Some topics in measure theory

1 The Egorov and Lusin theorems

1.1 Lemma

Let E be a subset of $\mathbf{R}^k$ with finite Lebesgue measure $m(E)$. Let $f_n : E \to \mathbf{R}^*$ be measurable functions converging a.e. on E to a function $f : E \to \mathbf{R}$. Let ε, δ be positive. Then there exist N and a subset A of E such that $m(A) < \delta$ and $|f_n(x) - f(x)| < \varepsilon$ for all $n \geq N$ and all $x \in E \backslash A$.

Proof. Since $\limsup f_n$ and $\liminf f_n$ are measurable, we can ensure that f is measurable by changing it to be 0 on the set of measure 0 on which f_n fails to tend to f. Let E_N be the set of all $x \in E$ for which there exists $n \geq N$ with $|f_n(x) - f(x)| \geq \varepsilon$. Then $E_{N+1} \subseteq E_N$ and the intersection P of the E_N has measure 0. Since

$$m(E) = m(E \backslash P) = \lim m(E \backslash E_N),$$

there exists N such that $A = E_N$ has $m(A) < \delta$.

1.2 Egorov's theorem

Let E be a subset of $\mathbf{R}^k$ of finite Lebesgue measure $m(E)$. Let $f_n : E \to \mathbf{R}^*$ converge a.e. on E to $f : E \to \mathbf{R}$. Let $\eta > 0$. Then there exists a subset A of E, with $m(A) < \eta$, such that $f_n \to f$ uniformly on $E \backslash A$.

Proof. Choose $A_n \subseteq E$ and p_n such that $m(A_n) < \eta/2^n$ and $|f_m(x) - f(x)| < 1/n$ for all $m \geq p_n$ and all $x \in E \backslash A_n$. Let $A = \bigcup A_n$. Then $E \backslash A \subseteq E \backslash A_n$, and so $m \geq p_n$ implies that $|f_m(x) - f(x)| < 1/n$ for all $x \in E \backslash A$.

Note that no such theorem holds for E of infinite measure. Let $f_n = \chi_{[n,\infty)}$, for $n \in \mathbf{N}$. Then $f_n \to 0$ pointwise, but not uniformly on the complement of any set of finite measure.

1.3 Definition

Let X be a Hausdorff space and let $f : X \to \mathbf{R}$ be a function. The support of f is the closure C of the set $\{x \in X : f(x) \neq 0\}$. We say f has compact support if this set is compact.

Note that if there exists a compact set F such that $f = 0$ off F then $X \backslash F$ is open and does not meet C. Thus C is a closed subset of F and so compact.

1.4 Urysohn's lemma

Let K be a compact subset of an open subset V of the metric space X. Then there exists a continuous $f : X \to [0, 1]$ such that $f = 1$ on K and $f = 0$ off V. If X is locally compact then f can be chosen to have compact support.

Proof. Define $f(x)$ by

$$1 - f(x) = \frac{d(x, K)}{d(x, K) + d(x, V^c)}.$$

Alternatively, take $d(K, V^c) > \delta > 0$ and put

$$f(x) = \delta^{-1} \min\{\delta, d(x, V^c)\}.$$

Now suppose that X is locally compact. This means that every x lies in an open set with compact closure. Hence $B(x, r)$ has compact closure for all sufficiently small r. Cover K by finitely many open V_j whose compact closures are contained in V, and apply the above construction with V replaced by the union of the V_j.

1.5 Lusin's theorem

Suppose that f is a measurable real-valued function on $\mathbf{R}^k$ and that E is a subset of $\mathbf{R}^k$ with $m(E) < \infty$ and $f(x) = 0$ off E. Suppose that $\varepsilon > 0$. Then there exists a function $g : \mathbf{R}^k \to \mathbf{R}$ such that g is uniformly continuous on $\mathbf{R}^k$ and vanishes outside a compact set, and $m(\{x : f(x) \neq g(x)\}) < \varepsilon$. If $|f| \leq M$ on $\mathbf{R}^k$, we can choose g so that $|g| \leq M$ on $\mathbf{R}^k$.

Proof. Assume first that $0 \leq f < 1$ on E and E is compact. Choose measurable simple functions with $0 \leq s_n \leq s_{n+1} \leq f$ such that $s_n \to f$ pointwise. Here s_n is formed in such a way that it takes only values $m/2^n$ with $m \in \mathbf{Z}$. In fact $s_n(x)$ is the largest $m/2^n$ which is $\leq f(x)$. So $s_0(x) = 0$ and $t_n(x) = s_n(x) - s_{n-1}(x)$ is always 0 or 2^{-n}. So $2^n t_n(x)$ is χ_{T_n} for some T_n. Finally, we have $f(x) = \sum_{n \in \mathbf{N}} t_n(x)$.

Fix a bounded open set V containing E. Choose compact K_n and open V_n such that

$$K_n \subseteq T_n \subseteq V_n \subseteq V, \quad m(V_n \backslash T_n) < 2^{-n}\varepsilon.$$

Using Urysohn's lemma, choose a continuous h_n which is 1 on K_n and 0 off V_n. Let $H = \bigcup_n (V_n \backslash T_n)$. Then $m(H) < \varepsilon$. Let $g(x) = \sum_n 2^{-n} h_n(x)$. Then g is continuous and so uniformly continuous on $\mathbf{R}^k$, since it vanishes off a compact set. For x not in H we have either $x \notin V_n$, in which case $h_n(x) = 0 = t_n(x)$, or $x \in T_n$, in which case $h_n(x) = t_n(x) = 1$. So $g(x) = f(x)$ off H.

If E is not compact, choose compact $A \subseteq E$ with $m(E \backslash A) < \varepsilon/2$ and apply the above with ε replaced by $\varepsilon/2$.

For bounded f, say $|f| \leq M$, look at $(f + M)/3M$.

For a general measurable f, just note that since $\bigcap_n \{x : |f(x)| > n\} = \emptyset$ and $m(E) < \infty$, we have $|f| \leq n$ off a set of measure $\rho_n \to 0$.

Finally, to show that if $|f| \leq M$ then g can be chosen so that the same is true of g, let $\phi(y) = y$ for $|y| \leq M$, with $\phi(y) = yM/|y|$ for $|y| > M$. Then $\phi(g)$ is continuous, bounded in modulus by M, and $x \notin H$ gives $f(x) = g(x) = \phi(g(x))$.

1.6 Corollary

Let f be a measurable real-valued function on $\mathbf{R}^k$ such that $\{x : f(x) \neq 0\}$ has finite measure. Then there exist continuous functions g_n, each vanishing off a compact set L_n, such that $g_n(x) \to f(x)$ for almost every x.

To prove this, choose g_n as in Lusin's theorem, such that $g_n = f$ off a set P_n of measure at most 2^{-n}. Then $Q_n = \bigcup_{m=n}^{\infty} P_m$ has measure at most 2^{1-n}. For almost all x we have x not in the intersection of the Q_n. Thus for large n we have $x \notin Q_n$ and $g_n(x) = f(x)$.

1.7 Lemma: partitions of unity

Let K be a compact subset of a locally compact metric space (X, d), and let $K \subseteq V_1 \cup \ldots \cup V_n$, with the V_j open. Then there exist functions $h_j : X \to [0, 1]$, with compact support and with $h_j = 0$ off V_j, such that $h_1 + \ldots + h_n = 1$ on K.

Proof. For each $x \in K$ choose open W_x with compact closure $\overline{W_x} \subseteq V_m$ for some m. Choose x_q such that $K \subseteq \bigcup_{q=1}^{M} W_{x_q}$. Let H_m be the union of those $\overline{W_{x_q}}$ which satisfy $\overline{W_{x_q}} \subseteq V_m$. Choose $g_m : X \to [0, 1]$, with compact support, such that $g_m = 1$ on H_m and $g_m = 0$ off V_m. Define

$$h_1 = g_1, \quad h_2 = (1 - g_1)g_2, \ldots, \quad h_n = (1 - g_1)(1 - g_2) \ldots (1 - g_{n-1})g_n.$$

Then $h_m = 0$ off V_m. We have

$$h_1 + \ldots + h_n = 1 - (1 - g_1) \ldots (1 - g_n).$$

This is true for $n = 1$ and is verified by induction on n, using

$$h_1 + \ldots + h_m + h_{m+1} = 1 - (1 - g_1) \ldots (1 - g_m) + (1 - g_1) \ldots (1 - g_m)g_{m+1}.$$

Now for $x \in K$ we have $x \in H_m$ and so $g_m(x) = 1$ for at least one m.

2 Fubini's theorem

2.1 Lemma

Let K be a compact subset of $\mathbf{R}^N$. Then $g = \chi_K$ is upper semi-continuous and there exists a sequence g_n of uniformly bounded continuous functions such that $g_1 \geq g_2 \geq \ldots \geq g$ and $g_n \to g$ pointwise on $\mathbf{R}^N$. If S is large enough we can choose the g_n to be zero outside the ball of centre the origin, radius S.

Proof. Suppose first that $g(x_0) < t$. If $t > 1$ then $g < t$ everywhere. If $t \leq 1$ then x_0 is in the open complement of K and so $g < t$ near to x_0.

Choose very large S and take $n \in \mathbf{N}$. Let E_n be the (closed) set of x whose distance from K is at least $1/n$. Set

$$g_n(x) = 1 - \frac{\mathrm{dist}\{x, K\}}{\mathrm{dist}\{x, E_n\} + \mathrm{dist}\{x, K\}}.$$

Then $g = 0$ on E_n and $g = 1$ on K and, if $x \notin K$ then $x \in E_n$ for large enough n. Since $E_n \subseteq E_{n+1}$, the denominators on the RHS decrease and so does $g_n(x)$.

2.2 Lemma

Let J be a closed interval in $\mathbf{R}^N$. Let g be continuous, real-valued on the closed interval $[a,b]\times J$ in $\mathbf{R}^{N+1}$. Then

$$\int_{[a,b]\times J} g(x,y)dxdy = \int_a^b \left(\int_J g(x,y)dy \right) dx.$$

Proof. Here we are using y to denote an N-tuple $(y_1,\ldots,y_N)$. Let $a < X_0 < b$ and let

$$H(X) = \int_{[a,X]\times J} g(x,y)dxdy, \quad L(X) = \int_a^X \left(\int_J g(x,y)dy \right) dx.$$

Let $\varepsilon > 0$. Then there exists $\delta > 0$ such that $|(x,y)-(x',y')| < 4\delta$ implies that $|g(x,y)-g(x',y')| < \varepsilon$, for any $(x,y),(x',y') \in [a,b]\times J$. Let $X_0-\delta < X_1 < X_2 < X_0+\delta$. Take a partition Q_0 of J so fine that for any refinement Q of Q_0 the following is true. If J_m are the closed intervals into which Q partitions J, then for any choice of $t_m \in J_m$ we have

$$\left| \sum_m g(X_0,t_m)|J_m| - \int_J g(X_0,y)dy \right| < \varepsilon.$$

Here $|J_m|$ denotes the N-dimensional volume. Assume also that Q_0 is so fine that, for every refinement Q of Q_0, each $[X_0-\delta, X_0+\delta]\times J_m$ has diameter less than 4δ. We form a partition P_0 of $[X_1,X_2]\times J$ by taking the product of $[X_1,X_2]$ with each closed interval of Q_0. Now let P be any refinement of P_0. The closed intervals of P_0 are then just $K_{m,k} = [x_{k-1},x_k]\times J_m$, in which $X_1 = x_0 < x_1 < \ldots < x_n = X_2$ is a partition of $[X_1,X_2]$ and the J_m are the closed intervals corresponding to some refinement Q of Q_0. Fix $y_m \in J_m$. If $(x_{m,k}, y_{m,k}) \in K_{m,k}$ we get a Riemann sum

$$\sum_{m,k} g(x_{m,k},y_{m,k})|K_{m,k}| = \sum_{m,k} g(x_{m,k},y_{m,k})|x_k - x_{k-1}||J_m|$$

and this is within $\rho = \varepsilon(X_2-X_1)|J|$ of

$$\sum_{m,k} g(X_0,y_m)|x_k - x_{k-1}||J_m| = (X_2-X_1)\sum_m g(X_0,y_m)|J_m|,$$

which is within $(X_2-X_1)\varepsilon$ of $(X_2-X_1)\int_J g(X_0,y)dy$. Thus as $H_2 - H_1 \to 0+$,

$$\frac{H(X_2)-H(X_1)}{X_2-X_1} \to \int_J g(X_0,y)dy.$$

Hence $H'(X) = L'(X)$.

2.3 Fubini-Tonelli theorem: first step

Let I be a closed interval in $\mathbf{R}$ and J a closed interval in $\mathbf{R}^N$. Denote p-dimensional Lebesgue measure by λ_p. Let K be a closed subset of the interval $I\times J$ in $\mathbf{R}^{N+1}$. Take uniformly bounded continuous g_n decreasing to $g = \chi_K$ as in Lemma 2.1 above. For fixed x, the function $g_x(y) = g(x,y)$ is a limit of continuous functions and so measurable. Also the DCT gives

$$\phi(x) = \int_J g(x,y)d\lambda_N(y) = \lim \phi_n(x), \quad \phi_n(x) = \int_J g_n(x,y)d\lambda_N(y)$$

so ϕ is again a limit of continuous functions and so measurable. So

$$\int_I \left(\int_J g(x,y)d\lambda_N(y)\right) d\lambda(x) = \int_I \lim \phi_n(x)d\lambda(x) = \lim \int_I \phi_n(x)d\lambda(x),$$

using the DCT again, since $\phi_n \to \phi$. The last limit is the same as

$$\lim \int_{I\times J} g_n(x,y)d\lambda_{N+1} = \int_{I\times J} g(x,y)d\lambda_{N+1}$$

by the DCT.

2.4 Corollary

Let E be a compact or relatively open subset of $I \times J$. Let $g(x,y) = \chi_E(x,y)$. Then for each $x \in I$, the function $g_x(y) = g(x,y)$ is a measurable function of y. Also

$$k(x) = \int_J g(x,y)d\lambda_N(y)$$

is a measurable function of x and we have

$$\int_{I\times J} g(x,y)d\lambda_{N+1} = \int_I \left(\int_J g(x,y)d\lambda_N(y)\right) d\lambda(x).$$

We have already established this for compact E and for relatively open E we just look at the complement in $I \times J$. The fact that k is a measurable function arises since k is a limit of continuous functions.

2.5 Lemma

Let E be a subset of $\mathbf{R}^{N+1}$, of measure 0. Then for almost all x the set $E_x = \{y \in \mathbf{R}^N : (x,y) \in E\}$ has measure 0.

Proof. Assume WLOG that E is bounded, say $E \subseteq I \times J$, with I, J as above. Let $\delta > 0, \varepsilon > 0$. Enclose E is the union V of open intervals I_m of total volume at most $\delta\varepsilon$. Let $g(x,y) = \chi_V(x,y)$. Then

$$\delta\varepsilon \geq \int_{I\times J} g(x,y)d\lambda_{N+1} = \int_I \left(\int_J g(x,y)d\lambda_N(y)\right) d\lambda(x).$$

Let F be the set of x such that

$$\int_J g(x,y)d\lambda_N(y) \geq \varepsilon.$$

Then F has measure at most δ. Let G be the set of x for which E_x has outer measure at least ε. Then $G \subseteq F$ and so G has outer measure at most δ. Since δ is arbitrary, G has measure 0. Since ε is arbitrary, the set of x for which E_x has positive outer measure has itself 0 outer measure.

2.6 Fubini-Tonelli theorem: second step

Next, let g be a bounded non-negative measurable function on $I \times J$, say $0 \le g \le M$. Then by Lusin's theorem there exists a sequence of continuous functions g_n, with $|g_n| \le M$, such that $g_n \to g$ a.e. on $I \times J$, say on $(I \times J) \backslash E$, where E has measure 0.

So for almost all x we have $g_n(x, y) \to g(x, y)$ a.e. on J. For these x, the function $g_x(y) = g(x, y)$ is thus a measurable function on J, and

$$k_n(x) = \int_J g_n(x, y) d\lambda_N(y) \to k(x) = \int_J g(x, y) d\lambda_N(y)$$

by the DCT. Since each k_n is a continuous function of x and $k_n \to k$ a.e. on I, we see that k is a measurable function of x. Finally, using the DCT,

$$\int_{I \times J} g(x, y) d\lambda_{N+1} = \lim \int_{I \times J} g_n(x, y) d\lambda_{N+1} =$$

$$= \lim \int_I \left(\int_J g_n(x, y) d\lambda_N(y) \right) d\lambda(x) =$$

$$= \int_I \lim \left(\int_J g_n(x, y) d\lambda_N(y) \right) d\lambda(x) = \int_I \left(\int_J g(x, y) d\lambda_N(y) \right) d\lambda(x).$$

2.7 Fubini-Tonelli theorem

Let $f(x, y) : \mathbf{R}^{N+1} \to [0, \infty]$ be Lebesgue measurable. Then for almost all x, the function $f_x(y) = f(x, y)$ is a measurable function of y. Also

$$\int_{\mathbf{R}^N} f(x, y) d\lambda_N(y)$$

is a measurable function of x, and

$$\int_{\mathbf{R}} \left(\int_{\mathbf{R}^N} f(x, y) d\lambda_N(y) \right) d\lambda(x) = \int_{\mathbf{R}^{N+1}} f(x, y) d\lambda_{N+1}.$$

Proof. We know the result is true when f is bounded, non-negative and measurable, and 0 outside a compact interval of $\mathbf{R}^{N+1}$. In general, write f as the non-decreasing limit of a sequence of non-negative simple functions s_n with compact support. For almost all x, we see that $f_x(y) = \lim s_n(x, y)$ is measurable, and

$$\int_{\mathbf{R}} s_n(x, y) d\lambda_N(y) \to \int_{\mathbf{R}} f(x, y) d\lambda_N(y) \quad a.e.(x), \quad \int_{\mathbf{R}^{N+1}} s_n d\lambda_{N+1} \to \int_{\mathbf{R}^{N+1}} f d\lambda_{N+1}.$$

3 L^p spaces

Let (X, Π, μ) be a measure space, and let $0 < p \le \infty$. For measurable complex-valued f, define

$$\|f\|_p = \left(\int_X |f|^p d\mu \right)^{1/p}, \quad 0 < p < \infty,$$

and let $\|f\|_\infty$ be the infimum of $M > 0$ such that $|f| \le M$ a.e. on X. We say that f is in L^p if $\|f\|_p < \infty$. Note that in this case f has to be finite almost everywhere, and that we identify elements of L^p if they agree a.e.

Since

$$|f+g|^p \leq 2^p \max\{|f|^p, |g|^p\} \leq 2^p(|f|^p + |g|^p),$$

each L^p is a vector space.

We shall prove that if $p \geq 1$ then L^p is a complete normed space. For $p \geq 1$, define $q = p' \geq 1$ by $1/p + 1/q = 1$.

3.1 Lemma

Let $0 \leq a, b < \infty$ and let $0 < t < 1$. Then

$$a^t b^{1-t} \leq ta + (1-t)b,$$

with equality iff $a = b$.

Proof. If $b = 0$ this is obvious. Let

$$g(x) = (1-t) + tx - x^t, \quad g'(x) = t - tx^{t-1}.$$

For $0 < x < 1$ we have $g'(x) < 0$ while for $x > 1$ we have $g'(x) > 0$. Thus g has a minimum at 1 and

$$g(x) \geq g(1) = 0, \quad x^t \leq (1-t) + tx, \quad x \geq 0.$$

Further, $g(x) > 0$ for $x \geq 0, x \neq 1$. Now put $x = a/b$, and multiply through by b.

3.2 Hölder's inequality

If $f \in L^p$ and $g \in L^q$ then $fg \in L^1$ and

$$\int_X |fg| d\mu \leq \|f\|_p \|g\|_q.$$

Equality holds only if either (i) $|f|^p$ is a constant multiple of $|g|^q$ or (ii) $g = 0$, a.e. in μ.

Proof. For $p = 1, q = \infty$ this is obvious. Now suppose $\infty > p > 1$. Assume first that

$$\|f\|_p = \|g\|_q = 1.$$

Apply the lemma with $a = |f|^p, b = |g|^q, t = 1/p, 1-t = 1/q$. We get

$$|fg| \leq t|f|^p + (1-t)|g|^q \tag{1}$$

and so

$$\int_X |fg| d\mu \leq t + (1-t) = 1.$$

Equality can only hold if equality holds in (1) a.e. and so only if $|f|^p = |g|^q$ a.e.

To prove the general case, note that the result is obvious if f or g is 0 a.e. Now look at the functions

$$F = f/\|f\|_p, \quad G = g/\|g\|_q.$$

We have $\int_X |FG| d\mu \leq 1$, with equality only if $|F|^p = |G|^q$ a.e.

3.3 Minkowski's inequality

For $p \geq 1$ we have

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p.$$

In particular L^p is a normed space.

Proof. For $p = 1, \infty$ this is easy. Assume $1 < p < \infty$. Assume also that $f, g \in L^p$ and that $f+g$ is not almost everywhere zero, because otherwise there is nothing to prove. We have $f+g \in L^p$ and

$$|f+g|^p \leq |f+g|^{p-1}|f| + |f+g|^{p-1}|g|.$$

Also

$$\||f+g|^{p-1}\|_q = \left(\int_X |f+g|^{(p-1)q} d\mu\right)^{1/q} = \left(\int_X |f+g|^p d\mu\right)^{1/q} = \|f+g\|_p^{p/q},$$

using $1/p + 1/q = 1$. Thus Hölder's inequality gives

$$\int_X |f+g|^{p-1}|f| d\mu \leq \|f\|_p \||f+g|^{p-1}\|_q = \|f\|_p \|f+g\|_p^{p/q}.$$

Thus

$$\|f+g\|_p^p \leq (\|f\|_p + \|g\|_p)\|f+g\|_p^{p/q}.$$

Since $p - p/q = 1$ the result follows.

3.4 Theorem

Let $p \geq 1$. Then L^p is complete.

Proof. Let f_n be a Cauchy sequence in L^p. Thus given $\varepsilon > 0$ there exists N such that $\|f_n - f_m\|_p < \varepsilon$ for all $n, m \geq N$. We can assume that $f_n(x)$ is finite for all n, x. Choose $n_1 < n_2 < \ldots$ such that $g_k = f_{n_k}$ has $\|g_{k+1} - g_k\|_p < 2^{-k}$. Set

$$F_n = |g_1| + \sum_{k=1}^{n-1} |g_{k+1} - g_k|.$$

Then F_n converges monotonely to a function F, and the MCT and Minkowski's inequality give $\|F\|_p < \infty$. Thus F is finite almost everywhere. Hence

$$g(x) = \lim_{n\to\infty} g_n(x)$$

exists and is finite a.e. Since $|g_n| \leq F_n \leq F$ we have $|g| \leq F$ and so $g \in L^p$. Further,

$$\|g_n - g\|_p \leq \|g_n\|_p + \|g\|_p \leq \|F\|_p + \|g\|_p < \infty,$$

and since $g_n \to g$ a.e. we have, using the DCT,

$$\|g_n - g\|_p \to 0.$$

Thus

$$\|f_n - g\|_p \leq \|g_n - g\|_p + \|f_n - g_n\|_p \to 0$$

and $f_n \to g$ in L^p.

4 Some Hilbert space facts

4.1 Lemma

Let H be a Hilbert space i.e. a vector space with an inner product $< x, y >= \overline{< y, x >}$ such that H, with norm $\|x\| =< x, x >^{1/2}$, is complete. Let E be a non-empty closed convex subset of H. Then E has a unique element x of smallest norm.

Proof. Since

$$\|x + y\|^2 + \|x - y\|^2 =< x + y, x + y > + < x - y, x - y >= 2\|x\|^2 + 2\|y\|^2,$$

we get

$$\|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 - \|x + y\|^2.$$

Let δ be the infimum of $\|u\|$ for $u \in E$, and take $x_n \in E, \|x_n\| \to \delta$. The limit x has norm δ, and is the only element of E of norm δ, since if $y \in E$ has norm δ, we have by convexity

$$(x + y)/2 \in E, \quad \|(x + y)\| \geq 2\delta$$

and so

$$\|x - y\|^2 \leq 2\delta^2 + 2\delta^2 - 4\delta^2 = 0.$$

4.2 Fact

Let M be a closed subspace of H. Then

$$M^\perp = \{y \in H :< x, y >= 0 \quad \forall x \in M\}$$

is closed.

This is because $M^\perp$ is an intersection of closed sets, as $\{y :< x, y >= 0\}$ is closed for each x.

4.3 Fact

Let M be a closed subspace of H and let $x \in H$. Then $x = P_x + Q_x$, where $P_x \in M, Q_x \in M^\perp$.

Proof. The set $x + M = \{x + y : y \in M\}$ is closed and convex. Let $Q_x = x + q$ be the element of smallest norm in $x + M$, and let $P_x = x - Q_x$.

We assert that $z = Q_x \in M^\perp$ (this is obvious if $M = \{0\}$). Let y be an element of M of norm 1. Let $\alpha \in \mathbf{C}$. Then $z - \alpha y = x + q - \alpha y \in x + M$ so, by the choice of z,

$$< z, z > \leq < z - \alpha y, z - \alpha y >,$$

which gives

$$0 \leq -\alpha < y, z > -\overline{\alpha} < z, y > +|\alpha|^2.$$

Choose $\alpha =< z, y >$. We get

$$0 \leq -|\alpha|^2$$

and so $< z, y >= 0$.

So $Q_x \in M^\perp$ and $P_x \in M$ since $Q_x \in x + M$.

4.4 Theorem

Let L be a continuous linear functional on the Hilbert space H. Then there exists a unique $y \in H$ such that $L(x) =< x, y >$ for all $x \in H$.

Proof. Let $M = \{x : L(x) = 0\}$. If $M = H$, set $y = 0$. If $M \neq H$, take $z \in H \backslash M$ and write $z = z_1 + z_2, z_1 \in M, z_2 \in M^{\perp}$. So $L(z) = L(z_2) \neq 0$, so we can choose $w \in M^{\perp}$ with norm 1. Put

$$u = L(x)w - L(w)x.$$

Then

$$< u, w >= L(x) < w, w > -L(w) < x, w > .$$

Also $L(u) = 0$, so $u \in M$. Hence $u \perp w$ so $< u, w >= 0$. Thus

$$L(x) = L(x) < w, w >= L(w) < x, w >=< x, y >, \quad y = \overline{L(w)}w.$$

5 Complex measures

Let Π be a σ-algebra of subsets of X. A complex measure is just a function $\mu : \Pi \to \mathbf{C}$ such that $\mu(E) = \sum \mu(E_j)$ whenever E is the countable union of pairwise disjoint elements E_j of Π. Note that since the sum has to be independent of the order of the terms it is implicit here that the series has to always be absolutely convergent.

Define

$$|\mu|(E) = \sup\{\sum |\mu(E_j)|\}$$

the sup taken over all partitions of E into countably many pairwise disjoint elements E_j of Π.

5.1 Theorem

$|\mu|$ is a measure.

Proof. Let E be the union of countably many pairwise disjoint elements E_j of Π. Take real t_j with $t_j < |\mu|(E_j)$. Then E_j can be written as the disjoint union of elements $A_{j,k}$ of Π, such that

$$t_j < \sum_k |\mu(A_{j,k})|.$$

Then the union of the (pairwise disjoint) $A_{j,k}$ is E so, by definition of $|\mu|$,

$$\sum_j t_j < \sum_{j,k} |\mu(A_{j,k}| \leq |\mu|(E).$$

Thus

$$\sum_j |\mu|(E_j) \leq |\mu|(E).$$

To get an inequality in the opposite direction let E be the pairwise disjoint union of elements A_k of Π. Then

$$\sum_k |\mu(A_k)| = \sum_k |\sum_j \mu(A_k \cap E_j)| \leq \sum_k \sum_j |\mu(A_k \cap E_j)| = \sum_j \sum_k |\mu(A_k \cap E_j)| \leq \sum_j |\mu|(E_j).$$

5.2 Signed measures

Given a *real* measure μ, we can put

$$\mu^+ = \frac{1}{2}(|\mu| + \mu), \quad \mu^- = \frac{1}{2}(|\mu| - \mu),$$

so that μ^+, μ^- are measures and $\mu = \mu^+ - \mu^-, |\mu| = \mu^+ + \mu^-$. This is the Jordan decomposition.

6 Absolute continuity of measures

Let μ be a measure (i.e. taking values in $[0, \infty]$ in the usual way) and let λ be a measure or a complex measure. We say that λ is absolutely continuous with respect to μ, written $\lambda \ll \mu$, if $\mu(E) = 0$ implies $\lambda(E) = 0$.

We say that λ is concentrated on A if $\lambda(E) = \lambda(A \cap E)$ for every E. This is true iff $\lambda(F) = 0$ whenever $F \cap A = \emptyset$.

Finally λ_1, λ_2 are mutually singular, written $\lambda_1 \perp \lambda_2$, if there exist disjoint A_1, A_2 such that λ_j is concentrated on A_j.

6.1 Elementary properties

We have:

(a) If λ is concentrated on A, so is $|\lambda|$. For if $F \cap A = \emptyset$ and $F = \bigcup F_j$ then $|\lambda(F_j)| = 0$.

(b) If $\lambda_1 \perp \lambda_2$ then $|\lambda_1| \perp |\lambda_2|$, by (a).

(c) If $\lambda_1 \perp \mu$ and $\lambda_2 \perp \mu$ then $\lambda_1 + \lambda_2 \perp \mu$. To see this, take A_j, B_j (disjoint for each j) such that λ_j is concentrated on A_j, while μ is concentrated on B_j. Then $\lambda_1 + \lambda_2$ is concentrated on $A_1 \cup A_2$, while μ is concentrated on $B_1 \cap B_2$.

(d) If $\lambda_1 \ll \mu$ and $\lambda_2 \ll \mu$ then $\lambda_1 + \lambda_2 \ll \mu$. This is obvious.

(e) If $\lambda \ll \mu$ then $|\lambda| \ll \mu$. Suppose $\mu(E) = 0$ and $E = \bigcup E_j$. Then $\mu(E_j) = 0$ so $\lambda(E_j) = 0$ so $|\lambda|(E) = 0$.

(f) If $\lambda_1 \ll \mu$ and $\lambda_2 \perp \mu$ then $\lambda_1 \perp \lambda_2$. To see this, we have λ_2 concentrated on a set A such that $\mu(A) = 0$ and hence $\lambda_1(A) = 0$ so that λ_1 is concentrated on the complement of A.

(g) If $\lambda \ll \mu$ and $\lambda \perp \mu$ then $\lambda \perp \lambda$ so clearly $\lambda = 0$.

6.2 Lemma

Let μ be a σ-finite measure on X (this means that $X = \bigcup E_n$ with $\mu(E_n) < \infty$). Then there exists $w : X \to (0, 1)$ with $w \in L^1(\mu)$.

Proof. Define w_n by

$$w_n(x) = \frac{1}{2^n(1 + \mu(E_n))}, \quad x \in E_n,$$

with $w_n = 0$ off E_n. Then put $w = \sum_n w_n$.

Note that here

$$\tilde{\mu}(E) = \int_E w d\mu$$

is a measure, and $\tilde{\mu}(E) = 0$ iff $\mu(E) = 0$, since if $\tilde{\mu}(E) = 0$ we get $\tilde{\mu}(E \cap E_n) = 0$ and so $\mu(E \cap E_n) = 0$ since w is constant on E_n.

6.3 The Lebesgue-Radon-Nikodym theorem

Let μ be a measure on a σ-algebra Π of subsets of X, and let λ be a complex measure on Π (N.B. it is then implicit that $\lambda(X) \in \mathbf{C}$ is finite).

Then there exist unique (complex) measures $\lambda_a \ll \mu, \lambda_s \perp \mu$, such that $\lambda = \lambda_a + \lambda_s$. If $\lambda \geq 0$, the same is true of λ_a, λ_s. Also there exists a unique $h \in L^1(\mu)$ such that

$$\lambda_a(E) = \int_E h d\mu, \quad E \in \Pi.$$

Proof. First we check the uniqueness. Given another such pair λ'_a, λ'_s, we have

$$\lambda'_a - \lambda_a = \lambda_s - \lambda'_s, \quad \lambda'_a - \lambda_a \ll \mu, \quad \lambda_s - \lambda'_s \perp \mu,$$

and so $\lambda_a - \lambda'_a = 0$.

To prove the existence, assume first that λ is a non-negative measure with $\lambda(X) < \infty$. Choose an $L^1(\mu)$ function $w : X \to (0, 1)$ as in the lemma above. Then

$$\phi(E) = \lambda(E) + \int_E w d\mu$$

defines a measure on Π, with $\phi(X)$ finite since $w \in L^1$, and

$$\int_X f d\phi = \int_X f d\lambda + \int_X f w d\mu \quad (2)$$

for every non-negative μ-measurable f on X (the formula is true for $f = \chi_E$ and so for simple functions and hence in general). We then have (2) for $f \in L^1(\phi)$.

Let $f \in L^2(\phi)$. Then Cauchy-Schwarz gives

$$|\int_X f d\lambda| \leq \int_X |f| d\lambda \leq \int_X |f| d\phi \leq (\int_X |f|^2 d\phi)^{1/2} \phi(X)^{1/2}$$

and so

$$f \to \int_X f d\lambda$$

is a bounded linear functional on $L^2(\phi)$. Now, $L^2(\phi)$ is a Hilbert space, with

$$< f, g >= \int_X f \overline{g} d\phi,$$

and so there exists $g \in L^2(\phi)$ (unique up to changing g on a set of ϕ-measure 0) such that

$$\int_X f d\lambda = \int_X f g d\phi \quad (3)$$

for all $f \in L^2(\phi)$.

Now let $\phi(E) > 0$, and put $f = \chi_E$. We get

$$\lambda(E) = \int_X f d\lambda = \int_X f g d\phi = \int_E g d\phi.$$

So

$$\frac{1}{\phi(E)} \int_E g d\phi = \frac{\lambda(E)}{\phi(E)} \in [0, 1]$$

by the definition of ϕ. So g is real a.e. and ≥ 0 a.e. (both w.r.t. ϕ). Also, since

$$0 \leq \phi(E) - \lambda(E) = \int_E (1 - g) d\phi$$

we have $0 \leq g \leq 1$ a.e. w.r.t. ϕ. We may therefore change g, if necessary, to be in $[0, 1]$ everywhere, and we still have (3).

Put

$$A = \{x : 0 \leq g(x) < 1\}, \quad B = \{x : g(x) = 1\}.$$

Let

$$\lambda_a(E) = \lambda(A \cap E), \quad \lambda_s(E) = \lambda(B \cap E).$$

These are measures. Now applying (2) to fg and using (3) give

$$\int_X f(1 - g) d\lambda = \int_X f g w d\mu, \quad f \in L^2(\phi). \tag{4}$$

Choose $f = \chi_B$. Then since $g = 1$ on B we get

$$0 = \int_B w d\mu$$

and so $\mu(B) = 0$. Since λ_s is concentrated on B we have $\lambda_s \perp \mu$.

Next choose $f = (1 + g + \ldots + g^n)\chi_E$. Then (4) gives

$$\int_E (1 - g^{n+1}) d\lambda = \int_E g(1 + g + \ldots + g^n) w d\mu. \tag{5}$$

If $x \in B$ then $g(x) = 1$ and so $1 - g(x)^{n+1} = 0$. If $x \in A$ then $1 - g(x)^{n+1} \to 1$. So the integrand on the LHS of (5) tends pointwise and non-decreasingly to $\chi_A(x)$, while that on the RHS tends pointwise to some non-negative $h(x)$. The MCT then gives

$$\lambda_a(E) = \lambda(A \cap E) = \int_E \chi_A d\lambda = \int_E h d\mu.$$

In particular $\int_X h d\mu = \lambda_a(X) \leq \lambda(X) < \infty$, so $h \in L^1(\mu)$.

In the general case we just write $\lambda = \rho + i\tau$ and consider separately the positive and negative parts of ρ, τ.

6.4 Remarks

We note:

(i) The theorem can be extended to the case where λ is a non-negative σ-finite measure, although h may not be in $L^1(\mu)$. To see this, write $X = \bigcup X_n$, where $X_n \subseteq X_{n+1}$ and $\lambda(X_n), \mu(X_n)$ are finite (take such sequences X'_n, X''_n for λ, μ respectively and put $X_n = X'_n \cap X''_n$). Now put

$H_1 = X_1, H_n = X_n \backslash X_{n+1}$. Put $\lambda_n(E) = \lambda(H_n \cap E)$ and write $\lambda_n = \alpha_n + \beta_n$ where $\alpha_n \ll \mu$ is given by

$$\alpha_n(E) = \int_{E \cap H_n} h_n d\mu$$

for some h_n defined on H_n, with

$$\int_{H_n} h_n d\mu < \infty,$$

and β_n is concentrated on a set B_n of zero μ-measure. Finally, put

$$\lambda_a(E) = \sum_n \alpha_n(E \cap H_n) = \int_E h d\mu, \quad \lambda_s(E) = \sum_n \beta_n(E \cap H_n).$$

Here $h = h_n$ on H_n. Clearly λ_s is concentrated on the union of the B_n, which has zero μ-measure.

(ii) The theorem is not true for μ Lebesgue measure and λ the counting measure on $(0, 1)$. For λ_s would be concentrated on a set A of Lebesgue measure 0. If $B \subseteq (0, 1) \backslash A$ has Lebesgue measure 0 we'd then get

$$\lambda(B) = \lambda_a(B) = 0$$

which is clearly not always true.

7 Non-decreasing functions

7.1 Vitali covering lemma

Let E be a subset of $\mathbf{R}$ of finite outer measure, and let H be a collection of intervals, each of positive length, which cover E in the sense of Vitali i.e. for every $\delta > 0$ and every $x \in E$ there exists $I_x \in H$ such that $x \in I_x$ and $m(I_x) < \delta$.

Then given $\delta > 0$ we can find pairwise disjoint $I_1, \ldots, I_N \in H$ such that

$$m^*(E \backslash (\bigcup_{j=1}^{N} I_j)) < \delta.$$

Proof. Assume WLOG that the intervals I of H are all closed. Enclose E in an open set U of finite measure. Assume WLOG that all $I \in H$ satisfy $I \subseteq U$ (we can discard those which do not, and still cover E).

We choose disjoint $I_j \in H$ as follows. Take any $I_1 \in H$. Now, assuming that $I_1, \ldots, I_n$ have been chosen, let k_n be the supremum of the lengths of intervals of H which do not meet $I_1, \ldots, I_n$ (if there are no such intervals, then $E \subseteq \bigcup_{j=1}^{n} I_j$ and we are finished). Certainly $k_n \leq m(U) < \infty$. So we choose $I_{n+1} \in H$, disjoint from $I_1, \ldots, I_n$, of length $> \frac{1}{2}k_n$.

This gives a sequence of disjoint $I_j \in H$, and since $I_j \subseteq U$ we get

$$\sum_{j=1}^{\infty} m(I_j) < \infty.$$

Choose N so that

$$\sum_{j=N+1}^{\infty} m(I_j) < \delta/5.$$

For each I_n, let J_n be the closed interval with the same mid-point as I_n, but with length $5m(I_n)$.

Let $R = E\backslash(\bigcup_{j=1}^{N} I_j)$. We need to show that $m^*(R) < \delta$. Let $x \in R$. Since the union of $I_1, \ldots, I_N$ is closed, we can find $I \in H$ with $x \in I$ such that I does not meet any of $I_1, \ldots, I_N$.

We claim that I meets some I_n. This holds, because if I fails to meet $I_1, \ldots, I_n$ then $0 < m(I) \leq k_n < 2m(I_{n+1}) \to 0$ as $n \to \infty$. So let n be the smallest integer such that I meets I_n. Then $n > N$ and $m(I) \leq k_{n-1} < 2m(I_n)$. Hence $x \in I \subseteq J_n$ (because if we attach two copies of I_n at each end, we enclose I and get J_n).

So

$$R \subseteq \bigcup_{j=N+1}^{\infty} J_n, \quad m^*(R) \leq \sum_{j=N+1}^{\infty} m(J_n) < \delta.$$

7.2 Derivates

Let $f : \mathbf{R} \to \mathbf{R}$ be non-decreasing. For each x define

$$D^R f(x) = \limsup_{h \to 0+} \frac{f(x+h) - f(x)}{h},$$

$$D_R f(x) = \liminf_{h \to 0+} \frac{f(x+h) - f(x)}{h},$$

$$D^L f(x) = \limsup_{h \to 0+} \frac{f(x) - f(x-h)}{h},$$

$$D_L f(x) = \liminf_{h \to 0+} \frac{f(x) - f(x-h)}{h}.$$

Obviously all of these are ≥ 0.

Note that if $g(x) = -f(-x)$ then g is non-decreasing and

$$D^L f(x) = D^R g(-x), \quad D_R f(x) = D_L g(-x).$$

7.3 Theorem

Let $f : [a, b] \to \mathbf{R}$ be non-decreasing. Then f is differentiable a.e.

Proof. It suffices to show that $D^R f, D^L f, D_R f, D_L f$ are equal a.e. on (a, b). Let u, v be positive rational numbers, with $u > v$, and let

$$E = \{x \in (a, b) : D^R f(x) > u > v > D_L f(x)\}.$$

Let $s = m^*(E)$. Take $\delta > 0$ and enclose E in an open U of measure $m(U) < s + \delta$.

For each $x \in E$ we can find $[x - h, x] \subseteq U$ with h arbitrarily small and positive, and

$$f(x) - f(x - h) < vh.$$

By Vitali's lemma there exist $I_n = [x_n - h_n, x_n], n = 1, \ldots, N$, whose disjoint interiors cover a subset A of E, with $m^*(A) > s - \delta$ (here we use the sub-additivity of m^*). Further,

$$\sum_{j=1}^{N} (f(x_n) - f(x_n - h_n)) < v \sum_{j=1}^{n} h_n < vm(U) < v(s + \delta).$$

Let $y \in A$. Then we can find n and $k > 0$ such that $[y, y+k] \subseteq (x_n - h_n, x_n)$ and

$$f(y+k) - f(y) > uk.$$

By Vitali again, there exist such intervals $J_p = [y_p, y_p + k_p], p = 1, \ldots, M$, whose union covers a subset B of A with $m^*(B) > s - 2\delta$. We have

$$\sum_{p=1}^{M} (f(y_p + k_p) - f(y_p)) > u \sum_{p=1}^{M} k_p > u(s - 2\delta).$$

Now each J_p is a subset of some I_n. Further,

$$\sum_{J_p \subseteq I_j} (f(y_p + k_p) - f(y_p)) \leq f(x_j) - f(x_j - h_j)$$

since f is non-decreasing. Summing over j we get

$$v(s+\delta) > \sum_{j=1}^{N} (f(x_j) - f(x_j - h_j)) \geq \sum_{p=1}^{M} (f(y_p + k_p) - f(y_p)) > u(s - 2\delta).$$

Thus $vs \geq us$ since δ is arbitrary, and this forces $s = 0$.

We have thus shown that a.e. $D_L f \geq D^R f$ and so

$$D^L f(x) \geq D_L f(x) \geq D^R f(x) \geq D_R f(x).$$

The same argument applied to $-f(-x)$ gives $D_R f(x) \geq D^L f(x)$ a.e. and the theorem is proved.

7.4 Theorem

Let $f : [a,b] \to \mathbf{R}$ be non-decreasing. Then f' is measurable and $\int_{[a,b]} f' dm \leq f(b) - f(a)$.

Proof. Extend f to a non-decreasing function on $\mathbf{R}$ by making it constant on $(-\infty, a], [b, \infty)$. Let $f_n(x) = n(f(x + 1/n) - f(x))$ for $n \in \mathbf{N}$. Then $f_n \to f'$ a.e., so f' is measurable. Also $f_n \geq 0$ and

$$\int_{[a,b]} f_n dm = n \int_{[b,b+1/n]} f dm - n \int_{[a,a+1/n]} f dm \leq f(b) - f(a).$$

The result now follows from the DCT.

8 Absolutely continuous and BV functions

A function $f : I = [a,b] \to \mathbf{R}^N$ is called absolutely continuous (AC) if the following is true: to each $\varepsilon > 0$ corresponds $\delta > 0$ such that

$$\sum_{j=1}^{n} |f(b_j) - f(a_j)| < \varepsilon$$

whenever we have pairwise disjoint (a_j, b_j) (with $a_j, b_j \in I$) of total length $\sum_{j=1}^{n} (b_j - a_j) < \delta$.

8.1 Theorem

Let $f \in L^1(I), I = [a, b]$. Then $F(x) = \int_{[a,x]} f(t)dt$ is absolutely continuous.

Proof. Assume WLOG that $f \geq 0$. Since f is finite a.e., the functions $f_n(x) = \min\{f(x), n\}$ converge a.e. to $f(x)$. Let $\varepsilon > 0$. Choose n so large that

$$\int_I (f(x) - f_n(x))dm < \varepsilon/2.$$

Since $0 \leq f_n \leq n$ there exists $\delta > 0$ such that if $E \subseteq I$ has measure $< \delta$ then $\int_E f_n(x)dm < \varepsilon/2$. Hence

$$\int_E f(x)dm \leq \int_E f_n(x)dm + \int_I (f(x) + f_n(x))dm < \varepsilon.$$

8.2 The total variation

Let $f : [a, b] \to \mathbf{R}^k$. Let $P = \{x_j\}$ be a partition of $[a, b]$ with vertices $a = x_0 < x_1 < \ldots < x_n < b$. Define

$$L(P, f) = \sum_{j=1}^{n} |f(t_j) - f(t_{j-1})|.$$

Set

$$v_f(a, b) = \sup L(P, f)$$

over all such partitions P. Obviously, $v_f(a, b) \geq |f(b) - f(a)|$.

8.3 Lemma

If $a < c < b$ we have

$$v_f(a, b) = v_f(a, c) + v_f(c, b).$$

Also $v_{f+g}(a, b) \leq v_f(a, b) + v_g(a, b)$.

Proof. Since adjoining an extra point to P can only increase $L(P, f)$ we have $v_f(a, b) = \sup L(P, f)$ with the sup over all partitions P of $[a, b]$ having c as a vertex. Such a P can be decomposed into a partition P_1 of $[a, c]$ and a partition P_2 of $[c, b]$ and

$$L(P, f) = L(P_1, f) + L(P_2, f).$$

Conversely, any such P_1, P_2 can be combined to make a partition of $[a, b]$.

The second assertion is easy since

$$L(P, f + g) \leq L(P, f) + L(P, g).$$

8.4 Bounded variation (BV)

We say that f has BV on $[a, b]$ if $v_f(a, b) < \infty$.

8.5 Lemma

Let $f : [a,b] \to \mathbf{R}^N$ have BV. Then $F(x) = v_f(a,x)$ is non-decreasing on $[a,b]$. If $N = 1$ then so is $f + F$.

Proof. The assertion for F is obvious. Take $a \leq x < y \leq b$. Take a partition P of $[a,x]$. Then

$$F(y) \geq |f(y) - f(x)| + L(P,f).$$

Thus

$$F(y) \geq |f(y) - f(x)| + F(x) \geq f(x) + F(x) - f(y).$$

8.6 Corollary

$f : [a,b] \to \mathbf{R}$ has BV iff $f = f_1 - f_2$, where the f_j are non-decreasing on $[a,b]$.

8.7 Lemma

Let $f : I = [a,b] \to \mathbf{R}^N$ be AC. Then f has BV. Also $F(x) = v_f(a,x)$ and (if $N = 1$) $f + F$ are AC on I.

Proof. We first show that f has BV. Take $\delta_1 > 0$ corresponding to $\varepsilon = 1$ in the definition of AC.

Take any partition $P = \{t_0, \ldots, t_n\}$ of $[a,b]$. Refine P if necessary so that each $t_k - t_{k-1}$ is less than $\delta_1/4$. We can divide the sub-intervals $[t_{k-1}, t_k]$ into blocks each of length $< \delta_1$ but, apart possibly from the last block, of length at least $\frac{3}{4}\delta_1$. The number Q of blocks satisfies

$$b - a \geq Q\frac{3}{4}\delta_1.$$

Thus $L(P) \leq (Q+1) \leq 1 + \frac{4}{3}(b-a)/\delta_1$.

Now we show that F is AC. Take $\varepsilon > 0$ and choose a corresponding δ for f. Let (a_q, b_q) be pairwise disjoint subintervals of I, of total length less than δ. Form a partition $Q_q = \{t_{j,q}\}$ of $[a_q, b_q]$. Then

$$\sum_q L(Q_q) < \varepsilon,$$

because the intervals $(t_{j-1,q}, t_{j,q})$ are pairwise disjoint subintervals of I, of total length $< \delta$. Thus

$$\sum_q (F(b_q) - F(a_q)) = \sum_q v_f(a_q, b_q) \leq \varepsilon.$$

8.8 Corollary

An absolutely continuous function f on $[a,b]$ can be written the difference of absolutely continuous non-decreasing functions, and f is differentiable a.e.

8.9 Theorem

Suppose that f is AC on $[a,b]$ and $f' = 0$ a.e. Then f is constant on $[a,b]$.

Proof. Take $c \in (a,b]$. Let E be the subset of (a,c) on which $f' = 0$, so that $m(E) = c - a$, and let $\varepsilon > 0, \eta > 0$. Let $\delta > 0$ correspond to ε in the definition of absolute continuity of f.

For each $x \in E$ we can find an interval $[x, x+h] \subseteq (a,c)$ with

$$|f(x+h) - f(x)| < \eta h.$$

By Vitali we can find finitely many such intervals $I_k = [x_k, y_k], y_k = x_k + h_k$, such that the interiors of the I_k are disjoint and cover all of E bar a set of measure less than δ. We can order x_k so that

$$y_0 = a \leq x_1 < y_1 \leq x_2 < \ldots < y_n \leq c = x_{n+1}.$$

Now

$$\sum_{k=0}^{n} (x_{k+1} - y_k) < \delta.$$

So

$$\sum_{k=0}^{n} |f(x_{k+1}) - f(y_k)| < \varepsilon,$$

by the definition of AC. But

$$\sum_{k=1}^{n} |f(y_k) - f(x_k)| \leq \eta \sum_{k=1}^{n} (y_k - x_k) \leq \eta(c-a).$$

Thus

$$|f(c) - f(a)| = \left| \sum_{k=0}^{n} (f(x_{k+1}) - f(y_k)) + \sum_{k=1}^{n} (f(y_k) - f(x_k)) \right| \leq \varepsilon + \eta(c-a).$$

Hence $f(c) = f(a)$.

8.10 Theorem

Let $f : I = [a,b] \to \mathbf{R}$ be AC. Then $f(b) - f(a) = \int_{[a,b]} f'(t)dm$.

Proof. It suffices to prove this when f is non-decreasing (otherwise write f as the difference of non-decreasing AC functions). Now

$$f(b) - f(a) \geq \int_{[a,b]} f' dm,$$

so f' (which is non-negative) is in L^1 on $[a,b]$. Thus

$$g(x) = f(a) + \int_{[a,x]} f' dm$$

is AC on $[a,b]$ and, since $(f-g)' = 0$ a.e., we get $f = g$.

8.11 Change of variables theorem

Let g be non-decreasing and AC on $[a,b]$. Let $c = g(a), d = g(b)$. Let f be a non-negative measurable function on $[c,d]$. Then $f(g)g'$ is measurable on $[a,b]$ and

$$\int_{[c,d]} f dm = \int_{[a,b]} f(g)g' dm.$$

Proof.

Let U be an open set. Then $V = g^{-1}(U \cap [c,d])$ is a relatively open subset of $[a,b]$ and can be written as a union of pairwise disjoint intervals I_j (all open, bar two, which are half-open). The $g(I_j)$ are disjoint, apart possibly from their end-points, and so

$$m(U \cap [c,d]) = m(g(V)) = \sum m(g(I_j)) = \sum \int_{I_j} g'(x)dm(x).$$

Thus

$$m(E) = \int_{g^{-1}(E)} g'(x)dm(x)$$

for every relatively open subset E of $[c,d]$.

Let $H = \{x : g'(x) \neq 0\}$ (note that g' is measurable). Let F be a measurable subset of $[c,d]$ with measure 0. Then we claim that the set $g^{-1}(F) \cap H$ has measure 0. Take $\delta, \varepsilon > 0$. Let U be a relatively open subset of $[c,d]$ of measure less than δ. Then

$$\int_{g^{-1}(U)} g'(x)dm(x) < \delta$$

and so the set $\{x : g(x) \in U, g'(x) > \varepsilon\}$ has measure at most δ/ε. Since δ is arbitrary, $\{x : g(x) \in F, g'(x) > \varepsilon\}$ has measure 0 and since ε is arbitrary our assertion is proved.

Next, let E be any measurable subset of $[c,d]$. Take relatively open V_n, with $E \subseteq V_n \subseteq [c,d]$ and $m(V_n) \to m(E)$. Let V be the intersection of the V_n. Then the DCT gives

$$m(E) = \lim m(V_n) = \lim \int_{[a,b]} \chi_{V_n}(g)g'(x)dm(x) = \int_{[a,b]} \chi_V(g)g'(x)dm(x).$$

But $\chi_E(g)g' = \chi_V(g)g'$ if $g' = 0$, and the set of x such that $g'(x) \neq 0$ and $g(x) \in V \backslash E$ has measure 0. Thus $\chi_E(g)g'$ is measurable, in the sense that it agrees a.e. with a measurable function, and we get

$$m(E) = \int_{[a,b]} \chi_E(g)g'(x)dm(x).$$

Now let f be a non-negative measurable function on $[c,d]$. We can write $f = \lim s_n$, in which s_n are non-negative measurable simple functions. We have $f(g)g' = \lim s_n(g)g'$ and so $f(g)g'$ is a measurable function and

$$\int_{[c,d]} f(y)dm(y) = \lim \int_{[c,d]} s_n(y)dm(y) = \lim \int_{[a,b]} s_n(g)g'(x)dm(x) = \int_{[a,b]} f(g)g'(x)dm(x)$$

by the MCT (note that $g' \geq 0$).

9 Differentiation of measures and set functions

Throughout this section we work in k dimensional real space $\mathbf{R}^k$, and we denote the k dimensional Lebesgue measure by m.

9.1 Lemma

Let $B_j = B(x_j, r_j)$ *and let* $W = \bigcup_{j=1}^{N} B_j$*. Then there exists a subset* S *of* $\{1, \ldots, N\}$ *such that:*

(i) the $B_j, j \in S$*, are pairwise disjoint;*

(ii) $W \subseteq \bigcup_{j \in S} B(x_j, 3r_j)$*;*

(iii) $m(W) \leq 3^k \sum_{j \in S} m(B_j)$*.*

Proof. We assume WLOG that $r_1 \geq r_2 \geq \ldots \geq r_N$. Put $j_1 = 1$ and discard all B_j which meet B_{j_1}. Let j_2 be the least remaining $j > j_1$ (if any)remaining. Then B_{j_2} fails to meet B_{j_1}. Discard all B_j which meet B_{j_2}, and let j_3 be the least remaining $j > j_2$ (if any). Again, B_{j_3} fails to meet B_{j_2}. Carrying on like this, S is the set of j_m. If $j > j_m$ and B_j meets B_{j_m} then, since $r_j \leq r_{j_m}$, we have $B_j \subseteq B(x_{j_m}, 3r_{j_m})$, and (ii) and (iii) follow. (i) is obvious, since we discard all B_j which meet B_{j_m}.

9.2 Lemma

Let μ *be a measure on* $\mathbf{R}^k$*, and let* $t \in (0, \infty)$*. Let*

$$\|\mu\| = \mu(\mathbf{R}^k) < \infty, \quad M_\mu(x) = \sup_{0<r<\infty} \frac{\mu(B(x,r))}{m(B(x,r))}.$$

Then $m(\{x : M_\mu(x) > t\}) \leq 3^{-k} t^{-1} \|\mu\|$*.*

Proof. Let K be a compact subset of the set $\Omega = \{x : M_\mu(x) > t\}$. Note that Ω is open because we can choose r with $\mu(B(x,r)) > tm(B(x,r))$ and $s < r$ but close to r with $\mu(B(x,s)) > tm(B(x,r))$. For x' close to x we have $\mu(B(x',r)) \geq \mu(B(x,s)) > tm(B(x,r)) = tm(B(x',r))$.

Now the compact set K may be covered by finitely many $B(x_j, r_j)$ each with $\mu(B(x_j, r_j)) > tm(B(x_j, r_j))$. Choose a set S as in the previous lemma such that the $B(x_j, r_j), j \in S$ are disjoint and

$$m(K) \leq m(\bigcup_{j=1}^{n} B(x_j, r_j)) \leq 3^k \sum_{j \in S} m(B(x_j, r_j)) \leq 3^k t^{-1} \sum_{j \in S} \mu(B(x_j, r_j)) \leq 3^k t^{-1} \|\mu\|.$$

9.3 Theorem

Let f *be in* $L^1(m)$*. Set*

$$T(x,r,f) = \frac{1}{m(B(x,r))} \int_{B(x,r)} |f - f(x)| dm, \quad T(x) = T_f(x) = \limsup_{r \to 0+} T(x,r,f).$$

Then $T(x) = 0$ *a.e. in* $\mathbf{R}^k$*.*

Proof. It suffices to prove this when f has compact support. For $p \in \mathbf{N}$, set $f_p(x) = f(x)$ if $|f(x)| \leq p$, with $f_p(x) = 0$ otherwise. Since $|f_p|$ tends a.e. increasingly to f, we have $\int_{\{x:|f(x)|>p\}} |f| dm \to 0$ and $\int_{\mathbf{R}^k} |f - f_p| dm \to 0$.

Let $y > 0$. Let $n \in \mathbf{N}$. Choose p large, and a continuous g such that $g = f_p$ off a set of small measure. This can be done by Lusin's theorem, with $|g| \leq p$, and so we can ensure that

$$\int_{\mathbf{R}^k} |f - g| dm < 1/n.$$

Put $h = f - g$. Since g is continuous we have $T_g(x) = 0$. Since

$$T(x, r, h) \le \frac{1}{m(B(x,r))} \int_{B(x,r)} |h| dm + |h(x)|$$

we get

$$T_f(x) = T_h(x) \le M_\sigma(x) + |h(x)|, \tag{6}$$

in which σ is the measure given by

$$\sigma(E) = \int_E |h| dm.$$

Thus if $T_f(x) > 2y$ then at least one of the terms on the RHS of (6) must exceed y. But

$$\|\sigma\| = \int_{\mathbf{R}^k} |h| dm \le 1/n,$$

so

$$m(\{x : M_\sigma(x) > y\}) \le 3^k (yn)^{-1}$$

by the previous lemma. Also,

$$m(\{x : |h(x)| > y\}) \le 1/yn.$$

So

$$E(y, n) = \{x : M_\sigma(x) > y \quad \text{or} \quad |h(x)| > y\}$$

has

$$m(E(y, n)) \le (3^k + 1)(yn)^{-1}.$$

But if $T_f(x) > 2y$ then x is in every $E(y, n)$, so the set $\{x : T_f(x) > 2y\}$ has measure 0. This is true for every $y > 0$, and so the theorem is proved.

9.4 Definition

We say that a sequence of Borel sets B_n shrink nicely to x if there exist $t \in (0, \infty)$ and $r_n \to 0+$ such that $E_n \subseteq B(x, r_n)$ and $m(B_n) \ge t m(B(x, r_n))$ for all n. It is not required that $x \in B_n$.

9.5 Theorem

Let $f \in L^1(m)$ on $\mathbf{R}^k$. Then for almost every x in $\mathbf{R}^k$ the following is true.

If B_n shrinks nicely to x then

$$\lim_{n\to\infty} \frac{1}{m(B_n)} \int_{B_n} f dm = f(x).$$

It is not required that the t associated with B_n be independent of x.

Proof. We have

$$|\frac{1}{m(B_n)} \int_{B_n} f dm - f(x)| = |\frac{1}{m(B_n)} \int_{B_n} f - f(x) dm| \le \frac{1}{m(B_n)} \int_{B_n} |f - f(x)| dm.$$

But

$$\frac{1}{m(B_n)} \int_{B_n} |f - f(x)| dm \le \frac{1}{t m(B(x, r_n))} \int_{B(x,r_n)} |f - f(x)| dm \to 0$$

for almost every x, independent of t.

9.6 Set functions

Let A, B be open subsets of $\mathbf{R}^k$ and let $f : A \to B$ be a homeomorphism. For a Borel subset E of A, set

$$\mu(E) = m(f(E)).$$

Then μ is a σ-finite measure. We may write

$$\mu(E) = \int_E hdm + \mu_s(E),$$

in which h is a non-negative measurable function, finite a.e., and μ_s is a measure, singular w.r.t. Lebesgue measure. Note that h is finite a.e. since we can write A as the union of compact sets H, for each of which $m(H)$ and $\mu(H)$ are finite.

Claim: *For Borel $E \subseteq A$ such that E and $f(E)$ have finite Lebesgue measure, $\mu_s(E)$ is the supremum of $\mu_s(H)$ over all compact subsets H of A.*

To prove this, take compact $F_n \subseteq F_{n+1} \subseteq f(E)$ such that $m(F_n) \to m(f(E))$, and compact $E_n \subseteq E_{n+1} \subseteq E$ such that $m(E_n) \to m(E)$. Let $H_n = E_n \cup f^{-1}(F_n)$. Then $m(F_n) \leq m(f(H_n)) \leq m(f(E))$, so $m(f(H_n)) \to m(f(E))$ i.e. $\mu(H_n) \to \mu(E)$. Also $m(E_n) \leq m(H_n) \leq m(E)$, so $m(H_n) \to m(E)$. Thus $H^* = \bigcup_n H_n$ has $m(E \backslash H^*) = 0$, so the MCT gives

$$\int_{H_n} hdm \to \int_{H^*} hdm = \int_E hdm.$$

Thus $\mu_s(H_n) \to \mu_s(E)$.

9.7 Theorem

Let f, μ, μ_s be as in the previous subsection. For almost all $x \in \mathbf{R}^k$, if B_n shrinks nicely to x then

$$\frac{m(f(B_n))}{m(B_n)} \to h(x).$$

Proof. Reducing A if necessary we can assume that $\mu(A)$ and $m(A)$ are finite. Define $\nu(E) = \mu_s(E \cap A)$ so that ν is a measure on $\mathbf{R}^k$, with $\|\nu\| = \nu(\mathbf{R}^k)$ finite, and ν is concentrated on a set C of Lebesgue measure 0. Take $\delta > 0$ and a compact subset K of C such that $\nu(K) > \nu(C) - \delta$. Let $\tau(E) = \nu(E) - \nu(K \cap E)$ so that $\|\tau\| < \delta$.

For a measure λ let

$$D_\lambda(x) = \lim_{n\to\infty} \left(\sup_{0<r<1/n} \frac{\lambda(B(x,r))}{m(B(x,r))} \right).$$

For $x \notin K$ we have $D_\nu(x) = D_\tau(x)$ and so $D_\nu(x) \leq M_\tau(x)$. But, if $t > 0$, the set $R_t = \{x : M_\tau(x) > t\}$ has measure at most $3^k t^{-1} \|\tau\| < 3^k t^{-1} \delta$. But $S_t = \{x : D_\nu(x) > t\} \subseteq K \cup R_t$, so $m(S_t) < 3^k t^{-1} \delta$. Since δ is arbitrary we get $m(S_t) = 0$ and so $D_\nu(x) = 0$ a.e., which gives

$$\lim_{r\to 0+} \frac{\mu_s(B(x,r))}{m(B(x,r))} = 0$$

for almost every x in A.

9.8 Theorem

Let A, B be open subsets of $\mathbf{R}^k$, and let $f : A \to B$ be a homeomorphism. Define a measure μ by $\mu(E) = m(f(A \cap E))$. Then there exist a non-negative function h and a measure ν, singular with respect to m, such that

$$\mu(E) = \int_E h dm + \nu(E).$$

The function h is finite a.e. Also, for almost every x we have

$$\frac{\mu(B_n)}{m(B_n)} \to h(x)$$

if B_n shrinks nicely to x. If f maps sets of measure zero to sets of measure zero then $\nu = 0$.

Proof. Taking compact X_n which expand to fill A, we see that μ is σ-finite. This gives h and ν as in the Lebesgue-Radon-Nikodym theorem. Since

$$\int_{X_n} h dm \leq \mu(X_n) < \infty,$$

we get h finite a.e.

9.9 Lemma

Let f, A, B be as in the previous theorem. If f is differentiable a.e. then $h = |J_f|$ a.e., in which J_f in the Jacobian determinant. If, in addition, f maps sets of measure zero to sets of measure zero, we have

$$\int_B g(y) dm(y) = \int_A g(f(x)) |J_f| dm(x)$$

for every non-negative measurable g.

Proof. Suppose that f is differentiable at x. Then

$$f(x+t) = f(x) + Mt + o(|t|), \quad t \to 0,$$

in which M is a matrix with determinant J_f. If $J_f = 0$ then $m(f(B(x,r)))/m(B(x,r)) \to 0$ as $r \to 0$. If M is non-singular and $\delta > 0$ we have

$$f(B(x,r)) \subseteq (1+\delta)M(B(x,r))$$

for sufficiently small r. Thus $h \leq |J_f|$. To get the reverse inequality, write $y = f(x), y+k = f(x+h), G = f^{-1}$, to get

$$G(y+k) = G(y) + M^{-1}y + o(|y|), \quad y \to 0.$$

Thus for small r, we get

$$G((1-\delta)M(B(x,r))) \subseteq M^{-1}M((B(x,r))) = B(x,r)$$

and so

$$(1-\delta)M(B(x,r)) \subseteq f(B(x,r)).$$

10 The Riesz representation theorem

10.1 Theorem

Let X be a locally compact metric space, and let L be a positive ($f \geq 0$ implies $L(f) \geq 0$) linear functional on the space $C_0 = C_0(X)$ of continuous real-valued functions with compact support. Then there exist a σ-algebra Π and a non-negative measure μ such that:

(i) $L(f) = \int_X f d\mu$ for every $f \in C_0$;

(ii) every Borel set of X is an element of Π;

(iii) $\mu(K) < \infty$ for every compact subset K of X;

(iv) if $E \in \Pi$ then $\mu(E)$ is the infimum of $\mu(V)$ over all open V containing E;

(v) if E is open or $\mu(E)$ is finite, then $\mu(E)$ is the supremum of $\mu(K)$ over all compact $K \subseteq E$;

(vi) if $A \subseteq B$ and $\mu(B) = 0$ then $A \in \Pi$ and $\mu(A) = 0$;

(vii) the μ satisfying (i) to (vi) is unique;

(viii) if V is open then $\mu(V)$ is the supremum of $L(f)$ over all C_0 functions $f : X \to [0, 1]$ such that $f = 0$ off V.

Note that simplifications (and a simpler proof!) arise if X is compact. In this case the function which is identically 1 is C_0, and $\mu(X) = L(1) < \infty$.

Proof. We first prove (vii). So assume that μ_1 and μ_2 both satisfy (i) to (vi). Let K be compact. Let $\delta > 0$. Using (iv), choose open V with $K \subseteq V$ and $\mu_2(V) < \mu_2(K)+\delta$. Using Urysohn's lemma, choose $f \in C_0$ with $f = 1$ on K and $f = 0$ off V. Then by (i),

$$\mu_1(K) \leq \int_X f d\mu_1 = \int_X f d\mu_2 \leq \mu_2(V) \leq \mu_2(K) + \delta.$$

Thus $\mu_1(K) \leq \mu_2(K)$ and hence they are equal. By (v) we have $\mu_1(E) = \mu_2(E)$ for every open E and by (iv) this holds for all E.

10.1.1 Construction of μ and Π

For open V define $\mu(V)$ to be the supremum of $L(f)$ over all $f \in C_0$ such that $0 \leq f \leq 1$ and $f = 0$ off V. Obviously, for open V_j, if $V_1 \subseteq V_2$ this gives $\mu(V_1) \leq \mu(V_2)$.

For any E we now define $\mu(E)$ to be the infimum of $\mu(V)$ over all open V with $E \subseteq V$. If E is open this agrees with the previous definition, and property (iv) is immediate.

Obviously if $A \subseteq B$ then $\mu(A) \leq \mu(B)$.

Next, let Π_1 be the set of all $E \subseteq X$ such that $\mu(E) < \infty$ and $\mu(E)$ is the supremum of $\mu(K)$ over all compact $K \subseteq E$.

Finally, let Π be the set of all $E \subseteq X$ such that $E \cap L \in \Pi_1$ for every compact L.

Clearly if $\mu(E) = 0$ then $\mu(E \cap K) = 0$ for every compact K. Thus $E \cap K \in \Pi_1$ and $E \in \Pi$.

Note that if $f \geq g$ and $f, g \in C_0$ then $L(f) = L(g) + L(f - g) \geq L(g)$.

10.1.2 Step I

For any $E_j \subseteq X$ we have $\mu(\bigcup_j E_j) \le \sum_j \mu(E_j)$.

We first show that $\mu(V_1 \cup V_2) \le \mu(V_1) + \mu(V_2)$ for open V_j. Let $g \in C_0$ with $0 \le g \le 1$ and $g = 0$ off $V_1 \cup V_2$. By the lemma on partitions of unity we can find C_0 functions $h_j : X \to [0,1]$ such that $h_j = 0$ off V_j and $h_1 + h_2 = 1$ on the compact support of g. Thus $g \le h_1 + h_2$ and

$$L(g) \le L(h_1 + h_2) = L(h_1) + L(h_2) \le \mu(V_1) + \mu(V_2).$$

To prove Step I, note that the result is obvious if $\mu(E_j) = \infty$ for some j. Assuming all $\mu(E_j)$ are finite, take $\delta > 0$ and open V_j such that

$$E_j \subseteq V_j, \quad \mu(V_j) < \mu(E_j) + \delta 2^{-j}.$$

Let V be the union of the V_j. Then V is open. If $f : X \to [0,1]$ is in C_0 with $f = 0$ off V, then the compact support of f is a subset of $V_1 \cup V_2 \cup \ldots V_n$ for some n. So

$$L(f) \le \mu(V_1 \cup \ldots \cup V_n) \le \sum_{j=1}^{n} \mu(V_j).$$

Thus $\mu(V) \le \sum_j \mu(V_j) \le \delta + \sum_j \mu(E_j)$. Hence $\mu(\bigcup_j E_j) \le \mu(V) \le \delta + \sum_j \mu(E_j)$.

10.1.3 Step II

If K is compact then $\mu(K)$ is the infimum of $L(f)$ over all C_0 functions $f : X \to [0,1]$ such that $f = 1$ on K. Thus $\mu(K) < \infty$, proving (iii). Also (obviously, by monotonicity) $K \in \Pi_1$.

Take f as in the statement, and $t \in (0,1)$. Then the set $V = \{x : f(x) > t\}$ is open, and $K \subseteq V$. Thus

$$\mu(K) \le \mu(V) \le t^{-1} L(f).$$

This gives $\mu(K) \le L(f)$.

Now let $\varepsilon > 0$ and take open V with $\mu(V) < \mu(K) + \varepsilon$. Choose a C_0 function $h : X \to [0,1]$ such that $h = 1$ on K and $h = 0$ off V. Then

$$L(h) \le \mu(V) < \mu(K) + \varepsilon.$$

This proves Step II.

10.1.4 Step III

If V is open then $\mu(V)$ is the supremum of $\mu(K)$ over all compact $K \subseteq V$. If, in addition, $\mu(V)$ is finite then $V \in \Pi_1$.

Proof. By monotonicity, we only need exhibit compact $K \subseteq V$ with $\mu(K)$ arbitrarily close to $\mu(V)$. Let $t < \mu(V)$. Then there exists a C_0 function $f : X \to [0,1]$ such that $t < L(f) \le \mu(V)$. Let K be the compact support of f. If W is open and $K \subseteq W$ we have $f = 0$ off W and so $L(f) \le L(W)$. This is true for every such W so $L(f) \le \mu(K)$. Hence $\mu(K) > t$.

10.1.5 Step IV

Let E_j be pairwise disjoint elements of Π_1, with union E. Then $\mu(E) = \sum_j \mu(E_j)$. If $\mu(E) < \infty$ then $E \in \Pi_1$.

Proof. We show first that if K_1, K_2 are compact and disjoint then $\mu(K_1 \cup K_2) = \mu(K_1) + \mu(K_2)$. Since X is Hausdorff, K_2 is closed, so there is a C_0 function f such that $f : X \to [0,1]$ with $f = 1$ on K_1 and $f = 0$ on K_2. Take $\delta > 0$. Then we know by Step II that there is a C_0 function $g : X \to [0,1]$ with $g = 1$ on $K_1 \cup K_2$ and $L(g) < \mu(K_1 \cup K_2) + \delta$. But $fg = 1$ on K_1 and $(1-f)g = 1$ on K_2. Using Step II again,

$$\mu(K_1 \cup K_2) + \delta > L(g) = L(fg) + L((1-f)g) \geq \mu(K_1) + \mu(K_2).$$

Using Step I and the fact that δ is arbitrary, we get $\mu(\bigcup K_j) = \sum \mu(K_j)$ for any finite collection of pairwise disjoint compact K_j.

To prove Step IV, note that by definition of Π_1 we know that $\mu(E_j)$ is finite for every j. Since $E_j \in \Pi_1$ we can take compact $H_j \subseteq E_j$ with $\mu(H_j) > \mu(E_j) - \delta 2^{-j}$. Thus

$$\mu(E) \geq \mu(\bigcup_{j=1}^{n} H_j) = \sum_{j=1}^{n} \mu(H_j) \geq -\delta + \sum_{j=1}^{n} \mu(E_j).$$

Using Step I again, we get the first assertion of Step IV.

Now suppose that $\mu(E) < \infty$. Take $\delta > 0$ and form H_j as before. We have $\mu(E) = \sum_j \mu(E_j)$ and so, for some N,

$$\mu(E) \leq \delta + \sum_{j=1}^{N} \mu(E_j) \leq 2\delta + \sum_{j=1}^{N} \mu(H_j) = 2\delta + \mu(\bigcup_{j=1}^{N} H_j).$$

Thus $E \in \Pi_1$.

10.1.6 Step V

Let $E \in \Pi_1$ and $\delta > 0$. Then there exist compact K and open V such that $\mu(V \backslash K) < \delta$ and $K \subseteq E \subseteq V$.

Proof. Choose compact K, open V with $K \subseteq E \subseteq V$ such that

$$\mu(V) - \delta/2 < \mu(E) < \mu(K) + \delta/2.$$

(The first is by definition of μ, the second by definition of Π_1). Now $V \backslash K$ is in Π_1, by Step III, and so is K, by Step II. So Step IV gives

$$\mu(K) + \mu(V \backslash K) = \mu(V) < \mu(K) + \delta.$$

10.1.7 Step VI

Let $A, B \in \Pi_1$. Then $A \backslash B, A \cup B, A \cap B$ are all in Π_1.

Proof. Take $\delta > 0$ and open V_j, compact K_j, such that $K_1 \subseteq A \subseteq V_1$ and $K_2 \subseteq B \subseteq V_2$ and $\mu(V_j \backslash K_j) < \delta$.

But

$$A \backslash B \subseteq V_1 \backslash K_2 \subseteq (V_1 \backslash K_1) \cup (K_1 \backslash V_2) \cup (V_2 \backslash K_2).$$

To see this, suppose $x \in V_1 \backslash K_2$. If x is in V_2 or not in K_1 then it is in the third or first set of the last three. Otherwise it is in $K_1 \backslash V_2$.

Since $K = K_1 \backslash V_2 = K_1 \cap V_2^c \subseteq A \cap B^c = A \backslash B$ is compact and $\mu(A \backslash B) < \mu(K) + 2\delta$ we get $A \backslash B \in \Pi_1$.

Since $A \cup B$ is the disjoint union of $A \backslash B$ and B, we get $A \cup B \in \Pi_1$, by Step IV. Also $A \backslash (A \backslash B) = A \backslash (A \cap B^c) = A \cap (A^c \cup B) = A \cap B$ is in Π, by the first part.

10.1.8 Step VII

Π *is a σ-algebra containing all Borel subsets of X.*

Proof. Let K be compact and let $A \in \Pi$. Then $A \cap K \in \Pi_1$, by definition of Π. But $K \cap A^c = K \cap (A^c \cup K^c) = K \backslash (A \cap K)$ and this is in Π_1, by Step VI. So $A^c \in \Pi$.

Now suppose that A_j are all in Π, with union A. Then $B_j = A_j \cap K \in \Pi_1$. Define

$$C_1 = B_1, \quad C_{n+1} = B_{n+1} \backslash B_n.$$

Then each C_n is in Π_1, by Step VI. So the union of the C_j is in Π_1, by Step IV. But $\bigcup C_j = \bigcup B_j = A \cap K$. So $A \in \Pi$.

Now note that if C is closed then $C \cap K$ is compact and in Π_1, by Step II. So $C \in \Pi$.

10.1.9 Step VIII

If $E \in \Pi$ and $\mu(E) < \infty$ then $E \in \Pi_1$. The converse is also true.

Note that together with Step III this proves (v).

Proof. If $E \in \Pi_1$ then $\mu(E) < \infty$ by definition and by II and VI we have $E \cap K \in \Pi_1$ for every compact K, so $E \in \Pi$.

Now suppose $E \in \Pi$ and $\mu(E) < \infty$. Take $\delta > 0$ and open V with $E \subseteq V$ and $\mu(V) < \infty$. Thus $V \in \Pi_1$, by III. Take compact $K \subseteq V$ with $\mu(V \backslash K) < \delta$. Since $E \in \Pi$ we have $E \cap K \in \Pi_1$ and so there is a compact $H \subseteq E \cap K$ with

$$\mu(E \cap K) < \mu(H) + \delta.$$

Thus, using Step I,

$$\mu(E) \leq \mu(E \backslash K) + \mu(E \cap K) \leq \mu(V \backslash K) + \mu(E \cap K) \leq 2\delta + \mu(H).$$

Thus $E \in \Pi_1$.

10.1.10 Step IX

μ is a measure on Π.

Proof. Suppose that E is the union of pairwise disjoint elements of Π. We need to show that $\mu(E) = \sum \mu(E_j)$. This is obvious if $\mu(E_j) = \infty$ for any j. If all $\mu(E_j)$ are finite, then each E_j is in Π_1 and the result follows from Step IV.

10.1.11 Step X

We have $L(f) = \int_X f d\mu$ for every $f \in C_0$.

Proof. It suffices to show that $L(f) \leq \int_X f d\mu$, because then we get $-L(f) = L(-f) \leq \int_X -f d\mu = -\int_X f d\mu$.

Let $f \in C_0$ have compact support K. Let $f(K) \subseteq [a,b]$. Choose y_j with $y_0 < a < y_1 < \ldots < y_n = b$ and $y_n - y_{n-1} < \delta$. Then $E_j = \{x \in K : y_{j-1} < f(x) \leq y_j\}$ is a Borel set.

Choose open V_j with $E_j \subseteq V_j$ and $\mu(V_j) < \mu(E_j) + \delta/n$. Since $f(x) < y_j + \delta$ on an open set containing E_j we can do this so that $f(x) < y_j + \delta$ on V_j. By the lemma on partitions of unity there exist C_0 functions h_j such that $h_j = 0$ off V_j and $h_1 + \ldots h_n = 1$ on K. Then $f = f\sum_{j=1}^n h_j$ and

$$\mu(K) \leq L(\sum_{j=1}^n h_j) = \sum_{j=1}^n L(h_j),$$

using Step II. Now

$$L(f) = \sum_{j=1}^n L(h_j f) \leq \sum_{j=1}^n L((y_j + \delta)h_j) =$$

$$= \sum_{j=1}^n (|a| + y_j + \delta)L(h_j) - |a|\sum_{j=1}^n L(h_j) \leq$$

$$\leq -|a|\mu(K) + \sum_{j=1}^n (|a| + y_j + \delta)L(h_j),$$

since $f < y_j + \delta$ on V_j. This gives, since $|a| + y_j \geq 0$ for $j \geq 1$ and $L(h_j) \leq \mu(V_j)$,

$$L(f) \leq -|a|\mu(K) + \sum_{j=1}^n (|a| + y_j + \delta)(\mu(E_j) + \delta/n) = \sum_{j=1}^n (y_j + \delta)\mu(E_j) + \delta|a| + \delta^2 + \sum_{j=1}^n y_j\delta/n.$$

But

$$\sum_{j=1}^n (y_j + \delta)\mu(E_j) \leq \sum_{j=1}^n \int_{E_j} f + 2\delta d\mu = \int_X f d\mu + 2\delta\mu(K),$$

since $f > y_j - \delta$ on E_j. Also,

$$\sum_{j=1}^n y_j\delta/n \leq b\delta,$$

and so the result follows since δ is arbitrary.