For $\mathcal{A}$ a σ-algebra, $E \in \mathcal{A}$, define $\mathcal{P}^*(E, \mathcal{A}) := \{\{E_k \in \mathcal{A}; k \in \mathbb{N}\}; E = \cup_k E_k, E_i \cap E_j = \phi \; \forall i \neq j\}$. Always $\{\phi, E\} \in \mathcal{P}^*(E, \mathcal{A})$, so $\mathcal{P}^*(E, \mathcal{A})$ is never empty, and also $\mathcal{P}^*(\phi, \mathcal{A}) = \{\phi\}$.

For $\mathcal{A}$ a σ-algebra, define
$S^{\pm}(\mathcal{A}) := \{\sum_{k=1}^n c_k \, \mathbf{1}_{E_k}; \{c_k \in \mathbb{R}; c_i \neq c_j \, \forall i \neq j\}, \{E_k \in \mathcal{A}; E_i \cap E_j = \phi \, \forall i \neq j\}, n \in \mathbb{N}\}$
$S^{+}(\mathcal{A}) := \{\sum_{k=1}^n c_k \, \mathbf{1}_{E_k}; \{c_k \in [0, \infty], c_i \neq c_j \, \forall i \neq j\}, \{E_k \in \mathcal{A}; E_i \cap E_j = \phi \, \forall i \neq j\}, n \in \mathbb{N}\}$

A complex or a signed and finite measure on a measurable space $(X, \mathcal{A})$ is a function, ν, from $\mathcal{A}$ to $\mathbb{R}$ or $\mathbb{C}$ such that
1) $\nu(\phi) = 0$
2) $\nu(\cup_{k \in \mathbb{N}} E_k) = \sum_{k \in \mathbb{N}} \nu(E_k)$, $E_k \in A$, disjoint.

Because the union in (2) is independent of the labeling of the $\{E_k\}$, the sum in (2) is rearangement-invariant, which implies that it converges iff it does so to absolutely, and does to the same number.

Alternatively, a complex measure ν on $(X, \mathcal{A})$ is a complex function on $\mathcal{A}$ such that
3) $\nu(E) = \sum_{k \in \mathbb{N}} \nu(E_k), \;$ for all $\{E_k\} \in \mathcal{P}^*(E, \mathcal{A})$.

$(3 \Rightarrow 1), \phi = E = \cup_k E_k \Rightarrow E_k = \phi. \Rightarrow \nu(\phi) = \sum_{k \in \mathbb{N}} \nu(\phi) \Rightarrow \nu(\phi) = 0.$ $(3 \Leftrightarrow 2), E := \cup_k E_k$.

Write $\mathcal{M}(X, \mathcal{A})$ or $\mathcal{M}(\mathcal{A})$ for the set of all complex or signed and finite measures on $\mathcal{A}$. Write $\mathcal{M}^{\pm}(X, \mathcal{A})$ or $\mathcal{M}^{\pm}(\mathcal{A})$ for the set of all signed and finite measures on $\mathcal{A}$. Write $\mathcal{M}^{+}(X, \mathcal{A})$ or $\mathcal{M}^{+}(\mathcal{A})$ for the set of all positive measures on $\mathcal{A}$. Positive measures need not be finite, so $\mathcal{M}^{+}(\mathcal{A}) \not\subset \mathcal{M}(\mathcal{A})$.

If $\mu_1, \mu_2 \in \mathcal{M}^{+}(\mathcal{A})$, then we say that $\mu_1 \leq \mu_2$ iff $\mu_1(E) \leq \mu_2(E)$ for all $E \in \mathcal{A}$.

Given $\nu \in \mathcal{M}(\mathcal{A})$, we wish to find the smallest $\mu \in \mathcal{M}^{+}(\mathcal{A})$ s.t.

$$\mu(E) \geq |\nu(E)| \text{ for all } E \in \mathcal{A} \; ^{(\dagger_1)},$$

smallest in the sense of the previous point. When $(\dagger_1)$ holds we say that μ dominates ν. Let $\{E_k\} \in \mathcal{P}^*(E, \mathcal{A})$ arbitrarily, we then have that $|\nu(E_k)| \leq \mu(E_k)$ for all E_k by $(\dagger_1)$, summing these gives

$$\mu(E) = \sum_{k \in \mathbb{N}} \mu(E_k) \geq \sum_{k \in \mathbb{N}} |\nu(E_k)| \geq |\nu(E)|, \;\; \text{for all } \{E_k\} \in \mathcal{P}^*(E, \mathcal{A}).$$

Thus, for any μ dominating ν, we can find a $\sum_{k \in \mathbb{N}} |\nu(E_k)|$ not strictly between any $\mu(E)$ and $|\nu(E)|$. So the best we could do, in the sense of minimizing $(\dagger_1)$, is $\sum_{k \in \mathbb{N}} |\nu(E_k)|$, for some $\{E_k\} \in \mathcal{P}^*(E, \mathcal{A})$ which minimizes this quantity. This suggests the definition

$$|\nu|(E) := \sup \left\{ \sum_{k \in \mathbb{N}} |\nu(E_k)|; \; \{E_k\} \in \mathcal{P}^*(E, \mathcal{A}) \right\}.$$

Briefly, $(\dagger_1)$ holds because this sup is an upper bound, and the "smallest" criterion holds because the sup is the smallest such upper bound. This quantity is called the total variation measure of ν.

(Rudin 6.2) $(X, \mathcal{A})$ measurable, $\nu \in \mathcal{M}(X, \mathcal{A})$, then $|\nu| \in \mathcal{M}^+(X, \mathcal{A})$, and $|\nu| \leq \mu$ for all $\mu \in \mathcal{M}^+(X, \mathcal{A})$ satisfying $\mu(E) \geq |\nu(E)|$ for all $E \in \mathcal{A}$ $^{(\dagger_1)}$.

Proof:

First, for $E \in \mathcal{A}$, let $F = \{\sum_{k\in\mathbb{N}} |\nu(E_k)|;\ \{E_k\} \in \mathcal{P}^*(E, \mathcal{A})\}$ is a well defined set, because $\mathcal{P}^*(E, \mathcal{A}) \subset \mathcal{A}$, so that these sums are well defined. Note that $F \subset \mathbb{R}$, if F is unbounded, then $|\nu|(E) = \infty$, otherwise F is bounded and this $\sup(F) \in \mathbb{R}$ exists.

$\mathcal{P}^*(\phi, \mathcal{A}) = \{\phi\} \Rightarrow |\nu|(\phi) = |\nu(\phi)| = 0$.
For any $\{E_i\} \in \mathcal{P}^*(E, \mathcal{A})$, that $|\nu|(E) = \sum_{i\in\mathbb{N}} |\nu|(E_i)$ follows by "$\leq$" and "$\geq$" cases.

"$\geq$": If $|\nu|(E) = \infty$ then this case always holds, so assume $|\nu|(E) < \infty$. $\{E_i\} \in \mathcal{P}^*(E, \mathcal{A})$ is given. Pick $\{t_i \in \mathbb{R}; i \in \mathbb{N}, t_i \geq 0\}$ such that $t_i < |\nu|(E_i)$, but if $|\nu|(E_i) = 0$, then let $t_i = 0$. Given each t_i we can find a partition of E_i, $\{A_{i,j}\} \in \mathcal{P}^*(E_i, \mathcal{A})$, such that $\sum_{j\in\mathbb{N}} |\nu(A_{i,j})| \geq t_i$. Each E_i has atleast one well defined partition; at a minimum $\{E_i, \phi\} \in \mathcal{P}^*(E_i, \mathcal{A})$. If this is the only partition in $\mathcal{P}^*(E_i, \mathcal{A})$, then $|\nu|(E_i) = |\nu(E_i)|$, and in this case $\sum_{j\in\mathbb{N}} |\nu(A_{i,j})| = |\nu(E_i)| = |\nu|(E_i) > t_i$. Now we have that $\{A_{i,j}; i, j \in \mathbb{N}\} \in \mathcal{P}^*(E, \mathcal{A})$, so that by the sup in the difinition of $|\nu|$, summing over i,

$$|\nu|(E) \geq \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} |\nu(A_{i,j})| \geq \sum_{i=1}^{\infty} t_i.$$

Lemma about $\mathbb{R}$, Let $L \in \mathbb{R}, \{a_k \in \mathbb{R}; k \in \mathbb{N}\}$, then

$$\left(\sum_{k=1}^{n} t_k \leq L \text{ for all } \{t_k\} \in \mathbb{R}^n,\ t_k \leq a_k,\ n \in \mathbb{N}\right) \Rightarrow \sum_{k=1}^{\infty} a_k \leq L.$$

If $n = 1$, suppose $a > L$, then can find some $t \in \mathbb{R}$ s.t. $L < t \leq a$, so the statement $(t \leq L\ \forall\, t \in \mathbb{R}, t \leq a)$ contradicts $a > L$, but either $a > L$ or $a \leq L$. Suppose the lemma is true for the case $n \in \mathbb{N}$, fixed, then $\sum_{k=1}^{n+1} t_k \leq L$ for all $\{t_k\} \in \mathbb{R}^{n+1}$, $t_k \leq a_k \Rightarrow \sum_{k=1}^{n} t_k \leq L - t_{k+1}$ for all $\{t_k\} \in \mathbb{R}^n, t_{k+1} \in \mathbb{R},\ t_k \leq a_k \Rightarrow$ (by statement is true for $n \in \mathbb{N}$) $\sum_{k=1}^{n} a_k \leq L - t_{k+1}, t_{k+1} \in \mathbb{R},\ t_{k+1} \leq a_{k+1} \Rightarrow t_{k+1} \leq L - \sum_{k=1}^{n} a_k, t_{k+1} \in \mathbb{R},\ t_{k+1} \leq a_{k+1} \Rightarrow$ (by statement is true for $n = 1$) $a_{k+1} \leq L - \sum_{k=1}^{n} a_k \Rightarrow \sum_{k=1}^{n+1} a_k \leq L$.

Now, using this lemma, with $L = |\nu|(E)$, $a_k = |\nu|(E_k)$, and t_k chosen so that $t_k < |\nu|(E_k)$ as previously, and relying on the result that $\sum_{k=1}^{\infty} t_k \leq |\nu|(E)$, we have that

$$\sum_{k=1}^{\infty} |\nu|(E_k) \leq |\nu|(E).$$

"$\leq$": $\{E_k\} \in \mathcal{P}^*(E, \mathcal{A})$ is given. Then for all $\{A_j\} \in \mathcal{P}^*(E, \mathcal{A})$, $\{A_j \cap E_k; k \in \mathbb{N}\} \in \mathcal{P}^*(A_j, \mathcal{A})$, then

$$\sum_{j\in\mathbb{N}} |\nu(A_j)| = \sum_{j\in\mathbb{N}} |\sum_{k\in\mathbb{N}} \nu(A_j \cap E_k)| \leq \sum_{j\in\mathbb{N}}\sum_{k\in\mathbb{N}} |\nu(A_j \cap E_k)| = \sum_{k\in\mathbb{N}}\sum_{j\in\mathbb{N}} |\nu(A_j \cap E_k)| \leq \sum_{k\in\mathbb{N}} |\nu|(E_k),$$

by $\{A_j \cap E_k; j \in \mathbb{N}\} \in \mathcal{P}^*(E_k, \mathcal{A})$. This was for all $\{A_j\} \in \mathcal{P}^*(E, \mathcal{A})$, so is true for the sup in the definition of $|\nu|$, so

$$\sum_{k=1}^{\infty} |\nu|(E_k) \geq |\nu|(E).$$

So we have that $|\nu| \in \mathcal{M}^+(X, \mathcal{A})$. That $|\nu|(E) \geq |\nu(E)|$ follows by noting that $\{E, \phi\} \in \mathcal{P}^*(E, \mathcal{A})$ so that $|\nu|(E) \geq |\nu(E)| + |\nu(\phi)| = |\nu(E)|$. Suppose $\mu \in \mathcal{M}^+(\mathcal{A})$ was another positive measure satisfying ($\dagger_1$), then for $\{E_k\} \in \mathcal{P}^*(E, \mathcal{A})$ arbitrarily, applying ($\dagger_1$) and summing, $\sum_{k\in\mathbb{N}} |\nu(E_k)| \leq \sum_{k\in\mathbb{N}} \mu(E_k) = \mu(E)$, now by its definition, $|\nu|(E)$ is the sup of numbers of the form on the LHS, and by this inequality, $\mu(E)$ is an upper bound for such numbers, thus $|\nu|(E) \leq \mu(E)$, for all $E \in \mathcal{A}$. $\square$

$\mathcal{M}(X, \mathcal{A})$ is a vector space, with respect to measure addition, $(\nu_1 + \nu_2)(E) = \nu_1(E) + \nu_2(E)$, scaling, $\lambda(\nu)(E) = (\lambda\nu)(E)$, the zero measure, $0(E) = 0$ for all $E \in \mathcal{A}$. The details to this are obvious.

$\nu_1, \nu_2 \in \mathcal{M}(X, \mathcal{A})$, $\lambda \in \mathbb{R}$ or $\mathbb{C}$ then $|\nu_1 + \nu_2| \leq |\nu_1| + |\nu_2|$, $|\lambda\nu_1| = |\lambda|\,|\nu_1|$.

Proof: For all $E \in \mathcal{A}$, $|(\nu_1 + \nu_2)(E)| = |\nu_1(E) + \nu_2(E)| \leq |\nu_1(E)| + |\nu_2(E)| \leq |\nu_1|(E) + |\nu_2|(E) = (|\nu_1| + |\nu_2|)(E)$. Scaling follows by $|\lambda\nu_1(E)| = |\lambda||\nu_1(E)|$, and for any $A \subset \mathbb{R}, a \in \mathbb{R}$, $A \neq \phi, a > 0$ $\sup\{ax : x \in A\} = a \sup A$.

Theorem 6.4 in Rudin: If $\nu \in \mathcal{M}(X, \mathcal{A})$, then $|\nu|(X) < \infty$.

$\mathcal{M}(X, \mathcal{A})$ is a normed space w.r.t. $||\nu|| := \nu(X)$ for all $\nu \in \mathcal{M}(X, \mathcal{A})$.

Proof: $||\nu_1+\nu_2|| = |\nu_1+\nu_2|(X) \leq |\nu_1|(X)+|\nu_2|(X) = ||\nu_1||+||\nu_2||$. $||\lambda\nu|| = |\lambda\nu|(X) = |\lambda||\nu|(X) = |\lambda|||\nu||$. $||\nu||(X) = |\nu|(X) \geq 0$, $||\nu|| = 0 \Rightarrow |\nu|(X) = 0 \Rightarrow 0 = |\nu|(X) \geq |\nu|(E) \geq |\nu(E)|$ for all $E \in \mathcal{A} \Rightarrow \nu = 0$.

$\mathcal{M}(X, \mathcal{A})$ is a complete metric space with respect to the canonical metric induced by the norm: $d(\nu_1, \nu_2) = (\nu_1 - \nu_2)(X)$. Thus $\mathcal{M}(X, \mathcal{A})$ is a Banach space.

Proof: ADD

Jordan decomposition: For all $\nu \in \mathcal{M}^\pm(X, \mathcal{A})$, , define $\nu^+ = \frac{1}{2}(|\nu| + \nu)$, $\nu^- = \frac{1}{2}(|\nu| - \nu)$. Then $\nu^+, \nu^- \in \mathcal{M}^+(X, \mathcal{A})$, $\nu = \nu^+ - \nu^-$, $|\nu| = \nu^+ + \nu^-$. This is the Jordan decomposition of ν, and is unique. Further, if $\nu \in \mathcal{M}(X, \mathcal{A})$, then define $\mathrm{Re}(\nu)(E) = \mathrm{Re}(\nu(E))$, $\mathrm{Im}(\nu)(E) = \mathrm{Im}(\nu(E))$ for all $E \in \mathcal{A}$, then $\mathrm{Re}(\nu), \mathrm{Im}(\nu) \in \mathcal{M}(X, \mathcal{A})$, and so $\nu = \sum_{k=0}^{3} i^k \nu_k$, where each $\nu_k \in \mathcal{M}^+(X, \mathcal{A})$, $i = \sqrt{-1}$.

Proof: ADD.

For all $f : X \to \mathbb{C}$, $\mathcal{A}$-measurable, $\nu \in \mathcal{M}(X, \mathcal{A})$, say that f is ν-integrable if it is $|\nu|$-integrable, so $f \in \mathcal{L}(X, |\nu|)$. Write $\nu_0 = \mathrm{Re}(\nu)^+, \nu_1 = \mathrm{Re}(\nu)^-, \nu_2 = \mathrm{Im}(\nu)^+, \nu_3 = \mathrm{Im}(\nu)^-$, then

$$\int_X f d\nu = \sum_{k=0}^{3} i^k \int_X f \, d\nu_k$$

and $f \in \mathcal{L}(X, |\nu|)$ iff $|f| \in \mathcal{L}(X, |\nu|)$ iff $|f| \in \mathcal{L}(X, |\nu|)$ iff $|f| \in \mathcal{L}(X, \nu_k)$ iff $f \in \mathcal{L}(X, \nu_k)$, for all $k \in \{0, 1, 2, 3\}$.

Proof: ADD

For $\nu \in \mathcal{M}(X, \mathcal{A})$, say that ν is concentrated on $A \in \mathcal{A}$ if $\nu(E) = \nu(A \cap E)$ for all $E \in \mathcal{A}$.

This is equivalent to $\nu(E) = 0$ for all $E \in \mathcal{A}, E \subset A^c$, by $\nu(E) = \nu(E \cap A) + \nu(E \cap A^c) = \nu(E \cap A) \Leftrightarrow \nu(E \cap A^c) = 0$. Not equivalently that $\nu(A^c) = 0$.

If $A, B \in \mathcal{A}$, and ν is concentrated on both A, B, then $\nu(A \setminus B) = \nu(A \cap B^c) = \nu((A \cap B^c) \cap B) = \nu(\phi) = 0$, thus $\nu(A \Delta B) = 0$. So sets on which a measure concentrate differ by at most sets of measure zero.

If $A, T \in \mathcal{A}$, and ν is concentrated on A, $A \cap T = \phi$, then $\nu(T) = 0$, and for all $E \in \mathcal{A}$, $\nu(E) = \nu(E \cap A) + 0 = \nu(E \cap A) + \nu(E \cap T) = \nu(E \cap (A \cup T))$. So if ν is concentrated on A, and $B \in A$ is any other set which contains A, then ν is concentrated on B also.

Then, if $\nu_1, \nu_2 \in \mathcal{M}(\mathcal{A})$, ν_1 concentrated on A_1, ν_2 on A_2, then ν_1 and ν_2 both concentrated on $A_1 \cup A_2$, so for $\nu = \nu_1 + \nu_2$, $\nu(E) = \nu_1(E) + \nu_2(E) = \nu_1(E \cap (A_1 \cup A_2)) + \nu_2(E \cap (A_1 \cup A_2)) = \nu(E \cap (A_1 \cup A_2))$, so ν concentrated on $A_1 \cup A_2$. Clearly sets of concentration don't change when scaling a measure by non-zero scalar.

This all suggests the following construction.

$$\bigcap \{A \in \mathcal{A};\ \nu(E) = \nu(E \cap A) \text{ for all } E \in \mathcal{A}\}$$

Is this set well defined? Need to show that it is in $\mathcal{A}$.

For $\nu \in \mathcal{M}(X, \mathcal{A})$, $\mu \in \mathcal{M}^+(X, \mathcal{A})$, say that ν is absolutely continuous w.r.t. μ if $\mu(E) = 0 \Rightarrow \nu(E) = 0$ for all $E \in \mathcal{A}$, and write $\nu \ll \mu$.

For $\nu_1, \nu_2 \in \mathcal{M}(X, \mathcal{A})$, say ν_1 and ν_2 are mutually singular if they are concentrated on disjoint sets, and write $\nu_1 \perp \nu_2$.

If $\nu \perp \nu$, then for any sets $A, B \in \mathcal{A}$ on which ν concentrates, $A \cap B = \phi$, but $\nu(A \Delta B) = 0$, so $\nu(A \cup B) = 0$, so ν concentrates only on ν-null sets, so $\nu = 0$.

(Rudin 6.8) For $\mu \in \mathcal{M}^+(X, \mathcal{A})$, $\nu, \nu_1, \nu_2 \in \mathcal{M}(X, \mathcal{A})$,

a) ν concentrated on $A \Rightarrow |\nu|$ concentrated on $\mathcal{A}$.
b) $\nu_1 \perp \nu_2 \Rightarrow |\nu_1| \perp |\nu_2|$.
c) $\nu_1 \perp \mu$, $\nu_2 \perp \mu \Rightarrow \nu_1 + \nu_2 \perp \mu$.
d) $\nu_1 \ll \mu, \nu_2 \ll \mu \Rightarrow \nu_1 + \nu_2 \ll \mu$.
e) $\nu \ll \mu \Rightarrow |\nu| \ll \mu$.
f) $\nu_1 \ll \mu, \nu_2 \perp \mu \Rightarrow \nu_1 \perp \nu_2$.
g) $\nu \ll \mu, \nu \perp \mu \Rightarrow \nu = 0$.

a) Let $E \subset A^c$, $\{E_k\} \in \mathcal{P}^*(E, \mathcal{A})$, then $\nu(E_k) = 0$, $\{E_k\}$ is arbitrary, so $|\nu|(E) = 0$.

b) If ν_1 is concentrated on A_1, ν_2 on A_2, then $|\nu_1|$ on A_1, and $|\nu_2|$ on A_2 by (a). The hypothesis is that A_1 and A_2 are disjoint, this is then what is needed for the conclusion.

c) ν_1 concentrated on A_1, μ on B_1, ν_2 on A_2, and μ on B_2, so the hypothesis is that that $A_1 \cap B_1 = \phi$, $A_2 \cap B_2 = \phi$, but $\nu_1 + \nu_2$ is concentrated on $A_1 \cup A_2$, and μ on $B_1 \cup B_2$, and $(A_1 \cup A_2) \cap (B_1 \cup B_2) = (A_1 \cap B_1) \cup (A_2 \cap B_2) = \phi$.

d) If $\mu(E) = 0$, then $\nu_1(E) = 0 = \nu_2(E)$, so $\nu_1(E) + \nu_2(E) = (\nu_1 + \nu_2)(E) = 0$.

e) If $\mu(E) = 0$, then for any $\{E_k\} \in \mathcal{P}^*(E, \mathcal{A})$, $\mu(E_k) = 0$, thus $\nu(E_k) = 0$ for all E_k bu hypothesis, and thus $|\nu|(E) = 0$.

f) Since $\nu_2 \perp \mu$, there is an $A \in \mathcal{A}$ with $\mu(A) = 0$, and ν_2 concentrated on A. $\nu_1(E) = 0$ for all $E \subset A, E \in \mathcal{A}$ by $\mu(E) = 0$, thus ν_1 concentrates on A^c.

g) by (f), $\nu \perp \nu$, so $\nu = 0$.

(Rudin 6.9) For $\mu \in \mathcal{M}^+(X, \mathcal{A})$, σ-finite, then there is a function $w \in L^1(\mu)$ s.t. $w(x) \in (0, 1)$ for all $x \in X$. Thus $\tilde{\mu}(E) := \int_E w\, d\mu \in \mathcal{M}^+(X, \mathcal{A})$, and $\mu(E) = 0 \Leftrightarrow \tilde{\mu}(E) = 0$ for all $E \in \mathcal{A}$ and $\tilde{\mu}(X) \le 1 < \infty$.

Proof: $X = \cup_k E_k, E_k \in \mathcal{A}, \mu(E_k) < \infty$. Let $w_k(x) = \chi_{E_k}(x)\ 2^{-k}/(1 + \mu(E_k))$. Then on $E_k, 1 + \mu(E_k) \in (1, \infty), 1/(1 + \mu(E_k)) \in (0, 1)$. Then $w(x) = \sum_{k=1}^\infty w_k(x) \in (0, 1)$. Each w_k is the product of a measurable characteristic function, and a number, and so is measurable, and then by Beppo-Levi,

$$\tilde{\mu}(X) = \int_X \sum_{k=1}^\infty w_k(x) d\mu(x) = \sum_{k=1}^\infty \int_X w_k(x) d\mu(x) = \sum_{k=1}^\infty \frac{2^{-k}}{(1 + \mu(E_k))} \mu(E_k) \le \sum_{k=1}^\infty 2^{-k} = 1.$$

$$\tilde{\mu}(E) = \sum_{k=1}^\infty \int_X \chi_E(x) \frac{\chi_{E_k}(x) 2^{-k}}{(1 + \mu(E_k))} d\mu(x) = \sum_{k=1}^\infty \frac{\mu(E \cap E_k)}{1 + \mu(E_k)} 2^{-k} = 0 \Leftrightarrow \mu(E) = 0$$

because $\mu(E) = \sum_{k=1}^\infty \mu(E \cap E_k)$.

(Rudin 1.40) $\mu \in \mathcal{M}^+(\mathcal{A})$, $\mu(X) < \infty$, $f : X \to \mathbb{F}$, $f \in L^1(\mu)$, $S \subset \mathbb{F}$, closed, and define the averages

$$A_E(f) = \frac{1}{\mu(E)} \int_E f\, d\mu.$$

If $A_E(f) \in S$ for all $E \in \mathcal{A}$ with $\mu(E) > 0$, then $f(x) \in S$ for μ-a.e. $x \in X$.

Proof: Let $r > 0$ and $z \in \mathbb{F}$ such that $B_{z,r} := \{y \in \mathbb{F}; |z - y| \le r\} \subset S^c$. Let $E = f^{-1}(B_{z,r})$, suppose that $\mu(E) > 0$. Then $|A_E(f) - z| = |A_E(f) - \frac{\mu(E)}{\mu(E)} z| = |\frac{1}{\mu(E)} \int_E f\, d\mu - \frac{1}{\mu(E)} z \int_E 1\, d\mu| = |\frac{1}{\mu(E)} \int_E (f(x) - z)\, d\mu(x)| \le \frac{1}{\mu(E)} \int_E |f(x) - z|\, d\mu(x) | \frac{1}{\mu(E)} \int_E (f(x) - z)\, d\mu(x)| \le \frac{1}{\mu(E)} \int_{\{x \in X; |f(x) - z| \le r\}} |f(x) - z|\, d\mu(x) \le \frac{1}{\mu(E)} \int_E r\, d\mu(x) = r$. So $\mu(E) > 0 \Rightarrow |A_E(f) - z| \le r \Rightarrow A_E(f) \in B_{z,r} \subset S^c$, but the hypothesis is that $A_E(f) \in S$ for all $E \in \mathcal{A}$, thus $\mu(E) = 0$. Since any open set in $\mathbb{F}$ is a union of open balls, and hence a countable union of closed balls like $B_{z,r}$, $\mu(f^{-1}(S^c)) = \mu(f^{-1}(\cup_{k \in \mathbb{N}} B_k)) \le \mu(\sum_{k \in \mathbb{N}} f^{-1}(B_k)) = 0$.

If $(X, \mathcal{A}, \mu)$ a (positive) measure space, $f, g : X \to \mathbb{F} \in L^1(\mu)$. If $\int_E f\, d\mu = \int_E g\, d\mu$ for all $E \in \mathcal{A}$, then $f = g$ $\mu-$a.e., this is also true for measurable $f, g : X \to [0, \infty]$ if μ is σ-finite.

Proof: f, g integrable implies $\int_E f\, d\mu - \int_E g\, d\mu = 0 \Rightarrow \int_E (f - g)\, d\mu = 0$. Let $h = f - g$, then let $A_E = \frac{1}{\mu(E)} \int_E h\, d\mu$, so $A_E = 0$ for all $E \in \mathcal{A}$, and $\{0\}$ is a closed subset of $\mathbb{F}$ (finish)

For measurable $f, g : X \to [0, \infty]$, and if $X = \cup_{k \in \mathbb{N}} X_k$, $\mu(X_k) < \infty$. Fix $U = E_k$, some k, so $\mu(U) < \infty$, let $F = U \cap f^{-1}(\{\infty\})$, then let $F'_n = F \cap g^{-1}([0, n))$. Suppose $F'_n \ne \phi$, then $\int_{F'_n} f\, d\mu = \infty$ because $f = \infty$ everywhere on F'_n, but $\int_{F'_n} g\, d\mu \le n\mu(F') < \infty$, but $\int_{F'_n} f\, d\mu = \int_{F'_n} g\, d\mu$, so $F'_n = \phi$ for all $n \in \mathbb{N}$. Now $(g^{-1}(\{\infty\}))^c = \cup_{n=1}^\infty g^{-1}([0, n))$, $F \cap (g^{-1}(\{\infty\}))^c = \cup_{n=1}^\infty F \cap g^{-1}([0, n)) = \cup_{n=1}^\infty \phi = \phi$, but

$F = F \cap (g^{-1}(\{\infty\}))^c \cup F \cap (g^{-1}(\{\infty\})) = F \cap (g^{-1}(\{\infty\}))$ (finish)

If $(X, \mathcal{A}, \mu)$ a measure space, $[f] \in L^1(\mathcal{A}, \mu)$, then $\nu \in \mathcal{M}(\mathcal{A})$ and $\nu \ll \mu$, where

$$\nu(E) = \int_E f \, d\mu \text{ for all } E \in \mathcal{A}.$$

Proof: First, $|f\mathbf{1}_E| \leq |f| \Rightarrow f\mathbf{1}_E \in L^1(\mathcal{A}, \mu)$ so $|\nu(E)| < \infty$ for all $E \in \mathcal{A}$, if $w \in [f]$ arbitrarily, then $\int_E f \, d\mu = \int_E w \, d\mu$ for all $E \in \mathcal{A}$, because $f = w$ μ-a.e., so ν is independent of the choice of functions in $[f]$, and so ν is finite and well defined. $\nu(\phi) = \int_X \mathbf{1}_\phi f \, d\nu = \int_X 0 \, d\nu = 0$. If $\{E_k\}_{k\in\mathbb{N}} \in \mathcal{A}$, disjoint, $E = \cup_{k\in\mathbb{N}} E_k$, then $\nu(E) = \int_X \mathbf{1}_E f \, d\mu = \int_X (\sum_{k\in\mathbb{N}} \mathbf{1}_{E_k}) f \, d\mu$. Now, let $h_n(x) = \sum_{k=1}^n \mathbf{1}_{E_k}(x) f(x)$, then each h_n is $\mathcal{A}$-measurable, $\lim_{n\to\infty} h_n(x) f(x) = \mathbf{1}_E(x) f(x)$, which is $\mathcal{A}$-measurable, and $|h_n(x)| \leq |f(x)|$, and $|f| \in L^1(\mathcal{A}, \mu)$. So by LDCT, $\nu(E) = \int_X (\sum_{k\in\mathbb{N}} \mathbf{1}_{E_k}) f \, d\mu = \sum_{k\in\mathbb{N}} \int_{E_k} f \, d\mu = \sum_{k\in\mathbb{N}} \nu(E_k)$, so $\nu \in \mathcal{M}(\mathcal{A})$. Next, $f(x) \leq \sup\{f(x); x \in E\}$ for all $x \in E \Rightarrow \nu(E) = \int_E f \, d\mu \leq \sup\{f(x); x \in E\} \int_E d\mu = \sup\{f(x); x \in E\} \mu(E)$, so $\mu(E) = 0 \Rightarrow \nu(E) = 0$, so $\nu \ll \mu$.

Rudin 6.10: $(X, \mathcal{A})$ a measurable space, $\mu \in \mathcal{M}^+(X, \mathcal{A})$, σ-finite.

a) Lebesgue Decomposition Theorem (LDT):

For all $\nu \in \mathcal{M}(X, \mathcal{A})$, there exist unique $\nu_a, \nu_s \in \mathcal{M}(X, \mathcal{A})$ such that $\nu = \nu_a + \nu_s$, $\nu_a \ll \mu$, $\nu_s \perp \mu$, $\nu_a \perp \nu_s$.

b) Radon-Nikodym Theorem (RNT):

For all $\nu \in \mathcal{M}(X, \mathcal{A})$ such that $\nu \ll \mu$, there exists a unique $[h] \in L^1(\mu)$ such that $\nu(E) = \int_E h\, d\mu$, $E \in \mathcal{A}$.

Remarks:

In the LDT, we can apply the RNT to ν_s, and if $\nu \ll \mu$, then $\nu_s = 0$.

In the LDT and RNT, $\nu \geq 0$, then $h, \nu_a, \nu_s \geq 0$ μ-a.e.

In the LDT and RNT, if $\nu \in \mathcal{M}^+(\mathcal{A})$, then these results hold with $\nu_a, \nu_s \in \mathcal{M}^+(X, \mathcal{A})$, but $h : X \to [0, \infty]$, $\mathcal{A}$-measurable, might not be integrable, but will be the (infinite) sum of locally integrable functions.

Notation:

Given $\nu \in \mathcal{M}(X, \mathcal{A})$, write $\frac{d\nu}{d\mu}$ for the μ-integrable function such that $\nu_a(E) = \int_E \frac{d\nu}{d\mu}\, d\mu$, for all $E \in \mathcal{A}$, where ν_a is the part in the LDT such that $\nu_a \ll \nu$, then $d\nu = \frac{d\nu}{d\mu} d\mu$. $\frac{d\nu}{d\mu}$ is called the Radon-Nikodym derivative of ν with respect to μ. More generally, take $d\nu = w\, d\mu$ to mean $\nu(E) = \int_E w\, d\mu$ for all $E \in \mathcal{A}$.

Proof:

Uniqueness: In LDT, if (ν_a', ν_s') another pair of measures from LDT, then $\nu_a' - \nu_a = \nu_s - \nu_s'$, $\nu_a' - \nu_a \ll \mu$, and $\nu_s - \nu_s' \perp \mu$, hence both sides here are 0, by c,d,g from preliminary propositions. In RNT, if $\nu(E) = \int_E h\, d\mu = \int_E h'\, d\mu$, then $\int_E (h - h')\, d\mu = 0$, $E \in \mathcal{A}$, and then by the vanishing principle, $h = h'$ μ-a.e..

Step 1: If $\nu \in \mathcal{M}^+(X, \mathcal{A})$, $\nu(X) < \infty$, then apply Rudin 6.9 to μ to obtain $w \in L^1(\mu), w(x) \in (0, 1)$ for all $x \in X$. Then $\varphi(E) := \nu(E) + \int_E w\, d\mu$ is a positive finite measure on $\mathcal{A}$, and $\varphi \geq \nu$. Then for any $\mathcal{A}$-measurable function $f : X \to [0, \infty]$,

$$\int_X f\, d\varphi = \int_x f\, d\nu + \int_x f w\, d\mu,$$

by following the standrad steps in the construction of the integral. If $f \in L^2(\mu)$,

$$\int_X |f|\, d\nu \leq \int_X |f|\, d\varphi \leq \left(\int_X 1\, d\varphi\right)^{1/2} \left(\int_X |f|^2\, d\varphi\right)^{1/2} = \varphi(X)^{1/2} \left(\int_X |f|^2\, d\varphi\right)^{1/2} < \infty$$

by the Schwarz inequailty, so $f \in L^1(\nu), f \in L^1(\varphi)$, similarly, $fw \in L^1(\mu)$. Thus, $f \mapsto \int_X f\, d\nu$ is a linear functional, bounded (by $\sqrt{\varphi(X)}$) on $L^2(\varphi)$. Hence by $L^2(\varphi)$ being a Hilbert space, and Riesz representation, there exists a $g \in L^2(\varphi)$ so that

$$\int_x f\, d\nu = \int_x f g\, d\varphi, \text{ for all } f \in L^2(\varphi).$$

Then, for $f = \chi_E$, for any $E \in \mathcal{A}$ with $\varphi(E) > 0$, $\lambda(E) = \int_E g\, d\varphi$, and because $0 \leq \lambda \leq \varphi$, $0 \leq \lambda(E)/\varphi(E) \leq \varphi(E)/\varphi(E) = 1$,

$$0 \leq \frac{1}{\varphi(E)} \int_E g\, d\varphi \leq 1, \text{ for all } E \in \mathcal{A}.$$

So by Rudin 1.40, $g \in [0,1]$ $\varphi-$a.e. so wlog, $g(x) \in [0,1]$ for all $x \in X$.

Define $A = g^{-1}([0,1))$, $B = g^{-1}(\{1\})$, then $A, B \in \mathcal{A}$ by $g \in L^2(\varphi) \Rightarrow g$ is $\mathcal{A}$-measurable, $A \cup B = X$, $A \cap B = \phi$. Define $\nu_a(E) = \nu(A \cap E)$, $\nu_s(E) = \nu(B \cap E)$ for all $E \in \mathcal{A}$. Notice $\nu_a(E) + \nu_s(E) = \nu(E \cap A) + \nu(E \cap B) = \nu(E \cap X) = \nu(E)$ for all $E \in \mathcal{A}$, so $\nu = \nu_a + \nu_s$, and by definition, ν_s is concentrated on A, ν_s on B so $\nu_a \perp \nu_s$.

Now, rewriting,

$$\int_X f\, d\nu = \int_X fg\, d\varphi = \int_X fg\, d\nu + \int_X fgw\, d\mu \rightarrow$$
$$\int_X (1-g) f\, d\nu = \int_X fgw\, d\mu.$$

Let $f = \chi_B$, then the LHS is 0, and the RHS is $\int_X w\, d\mu$, and since $w > 0$, $\mu(B) = 0$, so μ is concentrated on $B^c = A$, so that $\nu_s \perp \mu$. Next, let $f = \chi_E \sum_{k=0}^n g^k$, then $f \geq 0$, $f \in L^2(\varphi)$. Then,

$$\int_E (1-g) f\, d\nu = \int_E fgw\, d\mu \rightarrow \quad \int_E (1 - g^{n+1})\, d\nu = \int_E \sum_{k=0}^n g^{k+1} w\, d\mu.$$

Let

$$h(x) = w(x) g(x) \sum_{k=0}^{\infty} g^k(x).$$

For $x \in A$, $g^k(x)$ decreases monotonically, so the partial sums in h increase monotonically, and $h(x) = \frac{g(x)w(x)}{1-g(x)}$. So, taking the limit of the equation, gives by LMCT

$$\lim_{n\to\infty} \int_{E\cap A} (1 - g^{n+1})\, d\nu = \lim_{n\to\infty} \int_{E\cap A} \sum_{k=0}^n g^{k+1} w\, d\mu =$$

$$\int_{E\cap A} \lim_{n\to\infty} (1 - g^{n+1})\, d\nu = \int_{E\cap A} \lim_{n\to\infty} \sum_{k=0}^n g^{k+1} w\, d\mu =$$

$$\int_{E\cap A} 1\, d\nu = \int_{E\cap A} h\, d\mu = \nu(E \cap A) = \nu_a(E).$$

For $x \in B$, $g(x) = g^k(x) = 1$, so $1 - g^{n+1}(x) = 0$, so

$$\int_{E\cap B} h\, d\nu = \int_{E\cap B} 0\, d\nu = 0.$$

not finished.

Folland 1.29: $(X, \mathcal{A}, \mu)$ a σ-finite measure space, for $w : X \to [0, \infty]$, $\mathcal{A}$-measurable and ν as $d\nu = w\,d\mu$, then $\nu \in \mathcal{M}^+(\mathcal{A})$, and for all $f : X \to [0, \infty]$ $\mathcal{A}$-measurable, $\int_X f\,d\nu = \int_X fw\,d\mu$.

Proof: let $\{E_k\} \in \mathcal{P}^*(E, \mathcal{A})$. Notice that $\mathbf{1}_E w = \sum_{kj=1}^{\infty} \mathbf{1}_{E_k} w$, by Beppi-Levi, $\nu(E) = \sum_{k=1}^{\infty} \nu(E_k)$, and $\nu(\phi) = \int_X \mathbf{1}_\phi w\,d\mu = 0$, so $\nu \in \mathcal{M}^+(\mathcal{A})$. Then letting $f = \mathbf{1}_E, E \in \mathcal{A}$ the formula holds, thus for simple functions, and thus for general $f : X \to [0, \infty]$ using LMCT.

Folland 6.13: for $\nu \in \mathcal{M}(\mathcal{A})$, $\mu \in \mathcal{M}^+(\mathcal{A})$ σ-finite, and $d\nu = w\,d\mu$, then $d|\nu| = |w|\,d\mu$, or in other words, $\frac{d|\nu|}{d\mu} = |\frac{d\nu}{d\mu}|$.

Proof: Need to show that $|\nu|(E) = \int_E |w|\,d\mu$ for all $E \in \mathcal{A}$. Notice $|\nu(E)| = |\int_E w\,d\mu| \le \int_E |w|\,d\mu$. So by definition, $|\nu| \le |\lambda|$, with λ defined by $d\lambda = |w|\,d\mu$. For the reverse inequality, let $A = \{x \in X; w(x) \neq 0\}$, then $A \in \mathcal{A}$ because $[w] \in L^1(\mu)$ so w is $\mathcal{A}$-measurable. Define $K(x) = |w(x)| \div w(x)$, for $x \in \mathcal{A}$, and $K(x) = 0$ else, then K is measruable, beasue $|w(x)| \div w(x)$ is measurable w.r.t. $\mathcal{A}_A$, . Then $|K| \le 1$ and so there exists a sequence of simple functions $s_n \nearrow K$ pointwise, $|s_n| \le |K| \le 1$. Then $s_n(x)w(x) \to K(x)w(x) = |w(x)|$. By LDCT, $\int_E s_n h\,d\mu \to \int_E |w|\,d\mu$. Suppose $s_n = \sum_{k=1}^m c_k \mathbf{1}_{E_k}$ in its standard representation, then $|c_k| \le 1$ and

$$|\int_E s_n w\,d\mu| = |\sum_{k=1}^m c_k \int_{E \cap E_k} w\,d\mu| \le \sum_{k=1}^m |c_k||\nu(E \cap E_k)| \le \sum_{k=1}^m 1 \cdot ||\nu(E \cap E_k)| = |\nu|(E \cap (\cup_k E_k)) \le |\nu|(E),$$

so $\int_E |w|\,d\mu \le |\nu|(E)$. $\square$

For $\nu \in \mathcal{M}(\mathcal{A})$, $\mu \in \mathcal{M}^+(\mathcal{A})$ σ-finite, and $d\nu = w\,d\mu$, then $d\nu_k = w_k\,d\mu_k$, where $\nu = \sum_{k=0}^3 i^k \nu_k$, $w = \sum_{k=0}^3 i^k w_k$, and each $\nu_k, w_k \ge 0$.

$f^\pm = \frac{1}{2}f \pm \frac{1}{2}f$, $\operatorname{Re} f = \frac{1}{2}f + \frac{1}{2}\overline{f}$, $\operatorname{Im} f = \frac{1}{2}f - \frac{1}{2}\overline{f}$. $f_0 = \operatorname{Re} f^+$, $f_1 = \operatorname{Im} f^+$, $f_2 = \operatorname{Re} f^-$, $f_3 = \operatorname{Im} f^-$.

Proof: The Jordan decomposition of ν and the decomposition of w into positive, negative, real and imaginary parts take the same form, then use the previous theorem and the linearity of the integral.

For $\nu \in \mathcal{M}(\mathcal{A})$, $\mu \in \mathcal{M}^+(\mathcal{A})$ σ-finite, then $d\nu = w\,d\mu \Leftrightarrow \int_x f\,d\nu = \int_X fw\,d\mu$ for all $f \in L^1(X, |\nu|)$.

Proof: ($\Leftarrow$) For $E \in \mathcal{A}$, $f = \mathbf{1}_E$, then $\int_E 1 d\nu = \int_E w\,d\mu = \nu(E)$. ($\Rightarrow$) follows by decomposition into parts and linearity and defninition of the integral.

Folland 6.12: for $\nu \in \mathcal{M}(\mathcal{A})$ there is a $[w] \in L^1(\mathcal{A}, |\nu|)$ such that $d\nu = w\,d|\nu|$, with $|w| = 1$.

Proof: ADD

Using these results, for any $\nu \in \mathcal{M}(\mathcal{A})$, and w as $d\nu = w\,d|\nu|$, then for all $f \in L^1(\mathcal{A}, |\nu|)$, $\int_X f\,d\nu = \int_X fw\,d|\nu|$. This may be taking as the definition of $\int_X f\,d\nu$, by integrating fw in the usual way, i.e. $|fw| = |f|$ so $fw \in L^1(\mathcal{A}, |\nu|)$, then apply the usual steps.

Hahn decomposition: If $\nu \in \mathcal{M}^\pm(X, \mathcal{A})$, then there exist disjoint $E^+, E^- \in \mathcal{A}, X = E^+ \cup E^-$, so that ν^+ is concentrated on E^+ and ν^- on E^-, thus $\nu^+ \perp \nu^-$.

Proof: There exists $w : X \to \{-1, 1\}$ such that $\nu(E) = \int_E w\,d|\nu|$ for all $E \in \mathcal{A}$; let $E^\pm = w^{-1}(\{\pm\})$, then $E^+ \cap E^- = 0$ and $E^+ \cup E^- = X$ by properties of inverse images. Then $\nu(A \cap E^\pm) = \int_{A \cap E^\pm} w\,d|\nu| =$

Duality o L^p spaces:

Let $1 \leq p < \infty$ and let q be the Hölder conjugate index of p. So, $\frac{1}{p} + \frac{1}{q} = 1$, $q + p = pq$, $q = pq - p$, $p = q/(q-1) = \frac{1}{1-\frac{1}{p}}$, and all these are hold when switching p and q.

There is a canonical map:

$$\Phi : L^q(X,\mu) \to (L^p(X,\mu))^*,\ \Phi := g \mapsto \left(f \mapsto \int_X fg\, d\mu \right) \text{ for all } g \in L^q(X,\mu), f \in L^p(X,\mu).$$

Φ is linear follows by linearity fo the intregral, and by Hölder's inequality, $|\Phi(g)(f)| = |\int_X fg\, d\mu| \leq \int_X |fg|\, d\mu \leq ||f||_p ||g||_q \leq \infty$, so that $||\Phi(g)|| \leq ||g||_q$ for all $g \in L^q(X,\mu)$ so Φ is a bounded linear functional on $L^q(X,\mu)$.

If $1 < p < \infty$ then this Φ is an isometry, i.e. $||\Phi(g)|| = ||g||_q$ for all $g \in L^q(X,\mu)$. If $p = 1$ then Φ is an isometry if X is σ-finite.

Proof: for the first case, $p > 1$, given $g \in L^q$ let $f = \overline{\operatorname{sgn}(g)}|g|^{q-1}$, so $f(x) = \overline{g(x)}|g(x)|^{q-2}$ if $x \neq 0$, $f(x) = 0$ otherwise, f is measure by usual tricks. Then $|f(x)|^p = |g(x)|^p |g(x)|^{pq-2p} = |g(x)|^{pq-p} = |g|^q$, so $g \in L^q(x,\mu) \Rightarrow \int_X |g|^q\, d\mu < \infty \Rightarrow \int_X |f|^p\, d\mu < \infty \Rightarrow f \in L^p(X,\mu)$.

$\nu \in \mathcal{M}^{\pm}(X,\mathcal{A})$, then by the Hahn decomposition gives $A_+, A_- \in \mathcal{A}$, $A_+ \cap A_- = \phi, A_+ \cup A_- = X$, $\nu_{\pm}(E) = \frac{1}{2}(|\nu|(E) \pm \nu(E)) = \pm\nu(E \cap A_{\pm})$, for all $E \in \mathcal{A}$, and $\nu_{\pm}$, and then $\nu_+ \perp \nu_-$. Then $\nu_{\pm}$ are unique positive finite measures on $\mathcal{A}$.

Proof: $\nu_{\pm}$ are signed measures, need to show that they are positive. Write $w = \frac{d\nu}{d|\nu|}$, then $|w| = 1$, $w : X \to \{-1, 1\}$, and $A_{\pm} = w^{-1}(\{\pm 1\})$, then $\nu_{\pm}(E) = \frac{|\nu|(E) \pm \nu(E)}{2} = \int_E \frac{|h| \pm h}{2}\, d|\nu| = \int_E h^{\pm}\, d|\nu| \geq 0$ for all $E \in \mathcal{A}$ (integral of a positive functions wrt a positive measure). Uniqueness follows by supposing that $\tilde{A}_+, \tilde{A}_- \in \mathcal{A}$ another such pair, then $\nu(A_+ \cap \tilde{A}_-) = \nu_+(\tilde{A}_-)$, but ν_+ is concentrated on $\tilde{A}_+$ so $\nu(A_+ \cap \tilde{A}_-) = 0$, similarly, $\nu(A_- \cap \tilde{A}_+) = 0$.

Counter example to show that in the RNT, the positive measure needs to be σ-finite. Let μ be the counting measure on $(\mathbb{R}, \mathcal{P}(\mathbb{R}))$, then $L^1(\mathbb{R}, \mu) = \ell^1(\mathbb{R})$, and for any $\nu \in \mathcal{M}(\mathbb{R})$, E ν-measurable, $\mu(E) = 0 \Rightarrow \sum_{x \in E} 1 = 0 \Rightarrow E = \phi \Rightarrow \nu(E) = 0$, so $\nu \ll \mu$. Then, let $(\mathbb{R}, \overline{\mathcal{L}}, \overline{\lambda})$ be the complete Lebesgue measure space on $\mathbb{R}$. For all $h \in \ell^1(\mathbb{R})$, let $\nu_h(E) = \int_E h\, d\nu$. Now we've seen that $h(x) = 0$ for all but countably many $x \in \mathbb{R}$, which means that ν_h is concentrated on a countable set in $\mathbb{R}$, A_h, and $A_h \in \overline{\mathcal{L}}$ by completeness, with $\overline{\lambda}(A_h) = 0$. Finally, the RNT would say that for any $\gamma \in \mathcal{M}(\mathbb{R}, \overline{\mathcal{L}})$, there exists an $h \in \ell^1(\mathbb{R})$ such that $\gamma(E) = \nu_h(E)$ for all $E \in \overline{\mathcal{L}}$, but then there exists an $A \subset \mathbb{R}$ such that γ is concentrated on A; choose $\gamma(E) = \int_E w\, d\overline{\lambda}$ for all $E \in \overline{\mathcal{L}}$, some $[w] \in L^1(\overline{\lambda})$, so γ is concetrated on a $\overline{\lambda}$-null set, because that set is countable, and $\gamma \ll \overline{\lambda}$, so γ is concentrated on a γ-null set, so γ=0, a contratdiction for any $[w] \neq 0$.

Let $\mathcal{M}_{\mu-a.c.} = \{\nu \in \mathcal{M}(\mathcal{A}); \nu \ll \mu\}$, $\mathcal{M}_{\mu-a.c.}$ is a subspace of the Banach space $\mathcal{M}(\mathcal{A})$.

Let $0(E) = 0$ for all $E \in \mathcal{A}$, then $0 \ll \mu$ trivially, so $0 \in \mathcal{M}_{\mu-a.c.}$. Let $\nu, \sigma \in \mathcal{M}_{\mu-a.c.}$, $c \in \mathbb{F}$, then $\lambda := \nu + c \cdot \sigma \in \mathcal{M}(\mathcal{A})$, and suppose $E \in \mathcal{A}$, $\mu(E) = 0$, then $\nu(E) = 0 = \sigma(0) = c \cdot \sigma(0)$, so $\lambda(E) = 0$ and so $\lambda \in \mathcal{M}_{\mu-a.c.}$. $\mathcal{M}_{\mu-a.c.}$ inherits its norm from $\mathcal{M}(\mathcal{A})$. Suppose $(\nu_k)_{k \in \mathbb{N}}$ is sequence in $\mathcal{M}_{\mu-a.c.}$, then $\nu = \lim_{k \to \infty} \nu_k \in \mathcal{M}(\mathcal{A})$, because $\mathcal{M}(\mathcal{A})$ is a Banach space. Suppose $E \in \mathcal{A}$, $\mu(E) = 0$, then each $\nu_k \in \mathcal{M}_{\mu-a.c.} \Rightarrow \nu_k(0) = 0$, and then $\nu(E) = \lim 0 = 0$, so $\nu \in \mathcal{M}_{\mu-a.c.}$, and so $\mathcal{M}_{\mu-a.c.}$ is a complete metric space, and thus a Banach subspace.

(X, A, μ) a measure space, then the map $\Phi : L^1(\mathcal{A}, \mu) \to \mathcal{M}_{\mu-a.c.}$ defined by $\Phi([w])(E) = \int_E w\, d\mu$, is linear, bijective, and $||\Phi([w])|| = ||[w]||$. So $\mathcal{M}_{\mu-a.c.}$ and $L^1(\mathcal{A}, \mu)$ are isometrically isomorphic. The linearity of Φ does not depend of μ being σ-finite.

Proof: First, for any $[w] \in L^1(\mathcal{A}, \mu)$, then for any $E \in \mathcal{A}$, let $\nu(E) = \int_E w\, d\mu$. We have already shown that ν is well defined, and $\nu \in \mathcal{M}(\mathcal{A})$, so Φ is well defined.

Lemma: For $(X, \mathcal{A}, \mu)$ a measure space, $f, g : X \to \mathbb{R}$, $\mathcal{A}$-measurable. Then $\int_A f\, d\mu = \int_A g\, d\mu$ for all $A \in \mathcal{A} \Rightarrow f = g$, $\mu-$a.e. Proof: let $h = f - g$, $A^+ = h^{-1}([0, \infty])$, $A^- = h^{-1}([-\infty, 0))$. Then $h \geq 0$ on A^+ so $\int_{A^+} h\, d\mu = 0 \Rightarrow h = 0$ μ-a.e. on A^+ by the vanishing principle. Similarly, $h^- = 0$ μ-a.e. on A^-.

Next, $\Phi([f]) = \Phi([g]) \Rightarrow \int_E f\, d\mu = \int_E g\, d\mu$ for all $E \in \mathcal{A} \Rightarrow f = g \;\; \mu-$a.e. $\Rightarrow f, g \in [g] = [f]$, so Φ is injective, using the lemma. If $\nu \in \mathcal{M}_{\mu-a.c.}$, then the RNT says that there is a unique $[h] \in L^1(\mathcal{A}, \mu)$ s.t. $\nu(E) = \int_E h\, d\mu$ for all $E \in \mathcal{A}$, hence Φ is surjective.

Linearity: for all $f, g \in L^1(\mathcal{A}, \mu), c \in \mathbb{F}$, $\Phi([f] + c \cdot [g])(E) = \int_E (f + c \cdot g)\; d\mu = \int_E f d\mu + c \cdot \int_E g\, d\mu \;\; \forall E \in \mathcal{A}$, and $\Phi([f]) + c \cdot \Phi([g]) = \int_E f d\mu + c \cdot \int_E g\, d\mu \;\; \forall E \in \mathcal{A}$, so $\Phi([f] + c \cdot [g])(E) = \Phi([f]) + c \cdot \Phi([g])$.

Lastly, let $\nu = \Phi([f])$, then $||\Phi([f])|| = |\nu|(X) = \int_X |f|\, d\nu = ||f||_1$, by Rudin 6.13. $\square$

(X, A) a measurable space, $\nu, \sigma, \mu \in \mathcal{M}(\mathcal{A})^+$, σ-finite.

1. $\nu \ll \mu, f : X \to [0, \infty]$ $\mathcal{A}$-measurable, then $\int_X f\, d\nu = \int_X f \frac{d\nu}{d\mu}\, d\mu$.
2. $\nu \ll \mu, \sigma \ll \mu$ then $\frac{d}{d\mu}(\nu + \sigma) = \frac{d\nu}{d\mu} + \frac{d\nu}{d\mu}$, μ-a.e.
3. $\nu \ll \sigma \ll \mu$ then $\frac{d\nu}{d\sigma}\frac{d\sigma}{d\mu} = \frac{d\nu}{d\mu}$ μ-a.e.
4. $\nu \ll \mu$, $\mu \ll \nu$ then $\frac{d\nu}{d\mu} = \left(\frac{d\mu}{d\nu}\right)^{-1}$

Proof: 1) The RNT for $\mathcal{M}^+$ says that $\frac{d\nu}{d\mu}$ exists and $\frac{d\nu}{d\mu} : X \to [0, \infty]$, is $\mathcal{A}$-measurable, but might not be integrable wrt μ, then Folland 1.29 applies to give the result.

Proof: 2) Let $m = \nu + \sigma$, we showed that $\mathcal{M}_{\mu-a.c.}$ is a linear subspace of $\mathcal{M}(\mathcal{A})$, so $m \ll \mu$. Then the RNT for $\mathcal{M}^+$ says that $\frac{dm}{d\mu}$ exists and $\frac{dm}{d\mu} : X \to [0, \infty]$, is $\mathcal{A}$-measurable, but might not be integrable wrt μ, similarly for $\frac{d\nu}{d\mu}$ and $\frac{d\sigma}{d\mu}$, and by definition $\nu(E) + \sigma(E) = m(E) = \int_E \frac{dm}{d\mu} d\mu = \int_E \frac{d\nu}{d\mu} d\mu + \int_E \frac{d\sigma}{d\mu} d\mu = \int_E \frac{d\nu}{d\mu} + \frac{d\sigma}{d\mu} d\mu$ for all $E \in \mathcal{A}$, and as usual, $\int_E \frac{dm}{d\mu}\, d\,u = \int_E \frac{d\nu}{d\mu} + \frac{d\sigma}{d\mu} d\mu$ for all $E \in \mathcal{A} \Rightarrow \frac{dm}{d\mu} = \frac{d\nu}{d\mu} + \frac{d\sigma}{d\mu}$ $\mu-$a.e..

Proof: 3) he RNT for $\mathcal{M}^+$ says that $\frac{d\nu}{d\mu}, \frac{d\nu}{d\sigma}, \frac{d\sigma}{d\mu}$ exist and $\frac{d\nu}{d\mu}, \frac{d\nu}{d\sigma}, \frac{d\sigma}{d\mu} : X \to [0, \infty]$, are $\mathcal{A}$-measurable. Then applying (1) to $\mathbf{1}_E \frac{d\nu}{d\sigma}$ for all $E \in \mathcal{A}$, gets $\int_E \frac{d\nu}{d\sigma}\, d\sigma = \int_E \frac{d\nu}{d\sigma}\frac{d\sigma}{d\mu}\, d\mu$ for all $E \in \mathcal{A}$, so again $\frac{d\nu}{d\sigma}\frac{d\sigma}{d\mu} = \frac{d\nu}{d\sigma}$ μ-a.e..

Proof: 4) $\mu \ll \mu, \nu \ll \nu$, and $\frac{d\mu}{d\mu} = 1 = \frac{d\nu}{d\nu}$ μ, ν-a.e., because for all $E \in \mathcal{A}$, $\mu(E) = \int_E d\mu = \int_E \frac{d\mu}{d\mu}\, d\mu \Rightarrow \frac{d\mu}{d\mu} = 1$ $\mu-$a.e., similarly for ν. Let $A = \{x \in X; \frac{d\nu}{d\mu}(x) = 0\}$, then $\nu(A) = \int_A 0\, d\mu = 0 \Rightarrow \mu(A) = 0 = \int_A \frac{d\mu}{d\nu}\, d\nu \Rightarrow \frac{d\mu}{d\nu} = 0$ ν-a.e. on A, but $\nu-$a.e. on a ν-null set means everywhere, and similarly, $B = \{x \in X; \frac{d\mu}{d\nu}(x) = 0\}$, $\mu(B) = \nu(B) = 0$. So $\frac{d\nu}{d\mu}, \frac{d\mu}{d\nu} > 0$ a.e. Then, let $\{E_k\}_{k \in \mathbb{N}} \in \mathcal{A}$, $X = \cup_k E_k$, $\nu(E_k) < \infty, \mu(E_k) < \infty$, possible by both ν, μ σ-finite. Then let $\nu_k(E) = \nu(E \cap E_k)$, $\mu_k(E) = \mu(E \cap E_k)$, then $\mu_k(E_k) = \int_{E_k} \frac{d\mu_k}{d\nu_k}\, d\nu_k < \infty \Rightarrow \frac{d\mu_k}{d\nu_k} < \infty$ ν_k-a.e. and so μ_k-a.e. by the finiteness principle, similarly for $\frac{d\nu_k}{d\mu_k}$. So on E_k, $\left(\frac{d\mu_k}{d\nu_k}\right)^{-1}$ and $\left(\frac{d\mu_k}{d\nu_k}\right)^{-1}$ are well defined, i.e., take either to be 1 on null sets where they are ill-defined, and these won't affect integrals, and hence the associted measures; trying to avoid $1/0$ and $1/\infty$. Then, by (3), $\nu \ll \mu \ll \nu \Rightarrow \frac{d\nu}{d\nu} = \frac{d\nu}{d\mu}\frac{d\mu}{d\nu} = 1$ a.e., and dividing gives $\frac{d\mu}{d\nu} = \left(\frac{d\nu}{d\mu}\right)^{-1}$.

Notation:

A measure defined on a Borel σ-algebra is called a Borel measure. Write $\mathcal{M}(\mathbb{R}^n)$ for the Borel measures on $\mathbb{R}^n$. If $E \subset \mathbb{R}^n$, then $\mathcal{M}(E)$ are the borel measures on E, but $\mathcal{M}(E) \subset \mathcal{M}(\mathbb{R}^n)$, by zero padding.

Definition: a positive finite measure μ on $\mathbb{R}^n$ is regular if

1. $\mu(E) = \inf\{\mu(U); U \text{ open}, U \supset E\}$
2. $\mu(E) = \sup\{\mu(K); K \text{ compact}, K \subset E\}$

Every positive finite measure on $\mathbb{R}^n$ is regular. For every measure $\nu \in \mathcal{M}(\mathbb{R}^n)$, for all $E \in \mathcal{B}\ell(\mathbb{R}^n)$, there exists a sequence of open sets $U_k \supset E$, and compact sets $K_n \subset E$ such that $\nu(U_n) \to \nu(E)$ and $\nu(K_n) \to \nu(E)$.

Proof: ADD (add after looking at some topology, and reviewing metric space stuff)

Definition: we say a measure $\nu \in \mathcal{M}(\mathbb{R}^n)$ is regular if each positive ν_k, in the Jordan decomposition $\nu = \sum_{k=0}^{3} i^k \nu_k$, is regular. By the last result, every $\nu \in \mathcal{M}(\mathbb{R}^n)$ is regular and then the condition about sequences of sets holds.

If $\nu \in \mathcal{M}(\mathbb{R})$, we define its distribution function by $F_\nu(x) = \nu((\infty, x])$. $\nu \mapsto F_\nu$ is injective and linear on $\mathcal{M}(\mathbb{R})$.

Proof: Linearity is easy, to see that $\nu \mapsto F_\nu$ is surjective, suppose $F_\nu = 0$, this means that $F_\nu(x) = 0$ for all $x \in \mathbb{R}$, so $\nu((-\infty, x]) = 0$ for all x, so $\nu((a, b]) = \nu((-\infty, b] \setminus (-\infty, a]) = 0 - 0 = 0$. Then because every open set in $\mathbb{R}$ is a countable union of disjoint open invervals, and open intervals can written as a countable union of disjoint half open intervals, every open set in $\mathbb{R}$ has ν-measure zero, and then by Jordan decomposition, ν can be written as a sum of regular measures, which are then all zero measures, so ν is the zero measure. Thus $\text{Ker}(\nu \mapsto F_\nu) = \{\ 0\ \}$, and so $\nu \mapsto F_\nu$ is surjective.

$F : \mathbb{R} \to \mathbb{R}$ is of *bounded variation*, BV, or say $F \in BV$, if $\text{Var}(F) < \infty$, where $\text{Var}(F) := \sup\{V_F(x); x \in \mathbb{R}\}$, and $V_F(x)$ is the *total variation function* of F,

$$V_F(x) := \sup\left\{\sum_{k=1}^{n} |F(x_k) - F(x_{k-1})|;\ x_0 < x_1 < ... < x_n = x, (x_k) \in \mathbb{R}, n \in \mathbb{N}\right\}.$$

Def: $F : \mathbb{R} \to \mathbb{R}$ is in NBV if $F \in BV$, F is right continuous at all $x \in \mathbb{R}$, and $\lim_{x\to-\infty} F(x) = 0$; normalized BV.

If $\nu \in \mathcal{M}(\mathbb{R}, \mathcal{B}\ell(\mathbb{R}))$, then $F_\nu \in NBV$.

Proof: Let $x_0 < x_1 < ... < x_n = x \in \mathbb{R}$, then $\sum_{k=1}^{n} |F_\nu(x_k) - F_\nu(x_{k-1})| = \sum_{k=1}^{n} |\nu([x_{k-1}, x_k))| \leq \sum_{k=1}^{n} |\nu|([x_{k-1}, x_k)) = |\nu|([x_0, x_n)) \leq |\nu|(\mathbb{R}) = ||\nu|| < \infty$.

$$V_{F_\nu}(x) = \sup\left\{\sum_{k=1}^{n} |\nu((x_{k-1}, x_k])|;\ x_0 < x_1 < ... < x_n = x, (x_k) \in \mathbb{R}, n \in \mathbb{N}\right\} \leq$$

$$\sup\left\{\sum_{k=1}^{n}|\nu|((x_{k-1},x_k]);\ x_0<x_1<...<x_n=x,(x_k)\in\mathbb{R},n\in\mathbb{N}\right\}=V_{F_{|\nu|}}(x)\leq$$

$$\sup\left\{|\nu((-\infty,x_0])|+\sum_{k=1}^{n}|\nu|((x_{k-1},x_k]);\ x_0<x_1<...<x_n=x,(x_k)\in\mathbb{R},n\in\mathbb{N}\right\}\leq|\nu|((-\infty,x])$$

so $V_{F_\nu}(x)\leq V_{F_{|\nu|}}(x)\leq|\nu|((-\infty,x])\leq||\nu||<\infty$. So $V_{F_\nu}(x)$ is bounded by a constant in x, $||\nu||$, so $\sup(\{V_{F_\nu}(x);x\in\mathbb{R}\})\leq||\nu||$, so $\mathrm{Var}F_\nu|\leq||\nu||$, so $F_\nu\in BV$.

For $x\in[-\infty,\infty)$, pick $x_n\searrow x$, take Jordan decomposition of ν, to get $\nu=\nu_+-\nu_-$, with $\nu_\pm$ positive finite measures, then $\nu_\pm((-\infty,x_n])\to\nu_\pm(\cap_{k\in\mathbb{N}}(-\infty,x_k])=\nu_\pm((-\infty,x])$, by continuity from above, so $\nu_\pm((-\infty,x_n])\to\nu((-\infty,x_n])$. Thus $\lim_n F_\nu(x_n)=F_\nu(x)$, so $F_\nu(-\infty)=0$, and F_ν is right continuous by definition.

(Folland 3.28) If $F\in BV$ then $\lim_{x\to-\infty}V_F(x)=0$ and $F\in BV\Rightarrow V_F\in NBV$.
Proof: ADD

Properties of BV,

1) If $F,G:\mathbb{R}\to\mathbb{R}$, $c\in\mathbb{R}$, then $V_{F+G}(x)\leq V_F(x)+V_F(x)$ and $V_{cF}(x)=|c|V_F(x)$. Hence BV is a vector space and if $F,G\in BV$, then $\mathrm{Var}(F+G)\leq\mathrm{Var}(F)+\mathrm{Var}(G)$ and $\mathrm{Var}(cF)=|c|\mathrm{Var}(F)$. NBV is a subspace of BV.
2) If $F\in BV$, then $V_F(x)$ is an increasing function of x, bounded above by $\mathrm{Var}(F)$.
3) a) Moreover, if $x<y$, then $V_F(y)-V_F(x)=\sup\left(\{\sum_{k=1}^n|F(x_k)-F(x_{k-1})|;x\leq x_0<x_1<...<x_n=y\}\right)$.
   b) special case: $|F(y)-F(x)|\leq V_F(y)-V_F(x)\leq V_F(y)\leq\mathrm{Var}(F)$.
   c) consequence: $F\in BV\Rightarrow F$ is bounded.
4) An increasing $F:\mathbb{R}\to\mathbb{R}$ is in BV iff F is bounded.
5) $F:\mathbb{R}\to\mathbb{R}\in BV$ iff $F=F_1-F_2$, where $F_1,F_2:\mathbb{R}\to\mathbb{R}$ are bdd and increasing.
6) $F:\mathbb{R}\to\mathbb{C}\in BV$ iff $\mathrm{Re}\,F,\mathrm{Im}\,F\in BV$.
7) $F\in BV\Rightarrow F$ continuous except at countable many points, and for all $x\in\mathbb{R}$, $F(x+)=\lim_{t\to x^+}F(t)$ and $F(x-)=\lim_{t\to x^-}F(t)$, and $\lim_{x\to+\infty}F(x)$ and $\lim_{x\to-\infty}F(x)$ all exist and are in $\mathbb{R}$.
8) $F\in BV\Leftrightarrow F=F_1-F_2+iF_3-iF_4$, where $F_k:\mathbb{R}\to\mathbb{R}$, increasing, bounded, right continuous, and $\lim_{x\to-\infty}F_k(x)=0$ for all k.

Proof: ADD

The linear map $T=\nu\mapsto F_\nu$ from $\mathcal{M}(\mathbb{R})$ to NBV is an isomorphism. Thus it is bijective and $\mathrm{Var}(F_\nu)=||\nu||$ for all $\nu\in\mathcal{M}(\mathbb{R})$, which implies that NBV is a Banach space with norm $||F||=\mathrm{Var}(F)$, $||T(\nu)||=||\nu||$. This also applies to $\mathcal{M}([a,b])$ and $NBV([a,b]()$, by zero padding $F(x)$ and replacing ν by $\nu_{[a,b]}(E):=\nu(E\cap[a,b])$.

Proof: ADD

We say $F : \mathbb{R} \to \mathbb{R}$ is *absolutely continuous*, or say $F \in AC$ if given $\epsilon > 0$, there exists a $\delta > 0$ such that $a_1 < b_1 < a_2 < b_2 < ... < a_n < b_n$ and $\sum_{k=1}^n (b_k - a_k) < \delta \Rightarrow \sum_{k=1}^n |F(b_k) - F(a_k)| < \epsilon$. If $n = 1$ then this is uniformly continuous, so $F \in AC \Rightarrow F$ is uniformly continuous, also $AC \subset BV$. Define $NAC := NBV \cap AC$. Again this can apply to $F : [a, b] \to \mathbb{R}$.

Proof: ADD

$F \in NAC \Leftrightarrow \mu_F \ll \lambda$, where μ_F is the Lebesgue-Stieltjes measure from F, and λ is the Lebesgue measure.

Proof: ADD

(a Vitali covering lemma) Suppose $W \subset \mathbb{R}^k$, $W \subset \cup_{i=1}^n B(x_i, r_i)$, where $B(x, r)$ is the ball centered at $x \in \mathbb{R}^k$, with readius $r > 0$, then there exists $S \subset \{1, 2, ..., n\}$ such that:

a) $B(x_i, r_i) \cap B(x_j, r_j) = \phi$ if $i, j \in S$, $i \neq j$.

b) $W \subset \cup_{i \in S} B(x_i, 3r_i)$

c) $\lambda(w) \leq 3^k \sum_{i \in S} \lambda(B(x_i, r_i))$

Proof: ADD

For $\mu \in \mathcal{M}(\mathbb{R}^k)$, $x \in \mathbb{R}^k$, $r > 0$, define $(Q_r\mu)(x) = \frac{\mu(B(x,r)}{\lambda(B(x,r))}$. Call $M_\mu(x) := \sup\{(Q_r|\mu|)(x); 0 < r < \infty\}$, the *maximal function* of μ, $M_\mu : \mathbb{R}^k \to [0, \infty]$. A special case, for $F \in L^1(\mathbb{R}^k, \lambda)$, $\mu(E) := \int_E F \, d\lambda$, in this case write M_F for M_μ.

$F : \Omega \to [-\infty, \infty]$ is called *lower semi continuous* (lsc) if $F^{-1}((t, \infty])$ is open for all $t \in \mathbb{R}$, this makes sense if Ω is any topological space.

$\mu \in \mathcal{M}(\mathbb{R}^k) \Rightarrow M_\mu$ is lower semi continuous.
Proof: ADD

(Hardy Littlewood theorem) If $\mu \in \mathcal{M}(\mathbb{R}^k)$, $a < t < \infty$ then $\lambda(\{x \in \mathbb{R}^k; M_\mu(x) > t\}) \leq 3^k ||\mu|| \div t$.
Proof: ADD

A function $f : \mathbb{R}^k \to \mathbb{F}$ is called *locally integralbe*, or $f \in L^1_{loc.}(\mathbb{R}^k, \lambda)$ if $F|_K \in L^1(K, \lambda)$ for all compact $K \subset \mathbb{R}^k$.

If $f \in L^1_{loc.}(\mathbb{R}^k, \lambda)$, $x \in \mathbb{R}^k$ is called a *lebesque point* for f if

$$\lim_{r \to 0} \frac{1}{\lambda(B(x, r))} \int_{B(x,r)} |f(y) - f(x)| \, d\lambda(y) = 0.$$

If x is a Lebesgue point for f then

$$f(x) = \lim_{r \to 0} \frac{1}{\lambda(B(x, r))} \int_{B(x,r)} f(y) \, d\lambda(y).$$

Proof: ADD

For $\mu \in \mathcal{M}(\mathbb{R}^k)$, define the *symmetric derivative* of μ as $D_\mu(x) = \lim_{r\to 0} \frac{\mu(B(x,r))}{\lambda(B(x,r))}$, wherever this limit exists, $x \in \mathbb{R}^k$.

Define

$$f^*(x) = \limsup_{r\to 0} \frac{1}{\lambda(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\, d\lambda(y).$$

then

1) $(f+g)^* \leq \int f^* + g^*$ for all $f, g \in L^1(\mathbb{R}^k, \lambda)$

2) If g is continuous at x then $g^*(x) = 0$.

3) If $f, g \in L^1(\mathbb{R}^k, \lambda)$, g continuous then $f^* = (f^* - g + g) \leq (f-g)^* + g^* = (f-g)^*$

4) If $f \in L^1(\mathbb{R}^k, \lambda)$ then $f^* \leq |f| + M_f$.

Proof: ADD

(Lebesue's theorem)

a) If $f \in L^1_{loc.}(\mathbb{R}^k, \lambda)$, then a.e. $x \in \mathbb{R}^k$ is a Lebesgue point.

b) $\mu \in \mathcal{M}(\mathbb{R}^k)$, $\mu \ll \lambda \Rightarrow D_\mu = \frac{d\mu}{d\lambda}$, λ-a.e.

Proof: ADD

Corollary: If $[f] \in L^1_{loc.}(\mathbb{R}^k, \lambda)$ then for any $g \in [f]$, $f(x) = g(x)$ for all Lebesgue points x for f. Thus, for all $[f] \in L^1(\mathbb{R}^k, \lambda)$, there is a cononical $\hat{f} \in [f]$ such that all points in $\mathbb{R}^k$ are lebesgue points for $\hat{f}$.

Proof: ADD

For $x \in \mathbb{R}^k$, a sequence $(E_k)_{k=1}^\infty$ of measurable sets in $\mathbb{R}^k$ is said to *shrink nicely* to x if there exists a $C > 0$, scalars $r_k \searrow 0$ such that $E_k \subset B(x, r_k)$ and $\lambda(B(x,r_k)) \leq C\lambda(B(x,r_k))$ for all $k \in \mathbb{N}$. In this case we write $E_k \overset{s.n.}{\to} x$.

If $f \in L^1_{loc.}(\mathbb{R}^k, \lambda)$, and x is a lebesgue point of f then

$$f(x) = \lim_{r\to 0} \frac{1}{\lambda(E_k)} \int_{E_k} f(y)\, d\lambda(y).$$

Proof: ADD

(first fundamental theorem of calculus)

If $[g] \in L^1([a,b],\lambda)$ resp. $[g] \in L^1(\mathbb{R},\lambda)$, let $G(x) = \int_a^x g(t)\,dt = \int_{[a,b]} f\,d\lambda$ resp. $G(X) = \int_{-\infty}^x g(t)\,dt = \int_{(-\infty,x)} g\,d\lambda$, then $G \in NAC([a,b])$ resp. $F \in NAC$, and G is differentialbe a.e. and $G' = g$ a.e.

(second fundamental theorem of calculus, version 1)

a) $F \in AC \Leftrightarrow$ (F is diff'able a.e. on $[a,b]$ and $F' \in L^1([a,b],\lambda)$ and $F(x) - F(a) = \int_z^x F'(t)\,dt$ for all $x \in [a,b]$).

b) $F \in AC \Leftrightarrow$ (F is diff'able a.e. on $\mathbb{R}$ and $F' \in L^1(\mathbb{R},\lambda)$ and $F(x) - F(a) = \int_z^x F'(t)\,dt$ for all $x \in \mathbb{R}$).

c) $F \in NAC$ or $F \in NAC([a,b])$ then $F' = \frac{d\mu_F}{d\lambda}$.

Proof: ADD

If $\mu \in \mathcal{M}(\mathbb{R}^k)$ then

a) $D_\mu(x)E$ exists for a.e. $x \in \mathbb{R}^k$ and $D_\mu = \frac{d\mu_a}{d\lambda}$ λ-a.e., where μ_a is the absolutely continuous part in the LDT of μ.

b) If $\mu \perp \lambda$ Then $D_\mu(x) = 0$ λ-a.e. and for λ-a.e. x, $\lim_{k\to 0} \mu(E_k)/\lambda(E_k) = 0$ if $E_k \overset{s.n.}{\to} x$.

c) For λ-a.e. x, $\lim_{k\to 0} \mu(E_k)/\lambda(E_k) = 0$ if $E_k \overset{s.n.}{\to} x$. ??? ADD.

Proof: ADD

Corollary: If $F \in NBV$ then $F' = \frac{d\mu_a}{d\lambda} = D_\mu$ λ-a.e. where $\mu \in \mathcal{M}(\mathbb{R})$ as $F(x) = \mu((-\infty, x])$. So a bigger class of functions is differentiable with this forumla.
Proof: ADD

If $\mu \in \mathcal{M}(\mathbb{R}^k)$ then

a) If $F : \mathbb{R} \to \mathbb{R}$ is increasing, then F is differentiable λ-a.e.

b) if $F \in BV$, then F is differentiable λ-a.e.

c) if $F \in BV$ there exists a constant c, and $G \in NBV$ such that $F = C + G$ everywhere except at a countable number of points. May take $C = \lim_{x\to-\infty} F(x)$ and $G(x) = \lim_{y\to x^+} F(y) - C$ for all x. Then $F' = G' = D_\mu = \frac{d\mu_a}{d\lambda}$ a.e. where μ is the measure on $\mathbb{R}$ associated to G.

Proof: ADD

If $H \in BV$, $H \geq 0$ for all x, $H = 0$ except on a countable set, then H is differentiable a.e. and $H' = 0$ a.e.
Proof: ADD

Remark: A function $H : \mathbb{R} \to \mathbb{R}$ such that $H' = 0$ a.e. is called a *singular function*. Note: take any $\mu \in \mathcal{M}(\mathbb{R})$, $\mu \perp \lambda$, then defining $F_\mu(x) = \mu((-\infty, x])$, as usual, then $F_\mu \in NBV$, and $F' = \frac{d\mu_a}{d\lambda} = 0$ a.e. so F_μ is singular. Conversely, If $H \in NBV$ is singular, ADD.

$F \in BV[a, b]$

5,1) $F \in AC[a, b]$, take $F(x) = 0$ for $x \notin [a, b]$, then $\lim_{x \to -\infty} F(x) = 0$, and F is continuous because it's uniformly continuous, and $F \in BV[a, b]$, so $AC[a, b] \subset NBV$, and so $\mu_F \ll \lambda$ by theorem 5.1.6. $\mu_F \ll \lambda \Leftrightarrow |\mu_F| \ll \lambda$ by prop 4.1.5.e. F is not nescesarily in $AC[a, b]$ because $F(a)$ is not nescesarily 0.

Now that $|\mu_F| = \mu_{V_F}$. let $\alpha = \nu \mapsto |\nu|, \beta = F \mapsto V_F, \gamma = F \mapsto \mu_F$, and γ is invertible by 5.1.4. Want to show that $\alpha \circ \gamma = \gamma \circ \beta$, but then $\gamma^{-1} \circ (\alpha \circ \gamma) = \beta$, which says that $|\mu_F|((-\infty, x]) = V_F(x)$ for all x, which is true by last lines of proof of 5.1.4.

So $|\mu_F| = \mu_{V_F} \ll \lambda$. $F \in NBV \Rightarrow V_F \in NBV$ by lemma before 8 in 5.1.3, then again theorem 5.1.6 says that $\mu_{V_F} \ll \lambda \Rightarrow V_F \in NAC \Rightarrow V_F \in AC$, and so $V_F \in AC[a, b]$. $\square$

5,2) Suppose $F \in NBV[a, b]$, $F \geq 0$ and F is increasing, $F(a) = 0$. Let μ be the measure associated to F, then by corollary 5.2.5, $F' = \frac{d\mu_{F,a}}{d\lambda}$, where $\mu_F = \mu_{F,a} + \mu_{F,s}$, form the Lebesgue decomposition theorem, $\mu_{F,a} \ll \lambda$. By 5.1.4, $|\frac{d\mu_{F,a}}{d\lambda}| = \frac{d|\mu_{F,a}|}{d\lambda}$, so that $\int_{[a,b]} |F'| \, d\lambda = |\mu_{F,a}|((a, b])$. Now $F \geq 0 \Rightarrow \mu_F \geq= 0 \Rightarrow \mu_{F,a}, \mu_{F,s} \geq 0$, by the lebesgue decomposition theorem, so $\mu_{F,a} \leq \mu_F \Rightarrow |\mu_{F,a}|(a, b) \leq |\mu_F|(a, b) = V_F(a, b)$.

If $F \in BV[a, b]$, then by 5.2.6, $F(x) = F(a) + G(x)$, a.e. where $G \in NBV[a, b]$, and $F' = G'$ a.e. so that $\int_{[a,b]} |F'| \, d\lambda = \int_{[a,b]} |G'| \, d\lambda$.

5,3) In 5.2, have $\int_{[a,b]} |F'| \, d\lambda = |\mu_{F,a}|((a, b])$, and $\mu_F = \mu_{F,a} + \mu_{F,s}$, but $\mu_{F,s} = 0 \Leftrightarrow \mu_F \ll \lambda \Leftrightarrow F \in NAC[a, b]$.

5,4) If $F \in BV[a, b]$, then by 5.2.6, $F(x) = F(a) + G(x)$, a.e. where $G \in NBV[a, b]$, and $F' = G'$ a.e, so assume that $F \in NBV[a, b]$. In this case there is a $\mu \in \mathcal{M}(\mathbb{R})$ such that $\mu((-\infty, x]) = F(x)$ for all x. Apply lebesgue decomposition to obtain $\mu = \mu_a + \mu_s$, 5.2.6 says that $F' = \frac{d\mu_a}{d\lambda} = D_\mu$ a.e. So $|F'| = |D_\mu| = |\frac{d\mu_a}{d\lambda}| = \frac{d|\mu_a|}{d\lambda}$ by prop 1 in 4.1.7. Then again,

$$|F'| = \frac{d|\mu_a|}{d\lambda} = D_{|\mu|} = \lim_{r \to 0} \frac{|\mu_a|(B(x, r))}{\lambda(B(x, r))} = \lim_{r \to 0} \frac{V_F(x - r, x + r)}{2r} = V_F'(x) \quad \lambda - \text{a.e.}$$

$\square$

Folland 33) $F : \mathbb{R} \to \mathbb{R}$, By 5.2.6a F is differentiable a.e.. Given $a < b \in \mathbb{R}$, replace F with $F\mathbf{1}_{[a,b]}$. By problem 4, F' is measurable. If $F(x) < \infty$ then $F(y) < \infty$ for all $y \in [a, x]$ by F increasing; if $F(x) = \infty$ some $x \in [a, b]$ then $F(b) - F(a) = \infty >= \int_{[a,b]} F' \, d\lambda \in [-\infty, \infty]$ so the equality holds. assume $F(x) \in \mathbb{R}$ for all $x \in [a, b]$. Then F is bounded by $F(b)$, so by 5.1.3b $F \in BV[a, b]$. By 5.2.6c $F(x) = F(a) + G(x)$ a.e. where $G \in NBV[a, b]$, and $G' = F'$ a.e., also then by problem 5,2) $\int_{[a,b]} |F'| \leq V_F(a, b) < \infty$, so $G', F' \in L^1([a, b], \lambda)$. Let μ be the measure asssociated with F and G. This is the same measure for both F, G, because it is defined as the inf of sums of the numbers of the form $F(b_k) - F(a_k)$, but $F(b_k) - F(a_k) = G(b_k) - G(a_k)$. Now $G(a) = 0$ and G is increasing, so $G(x) \geq 0$ for $x \in [a, b]$, and $G \in NBV[a, b]$, so by the arguement in problem 5,2 , $\mu_a \leq \mu$ and $F' = \frac{d\mu_a}{d\lambda}$ a.e., so $\int_{[a,b]} F' \, d\lambda = \mu_a([a, b])$. On the other hand, $F(b) - F(a) = \mu((a, b])$, and $\mu \geq \mu_a$ from the lebesgue decomposition, i.e. $\mu = \mu_a + \mu_s$ and all these are positive finite measures, because $G \geq 0$. This gives $F(b) - F(a) = \mu((a, b]) \geq \mu_a([a, b]) = \int_{[a,b]} F' \, d\lambda$. a, b are arbitrary so this holds for all $a, b \in \mathbb{R}$. $\square$

Main idea:

$(\Omega, \mathcal{F}, P)$, $P(\Omega) = 1$ a probability space.

If picking n points, $\{\omega_{n,k}\}_{k=1}^{n}$ "at random" from Ω, so all $\omega_{n,k} \in \Omega$, then the following will be true

$$\lim_{n\to\infty} \frac{\#\{k \in \{1,2,...,n\}; \omega_{n,k} \in E\}}{n} = P(E), \text{ for all } E \in \mathcal{F},$$

where $\#$ is the counting measure.

Nonsense:

$\alpha : \mathbb{N} \to \Omega$, $n \in \mathbb{N}$, onto, but not one to one. Define $\#_n = \frac{1}{n}\#$, $N = \{1, 2, ..., n\}$. Then $(N, \mathcal{P}(N), \#_n)$ is a porobability space. Let $\alpha_n = \alpha|_N$, then this is an Ω valued random varaible. Now we can define

$$P(E) = \lim_{n\to\infty} \#_n \alpha_n^{-1}(E), \text{ for all } E \in \mathcal{F}$$

Random varaibles:

$(S, \mathcal{S})$ a measurable space, $X : \Omega \to S$ is called a random variable when it is $(\mathcal{F}, \mathcal{S})$-measurable.

Define $P_X : \mathcal{S} \to [0, +\infty]$ by $P_X(E) = P(\{\omega \in \Omega; X(\omega) \in E\})$, this is the image measure by X.

$$P_X(S) = P(\{\omega \in \Omega; X(\omega) \in S\}) = P(\Omega) = 1$$

So the image measure induced by a random variable is a probability measure on its state space.

P_X is called the dirstribution of X.

Define $F_X : \mathbb{R} \to [0, 1] = x \mapsto P_X((-\infty, x])$, this is called the cumulative distribution function.

From wikipedia:
"The probability density function of a random variable is the RadonNikodym derivative of the induced measure with respect to some base measure (usually the Lebesgue measure for continuous random variables)."
ADD many details here

Expectation:

Define the expectation value of X as $E(X) = \int_\Omega X\, dP$, the integral of X.

Suppose X is a simple function, then $X(\omega) = \sum_{k=1}^{n} c_k \chi_{E_k}(\omega)$, $c_k \in S$, unique, and $E_k \in \mathcal{F}$ disjoint.

$$E(X) = \sum_{k=1}^{n} c_k P(E_k)$$

Discrete rv:

A discrete random variable X is one whose state space is countable. In this case there is a bijective map, $\gamma : S \to \mathbb{N}$, and clearly the function $\gamma \circ X$ is $(\mathcal{F}, \mathcal{P}(\mathbb{N}))$-measurable.
We may write $S = \{x_k := \gamma^{-1}(k)\}_{k=1}^{\infty}$, and may define $E_k := X^{-1}(x_k)$, $X(\omega) = \sum_{k=1}^{\infty} x_k \chi_{E_k}$.
If we temporarily adopt the notation "$p(x_k) = P(X = x_k)$"$:= P(E_k)$, then

$$E(X) = \sum_{k \in \mathbb{N}} x_k p(x_k)$$

In this simple case Ω may not really be nescesary, as $(\{x_k\}, \mathcal{P}(\{x_k\}), p)$ is a probability space in it's own right, and note, with $\beta_n := X \circ \alpha_n$, α_n as in the above nonsense,

$$p(x_k) = \lim_{n \to \infty} \#_n \beta_n^{-1}(x_k), \text{ for all } x_k$$

Conditional Expectation

$(X, \mathcal{A}, \mu)$ a σ-finite measure space, and $(X, \mathcal{F})$ a measurable space, $\mathcal{F} \subset \mathcal{A}$. For any $f : X \to \mathbb{R}$, $\mathcal{A}$-measurable, $h : X \to \mathbb{R}$ is the conditional expectation of f with respect to $\mathcal{F}$ if it is $\mathcal{F}$-measurable and for all $A \in \mathcal{F}$,

$$\int_A f \, d\mu = \int_A h \, d\mu,$$

we write $E[f|\mathcal{F}]$ for the conditional expectation of f w.r.t. $\mathcal{F}$.

If in addition, $f \in L^1(X, \mathcal{A}, \mu)$, then $\nu \in \mathcal{M}(X, \mathcal{A})$, where $\nu(E) := \int_E f \, d\mu$. $(X, \mathcal{F}, \mu|_{\mathcal{F}})$ is a measure space, $\nu|_{\mathcal{F}} \in \mathcal{M}(X, \mathcal{F})$, and $\nu|_{\mathcal{F}} \ll \mu|_{\mathcal{F}}$, then by the RNT, there exists a unique $h \in L^1(X, \mathcal{F}, \mu|_{\mathcal{F}})$ such that $\nu(E) = \int_E h \, d\mu|_{\mathcal{F}}$.

$$E[f|\mathcal{F}] = \frac{d\nu}{d\mu|_{\mathcal{F}}}, \text{ where } \nu(E) = \int_E f \, d\mu \text{ for all } E \in \mathcal{A}, \; f \in L^1(X, \mathcal{A}, \mu), \; \mathcal{F} \subset \mathcal{A}.$$

Remarks:

1. For $(X, \mathcal{A}, \mu)$ a measure space, $f, g : X \to \mathbb{R}$, $\mathcal{A}$-measurable. Then $\int_A f \, d\mu = \int_A g \, d\mu$ for all $A \in \mathcal{A} \Rightarrow f = g$, $\mu-$a.e. Proof: let $h = f - g$, $A^+ = h^{-1}([0, \infty])$, $A^- = h^{-1}([-\infty, 0))$. Then $h \geq 0$ on A^+ so $\int_{A^+} h \, d\mu = 0 \Rightarrow h = 0$ μ-a.e. on A^+ by the vanishing principle. Similarly, $h^- = 0$ μ-a.e. on A^-.

2. In (1), what is really needed for the result is that the integrals of f and g agree on $(f-g)^{-1}(\mathbb{R}^+) \in \mathcal{A}$ and $(f-g)^{-1}(\mathbb{R}^-) \in \mathcal{A}$.

3. Let h be the conditional expectation of f w.r.t. $\mathcal{F}$, with $F \subset \mathcal{A}$, $\Delta := f - h$. h is $\mathcal{F}$-measurable, so is $\mathcal{A}$-measurable, so Δ is too, and so $\Delta^{-1}(\mathbb{R}^{+,0}) \in \mathcal{A}$, and $\Delta^{-1}(\mathbb{R}^-) \in \mathcal{A}$. By (2), if in addition, $\Delta^{-1}(\mathbb{R}^{+,0}) \in \mathcal{F}$, and $\Delta^{-1}(\mathbb{R}^-) \in \mathcal{F}$, then $f = h$ $\mu-$a.e., so in some sense, the larger $\mathcal{F}$ is, the closer an approximation h is of f.

Call $C_c(X,Y)$ the space of all continuous functions on X, to Y, if Y is not specified, assume $Y=\mathbb{R}$. Write $||f||_u$ for the uniform norm of f, i.e., $||f||_u = \sup(\{|f(x)|; x \in X\})$ A linear functional I on $C_c(X)$ is positive is $I(f) \geq 0$ when $f \geq 0$. If U is open in X, and $f \in C_c(X)$, write $f \prec U$ to mean that $f(x) \in [0,1]$ and $\mathrm{supp}(f) \subset U$, this is a stronger statement than $0 \leq f \leq \mathbf{1}_U$, which only implies that $\mathrm{supp}(f) \subset \overline{U}$.

Proposition (Folland 7.1), if I is a positive linear functional on $C_c(X)$, for each compact $K \subset X$, there is a constant, C_K, such that $|I(f)| \leq C_K ||f||_u$ for all $f \in C_K(X)$ such that $\mathrm{supp}(f) \subset K$.

If μ is a Borel measure on X such that $\mu(K) < \infty$ for every compact $K \subset X$, then $C_c(X) \subset L^1(\mu)$, so that $f \mapsto \int f \, d\mu$ is a positive linear functional on $C_c(X)$.

If I is a positive linear functional on $C_c(X)$, there is a unique Radon measure, μ on $\mathcal{B}\ell(X)$ such that $I(f) = \int_X f \, d\mu$ for all $f \in C_c(X)$. Moreover , μ satisfies

a) $\mu(U) = \sup(\{I(f); f \in C_c(X), f \prec U\})$ for all open $U \subset X$

b) $\mu(K) = \inf(\{I(f); f \in C_c(X), f \geq \mathbf{1}_K\})$ for all compact $K \subset X$.

Proof: ADD

Definitions:

For a normed space X, write $\mathrm{Ball}(X) := \{x \in X; ||x|| \leq 1\}$, and $\mathcal{U}_X := \{x \in X; ||X|| < 1\}$.

A function $f : X \to Y$, between topological spaces is continuous if $f^{-1}(U)$ is open for all open $U \subset Y$.

A function $f : X \to Y$, between topological spaces is an open map if $f(U)$ is open for all open $U \subset X$.

A homeomorphism or bicontinuous map is a bijective map which is also open, or f^{-1} is also continuous.

Prop: for a bijective map, $f : X \to Y$, f is an open map iff f^{-1} is continuous.