)}{b} \Rightarrow (\frac{-z}{\beta}, \frac{\beta s + \varphi(z)}{b}) \in \mathcal{R}.$$

In particular, for any arbitrary $s < f(\varphi)$, we have:

$$\sup_{(x,\alpha) \in \mathcal{R}} \varphi(x) + \alpha \geqslant \varphi(\frac{-z}{\beta}) + \frac{\beta s + \varphi(z)}{\beta} = s \Rightarrow f(\varphi) = \sup\{\varphi(x) + \alpha \mid (x, \alpha) \in \mathcal{R}\}.$$

For a general functional $f\colon X^* \to \mathbb{R} \cup \{\infty\}$, choose $\varphi \in \{f < \infty\}$ and $s < f(\varphi)$ arbitrary. Apply the functional separation theorem as above to get $z \in X$ and $\beta \in \mathbb{R}$ with $\psi(z) + \beta t > \varphi(z) + \beta s$ for any pair $(\psi, t) \in \mathrm{Epi}(f)$. In particular, we have $(\varphi, f(\varphi)) \in \mathrm{Epi}(f)$ by our choice, thus giving us $\beta > 0$. Now set:

$$\tilde{f}\colon X^* \to \mathbb{R} \cup \{\infty\}, \quad \phi \longmapsto f(\phi) + \frac{\phi(z)}{\beta} - \frac{\beta s + \varphi(z)}{\beta},$$

which is a positive linear functional on X^*, since we have for fixed $\phi \in X^*$:

$$\phi(z) + \beta f(\phi) = \phi(z) + \inf\{\beta t \mid (\phi, t) \in \mathrm{Epi}(f)\} \geqslant \varphi(z) + \beta s \Leftrightarrow \beta \tilde{f}(\phi) = \beta f(\phi) + \phi(z) - (\beta s + \varphi(z)) \geqslant 0.$$

By the case for positive linear functionals on X^*: for any $\phi \in X^*$, there are $(z_i) \in X$ and $(\lambda_i) \in \mathbb{R}$ such that:

$$\tilde{f}(\phi) = \sup_{i\in\mathbb{N}} \phi(z_i) + \lambda_i = f(\phi) + \frac{\phi(z)}{\beta} - \frac{\beta s + \varphi(z)}{\beta} \Leftrightarrow f(\phi) = \sup_{i\in\mathbb{N}} \phi(\tilde{z}_i) + \tilde{\lambda}_i,$$

where $\tilde{z}_i = z_i - \beta^{-1}z$ and $\tilde{\lambda}_i = \lambda_i + \beta^{-1}(\varphi(z) + \beta s)$. Suppose now (z_i, λ_i) is the defining sequence for f. Then by linearity $f(0) = 0 = \sup_{i\in\mathbb{N}} \lambda_i$, which gives $\lambda_i \leqslant 0$. In particular, for any $\phi \in X^*$, $f(\phi) \leqslant \sup_{i\in\mathbb{N}} \phi(z_i)$. Since f is positive 1-homogeneous, we have for any $i \in \mathbb{N}$, $f(\phi) \geqslant \phi(z_i) + t\lambda_i$ for all $t > 0$. Sending $t \downarrow 0$ and taking supremum over the indexing set gives: $f(\phi) \geqslant \sup_{i\in\mathbb{N}} \phi(z_i)$. □

Definition 5.5 (Recession function)**.** *Let $f\colon \mathbb{R}^m \to \mathbb{R}\cup\{\infty\}$ be convex, lower-semicontinuous such that $f \not\equiv \infty$. We define the recession functional of f to be, for an arbitrary $p_0 \in \mathbb{R}^m$ with $f(p_0) < \infty$:*

$$f_\infty\colon \mathbb{R}^m \to \mathbb{R}\cup\{\infty\}, \quad f_\infty(p) := \lim_{t\uparrow\infty} \frac{f(p_0 + tp) - f(p_0)}{t}.$$

If $f \equiv \infty$, we set for completeness $f_\infty(p) = \infty$ for any $p \neq 0$ and $f_\infty(0) = 0$.

Remark 5.4. The values of the recession function is identified by its value on the unit sphere in the domain by positive 1-homogeneity. Indeed, we have for any $t > 0$ and $p \in \mathbb{R}^m$:

$$f_\infty(tp) = \lim_{h\uparrow\infty} \frac{f(p_0 + h(tp)) - f(p_0)}{h} = t \cdot \lim_{h\uparrow\infty} \frac{f(p_0 + (th)p) - f(p_0)}{th} = tf_\infty(p).$$

Thus we can expect f_∞ to tell us about the behaviour of the convex function f at infinity along each direction in $\mathbb{R}^m$. In particular, f_∞ is finite along any direction of at most linear growth of f and infinite along directions of growth faster than linear. We have the following function:

$$f_{p_0,p}\colon (0,\infty) \to \mathbb{R}\cup\{\infty\}, \quad t \longmapsto \frac{f(p_0 + tp) - f(p_0)}{t},$$

is monontone increasing on $(0,\infty)$ for any $p_0 \in \{f < \infty\}$ and $p \in \mathbb{S}^{m-1}$ by convexity. Indeed, for any $t > s > 0$, $p_0 + sp = \frac{s}{t}(p_0 + tp) + \frac{t-s}{t}p_0 \in [p_0, p_0 + tp]$:

$$f(p_0 + sp) \leqslant f(p_0) + \frac{s}{t}(f(p_0 + tp) - f(p_0)) \Leftrightarrow \frac{f(p_0 + sp) - f(p_0)}{s} \leqslant \frac{f(p_0 + tp) - f(p_0)}{t}.$$

In particular, the limit exists in $\mathbb{R}\cup\{\infty\}$ and f_∞ is convex and lower-semicontinuous by the same argument.

Lemma 5.2 (Well-posedness of recession)**.** *Let $f\colon \mathbb{R}^m \to \cup\{\infty\}$ be convex and lower-semicontinuous such that $f \not\equiv \infty$. Choose, by lower-semicontinuity, $(x_i) \in \mathbb{R}^m$ and $(\lambda_i) \in \mathbb{R}$ such that $f(p) = \sup_{i\in\mathbb{N}}\langle x_i, p\rangle_{Euc} + \lambda_i$. Then:*

$$f_\infty(p) = \sup \langle x_i, p\rangle_{Euc} \text{ for any } p \in \mathbb{R}^m.$$

Proof. Fix $p_0 \in \mathbb{R}^m$ such that $f(p_0) < \infty$. Note since $\langle x_i, p_0\rangle_{Euc} + \lambda_i - f(p_0) \leqslant 0$:

$$f_\infty(p) = \sup_{k\in\mathbb{N}} \frac{f(p_0 + kp) - f_(p_0)}{k} = \sup_{i,k\in\mathbb{N}} \frac{\langle x_i, p_0\rangle_{Euc} + \lambda_i - f(p_0)}{k} + \langle x_i, p\rangle_{Euc} = \sup_{i\in\mathbb{N}} \langle x_i, p\rangle_{Euc}.$$

□

Using the recession function, we can define functionals on pairs of Radon measures relative to some convex function. We first recall some concepts from abstract measure theory. Let X be a topological space and consider two signed Borel measures $\mu, \nu \in \mathcal{M}(X;\mathbb{R})$ on X. We say ν is singular with respect to μ, written $\nu \perp \mu$, if there exists $E, F \in \mathcal{B}(X)$ partitioning the whole space such that E is μ-negligible and F is ν-negligible: i.e.

$$E \cup F = E\Delta F = (E \setminus F) \cup (F \setminus E) = X, \quad \mu E = 0, \quad \nu F = 0.$$

Theorem 5.2 (Lebesgue-Radon-Nikodym decomposition)**.** *Let $\nu \in \mathcal{M}(X;\mathbb{R})$ and $\mu \in \mathcal{M}^+(X)$ be σ-finite Borel measures. Then there exists unique σ-finite signed Borel measures $\lambda, \eta \in \mathcal{M}(X;\mathbb{R})$ such that:*

$$\lambda \perp \mu, \quad \eta << \mu, \quad \text{and } \nu = \lambda + \eta,$$

which we refer to as the Lebesgue singular decomposition of ν with respect to μ. Furthermore, there exists an extended μ-integrable function $f\colon X \to \mathbb{R}$, called the Radon-Nikodym derivative $f = \frac{d\nu}{d\mu}$ of ν with respect to μ, such that $d\eta = fd\mu$ unique upto μ-a.e. equality. We denote the decomposition by $\nu = \nu^s + \frac{d\nu}{d\mu}$.

Proof. See [Fol13]. □

Lemma 5.3. *Let $\lambda \in \mathcal{M}^+(\Omega)$ be σ-finite measure and $(\psi_i)_{i\in\mathbb{N}}$ non-negative Borel functions on Ω. Then:*

$$\int_\Omega \sup_{i\in\mathbb{N}} \psi_i d\lambda = \sup_{\mathcal{I},(A_i|i\in\mathcal{I})} \sum_{i\in\mathcal{I}} \int_{A_i} \psi_i d\lambda,$$

where the supremum is taken over the collection of all finite indexing set $\mathcal{I} \subset \mathbb{N}$ and pairwise disjoint precompact open subsets $(A_i)_{i\in\mathcal{I}}$ of Ω.

Proof. Consider the iterated maximum function $\max_{i=1,\dots,k} \psi_i$. Assuming we can show for each $k \in \mathbb{N}$:

$$\int_\Omega \max_{i=1,\dots,k} \psi_i d\lambda = \sup\{\sum_{i=1}^k \int_{A_i} \psi_i d\lambda \mid (A_i)_{i=1}^k \text{ disjoint open precompact}\},$$

then since $\max_{i=1,\dots,k} \psi_i \nearrow \sup_{i\in\mathbb{N}} \psi_i$ as $k \to \infty$, we can apply the monotone convergence theorem to get:

$$\int_\Omega \sup_{i\in\mathbb{N}} \psi_i d\lambda = \int_\Omega \lim_{k\to\infty} \max_{i=1,\dots,k} \psi_i d\lambda = \lim_{k\to\infty} \sup_{(A_i)_{i=1}^k} \sum_{i=1}^k \int_{A_i} \psi_i d\lambda = \sup_{\mathcal{I},(A_i)_{i\in\mathcal{I}}} \int_{A_i} \psi_i d\lambda.$$

Fix $k \in \mathbb{N}$ and consider the sets $B_i := \{x \in \Omega \mid \psi_i(x) = \max_{j\leqslant k} \psi_j(x)\}$ and for convenience, take $B_0 = \emptyset$. Now set $C_i = B_i \setminus \bigcup_{j<i} B_j$, which gives us a sequence of pairwise disjoint sets. The sets $(C_i)_{i=1}^k \in \mathcal{B}(\Omega)$ partition the measure space $\Omega = \dot{\bigcup}_{i\leqslant k} C_i$ and satisfy the condition:

$$\int_\Omega \max_{i=1,\dots,k} \psi_i d\lambda = \sum_{i=1}^k \int_{C_i} \psi_i d\lambda.$$

By the interior regularity of the measure λ, we can use compact sets $K_i \subset C_i$ to approximate the collection (C_i). Then we can pick open neighbourhoods $A_i \supset K_i$ which remain pairwise disjoint and have measure sufficiently close to C_i. In particular, by the means of monotone convergence theorem again, the supremum over disjoint Borel families is equal to the supremum over families of pairwise disjoint compact sets. This supremum is then equal to the supremum over disjoint precompact families. □

Theorem 5.3 (l.s.c. criterion)**.** *Let $\Omega \subset \mathbb{R}^n$ be open, $\nu, (\nu_i) \in \mathcal{M}(\Omega; \mathbb{R}^m)$ and $\mu, (\mu_i) \in \mathcal{M}^+(\Omega)$ all Radon measures. Let $f\colon \mathbb{R}^m \to [0, \infty]$ be convex, we define the functional on pairs of Radon measures by:*

$$G(\mu, \nu) := \int_\Omega f \circ \frac{d\nu}{d\mu}(x) d\mu(x) + \int_\Omega f_\infty \circ \frac{d\nu^s}{d|\nu^s|}(x) d|\nu^s|(x),$$

where ν^s is the singular part of ν with respect to μ.
If f is also lower-semicontinuous and $\nu_i \overset{*}{\rightharpoonup} \nu$, $\mu_i \overset{*}{\rightharpoonup} \mu$ locally on Ω, then G weak- lower-semicontinuous:*

$$G(\nu, \mu) \leqslant \liminf_{i\to\infty} G(\nu_i, \mu_i).$$

Proof. Suppose without loss of generality there exists $z \in \mathbb{R}^m$ with $f(z) < \infty$. We can thus find $(\lambda_i) \in \mathbb{R}$ and $(x_i) \in \mathbb{R}^m$ such that for any points $p \in \mathbb{R}^m$:

$$f(p) = \sup\{L_i(p) \mid i \in \mathbb{N}\} = \sup_{i\in\mathbb{N}} \langle x_i, p\rangle_{Euc} + \lambda_i \quad \text{and} \quad f_\infty(p) = \sup_{i\in\mathbb{N}} \langle x_j, p\rangle_{Euc}.$$

Fix some arbitrary $k \in \mathbb{N}$ and pairwise disjoint open subsets $(A_j)_{j=1}^k$ in Ω which are also precompact.
Then for any $j = 1, \dots, k$, $\varphi_j \in C_c^1(A_j; [0, 1])$, we have:

$$\begin{aligned}\int_{A_j} \lambda_j \varphi_j d\mu_i + \langle x_i, \int_{A_j} \varphi_j d\nu_i\rangle_{Euc} &= \int_{A_j} \varphi_j L_j \circ \frac{d\nu_i}{d\mu_i} d\mu_i + \int_{A_j} \varphi_j \cdot \langle x_i, \frac{d\nu_i^s}{d|\nu_i^s|}\rangle_{Euc} d|\nu_i^s| \\ &\leqslant \int_{A_j} f \circ \frac{d\nu_i}{d\mu_i} d\mu_i + \int_{A_j} f_\infty \circ \frac{d\nu_i^s}{d|\nu_i^s|} d|\nu_i^s| = G(\mu_i, \nu_i).\end{aligned}$$

Summing over $j = 1, \dots, k$ and passing to the limit $i \to \infty$ using the assumed weak-* convergence gives:

$$\sum_{j=1}^{k} \int_{A_j} \lambda_j \varphi_j d\mu + \langle x_i, \int_{A_j} \varphi_j d\nu \rangle_{Euc} \leqslant \liminf_{i \to \infty} G(\mu_i, \nu_i).$$

Consider the Lebesgue decomposition of ν with respect to μ:

$$\nu = \frac{d\nu}{d\mu}\mu + \nu^s, \quad \Omega = M \dot{\cup} N \quad \text{with} \begin{cases} \mu N = 0; \\ \nu M = 0, \end{cases}$$

where ν^s is concentrated on the μ-negligible set N. We define the following functions accordingly:

$$\psi_j(x) := \begin{cases} L_j \circ \frac{d\nu}{d\mu}(x) & x \in \Omega \setminus N \\ \langle x_j, \frac{d\nu^s}{d|\nu^s|}(x) \rangle_{Euc} & x \in N \end{cases} \text{ and } \psi(x) := \begin{cases} f \circ \frac{d\nu}{d\mu}(x) & x \in \Omega \setminus N \\ f_\infty \circ \frac{d\nu^s}{d|\nu^s|}(x) & x \in N \end{cases},$$

and define further $\lambda = \mu + |\nu^s|$ which gives us via taking the supremum over $[0,1]$-valued functions:

$$\sum_{j=1}^{k} \int_{A_j} \psi_j \varphi_j d\lambda \leqslant \sup_{\varphi_j \in C_c^1(A_j;[0,1])} \sum_{j=1}^{k} \int_{A_j} \psi_j \varphi_j d\lambda = \sum_{j=1}^{k} \int_{A_j} \psi_j^+ d\lambda \leqslant \liminf_{i \to \infty} G(\mu_i, \nu_i).$$

Clearly, we have $\sup_{j \in \mathbb{N}} \psi_j = \psi$ and the recession function of a non-negative convex function remains non-negative, thus giving us $\sup_{j \in \mathbb{N}} \psi_j^+ = \psi \geqslant 0$. Since $k \in \mathbb{N}$ and $(A_j)_{j=1}^k$ were chosen arbitrarily, we can approximate $G(\mu, \nu) = \int_\Omega \psi d\lambda$ by integrating over finite disjoint families of precompact open subsets: i.e.

$$G(\mu, \nu) = \int_\Omega \psi d\lambda = \sup\{ \sum_{j \in \mathcal{J}} \int_{A_j} \psi_j d\lambda \mid |\mathcal{J}| < \infty; \, (A_j)_{j \in \mathcal{J}} \text{ compact } \} \leqslant \liminf_{i \to \infty} G(\mu_i, \nu_i),$$

which is the desired lower-semicontinuity with respect to the weak-* topology. □

Remark 5.5. Suppose f has growth faster than linear, which implies that $f_\infty(z) = \infty$ for all $z \neq 0$. Then:

$$G(\mu, \nu) < \infty \Rightarrow \frac{d\nu^s}{d|\nu^s|} \equiv 0 \text{ on } \Omega \overset{\nu^s << |\nu^s|}{\Rightarrow} \nu^s \equiv 0 \Rightarrow \nu << \mu \Rightarrow G(\mu, \nu) = \int_\Omega f \circ \frac{d\nu}{d\mu}(x) d\mu(x).$$

The theorem and remark above give rise to a method to show stability of absolute continuity of measures under weak-* convergence. Indeed, let $f \colon \mathbb{R}^m \to [0, \infty]$ be convex, lower-semicontinuous of growth faster than linear. Consider sequences of Radon measures $(\nu_i) \in \mathcal{M}(\Omega; \mathcal{R}^m)$ and $(\mu_i) \in \mathcal{M}^+(\Omega)$ such that $\mu_i \overset{*}{\rightharpoonup} \mu$ and $\nu_i \overset{*}{\rightharpoonup} \nu$. Now suppose we have boundedness of the integral and absolute continuity in the sequence, i.e.

$$\sup_{i \in \mathbb{N}} | \int_\Omega f \circ \frac{d\nu_i}{d\mu_i} d\mu_i | < \infty \quad \text{and} \quad \nu_i << \mu_i \text{ for all } i \in \mathbb{N}.$$

It is clear that $\nu << \mu$ by approximating indicator functions by compactly supported continuous functions. Now using the l.s.c. criterion and Remark 5.5, we can deduce:

$$\int_\Omega f \circ \frac{d\nu}{d\mu}(x) d\mu(x) \leqslant \liminf_{i \to \infty} \int_\Omega f \circ \frac{d\nu_i}{d\mu_i}(x) d\mu_i(x).$$

Recall for any Radon measure $\mu \in \mathcal{M}(\Omega; \mathbb{R}^m)$, we always have absolute continuity with respect to its total variation measure $\mu << |\mu|$. It is thus natural to study properties of the following analogous function under weak-* convergence, which we define below:

$$H \colon \mathcal{M}(\Omega; \mathbb{R}^m) \to \mathbb{R}, \quad H(\mu) := \int_\Omega f(x, \frac{d\mu}{d|\mu|}(x)) d|\mu|(x),$$

for some Borel function $f \colon \Omega \times \mathbb{R}^m \to [0, \infty]$. We state some basic properties below, due to Y.G. Reshetnyak.

Proposition 5.6 (Elementary properties)**.** *Let $f \colon \Omega \times \mathbb{R}^m \to [0, \infty]$ be a Borel function which is positively 1-homogeneous and convex in the second variable. Then H satisfies the following properties:*

(i) H *is convex and positively* 1*-homogeneous;*

(ii) $H(\mu) = \int_\Omega f(x, \frac{d\mu}{d\lambda}(x))d\lambda(x)$ *for any* $\lambda \in \mathcal{M}^+(\Omega)$ *such that* $|\mu| << \lambda$*;*

(iii) $H(\mu + \nu) = H(\mu) + H(\nu)$ *for any* $|\mu| \perp |\nu|$.

Theorem 5.4 (Reshetnyak lower-semicontinuity)**.** *Let* $\Omega \subset \mathbb{R}^n$ *be open and* $\mu, (\mu_i) \in \mathcal{M}(\Omega; \mathbb{R}^m)$ *finite Radon measures such that* $\mu_i \overset{*}{\rightharpoonup} \mu$*. Then for every lower-semicontinuous function* $f\colon \Omega \times \mathbb{R}^m \to [0, \infty]$ *which is positively* 1*-homogeneous and convex in the second variable:*

$$\int_\Omega f(x, \frac{d\mu}{d|\mu|}(x))d|\mu|(x) \leqslant \liminf_{i\to\infty} \int_\Omega f(x, \frac{d\mu_i}{d|\mu_i|}(x))d|\mu_i|(x).$$

Proof. Consider the polar decompositions $\mu_i = h_i|\mu_i|$ and $\mu = h|\mu|$ for Borel functions $h, h_i\colon \Omega \to \mathbb{S}^{m-1}$. Define new measures $\nu_i := |\mu_i| \otimes \delta_{h_i(x)} \in \mathcal{M}(\Omega \times \mathbb{S}^{m-1}; \mathbb{R}^m \times \{0, 1\})$ which can be represented as linear functionals on $C_0(\Omega \times \mathbb{S}^{m-1}; \mathbb{R})$, i.e.

$$\int_{\Omega\times\mathbb{S}^{m-1}} \varphi(x, y)d\nu_i(x, y) = \int_\Omega \varphi(x, h_i(x))d|\mu_i|(x) \text{ for all } \varphi \in C_0(\Omega \times \mathbb{S}^{m-1}; \mathbb{R}).$$

Recall weak-* convergence of measures $\mu_i \overset{*}{\rightharpoonup} \mu$ implies boundedness of the sequence $\sup_{i\in\mathbb{N}} |\mu_i|(\Omega) < \infty$. In particular, we have $\sup_{i\in\mathbb{N}} |\nu_i|(\Omega \times \mathbb{S}^{m-1}) = \sup_{i\in\mathbb{N}} |\mu_i|(\Omega) < \infty$. By the De La Vallée Poussin weak-* compactness criterion, we can assume, upto passing to a subsequence, $\nu_i \overset{*}{\rightharpoonup} \nu$, where ν is some finite Radon measure on $\Omega \times \mathbb{S}^{m-1}$. Denoting the projection onto Ω by $\pi\colon \Omega \times \mathbb{S}^{m-1} \to \Omega$, it is clear that $|\mu_i| = \pi_*\nu_i$ for each $i \in \mathbb{N}$. Since the projection map is continuous, weak-* continuity of the push-forward gives us $|\mu_i| \overset{*}{\rightharpoonup} \pi_*\nu$. In particular, we can deduce $\lambda := \pi_*\nu \geqslant |\mu|$. Now using the method of disintegration, we can find λ-measurable map $x \mapsto \nu_x$ with $\nu_x(\mathbb{S}^{m-1}) = 1$ and $\nu = \lambda \otimes \nu_x$. It then remains to show:

$$\int_{\mathbb{S}^{m-1}} y d\nu_x(y) = h(x)\frac{d|\mu|}{d\lambda}(x) \quad \lambda\text{-a.e. } x \in \Omega.$$

Indeed, then the result follows from lower weak-* semicontinuity and Jensen's inequality:

$$\begin{aligned}\liminf_{i\to\infty} \int_{\Omega\times\mathbb{S}^{m-1}} f(x, y)d\nu_i(x, y) &\geqslant \int_{\Omega\times\mathbb{S}^{m-1}} f(x, y)d\nu(x, y) \overset{\text{Fubini}}{=} \int_\Omega \int_{\mathbb{S}^{m-1}} f(x, y)d\nu_x(y)d\lambda(x)\\ &\overset{\text{Jensen}}{\geqslant} \int_\Omega f(x, h(x)\frac{d|\mu|}{d\lambda}(x))d\lambda(x) \overset{\text{Radon-Nikodym}}{=} \int_\Omega f(x, g(x))d|\mu|(x).\end{aligned}$$

The above claim essentially follows from the weak-* convergence using test function of form $\varphi(x, y) = \psi(x)y$:

$$\begin{aligned}\int_\Omega \left(\int_{\mathbb{S}^{m-1}} y d\nu_x(y)\right) \psi(x)d\lambda(x) &\overset{\text{Fubini}}{=} \int_{\Omega\times\mathbb{S}^{m-1}} \psi(x)y d\nu(x, y) = \lim_{i\to\infty} \int_{\Omega\times\mathbb{S}^{m-1}} \psi(x)y d\nu_i(x, y)\\ &= \lim_{i\to\infty} \int_\Omega \left(\int_{\mathbb{S}^{m-1}} y d\delta_{h_i(x)}\right) \psi(x)d\mu_i(x) = \lim_{i\to\infty} \int_\Omega \psi(x)h_i(x)d|\mu_i|(x)\\ &= \lim_{i\to\infty} \int_\Omega \psi(x)d\mu_i(x) = \int_\Omega \psi(x)d\mu(x) = \int_\Omega \psi(x)h(x)\frac{d|\mu|}{d\lambda}(x)d\lambda(x),\end{aligned}$$

where $\psi \in C_0(\Omega)$ arbitrary. Then the claim follows from the Riesz representation theorem for $C_0(\Omega)$. ☐

Theorem 5.5 (Reshetnyak continuity)**.** *Let* $\Omega \subset \mathbb{R}^n$ *be open and* $\mu, (\mu_i) \in \mathcal{M}(\Omega; \mathbb{R}^m)$ *finite Radon measures such that* $\mu_i \overset{*}{\rightharpoonup} \mu$*. Suppose further the convergence of the total variation* $|\mu_i|(\Omega) \to |\mu|(\Omega)$ *as* $i \to \infty$*. Then:*

$$\lim_{i\to\infty} \int_\Omega f(x, \frac{d\mu_i}{d|\mu_i|}(x))d|\mu_i|(x) = \int_\Omega f(x, \frac{d\mu}{d|\mu|}(x))d|\mu|(x) \textit{ for all } f \in C_0(\Omega \times \mathbb{S}^{m-1}; \mathbb{R}).$$

Proof. Choose again the polar decompositions $\mu_i = h_i|\mu_i|$ and $\mu = h|\mu|$. We can define as before a sequence of Radon measures $\nu_i := |\mu_i| \otimes \delta_{h_i(x)}$, which upto subsequences converges to another finite Radon measure $\nu \in \mathcal{M}(\Omega\times\mathbb{S}^{m-1}; \mathbb{R}^m\times\{0, 1\})$ with respect to the weak-* topology $\sigma(\mathcal{M}, C_0)$. Using the analogous argument, we can again deduce $|\mu_i| \xrightharpoonup{*} \lambda$ for some finite Radon measure $\lambda \in \mathcal{M}^+(\Omega)$. Note since $|\mu_i|(\Omega) \to |\mu|(\Omega)$, we necessarily have $\lambda = \lim_{i\to\infty} |\mu_i| = |\mu|$. Thus, from $\int_{\mathbb{S}^{m-1}} y d\nu_\cdot(y) = h\frac{d|\mu|}{d\lambda}$ λ-a.e. we can deduce:

$$0 = \frac{1}{2}\int_{\mathbb{S}^{m-1}} |y - h(x)| d\nu_x(y) = 1 - \langle h(x), \int_{\mathbb{S}^{m-1}} y d\nu_x(y)\rangle_{Euc} \Rightarrow \nu_x = \delta_{h(x)} \ |\mu|\text{-a.e.} x \in \Omega$$

Now $\nu_i(\Omega\times\mathbb{S}^{m-1}) = |\mu_i|(\Omega) \to |\mu|(\Omega) = \nu(\Omega\times\mathbb{S}^{m-1})$ as $i \to \infty$, we can thus conclude $\nu_i \xrightharpoonup{*} \nu$ analogously:

$$\lim_{i\to\infty}\int_\Omega f(x, h_i(x)) d|\mu_i|(x) = \lim_{i\to\infty}\int_{\Omega\times\mathbb{S}^{m-1}} f(x,y) d\nu_i(x,y) = \int_{\Omega\times\mathbb{S}^{m-1}} f d\nu = \int_\Omega f(x, h(x)) d|\mu|(x),$$

for all $f \in C_0(\Omega \times \mathbb{S}^{m-1})$ since $\nu = |\mu| \otimes \delta_h(x)$ by the arguments above. □

6 Daniell integrals

One of the main difficulties with the traditional formulation of the Lebesgue integration theory is the requirement of an adequate measure theory a priori. An alternative, but equivalent construction is developed initially by P.J. Daniell via axiomatising the integral. This approach alleviates practical difficulties for the application of integration theory in the field of functional analysis.

Definition 6.1 (Lattices)**.** *A lattice of functions on a set X is a non-empty collection L of functions satisfying:*

(i) $f + h, \lambda f, \inf\{f, h\}, \inf\{f, \lambda\} \in L$ *whenever* $f, h \in L$ *and* $\lambda \in [0, \infty)$;

(ii) $f - h \in L$ *whenever* $f, h \in L$ *such that* $h \leqslant f$.

In particular, any vector space of functions $X \to \mathbb{R}$ closed under taking infimum is a lattice.

Definition 6.2 (Axioms of Daniell integral)**.** *Let L be a lattice over X and set $L^+ := \{f \in L \mid f \geqslant 0\}$. We say an operator $T\colon L \to \mathbb{R}$ is:*

(i) linear, if for all $f, h \in L$ and $\lambda \in [0,\infty)$, $T(f + \lambda h) = Tf + \lambda Th$;

(ii) monotone, if for all $f, h \in L$ with $h \leqslant f$, we have $Th \leqslant Tf$;

(iii) continuous with respect to monotone convergence, if $Tf_i \to Tf$ whenever $f_i \to f$ monotonically with $(f_i) \in L$ (and consequently $f \in L$ by monotone convergence and closure under taking infimum);

(iv) bounded, if for all $f \in L$, $\sup\{Th \mid 0 \leqslant h \leqslant f\} < \infty$.

We define a Daniell integral to be a linear, bounded operator $T\colon L \to \mathbb{R}$ which is continuous with respect to monotone convergence.

Lemma 6.1 (Monotonicity of Daniell integral)**.** *Let $T\colon L \to \mathbb{R}$ be a linear, monotone operator which is continuous with respect to monotone convergence. Then T is necessarily a bounded operator and thus define a Daniell integral. We refer to such operators as a monotone Daniell integral.*

Proof. Suppose without loss of generality that $f \in L^+$ with $Tf < \infty$. Then by monotonicity, for any $h \in L^+$ such that $h \leqslant f$, we must have $Tf \geqslant Th$. Taking supremum over all such functions gives us the result. □

Lemma 6.2 (Monotone decomposition of Daniell integral)**.** *Let L be a lattice of functions on X and $T\colon L \to \mathbb{R}$ a Daniell integral. Define the following functionals on L:*

$$T^+\colon L^+ \longrightarrow \mathbb{R}, \quad f \longmapsto \sup_{h\in L^+; h\leqslant f} Th; T^-\colon L^+ \longrightarrow \mathbb{R}, \quad f \longmapsto -\inf_{h\in L^+; h\leqslant f} Th.$$

Then both T^+, T^- define monotone Daniell integrals on L^+ and $T \equiv T^+ - T^-$.

Proof. Note first that L^+ also defines a lattice over X. Let $f \in L^+$ and pick an arbitrary $h \in L^+$ such that $h \leqslant f$. Then any function in L^+ bounded above by h is also bounded above by f. This implies that we are

taking a supremum over a larger collection, which therefore gives $T^+h \leqslant T^+f$. Thus T^+ is monotone.
For arbitrary functions $f, h \in L^+$ and $\lambda \in [0, \infty)$, we have $f + \lambda h \in L^+$ by linearity and note:

$$T^+(f + \lambda h) = \sup_{\varsigma \in L^+; \varsigma \leqslant f + \lambda h} T\varsigma \geqslant \sup_{\varsigma \in L^+; \varsigma \leqslant f} T\varsigma + \sup_{\varsigma \in L^+; \varsigma \leqslant h} \lambda T\varsigma = T^+f + \lambda T^+h.$$

Now fix $\varphi, \psi \in L^+$ such that $\varphi \leqslant f$ and $\psi \leqslant h$ respectively. Then by linearity of T:

$$T^+f \geqslant T\varphi = T(\varphi + \lambda\psi) - \lambda T\psi \geqslant T(\varphi + \lambda\psi) - \lambda \sup_{\psi \in L^+; \psi \leqslant h} T\psi = T(\varphi + \lambda\psi) + \lambda T^+h.$$

Now taking the supremum over all functions in L^+ with an upper bound $f + \lambda h$ gives us the linearity of T^+. Continuity with respect to monotone convergence follows via a diagonal argument. Thus T^+ defines a monotone Daniell integral on L^+ and similarly for T^-. Now for $f, h \in L^+$ with $h \leqslant f$, we have $f \geqslant f - h \in L^+$ and:

$$Th - T^-f \leqslant Th + T(f - h) = Tf \leqslant Th + T^+f,$$

by monotonicity. We can easily deduce $T^+f - T^-f \leqslant Tf$ from above LHS inequality. On the other hand, we have for all $h \in L^+$ satisfying $h \leqslant f$:

$$Tf \leqslant Th + T^+f \Leftrightarrow T^+f \geqslant Tf - Th \geqslant Tf + (-\inf_{\psi \in L^+; \psi \leqslant f} T\psi) = Tf + T^-f,$$

which gives us the reverse inequality and hence the decomposition $T \equiv T^+ - T^-$ holds. $\square$

It is important to see that integration against an outer measure μ induces a positive linear functional on the space of μ-measurable functions. Continuity of μ-integrals follows from the monotone convergence theorem. Thus, μ-integrals indeed defines a monotone Daniell integral. Conversely, we also have the following:

Theorem 6.1 (Representation theorem for Daniell integral)**.** *Let L be a lattice of functions on X and $T\colon L \to \mathbb{R}$ a monotone Daniell integral. Then there exists $\mu \in \mathcal{M}^+(X)$ such that each $f \in L^+$ is μ-measurable and*

$$Tf = \int_X f d\mu \quad \textit{for all } f \in L.$$

Proof. Recall that any monotone Daniell integral is a positive linear functional. Define the set function on X:

$$\mu\colon \mathscr{P}(X) \longrightarrow [0, \infty], \quad E \longmapsto \mu E := \inf\left\{\lim_{i\to\infty} Tf_i \mid (f_i) \in L^+ \text{ increasing, } \lim_{i\to\infty} f_i(x) \geqslant 1 \text{ for all } x \in E\right\}.$$

Then note $\mu\emptyset = 0$ and μ is clearly monotone by construction. If $(E_i) \in \mathscr{P}(X)$ and $(f_i^j) \in L^+$ is an increasing sequence as described above for each $j \in \mathbb{N}$, then define $h_i := \sum_{j=1}^{i} f_i^j \in L^+$ for each $i \in \mathbb{N}$. Now (h_i) is monotone increasing in i and satisfies for all fixed $k \in \mathbb{N}$:

$$\lim_{i\to\infty} h_i(x) = \lim_{i\to\infty} \sum_{j=1}^{i} f_i^j(x) = \lim_{i\to\infty} \sum_{j\neq k} f_i^j(x) + \lim_{i\to\infty} f_i^k(x) \geqslant 1 + \lim_{i\to\infty} \sum_{j\neq k} f_i^j(x) \geqslant 1 \quad \text{for all } x \in E_k$$

Furthermore, we can see that by monotonicity and linearity of T:

$$Th_i = \sum_{j=1}^{i} Tf_i^j \leqslant \sum_{j=1}^{\infty} \lim_{i\to\infty} Tf_i^j,$$

where we deduce $\mu(\bigcup_{j=1}^{\infty} E_j) \leqslant \lim_{i\to\infty} Th_i \leqslant \sum_{j=1}^{\infty} \mu E_j$ by taking the infimum over the collection of all such $(f_i^j) \in L^+$ for all $j \in \mathbb{N}$. In particular, $\mu \in \mathcal{M}^+(X)$ indeed defines an outer measure on X.
Now let $f \in L^+$ and $t < s \in \mathbb{R}$ be arbitrarily chosen. We will show the following:

$$\mu E \geqslant \mu(E \cap A) + \mu(E \cap B), \quad \text{where } \begin{cases} A = f^{-1}((-\infty, t)); \\ B = f^{-1}((s, \infty)); \end{cases}$$

which would then give us the μ-measurability of f. Now choose $(h_i) \in L^+$ a monontone increasing sequence such that $\lim_{i\to\infty} h_i(x) \geqslant 1$ for all $x \in E$. Then define:

$$\varphi := \frac{\inf\{f, s\} - \inf\{f, t\}}{s - t} \quad \text{and} \quad \psi_i := \inf\{h_i, \varphi\},$$

which gives us $0 \leqslant \psi_{i+1} - \psi_i \leqslant h_{i+1} - h_i$ for all $i \in \mathbb{N}$. Note also $\varphi \equiv 1$ on $\{f \geqslant s\}$ and $\varphi \equiv 0$ on $\{f \leqslant t\}$. By construction, (ψ_i) and $(h_i - \psi_i)$ are functions in L^+ satisfying the following:

$$\lim_{i\to\infty} \psi_i(x) = \min\{\lim_{i\to\infty} h_i(x), 1\} \geqslant 1 \qquad \text{for all } x \in E \cap B;$$
$$\lim_{i\to\infty} (h_i - \psi_i)(x) \geqslant 1 - \min\{\lim_{i\to\infty} h_i(x), 0\} \geqslant 1 \qquad \text{for all } x \in E \cap A.$$

In particular, by the definition of our choice of measure μ, we deduce:

$$\lim_{i\to\infty} Th_i = \lim_{i\to\infty} T\psi_i + T(h_i - \psi_i) \geqslant \mu(E \cap A) + \mu(E \cap B),$$

whereas we can take the infimum over all such (h_i) on the RHS to conclude $\mu E \geqslant \mu(E \cap B) + \mu(E \cap A)$. Hence any $f \in L^+$ is necessarily μ-measurable. Now it remains to show μ is a representation measure for T. Let $E \subset X$ and $(f_i) \in L^+$ a monotone increasing sequence such that $\lim_{i\to\infty} f_i(x) \geqslant 1$ for all $x \in E$. Then for all $h \in L^+$ with $h \leqslant \chi_E$, we can define $h_i := \inf\{h, f_i\} \in L^+$ such that $h_i \nearrow h$. In particular, by monotonicity and continuity of T with respect to monotone convergence:

$$Th = \lim_{i\to\infty} Th_i \leqslant \lim_{i\to\infty} Tf_i,$$

which gives us $\mu E \geqslant Th$. Now fix any $f \in L^+$ and define $f_s := \inf\{f, s\}$. Let $\varepsilon > 0$ be arbitrary. Then notice for all $k \in \mathbb{N}$, we have $f_{k\varepsilon} - f_{(k-1)\varepsilon} \in [0, \varepsilon]$, where $f_{k\varepsilon}(x) - f_{(k-1)\varepsilon}(x) = \varepsilon$ whenever $f(x) \geqslant k\varepsilon$ and $f_{k\varepsilon}(x) - f_{(k-1)\varepsilon}(x) = 0$ whenever $f(x) \leqslant (k-1)\varepsilon$. Thus, $\varepsilon^{-1}(f_{k\varepsilon} - f_{(k-1)\varepsilon}) \geqslant \chi_{\{f \geqslant k\varepsilon\}}$ and we can set $f_i = \varepsilon^{-1}(f_{k\varepsilon} - f_{(k-1)\varepsilon})$ for all $i \in \mathbb{N}$ to obtain by definition of μ:

$$\varepsilon^{-1}T(f_{k\varepsilon} - f_{(k-1)\varepsilon}) = \varepsilon^{-1}(f_{k\varepsilon} - f_{(k-1)\varepsilon}) \geqslant \mu\{f \geqslant k\varepsilon\}.$$

Now by the above observations, we deduce $\operatorname{supp}(f_{(k+1)\varepsilon} - f_{k\varepsilon}) \subset \{f \geqslant k\varepsilon\}$ and thus:

$$T(f_{k\varepsilon} - f_{(k-1)\varepsilon}) \geqslant \varepsilon\mu\{f \geqslant k\varepsilon\} \geqslant \int_X f_{(k+1)\varepsilon} - f_{k\varepsilon} d\mu \geqslant \varepsilon\mu\{f \geqslant (k+1)\varepsilon\} \geqslant T(f_{(k+2)\varepsilon} - f_{(k+1)\varepsilon}).$$

The telescoping sum over k from 1 to i gives us by noting $f_0 \equiv 0$:

$$Tf_{i\varepsilon} = \sum_{k=1}^{i} T(f_{k\varepsilon} - f_{(k-1)\varepsilon}) \geqslant \int_X f_{(i+1)\varepsilon} - f_\varepsilon \geqslant \sum_{k=1}^{i} T(f_{(k+2)\varepsilon} - f_{(k+1)\varepsilon}) = T(f_{(i+2)\varepsilon} - f_{(i+1)\varepsilon}).$$

Notice by construction, we have $f_t \nearrow f$ as $t \to \infty$. In particular, by the monotone convergence theorem and the continuity of T with respect to monotone convergence, the above inequality yields:

$$Tf \geqslant \int_X f - f_\varepsilon d\mu \geqslant T(f - f_\varepsilon).$$

Now sending $\varepsilon \searrow 0$ and using the monotone convergence properties again gives us:

$$Tf = \int_X f d\mu \qquad \text{for all } f \in L^+.$$

For a general $f \in L$, split f into positive and negative parts, then using linearity of (Daniell) integral, we can deduce that μ is the (outer) measure representation of T, i.e. $Tf = \langle f, \mu \rangle$. $\square$

Let X be a locally compact Hausdorff topological space. The function space $C_c^0(X)$ of compactly supported continuous functions is a vector space closed under taking infimum since functions in this space have compact supports, which then implies that $C_c^0(X)$ defines a lattice over X. Now any monotone Daniell integral clearly induces a positive linear functional on $C_C^0(X)$. In particular, we can view the above representation theorem as a generalisation of the Riesz-Markov-Kakutani theorem for Radon measures.

Theorem 6.2 (Weak-* compactness of Radon measures)**.** *Let X be a locally compact Hausdorff topological space and $(\mu_i) \in \mathcal{M}(X)$ with a uniform bound on the corresponding sequence of total variation measures $(|\mu_i|)$, i.e. $\sup_{i\in\mathbb{N}} |\mu_i|(X) < \infty$. Then (μ_i) admits a weak-* convergenet subsequence in $\mathcal{M}(X)$.*

Proof. This is a consequence of the Banach-Alaoglu theorem by recalling $\|\mu\|_{\mathcal{M}} := |\mu|(X)$ defines a norm on the space of vector-valued measures $\mathcal{M}(X;\mathbb{R})$. The isometric embedding $(\mathcal{M}(X;\mathbb{R}), \|\cdot\|_{\mathcal{M}}) \hookrightarrow ((C_c^0(X))^*, \|\cdot\|_{op})$ implies that (μ_i) induces a sequence of bounded linear functionals on $C_C^0(X)$, which admits a uniform bound on the opertor norm. Now the result follows by weak-* compactness of $\overline{\mathbb{B}}_{(C_c^0(X))^*}$. $\square$

It is common, although somewhat unnecessary for practical purposes, to deduce a measure-theoretic concepts based Daniell integrals in terms of a class of the so-called elementary functions. Indeed, if we choose the collection of all simple functions on X to be elementary functions, then we can agree to call $T\chi_E$ the measure of $E \subset X$ with respect to the monotone Daniell integral $T\colon L \to \mathbb{R}$.

Example 6.1. *Let $X = [a,b]$ and L the lattice of simple functions defined on X. Then the Daniell integral agrees with integration against the 1-dimensional Lebesgue measure on $\mathcal{L}^1$-summable functions over $[a,b]$.*

Definition 6.3 (Sets of measure zero)**.** *Let X be a set and $T\colon L \to \mathbb{R}$ a monotone Daniell integral on some lattice L over X. A subset $N \subset X$ is said to have measure zero in the Daniell integral sense with respect to T if for all $\varepsilon > 0$, there exists a monotone increasing sequence $(\psi_i) \in L^+$ such that:*

$$\sup_{i\in\mathbb{N}} T\psi_i \leqslant \varepsilon \quad \textit{and} \quad \lim_{i\to\infty} \psi_i(x) \geqslant \textit{ for all } x \in N.$$

A property is said to hold a.e. on X with respect to T if it holds everywhere away from a set N of measure zero.

Example 6.2 (Generalisation of Stieltjes integral)**.** *Recall that the usual Stieltjes integral is defined for integrating continuous functions on intervals against a function of bounded variation. Suppose now $\gamma\colon [a,b] \to \mathbb{R}$ is not of bounded variation but satisfies:*

$$\sum_{i=1}^{N} t_i|\gamma(t_i) - \gamma(t_{i-1})| \leqslant M < \infty \quad \textit{for all } (t_i \mid i = 0,1,\ldots,N) \in \mathscr{P}_{[a,b]},$$

where $\mathscr{P}_{[a,b]}$ denotes the collection of all partitions of $[a,b]$. The above implies $\alpha(t) := t\gamma(t) \in BV([a,b])$ and thus the Stieltjes integral with respect to γ can be defined in terms of the singular integral with respect to α:

$$\int_{[a,b]} f d\gamma := \int_a^b \frac{f(t)}{t} d\alpha(t) = \int_a^b \frac{f(t)}{t} t d\gamma(t).$$

7 Covering and differentiation theorems

The idea of the most fundamental covering theorems is to select, from a family of sets, a sub-collection with controlled overlap or some disjointness property in the measure-theoretic sense. The covering theorems yields significant consequences in terms of the geometry of some certain classes of Radon measures. Covering results with a uniform bound on the size of elements in the sub-family are often derived in the more convenient setting of metric spaces. A prototypical result is as follows:

Lemma 7.1. *Let (X,d) be a locally compact separable metric space. Then for any bounded subset $E \subset X$ and $\rho > 0$, there exists a finite string of elements $(x_i \mid i = 1,\ldots,N) \in E$ such that:*

$$E \subset \bigcup_{i=1}^{N} \mathbb{B}(x_i,\rho).$$

If $X = \mathbb{R}^n$, then each $x \in E$ is contained in at most 3^n of the balls $(\mathbb{B}(x_i,\rho) \mid i = 1,\ldots,N)$.

Proof. Let $x_1 \in E$ be arbitrary and denote by $E_1 := \mathbb{B}(x_1,\rho)$. Choose further $x_i \in E \setminus (\bigcup_{k=1}^{i-1} E_k)$ until we obtain a covering of E by open ρ-balls. Note this can be achieved in finitely many steps since $\mathrm{d}(x_k,x_l) \geqslant \rho$ for any $k \neq l$ and $E \subset \mathbb{B}(y,R)$ for some $y \in X$ and $R \in (0,\infty)$ sufficiently large.
Now let $X = \mathbb{R}^n$ and fix an arbitrary $x \in E$. Denote by $\mathcal{I}(x)$ the set of indices such that $x \in \bigcap_{i\in\mathcal{I}(x)} \mathbb{B}(x_i,\rho)$. Then by construction, we have $\mathbb{B}(x_k,\frac{\rho}{2}) \subset \mathbb{B}(x,\frac{3\rho}{2})$ for all $k \in \mathcal{I}(x)$ (since $\mathrm{d}_{Euc}(x_k,x) \leqslant \rho$) and the collection of balls $(\mathbb{B}(x_i,\frac{\rho}{2}) \mid i = 1,\ldots,N)$ are pairwise disjoint. In particular, we have the following volume inequality:

$$|\mathcal{I}(x)|2^{-n}\omega_n\rho^n = \sum_{i=1}^{N} \mathcal{L}^n(\mathbb{B}(x_i,\frac{\rho}{2})) = \mathcal{L}^n\left(\bigcup_{i\in\mathcal{I}(x)} \mathbb{B}(x_i,\frac{\rho}{2})\right) \leqslant \mathcal{L}^n\left(\mathbb{B}(x,\frac{3\rho}{2})\right) = 3^n 2^{-n}\omega_n\rho^n,$$

where ω_n denotes the volume of the n-dimensional unit ball. Thus $|\mathcal{I}(x)| \leqslant 3^n$. $\square$

The control on the intersections simply required an explicit form of the measure of covering ρ-balls, which can be mimicked on other metric spaces with a generic measure. We describe some properties of coverings below:

Definition 7.1 (Properties of coverings)**.** *Let $\mathscr{F}$ be a collection of subsets in X. We say that $\mathscr{F}$ is a disjoint family of subsets if for all $E, F \in \mathscr{F}$, we have $E \cap F = \emptyset$ unless $E = F$.*
Let $E \subset X$ be a subset of a metric space and $\mathscr{B}$ a covering of E by closed balls in X. We say $\mathscr{B}$ is a fine cover of E if for all $\varepsilon > 0$ arbitrarily small and $x \in E$, there exists $B \in \mathscr{B}$ such that $\operatorname{diam}(B) \leqslant 2\varepsilon$ and $x \in B$:

$$i.e. \quad \inf\{\operatorname{diam}(B) \mid B \in \mathscr{B},\ x \in B\} = 0 \quad \textit{for all } x \in E.$$

A fine cover is sometimes referred to as a Vitali cover. Fix a metric measure space (X, d, μ) and $E \subset X$, we say a collection $\mathscr{F}$ of subsets in X is a μ-a.e. covering for E if $E \setminus (\bigcup \mathscr{F})$ is μ-negligible.

Theorem 7.1 (Fundamental covering theorem)**.** *Let $\mathscr{B}$ be a family of closed balls in a general metric space (X, d). Suppose $\sup_{B\in\mathscr{B}} \operatorname{diam}(B) < \infty$. Then $\mathscr{B}$ admits a disjoint sub-family $\mathscr{G}$ such that:*

$$\bigcup \mathscr{B} \subset \bigcup_{B\in\mathscr{G}} 5B.$$

Furthermore, every ball $B \in \mathscr{B}$ intersects with a ball in $\mathscr{G}$ with radius at least $2^{-1}\operatorname{rad}(B)$.

Proof. Consider the following collections of disjoint sub-families $\varpi \subset \mathscr{B}$:

$$\mathscr{F} := \{\varpi \subset \mathscr{B} \text{ disjoint } \mid \forall B \in \mathscr{B} \text{ with } B \cap B_0 \neq \emptyset,\ B_0 \in \varpi\ \exists B_1 \in \varpi \text{ s.t. } B_1 \cap B \neq \emptyset,\ \operatorname{rad}(B_1) \geqslant \frac{1}{2}\operatorname{rad}(B)\},$$

equipped with the partial order by inclusion. Note for any $B \in \mathscr{B}$ with $\operatorname{rad}(B)$ sufficiently close to the supremal radius over $\mathscr{B}$, we have $\{B\} \in \mathscr{F}$ and thus $\mathscr{F}$ is nonempty. Now for any chain $\mathcal{C} \subset \mathscr{F}$, it is clear to see that $\varpi_{\mathcal{C}} := \bigcup \mathcal{C} = \bigcup_{\varpi\in\mathcal{C}} \varpi \in \mathscr{F}$ defines an upper bound for $\mathcal{C}$. Now Zorn's lemma implies the existence of an maximal element $\mathscr{G}$. We claim that $\mathscr{G}$ is the desired disjoint sub-family.
Suppose now there exists $B \in \mathscr{B}$ such that $B \cap E = \emptyset$ for all $E \in \mathscr{G}$. Then we can pick $B_0 \in \mathscr{B}$ with radius greater than half the radius of any other ball disjoint from elements in $\mathscr{G}$. Thus for any $B_1 \in \mathscr{B}$ intersecting some ball in the collection $\mathscr{G}_0 := \mathscr{G} \cup \{B_0\}$, B_1 necessarily intersect with some ball in $\mathscr{G}_0$ with radius at least half of $\operatorname{rad}(B)$. In particular, $\mathscr{G}_0 \in \mathscr{F}$ with $\mathscr{G}_0 \succ \mathscr{G}$, which then contradicts the maximality of $\mathscr{G}$. Hence every $B \in \mathscr{B}$ intersects with some $B' \in \mathscr{G}$ such that $\operatorname{rad}(B) \leqslant 2\operatorname{rad}(B')$. Writing $B = \mathbb{B}(x, r)$ and $B' = \mathbb{B}(x', r')$, then we necessarily have for all $y \in B$:

$$\mathrm{d}(x', y) \leqslant \mathrm{d}(x', z) + \mathrm{d}(z, x) + \mathrm{d}(x, y) \leqslant r' + 2r \leqslant 5r',$$

where $z \in B \cap B'$ is an arbitrary intersection point. Thus $B \subset 5B'$ and the result follows. □

Remark 7.1. The name "doubling measure" is somewhat misleading. Of course, it is simply a measure which works nicely with scaling of balls in X. In particular, we have for any $\lambda \geqslant 2$ and balls $B \subset X$, we can apply the doubling property to obtain $\mu(\lambda B) = \mu(2^{\frac{\log\lambda}{\log 2}} B) \leqslant C(\mu)^{\frac{\log\lambda}{\log 2}} \mu B \leqslant C(\lambda, \mu)\mu B$.

Proposition 7.1 (Fine cover lemma)**.** *Let (X, d) be a metric space and $E \subset X$ with a fine cover $\mathscr{B}$. Then there exists a disjoint sub-family $\mathscr{B}' \subset \mathscr{B}$ such that:*

$$E \setminus \left(\bigcup_{B\in\mathcal{I}} B\right) \subset \bigcup_{B\in\mathscr{B}'\setminus\mathcal{I}} 5B = \bigcup(\mathscr{B}' \setminus \mathcal{I}).$$

Proof. Suppose without loss of generality that $\sup_{B\in\mathscr{B}} \operatorname{diam}(B) < \infty$ since $\mathscr{B}$ is a fine cover. Then denote by $\mathscr{G} \subset \mathscr{B}$ the disjoint sub-family obtained by applying Theorem 7.1. Now for any finite sub-family $\mathcal{I} \subset \mathscr{G}$, $B_{\mathcal{I}} := \bigcup \mathcal{I}$ is a closed subset. Thus by property of fine covers, for any $x \in E \setminus B_{\mathcal{I}}$, there exists $B \in \mathscr{B}$ containing x such that $\operatorname{diam}(B) \leqslant 2\varepsilon$ where we pick $\varepsilon = 2^{-1}\operatorname{dist}(x, B_{\mathcal{I}}) > 0$. In particular, $B \cap B_{\mathcal{I}} = \emptyset$. By properties of the sub-family $\mathscr{G} \subset \mathscr{B}$, B must intersect some $B_0 \in \mathscr{G}$ with $2\operatorname{rad}(B_0) \geqslant \operatorname{rad}(B)$. Using triangle inequality as in the proof for Theorem 7.1, we deduce $x \in 5B_0 \subset \bigcup(\mathscr{G} \setminus \mathcal{I})$ as required. □

Theorem 7.2 (Vitali covering theorem)**.** *Let $\mu \in \mathcal{M}(X)$ be a doubling measure on the metric space (X, d), i.e. a Borel measure μ admitting some uniform constant $C(\mu) \in (0, \infty)$ such that:*

$$0 < \mu(2B) \leqslant C(\mu)\mu B < \infty \quad \textit{for all } B = \mathbb{B}(x, r) \subset X.$$

Suppose $\mathscr{B}$ is a fine cover for some subset $E \subset X$. Then there exists a countable disjoint sub-family $\mathscr{B}' \subset \mathscr{B}$ such that $E \setminus (\bigcup \mathscr{B}')$ is μ-negligible, i.e. $\mathscr{B}'$ defines a μ-a.e. covering by closed balls in X.

Proof. Note that the definition of fine covers is a local property (i.e. every point in E can be covered by a ball in $\mathscr{B}$ of sufficiently small radius). In particular, for any bounded subset of $E' \subset E$, $\mathscr{B}$ clearly remains a fine cover for E'. Taking a compact exhaustion and recalling that any countable union of μ-negligible sets remains μ-negligible, we notice it suffices to assume E is bounded. Then there exists some ball $B_0 \subset X$ large such that $E \subset B_0$ and without loss of generality $B \subset 2B_0$ for all $B \in \mathscr{B}$.
Let $\mathscr{B}_0 \subset \mathscr{B}$ be the sub-family as described in Proposition 7.1. Now for all $N \in \mathbb{N}$, denote by $\mathscr{B}_N \subset \mathscr{B}$ the collection of balls such that $\mu B \geqslant N^{-1}$. Notice $\mathscr{B}_0 = \bigcup_{N=1}^{\infty} \mathscr{B}_N$. Then by assumption and disjointness:

$$\frac{1}{N}|\mathscr{B}_N| \leqslant \sum_{B \in \mathscr{B}_N} \mu B = \mu\left(\bigcup_{B \in \mathscr{B}_N} B\right) \leqslant \mu(2B_0) < \infty.$$

Thus the cardinality of each $\mathscr{B}_N$ is at most $N\mu(2B_0)$ and $\mathscr{B}_0$ is necessarily countable, i.e. we can enumerate $\mathscr{B}_0 = \{B_i \mid i \in \mathbb{N}\}$. By a similar argument using boundedness and disjointness as above, we have:

$$\sum_{i=1}^{\infty} \mu B_i = \mu\left(\bigcup_{i=1}^{\infty} B_i\right) \leqslant \mu(2B_0) \leqslant C(\mu)\mu B_0 < \infty,$$

by the property of doubling measures. Therefore $\sum_{i>N} \mu B_i \to 0$ as $N \to \infty$. Now by the conclusion of the fine covering lemma and the fact that μ is a doubling measure on X, we can easily deduce:

$$\mu\left(E \setminus \left(\bigcup_{i=1}^{N} B_i\right)\right) \leqslant \mu\left(\bigcup_{i \geqslant N+1} 5B_i\right) \leqslant \sum_{i \geqslant N+1} \mu(5B_i) \leqslant C(\mu) \sum_{i \geqslant N+1} \mu B_i \longrightarrow 0,$$

as $N \to \infty$. Thus $\mathscr{B}_0$ is the required countable disjoint μ-a.e. covering for E. $\square$

Vitali's theorem allows us to give a proof to the celebrated Lebesgue's differentiation theorem on a doubling metric measure space using the approach of maximal functions. The idea goes back to Hardy and Littlewood:

Definition 7.2 (Maximal functions)**.** *Let (X, d, μ) be a metric measure space. Suppose further $\mu\mathbb{B}(x,\rho) \in (0,\infty)$ for all $x \in X$ and $\rho \in (0,\infty)$. We define the Hardy-Littlewood maximal operator on the space of μ-measurable functions on X:*

$$\mathscr{M} : f \longmapsto \mathscr{M}f(x) := \sup_{\rho>0} \fint_{\mathbb{B}(x,\rho)} |f| d\mu = \sup_{\rho>0} \frac{1}{\mu\mathbb{B}(x,\rho)} \int_{\mathbb{B}(x,\rho)} |f| d\mu.$$

Proposition 7.2 (Borel-measurability of maximal function)**.** *Let (X, d, μ) be a metric measure space such that $\mu\mathbb{B}(x,\rho) \in (0,\infty)$. Then for any $f \in L^1_{loc}(\mu)$, $\mathscr{M}f$ is Borel-measurable.*

Proof. Let $x \in X$ be arbitrary and choose $(x_i) \in X$ such that $\mathrm{d}(x_i, x) \searrow 0$. Fix some $\rho > 0$, we claim that $\mathbb{B}(x,\rho) \setminus \mathbb{B}(x_i,\rho)$ decreases to the empty set as $i \to \infty$. Indeed, for any $y \in \mathbb{B}(x,\rho) \setminus \{x\}$, denote by $\delta(y) := \mathrm{d}(x,y) \in (0,\rho)$. Now for $i \in \mathbb{N}$ sufficiently large, we have:

$$\mathrm{d}(x_i, y) \leqslant \mathrm{d}(x_i, x) + \mathrm{d}(x,y) \leqslant \varepsilon + \delta(y),$$

for some $\varepsilon > 0$ arbitrarily small. Therefore $y \in \mathbb{B}(x_i,\rho)$ eventually as $i \to \infty$. In particular, by the means of the monotone convergence theorem, we have:

$$\mu\mathbb{B}(x,\rho) \leqslant \liminf_{i\to\infty} \int \chi_{\mathbb{B}(x_i,\rho)} d\mu + \int \chi_{\mathbb{B}(x,\rho)\setminus\mathbb{B}(x_i,\rho)} d\mu = \liminf_{i\to\infty} \mu\mathbb{B}(x_i,\rho),$$

which gives us the lower-semicontinuity and thus the Borel-measurability of the map $y \mapsto \mu\mathbb{B}(y,\rho)$ for all fixed $\rho > 0$. On the other hand, fix $\rho > 0$ and $f \in L^1_{loc}(\mu)$. Since $\chi_{\mathbb{B}(x_i,\rho)} \to \chi_{\mathbb{B}(x,\rho)}$ in $L^1(\mu)$ and $f\chi_{\mathbb{B}(x,2\rho)} \in L^1(\mu)$, we can apply Lebesgue's dominated convergence theorem such that:

$$\int_{\mathbb{B}(x_i,\rho)} f d\mu = \int f\chi_{\mathbb{B}(x_i,\rho)} d\mu \longrightarrow \int f\chi_{\mathbb{B}(x,\rho)} d\mu = \int_{\mathbb{B}(x,\rho)} f d\mu \quad \text{as } i \to \infty.$$

Now recall that the space of Borel-measurable functions forms an algebra which is closed under taking supremum over a countable collection, we have that:

$$\mathscr{M}f(x) := \sup_{\rho>0} \frac{1}{\mu\mathbb{B}(x,\rho)} \int_{\mathbb{B}(x,\rho)} f d\mu,$$

defines a Borel-measurable function. $\square$

Corollary 7.1. *Let $\mu \in \mathcal{M}(X)$ be a doubling measure on a metric space (X, d). Then for any $f \in L^1_{loc}(\mu)$,*

$$\mathscr{M}_\rho \colon x \longmapsto \frac{1}{\mu\mathbb{B}(x,\rho)} \int_{\mathbb{B}(x,\rho)} f d\mu,$$

is Borel-measurable for any $\rho > 0$ fixed and thus $\mathscr{M} f$ is also Borel-measurable.

Theorem 7.3 (Hardy-Littlewood maximal inequality)**.** *Let (X, d, μ) be a metric measure space equipped with a doubling measure $\mu \in \mathcal{M}(X)$. Then there exists a uniform constant $C \in (0,\infty)$ such that for all $f \colon X \to \mathbb{R}$ μ-measurable and $\lambda > 0$, we have:*

$$\mu\{\mathscr{M} f \geqslant \lambda\} \leqslant \frac{C}{\lambda} \int_X |f| d\mu.$$

Proof. For any fixed, arbitrary $\lambda > 0$, denote by $E_\lambda := \{\mathscr{M} f > \lambda\}$ and seting for all $\rho > 0$:

$$E_{\lambda,\rho} := \left\{ x \in X \mid \exists r \in (0,\rho) \text{ with} \int_{\mathbb{B}(x,r)} |f| d\mu \geqslant \lambda \mu\mathbb{B}(x,r) \right\} = \bigcup_{r \in (0,\rho)} \{\mathscr{M}_r f \geqslant \lambda\}.$$

We can see that $E_{\lambda,\rho} \in \mathcal{B}(X)$ for all $\rho > 0$ such that $E_{\lambda,\rho}$ increases monotonically to E_λ, i.e. $\chi_{E_{\lambda,\rho}} \nearrow \chi_{E_\lambda}$ pointwise μ-a.e. as $\rho \to \infty$. Denote by $\mathscr{B}_\rho$ the collection of balls $\mathbb{B}(x,r)$ with $x \in E_{\lambda,\rho}$ and $r \in (0,\rho)$ satisfying the condition above. Noting that there is a uniform diameter bound on elements in $\mathscr{B}_\rho$ for any fixed $\rho > 0$, we can thus use Theorem 7.1 to find a disjoint sub-family $\mathscr{B}'_\rho$ such that:

$$\int_X |f| d\mu \geqslant \sum_{B \in \mathscr{B}'_\rho} \int_B |f| d\mu \geqslant \lambda \sum_{B \in \mathscr{B}'_\rho} \mu B \geqslant \frac{\lambda}{C(\mu)} \sum_{B \in \mathscr{B}'_\rho} \mu(5B) \geqslant \frac{\lambda}{C(\mu)} \mu E_{\lambda,\rho},$$

where the inequalities follow by disjointness and the doubling property. Now sending $\rho \to \infty$ gives us the desired statement (with $C = C(\mu)^{\frac{\log 5}{\log 2}} \in (0,\infty)$) by monotone convergence. □

We are now ready to prove the Lebesgue's differentiation theorem.

Theorem 7.4 (Lebesgue's differentiation theorem)**.** *Let (X, d, μ) be a metric measure space where $\mu \in \mathcal{M}(X)$ a doubling measure on X. Then for all $f \colon X \to \mathbb{R}$ locally μ-integrable we have:*

$$⨍_{\mathbb{B}(x,\rho)} |f(y) - f(x)| d\mu(y) = \frac{1}{\mu\mathbb{B}(x,\rho)} \int_{\mathbb{B}(x,\rho)} |f - f(x)| d\mu \longrightarrow 0 \quad \textit{for } \mu\textit{-a.e. } x \in X.$$

Proof. We can first show this for continuous functions. If $f \in C^0(X)$, then we have for all $x \in X$:

$$⨍_{\mathbb{B}(x,\rho)} |f(y) - f(x)| d\mu(y) \leqslant \sup_{y \in \mathbb{B}(x,\rho)} |f(y) - f(x)| \longrightarrow 0 \quad \text{as } \rho \searrow 0,$$

by exploiting continuity of f. Now recall that $C^0(X) \cap L^p(\mu) \subset L^p(\mu)$ is a dense subspace. Consider an arbitrary $f \in L^1(\mu)$ and fix some $\varepsilon > 0$, there exists some $f_\varepsilon \in C^0(X) \cap L^1(X)$ by density such that $\|f_\varepsilon - f\|_{L^1(\mu)} \leqslant \varepsilon^2$. Now by the Hardy-Littlewood maximal inequality, we have:

$$\mu\{\mathscr{M}(|f_\varepsilon - f|) \geqslant \varepsilon\} \leqslant C\varepsilon^{-1} \|f_\varepsilon - f\|_{L^1(\mu)} \leqslant C\varepsilon \quad \text{for all } \varepsilon > 0.$$

In particular, writing $E_\varepsilon := \{\mathscr{M}(f_\varepsilon - f) \geqslant \varepsilon\}$, for any $x \in X \setminus E_\varepsilon$, we have by definition:

$$\limsup_{\rho \downarrow 0} \mathscr{M}_\rho(|f_\varepsilon - f|)(x) = \limsup_{\rho \downarrow 0} ⨍_{\mathbb{B}(x,\rho)} |f_\varepsilon - f| d\mu \leqslant \mathscr{M}(|f_\varepsilon - f|)(x) \leqslant \varepsilon.$$

Now recall L^1-convergence implies μ-a.e. convergence, thus at each $x \notin E_\varepsilon$ where $|f_\varepsilon(x) \to f(x)| \leqslant \varepsilon$:

$$\begin{aligned} \limsup_{\rho \downarrow 0} ⨍_{\mathbb{B}(x,\rho)} |f - f(x)| d\mu \leqslant & \limsup_{\rho \downarrow 0} ⨍_{\mathbb{B}(x,\rho)} |f - f_\varepsilon| + |f_\varepsilon(x) - f(x)| + |f_\varepsilon - f_\varepsilon(x)| d\mu \\ \leqslant & \mathscr{M}(|f_\varepsilon - f|)(x) + \varepsilon + \lim_{\rho \downarrow 0} \sup_{y \in \mathbb{B}(x,\rho)} |f_\varepsilon(y) - f_\varepsilon(x)| \leqslant 2\varepsilon \end{aligned}$$

Therefore, we obtain a subset $F_\varepsilon \subset E_\varepsilon$ satisfying $\mu F_\varepsilon \leqslant \mu E_\varepsilon \leqslant C\varepsilon$ and:

$$\limsup_{\rho\downarrow 0} \frac{1}{\mu\mathbb{B}(x,\rho)} \int_{\mathbb{B}(x,\rho)} |f - f(x)| d\mu \leqslant \varepsilon \quad \text{for all } x \in X \setminus F_\varepsilon.$$

Thus $(F_\varepsilon \mid \varepsilon > 0)$ monotonically decreases to a set N of measure zero. Then by construction:

$$\frac{1}{\mu\mathbb{B}(x,\rho)} \int_{\mathbb{B}(x,\rho)} |f - f(x)| d\mu \longrightarrow 0 \quad \text{as } \rho \downarrow 0,$$

for all $x \in X \setminus N$ as required. For $f \in L^1_{loc}(\mu)$, choose a compact exhaustion (K_i) of X and notice that $f\chi_{K_i} \in L^1(\mu)$ for all $i \in \mathbb{N}$. Then the convergence holds μ-a.e. on each K_i, whence the result follows by sending $i \to \infty$ since K_i increases monotonically to X. □

Remark 7.2. An alternative approach to derive the Lebesgue's differentiation theorem on Euclidean spaces, as we will see later, will be to discuss splitting properties of vector-valued measures on $\mathbb{R}^n$.

Corollary 7.2 (Lebesgue's density theorem)**.** *Let (X, d, μ) be a doubling metric measure space. Then for any $E \in \mathcal{B}_\mu(X)$ such that $\mu E < \infty$:*

$$\frac{\mu(E \cap \mathbb{B}(x,\rho))}{\mu\mathbb{B}(x,\rho)} \overset{\rho\downarrow 0}{\longrightarrow} \delta_{x\in E} \quad \textit{for } \mu\textit{-a.e. } x \in X.$$

Proof. This follows by applying Lebesgue's differentiation theorem to the indicator function $\chi_E \in L^1(\mu)$. □

Another important covering theorem is due to Besicovitch, which is a fundamental tool to study Radon measures on Euclidean spaces. Geometrically, as opposed to Vitali's covering theorem, Besicovitch's theorem gives us a subcover instead of covering by enlarged balls. The trade-off here is the requirement of a control on the overlaps.

Lemma 7.2. *Let $\overline{\mathbb{B}}_i := \overline{\mathbb{B}}(x_i, r_i) \subset \mathbb{R}^n$ for $i = 1, \ldots, m$. Suppose for all $i \neq j$, we have $x_j \notin \overline{\mathbb{B}}_i$ and $\bigcap_{i=1}^m \overline{\mathbb{B}}_i \neq \emptyset$. Then there exists a dimensional constant $k(n) \in \mathbb{N}$ such that $m \leqslant k(n)$.*

Proof. Suppose without loss of generality (by translation) $0 \in \bigcap_{i=1}^m \overline{\mathbb{B}}_i$ and $x_i \neq 0$ for all i. Then for all $i \neq j$, we have $|x_i| \leqslant r_i \leqslant |x_i - x_j|$. Consider the 2-plane $\Pi_{ij} \subset \mathbb{R}^n$ containing $x_i, x_j, 0$. Observe the triangle on Π_{ij} spanned by the vectors x_i and x_j, we can deduce via some elementary binomial identities:

$$|x_i - x_j|^2 - |x_j|^2 \geqslant r_i^2 - r_j^2 = (r_i + r_j)(r_i - r_j) \geqslant r_i(r_i - r_j) = r_i^2 - r_i r_j \geqslant |x_i|^2 - |x_i||x_j|,$$

where we have assumed without loss of generality $r_i \geqslant r_j$. Now rearranging gives us via the cosine rule:

$$|x_i||x_j| \geqslant |x_i|^2 + |x_j|^2 - |x_i - x_j|^2 \Leftrightarrow \frac{1}{2} \geqslant \frac{|x_i|^2 + |x_j|^2 - |x_i - x_j|^2}{2|x_i||x_j|} = \cos\vartheta,$$

where we have denoted by $\vartheta \in (0, \pi]$ the angle between the vectors x_i and x_j. Then $\vartheta \leqslant \arccos(\frac{1}{2}) = \frac{\pi}{2}$. Now projecting x_i and x_j onto the unit circle on Π_{ij} gives us, by using the cosine rule again:

$$\left| \frac{x_i}{|x_i|} - \frac{x_j}{|x_j|} \right| = \sqrt{2 - 2\cos(\vartheta)} \geqslant \sqrt{2 - 2\cos\left(\frac{\pi}{3}\right)} = 1.$$

Now since $\mathbb{S}^{n-1}$ is compact, there exists an upper bound on the cardinality of 1-separated subsets of $\mathbb{S}^{n-1}$. □

Remark 7.3. It is often difficult to remember the proof for the geometric lemma above (if you are not less familiar with elementary trigonometry). It is probably not difficult to realise the necessity to project the centres of the balls to the unit sphere and then use a compactness argument to see that the number of strictly separated centres must be finite. The geometric argument here is to see the unit sphere restricted to a 2-plane as a length space, which then leads us to a use of the cosine rule.

Theorem 7.5 (Besicovitch covering theorem)**.** *There exists a uniform dimensional constant $\beta(n) \in \mathbb{N}$ such that for all family $\mathscr{B}$ of closed balls in $\mathbb{R}^n$ with the set of centres $E \subset\subset \mathbb{R}^n$, there exists $\beta(n)$ disjoint, (at most) countable sub-families $\mathscr{B}_1, \ldots, \mathscr{B}_{\beta(n)} \subset \mathscr{B}$ such that:*

$$E \subset \bigcup_{i=1}^{\beta(n)} \left(\bigcup \mathscr{B}_i \right).$$

In particular, there exists an at most countable covering of E by closed balls $\mathscr{B}_E = \{B_i \mid i \in \mathbb{N}\} \subset \mathscr{B}$ such that every $x \in \mathbb{R}^n$ lies in at most $\beta(n)$ of the closed balls $(B_i \mid i \in \mathbb{N})$.

Proof. Fix for each $x \in E$, some closed ball $\overline{\mathbb{B}}(x, \rho(x))$. Notice that E is precompact in $\mathbb{R}^n$, the open cover $\{\mathbb{B}(x, 2^{-1}r(x)) \mid x \in E\}$ admits a finite subcover. Thus we can assume $M_1 := \sup_{x\in E} \rho(x) < \infty$ without loss of generality by passing to the finite subcover. Pick an arbitrary $x_1^1 \in E$ with $\rho(x_1^1) \geqslant 2^{-1}M_1$ and then select x_{k+1}^1 inductively such that $x_{k+1}^1 \in E \setminus (\bigcup_{i=1}^k \overline{\mathbb{B}}(x_i^1, \rho(x_i)))$ and $\rho(x_i^1) \geqslant 2^{-1}M_1$ until $\rho(x) < 2^{-1}M_1$ for all $x \in E \setminus \{x_i^1\}_{i=1}^{K_1}$ for some finite number $K_1 \in \mathbb{N}$. Now we proceed with the inductive process by setting $K_k := \sup\{\rho(x) \mid x \in E \setminus \bigcup_{i=1}^k \{x_j^i\}_{j=1}^{K_i}\}$. This gives rise to an at most countable increasing sequence $(K_i \mid i \in \mathbb{N}) \in \mathbb{N}$, a decreasing sequence of lower bounds $(M_i \mid i \in \mathbb{N}) \in \mathbb{R}_+$ satisfying $2M_{i+1} \leqslant M_i$ for all $i \in \mathbb{N}$, and collections $(\mathcal{I}_i)$ of indices with $\mathcal{I}_i = \{K_i + 1, \ldots, K_{i+1}\}$ (where we write $K_0 = 0$ for convenience). Thus we obtain a countable sub-family by adjoining all collections from above and relabelling, i.e. :

$$\mathscr{B}_0 = \left\{ \overline{\mathbb{B}}_i = \overline{\mathbb{B}}(x_i, \rho(x_i)) \mid i \in \bigcup_{k=1}^{\infty} \mathcal{I}_k \right\}, \quad \text{where for each } k \in \mathbb{N} \begin{cases} \rho(x_k) \in [\frac{M_i}{2}, M_i] \text{ for some } i \in \mathbb{N}; \\ x_k \in E \setminus (\bigcup_{j=1}^{k-1} \overline{\mathbb{B}}_j). \end{cases}$$

Now if $m < i$, by above it yields for all $k \in \mathcal{I}_i$, $x_k \in E \setminus (\bigcup_{j=1}^{k-1} \overline{\mathbb{B}}_j) \subset E \setminus (\bigcup_{l\in\mathcal{I}_m} \overline{\mathbb{B}}_l)$ by the ordering of the indices collections $(\mathcal{I}_i)$. On the contrary, if $m > i$, then $\rho(x_l) < \rho(x_k)$ for any $l \in \mathcal{I}_m$, which then implies $x_k \in \overline{\mathbb{B}}_l$. In particular, for any $k \in \mathcal{I}_i$, $x_k \in E \setminus (\bigcup_{m\neq i} \bigcup_{l\in\mathcal{I}_m} \overline{\mathbb{B}}_l)$.
Note since (M_i) is decreasing, we have either there are only finitely many M_i or $M_i \to 0$ which then implies $\rho(x_i) \to 0$ as $i \to \infty$. In either case, we would have $E \subset \bigcup_{i\in\mathbb{N}} \overline{\mathbb{B}}_i$.
Now suppose $x \in \bigcap_{j=1}^J \overline{\mathbb{B}}_{k_j}$ for some $(k_j)_{j=1}^J \in \mathbb{N}$. By the geometric technical lemma above, $k_j \in \mathcal{I}_m$ for at most $k(n)$ of $m \in \mathbb{N}$. Writing $\mathcal{I}_m \cap \{k_j\}_{j=1}^J = \{l_1, \ldots, l_N\}$, we deduce by the properties of the collection $\mathscr{B}_0$ that $(\overline{\mathbb{B}}(x_{l_i}, \frac{\rho(x_i)}{4}) \mid i = 1, \ldots, N)$ forms a disjoint collection and $\overline{\mathbb{B}}(x_{l_i}, r(x_{l_i})) \subset \overline{\mathbb{B}}(x, 2M_m)$. Then by disjointness, a volume comparison argument gives us:

$$\begin{aligned} N\omega_n \left(\frac{M_m}{8}\right)^n &\leqslant \sum_{k=1}^N \mu\overline{\mathbb{B}}(x_{l_k}, \frac{\rho(x_{l_k})}{4}) = \mu\left(\bigcup_{k=1}^N \overline{\mathbb{B}}(x_{l_k}, \frac{\rho(x_{l_k})}{4})\right) \\ &\leqslant \mu\left(\bigcup_{k=1}^N \overline{\mathbb{B}}(x_{l_k}, \rho(x_{l_k}))\right) \leqslant \mu\overline{\mathbb{B}}(x, 2M_m) = 2^n\omega_n M_m^n, \end{aligned}$$

since $\frac{M_m}{8} \leqslant \frac{\rho(x_{l_k})}{4} \leqslant 2M_m$. Thus we can pick $\beta(n) := 16^n k(n) < \infty$ for controlling intersection.
Now order the collection $\mathscr{B}_0 = \{\overline{B}_k = \overline{\mathbb{B}}(x_k, \rho(x_k)) \mid k \in \mathbb{N}\}$ with $\rho_k = \rho(x_k) \geqslant \rho_{k+1}$ for all k. Then define in an inductive manner:

$$\overline{B}_{1,1} := \overline{B}_k \quad \text{and} \overline{B}_{1,j+1} = \overline{B}_{k_0}, \quad \text{where} k_0 = \inf\{k \mid \overline{B}_k \cap \left(\bigcup_{i=1}^j \overline{B}_{1,i}\right) = \emptyset\}.$$

Then $\mathscr{B}_1 := \{\overline{B}_{1,j} \mid j \in \mathbb{N}\} \subset \mathscr{B}_0$ is a countable sub-family. Now choose the minimal $k \in \mathbb{N}$ such that $\overline{B}_k \notin \mathscr{B}_1$ (otherwise $\mathscr{B}_1$ a subcover for E) and proceed iteratively by setting $\overline{B}_{2,j} := \overline{B}_{k_2}$ where we have $k_2 := \inf\{k \in \mathbb{N} \mid \overline{B}_k \cap (\bigcup_{i=1}^j \overline{B}_{2,j}) = \emptyset\}$. Iterating the above gives us a disjoint sequence of subfamilies $(\mathscr{B}_i \mid i \in \mathbb{N})$ for $\mathscr{B}_0$, where the sequence is at most countable. The theorem will follow once we can show the existence of some $m \leqslant 64^n k(n) + 1 = 4^d\beta(n) + 1$ such that $\bigcup_{i=1}^m \mathscr{B}_i$ forms a covering for E.
Suppose there is some $x \in E \setminus (\bigcup_{i=1}^m \bigcup \mathscr{B}_i)$ for some $m \in \mathbb{N}$. Since $\mathscr{B}_0 = \{\overline{B}_k\}_k$ indeed forms a cover, we can find some $k \in \mathbb{N}$ such that $x \in \overline{B}_k$ with $\overline{B}_k \notin \mathscr{B}_i$ for all $i = 1, \ldots, m$. Then by construction, for all $j \in \mathbb{N}$, we can find some index i_j such that $\overline{B}_k \cap \overline{B}_{j,i_j} \neq \emptyset$ and $\rho_k \leqslant \rho_{j,i_j}$. Thus there is some open ball $B(y_j, \frac{\rho_k}{2}) \subset 2\overline{B}_k \cap \overline{B}_{j,i_j}$. Now points in $\mathbb{R}^n$ can be contained in at most $\beta(n)$ of $\overline{B}_{j,i_j}$ and thus the same holds for the collection of smaller balls $(\mathbb{B}(y_j, \frac{\rho_k}{2}) \mid j, k)$. Then again using a volume comparison argument:

$$\frac{m2^{-n}\omega_n\rho_k^n}{\beta(n)} = \frac{1}{\beta(n)} \sum_{j=1}^m \mathcal{L}^n(\mathbb{B}(y_j, \frac{\rho_k}{2})) \leqslant \mathcal{L}^n\left(\bigcup_{j=1}^m \mathbb{B}(y_j, \frac{\rho_k}{2})\right) \leqslant \mathcal{L}^n(2\overline{B}_k) \leqslant 2^n\omega_n\rho_k^n.$$

In particular, we have $m \leqslant 4\beta(n)$. Since E is precompact and $\mathscr{B}_0$ forms a covering for E, there is some $m \leqslant 4\beta(n) + 1$ such that $\bigcup_{i=1}^m \mathscr{B}_i$ forms a subcover. □

Theorem 7.6 (Vitali-Besicovitch covering theorem)**.** *Let $E \subset\subset \mathbb{R}^n$ be a Borel set and $\mathscr{B}$ a fine cover for E. Then for each vector-valued Radon measure $\mu \in \mathcal{M}(\mathbb{R}^n; V)$, there exists some countable disjoint sub-family $\mathscr{B}' \subset \mathscr{B}$ covering E upto μ-negligible sets, i.e. $\mu(E \setminus \bigcup \mathscr{B}') = 0$.*

Proof. By the Besicovitch covering theorem, there is some disjoint collection of sub-families $\{\mathscr{B}_i \mid i = 1, \dots, \beta(n)\}$ such that $E \subset \bigcup_{i=1}^{\beta(n)} \bigcup \mathscr{B}_i$. Now by σ-additivity of the total variation measure $|\mu|$:

$$|\mu|(E) = |\mu| \left(E \cap \left(\bigcup_{i=1}^{\beta n} \bigcup \mathscr{B}_i \right) \right) = \sum_{i=1}^{\beta(n)} |\mu|(E \cap \bigcup \mathscr{B}_i).$$

In particular, there must exist some $i = 1, \dots, \beta(n)$ such that:

$$|\mu|(E \cap \bigcup \mathscr{B}_i) \geqslant \frac{1}{\beta(n)} |\mu|(E),$$

which allows us to pick a further sub-family $\mathscr{G}_1 \subset \mathscr{B}_i$ such that:

$$|\mu|(E \cap \bigcup \mathscr{G}_1) \geqslant \frac{1}{2\beta(n)} |\mu|(E).$$

The above can be repeated on subsets of $E_1 := E \setminus (\bigcup \mathscr{G}_1)$ and the corresponding fine cover $\{\overline{B} \in \mathscr{B} \mid \overline{B} \cap \bigcup \mathscr{G}_1\}$ of E_1. This gives rise to a decreasing sequence of sets $(E_i \mid i \in \mathbb{N})$ and subfamilies $(\mathscr{G}_i \mid i \in \mathbb{N})$ satisfying:

$$|\mu|(E_{i+1}) \leqslant (1 - \frac{1}{2\beta(n)}) |\mu|(A_i) \quad \text{for all } i \in \mathbb{N}.$$

Note also by construction $\mathscr{B}' := \bigcup_{i \in \mathbb{N}} \mathscr{G}$ is a disjoint countable sub-family such that:

$$E \setminus \bigcup \mathscr{B}' = E \setminus \bigcup_{i \in \mathbb{N}} \bigcup \mathscr{G}_i = \bigcap_{i \in \mathbb{N}} (E \setminus \bigcup \mathscr{G}_i) \subset \bigcap_{i \in \mathbb{N}} E_i,$$

where $\bigcap_{i=1}^{\infty} E_i$ defines a $|\mu|$-negligible and thus μ-negligible set. □

Definition 7.3 (Density functions)**.** *Let $\mu \in \mathcal{M}^+(\mathbb{R}^n)$ and $\nu \in \mathcal{M}(\mathbb{R}^n; \mathbb{R})$. We define the lower and upper density of ν with respect to μ at each $x \in \operatorname{supp} \mu$ via:*

$$D_\mu^- \nu(x) := \liminf_{\rho \downarrow 0} \frac{\nu \mathbb{B}(x, \rho)}{\mu \mathbb{B}(x, \rho)} \quad \textit{and} \quad D_\mu^+ \nu(x) := \limsup_{\rho \downarrow 0} \frac{\nu \mathbb{B}(x, \rho)}{\mu \mathbb{B}(x, \rho)}.$$

If the lower and upper densities at $x \in \operatorname{supp} \mu$ coincide, then we say that blow-up quotient $\frac{d\nu}{d\mu}$ exists at x.

Lemma 7.3 (Density lemma)**.** *Let $\mu, \nu \in \mathcal{M}(\mathbb{R}^n; \mathbb{R})$ be Radon measures and $E \subset \operatorname{supp} \mu$ a Borel subset. Then $\{D_\mu^+ \nu = \infty\}$ is μ-negligible and moreover for any $t \in \mathbb{R}$, we have:*

- *(i) if $D_\mu^- \nu \leqslant t$ on E, then $\nu E \leqslant t \mu E$;*
- *(ii) if $D_\mu^+ \nu \geqslant t$ on E, then $\nu E \geqslant t \mu E$.*

Proof. (i) Assume without loss of generality that E is bounded. For any open neighbourhood $U \supset E$ and $\varepsilon > 0$, we want to apply the Besicovitch-Vitali theorem on the following collection of closed balls:

$$\mathscr{F}_\varepsilon := \{\overline{\mathbb{B}}(x, \rho) \mid x \in E, \, \mathbb{B}(x, \rho) \subset U, \, \nu \overline{\mathbb{B}}(x, \rho) \leqslant (t + \varepsilon) \mu \overline{\mathbb{B}}(x, \rho)\}.$$

Indeed, suppose $D_\mu^- \nu(x) = \liminf_{\rho \downarrow 0} \frac{\nu \mathbb{B}(x,\rho)}{\mu \mathbb{B}(x,\rho)} \leqslant t$ for all $x \in E$. Then for any $\rho > 0$ sufficiently small, we must have for all $x \in E$, $\nu \overline{\mathbb{B}}(x, \rho) \leqslant (t + \varepsilon) \mu \overline{\mathbb{B}}(x, \rho)$. In particular $\mathscr{F}_\varepsilon$ defines a fine cover for E. Therefore, the Besicovitch-Vitali theorem applies and there is an at most countable, disjoint subfamily $\mathscr{F}'_\varepsilon \subset \mathscr{F}_\varepsilon$ such that:

$$\nu(E \setminus \bigcup \mathscr{F}'_\varepsilon) = 0 \Rightarrow \nu E \leqslant \sum_{B \in \mathscr{F}'_\varepsilon} \nu B \leqslant (t + \varepsilon) \sum_{B \in \mathscr{F}'_\varepsilon} \mu B \leqslant (t + \varepsilon) \mu U,$$

where the last inequality follows from disjointness and $B \subset U$ for all $B \in \mathscr{F}'_\varepsilon \subset \mathscr{F}_\varepsilon$ by definition. Sending $\varepsilon \searrow 0$ allows us to conclude by outer regularity of μ on Borel sets.
(ii) Similarly, suppose now $D_\mu^- \nu(x) = \limsup_{\rho \downarrow 0} \frac{\nu \mathbb{B}(x,\rho)}{\mu \mathbb{B}(x,\rho)} \geqslant t$ for all $x \in E$. Then $\nu \overline{\mathbb{B}}(x, \rho) \geqslant (t - \varepsilon) \mu \overline{\mathbb{B}}(x, \rho)$ for all $\rho > 0$ small. We can again define the fine cover for E for any arbitrary open neighbourhood $U \supset E$:

$$\mathscr{F}^\varepsilon := \{\overline{\mathbb{B}}(x, \rho) \mid x \in E, \, \mathbb{B}(x, \rho) \subset U, \, \nu \overline{\mathbb{B}}(x, \rho) \geqslant (t - \varepsilon) \mu \overline{\mathbb{B}}(x, \rho)\},$$

from which we can find a countable, disjoint sub-collection $\mathscr{F}^\varepsilon_* \subset \mathscr{F}^\varepsilon$ satisfying:

$$\mu(E \setminus \bigcup \mathscr{F}^\varepsilon_*) = 0 \to (t-\varepsilon)\mu E \leqslant (t-\varepsilon) \sum_{B \in \mathscr{F}^\varepsilon_*} \mu B \leqslant \sum_{B \in \mathscr{F}^\varepsilon_*} \nu B = \nu U.$$

Then we can conclude as in (i). Let $E := \{D^+_\mu \nu = \infty\}$ and notice $E = \bigcap_k E_k$ where $E_k := \{D^+_\mu \nu \geqslant k\}$ for $k \in \mathbb{N}$. In particular, for each fixed k, applying (ii) to $E_k \in \mathcal{B}(\mathbb{R}^n)$ gives us $\nu E_k \geqslant t\mu E_k$ for all $t \in \mathbb{R}$. Now sending $t \to \pm\infty$ forces $\mu E_k = 0$ for all k. Thus $E \subset E_k$ is μ-negligible. $\square$

Theorem 7.7 (Besicovitch differentiation theorem)**.** *Let $\mu \in \mathcal{M}^+(\mathbb{R}^n)$ and $\nu \in \mathcal{M}(\mathbb{R}^n; V)$ be Radon measures where V is some finite-dimensional normed vector space. Then we have:*

$$\frac{d\nu}{d\mu}(x) = \lim_{\rho \downarrow 0} \frac{\nu \mathbb{B}(x,\rho)}{\mu \mathbb{B}(x,\rho)} \textit{ exists in } V \quad \mu\textit{-a.e. on } \operatorname{supp}\mu.$$

In particular, we obtain the Lebesgue-Radon-Nikodym decomposition of ν:

$$\nu = \frac{d\nu}{d\mu}\mu + \nu^s, \quad \textit{where} \begin{cases} \nu^s = \nu|_E \in \mathcal{M}(\mathbb{R}^n; V); \\ E := (\mathbb{R}^n \setminus \operatorname{supp}\mu) \cup \{x \in \operatorname{supp}\mu \mid \lim_{\rho\downarrow 0} \frac{|\nu|(\mathbb{B}(x,\rho))}{\mu\mathbb{B}(x,\rho)} = \infty\}. \end{cases}$$

Proof. Arguing componentwise, we may reduce to the case where ν is a signed measure. Then by the density lemma, $D^+_\mu \nu$ is finite μ-a.e. We consider the following set functions defined on $\mathcal{B}(\mathbb{R}^n)$:

$$\lambda^+(B) := \int_B D^+_\mu \nu d\mu, \quad \lambda^-(B) := \int_B D^-_\mu \nu d\mu \quad \text{for all } B \in \mathcal{B}(\mathbb{R}^n).$$

Fix $t > 1$ and choose a Borel set $B \subset \operatorname{supp}\mu$ with $D^+_\mu \nu < \infty$ everywhere on B. Now partition our choice of B as the disjoint union of $B_k := \{x \in B \mid D^+_\mu \nu(x) \in (t^k, t^{k+1}]\} \subset \mathbb{R}^n \setminus E$ for $k \in \mathbb{Z}$. By the density lemma:

$$\lambda^+(B_k) \leq t^{k+1}\mu B_k \leqslant t\nu B_k \Rightarrow \lambda^+(B) = \sum_{k\in\mathbb{Z}} \lambda^+(B_k) \leqslant t \sum_{k \in \mathbb{Z}} \nu B_k = t(\nu|_{\mathbb{R}^\kappa \setminus E})(B),$$

which follows from the disjointness property. Sending $t \searrow 1$ gives us $\lambda^+(B) \leqslant t(\nu|_{\mathbb{R}^n \setminus E})(B)$ for any Borel subset $B \subset \mathbb{R}^n \setminus E$. Using an analogous approach, for any Borel set $B \subset \operatorname{supp}\mu$ with $D^-_\mu \nu \in (0,\infty)$ on B, we define $B_j := \{x \in B \mid D^-_\mu \nu(x) \in (t^k, t^{k+1}]\}$. We deduce again from the density lemma:

$$\nu B_k \leqslant t^{k+1}\mu B_k \Rightarrow \lambda^-(B) = \sum_{k\in\mathbb{Z}} \lambda^-(B_k) \geqslant \frac{1}{t} \sum_{k\in\mathbb{Z}} \nu B_k = \frac{1}{t}(\nu|_{\mathbb{R}^n\setminus E})(B).$$

In particular, we have $\lambda^+ \leqslant \nu|_{\mathbb{R}^n \setminus E} \leqslant \lambda^-$ on $\mathbb{R}^n \setminus E$, which allows us to conclude:

$$D^+_\mu \nu = D^-_\mu \nu = \frac{d\nu}{d\mu} \quad \mu\text{-a.e. on } \operatorname{supp}\mu \setminus E.$$

Furthermore, the decomposition of measure holds with $d\nu|_{\mathbb{R}^d \setminus E} = \frac{d\nu}{d\mu} d\mu$ and $\nu^s = \nu|_E$. $\square$

Remark 7.4. If V is in fact an inner product space, we can use a duality argument, in the sense of the Riesz-Markov-Kakutani theorem, to give the following alternative characterisation of the Lebesgue-Besicovitch differentiation theorem:

$$\langle f, \nu \rangle = \int \langle f, \nu \rangle_V = \int \langle f, \frac{d\nu}{d\mu} \rangle_V d\mu + \int \langle f, d\nu^s \rangle_V \quad \text{for all } f \in C^0_c(\mathbb{R}^n; V).$$

Corollary 7.3 (Radon-Nikodym)**.** *Let $\mu \in \mathcal{M}^+(\mathbb{R}^n)$ and $\nu \in \mathcal{M}(\mathbb{R}^n; V)$ a vector-valued measure where $(V, \|\cdot\|)$ a normed vector space. If $\nu << \mu$, then the Radon-Nikodym derivative of ν with respect to μ is μ-integrable and:*

$$\nu = \frac{d\nu}{d\mu}\mu, \quad \textit{i.e. } \nu E = \int \chi_E \frac{d\nu}{d\mu} d\mu \quad \textit{for all } E \in \mathcal{B}(\mathbb{R}^n).$$

Proof. The absolute continuity gives us $\operatorname{supp}\nu \subset \operatorname{supp}\mu$. Note also that any Radon measure is necessarily absolutely continuous with respect to its own total variation measure. If the support $E = \operatorname{supp}(\nu^s) \subset \operatorname{supp}\nu \subset \operatorname{supp}\mu$ of the singular part ν^s (as defined in Theorem 7.7) has nonzero $|\nu|$-measure, using the density lemma we can deduce $0 = |\nu|(E) \geqslant t\mu E$ for any $t \in (0, \infty)$ since we have $\mu \llcorner E$ is supported on a subset where $D^+_\mu|\nu| = \infty$. However, the absolute continuity gives us $\nu E = 0$ and thus $|\nu|(E) = 0$ which is a contradiction. Hence the singular part ν^s has ν-negligible support and the result follows from Besicovitch's theorem. $\square$

Remark 7.5. An application of the Radon-Nikodym theorem, by noting $\nu << |\nu|$, gives us the existence of the polar $\frac{d\nu}{d|\nu|}$ pointwise $|\nu|$-a.e. and it is rather clear to see that $\|\frac{d\nu}{d|\nu|}\| = 1$ $|\nu|$-a.e. Indeed, consider some auxillary Radon measure $\mu \in \mathcal{M}^+(\mathbb{R}^n)$ such that $\nu << \mu$. Then the Radon-Nikodym decomposition gives us:

$$\frac{d\nu}{d\mu}d\mu = d\nu = \frac{d\nu}{d|\nu|}d|\nu| = \frac{d\nu}{d|\nu|}\left|\frac{d\nu}{d\mu}\right|d\mu \quad \mu\text{-a.e. on } \mathbb{R}^n,$$

where the last equality follows by Remark 5.1. Thus $|\frac{d\nu}{d|\nu|}| = 1$ pointwise $|\nu|$-a.e. since $|\frac{d\nu}{d\mu}| > 0$ $|\nu|$-a.e. on $\mathbb{R}^n$. We can now derive the Lebesgue's differentiation theorem on Euclidean spaces as a consequence of Besicovitch's differentiation theorem for Radon measures.

Corollary 7.4 (Lebesgue's differentiation theorem on $\mathbb{R}^n$)**.** *Let $\mu \in \mathcal{M}^+(\mathbb{R}^n)$ be a positive Radon measure. Then we have for any $f \in L^1(\mu)$:*

$$⨍_{\mathbb{B}(x,\rho)} |f(y) - f(x)|d\mu(y) \overset{\rho\downarrow 0}{\longrightarrow} 0 \quad \textit{for } \mu\textit{-a.e. } x \in \mathbb{R}^n.$$

Proof. For any fixed $q \in \mathbb{Q}^d$, by Besicovitch's theorem with $d\nu = |f - q|d\mu$, we obtain for all $\rho > 0$:

$$\nu\mathbb{B}(x,\rho) = \int_{\mathbb{B}(x,\rho)} |f(y) - q|d\mu(y) \text{ for all } x \in \operatorname{supp}\mu \setminus N_q,$$

where N_q is μ-negligible. In particular, sending $\rho \searrow 0$ gives us:

$$⨍_{\mathbb{B}(x,\rho)} |f(y) - q|d\mu(y) = \frac{\nu\mathbb{B}(x,\rho)}{\mu\mathbb{B}(x,\rho)} \overset{\rho\downarrow 0}{\longrightarrow} |f(x) - q|.$$

Now notice that $\bigcup_{q\in\mathbb{Q}^n} N_q$ remains μ-negligible and for all $x \in \operatorname{supp}\mu \setminus (\bigcup_{q\in\mathbb{Q}^n} N_q)$:

$$\limsup_{\rho\downarrow 0} ⨍_{\mathbb{B}(x,\rho)} |f(y) - f(x)|d\mu(y) \leqslant |f(x) - q| + \lim_{\rho\downarrow 0} ⨍_{\mathbb{B}(x,\rho)} |f(y) - q|d\mu(y) = 2|f(x) - q|.$$

Now the theorem follows from the density of $\mathbb{Q}^n \subset \mathbb{R}^n$. $\square$

We will now turn to the Rademacher's theorem, which guarantees the classical differentiability of a locally Lipschitz function upto a $\mathcal{L}^n$-negligible set. The Lipschitz condition on a local neighbourhood gives us a uniform bound on the difference quotients. In the 1-dimensional case, we might expect via some compactness argument to deduce the pointwise derivative at certain points. This intuition in fact extends to the higher-dimensional scenarios as well.

Theorem 7.8 (Rademacher's theorem)**.** *Let $f \in C^{0,1}_{loc}(\mathbb{R}^n; \mathbb{R}^m)$. Then f is classically differentiable $\mathcal{L}^n$-a.e.*

Proof. Arguing componentwise or realising that differentiability is a local property, we may suppose without loss of generality that $m = 1$ and f is uniformly Lipschitz. Consider first the 1-dimensional case and define the set functions μ on half-rays on $\mathbb{R}$ via $\mu(-\infty, t) := f(t)$. Then $\mu \in \mathcal{M}(\mathbb{R})$ becomes a signed measure after extension by the Hahn-Carathéodory construction. Now we notice that for any $s < t$:

$$\mu[s,t) = f(t) - f(s) \leqslant [f]_{C^{0,1}(\mathbb{R})}|t - s|,$$

which clearly implies $\mu << \mathcal{L}^1$. By the Radon-Nikodym theorem, there exists some $h \in L^1(\mathbb{R})$ such that:

$$f(t) - f(s) = \mu[s,t) = \int_s^t h(\tau)d\mathcal{L}^1(\tau) \;\Rightarrow\; \frac{f(t+h) - f(t)}{h} ⨍_t^{t+h} h(\tau)d\mathcal{L}^1(\tau) \longrightarrow h(t),$$

for $\mathcal{L}^1$-a.e. $t \in \mathbb{R}$ by an application of Lebesgue's differentiation theorem. Thus f is classically differentiable at every Lebesgue point on $\mathbb{R}$ with derivative coinciding with the Radon-Nikodym derivative.
Now consider the remaining cases where $n > 1$. For fixed $v \in \mathbb{S}^{n-1} \subset \mathbb{R}^n$ and consider the set given by:

$$N_v := \{x \in \mathbb{R}^n \mid D_v f(x) := \frac{d}{dt}\bigg|_{t=0} f(x+tv) \text{ does not exist}\} = \mathbb{R}^n \setminus \left\{ \lim_{t \downarrow 0} \frac{f(x+tv) - f(x) - tD_v f(x)}{t} = 0 \right\},$$

where the directional derivative does not exist. The second equality above shows us N_v is $\mathcal{L}^n$-measurable. Notice that for all fixed $x \in \mathbb{R}^d$, the mapping $t \mapsto f(x+tv)$ is uniformly Lipschitz and thus admits a classical derivative for $\mathcal{L}^1$-a.e. $t \in \mathbb{R}$. In particular, writing $v_1 = v$ and completing it to a $g_{\mathbb{R}^n}$-orthonormal basis $(v_i \mid i = 1, \ldots, n)$, we can deduce from above that for all $v \in \mathbb{S}^{n-1}$ and $x \in \mathbb{R}^n$:

$$\mathcal{L}^1(N_v \cap (x + \mathbb{R}v)) = 0 \ \Rightarrow \ \mathcal{L}^n(N_v) = \int_{x+\mathbb{R}v_n} \cdots \int_{x+\mathbb{R}v_2} \mathcal{L}^1(N_v + \mathbb{R}v_1) d\tau_2 \ldots d\tau_n = 0,$$

by using the Fubini's theorem to integrating along each unit direction. Thus for any $v \in \mathbb{S}^{n-1}$, the directional derivative $D_v f$ exists pointwise $\mathcal{L}^n$-a.e. In particular, f admits classical partial derivatives on $\mathbb{R}^n$ upto a $\mathcal{L}^n$-negligible set $N = \bigcup_{i=1}^n N_{e_i}$ and thus gives classical differentiability along any axis. Now it remains to show the existence of a Frechét derivative on a set of full measure.
Let $\varphi \in C_c^\infty(\mathbb{R}^n)$ be arbitrary. Then using the $\mathcal{L}^n$-a.e. existence of partial derivatives and Lebesgue's dominated convergence theorem, we can deduce by writing $D_i^h f = h^{-1}(f(\cdot + he_i) - f)$:

$$\int \varphi D_i^h f = -\int f D_i^{-h} \varphi \xrightarrow{h \downarrow 0} -\int f D_i \varphi = \int \varphi D_i f \quad \text{ for all } i = 1, \ldots, n.$$

By linearity of directional derivatives, we have for any $v \in \mathbb{S}^{n-1}$, writing $D_v^h f = h^{-1}(\tau_{hv} f - f)$

$$\int \varphi D_v^h f = -\int f D_v^h \varphi \xrightarrow{h \downarrow 0} -\int f D_v \varphi = -\int f v^i D_i \varphi = \int v^i \varphi D_i f = \int \varphi \langle v, Df \rangle_{Euc},$$

where as usual, we have adopted the Einstein summation convention and $v = v^i e_i$. Since $\varphi \in C_c^\infty(\mathbb{R}^n)$ was chosen arbitrarily, we can deduce $D_v f = \langle v, Df \rangle_{Euc}$ $\mathcal{L}^n$-a.e. by using the fundamental lemma of calculus of variations. This gives us a basic pointwise estimate on the derivative since f is uniformly Lipschitz:

$$|Df(x)| = \left(\sum_{i=1}^n \langle e_i, Df(x) \rangle_{Euc}^2 \right)^{\frac{1}{2}} \leqslant \limsup_{h \downarrow 0} \left(\sum_{i=1}^n \left(\frac{f(x + he_i) - f(x)}{h} \right)^2 \right)^{\frac{1}{2}} \leqslant \sqrt{n} [f]_{C^{0,1}(\mathbb{R}^n)},$$

for $\mathcal{L}^n$-a.e. $x \in \mathbb{R}^n$. Now recall that $\mathbb{S}^{n-1}$ is a separable compact metric space. Let $(v_i \mid i \in \mathbb{N})$ be a dense subset of $\mathbb{S}^{n-1}$ and setting $Z_i := N_{v_i} \cup N$ (which remains $\mathcal{L}^n$-negligible) for each $i \in \mathbb{N}$. We will show that f is Frechét differentiable at each $x \in \mathbb{R}^n \setminus (\bigcup_{i=1}^\infty Z_i)$. Indeed, fix one such x, then for any arbitrary $v, w \in \mathbb{S}^{n-1}$:

$$\begin{aligned} |(D_v^h f(x) - \langle v, Df(x) \rangle) - (D_w^h f(x) - \langle w, Df(x) \rangle)| &\leqslant \frac{|f(x + hv) - f(x + hw)|}{|h|} + |\langle v - w, Df(x) \rangle| \\ &\leqslant [f]_{C^{0,1}(\mathbb{R}^n)} |v - w| + |Df(x)||v - w| \leqslant (\sqrt{n} + 1)[f]_{C^{0,1}(\mathbb{R}^n)} |v - w|. \end{aligned}$$

Now fix an arbitrarily small $\varepsilon > 0$ and set $\delta(n) = \frac{\varepsilon}{(1+\sqrt{n})[f]_{C^{0,1}(\mathbb{R}^n)}}$. By compactness, we can cover $\mathbb{S}^{n-1}$ by finitely many balls $(B_i \mid i = 1, \ldots, N)$ of radius $\delta(n)$. By density of (v_i) in $\mathbb{S}^{n-1}$, we may as well assume (upto reordering) $v_i \in B_i$ for all i. For any $v \in \mathbb{S}^{n-1}$, there exists some $i = 1, \ldots, n$ with $v \in B_i$. In particular:

$$\begin{aligned} \limsup_{h \downarrow 0} |D_v^h f(x) - \langle v, Df(x) \rangle_{Euc}| &\leqslant \limsup_{h \downarrow 0} |D_{v_i}^h f(x) - \langle v_i, Df(x) \rangle_{Euc}| + (1 + \sqrt{n})[f]_{C^{0,1}(\mathbb{R}^n)} |v - v_i| \\ &\leqslant (1 + \sqrt{n})[f]_{C^{0,1}(\mathbb{R}^n)} \delta(n) \leqslant (1 + \sqrt{n})[f]_{C^{0,1}(\mathbb{R}^n)} \frac{\varepsilon}{(1 + \sqrt{n})[f]_{C^{0,1}(\mathbb{R}^n)}} \leqslant \varepsilon. \end{aligned}$$

Since $(B_i \mid i = 1, \ldots, N)$ is a finite covering, we can make the above estimate uniform over $v \in \mathbb{S}^{n-1}$. The choice of ε above is arbitrary, which then allows us to deduce:

$$D_v^h f - \langle v, Df \rangle_{Euc} = \frac{f(\cdot + hv) - f}{h} - Df \cdot v \xrightarrow{h \downarrow 0} 0 \quad \text{ uniformly in } v \in \mathbb{S}^{n-1}.$$

This gives us the Frechét differentiability of f on a set of full measure. $\square$

We close this section with some technical consequences of Rademacher's theorem on level sets.

Theorem 7.9 (Differentiability on level sets)**.** *Let $f, h \colon \mathbb{R}^n \to \mathbb{R}^m$ be locally Lipschitz continuous. Then:*

(i) $Df(x) = 0$ *for $\mathcal{L}^n$-a.e. $x \in \{f = 0\}$;*

(ii) *if $m = n$, then $Dh(f(x))Df(x) = \mathrm{Id}$ for $\mathcal{L}^n$-a.e. $x \in \{h \circ f = \mathrm{Id}\}$.*

Proof. (i) Let $m = 1$ and $x \in \{f = 0\}$ a Lebesgues density point for the zero level-set where $Df(x)$ exists. If $Df(x) \neq 0$, then consider the following subset:

$$S = \left\{ v \in \mathbb{S}^{n-1} \mid \langle v, Df(x) \rangle_{Euc} \geqslant \frac{1}{2}|Df(x)| \right\},$$

which is necessarily non-empty. Now for all $v \in S$ and $h > 0$, we have:

$$f(x+hv) > f(x+hv) - \frac{h}{2}|Df(x)| \geqslant f(x+hv) - h\langle v, Df(x) \rangle_{Euc} = f(x+hv) - f(x) - \langle hv, Df(x) \rangle_{Euc} \xrightarrow{h \downarrow 0} 0,$$

which contradict the fact that x is a point of density 1 on $\{f = 0\}$. Thus $Df(x) = 0$ for any Lebesgue density point of the zero level-set.
(ii) Consider the following sets of full measures:

$$E_f := \{x \in \mathbb{R}^n \mid Df(x) \text{ exists}\} \quad \text{and} \quad E_h := \{x \in \mathbb{R}^n \mid Dh(x) \text{ exists}\}.$$

Then denote by $Y := \{h \circ f = \mathrm{Id}\}$ and consider the set given by:

$$F := Y \cap E_f \cap f^{-1}(E_h) = \{x \in \mathbb{R}^n \mid h \circ f(x) = x,\ Df(x), Dh(f(x)) \text{ exists}\}.$$

Note for all $x \in Y \setminus (f^{-1}(B))$, we have $f(x) \in \mathbb{R}^n \setminus B$ which then implies that $x = h(f(x)) \in h(\mathbb{R}^n \setminus E_h)$. Thus we obtain the following inclusion:

$$Y \setminus F \subset (\mathbb{R}^n \setminus E_f) \cup h(\mathbb{R}^n \setminus E_h),$$

which then implies by Rademacher's theorem and the scaling of Lebesgue measure via Lipschitz mappings:

$$\mathcal{L}^n(Y \setminus F) \leqslant \mathcal{L}^n(\mathbb{R}^n \setminus E_f) + \mathcal{L}^n(h(\mathbb{R}^n \setminus E_h)) \leqslant \mathcal{L}^n(\mathbb{R}^n \setminus E_f) + [h]^n_{C^{0,1}(\mathbb{R}^n)} \mathcal{L}^n(\mathbb{R}^n \setminus E_h) = 0.$$

Thus F is a subset of full measure in Y. Now the theorem follows aftering realising that for any $x \in F$, both derivatives $Df(x)$ and $Dh(f(x))$ exist and then via chain rule, the derivative of the composition exists and:

$$Dh(f(x))Df(x) = D(h \circ f)(x) = D(\mathrm{Id}),$$

since $h \circ f = \mathrm{Id}$ on Y. Thus it yields $D(h \circ f) = \mathrm{Id}$ pointwise $\mathcal{L}^n$-a.e. on Y. □

References

[AFP00] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs, 2000.

[Fol13] G.B. Folland. *Real Analysis: Modern Techniques and Their Applications*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 2013.