KØBENHAVNS UNIVERSITET

MATEMATISK INSTITUT

Selected topics from measure theory

September 1992

Christian Berg

Selected topics from measure theory

Contents

Preface

The present notes should be seen as a continuation of the basic measure theory course in 2MA, which is given for second year students. The intention was to attract some of these students who have become fascinated by measure theory, but on the other hand also more advanced students followed the course. Consequently I could not assume very much knowledge of functional analysis, and in particular I could not give the beautiful short proof of Liapounov's theorem due to Lindenstrauss, but followed Halmos 1948 paper from Bull. A.M.S. I want to thank Søren Eilers for having prepared careful seminars covering the two appendices. One of the reasons for including Liapounov's theorem was that a simple proof of the inequality in §4 could be based on it.

Copenhagen June 1991, Christian Berg

§1. Complex measures

Let (X, E) be a measurable space, i.e. a set X equipped with a σ-algebra E of subsets of X.

DEFINITION 1.1: A complex measure μ on (X, E) is a function $\mu : \mathsf{E} \to \mathbb{C}$ such that

$$\mu(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \mu(E_n) \tag{1}$$

for any sequence $E_1, E_2, \dots \in \mathsf{E}$ of pairwise disjoint sets.

For a complex measure μ we necessarily have $\mu(\emptyset) = 0$.

Since the value of the left-hand side of (1) is a complex number, it is tacitly assumed that the series on the right-hand side converges, and its sum is equal to $\mu(\bigcup_1^\infty E_n)$. If $\sigma : \mathbb{N} \to \mathbb{N}$ is any permutation (i.e. a bijection) then (1) implies that the rearranged series

$$\sum_{n=1}^{\infty} \mu(E_{\sigma(n)})$$

converges to $\mu(\bigcup_{n=1}^\infty E_{\sigma(n)}) = \mu(\bigcup_{n=1}^\infty E_n)$. We express this by saying that the series in (1) is *unconditionally convergent.*

We recall the classical fact about an infinite series with complex members: *It is absolutely convergent if and only if it is unconditionally convergent.* We obtain this as a corollary later.

An ordinary finite measure is of course a complex measure, but a measure μ with infinite total mass is not a complex measure. Measures from 2 MA will often be called positive measures.

Given a complex measure μ, we define a new set function $|\mu|$ which turns out to be a positive measure:

$$|\mu|(E) := \sup\{\sum_{n=1}^{\infty} |\mu(E_n)| \,\big|\, (E_n) \text{ partition of } E\}, \quad E \in \mathsf{E}$$

where "(E_n) partition of E" means that $E_1, E_2, \dots$ is a sequence of pairwise disjoint sets from E such that $\bigcup_1^\infty E_n = E$.

The measure $|\mu|$ is called the *total variation (measure)* of μ, or the *absolute value* of μ. □

THEOREM 1.2. *The total variation $|\mu|$ of a complex measure μ is a measure.*

PROOF: Clearly $|\mu|(\emptyset) = 0$ and $|\mu|(E) \le |\mu|(F)$ when $E \subseteq F$.

Let (E_i) be a sequence of pairwise disjoint sets from E with $E = \bigcup_1^\infty E_i$.

Let us first prove

$$\sum_{i=1}^{\infty} |\mu|(E_i) \le |\mu|(E)\,. \tag{2}$$

For this we can assume $|\mu|(E) < \infty$. Let $\varepsilon > 0$ be given and let us for each $i \in \mathbb{N}$ choose a partition $(A_{ij})_{j\ge 1}$ of E_i such that

$$|\mu|(E_i) - \frac{\varepsilon}{2^i} < \sum_{j=1}^{\infty} |\mu(A_{ij})|\,.$$

Then $(A_{ij}), i, j = 1, 2, \dots$ is a partition of E and summing over i we find

$$\sum_{i=1}^{\infty} |\mu|(E_i) - \varepsilon \le \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} |\mu(A_{ij})| \le |\mu|(E)\,,$$

and (2) follows since $\varepsilon > 0$ was arbitrary.

To prove equality in (2) we choose an arbitrary partition (A_i) of E. For fixed i, $(A_i \cap E_j)_{j\ge 1}$ is a partition of A_i, and for fixed j, $(A_i \cap E_j)_{i \ge 1}$ is a partition of E_j. Therefore

$$\begin{aligned}\sum_{i=1}^{\infty} |\mu(A_i)| &= \sum_{i=1}^{\infty} \Big|\sum_{j=1}^{\infty} \mu(A_i \cap E_j)\Big| \\ &\le \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} |\mu(A_i \cap E_j)| \\ &= \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} |\mu(A_i \cap E_j)| \\ &\le \sum_{j=1}^{\infty} |\mu|(E_j)\,,\end{aligned}$$

and taking supremum over all partitions (A_i) of E we get

$$|\mu|(E) \le \sum_{j=1}^{\infty} |\mu|(E_j)\,,$$

showing equality in (2). $\square$

LEMMA 1.3. *If $z_1, \dots, z_N \in \mathbb{C}$ then there is a subset S of $\{1, \dots, N\}$ for which*

$$\Big|\sum_{k\in S} z_k\Big| \ge \frac{1}{\pi}\sum_{k=1}^{N} |z_k|\,.$$

PROOF: Write $z_k = |z_k|e^{i\alpha_k}$. For $-\pi \le \theta \le \pi$ let $S(\theta)$ be the set of $k \in \{1,\dots,N\}$ for which $\cos(\alpha_k - \theta) > 0$. Then

$$|\sum_{k\in S(\theta)} z_k| = |e^{-i\theta}\sum_{k\in S(\theta)} z_k| \ge \mathrm{Re}(\sum_{k\in S(\theta)} e^{-i\theta} z_k) = \sum_{k\in S(\theta)} |z_k|\cos(\alpha_k - \theta)$$
$$= \sum_{k=1}^{N} |z_k|\cos^+(\alpha_k - \theta)\,.$$

Note that $S(\theta)$ can be empty. The empty sum is 0 by definition. The function $\varphi(\theta) = \sum_{k=1}^{N} |z_k|\cos^+(\alpha_k - \theta)$ is continuous on $[-\pi,\pi]$ and attains its maximum at $\theta = \theta_0$. Put $S = S(\theta_0)$. Then

$$|\sum_{k\in S} z_k| \ge \max_{\theta\in[-\pi,\pi]} \varphi(\theta) \ge \frac{1}{2\pi}\int_{-\pi}^{\pi} \varphi(\theta)d\theta = \sum_{k=1}^{N} |z_k| \frac{1}{2\pi}\int_{-\pi}^{\pi} \cos^+(\alpha_k - \theta)d\theta\,,$$

but

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \cos^+(\alpha_k - \theta)d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} \cos^+(\theta)d\theta = \frac{1}{2\pi}\int_{-\pi/2}^{\pi/2} \cos\theta d\theta = \frac{1}{\pi}$$

so finally

$$|\sum_{k\in S} z_k| \ge \frac{1}{\pi}\sum_{k=1}^{N} |z_k|\,.$$

□

REMARK. The constant $1/\pi$ in Lemma 1.3 is best possible. See Bledsoe: Amer. Math. Monthly 77 (1970), 180-182.

THEOREM 1.4. *If μ is a complex measure on $(X, \mathbb{E})$, then $|\mu|$ is a finite measure, i.e. $|\mu|(X) < \infty$.*

PROOF: We first show the following statement (∗):

If $E \in \mathbb{E}$ is such that $|\mu|(E) = \infty$, then we can split E as $E = A \cup B$, $A, B \in \mathbb{E}$, $A \cap B = \emptyset$ such that $|\mu(A)| > 1$, $|\mu|(B) = \infty$.

In fact put $t = \pi(1 + |\mu(E)|)$. Since $|\mu|(E) = \infty$ there exists a partition (E_i) of E such that

$$\sum_{1}^{\infty} |\mu(E_i)| > t\,,$$

and we can choose $N \in \mathbb{N}$ such that

$$\sum_{1}^{N} |\mu(E_i)| > t\,.$$

By Lemma 1.3 with $z_i = \mu(E_i)$,$i = 1, \dots, N$ we find $S \subseteq \{i, \dots N\}$ such that

$$|\sum_{i \in S} \mu(E_i)| \geq \frac{1}{\pi} \sum_{1}^{N} |\mu(E_i)| > \frac{t}{\pi} = 1 + |\mu(E)| \,, \tag{3}$$

so with $A = \bigcup_{i \in S} E_i, B \in \bigcup_{i \notin S} E_i$ we certainly have $|\mu(A)| \geq 1$ and

$$|\mu(B)| = |\mu(E) - \mu(A)| \geq |\mu(A)| - |\mu(E)| > 1$$

by (3). Evidently at least one of $|\mu|(A)$ and $|\mu|(B)$ is ∞ since their sum is ∞ by Theorem 1.2.

Assume now that $|\mu|(X) = \infty$. Using $(*)$ we first split $X = A_1 \cup B_1$ with $|\mu(A_1)| > 1$, $|\mu|(B_1) = \infty$, then we split $B_1 = A_2 \cup B_2$ with $|\mu(A_2)| > 1$ $|\mu|(B_2) = \infty$. Continuing in this way, we get a countable pairwise disjoint collection $(A_i)_{i \geq 1}$ with $|\mu(A_i)| > 1$. The countable additivity of μ implies

$$\mu(\bigcup_{1}^{\infty} A_i) = \sum_{1}^{\infty} \mu(A_i) \,,$$

but the series cannot converge since each term is of absolute value > 1. This contradiction shows that $|\mu|(X) < \infty$. □

THEOREM 1.5. *Let μ be a complex measure. Then $|\mu|$ is the smallest positive measure σ satisfying*

$$|\mu(E)| \leq \sigma(E) \text{ for } E \in \mathsf{E} \,.$$

PROOF: We clearly have $|\mu(E)| \leq |\mu|(E)$ for $E \in \mathsf{E}$ since $E = E \cup \emptyset \cup \emptyset \cup \dots$ is a partition of E.

Let σ be a positive measure on (X, E) satisfying

$$|\mu(E)| \leq \sigma(E) \text{ for } E \in \mathsf{E} \,. \tag{4}$$

We claim that $|\mu|(E) \leq \sigma(E)$ for $E \in \mathsf{E}$, and to see this it is enough to prove that

$$\sum_{1}^{\infty} |\mu(E_i)| \leq \sigma(E)$$

for any partition (E_i) of E, but this is clear since the left-hand side by (4) is majorized by

$$\sum_{1}^{\infty} \sigma(E_i) = \sigma(E) \,.$$

□

Corollary 1.6: Let μ be a complex measure and $E \in \mathsf{E}$. Then

$$|\mu|(E) = \sup\{\sum_{i=1}^{N} |\mu(E_i)| \mid N \in \mathbb{N}, \{E_1, \dots, E_N\} \text{ partition of } E\}$$
$$= \sup\{\sum_{i=1}^{N} |\mu(E_i)| \mid N \in \mathbb{N}, E_1, \dots, E_N \in \mathsf{E} \text{ pairwise disjoint subsets of } E\} ,$$

where "$\{E_1, \dots, E_N\}$ partition of E" means that $E_1, \dots, E_N \in \mathsf{E}$ are pairwise disjoint subsets of E with union E. □

Proof: For $\varepsilon > 0$ there exists a countable partition (E_i) of E such that

$$\sum_{1}^{\infty} |\mu(E_i)| > |\mu|(E) - \frac{\varepsilon}{2} ,$$

and since the series $\sum_1^\infty |\mu(E_i)|$ is convergent (*viz.* $\leq |\mu|(E) < \infty$), there exists $N \in \mathbb{N}$ such that

$$\sum_{N+1}^{\infty} |\mu(E_i)| < \frac{\varepsilon}{2} .$$

Then $E_1, \dots, E_N$ are pairwise disjoint subsets of E satisfying

$$\sum_{i=1}^{N} |\mu(E_i)| = \sum_{i=1}^{\infty} |\mu(E_i)| - \sum_{i=N+1}^{\infty} |\mu(E_i)| > |\mu|(E) - \frac{\varepsilon}{2} - \frac{\varepsilon}{2} ,$$

and

$\{E_1, \dots, E_N, \tilde{E}\}$ with $\tilde{E} = \bigcup_{N+1}^{\infty} E_i$ is a finite partition of E with

$$\sum_{i=1}^{N} |\mu(E_i)| + |\mu(\tilde{E})| > |\mu|(E) - \varepsilon .$$

□

From Lemma 1.3 we can also establish the result about "unconditional convergence ⇒ absolute convergence". (The converse implication is well-known). We prove an apparently stronger statement:

PROPOSITION 1.7. *Let (z_n) be a sequence of complex numbers such that $\sum_1^\infty z_{\sigma(n)}$ is convergent for any permutation $\sigma : \mathbb{N} \to \mathbb{N}$ with a sum $s_\sigma \in \mathbb{C}$. Then $\sum_1^\infty |z_n| < \infty$ and s_σ is independent of σ.*

PROOF: Assume $\sum_1^\infty |z_n| = \infty$. We first choose $N_1 \in \mathbb{N}$ such that

$$\sum_{j=1}^{N_1} |z_j| > \pi \,.$$

By Lemma 1.3 there exists $S_1 \subseteq \{1, \dots, N_1\}$ such that

$$|\sum_{j \in S_1} z_j| \geq \frac{1}{\pi} \sum_{j=1}^{N_1} |z_j| > 1 \,.$$

Since $\sum_{N_1+1}^\infty |z_j| = \infty$ we choose $N_2 > N_1$ such that $\sum_{N_1+1}^{N_2} |z_j| > \pi$ and (by Lemma 1.3) $S_2 \subseteq \{N_1 + 1, \dots, N_2\}$ such that

$$|\sum_{j \in S_2} z_j| > 1 \,.$$

Continuing in this way we get a sequence $S_1, S_2, \dots$ of pairwise disjoint and finite subsets of $\mathbb{N}$ such that

$$|\sum_{j \in S_n} z_j| > 1 \text{ for all } n \,. \tag{5}$$

If $\mathbb{N} \setminus \bigcup_1^\infty S_j = \{n_1 < n_2 < \dots\}$ we define a bijection $\sigma : \mathbb{N} \to \mathbb{N}$ by arranging $\mathbb{N}$ in the following way: $S_1, n_1, S_2, n_2, \dots$, where the elements in each S_n are ordered according to the standard order. (If $\mathbb{N} \setminus \bigcup_1^\infty S_j$ is a finite set $\{n_1 < n_2 < \cdots < n_k\}$ we arrange $\mathbb{N}$ by taking $S_1, n_1, \dots, S_{k-1}, n_{k-1}, S_k, n_k, S_{k+1}, \dots$). Now $\sum z_{\sigma(n)}$ is not convergent, because if it was, there would exist $N \in \mathbb{N}$ such that

$$\forall n \geq N \forall p \in \mathbb{N} : |\sum_{j=n+1}^{n+p} z_{\sigma(j)}| \leq \frac{1}{2} \,,$$

contradicting (5). □

EXAMPLES 1.8.

(a) Let $(z_n) \in \mathbb{C}$ be such that $\sum_1^\infty |z_n| < \infty$. We define a complex measure on the σ-algebra of all subsets of $\mathbb{N}$ by setting $\mu(\{n\}) = z_n$, $n \in \mathbb{N}$, so that

$$\mu(E) = \sum_{n \in E} z_n \text{ for } E \subseteq \mathbb{N} \,.$$

Clearly

$$|\mu|(E) = \sum_{n\in E} |z_n| \,.$$

(b) Let (X, E, μ) be a measure space with a positive measure μ and let $f : X \to \mathbb{C}$ be an integrable function. Then

$$\sigma(E) = \int_E f d\mu \;,\; E \in \mathsf{E}$$

is a complex measure by Lebesgue's dominated convergence theorem.

We claim that

$$|\sigma|(E) = \int_E |f| d\mu \quad , E \in \mathsf{E} \,. \tag{6}$$

PROOF: By Theorem 1.5 we have $|\sigma|(E) \le \int_E |f| d\mu$. Let f_1, f_2 be the real and imaginary part of f. Assume first that $E \in \mathsf{E}$ is a set such that $\mu(E) < \infty$ and that

$$a_1 \le f_1(x) \le b_1 \,,\; a_2 \le f_2(x) \le b_2 \text{ for } x \in E \,.$$

Then there exist $\xi_1 \in [a_1, b_1], \xi_2 \in [a_2, b_2]$ such that

$$\int_E f d\mu = (\xi_1 + i\xi_2)\mu(E) \,. \tag{7}$$

In fact,

$$a_1\mu(E) \le \int_E f_1 d\mu \le b_1\mu(E)$$

and therefore we can write $\int_E f_1 d\mu = \xi_1\mu(E)$ for some $\xi_1 \in [a_1, b_1]$ and similarly with f_2.

Assume next that $\mu(E) < \infty$ and let $\varepsilon > 0$ be given. We divide $\mathbb{C}$ in the countably many standard squares

$$S_{n,m} = \{z \in \mathbb{C} \,|\, n\varepsilon < \operatorname{Re} z \le (n+1)\varepsilon, m\varepsilon < \operatorname{Im} z \le (m+1)\varepsilon\} \,,$$

where $n, m \in \mathbb{Z}$, and define the partition of E

$$E_{n,m} = \{x \in E \,|\, f(x) \in S_{n,m}\} \,,\; n, m \in \mathbb{Z} \,.$$

Let $\xi_{n,m} \in \overline{S_{n,m}}$ be such that in accordance with (7)

$$\int_{E_{n,m}} f d\mu = \xi_{n,m}\mu(E_{n,m}) \,.$$

We then find

$$|\int_{E_{n,m}} f d\mu| = |\xi_{n,m}|\mu(E_{n,m}) ,$$

and therefore

$$\int_{E_{n,m}} |f - \xi_{n,m}| d\mu \leq \varepsilon\sqrt{2}\,\mu(E_{n,m}) .$$

This shows that

$$\begin{aligned}\int_E |f| d\mu - \sum_{n,m} |\int_{E_{n,m}} f d\mu| &= \sum_{n,m} \int_{E_{n,m}} (|f| - |\xi_{n,m}|) d\mu \\ &\leq \sum_{n,m} \int_{E_{n,m}} |f - \xi_{n,m}| d\mu \leq \varepsilon\sqrt{2} \sum_{n,m} \mu(E_{n,m}) = \varepsilon\sqrt{2}\,\mu(E)\end{aligned}$$

hence

$$\int_E |f| d\mu - |\sigma|(E) \leq \varepsilon\sqrt{2}\,\mu(E) ,$$

and (6) follows.

Finally let $E \in \mathbb{E}$ be arbitrary and define

$$E_0 = \{x \in E \,|\, f(x) = 0\} ,\ E_n = \{x \in E \,|\, |f(x)| \geq \frac{1}{n}\} ,\ n = 1, 2, \ldots .$$

Clearly

$$\sigma(A) = 0 \text{ for any } A \in \mathbb{E}, A \subseteq E_0$$

and

$$\infty > \int |f| d\mu \geq \int_{E_n} |f| d\mu \geq \frac{1}{n}\mu(E_n) ,$$

so

$$|\sigma|(E_0) = 0 ,\ |\sigma|(E_n) = \int_{E_n} |f| d\mu .$$

Since $E = E_0 \cup \bigcup_{n=1}^{\infty} E_n$ we get

$$\begin{aligned}|\sigma|(E) &= |\sigma|(E_0) + \lim_{n\to\infty} |\sigma|(E_n) \\ &= \lim_{n\to\infty} \int_{E_n} |f| d\mu = \int_E |f| d\mu .\end{aligned}$$

□

The set of complex measures on $(X, \mathbb{E})$ is denoted $M(X, \mathbb{E})$. This set is organized as a complex vector space under the operations

$$(\mu + \nu)(E) = \mu(E) + \nu(E)$$
$$(a\mu)(E) = a\mu(E)$$

for $\mu, \nu \in M(X, \mathsf{E})$, $a \in \mathbb{C}$ and $E \in \mathsf{E}$.

It is easy to verify that

$$|\mu + \nu| \leq |\mu| + |\nu|$$
$$\big| |\mu| - |\nu| \big| \leq |\mu - \nu|$$
$$|a\mu| = |a||\mu| \, .$$

We finally put

$$\|\mu\| = |\mu|(X) \, ,$$

and it is clear that $\| \cdot \|$ is a norm on $M(X, \mathsf{E})$, called the *total variation.*

LEMMA 1.9. *For $\mu \in M(X, \mathsf{E})$ and $E \in \mathsf{E}$ we have*

$$\sup_{E \in \mathsf{E}} |\mu(E)| \leq \|\mu\| \leq \pi \sup_{E \in \mathsf{E}} |\mu(E)| \, .$$

PROOF: We have $|\mu(E)| \leq |\mu|(E) \leq |\mu|(X) = \|\mu\|$, so the first inequality follows. Let $E_1, \ldots, E_n \in \mathsf{E}$ be pairwise disjoint. By Lemma 1.3 there exists $S \subseteq \{1, \ldots, n\}$ such that

$$\sum_{i=1}^{n} |\mu(E_i)| \leq \pi \Big| \sum_{i \in S} \mu(E_i) \Big| = \pi \Big| \mu\Big(\bigcup_{i \in S} E_i\Big) \Big|$$
$$\leq \pi \sup_{E \in \mathsf{E}} |\mu(E)| \, ,$$

and by Corollary 1.6 we get

$$|\mu|(X) \leq \pi \sup_{E \in \mathsf{E}} |\mu(E)| \, .$$

REMARK 1.10. A complex measure μ on (X, E) is a bounded function on E so we can consider $M(X, \mathsf{E})$ as a subspace of the vector space $B(\mathsf{E}, \mathbb{C})$ of bounded functions on E, and this is a Banach space under the uniform norm. Lemma 1.9 states that the restriction of the uniform norm to $M(X, \mathsf{E})$ is equivalent to the norm $\| \cdot \|$. In order to show that $M(X, \mathsf{E})$ is a Banach space, it is therefore sufficient to prove that $M(X, \mathsf{E})$ is a closed subspace of $B(\mathsf{E}, \mathbb{C})$ under the uniform norm.

THEOREM 1.11. *$M(X, \mathsf{E})$ is a Banach space under the total variation norm.*

PROOF: Let (μ_n) be a sequence of complex measures which converges uniformly to a function $\mu \in B(\mathsf{E}, \mathbb{C})$. We shall show that μ is a complex measure.

We first remark that μ is *finitely additive*: If $E_1, \ldots, E_m$ are pairwise disjoint sets from E then

$$\mu(\bigcup_{i=1}^{m} E_i) = \sum_{i=1}^{m} \mu(E_i)$$

because

$$\mu(\bigcup_{i=1}^{m} E_i) = \lim_{n\to\infty} \mu_n(\bigcup_{i=1}^{m} E_i) = \lim_{n\to\infty} (\sum_{i=1}^{m} \mu_n(E_i)) = \sum_{i=1}^{m} \mu(E_i) \ .$$

To see that μ is *countably additive* let (E_i) be a sequence of pairwise disjoint sets from E. For $\varepsilon > 0$ there exists n_0 such that

$$|\mu(A) - \mu_{n_0}(A)| \le \varepsilon \text{ for all } A \in \mathsf{E} \ .$$

By the finite additivity we get

$$\begin{aligned} &|\mu(\bigcup_{i=1}^{\infty} E_i) - \sum_{i=1}^{m} \mu(E_i)| = |\mu(\bigcup_{i=m+1}^{\infty} E_i)| \\ &\le |\mu(\bigcup_{i=m+1}^{\infty} E_i) - \mu_{n_0}(\bigcup_{i=m+1}^{\infty} E_i)| + |\mu_{n_0}(\bigcup_{i=m+1}^{\infty} E_i)| \\ &\le \varepsilon + |\mu_{n_0}|(\bigcup_{i=m+1}^{\infty} E_i) \ . \end{aligned}$$

The last term above tends to zero for $m \to \infty$ because $|\mu_{n_0}|$ is a finite positive measure. This shows that

$$|\mu(\bigcup_{i=1}^{\infty} E_i) - \sum_{i=1}^{m} \mu(E_i)| \le 2\varepsilon$$

for m sufficiently large, i.e. μ is a complex measure. □

Let us now specialize and consider a *real measure* μ on (X, E), i.e. a function $\mu : \mathsf{E} \to \mathbb{R}$ satisfying (1). Such measures are often called *signed measures* in contrast to positive measures.

If μ is a complex measure then $\mu_1 = \operatorname{Re}\mu$ and $\mu_2 = \operatorname{Im}\mu$ are real measures and

$$\mu = \operatorname{Re}\mu + i \operatorname{Im}\mu = \mu_1 + i\mu_2 \ . \tag{8}$$

Definition 1.12: For a real measure μ we define

$$\mu^+ = \frac{1}{2}(|\mu| + \mu)\ ,\ \mu^- = \frac{1}{2}(|\mu| - \mu)\ . \tag{9}$$

Since

$$-|\mu|(E) \le |\mu(E)| \le |\mu|(E) \text{ for } E \in \mathsf{E}$$

we see that μ^+ and μ^- are positive measures on (X, E). They are called the *positive* and *negative* part of μ.

Clearly

$$\mu = \mu^+ - \mu^-\ ,\ |\mu| = \mu^+ + \mu^-\ . \tag{10}$$

This formula shows that any real measure is a difference of positive measures. The decomposition $\mu = \mu^+ - \mu^-$ is called the *Jordan decomposition.* (C. Jordan, french mathematician, 1838-1921). There are of course many decompositions of μ as difference of finite positive measures because we can add any finite positive measure σ to μ^+ and to μ^-. We shall later see that the Jordan decomposition is the smallest such decomposition in the sense that if $\mu = \lambda_1 - \lambda_2$ with positive measures λ_1, λ_2, then $\lambda_1 \ge \mu^+$, $\lambda_2 \ge \mu^-$. This shows that $\sigma = \lambda_1 - \mu^+ = \lambda_2 - \mu^-$ is a positive measure such that $\lambda_1 = \mu^+ + \sigma$, $\lambda_2 = \mu^- + \sigma$.

It is clear that the set of real measures is a real vector space and a Banach space under the norm $\|\mu\|$. □

Appendix 1. The best constant in Lemma 1.3

Søren Eilers

Lemma 1.3 tells us that for all finite sets of complex numbers $z_1, \cdots, z_N$ there is a subset

$$S \subseteq \{1, \cdots, N\}$$

so that

$$|\sum_{k \in S} z_k| \geq \frac{1}{\pi} \sum_{k=1}^{N} |z_k| \,.$$

Let us realize that $1/\pi$ is the best possible constant in this inequality, i.e.

$$\forall \varepsilon > 0 \ \exists z_1, \cdots, z_N \in \mathbb{C} \ \forall S \subseteq \{1, \cdots, N\} : \frac{|\sum_{k \in S} z_k|}{\sum_{k=1}^{N} |z_k|} < \frac{1}{\pi} + \varepsilon \,. \qquad (*)$$

We can assume that all $z_k \neq 0$.

When S is chosen so that $|\sum_{k \in S} z_k|$ is maximal, S is of the form

$$S_\xi = \{k \in \{1, \cdots, N\} \mid (z_k|\xi) > 0\}$$

where $(z|w) = \mathrm{Re}(z\overline{w})$ for a suitable unit vector ξ in $\mathbb{C}$. Indeed, when

$$\sum_{k \in S} z_k = r\xi \,, \; r > 0 \,, \; \|\xi\| = 1$$

and S is optimal, we have

$$(z_k \mid \xi) > 0 \Rightarrow k \in S$$

since the Cauchy-Schwarz inequality gives us, for $(z_k \mid \xi) > 0$

$$|r\xi + z_k| \geq |(r\xi + z_k|\xi)| = |r + (z_k|\xi)| > r \,.$$

If $k \notin S$ then $r\xi + z_k$ is the sum corresponding to $S \cup \{k\}$ contradicting the maximality of the sum corresponding to S.

When $(z_k \mid \xi) < 0$,

$$|r\xi - z_k| \geq |(r\xi - z_k|\xi)| = |r - (z_k|\xi)| > r \,,$$

then $k \in S$ leads to a contradiction.

When $(z_k \mid \xi) = 0$ we have by the Pythagorean theorem

$$|r\xi \pm z_k|^2 = r^2 + |z_k|^2 \,.$$

The assumption $|z_k| \neq 0$ shows that both $k \in S$ and $k \notin S$ lead to a contradiction, as above.

For each N let us look at the roots in $z^{2N} = 1$,

$$e^{2\pi i k/2N} \;;\; k = 0, \cdots, 2N-1 \,.$$

From the discussion above we see that a subset of $\{0, \cdots, 2N-1\}$ yielding a sum with maximum modulus must correspond to N consecutive roots. Since

$$|\sum_{k=k_0}^{k_0+N-1} e^{2\pi i k/2N}| = |e^{2\pi i k_0/2N}||\sum_{k=0}^{N-1} e^{2\pi i k/2N}| = |\sum_{k=0}^{N-1} e^{2\pi i k/2N}| \,,$$

the maximum modulus is obtained by adding the first N roots.

Thus, showing that

$$\frac{|\sum_{k=0}^{N-1} e^{2\pi i k/2N}|}{\sum_{k=0}^{2N} |e^{2\pi i k/2N}|} \to \frac{1}{\pi} \,,\; N \to \infty$$

gives $(*)$. But since

$$\frac{|\sum_{k=0}^{N-1} e^{2\pi i k/2N}|}{\sum_{k=0}^{2N} |e^{2\pi i k/2N}|} = \frac{1}{2N}|\sum_{k=0}^{N-1} e^{i(\frac{\pi k}{N})}| = \frac{1}{2\pi}|\sum_{k=0}^{N-1} \frac{\pi}{N} e^{i(\frac{\pi k}{N})}|$$

and the sum to the right obviously is a mean sum for

$$e^{i\varphi} : [0, \pi] \to \mathbb{C}$$

this follows from the basic calculus observation (1MA IV.1.4) that mean sums converge to integrals when the maximal distance between the consecutive points tends to zero, and

$$|\int_0^\pi e^{i\varphi} d\varphi| = |[-ie^{i\varphi}]_0^\pi| = |i + i| = 2 \,.$$

§2. Absolute continuity

If μ is a positive measure on (X, E) and $f \in \mathcal{L}_1(\mu)$ we have seen that

$$\lambda(E) = \int_E f d\mu \ , \ E \in \mathsf{E} \tag{1}$$

is a complex measure.

We remark that if $\mu(E) = 0$ then $\lambda(E) = 0$ and this leads to an important new concept:

DEFINITION 2.1 Let μ be a positive measure and let λ be an arbitrary measure on (X, E), i.e. λ is either a complex measure or a positive measure.

We say that λ is *absolutely continuous* with respect to μ and write $\lambda \ll \mu$ if $\lambda(E) = 0$ for every $E \in \mathsf{E}$ with $\mu(E) = 0$, i.e.

$$\lambda \ll \mu \Leftrightarrow \forall E \in \mathsf{E}\big(\mu(E) = 0 \Rightarrow \lambda(E) = 0\big) \ . \tag{2}$$

The following theorem explains why the word "continuity" is used in connection with the relation $\lambda \ll \mu$.

THEOREM 2.2. *Suppose μ is a positive measure and λ is a complex measure on (X, E). Then the following conditions are equivalent*

(i) $\lambda \ll \mu$.

(ii) $\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall E \in \mathsf{E}\big(\mu(E) < \delta \Rightarrow |\lambda(E)| < \varepsilon\big)$.

PROOF: Suppose (ii) holds and that $\mu(E) = 0$. Then $|\lambda(E)| < \varepsilon$ for every $\varepsilon > 0$, hence $\lambda(E) = 0$.

Suppose next that (ii) does not hold. Then there exists an $\varepsilon > 0$ and there exist sets $E_n \in \mathsf{E}$, $n = 1, 2, \cdots$ such that

$$\mu(E_n) < 2^{-n} \ \text{ but } \ |\lambda(E_n)| \geq \varepsilon \ .$$

We define

$$A_n = \bigcup_{i=n}^{\infty} E_i \ , \ A = \bigcap_{n=1}^{\infty} A_n$$

and get

$$\mu(A_n) \leq \sum_{i=n}^{\infty} \mu(E_i) < 2^{-n+1} \ .$$

Since $A_1 \supseteq A_2 \supseteq \cdots$ we have $\mu(A) = 0$ and

$$|\lambda|(A) = \lim_{n \to \infty} |\lambda|(A_n) \ ,$$

which is $\geq \varepsilon$ because $|\lambda|(A_n) \geq |\lambda|(E_n) \geq |\lambda(E_n)| \geq \varepsilon$. On the other hand if (i) holds, then we have $\lambda(B) = 0$ for any $B \in \mathbb{E}, B \subseteq A$ and hence $|\lambda|(A) = 0$, which is a contradiction. □

REMARK 2.3. (i) does not imply (ii) if λ is a positive measure with infinite total mass. For instance, let μ be Lebesgue measure on $]0, \infty[$ and put

$$\lambda(E) = \int_E \frac{dt}{t} \text{ for } E \in \mathbb{B}(]0, \infty[) .$$

Then (i) holds but $\lambda(]0, \varepsilon]) = \infty$ for all $0 < \varepsilon$.

The main theorem of this section is the *Radon-Nikodym* theorem which essentially tells us that if $\lambda \ll \mu$ then λ has the form (1). We need however an extra condition on μ, namely σ-finiteness: A measure μ is called *σ-finite* if there exists a partition (E_n) of X such that $\mu(E_n) < \infty$ for $n = 1, 2, \cdots$.

THEOREM 2.4. *(Radon-Nikodym) Let λ be a complex (resp. positive) measure on $(X, \mathbb{E})$ and let μ be a σ-finite positive measure. Then the following conditions are equivalent:*

(i) $\lambda \ll \mu$.

(ii) *There exists $f \in \mathcal{L}_1(\mu)$ (resp. $f \in \mathcal{M}^+(X, \mathbb{E})$) such that*

$$\lambda(E) = \int_E f d\mu \text{ for } E \in \mathbb{E} .$$

Before we give the proof we need some lemmas.

LEMMA 2.5. *Let μ be a positive measure on $(X, \mathbb{E})$. Then μ is σ-finite if and only if there exists $f \in \mathcal{L}_1(\mu)$ satisfying $0 < f(x) < \infty$ for all $x \in X$.*

PROOF: If μ is σ-finite and (E_n) is a partition of X with $\mu(E_n) < \infty$, then

$$f = \sum_{n=1}^{\infty} \varepsilon_n 1_{E_n}$$

satisfies the required conditions if $\varepsilon_n > 0$, $\sum \varepsilon_n < \infty$ and $\sum \varepsilon_n \mu(E_n) < \infty$, and this is easily achieved.

Conversely, if f is integrable and $0 < f < \infty$ then

$$A_n = \{x \in X \mid f(x) \geq \frac{1}{n}\}, \; n = 1, 2, \cdots$$

is an increasing sequence of sets from E with $\bigcup_1^\infty A_n = X$ and

$$\infty > \int f d\mu \geq \int_{A_n} f d\mu \geq \frac{1}{n}\mu(A_n)$$

so $\mu(A_n) < \infty$. From this it is easy to see that μ is σ-finite. □

LEMMA 2.6. *Let σ and τ be positive measures on (X, E) with $\sigma(X) < \tau(X) < \infty$.*

Then there exists $Y \in \mathsf{E}$ such that

(a) $\sigma(Y) < \tau(Y)$.

(b) $\sigma|Y \leq \tau|Y$, *i.e.* $\sigma(E) \leq \tau(E)$ *for all* $E \in \mathsf{E}, E \subseteq Y$.

PROOF: Let $\delta = \tau - \sigma$. Then δ is a real measure with $-\sigma(X) \leq \delta(E) \leq \tau(X)$ for $E \in \mathsf{E}$. We shall construct two sequences (A_n), (X_n), $n \geq 0$ from E inductively. We put $A_0 = \emptyset$, $X_0 = X = X \setminus A_0$. If $A_0, \cdots, A_n$ and $X_0, \cdots X_n$ are constructed we consider the quantity

$$\alpha_n := \inf\{\delta(A) \mid A \in \mathsf{E},\, A \subseteq X_n\}\,. \tag{3}$$

Since $A = \emptyset$ occurs above $\alpha_n \leq \delta(\emptyset) = 0$. If $\alpha_n = 0$ we define $A_{n+1} = \emptyset$, $X_{n+1} = X_n \setminus A_{n+1} (= X_n)$. If $\alpha_n < 0$ are choose $A_{n+1} \in \mathsf{E}$, $A_{n+1} \subseteq X_n$ such that

$$\delta(A_{n+1}) \leq \frac{1}{2}\alpha_n\,,$$

which is possible by (3). Define $X_{n+1} = X_n \setminus A_{n+1}$.

The sequence (A_n) consists of pairwise disjoint sets, so the series $\sum_1^\infty \delta(A_n)$ is convergent and in particular $\delta(A_n) \to 0$. This implies $\lim_{n\to\infty} \alpha_n = 0$. The sequence (X_n) is decreasing so for

$$Y = \bigcap_0^\infty X_n \in \mathsf{E}$$

we find

$$\begin{aligned}\delta(Y) = \tau(Y) - \sigma(Y) &= \lim_{n\to\infty} \left(\tau(X_n) - \sigma(X_n)\right)\\ &= \lim_{n\to\infty} \delta(X_n)\,.\end{aligned}$$

Using that $\delta(A_{n+1}) \leq 0$ for all n we have

$$\delta(X_{n+1}) = \delta(X_n) - \delta(A_{n+1}) \geq \delta(X_n)$$

and hence

$$\delta(Y) = \lim_{n\to\infty} \delta(X_n) \geq \delta(X_0) = \delta(X)$$

which is > 0 by assumption, so (a) is verified. To verify (b) let $E \in \mathbb{E}$, $E \subseteq Y$. Then $E \subseteq X_n$ for all n so by (3) we have $\alpha_n \leq \delta(E)$. Since $\alpha_n \to 0$ we get (b). □

PROOF OF THEOREM 2.4: Only the implication (i) $\Rightarrow$ (ii) remains to be proved. This will be done in four steps.

1°. $\lambda \geq 0\,,\ \lambda(X) < \infty\,,\ \mu(X) < \infty\,.$

We consider the set G of measurable functions $g : X \to [0,\infty]$ such that $g\mu \leq \lambda$, i.e. such that

$$\int_A g\,d\mu \leq \lambda(A) \text{ for all } A \in \mathbb{E}\,. \tag{4}$$

The idea is to take some "maximal" function $f \in G$ and obtain equality in (4).

The function $g = 0$ belongs to G so $G \neq \emptyset$.

The set G is max-stable: $g_1, g_2 \in G \Rightarrow g_1 \vee g_2 \in G$. In fact, if we put $E = \{g_1 \geq g_2\}$ and $F = \{g_1 < g_2\}$, then $\{E, F\}$ is a partition of X, and for any $A \in \mathbb{E}$ we find

$$\begin{aligned}\int_A g_1 \vee g_2\,d\mu &= \int_{A\cap E} g_1 \vee g_2\,d\mu + \int_{A\cap F} g_1 \vee g_2\,d\mu \\ &= \int_{A\cap E} g_1\,d\mu + \int_{A\cap F} g_2\,d\mu \\ &\leq \lambda(A\cap E) + \lambda(A \cap F) = \lambda(A)\,.\end{aligned}$$

Defining

$$\gamma := \sup_{g\in G} \int g\,d\mu \tag{5}$$

we have $\gamma \leq \lambda(X) < \infty$, and we can choose a sequence (g_n') from G with $\lim \int g_n'\,d\mu = \gamma$. From the just shown we know that $g_n := \max(g_1', \cdots, g_n') \in G$, and (g_n) is an increasing sequence with $\int g_n'\,d\mu \leq \int g_n\,d\mu \leq \gamma$. By the monotone convergence theorem it follows that $f := \lim g_n$ belongs to G and $\gamma = \int f\,d\mu$. This can be expressed by saying that the function $g \mapsto \int g\,d\mu$ on G attains its maximum at $f \in G$.

We will show that $f\mu = \lambda$, and by (4) we already know that $f\mu \leq \lambda$, so $\tau := \lambda - f\mu$ is a positive measure. Step 1 will be completed by showing $\tau(X) = 0$.

If we assume $\tau(X) > 0$ then necessarily $\mu(X) > 0$ since $\lambda \ll \mu$. Let

$$\beta := \frac{1}{2}\frac{\tau(X)}{\mu(X)} > 0 \, .$$

Then $\tau(X) = 2\beta\mu(X) > \beta\mu(X)$, and we can apply Lemma 2.6 with $\sigma = \beta\mu$, so there exists $Y \in \mathsf{E}$ with $\tau(Y) > \beta\mu(Y)$ and $\beta\mu|Y \leq \tau|Y$. The first inequality implies in particular that $\tau(Y) > 0$ and hence $\mu(Y) > 0$ because of (i). The second inequality implies that the measurable function $f_0 := f + \beta 1_Y$ belongs to G, because for $A \in \mathsf{E}$ we have

$$\int_A f_0 d\mu = \int_A f d\mu + \beta\mu(A \cap Y) \leq \int_A f d\mu + \tau(A \cap Y) \leq \lambda(A) \, .$$

On the other hand

$$\int f_0 d\mu = \int f d\mu + \beta\mu(Y) = \gamma + \beta\mu(Y) > \gamma \, ,$$

which is in contradiction with (5).

2°. $\lambda \geq 0 \, , \, \lambda(X) = \infty \, , \, \mu(X) < \infty$.

We will show the existence of a partition $(X_n)_{n\geq 0}$ of X with the following properties

a) For any $A \in \mathsf{E} \, , \, A \subseteq X_0$ we have the alternatives

$$\mu(A) = \lambda(A) = 0 \quad \text{or} \quad \mu(A) > 0 \, , \, \lambda(A) = \infty$$

b) $\lambda(X_n) < \infty \, , \, n = 1, 2, \cdots$.

To do this, let F be the system of all sets $F \in \mathsf{E}$ with $\lambda(F) < \infty$ and define

$$\alpha := \sup_{F \in \mathsf{F}} \mu(F) \, .$$

Then there exists a sequence $(F_n)_{n\geq 1}$ from F with $\alpha = \lim \mu(F_n)$, and since F is stable under finite unions, we can assume that (F_n) is increasing. Defining

$$F_0 = \bigcup_1^\infty F_n \, , \; X_0 = \complement F_0 \, , \; X_1 = F_1 \, , \; X_{n+1} = F_{n+1} \setminus F_n \text{ for } n \geq 1 \, ,$$

we then have $\alpha = \mu(F_0)$ and $(X_n)_{n\geq 0}$ is a partition of X satisfying (b). To verify (a) all we have to show is

$$\forall A \in \mathsf{E} \, , \, A \subseteq X_0 \big(\lambda(A) < \infty \Rightarrow \mu(A) = 0\big) \, .$$

If $A \in \mathbb{E}, A \subseteq X_0$ satisfies $\lambda(A) < \infty$ then $A \in \mathbb{F}$ and hence $A \cup F_n \in \mathbb{F}$ for all $n \geq 1$. Since $A \cap F_n = \emptyset$ we have

$$\mu(A) + \mu(F_n) = \mu(A \cup F_n) \leq \alpha \ ,$$

and letting $n \to \infty$ we find $\mu(A) + \mu(F_0) \leq \alpha$ showing that $\mu(A) = 0$.

Let now $\mu_n = \mu|X_n$, $\lambda_n = \lambda|X_n$, $n \geq 0$ be the restrictions to the measurable space $(X_n, \mathbb{E}_{X_n})$, where $\mathbb{E}_{X_n} = \{E \in \mathbb{E} \mid E \subseteq X_n\}$. Then $\lambda_n \ll \mu_n$, $n \geq 0$ and by step 1° which can be applied for $n \geq 1$ there exists an $\mathbb{E}_{X_n}$-measurable function $f_n : X_n \to [0, \infty]$ such that $\lambda_n = f_n \mu_n$. Because of (a) we also have $\lambda_0 = f_0 \mu_0$ on $(X_0, \mathbb{E}_{X_0})$ if $f_0 \equiv \infty$. By putting the pieces $(f_n)_{n \geq 0}$ together via the partition $(X_n)_{n \geq 0}$, i.e. by defining $f : X \to [0, \infty]$ as being equal to f_n on X_n, elementary measure theory implies that $\lambda = f\mu$ because for $E \in \mathbb{E}$

$$\begin{aligned}\lambda(E) &= \sum_{n=0}^{\infty} \lambda(E \cap X_n) = \sum_{n=0}^{\infty} \lambda_n(E \cap X_n) = \sum_{n=0}^{\infty} \int_{E \cap X_n} f_n d\mu_n \\ &= \sum_{n=0}^{\infty} \int_{E \cap X_n} f d\mu = \int_E f d\mu \ .\end{aligned}$$

3°. $\lambda \geq 0$, μ is σ-finite.

By Lemma 2.5 there exists $h \in \mathcal{L}_1(\mu)$, $0 < h(x) < \infty$ and $h\mu$ is a finite measure with the same null sets as μ. Therefore $\lambda \ll h\mu$ and by step 1° and 2° there exists $f : X \to [0, \infty]$ such that $\lambda = f(h\mu) = (fh)\mu$, i.e. λ has the density fh with respect to μ.

If $\lambda = f\mu$ and $\lambda(X) < \infty$ then $\int f d\mu < \infty$ so $N = \{x \mid f(x) = \infty\}$ is a μ-null set. Replacing f by

$$\tilde{f}(x) = \begin{cases} 0 & , \quad x \in N \\ f(x) & , \quad x \notin N \end{cases}$$

we obtain an integrable function $\tilde{f}$ such that $\lambda = \tilde{f}\mu$.

4°. λ complex, μ is σ-finite.

Let $\lambda = \lambda_1 - \lambda_2 + i(\lambda_3 - \lambda_4)$ be the decomposition with $\lambda_1 = (\operatorname{Re}\lambda)^+$, $\lambda_2 = (\operatorname{Re}\lambda)^-$, $\lambda_3 = (\operatorname{Im}\lambda)^+$, $\lambda_4 = (\operatorname{Im}\lambda)^-$. It is easy to see that $\operatorname{Re}\lambda, \operatorname{Im}\lambda \ll \mu$ and hence $\lambda_j \ll \mu$, $j = 1, \cdots, 4$. By step 3° there exist non-negative integrable functions $f_j, j = 1, \cdots, 4$ such that $\lambda_j = f_j\mu$ and then $\lambda = f\mu$, where $f = f_1 - f_2 + i(f_3 - f_4)$ belongs to $\mathcal{L}_1(\mu)$. □

REMARK 2.7. The density f in Theorem 2.4 (ii) is uniquely determined up to a μ-null set. This statement amounts to the assertions:

(c) Let $f \in \mathcal{L}_1(\mu)$ and assume that $\int_E f d\mu = 0$ for all $E \in \mathbb{E}$. Then $f = 0$ μ-a.e.

(d) Let $f_1, f_2 : X \to [0, \infty]$ be measurable functions such that

$$\int_E f_1 d\mu = \int_E f_2 d\mu \text{ for all } E \in \mathbb{E} .$$

Then $f_1 = f_2$ μ-a.e.

PROOF OF (c): The complex case follows from the real case, so assume that f is real. Using $E = \{f > 0\}$ and $E = \{f < 0\}$ we find $f = 0$ μ-a.e.

For the statement in (c) we do not use that μ is σ-finite, but this is essential in (d). □

PROOF OF (d): If (d) holds for finite μ, it also holds for σ-finite μ, so we shall assume $\mu(X) < \infty$.

Let $E = \{f_1 > f_2\}$ and $E_n = \{x \in E \mid f_2(x) \leq n\}$. Then (E_n) is an increasing sequence with union E and

$$\begin{aligned}\int_E (f_1 - f_2) d\mu &= \lim_{n\to\infty} \int_{E_n} (f_1 - f_2) d\mu \\ &= \lim_{n\to\infty} \Big(\int_{E_n} f_1 d\mu - \int_{E_n} f_2 d\mu\Big) = 0 .\end{aligned}$$

The first equality sign follows from the monotone convergence theorem and in the second we use that

$$0 \leq \int_{E_n} f_2 d\mu \leq n\mu(E_n) < \infty .$$

It follows that $\mu(E) = 0$ and similarly $\mu(\{f_1 < f_2\}) = 0$ so $f_1 = f_2$ μ-a.e. □

DEFINITION 2.8. Let $\lambda \ll \mu$, where λ and μ are as in Theorem 2.4. The density f such that $\lambda = f\mu$ which is determined μ-a.e. is called the *Radon-Nikodym derivative* of λ with respect to μ, and it is often denoted

$$f = \frac{d\lambda}{d\mu} .$$

(Radon proved the theorem in 1913 for euclidean spaces with μ equal to Lebesgue measure. The general version is due to Nikodym 1930.)

The following example shows that Theorem 2.4 may fail if μ is not σ-finite.

EXAMPLE 2.9. On the measurable space $(\mathbb{R}, \mathbb{B})$ we consider Lebesgue measure λ and the counting measure μ. Then $\lambda \ll \mu$, but there is no measurable function $f : \mathbb{R} \to [0, \infty]$ such that $\lambda = f\mu$, because this equation implies

$$0 = \lambda(\{a\}) = \int_{\{a\}} f \, d\mu = f(a)$$

for all $a \in \mathbb{R}$.

We will now give some consequences of the Radon-Nikodym theorem.

THEOREM 2.10. *Let μ be a complex measure on $(X, \mathbb{E})$. Then there is a measurable function $a : X \to \mathbb{R}$ such that*

$$\frac{d\mu}{d|\mu|} = e^{ia}$$

i.e.

$$\mu(E) = \int_E e^{ia(x)} d|\mu|(x) \text{ for } E \in \mathbb{E}\,.$$

PROOF: The absolute value $|\mu|$ is a finite positive measure and clearly $\mu \ll |\mu|$. Let $h \in \mathcal{L}_1(\mu)$ be such that $\mu = h|\mu|$. By Example 1.8 (b) we know that $|\mu| = |h|\,|\mu|$ so by the uniqueness of the Radon-Nikodym derivative (Remark 2.7) $|h| = 1$ $|\mu|$-a.e. We can therefore assume that $|h| = 1$, and if $\mathrm{Arg}(z)$ is the principal argument $\in]-\pi, \pi]$ of $z \in \mathbb{C} \setminus \{0\}$ then Arg is measurable and so is $a = \mathrm{Arg} \circ h$. □

THEOREM 2.11. *Let μ be a real measure on $(X, \mathbb{E})$. Then there exists a partition $\{P, N\}$ of X, called a Hahn decomposition, such that*

$$\mu^+(E) = \mu(P \cap E)\,,\ \mu^-(E) = -\mu(N \cap E) \text{ for } E \in \mathbb{E}\,.$$

If $\mu = \lambda_1 - \lambda_2$ with positive measures λ_1, λ_2, then $\mu^+ \leq \lambda_1$, $\mu^- \leq \lambda_2$.

In other words X can be split into disjoint sets P and N such that P carries the positive mass and N the negative mass of μ.

PROOF: By Theorem 2.10 we have $\mu = h|\mu|$ for a measurable function h with $|h| = 1$ and we can assume that h is real-valued, hence $h = \pm 1$. Put

$$P = \{x \in X \mid h(x) = 1\}\,,\ N = \{x \in X \mid h(x) = -1\}\,.$$

Since $\mu^\pm = \frac{1}{2}(|\mu| \pm \mu)$ we have

$$\mu^\pm = \frac{1}{2}(1 \pm h)|\mu| = \begin{cases} 1_P|\mu| & = 1_P\mu \\ 1_N|\mu| & = -1_N\mu\,, \end{cases}$$

so for $E \in \mathbb{E}$ we get

$$\mu^+(E) = \mu(P \cap E) \,,\; \mu^-(E) = -\mu(N \cap E) \,.$$

If μ is the difference of two positive measures λ_1, λ_2 we find

$$\mu^+(E) = \mu(P \cap E) = \lambda_1(P \cap E) - \lambda_2(P \cap E) \leq \lambda_1(P \cap E) \leq \lambda_1(E)$$

and

$$\mu^-(E) = -\mu(N \cap E) = \lambda_2(N \cap E) - \lambda_1(N \cap E) \leq \lambda_2(N \cap E) \leq \lambda_2(E) \,.$$

□

LEMMA 2.12. *For a complex measure μ on $(X, \mathbb{E})$ and $A \in \mathbb{E}$ the following conditions are equivalent:*

(i) $\mu(E) = 0$ *for all* $E \in \mathbb{E}$ *with* $E \cap A = \emptyset$

(ii) $\mu(E) = \mu(E \cap A)$ *for all* $E \in \mathbb{E}$

(iii) $|\mu|(\complement A) = 0$.

If μ and A satisfy the above conditions we say that μ is *concentrated* on A. The proof of Lemma 2.12 is straightforward and left as an exercise.

DEFINITION 2.13. Let λ, ν be arbitrary measures on $(X, \mathbb{E})$. We say that λ and ν are *mutually singular* and write $\lambda \perp \nu$ if there exist disjoint sets $A, B \in \mathbb{E}$ such that λ is concentrated on A, ν is concentrated on B.

LEMMA 2.14. *Let λ, ν be arbitrary measures on $(X, \mathbb{E})$ and μ a positive measure on $(X, \mathbb{E})$.*

(a) *If $\lambda \perp \nu$, then $|\lambda| \perp |\nu|$.*

(b) *If $\lambda \perp \mu$ and $\nu \perp \mu$, then $\lambda + \nu \perp \mu$.*

(c) *If $\lambda \ll \mu$ and $\lambda \perp \mu$, then $\lambda = 0$.*

PROOF: (a) is clear since λ and $|\lambda|$ are concentrated on the same sets by Lemma 2.12.

(b) If $\{A_1, B_1\}$ and $\{A_2, B_2\}$ denote pairs of disjoint sets from $\mathbb{E}$ such that λ is concentrated on A_1, ν concentrated on A_2 and μ concentrated on B_1 and on B_2, then μ is concentrated on $B_1 \cap B_2$ and $\lambda + \nu$ is concentrated on $A_1 \cup A_2$ by Lemma 2.12 (iii)

(c) Let $\{A, B\}$ be a pair of disjoint sets such that λ is concentrated on A, μ is concentrated on B. For $E \in \mathbb{E}$ we find

$$\lambda(E) = \lambda(E \cap A) + \lambda(E \setminus A) = 0$$

where the first term vanishes because $\mu(E \cap A) = 0$. □

LEMMA 2.15. *Let λ, μ be positive σ-finite measures on (X, E) and assume $\lambda \leq \mu$. Then there exists a measurable function $f : X \to [0,1]$ such that $\lambda = f\mu$.*

PROOF: Clearly $\lambda \ll \mu$, so by the Radon-Nikodym theorem there exists a measurable function $f : X \to [0, \infty]$ so that $\lambda = f\mu$. Let $E \in \mathsf{E}$ with $\mu(E) < \infty$ and let $a > 1$. Then $E_a := \{x \in E \mid f(x) \geq a\} \in \mathsf{E}$ and

$$\mu(E_a) \geq \lambda(E_a) = \int_{E_a} f d\mu \geq a\mu(E_a) \,,$$

hence $\mu(E_a) = 0$. Since

$$E_{1+\frac{1}{n}} \nearrow \{x \in E \mid f(x) > 1\}$$

we see that $\mu(\{x \in E \mid f(x) > 1\}) = 0$, and using the σ-finiteness of μ we get $f \leq 1$ μ–a.e. . □

THEOREM 2.16. *Let λ be a complex measure and let μ be a positive σ-finite measure on (X, E).*

There exists a uniquely determined decomposition $\lambda = \lambda_a + \lambda_s$ as sum of complex measures λ_a and λ_s satisfying

(i) $\lambda_a \ll \mu$,

(ii) $\lambda_s \perp \mu$.

If λ is real (resp. positive) then λ_a and λ_s are real (resp. positive).

The decomposition above is called the *Lebesgue decomposition of λ with respect to μ* and λ_a is called the *μ-absolutely continuous part*, λ_s the *μ-singular part* of λ.

PROOF: The decomposition is unique, because if we consider two decompositions $\lambda = \lambda_a + \lambda_s = \lambda'_a + \lambda'_s$ as above and define $\tau = \lambda_a - \lambda'_a = \lambda'_s - \lambda_s$, then $\tau \ll \mu$ and $\tau \perp \mu$. By Lemma 2.14 (c) we get $\tau = 0$.

It is enough to prove the decomposition for positive finite measures λ because afterwards we can apply the decomposition to each of the positive terms in the decomposition $\lambda = \lambda_1 - \lambda_2 + i(\lambda_3 - \lambda_4)$.

Assume $\lambda \geq 0$, $\lambda(X) < \infty$. Then $\lambda + \mu$ is a positive σ-finite measure and clearly $\lambda \ll \lambda + \mu$. By Lemma 2.15 there exists a measurable function $f : X \to [0,1]$ such that $\lambda = f(\lambda + \mu)$. Let

$$A = \{x \in X \mid f(x) < 1\} \,, \; S = \{x \in X \mid f(x) = 1\}$$

then $\{A, S\}$ is a partition of X and we define

$$\lambda_a(E) = \lambda(E \cap A)\ ,\ \lambda_s(E) = \lambda(E \cap S) \quad \text{for } E \in \mathsf{E}\ .$$

Then $\lambda = \lambda_a + \lambda_s$ is the desired composition. If $\mu(E) = 0$ then

$$\lambda_a(E) = \lambda(E \cap A) = \int_{E\cap A} f d(\lambda + \mu) = \int_{E\cap A} f d\lambda$$

and hence

$$\int_{E\cap A} (1 - f) d\lambda = 0\ .$$

Since $1 - f(x) > 0$ for $x \in E \cap A$ this implies $\lambda(E \cap A) = 0$ showing that $\lambda_a \ll \mu$. To see that $\lambda_s \perp \mu$ we remark that λ_s is concentrated on S, and μ is concentrated on $\complement S$ because

$$\lambda(S) = \int_S f d(\lambda + \mu) = \lambda(S) + \mu(S)\ ,$$

so $\mu(S) = 0$. □

As an application of the Radon-Nikodym theorem we shall prove that the dual space of $L_p(\mu) = L_p(X, \mathsf{E}, \mu)$ can be identified with $L_q(\mu)$ if $1 < p < \infty$ and $1/p + 1/q = 1$. As usual $L_p(\mu)$ is the Banach space of equivalence classes $[f]$ of functions $f \in \mathcal{L}_p(\mu)$.

Let E be a complex Banach space. We recall that the dual space E' of continuous linear functionals $T : E \to \mathbb{C}$ is a Banach space under the norm

$$\|T\| = \sup\{|T(x)| \mid \|x\| \leq 1\}\ .$$

THEOREM 2.17. *Let μ be a positive measure on (X, E) and let $1 < p < \infty$, $1/p + 1/q = 1$. For $\varphi \in \mathcal{L}_q(X, \mathsf{E}, \mu)$ the expression*

$$T_\varphi([f]) = \int f\varphi d\mu \quad ,\ f \in \mathcal{L}_p(X, \mathsf{E}, \mu)$$

defines a continuous linear functional $T_\varphi : L_p(X, \mathsf{E}, \mu) \to \mathbb{C}$ satisfying

(i) $T_\varphi = T_\psi \Leftrightarrow \varphi = \psi \quad \mu - \text{a.e.}$

(ii) $\|T_\varphi\| = \|\varphi\|_q$.

The mapping $[\varphi] \mapsto T_\varphi$ is an isometric isomorphism of $L_q(X, \mathsf{E}, \mu)$ onto $L_p(X, \mathsf{E}, \mu)'$.

PROOF: By Hölder's inequality it is clear that T_φ defines a linear functional such that

$$|T_\varphi([f])| \leq \|f\|_p \|\varphi\|_q\ ,$$

which shows that T_φ is continuous and $\|T_\varphi\| \le \|\varphi\|_q$. The property (i) is a consequence of (ii), if we remark that $\varphi \mapsto T_\varphi$ is linear.

It remains to be shown that if Φ is a continuous linear functional on $L_p(X, \mathsf{E}, \mu)$, there exists $\varphi \in \mathcal{L}_q(X, \mathsf{E}, \mu)$ such that $T_\varphi = \Phi$ and $\|\varphi\|_q \le \|\Phi\|$.

If $\Phi = 0$ we can choose $\varphi = 0$, so assume $\|\Phi\| > 0$.

1° $\mu(X) < \infty$.

For $E \in \mathsf{E}$ we define $\lambda(E) = \Phi([1_E])$; note that $1_E \in \mathcal{L}_p(\mu)$ because $\mu(X) < \infty$. Then $\lambda : \mathsf{E} \to \mathbb{C}$ is finitely additive because Φ is linear. To prove countable additivity, suppose E is the union of countably many disjoint sets $E_n \in \mathsf{E}$, put $A_k = E_1 \cup \cdots \cup E_k$ and note that

$$\|1_E - 1_{A_k}\|_p = \mu(E \setminus A_k)^{1/p} \to 0 \quad \text{for} \quad k \to \infty\,.$$

By continuity of Φ we then have $\lambda(A_k) = \Phi([1_{A_k}]) \to \Phi([1_E]) = \lambda(E)$. This shows that λ is a complex measure. It is clear that $\lambda(E) = 0$ if $\mu(E) = 0$ because then $\|1_E\|_p = 0$. Thus $\lambda \ll \mu$, and by the Radon–Nikodym theorem there exists $\varphi \in \mathcal{L}_1(\mu)$ such that $\lambda = \varphi\mu$, i.e.

$$\Phi([1_E]) = \int_E \varphi \, d\mu \quad \text{for} \quad E \in \mathsf{E}\,. \tag{6}$$

By linearity (6) implies

$$\Phi([f]) = \int f\varphi \, d\mu \tag{7}$$

for all simple measurable functions, and we see next that (7) also holds for all bounded measurable functions f. In fact for any measurable function f satisfying $|f| \le K$ there exists a sequence (f_n) of simple measurable functions such that $\|f - f_n\|_u \le 1/n$ and then $\|f - f_n\|_p = (1/n)\,\mu(X)^{1/p}$. The first inequality shows that $\int f_n\varphi\, d\mu \to \int f\varphi\, d\mu$, the second that $\Phi([f_n]) \to \Phi([f])$. (The existence of (f_n): For each n let us consider a finite partition (D_i) of $\{z \in \mathbb{C} \mid |z| \le K\}$ in sets of diameter $\le \dfrac{1}{n}$, choose $d_i \in D_i$ and put

$$f_n = \sum_i d_i \, 1_{f^{-1}(D_i)}\,,$$

then $\|f - f_n\|_u \le 1/n$.)

Choose now a measurable function α with $|\alpha| = 1$ so that $\alpha\varphi = |\varphi|$ and define

$$E_n = \{x \in X \mid |\varphi(x)| \le n\}\,, \quad f = \alpha|\varphi|^{q-1}1_{E_n}\,.$$

Applying (7) to the bounded measurable function f we get

$$\int_{E_n} |\varphi|^q \, d\mu = \int f\varphi \, d\mu = \Phi([f]) \le \|\Phi\|\|f\|_p\,,$$

and dividing by

$$\|f\|_p = \Big(\int_{E_n} |\varphi|^q \, d\mu\Big)^{1/p}$$

we find

$$\int_{E_n} |\varphi|^q \, d\mu \le \|\Phi\|^q\,, \quad n \ge 1\,.$$

Letting $n \to \infty$ the monotone convergence theorem implies $\varphi \in \mathcal{L}_q(\mu)$ and $\|\varphi\|_q \le \|\Phi\|$. We know by (7) that Φ and T_φ agree on simple $\mathcal{L}_p$-functions which form a dense subset of $\mathcal{L}_p$ by a theorem in 2MA, and therefore the continuous functionals Φ and T_φ agree.

$2°$ $\mu(X) = \infty$.

Let $\mathbb{E}_0 = \{E \in \mathbb{E} \mid \mu(E) < \infty\}$. For $E \in \mathbb{E}_0$ we consider the space $L_p(\mu|E) = L_p(E, \mathbb{E}_E, \mu|E)$, and the embedding $I_E : L_p(\mu|E) \to L_p(\mu)$ given by

$$I_E([f]) = [\tilde{f}]\,, \quad \text{where} \quad \tilde{f}(x) = \begin{cases} f(x), & x \in E \\ 0, & x \notin E\,. \end{cases}$$

Clearly I_E is a linear isometry, and $\Phi \circ I_E : L_p(\mu|E) \to \mathbb{C}$ is a continuous linear functional on $L_p(\mu|E)$, so by part $1°$ there exists $\varphi_E \in \mathcal{L}_q(\mu|E)$ so that

$$\Phi(I_E([f])) = \int f\,\varphi_E \, d(\mu|E) \quad \text{for} \quad f \in \mathcal{L}_p(\mu|E)\,, \tag{8}$$

and

$$\tau(E) := \|\Phi \circ I_E\|^q = \int |\varphi_E|^q \, d(\mu|E)\,. \tag{9}$$

Now $\varphi_E : E \to \mathbb{C}$ is unique up to μ-null sets so it is clear that if $E, F \in \mathbb{E}_0$, $E \subseteq F$ then $\varphi_E = \varphi_F|E$ μ-a.e. It follows that $\tau : \mathbb{E}_0 \to [0, \infty[$ is *additive* and *increasing*. Furthermore $\tau(E) \le \|\Phi\|^q$ for all $E \in \mathbb{E}_0$.

We claim:

$$\sup_{E \in \mathbb{E}_0} \tau(E) = \|\Phi\|^q\,. \tag{10}$$

Let $0 < \lambda < \|\Phi\|$ be given. By definition of $\|\Phi\|$ there exists $f \in \mathcal{L}_p(\mu)$ with $\|f\|_p \leq 1$ and $|\Phi([f])| > \lambda$. For $r > 0$ define

$$A_r = \{x \in X \mid |f(x)| \geq r\}.$$

Since $f \in \mathcal{L}_p(\mu)$ we have $\mu(A_r) < \infty$; and $f\,1_{A_r} \to f$ in $\mathcal{L}_p(\mu)$ for $r \to 0$ by dominated convergence, hence

$$\Phi([f\,1_{A_r}]) \to \Phi([f])$$

by continuity of Φ, so

$$\|\Phi \circ I_{A_r}\| \geq |\Phi \circ I_{A_r}([f|A_r])| = |\Phi([f\,1_{A_r}])| > \lambda,$$

where the last inequality holds for r sufficiently small. This establishes (10).

By (10) we can find $E_1 \subseteq E_2 \subseteq \cdots \in \mathsf{E}_0$ such that $\tau(E_n) \nearrow \|\Phi\|^q$. Let φ_n be chosen according to (8) corresponding to the set E_n, and we can assume that $\varphi_{n+1}|E_n = \varphi_n\,, \quad n = 1, 2, \ldots$. The function

$$\varphi(x) = \begin{cases} \varphi_n(x) & \text{for} \quad x \in E_n \\ 0 & \text{for} \quad x \notin \bigcup_1^\infty E_n\,, \end{cases} \tag{11}$$

is measurable and

$$\int_{E_n} |\varphi|^q \, d\mu = \int_{E_n} |\varphi_n|^q \, d(\mu|E_n) = \tau(E_n) \nearrow \|\Phi\|^q.$$

By the monotone convergence theorem $\varphi \in \mathcal{L}_q(\mu)$ and $\|\varphi\|_q = \|\Phi\|$. We claim that $T_\varphi = \Phi$ and for this it suffices to prove that

$$\Phi([1_F]) = \int_F \varphi \, d\mu$$

for all $F \in \mathsf{E}_0$ since the linear span of the functions $1_F\,,\ F \in \mathsf{E}_0$ is dense in $\mathcal{L}_p(\mu)$. By additivity it is enough to consider the following two cases:

a) $F \in \mathsf{E}_0\,,\ F \subseteq \bigcup_1^\infty E_n$.

Since $1_{F\cap E_n} \to 1_F$ in $\mathcal{L}_p(\mu)$ we have

$$\Phi([1_F]) = \lim_{n\to\infty} \Phi([1_{F\cap E_n}]) = \lim_{n\to\infty} \int_{F\cap E_n} \varphi \, d\mu = \int_F \varphi \, d\mu.$$

b) $F \in \mathsf{E}_0\,,\ F \cap \left(\bigcup_1^\infty E_n\right) = \varnothing$.

Since $\varphi = 0$ on F by (11), we have to show that $\Phi([1_F]) = 0$ which is fulfilled since $\Phi \circ I_F = 0$, and this can be seen in the following way:

For $n \in \mathbb{N}$ we have

$$\tau(F) + \tau(E_n) = \tau(F \cup E_n) \leq \|\Phi\|^q$$

and for $n \to \infty$ we get

$$\tau(F) + \|\Phi\|^q \leq \|\Phi\|^q$$

hence

$$0 = \tau(F) = \|\Phi \circ I_F\|^q .$$

□

REMARK 2.18. (a) For $p = q = 2$ we get in particular that any continuous linear functional $\Phi : L_2(\mu) \to \mathbb{C}$ is given as the scalar product with an $\mathcal{L}_2$-function $\overline{\varphi}$:

$$\Phi([f]) = \int f\varphi \, d\mu = (f|\overline{\varphi}) \quad \text{for} \quad f \in \mathcal{L}_2(\mu) .$$

This is a general fact about Hilbert spaces and is due to F. Riesz: *Let H be a Hilbert space. For any continuous linear functional $\Phi : H \to \mathbb{C}$ there exists a unique vector $y \in H$ so that*

$$\Phi(x) = (x|y) \quad \text{for all} \quad x \in H .$$

It is possible to deduce the Radon–Nikodym theorem from the Riesz theorem. This is done in Rudin's book.

(b) For a Banach space X the double dual space X'' is defined as the dual space of X'. If X and X'' are the "same" we call X *reflexive.* We see that $L_p(\mu)$ is a reflexive Banach space for $1 < p < \infty$.

(c) As an application of Theorem 2.17 let $X = \{1, \cdots, n\}$ and let μ be the counting measure on X. Then $\mathcal{L}_p(X, \mu) = L_p(X, \mu)$ can be identified with $\mathbb{C}^n$ equipped with the norm

$$\|x\|_p = \begin{cases} \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} & , \quad 1 \leq p < \infty \\ \max\limits_{i=1,\cdots,n} |x_i| & , \quad p = \infty \end{cases}$$

Linear functionals $\varphi : \mathbb{C}^n \to \mathbb{C}$ are automatically continuous, and if $\varphi(e_i) = y_i, i = 1, \cdots, n$, where $e_1, \cdots, e_n$ is the standard basis of $\mathbb{C}^n$, then $y = (y_1, \cdots, y_n) \in \mathbb{C}^n$ determines φ because of the equation

$$\varphi(x) = \varphi\Big(\sum_1^n x_i e_i\Big) = \sum_1^n x_i y_i .$$

This shows that the dual space of $(\mathbb{C}^n, \|\cdot\|_p)$ is isomorphic with $\mathbb{C}^n$, and the norm on $\mathbb{C}^n$ as the dual space of $(\mathbb{C}^n, \|\cdot\|_p)$ is

$$\sup\left\{|\sum_1^n x_i y_i| \;\middle|\; \|x\|_p \le 1\right\} \tag{12}$$

which is equal to $\|y\|_q$ for $1 < p < \infty, 1/p + 1/q = 1$. In this simple case the result also holds for $p = 1$ and $p = \infty$. These results can be verified directly. In fact, the expression in (12) is $\le \|y\|_q$ by *Hölder's* inequality, and for $y \neq 0$ the above supremum is a maximum attained for

$$x = (x_1, \cdots, x_n) \text{ where } x_i = \overline{\operatorname{sgn}(y_i)}\left(\frac{|y_i|}{\|y\|_q}\right)^{q-1} \quad \text{for } 1 < p \le \infty,$$

and for $p = 1$ we choose $i_0 \in \{1, \cdots, n\}$ so that $|y_{i_0}| = \|y\|_\infty$ and put

$$x_i = \begin{cases} \overline{\operatorname{sgn}(y_{i_0})} & \text{if } i = i_0, \\ 0 & \text{if } i \neq i_0. \end{cases}$$

In general $L_\infty(\mu)' \neq L_1(\mu)$, but $L_1(\mu)' = L_\infty(\mu)$ at least when μ is σ-finite. We use the terminology of Theorem 2.17 in the precise statement below.

THEOREM 2.19. *Let μ be a positive σ-finite measure. Then $[\varphi] \mapsto T_\varphi$ is an isometric isomorphism of $L_\infty(\mu)$ onto $L_1(\mu)'$, where*

$$T_\varphi([f]) = \int f\varphi d\mu \quad \text{for } f \in \mathcal{L}_1(\mu), \varphi \in \mathcal{L}_\infty(\mu).$$

PROOF: As in the proof of Theorem 2.17 it suffices to establish the following: Let $\Phi : L_1(\mu) \to \mathbb{C}$ be a continuous linear functional with $\|\Phi\| > 0$. Then there exists $\varphi \in \mathcal{L}_\infty(\mu)$ such that $\Phi = T_\varphi$ and $\|\varphi\|_\infty \le \|\Phi\|$.

1° $\mu(X) < \infty$.

As before $\lambda(E) = \Phi([1_E]), E \in \mathbb{E}$ is a complex measure which is absolutely continuous with respect to μ. By the *Radon-Nikodym* theorem there exists $\varphi \in \mathcal{L}_1(\mu)$ such that (6) and (7) hold.
In particular

$$\left|\int_E \varphi d\mu\right| \le \|\Phi\| \, \|1_E\|_1 = \|\Phi\|\mu(E) \quad \text{for } E \in \mathbb{E}$$

showing that $|\varphi\mu| = |\varphi|\mu \le \|\Phi\|\mu$, and it follows easily that $|\varphi| \le \|\Phi\|$ μ-a.e., i.e. $\varphi \in \mathcal{L}_\infty(\mu)$ and $\|\varphi\|_\infty \le \|\Phi\|$. By (7) we know that Φ and T_φ agree on simple $\mathcal{L}_1$-functions hence $\Phi = T_\varphi$ by the denseness in $\mathcal{L}_1$ of these functions.

2° $\mu(X) = \infty$.

By Lemma 2.5 there exists $h \in \mathcal{L}_1(\mu)$ so that $0 < h(x) < \infty$ for all $x \in X$. Then $\sigma = h\mu$ is a finite measure and $[f] \mapsto [hf]$ is a linear isometry I of $L_1(\sigma)$ onto $L_1(\mu)$.

Then $\Psi = \Phi \circ I$ is a continuous linear functional on $L_1(\sigma)$ with $\|\Psi\| = \|\Phi\|$, and by 1° there exists $\psi \in \mathcal{L}_\infty(\sigma)$ with $\|\psi\|_\infty = \|\Psi\|$ such that

$$\Psi([g]) = \int g\psi d\sigma \quad \text{for } g \in \mathcal{L}_1(\sigma).$$

Noting that ψ also belongs to $\mathcal{L}_\infty(\mu)$ and with the same norm in $\mathcal{L}_\infty(\mu)$ as in $\mathcal{L}_\infty(\sigma)$ we find for $f \in \mathcal{L}_1(\mu)$:

$$\Phi([f]) = \Psi([f/h]) = \int (f/h)\psi h d\mu = \int f\psi d\mu.$$

□

REMARK 2.20. For a positive measure μ there is a *Radon-Nikodym* theorem with respect to μ if and only if the dual space of $L_1(\mu)$ is $L_\infty(\mu)$ and measure spaces with these properties have been characterized by I.E.Segal: *Equivalences in measure spaces.* Amer. J. Math.73 (1951), 275–313.

§3. Differentiation theory

In this section we shall develop a differentiation theory of measures with respect to Lebesgue measure and thus obtain a procedure for finding the Radon-Nikodym derivative of a measure which is absolutely continuous with respect to Lebesgue measure.

In $\mathbb{R}^k$ we write $|x|$ for the euclidean norm

$$|x| = (\sum x_i^2)^{\frac{1}{2}} ,$$

and $K(x,r)$ denotes the open ball

$$K(x,r) = \{y \in \mathbb{R}^k \mid |x-y| < r\} .$$

Let μ be a complex measure on $(\mathbb{R}^k, \mathbb{B}_k)$ and define

$$Q_r(\mu)(x) = \frac{\mu(K(x,r))}{m(K(x,r))} \quad , \; x \in \mathbb{R}^k \; , \; r > 0 \; , \tag{1}$$

where $m = m_k$ denotes Lebesgue measure.

If $\mu = fm$ and f is continuous for $x = a$, then it is clear that $Q_r(\mu)(a) \to f(a)$ for $r \to 0$, cf. 2MA, and we are going to examine $\lim_{r\to 0} Q_r(\mu)(x)$ under weaker assumptions on μ.

The *symmetric derivative* of μ is defined as

$$D(\mu)(x) = \lim_{r\to 0} Q_r(\mu)(x) \tag{2}$$

at those points $x \in \mathbb{R}^k$ at which the limit exists.

In the study of $D(\mu)$ we shall make use of the *maximal function* $M(\mu)$ introduced by Hardy and Littlewood in 1930:

$$M(\mu)(x) = \sup_{r>0} Q_r(|\mu|)(x) = \sup_{r>0} \frac{|\mu|(K(x,r))}{m(K(x,r))} \; . \tag{3}$$

Note that $m(K(x,r)) = V_k r^k$, where V_k is Lebesgue measure of the unit ball, so we have

$$|\mu(K(x,r))| \le |\mu|(K(x,r)) \le M(\mu)(x) V_k r^k \; . \tag{4}$$

This is only interesting for small $r > 0$ since μ is bounded by $\|\mu\|$. In contrast to $D(\mu)$, where we do not know its set of definition, the maximal function is defined at all points of $\mathbb{R}^k$ possibly having the value ∞ at certain points.

LEMMA 3.1. *The maximal function is lower semicontinuous, i.e. the set* $\{x \in \mathbb{R}^k \mid M(\mu)(x) > \lambda\}$ *is open for any* $\lambda \in \mathbb{R}$. *In particular* $M(\mu)$ *is a Borel function.*

PROOF: Assume $M(\mu)(x_0) > \lambda$. We shall find $\delta > 0$ such that $M(\mu)(x) > \lambda$ for all $x \in K(x_0, \delta)$. By (3) there exists $r > 0$ such that $Q_r(|\mu|)(x_0) > \lambda$ and we choose $\delta > 0$ such that

$$(r+\delta)^k < r^k \frac{Q_r(|\mu|)(x_0)}{\lambda} . \tag{5}$$

For $x \in K(x_0, \delta)$ we have $K(x_0, r) \subseteq K(x, r+\delta)$ by the triangle inequality, and therefore

$$|\mu|(K(x, r+\delta)) \geq |\mu|(K(x_0, r)) = Q_r(|\mu|)(x_0) V_k r^k$$

which is bigger than $\lambda m(K(x, r+\delta))$ by (5), hence

$$M(\mu)(x) > \lambda .$$

□

LEMMA 3.2. *(Wiener's covering Lemma.)*

If W *is the union of a finite collection of balls* $K_i = K(x_i, r_i)$, $1 \leq i \leq N$ *in* $\mathbb{R}^k$, *then there is a set* $S \subseteq \{1, \cdots, N\}$ *so that*

(a) *the balls* $K(x_i, r_i)$, $i \in S$, *are disjoint,*

(b) $W \subseteq \bigcup_{i \in S} K(x_i, 3r_i)$,

(c) $m(W) \leq 3^k \sum_{i \in S} m(K(x_i, r_i))$.

PROOF: We can assume that the balls are numbered so that $r_1 \geq r_2 \geq \cdots \geq r_N$. Put $i_1 = 1$. Remove all balls K_j that intersect K_{i_1}. In particular K_{i_1} is removed. Let K_{i_2} be the first of the remaining balls if there are any. Among these remove all K_j that intersect K_{i_2}, and let K_{i_3} be the first of the remaining ones, and so on as long as possible. In this way we get the finite set $S = \{i_1, i_2, \cdots\}$.

Clearly (a) holds. Every $K_j, j \notin S$, is a subset of $K(x_i, 3r_i)$ for some $i \in S$. In fact, $K_j \cap K_i \neq \emptyset$ for some $i \in S$ for which $r_i \geq r_j$ and hence $|x_i - x_j| < r_i + r_j \leq 2r_i$, so that

$$K_j = K(x_j, r_j) \subseteq K(x_i, 3r_i) .$$

This proves (b), and (c) is an immediate consequence of (b). □

The next theorem says, roughly speaking, that the maximal function cannot be large on a large set.

THEOREM 3.3. *Let μ be a complex measure on $\mathbb{R}^k$ and $\lambda > 0$, then*

$$m(\{x \in \mathbb{R}^k \mid M(\mu)(x) > \lambda\}) \le 3^k \frac{\|\mu\|}{\lambda} \,. \tag{6}$$

PROOF: Let $\Omega = \{x \in \mathbb{R}^k \mid M(\mu)(x) > \lambda\}$. By Lemma 3.1 we know that Ω is open. Let $C \subseteq \Omega$ be an arbitrary compact subset. Each $x \in C$ is centre of a ball $K(x, r_x)$ for which

$$|\mu|(K(x,r_x)) > \lambda m(K(x,r_x)) \,.$$

By compactness a finite collection of these balls covers C and Wiener's covering lemma gives us a disjoint subcollection $\{K_1, \cdots, K_n\}$ satisfying

$$m(C) \le 3^k \sum_{j=1}^{n} m(K_j) \le 3^k \lambda^{-1} \sum_{j=1}^{n} |\mu|(K_j) \le 3^k \frac{\|\mu\|}{\lambda} \,.$$

Since every compact subset of Ω has Lebesgue measure bounded by the right-hand side of (6), also $m(\Omega)$ is bounded by this number. In fact, $\Omega \ne \mathbb{R}^k$ and

$$C_n = \{x \in \Omega \mid |x| \le n \,,\ \operatorname{dist}(x, \complement\Omega) \ge \frac{1}{n}\} \quad , \ n \in \mathbb{N}$$

gives an increasing sequence of compact subsets of Ω for which $\bigcup_1^\infty C_n = \Omega$. □

For $f \in \mathcal{L}_1(\mathbb{R}^k, m)$ we define the *maximal function* $M(f)$ of f as the maximal function of the complex measure fm, i.e.

$$M(f)(x) = \sup_{r>0} \frac{1}{m(K_r)} \int_{K(x,r)} |f(y)| dy \,,$$

where we write $K_r = K(x,r)$ when the centre is without importance.

Let $\mathcal{L}_{\text{loc}}(\mathbb{R}^k)$ denote the vector space of Borel measurable functions $f : \mathbb{R}^k \to \mathbb{C}$ which are locally integrable, i.e.

$$\int_{K(x,r)} |f(y)| dy < \infty$$

for any ball $K(x,r)$. Note that $M(f)$ is well defined by (7) for any $f \in \mathcal{L}_{\text{loc}}$, and as in Lemma 3.1 it is seen to be lower semicontinuous.

DEFINITION 3.4: Let $f \in \mathcal{L}_{\text{loc}}(\mathsf{R}^k)$. A point $x_0 \in \mathsf{R}^k$ is called a *Lebesgue point* of f if

$$\lim_{r \to 0} \frac{1}{m(K_r)} \int_{K(x_0,r)} |f(x) - f(x_0)| dx = 0 . \tag{7}$$

Clearly, a point x_0 of continuity is a Lebesgue point. In general (7) means that the average of $|f - f(x_0)|$ is small on small balls centered at x_0. Lebesgue points are thus points, where f does not oscillate too much, in an average sense.

It is far from obvious that every $f \in \mathcal{L}_{\text{loc}}(\mathsf{R}^k)$ has Lebesgue points. Lebesgue showed that almost all points are Lebesgue points. □

THEOREM 3.5. *Let $f \in \mathcal{L}_{\text{loc}}$. The set $L(f)$ of Lebesgue points is a Borel set and $\complement L(f)$ is a Lebesgue null set.*

PROOF: Define

$$T_r(f)(x) = \frac{1}{m(K_r)} \int_{K(x,r)} |f - f(x)| dm$$

for $x \in \mathsf{R}^k$, $r > 0$, and

$$T(f)(x) = \limsup_{r \to 0} T_r(f)(x) .$$

Notice that

$$T_r(f)(x) \le \frac{1}{m(K_r)} \int_{K(x,r)} |f| dm + |f(x)| ,$$

and hence

$$T(f) \le M(f) + |f| . \tag{8}$$

Furthermore we have

$$\left.\begin{aligned} T_r(f+g) &\le T_r(f) + T_r(g) \\ T(f+g) &\le T(f) + T(g) \end{aligned}\right\} \tag{9}$$

for $f, g \in \mathcal{L}_{\text{loc}}(\mathsf{R}^k)$.

The Lebesgue set $L(f)$ is given by

$$L(f) = \{x \in \mathsf{R}^k \mid T(f)(x) = 0\} . \tag{10}$$

The function

$$F_r(x,y) = \begin{cases} \frac{1}{m(K_r)} |f(x) - f(y)|, & \text{if } |x-y| < r \\ 0, & \text{if } |x-y| \ge r \end{cases}$$

is non-negative and Borel measurable on $\mathbb{R}^{2k}$, so by Tonelli's theorem

$$T_r(f)(x) = \int F_r(x,y)dm(y)$$

is Borel measurable. To see that $T(f)$ is Borel measurable, we remark that

$$T(f)(x) = \limsup_n T_{\frac{1}{n}}(f)(x)$$

because of the inequality

$$T_r(f)(x) \leq (1+\frac{1}{n})^k T_{\frac{1}{n}}(f)(x)$$

for $r \in [\frac{1}{n+1}, \frac{1}{n}[$, and it follows that $L(f) \in \mathbb{B}_k$. We prove $m(\complement L(f)) = 0$ in two steps.

1° Assume $f \in \mathcal{L}_1(\mathbb{R}^k)$.

Let $\lambda > 0$ and $n \in \mathbb{N}$. By an approximation theorem (2 MA) there exists $g \in C_c(\mathbb{R}^k)$ so that $\|f-g\|_1 < 1/n$. Putting $h = f - g$ we have by (9)

$$T(f) = T(h+g) \leq T(h) + T(g) = T(h)$$

because $T(g) \equiv 0$ since g is continuous. Combined with (8) we get

$$T(f) \leq M(h) + |h| \ ,$$

hence

$$\{T(f) > 2\lambda\} \subseteq \{M(h) > \lambda\} \cup \{|h| > \lambda\} \ .$$

Combining the inequality

$$\|h\|_1 \geq \int_{\{|h|>\lambda\}} |h|dm \geq \lambda m(\{|h| > \lambda\})$$

with Theorem 3.3 (applied to $\mu = hm$) we find

$$m(\{T(f) > 2\lambda\}) \leq 3^k \frac{\|h\|_1}{\lambda} + \frac{\|h\|_1}{\lambda} < \frac{3^k+1}{n\lambda} \ . \tag{11}$$

The left-hand side of (11) is independent of n so $m(\{T(f) > 2\lambda\}) = 0$ for all $\lambda > 0$ showing that

$$m(\complement L(f)) = m(\{T(f) > 0\}) = 0 \ .$$

2° For $f \in \mathcal{L}_{\text{loc}}(\mathbb{R}^k)$ we note that

$$\complement L(f) \subseteq \bigcup_{n=1}^{\infty} \complement L(f 1_{K(0,n)}) \ ,$$

so $\complement L(f)$ is a null set by 1°. □

For every Lebesgue point x_0 of $f \in \mathcal{L}_{\mathrm{loc}}$ we have

$$\left|\frac{1}{m(K_r)}\int_{K(x_0,r)} f dm - f(x_0)\right| \leq \frac{1}{m(K_r)}\int_{K(x_0,r)} |f - f(x_0)| dm \to 0$$

and hence

$$\lim_{r\to 0}\frac{1}{m(K_r)}\int_{K(x_0,r)} f dm = f(x_0)\,. \tag{12}$$

In particular (12) holds for Lebesgue almost all $x_0 \in \mathbb{R}^k$.

Let $R(f)$ denote the set of points $x_0 \in \mathbb{R}^k$ for which the limit in (12) exists. Then $L(f) \subseteq R(f)$ and $R(f)$ remains the same if f is replaced by an equivalent function. However, if $x_0 \in L(f)$ and $f \sim f_1$, $f(x_0) \neq f_1(x_0)$ then (12) shows that $x_0 \notin L(f_1)$.

The function $f_{reg} : R(f) \to \mathbb{C}$ defined by

$$f_{reg}(x_0) = \lim_{r\to 0}\frac{1}{m(K_r)}\int_{K(x_0,r)} f dm \quad , \ x_0 \in R(f)$$

is "the most regular representative" for f. If we want a globally defined representative, we could define f_{reg} to be zero on $\complement R(f)$.

THEOREM 3.6. *Suppose μ is a complex measure on $(\mathbb{R}^k, \mathbb{B}_k)$ and $\mu \ll m$. Then the symmetric derivative*

$$D(\mu)(x) = \lim_{r\to 0}\frac{\mu(K(x,r))}{m(K(x,r))}$$

exists for Lebesgue almost all points of $\mathbb{R}^k$ and is a Radon-Nikodym derivative of μ with respect to m, i.e.

$$D(\mu) = \frac{d\mu}{dm}\,.$$

REMARK: A precise version of the last statement is: Every measurable extension $\tilde{D}(\mu)$ of $D(\mu)$ to $\mathbb{R}^k$ belongs to $\mathcal{L}(m)$ and $\mu = \tilde{D}(\mu)m$.

PROOF: Let $f \in \mathcal{L}(m)$ be so that $\mu = fm$. By (12) $D(\mu)$ exists at every point of $L(f)$ and is equal to f, so the assertion follows. □

DEFINITION 3.7: Suppose $x_0 \in \mathbb{R}^k$. A Sequence (E_n) of Borel sets in $\mathbb{R}^k$ is said to *shrink to x_0 nicely*, if there are a number $\alpha > 0$ and a sequence $r_n > 0$ tending to zero so that for $n \geq 1$

$$\left.\begin{array}{c} E_n \subseteq K(x_0, r_n) \\ m(E_n) \geq \alpha m(K(x_0, r_n)) \end{array}\right\}. \tag{13}$$

Note that we do not require x_0 to belong to E_n. The condition (13) means that E_n shall constitute a portion of $K(x_0, r_n)$ which is bounded below. □

Examples: (a) The standard cubes

$$]-\frac{1}{n}, \frac{1}{n}]^k \quad ,]0, \frac{1}{n}]^k$$

shrink to 0 nicely.

(b) Let $a_1, \cdots, a_k > 0$. The intervals

$$\prod_{i=1}^{k} [-\frac{a_i}{n}, \frac{a_i}{n}]$$

shrink to 0 nicely.

(c) The rectangles $[0, \frac{1}{n}] \times [0, \frac{1}{n^2}]$ do not shrink nicely to zero.

THEOREM 3.8. *Assume that* $(E_n(x))$ *shrink to* x *nicely for every* $x \in \mathbb{R}^k$ *and let* $f \in \mathcal{L}_{\text{loc}}$. *Then*

$$f(x) = \lim_{n \to \infty} \frac{1}{m(E_n(x))} \int_{E_n(x)} f dm \tag{14}$$

for any $x \in L(f)$, *in particular for Lebesgue almost all* $x \in \mathbb{R}^k$.

PROOF: Let $x \in L(f)$ and let $\alpha = \alpha(x)$ and $r_n = r_n(x)$ be the positive number and the radii associated to $E_n(x)$ according to (13). Then we have

$$\frac{\alpha(x)}{m(E_n(x))} \int_{E_n(x)} |f - f(x)| dm \leq \frac{1}{m(K(x, r_n))} \int_{K(x, r_n)} |f - f(x)| dm ,$$

and since the right-hand side converges to zero, so does the left-hand side, and (14) follows. □

We will now give two applications of Theorem 3.8.

THEOREM 3.9. *Let* $f \in \mathcal{L}_{\text{loc}}(\mathbb{R})$, $a \in \mathbb{R}$ *and define the primitive*

$$F(x) = \int_a^x f(t) dt \quad , x \in \mathbb{R} .$$

Then F *is differentiable at any* $x \in L(f)$ *with* $F'(x) = f(x)$. *In particular* F *is differentiable a.e.*

PROOF: Let $r_n > 0$ tend to zero and define $E_n(x) = [x, x + r_n]$. Then $(E_n(x))$ shrinks to x nicely, and we get by Theorem 3.8 that

$$\frac{F(x + r_n) - F(x)}{r_n} = \frac{1}{m(E_n(x))} \int_{E_n(x)} f dm \to f(x)$$

for every $x \in L(f)$. Letting $E_n(x) = [x - r_n, x]$ we similarly get that the left-hand derivative of F at x is $f(x)$. □

We next apply Theorem 3.8 to $f = 1_E$, where $E \in \mathbb{B}_k$, and find:

THEOREM 3.10. *For every Borel set $E \subseteq \mathbb{R}^k$ we have*

$$\lim_{r \to 0} \frac{m(E \cap K(x,r))}{m(K(x,r))} = 1$$

for Lebesgue almost all $x \in E$.

The same statement holds with nicely shrinking sequences instead of balls. The *metric density* of $E \in \mathbb{B}_k$ at $x \in \mathbb{R}^k$ is defined to be

$$\lim_{r \to 0} \frac{m(E \cap K(x,r))}{m(K(x,r))}$$

provided that this limit exists.

We see that the metric density of E is 1 at almost all points of E and 0 at almost all points of the complement of E.

We shall leave the differentiation theory for a while in order to examine to which extent a Borel measure μ on $\mathbb{R}^k$, i.e. a measure on $(\mathbb{R}^k, \mathbb{B}_k)$, is determined by its values on the families $\mathcal{F}$ and $\mathcal{G}$ of closed and open subsets of $\mathbb{R}^k$. Our first result is:

PROPOSITION 3.11. *Let μ be a Borel measure on $\mathbb{R}^k$ and assume $\mu(\mathbb{R}^k) < \infty$. For every $B \in \mathbb{B}_k$ we have*

$$\forall \varepsilon > 0 \; \exists F \in \mathcal{F} \; \exists G \in \mathcal{G} [F \subseteq B \subseteq G \, , \; \mu(G \setminus F) < \varepsilon] \, . \tag{15}$$

PROOF: Define $\mathbb{E}$ to be the family of Borel sets B for which (15) holds.

We claim that $\mathbb{E}$ is a σ-algebra, i.e. satisfying (i)–(iii) below:

(i) $\mathbb{R}^k \in \mathbb{E}$.

(ii) If $E \in \mathbb{E}$ then $\complement E \in \mathbb{E}$.

(iii) If $E_1, E_2, \cdots \in \mathbb{E}$ then $\bigcup_1^\infty E_n \in \mathbb{E}$.

Here (i) and (ii) are clear. To see (iii) let $\varepsilon > 0$ be given and assume that $E_1, E_2, \cdots \in \mathbb{E}$. For each $n \in \mathbb{N}$ we choose

$$F_n \in \mathcal{F} \, , \; G_n \in \mathcal{G}$$

so that $F_n \subseteq E_n \subseteq G_n$ and $\mu(G_n \setminus F_n) < \frac{\varepsilon}{2^n}$. Then

$$\Omega := \Big(\bigcup_1^\infty G_n\Big) \setminus \Big(\bigcup_1^\infty F_n\Big) \subseteq \bigcup_1^\infty (G_n \setminus F_n)$$

and hence

$$\mu(\Omega) \leq \sum_1^\infty \mu(G_n \setminus F_n) < \varepsilon \, .$$

However

$$\Omega_n := (\bigcup_{i=1}^{\infty} G_i) \setminus (\bigcup_{i=1}^{n} F_i) \searrow \Omega$$

and since $\mu(\mathbb{R}^k) < \infty$ we know that $\mu(\Omega_n) \searrow \mu(\Omega)$ and hence

$$\mu(\Omega_n) < \varepsilon$$

for n sufficiently big, showing that $\bigcup_1^\infty E_n \in \mathbb{E}$.

We finally claim that $\mathcal{G} \subseteq \mathbb{E}$ from which we conclude that $\mathbb{E} = \mathbb{B}_k$. To see that (15) holds for any open set G, it suffices to construct an increasing sequence (F_n) of closed subsets of G such that $\bigcup F_n = G$. This is clearly possible if $G = \mathbb{R}^k$, and if $G \neq \mathbb{R}^k$ we can define

$$F_n = \{x \in \mathbb{R}^k \mid \operatorname{dist}(x, \complement G) \geq \frac{1}{n}\}\ .$$

□

Let μ be a Borel measure on $\mathbb{R}^k$ with $\mu(\mathbb{R}^k) < \infty$. Then (15) is equivalent to the assertion

$$\mu(B) = \sup\{\mu(F) \mid F \in \mathcal{F}, F \subseteq B\} = \inf\{\mu(G) \mid G \in \mathcal{G}, B \subseteq G\} \tag{16}$$

for any $B \in \mathbb{B}_k$. Since any closed set $F \subseteq \mathbb{R}^k$ is the union of an increasing sequence of compact sets

$$F \cap \{x \in \mathbb{R}^k \mid |x| \leq n\}\ ,$$

we see that μ is *inner regular* (or *tight*) in the following sense:

$$\mu(B) = \sup\{\mu(K) \mid K \text{ compact}, K \subseteq B\} \text{ for any } B \in \mathbb{B}_k\ . \tag{17}$$

By the second equality in (16) we have

$$\mu(B) = \inf\{\mu(G) \mid G \in \mathcal{G}\ ,\ B \subseteq G\} \text{ for any } B \in \mathbb{B}_k\ , \tag{18}$$

and we say that μ is *outer regular.*

A Borel measure μ on $\mathbb{R}^k$ is called *regular* if (17) and (18) hold. We have seen that any finite Borel measure is regular.

DEFINITION 3.12: A Borel measure μ on $\mathbb{R}^k$ is called a *Radon measure* if

$$\mu(K) < \infty \text{ for any compact } K \subseteq \mathbb{R}^k\ .$$

THEOREM 3.13. *A Radon measure μ on $\mathbb{R}^k$ is σ-finite and regular.*

PROOF: A Radon measure is clearly σ-finite. For each $n \in \mathbb{N}$ let μ_n denote the finite Borel measure

$$\mu_n(B) = \mu(B \cap \{x \in \mathbb{R}^k \mid |x| \leq n\}) \ , \ B \in \mathbb{B}_k \ .$$

Let $B \in \mathbb{B}_k$ and $\lambda < \mu(B)$. Clearly $\lambda < \mu_n(B)$ for n sufficiently big, and by (17) there exists a compact set $K \subseteq B$ so that

$$\lambda < \mu_n(K) = \mu(K \cap \{x \in \mathbb{R}^k \mid |x| \leq n\}) \ ,$$

which shows that μ is tight.

To see the outer regularity we can assume $\mu(B) < \infty$. Let $\varepsilon > 0$ be given. Applying the outer regularity to each of the finite Borel measures

$$E \mapsto \mu(E \cap K(0,n)) \ ,$$

we can find open sets $G_n \supseteq B$ so that

$$\mu(G_n \cap K(0,n)) < \mu(B \cap K(0,n)) + \frac{\varepsilon}{2^n} \ .$$

The open set $G := \bigcup_{n=1}^{\infty} G_n \cap K(0,n)$ contains B and

$$\mu(G \setminus B) \leq \sum_{n=1}^{\infty} \mu((G_n \setminus B) \cap K(0,n)) < \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon \ ,$$

showing that $\mu(G) < \mu(B) + \varepsilon$. □

THEOREM 3.14. *Associate to each $x \in \mathbb{R}^k$ a sequence $(E_n(x))$ which shrinks to x nicely. If μ is a complex Borel measure and $\mu \perp m$ then*

$$\lim_{n\to\infty} \frac{\mu(E_n(x))}{m(E_n(x))} = 0 \quad m-a.e. . \tag{19}$$

PROOF: By the Jordan decomposition of μ it suffices to prove (19) under the assumption $\mu \geq 0$. Arguing as in the proof of Theorem 3.8 we find

$$\frac{\alpha(x)\mu(E_n(x))}{m(E_n(x))} \leq \frac{\mu(E_n(x))}{m(K(x,r_n))} \leq \frac{\mu(K(x,r_n))}{m(K(x,r_n))} \ ,$$

so (19) follows once it is established that $D(\mu)(x) = 0$ m-a.e..

For each $n \in \mathbb{N}$ the function

$$M_n(\mu)(x) = \sup_{0<r<\frac{1}{n}} Q_r(\mu)(x) ,$$

which is similar to the maximal function, is lower semicontinuous, and this is seen as in Lemma 3.1. Since $M_n(\mu)(x)$ is decreasing in n, the following limit exists

$$\overline{D}(\mu)(x) = \lim_{n\to\infty} M_n(\mu)(x)$$

and defines a Borel function $\overline{D}(\mu) : \mathbb{R}^k \to [0,\infty]$ called the *upper symmetric derivative* of μ.

Let $\lambda > 0$ and $\varepsilon > 0$ be given. Since $\mu \perp m$, μ is concentrated on a Borel set B for which $m(B) = 0$. By inner regularity of μ there exists a compact subset $K \subseteq B$ with $\mu(K) > \mu(B) - \varepsilon$. Let $\mu_1(E) = \mu(E \cap K)$ and $\mu_2(E) = \mu(E \setminus K)$ for $E \in \mathbb{B}_k$. Then $\mu = \mu_1 + \mu_2$ and $\mu_2(\mathbb{R}^k) = \mu(\mathbb{R}^k \setminus K) = \mu(B \setminus K) < \varepsilon$. For any $x \notin K$ and $r < \operatorname{dist}(x, K)$ we have

$$Q_r(\mu)(x) = Q_r(\mu_2)(x)$$

and hence

$$\overline{D}(\mu)(x) = \overline{D}(\mu_2)(x) \le M(\mu_2)(x) \quad \text{for } x \notin K .$$

This implies

$$\{\overline{D}(\mu) > \lambda\} \subseteq K \cup \{M(\mu_2) > \lambda\} ,$$

so by Theorem 3.3 we have

$$m(\{\overline{D}(\mu) > \lambda\}) \le m(K) + \frac{3^k}{\lambda}\|\mu_2\| < \frac{3^k\varepsilon}{\lambda} . \tag{20}$$

Since $\varepsilon > 0$ was arbitrary $m(\{\overline{D}(\mu) > \lambda\}) = 0$ for every $\lambda > 0$, and this shows that $\overline{D}(\mu) = 0$ m-a.e.. □

Theorems 3.8 and 3.14 can be combined to the following:

THEOREM 3.15. *Associate to each $x \in \mathbb{R}^k$ a sequence $(E_n(x))$ which shrinks to x nicely. Let μ be a complex Borel measure on $\mathbb{R}^k$ with Lebesgue decomposition $\mu = fm + \mu_s$ where $f \in \mathcal{L}_1(m)$ and $\mu_s \perp m$.*
Then

$$\lim_{n\to\infty} \frac{\mu(E_n(x))}{m(E_n(x))} = f(x) \qquad m - \text{a.e.} .$$

In particular $\mu \perp m$ if and only if $D(\mu)(x) = 0$ m-a.e..

The regularity of a Radon measure μ on $\mathbb{R}^k$ leads to various approximation results for integrable functions. We recall that the space $C_c(\mathbb{R}^k)$ of continuous complex valued functions with compact support is dense in $\mathcal{L}_p(\mathbb{R}^k, \mu)$ for any $1 \le p < \infty$. (Cf. Mat 2MA, II,7). In the next result we use definitions from exercise 13.

THEOREM 3.16. *(Vitali-Carathéodory). Let $f : \mathbb{R}^k \to \mathbb{R}$ be integrable with respect to a Radon measure μ on $\mathbb{R}^k$. For any $\varepsilon > 0$ there exist functions $u \leq f \leq v$ so that*

(i) *u is upper semicontinuous and bounded above,*

(ii) *v is lower semicontinuous and bounded below,*

(iii) *$\int(v-u)d\mu < \varepsilon$.*

PROOF: 1° Assume first $f \geq 0$. We know that there exists a sequence (s_n) of simple non-negative measurable functions so that $s_n \nearrow f$. Defining $t_n = s_n - s_{n-1}$, $n \geq 1$ with $s_0 := 0$, we see that $f = \sum_1^\infty t_n$, and since t_n is simple there exists a sequence of Borel sets (E_j), (not necessarily disjoint) and constants $c_j > 0$ so that

$$f(x) = \sum_{j=1}^\infty c_j 1_{E_j}(x) ,$$

and hence

$$\int f d\mu = \sum_{j=1}^\infty c_j \mu(E_j) , \tag{21}$$

where the series in (21) is convergent. By the regularity of μ there exist compact sets K_j and open sets G_j such that $K_j \subseteq E_j \subseteq G_j$ and

$$\mu(G_j \setminus K_j) < \frac{\varepsilon}{c_j} 2^{-j-1} . \tag{22}$$

Put

$$v = \sum_{j=1}^\infty c_j 1_{G_j} , \quad u = \sum_{j=1}^N c_j 1_{K_j} ,$$

where N is chosen so that

$$\sum_{j=N+1}^\infty c_j \mu(E_j) < \frac{\varepsilon}{2} . \tag{23}$$

Then $u \leq f \leq v$ satisfy (i) and (ii), and (iii) holds because

$$v - u = \sum_{j=1}^N c_j 1_{G_j \setminus K_j} + \sum_{j=N+1}^\infty c_j 1_{G_j} ,$$

so

$$\int (v-u)d\mu \le \sum_{j=1}^{N} c_j\mu(G_j \setminus K_j) + \sum_{j=N+1}^{\infty} c_j\mu(G_j)$$
$$= \sum_{j=1}^{\infty} c_j\mu(G_j \setminus K_j) + \sum_{j=N+1}^{\infty} c_j\mu(K_j) < \varepsilon$$

by (22) and (23).

2°. In the real case we write $f = f^+ - f^-$ and apply 1° to f^+ and f^-. If $u_1 \le f^+ \le v_1$, $u_2 \le f^- \le v_2$ where u_i satisfy (i), v_i satisfy (ii) and $\int (v_i - u_i)d\mu < \varepsilon/2$ then $u_1 - v_2 \le f \le v_1 - u_2$ and

$$\int ((v_1 - u_2) - (u_1 - v_2))\, d\mu < \varepsilon \, .$$

Furthermore $u_1 - v_2$ is sum of the upper semicontinuous functions u_1 and $-v_2$ and hence upper semicontinuous. It is also bounded above. Similarly $v_1 - u_2$ satisfies (ii). □

The fundamental theorem of calculus

This theorem says roughly speaking that differentiation and integration with respect to Lebesgue measure are inverse operations on functions. It has therefore two parts. The first part says classically that if f is continuous then

$$F(x) = \int_a^x f(t)dt$$

is differentiable with $F'(x) = f(x)$. Theorem 3.9 extends this part to $f \in \mathcal{L}_{\mathrm{loc}}(\mathbb{R})$.

The second part says classically, that if $f : [a,b] \to \mathbb{C}$ is continuously differentiable then

$$f(x) - f(a) = \int_a^x f'(t)dt \, , \; x \in [a,b] \, . \tag{24}$$

There are several guesses for generalizations of this part:

(a) Is it enough to assume differentiability of f? The answer is *no*, because although f' is measurable, it is not necessarily locally integrable.

Put $f(x) = x^2 \sin(1/x^2)$ for $x \neq 0$ and $f(0) = 0$. Then f is differentiable but

$$\int_0^1 |f'(t)|dt = \infty$$

because

$$\int_0^1 \frac{|\cos(1/t^2)|}{t}dt = \frac{1}{2}\int_1^{\infty}\frac{|\cos x|}{x}dx$$

$$> \frac{1}{2}\sum_{n=1}^{\infty}\int_{(2n-1)\frac{\pi}{2}}^{(2n+1)\frac{\pi}{2}}\frac{|\cos x|}{x}dx = \frac{1}{2}\sum_{n=1}^{\infty}\int_0^{\pi}\frac{\sin u}{u+(n-\frac{1}{2})\pi}du > \frac{1}{\pi}\sum_{n=1}^{\infty}\frac{1}{n+\frac{1}{2}} = \infty\,.$$

(b) Suppose $f : [a,b] \to \mathbb{C}$ is continuous on $[a,b]$ and differentiable for almost all $x \in [a,b]$, and that f' is equal a.e. to an integrable function on $[a,b]$. Is it then true that (24) holds?

Again the answer is *no* as the following example shows.

EXAMPLE 3.17. This example is connected to the Cantor set C.

We shall construct a sequence $(E_n)_{n\geq 0}$ of closed sets and a sequence of continuous functions $f_n : [0,1] \to [0,1]$, $n \geq 0$, each being piecewise linear.

We put $E_0 = [0,1]$ and $f_0(x) = x$.

We divide E_0 in 3 compact intervals E_{00}, E_{01}, E_{02} of equal length 1/3 and define f_1 as the piecewise linear function which is constant 1/2 on the middle interval E_{01} and connects $(0,0)$ to $(1/3, 1/2)$ on E_{00} and connects $(2/3, 1/2)$ to $(1,1)$ on E_{02}.

We put $E_1 = E_{00} \cup E_{02}$.

In the next step we divide each of the two intervals E_{00} and E_{02} in 3 compact intervals of equal length $1/3^2$ and remove the middle open intervals in order to get E_2 as union of 2^2 compact intervals of length $1/3^2$. To get f_2 from f_1 we change f_1 on E_{00} to be constant $1/2^2$ on the middle interval and linear on the remaining two intervals. Similarly we change f_1 on E_{02} to be constant $3/2^2$ on the middle interval, linear on the remaining two intervals. On E_{01} we have $f_2 = f_1 = 1/2$.

In the next step each of the 2^2 intervals is divided in 3 compact intervals of equal length. The middle intervals are removed and E_3 is the union of the 2^3 remaining intervals. We define $f_3 = f_2$ on $[0,1] \setminus E_2$ and on each of the 2^2 intervals of E_2 we change f_2 to be constant on the middle interval and linear on the two others. The constant value is the mean value of the value in the end points. Continuing in this way, we have

$$m(E_n) = (\frac{2}{3})^n\,,$$

and for $x \in [0,1]$

$$|f_0(x) - f_1(x)| \leq \frac{1}{2} - \frac{1}{3} = \frac{1}{6}\,,\ |f_n(x) - f_{n+1}(x)| \leq \frac{1}{6}\cdot\frac{1}{2^n}\,.$$

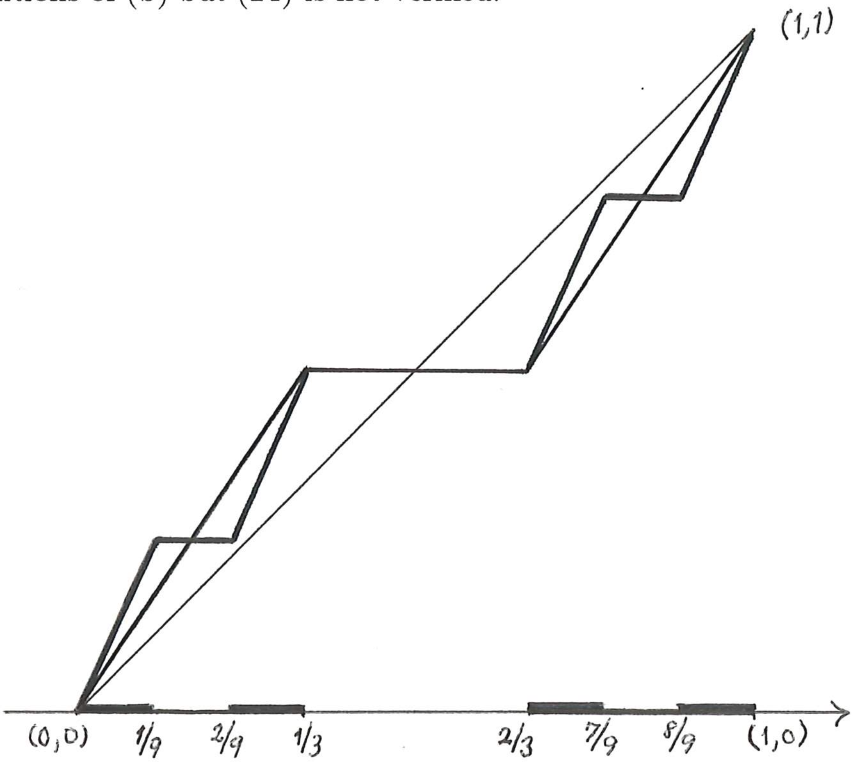
The Cantor set is defined as $C = \bigcap_0^\infty E_n$, which is compact with $m(C) = 0$.

The sequence (f_n) is a Cauchy sequence in the Banach space $C([0,1])$ with the uniform norm, because

$$\|f_n - f_{n+p}\|_u \le \sum_{j=n}^{n+p-1} \|f_j - f_{j+1}\|_u \le \frac{1}{6}\sum_{j=n}^{n+p-1} 2^{-j} < \frac{1}{3}2^{-n} ,$$

and therefore (f_n) converges uniformly to a continuous function $f : [0,1] \to [0,1]$ which is increasing since every f_n is so. By construction f is constant $1/2$ in the open interval $]1/3, 2/3[$, and similarly it is constant in each of the open intervals which are removed in order to get the Cantor set, the constant values being of the form $k/2^n$. It follows that f is differentiable in all points of $[0,1] \setminus C$ with $f'(x) = 0$, i.e. in Lebesgue almost all points.

The function f is called *Cantor-Lebesgue's singular function.* It satisfies the conditions of (b) but (24) is not verified.



Using dyadic and triadic numbers we can describe Cantor-Lebesgue's function. If $p \ge 2$ is an integer, a p-adic fraction is a symbol of the form ${}^p0, a_1a_2\cdots$, where $a_i \in \{0, 1, \cdots, p-1\}$, and it represents the real number

$$\sum_{i=1}^{\infty} \frac{a_i}{p^i} .$$

From E_n we remove 2^n open intervals of the form

$$]^30,a_1a_2\cdots a_n022\cdots,{}^30,a_1a_2\cdots a_n200\cdots[\,,$$

where $a_i \in \{0,2\}$, $i=1,\cdots,n$ and

$$f_{n+1}({}^30,a_1a_2\cdots a_n022\cdots) = f_{n+1}({}^30,a_1a_2\cdots a_n200\cdots) \\ = {}^20,\frac{a_1}{2}\frac{a_2}{2}\cdots\frac{a_n}{2}100\cdots\,.$$

Note that

$${}^30,a_1a_2\cdots a_n022\cdots \;=\; {}^30,a_1a_2\cdots a_n100\cdots\,.$$

On such an interval f is constant with the same value as f_{n+1} has at the end points.

Every point $x \in C$ has a unique representation as triadic fraction $x = {}^30,a_1a_2\cdots$ where $a_i \in \{0,2\}$ and by continuity of f we get

$$f({}^30,a_1a_2\cdots) = \lim_{n\to\infty} f({}^30,a_1\cdots a_n200\cdots) = \lim_{n\to\infty} f_{n+1}({}^30,a_1\cdots a_n200\cdots) \\ = {}^20,\frac{a_1}{2}\frac{a_2}{2}\cdots\,.$$

We shall finally give a true form of the second part of the fundamental theorem of calculus:

THEOREM 3.18. *If $f:[a,b]\to\mathbb{C}$ is differentiable at every point of $[a,b]$ and $f'\in\mathcal{L}([a,b])$ then*

$$f(x)-f(a) = \int_a^x f'(t)dt \quad \text{for } x\in[a,b]\,.$$

Note that differentiability is required at *every* point of $[a,b]$, implying continuity of f.

PROOF: It is clearly enough to consider real functions and to prove the formula for $x=b$. Fix $\varepsilon>0$. By Theorem 3.16 there exists a lower semicontinuous function $g\geq f'$ such that

$$\int_a^b g(t)dt < \int_a^b f'(t)dt + \varepsilon\,. \tag{25}$$

By adding a sufficiently small positive constant to g we can still have (25) and furthermore $g(x)>f'(x)$ for all $x\in[a,b]$. Let $\eta>0$ be arbitrary and define

$$F_\eta(x) = \int_a^x g(t)dt - f(x) + f(a) + \eta(x-a)\,,\ x\in[a,b]\,. \tag{26}$$

To each $x \in [a, b[$ there corresponds a $\delta_x > 0$ such that

$$g(t) > f'(x) \quad \text{and} \quad \frac{f(t) - f(x)}{t - x} < f'(x) + \eta$$

for all $t \in]x, x + \delta_x[$, simply because g is lower semicontinuous at x and f is differentiable (from the right) at x. For these values of t we then have

$$\begin{aligned} F_\eta(t) - F_\eta(x) &= \int_x^t g(s)ds - (f(t) - f(x)) + \eta(t - x) \\ &> (t - x)f'(x) - (t - x)(f'(x) + \eta) + \eta(t - x) = 0 \, , \end{aligned}$$

and we have shown the following:

$$\forall x \in [a, b[\; \exists \delta_x > 0 \; \forall t \in]x, x + \delta_x[\, : F_\eta(x) < F_\eta(t) \, . \tag{27}$$

The set $A = \{x \in [a, b] \mid F_\eta(x) \geq 0\}$ contains a because $F_\eta(a) = 0$, and A is closed since F_η is continuous. Putting $x_0 = \sup A$ then $x_0 \in A$. We claim that $x_0 = b$ for otherwise $x_0 < b$ and by (27) $F_\eta(t) > F_\eta(x_0) \geq 0$ for $t \in]x_0, x_0 + \delta_{x_0}[$ contradicting the definition of x_0.

It follows that $F_\eta(b) \geq 0$, so by (26) we have

$$f(b) - f(a) \leq \int_a^b g(t)dt + \eta(b - a) \, .$$

Since this holds for any $\eta > 0$ we get

$$f(b) - f(a) \leq \int_a^b g(t)dt < \int_a^b f'(t)dt + \varepsilon \, ,$$

and since $\varepsilon > 0$ was arbitrary

$$f(b) - f(a) \leq \int_a^b f'(t)dt \, .$$

Applying this inequality to $-f$, which also satisfies the conditions of the theorem, we get the opposite inequality. $\square$

Appendix 2. Liapounov's theorem

Søren Eilers

In this section we are going to prove that the range of a finite dimensional vector measure is always closed, and, under certain restrictions, convex.

Let (X, E) be a measurable space. An *m-dimensional real vector measure*, as defined in exercise 4, is a countably additive mapping $\mu : \mathsf{E} \to \mathbf{R}^m$.

Most of the theory we have developed for complex measures also applies to m-dimensional vector measures, since if we write

$$\mu = (\mu_1, \ldots, \mu_m)$$

μ is countably additive if and only if all the μ_i's are so. Thus μ is nothing but an m-dimensional vector function whose m components are real measures.

We equip $\mathbf{R}^m$ with the 1-norm and define the *absolute value* of μ as

$$|\mu|(E) = \sup\{\sum_{i=1}^{\infty} |\mu(E_i)|_1 \mid (E_i) \text{ partition of } E\} .$$

Using $|\cdot|_1$ instead of the euclidean norm has the advantage that

$$|\mu| = \sum_{j=1}^{m} |\mu_j| \tag{1}$$

where $|\mu_j|$ are the ordinary absolute values.

To see this, note first that

$$\sum_{i=1}^{\infty} |\mu(E_i)|_1 = \sum_{j=1}^{m} \sum_{i=1}^{\infty} |\mu_j(E_i)| \le \sum_{j=1}^{m} |\mu_j|(E) ,$$

so $|\mu|(E) \le \sum_j^m |\mu_j|(E)$. For the other inequality, for any $\varepsilon > 0$ and $j \in \{1, \ldots, m\}$ take a partition $\{E_i^j\}_{i\in\mathbf{N}}$ of E, so that

$$\sum_{i=1}^{\infty} |\mu_j(E_i^j)| \ge |\mu_j|(E) - \varepsilon/m .$$

Then $\{E_{i_1}^1 \cap \cdots \cap E_{i_m}^m \mid (i_1, \ldots, i_m) \in \mathbf{N}^m\}$ is also a partition of E, and writing it $\{F_\alpha \mid \alpha \in \mathbf{N}^m\}$ gives us

$$\sum_{i=1}^{\infty} |\mu_j(E_i^j)| \leq \sum_{\alpha} |\mu_j(F_\alpha)|$$

and hence

$$\sum_{j=1}^{m} |\mu_j|(E) \leq \varepsilon + \sum_{j=1}^{m} \sum_{\alpha} |\mu_j(F_\alpha)|$$
$$= \varepsilon + \sum_{\alpha} |\mu(F_\alpha)|_1 \leq |\mu|(E) + \varepsilon \ .$$

Since $\varepsilon > 0$ was arbitrary (1) follows.

For μ non-negative, i.e. with each $\mu_i \geq 0$, we thus have $|\mu| = \sum_{i=1}^{m} \mu_i$.

For vector measures $\mu : (X, \mathsf{E}) \to \mathbb{R}^m$ and $\nu : (X, \mathsf{E}) \to \mathbb{R}^n$ we say that ν is *absolutely continuous* with respect to μ (in symbols $\nu \ll \mu$), if

$$\forall E \in \mathsf{E}(|\mu|(E) = 0 \Rightarrow \nu(E) = 0) \ .$$

DEFINITION 1: A vector measure $\mu : (X, \mathsf{E}) \to \mathbb{R}^m$ is called *semi-convex* if for every $E \in \mathsf{E}$ there exists $F \subseteq E$, $F \in \mathsf{E}$ so that

$$\mu(F) = \frac{1}{2}\mu(E) \ .$$

Given $E \in \mathsf{E}$, we say that $\varphi \in K(\mu, E)$ if $\varphi : E \to [0, 1[$ is measurable and

$$\forall \lambda \in [0, 1] : \mu(\{\varphi < \lambda\}) = \lambda \mu(E) \ .$$

We say that μ is *convex*, when $K(\mu, E) \neq \emptyset$ for every $E \in \mathsf{E}$.

A convex measure is clearly semi-convex. The converse is far from obvious but will be proved in Lemma 4 below. Furthermore, if μ is a semi-convex vector measure and $A \in \mathsf{E}$, then the restriction $\mu_A : \mathsf{E} \to \mathbb{R}^m$ defined by

$$\mu_A(E) = \mu(A \cap E)$$

is semi-convex.

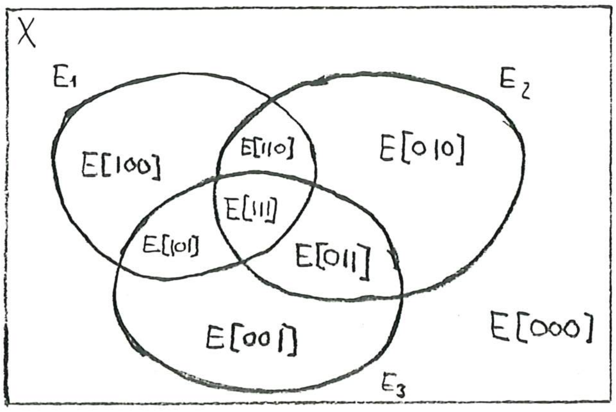
DEFINITION 2: The following notation will be useful in the sequel. When $E_1, \ldots, E_k$ is a finite family of subsets of E, to each binary number

$$b = (b_1, \ldots, b_k) \in B_k := \{0, 1\}^k$$

we associate a set

$$E[b] := \bigcap_{i=1}^{k} E_i^{b_i} \ ,$$

where $A^0 = E \setminus A$, $A^1 = A$. The sets $E[b], b \in B_k$ form a partition of E.



LEMMA 3. *Let $\mu : (X, \mathsf{E}) \to \mathbb{R}^m$ be a semi-convex vector measure. For each $E \in \mathsf{E}$ there exists a sequence $(E_n)_{n \in \mathbb{N}}$ from E of subsets of E so that for all $k \in \mathbb{N}$, any numbers $1 \leq n_1 < n_2 < \cdots < n_k$ and any $b \in B_k$*

$$\mu\left(\bigcap_{i=1}^{k} E_{n_i}^{b_i}\right) = \frac{1}{2^k}\mu(E) \ . \tag{2}$$

PROOF: From the semi-convexity we can take $E_1 \subseteq E$, $E_1 \in \mathsf{E}$ so that $\mu(E_1) = \frac{1}{2}\mu(E)$. Thus, (2) is satisfied for $n_1 = 1$, $k = 1$.

Now assume that $E_1, \ldots, E_n$ has been chosen so that (2) holds, when $1 \leq n_1 < \cdots < n_k \leq n$. For $b \in B_n$ write

$$E[b] := \bigcap_{i=1}^{n} E_i^{b_i}$$

and pick $F[b] \subseteq E[b]$ from E so that $\mu(F[b]) = \frac{1}{2}\mu(E[b])$. We can take

$$E_{n+1} := \bigcup_{b \in B_n} F[b] \ .$$

To see this we must prove that if $1 \leq n_1 < \cdots < n_{k-1} \leq n$ and if $n_k = n+1$, then (2) holds for each $b \in B_k$. Fix $b \in B_k$ and define

$$\tilde{B} = \{c \in B_n \mid \forall i \in \{1, \ldots, k-1\} : c_{n_i} = b_i\} \ ,$$

which has exactly 2^{n-k+1} elements. It is easy to see that

$$\bigcup_{i=1}^{k-1} E_{n_i}^{b_i} = \bigcup_{c\in\tilde{B}} E[c]\,.$$

When $b_k = 1$ we have

$$\bigcap_{i=1}^{k} E_{n_i}^{b_i} = (\bigcup_{c\in\tilde{B}} E[c]) \cap (\bigcup_{b\in B_n} F[b]) = \bigcup_{c\in\tilde{B}} F[c]\,,$$

and when $b_k = 0$ we have

$$\bigcap_{i=1}^{k} E_{n_i}^{b_i} = (\bigcup_{c\in\tilde{B}} E[c]) \setminus (\bigcup_{b\in B_n} F[b]) = \bigcup_{c\in\tilde{B}} E[c] \setminus F[c]\,,$$

so in both cases we get

$$\begin{aligned}\mu(\bigcap_{i=1}^{k} E_{n_i}^{b_i}) &= \frac{1}{2}\sum_{c\in\tilde{B}} \mu(E[c]) \\ &= \frac{1}{2}\frac{1}{2^n}\sum_{c\in\tilde{B}} \mu(E) = \frac{1}{2^k}\mu(E)\,.\end{aligned}$$

□

LEMMA 4. *A semi-convex vector measure $\mu : (X, \mathsf{E}) \to \mathsf{R}^m$ is convex.*

PROOF: For given $E \in \mathsf{E}$ choose $(E_n)_{n\in\mathsf{N}}$ from Lemma 3. Put

$$E_* = \liminf_n E_n = \bigcup_{n\geq 1}\bigcap_{p\geq 0} E_{n+p}$$

($x \in E_*$ if x is in E_n eventually) and define

$$\varphi(x) = \left(\sum_{n=1}^{\infty} \frac{1}{2^n} 1_{E_n}(x)\right) \cdot 1_{E\setminus E_*}(x)\,.$$

The function φ is obviously well defined and measurable with $\varphi(X) \subseteq [0,1[$. We will prove $\varphi \in K(\mu, E)$. For a dyadic number $k/2^n$, $k \in \{1,\ldots,2^n\}$ we have

$$\{\varphi < k/2^n\} = E_* \cup \{\sum_{i=1}^{n} \frac{1}{2^i} 1_{E_i} < k/2^n\}\,.$$

The inclusion from left to right is clear.

If an x in the rightmost set does not come from E_*, x must satisfy

$$\sum_1^n \frac{1}{2^i} 1_{E_i}(x) \leq \frac{k-1}{2^n} < \frac{k}{2^n}$$

since the sum must have the form $l/2^n$. Since $x \notin E_*$ we have

$$\sum_{n+1}^{\infty} \frac{1}{2^i} 1_{E_i}(x) < \frac{1}{2^n} ,$$

so that $\varphi(x) < k/2^n$.

Putting

$$B^* = \{b \in B_n \mid \sum_{i=1}^{n} 2^{n-i} b_i < k\} ,$$

we have in the terminology from Lemma 3

$$\{\sum_1^n \frac{1}{2^i} 1_{E_i} < \frac{k}{2^n}\} = \{\sum_1^n 2^{n-i} 1_{E_i} < k\} = \bigcup_{b \in B^*} E[b] ,$$

where B^* has exactly k elements (think binarily!). From (2) above, we see that E_* is a null set and that

$$\begin{aligned}\mu\left(\{\varphi < \frac{k}{2^n}\}\right) &= \mu\left(\{\sum_1^n \frac{1}{2^i} 1_{E_i} < \frac{k}{2^n}\}\right) = \mu\left(\bigcup_{b \in B^*} E[b]\right) \\ &= \sum_{b \in B^*} \mu(E[b]) = \sum_{b \in B^*} \frac{1}{2^n} \mu(E) = \frac{k}{2^n} \mu(E) .\end{aligned}$$

Any $\lambda \in]0,1]$ is limit of an increasing sequence of dyadic numbers λ_i. From the countable additivity

$$\mu(\{\varphi < \lambda\}) = \mu\left(\bigcup_1^{\infty} \{\varphi < \lambda_i\}\right) = \lim_i \mu(\{\varphi < \lambda_i\}) = \lim_i \lambda_i \mu(E) = \lambda \mu(E) .$$

□

LEMMA 5. *Let $\mu : (X, \mathsf{E}) \to \mathsf{R}^m$ be a convex and non-negative vector measure and let $E \in \mathsf{E}$, $\varphi \in K(\mu, E)$. If $\nu : (X, \mathsf{E}) \to \mathsf{R}^n$ is a vector measure such that $\nu \ll \mu$, then the function $N : [0,1] \to \mathsf{R}^n$ given by*

$$N(\lambda) = \nu(\{\varphi < \lambda\})$$

is continuous.

PROOF: By definition each component ν_i of ν satisfies $\nu_i \ll |\mu|$. From n applications of Theorem 2.2 we get, using the fact that $|\mu|(F) = |\mu(F)|_1$

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall F \in \mathsf{E} : |\mu(F)|_1 < \delta \Rightarrow |\nu(F)|_1 < \varepsilon\,.$$

Given ε, when $|\lambda_1 - \lambda_2| < \delta / |\mu(E)|$ with δ as above and $\lambda_2 \leq \lambda_1$, we have

$$\begin{aligned} |\mu(\{\lambda_2 \leq \varphi < \lambda_1\})|_1 &= |\mu(\{\varphi < \lambda_1\}) - \mu(\{\varphi < \lambda_2\})|_1 \\ &= |\lambda_1 \mu(E) - \lambda_2 \mu(E)|_1 < \delta \end{aligned}$$

so that

$$\begin{aligned} |N(\lambda_1) - N(\lambda_2)|_1 &= |\nu(\{\varphi < \lambda_1\}) - \nu(\{\varphi < \lambda_2\})|_1 \\ &= |\nu(\{\lambda_2 \leq \varphi < \lambda_1\})|_1 < \varepsilon\,. \end{aligned}$$

THEOREM 6. *When $\mu : (X, \mathsf{E}) \to \mathsf{R}^m$ is a convex vector measure and $E, F \in \mathsf{E}$, there exists a family $\{C(\lambda) | \lambda \in [0,1]\}$ of sets from E so that*

(i) $C(0) = E\,,\ C(1) = F$.

(ii) $\mu(C(\lambda)) = (1 - \lambda)\mu(E) + \lambda\mu(F)$.

Furthermore, if μ is non-negative and $\nu \ll \mu$ then

(iii) $\lambda \mapsto \nu(C(\lambda))$ *is continuous.*

In particular, the range of a convex vector measure is convex.

PROOF: Let $\varphi \in K(\mu, E \setminus F)$, $\psi \in K(\mu, F \setminus E)$ and define

$$C(\lambda) = (E \cap F) \cup \{\varphi < 1 - \lambda\} \cup \{\psi < \lambda\}\,.$$

Then (i) follows by definition, and (iii) follows by Lemma 5. For (ii) we have

$$\begin{aligned} \mu(C(\lambda)) &= \mu(E \cap F) + \mu(\{\varphi < 1 - \lambda\}) + \mu(\{\psi < \lambda\}) \\ &= \mu(E \cap F) + (1 - \lambda)\mu(E \setminus F) + \lambda\mu(F \setminus E) \\ &= (1 - \lambda)[\mu(E \cap F) + \mu(E \setminus F)] + \lambda[\mu(E \cap F) + \mu(F \setminus E)] \\ &= (1 - \lambda)\mu(E) + \lambda\mu(F)\,. \end{aligned}$$

□

DEFINITION 7: If ν is a real measure on (X, E) and $E \in \mathsf{E}$ has the properties

$$\begin{aligned} &\nu(E) \neq 0 \\ &\nu(F) \in \{0, \nu(E)\} \quad \forall F \subseteq E\,,\ F \in \mathsf{E}, \end{aligned}$$

then E is called an *atom* of ν.

A vector measure $\mu : (X, \mathsf{E}) \to \mathsf{R}^m$, is called *purely non-atomic* if none of the coordinate measures μ_i has any atoms. It is called *purely atomic*, if $(E_i)_{i \in I}$ is a countable family of disjoint sets that are atoms for all μ_i, so that

$$X = \bigcup_{i \in I} E_i \,.$$

EXAMPLES. The Dirac measure ε_a on (R, B) has an atom $\{a\}$. If the σ-algebra E on X consists of $\emptyset$ and X only, then X is an atom for every non-zero measure μ. The measure

$$\mu = \sum_{n=1}^{\infty} \frac{1}{2^n} \varepsilon_n$$

on $(\mathsf{N}, \mathcal{P}(\mathsf{N}))$ is purely atomic.

LEMMA 8. *A purely non-atomic (finite) measure* $\mu : (X, \mathsf{E}) \to [0, \infty[$ *is semi-convex.*

PROOF: Given $E \in \mathsf{E}$ we must find $F \in \mathsf{E}\,,\ F \subseteq E$, so that

$$\mu(F) = \frac{1}{2}\mu(E) \,.$$

We may assume $\mu(E) \neq 0$. For any $\varepsilon \in]0, 1]$ there exists $F_\varepsilon \subseteq E\,,\ F_\varepsilon \in \mathsf{E}$ so that

$$0 \leq \mu(F_\varepsilon) \leq \varepsilon \mu(E) \,. \tag{3}$$

If $\varepsilon < 2^{-n}$ this follows from n applications of the fact that μ is purely non-atomic, i.e. for every $G \in \mathsf{E}$ with $\mu(G) > 0$ there exists a partition $\{G_1, G_2\}$ of G so that

$$\mu(G) = \mu(G_1) + \mu(G_2) \text{ with } \mu(G_i) > 0\,,\ i = 1, 2,$$

so that every $G \in \mathsf{E}$ contains a subset $G' \in \mathsf{E}$ with $\mu(G') \leq \frac{1}{2}\mu(G)$.

If $\{E_i \mid i \in I\}$ is a family of disjoint sets from E with positive measure, I must be countable. In fact, each of the sets

$$I_n := \{i \in I \mid \mu(E_i) > \frac{1}{n}\}$$

is finite, since $\mu(X) < \infty$, and $I = \bigcup_1^\infty I_n$. The set I being countable we have $\bigcup_I E_i \in \mathsf{E}$.

Let us call a family $\mathsf{A} \subseteq \mathsf{E}$ *disjoint* if it consists of disjoint sets, and let us define

$$\mathcal{A} = \{\mathsf{A} \subseteq \mathsf{E} \mid \mathsf{A} \text{ disjoint}, \forall A \in \mathsf{A}(A \subseteq E, \mu(A) > 0), \mu(\bigcup_{A \in \mathsf{A}} A) \leq \frac{1}{2}\mu(E)\} .$$

By inclusion $\mathcal{A}$ is inductively ordered, for if $(\mathsf{A}_j)_{j \in J}$ is a totally ordered subset of $\mathcal{A}$, then

$$\mathsf{A} := \bigcup_{j \in J} \mathsf{A}_j$$

is a majorant for $(\mathsf{A}_j)_{j \in J}$. We must prove that $\mathsf{A} \in \mathcal{A}$. If $E_1, E_2 \in \mathsf{A}$ there exist $j_1, j_2 \in J$ so that $E_k \in \mathsf{A}_{j_k}, k = 1, 2$. By the total ordering we may assume $\mathsf{A}_{j_1} \supseteq \mathsf{A}_{j_2}$ and therefore the sets E_1, E_2 are disjoint, if they are different.

From the observation above A is countable. Write $\mathsf{A} = \{E_1, E_2, \ldots\}$. From the total ordering $\{E_1, \ldots, E_n\} \subseteq \mathsf{A}_{j_n}$ for some $j_n \in J$, and therefore

$$\mu\left(\bigcup_1^n E_i\right) \leq \frac{1}{2}\mu(E) ,$$

whence by the countable additivity

$$\mu\left(\bigcup E_i\right) = \sum \mu(E_i) = \lim_n \sum_1^n \mu(E_i) = \lim_n \mu\left(\bigcup_1^n E_i\right) \leq \frac{1}{2}\mu(E)$$

as required.

With F_ε chosen so that (3) holds for $\varepsilon \leq \frac{1}{2}$, $\{F_\varepsilon\} \in \mathcal{A}$, so Zorn's Lemma applies. Choose $\mathsf{A}_0 = \{E_i \mid i \in I\}$ maximal from $\mathcal{A}$. We must have

$$\mu\left(\bigcup_i E_i\right) = \frac{1}{2}\mu(E) ,$$

for if $\mu(\bigcup_I E_i) \leq (\frac{1}{2} - \varepsilon)\mu(E)$ we could find $F_\varepsilon \subseteq E \setminus \bigcup_I E_i$ so that $\mu(F_\varepsilon) \leq \varepsilon\mu(E)$, and then

$$\mathsf{A}_0 \cup \{F_\varepsilon\} \in \mathcal{A}$$

which would contradict the maximality of A_0. □

LEMMA 9. *Let $\mu = (\mu_1, \ldots, \mu_m) : (X, \mathsf{E}) \to \mathbb{R}^m$ be a non-negative and purely non-atomic vector measure and assume that*

$$\mu_i \ll \mu_{i-1} , \; i \in \{2, \ldots, m\} .$$

Then μ is convex.

PROOF: We proceed by induction in m. In view of Lemma 4 it is enough to find $F \subseteq E$, $F \in \mathsf{E}$ so that

$$\mu(F) = \frac{1}{2}\mu(E)$$

for a given set $E \in \mathsf{E}$.

For $m = 1$ F is given from Lemma 8. For $m > 1$ assume that $\mu' = (\mu_1, \dots, \mu_{m-1})$ is already known to be convex. With $\nu' = \mu_m$ we have $\nu' \ll \mu'$ since $\mu_m \ll \mu_{m-1}$. Since μ' is convex, we can find $E_0 \subseteq E$, $E_0 \in \mathsf{E}$ so that

$$\mu'(E_0) = \frac{1}{2}\mu'(E) .$$

If $\nu'(E_0) = \frac{1}{2}\nu'(E)$ we may take $F = E_0$, otherwise put $F_0 = E \setminus E_0$ and assume

$$\nu'(E_0) < \nu'(E)/2 , \ \nu'(F_0) > \nu'(E)/2 .$$

Theorem 6 (iii) applied to μ', ν', E_0 and F_0 gives us $\lambda \in [0,1]$ so that

$$\nu'(C(\lambda)) = \nu'(E)/2$$

since $[0,1]$ is connected. We have, by Theorem 6 (ii),

$$\begin{aligned}\mu'(C(\lambda)) &= (1-\lambda)\mu'(E_0) + \lambda\mu'(F_0) = \frac{1}{2}(1-\lambda)\mu'(E) + \frac{1}{2}\lambda\mu'(E) \\ &= \frac{1}{2}\mu'(E)\end{aligned}$$

and $F = C(\lambda)$ shows the semi-convexity. □

THEOREM 10. *(Liapounov, 1940). A purely non-atomic vector measure $\mu : (X, \mathsf{E}) \to \mathsf{R}^m$ is convex, and in particular the range $\mu(\mathsf{E})$ is a convex subset of R^m.*

PROOF: By Lemma 4 it is enough to prove that μ is semi-convex.

1° Assume first that μ is non-negative.

The vector measure μ' defined by

$$\mu'_i = \sum_{j=i}^{m} \mu_j , \ i = 1, \dots, m$$

is easily seen to be purely non-atomic, and it satisfies the conditions of Lemma 9 because

$$\mu'_i = \sum_{j=i}^{m} \mu_j \leq \sum_{j=i-1}^{m} \mu_j = \mu'_{i-1} \, .$$

Thus μ' is convex. With

$$T = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 0 & 1 & \cdots & 1 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

we have $\mu' = T\mu$, but T is invertible, and for each $E \in \mathsf{E}$ the (semi)-convexity of μ' gives us $F \subseteq E$, $F \in \mathsf{E}$ so that $\mu'(F) = \frac{1}{2}\mu'(E)$, and hence

$$\mu(F) = T^{-1}\mu'(F) = T^{-1}(\frac{1}{2}\mu'(E)) = \frac{1}{2}T^{-1}\mu'(E) = \frac{1}{2}\mu(E) \, .$$

2° When μ is any purely non-atomic vector measure, we consider the vector measure

$$\mu'' = (\mu_1^+, \mu_1^-, \ldots, \mu_m^+, \mu_m^-) \, .$$

By definition μ'' is non-negative, and since the Hahn decomposition theorem 2.11 gives us partitions $\{P_i, N_i\}$ of X so that

$$\mu_i^+(F) = \mu_i(F \cap P_i) \; , \; \mu_i(F) = -\mu_i(F \cap N_i) \; ,$$

an atom of μ_i^+ or μ_i^- would be an atom of μ_i as well. Thus, μ'' is purely non-atomic and for any given $E \in \mathsf{E}$ 1° gives us $F \subseteq E$, $F \in \mathsf{E}$ so that

$$\mu''(F) = \frac{1}{2}\mu''(E),$$

whence

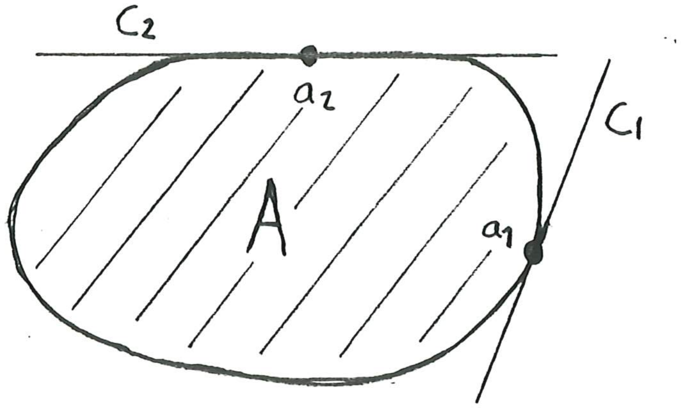
$$\begin{aligned} \mu(F) &= (\mu_1^+(F) - (\mu_1^-(F), \ldots, \mu_m^+(F) - \mu_m^-(F)) \\ &= \frac{1}{2}(\mu_1^+(E) - \mu_1^-(E), \ldots, \mu_m^+(E) - \mu_m^-(E)) \\ &= \frac{1}{2}\mu(E) \, . \end{aligned}$$

□

DEFINITION 11: When dealing with closed convex sets in $\mathbf{R}^m$ one often turns to the notion of supporting hyperplanes. A hyperplane C *supports* the closed convex set A, when $A \cap C \neq \emptyset$ and A is contained in one of the closed half-spaces determined by C.

A hyperplane C has the form $\{\varphi = \alpha\}$ for some (continuous) linear functional $\varphi : \mathsf{R}^m \to \mathsf{R}$. We may assume (replacing, if necessary, (α, φ) by $(-\alpha, -\varphi)$) that $\varphi(A) \subseteq [\alpha, \infty[$.

There are many supporting hyperplanes in the following sense:



THEOREM 12. *If A is a closed convex set in R^m and $a \in \partial A$, there is a supporting hyperplane C of A so that $a \in C$.*

PROOF: See Arne Brøndsted: An introduction to Convex Polytopes, theorem 4.3. □

LEMMA 13. *If a vector measure $\mu : (X, \mathsf{E}) \to \mathsf{R}^m$ has convex range $R := \mu(\mathsf{E})$ and C is an supporting hyperplane of $\overline{R}$, then $C \cap R \neq \emptyset$.*

PROOF: Let $C = \{\varphi = \alpha\}$, $\overline{R} \subseteq \{\varphi \geq \alpha\}$. Since by definition $C \cap \overline{R} \neq \emptyset$

$$\alpha = \inf\{\varphi(\mu(E)) \mid E \in \mathsf{E}\} .$$

Writing $\varphi(x) = \sum_1^m a_i x_i$, then $\varphi \circ \mu = \sum_1^m a_i \mu_i$ is a real measure. As a consequence of Hahn's decomposition theorem, $\varphi \circ \mu$ attains its minimum (in the terminology of 2.11 the minimum is realized by N), and $\alpha \in \varphi \circ \mu(\mathsf{E}) = \varphi(R)$ as required. □

LEMMA 14. *If $\mu : (X, \mathsf{E}) \to \mathsf{R}^m$ is a vector measure with range R and $\xi \in R$, there exists a vector measure μ' with range $R' = R - \xi$. If μ is convex, then so is μ'.*

PROOF: With $\xi = \mu(A)$ put

$$\mu'(E) = \mu(E \setminus A) - \mu(E \cap A) , \; E \in \mathsf{E} . \tag{4}$$

Then μ' is a vector measure and

$$\mu'(E) = \mu(E \setminus A) - (\mu(A) - \mu(A \setminus E)) = \mu((E \setminus A) \cup (A \setminus E)) - \xi$$

so the range R' of μ' satisfies $R' \subseteq R - \xi$. Furthermore

$$\begin{aligned} \mu(E) - \xi &= \mu(E) - \mu(A) = \mu(E \setminus A) - \mu(A \setminus E) \\ &= \mu'((E \setminus A) \cup (A \setminus E)) \,, \end{aligned}$$

where the last equality follows, when (4) is applied to $(E \setminus A) \cup (A \setminus E)$. This shows the opposite inclusion $R - \xi \subseteq R'$.

If μ is convex, for a given $E \in \mathsf{E}$ we may choose $F_-, F_+ \in \mathsf{E}$ so that

$$\begin{aligned} F_- &\subseteq E \setminus A \,,\ \mu(F_-) = \frac{1}{2}\mu(E \setminus A) \,; \\ F_+ &\subseteq E \cap A \,,\ \mu(F_+) = \frac{1}{2}\mu(E \cap A) \,. \end{aligned}$$

By (4) we have

$$\mu'(F_-) = \mu(F_-) \,,\ \mu'(F_+) = -\mu(F_+)$$

and hence

$$\mu'(F_- \cup F_+) = \mu(F_-) - \mu(F_+) = \frac{1}{2}(\mu(E \setminus A) - \mu(E \cap A)) = \frac{1}{2}\mu'(E)$$

so μ' is semi-convex and hence convex. □

LEMMA 15. *The range of a convex vector measure $\mu : (X, \mathsf{E}) \to \mathbb{R}^m$ is closed.*

PROOF: We proceed by induction after the affine dimension N of $R := \mu(\mathsf{E})$, i.e. the dimension of the smallest affine set containing R.

When $N = 1$ μ has the form

$$\mu(E) = \mu_0(E)\xi \,,\ E \in \mathsf{E}$$

for a suitable $\xi \in \mathbb{R}^m$ and μ_0 a real measure. Since R is convex, $\mu_0(\mathsf{E})$ is an interval of $\mathbb{R}$. Since real measures attain their extrema, the interval is closed.

Assume that ranges of convex vector measures with affine dimension less than N are closed. We will use this to show that for every supporting hyperplane C of $\overline{R}$ we have $\overline{R} \cap C \subseteq R$. From Theorem 12 cited above it follows that $\partial(\overline{R}) \subseteq R$, but for convex sets in $\mathbb{R}^m$ it is known that $\partial(\overline{R}) = \partial R$, and it follows that R is closed.

By Lemma 13 we can choose $\xi \in R \cap C$, and since $R - \xi$ is also the range of a convex vector measure by Lemma 14, we may assume that $0 \in C$ (replace

R, C, μ by $R - \xi, C - \xi, \mu'$). Take a continuous linear functional φ so that $C = \{\varphi = 0\}$, $\overline{R} \subseteq \{\varphi \geq 0\}$. Then $\nu = \varphi \circ \mu$ is a positive measure, and from the Lebesgue decomposition, for each coordinate measure μ_i there is a partition $\{A_i, S_i\}$ of X so that, letting

$$\mu_i^a(E) = \mu_i(A_i \cap E), \; \mu_i^s(E) = \mu_i(S_i \cap E) \text{ for } E \in \mathsf{E},$$

we have

$$\mu_i^a \ll \nu, \; \mu_i^s \perp \nu, \nu(S_i) = 0,$$

cf. (the proof of) 2.16. If we write

$$S = \bigcup_{i=1}^{m} S_i, \; A = X \setminus S$$

and

$$\mu_A(E) = \mu(E \cap A), \; \mu_S(E) = \mu(E \cap S)$$

we have $\mu = \mu_A + \mu_S$ and (since $A \subseteq A_i$ for $i = 1, \ldots, m$)

$$\mu_A \ll \nu, \; \nu(S) = 0.$$

With $R_S := \mu_S(\mathsf{E})$ we have $R_S \subseteq C$ since

$$\varphi(\mu_S(E)) = \nu(E \cap S) = 0.$$

The vector measure μ_S is obviously semi-convex, hence convex, and since the dimension of R_S is at most $N - 1$, R_S is closed.

Now take $\xi \in \overline{R} \cap C$. There is a sequence (E_n) from E, so that $\mu(E_n) \to \xi$, whence $\nu(E_n) \to \varphi(\xi) = 0$. Since $\mu_A \ll \nu$ we get from Theorem 2.2 that $\mu_A(E_n) \to 0$, and hence

$$\xi = \lim_{n\to\infty} \mu(E_n) = \lim_{n\to\infty} (\mu_A(E_n) + \mu_S(E_n)) = \lim_{n\to\infty} \mu_S(E_n) \in \overline{R_S} = R_S \subseteq R.$$

THEOREM 16. *(Liapounov, 1940). The range of a vector measure*

$$\mu : (X, \mathsf{E}) \to \mathbb{R}^m$$

is closed.

PROOF:
1° First assume that μ is non-negative.

Let $A \in \mathbb{R}_m^m$ be a regular matrix with positive elements (e.g. $a_{ij} = 2$ for $i = j$, $a_{ij} = 1$ elsewhere); replacing μ by $A \circ \mu$ gives us a vector measure, where every coordinate measure is absolutely continuous with respect to every other (the common null-sets are the sets for which $\mu_i(E) = 0$, $\forall i$). If the range of $A \circ \mu$ is closed, so is the range of μ, since A is a homeomorphism. We can thus assume that μ has this property of mutual absolute continuity of the components. In particular, every atom of a coordinate of μ is an atom of every other.

From a Zorn's Lemma argument, we can pick a maximal family $(E_i)_{i \in I}$ of disjoint common atoms. Since $\mu_j(E_i) > 0$ for every i, j, I is countable, and putting

$$A = \bigcup_{i \in I} A_i\,,\; N = X \setminus A$$

gives us a decomposition of μ

$$\mu_A(E) = \mu(A \cap E)\,,\; \mu_N(E) = \mu(N \cap E)$$

with μ_A purely atomic, μ_N purely non-atomic.

The range $R_N = \mu_N(\mathsf{E})$ is closed by Theorem 10 and Lemma 15. For $E \in \mathsf{E}$ we have

$$\mu_A(E) = \mu\left(\bigcup_I (A_i \cap E)\right) = \sum_I \mu(A_i \cap E) = \sum_I b_i(E)\mu(A_i)$$

with $b_i(E) \in \{0, 1\}$. Thus,

$$R_A = \{\sum_I b_i \mu(A_i) \mid (b_i) \in \{0,1\}^I\}\,.$$

Equipped with the discrete metric, $\{0, 1\}$ is compact. Hence $\{0, 1\}^I$ with the product topology is also compact (Tychonoff's theorem, see 2MA exercise I.6.10). The space $\{0, 1\}^I$ is metrizable under the metric

$$d(b, c) = \sum_{j=1}^{\infty} 2^{-j} |c_{i_j} - b_{i_j}|\,,$$

where $\{i_1, i_2, \ldots\}$ is some ordering of I. The mapping $\varphi : \{0, 1\}^I \to \mathbb{R}^m$ given by

$$\varphi(b) = \sum_I b_i \mu(A_i)$$

can be seen to be continuous with respect to this metric. (Given $b \in \{0, 1\}^I$ and $\varepsilon > 0$ take j_0 so that

$$\sum_{j=j_0+1}^{\infty} |\mu(A_{i_j})|_1 < \varepsilon .$$

When $d(b,c) < 2^{-j_0}$ then the first j_0 coordinates of b and c agree, and hence $|\varphi(b) - \varphi(c)|_1 < \varepsilon$.)

This shows that $R_A = \varphi(\{0,1\}^I)$ is compact, and since

$$R = R_A + R_N,$$

R is compact, hence closed.

2° In the general case we consider the vector measure

$$\nu = (\mu_1^+, \mu_1^-, \cdots, \mu_m^+, \mu_m^-),$$

which is non-negative with values in $\mathbb{R}^{2m}$, so its range is compact by 1°. If we define $f : \mathbb{R}^{2m} \to \mathbb{R}^m$ by

$$f(x_1, y_1, \cdots, x_m, y_m) = (x_1 - y_1, \cdots, x_m - y_m),$$

then $\mu(\mathsf{E}) = f(\nu(\mathsf{E}))$ is also compact.

HISTORICAL COMMENTS. Liapounov's theorem can be stated that the range of a finite dimensional vector measure is compact. It is convex if the vector measure is purely non-atomic. The theorem was proved in the paper "Sur les fonctions-vecteurs complètements additives", Bull. Acad. Sci. URSS vol.4 (1940), 465-478. In a subsequent paper with the same title, in the same journal, vol.10 (1946), 277-279, Liapounov gave a very elegant example to show that neither convexity nor closure can always be asserted for infinite dimensional vector spaces.

§4. An inequality

We shall now generalize Lemma 1.3 to an arbitrary euclidean space E, i.e. a real finite dimensional vector space with inner product. The complex plane $\mathbb{C}$ is of course a two-dimensional euclidean space.

Let E be of dimension d. The Lebesgue measure of E is denoted m_E. The Lebesgue measure of the unit ball in E is V_d and we know that $V_{d+2} = \frac{2\pi}{d+2} V_d$ (cf. Mat 2 MA) and hence

$$V_{2d} = \frac{(2\pi)^{d-1}}{2d(2d-2)\cdots 4} V_2 = \frac{\pi^d}{d!}$$
$$V_{2d+1} = \frac{(2\pi)^d}{(2d+1)(2d-1)\cdots 3} V_1 = \frac{2^{2d+1} d!}{(2d+1)!} \pi^d$$

or

$$V_d = \frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)} \quad \text{with } \Gamma(z) = \int_0^\infty x^{z-1} e^{-x}\, dx \ .$$

The surface measure ω of the unit sphere S satisfies

$$m_E(B_t(\Omega)) = \frac{1}{d} t^d \omega(\Omega)$$

for $\Omega \in \mathbb{B}(S)$, where $B_t(\Omega)$ for $t > 0$ is the sector

$$B_t(\Omega) = \{ s\xi \mid 0 < s < t \,,\, \xi \in \Omega \} \ .$$

In particular the total mass of ω is

$$\omega(S) = dV_d = \frac{2\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})} \ . \tag{1}$$

(If we want to emphasize the dimension we write ω_d for ω, and S^{d-1} for S because it is of dimension $d-1$.)

We recall that ω is invariant under orthogonal transformations. Therefore, for $a \in S$ the integral

$$k_d := \frac{1}{\omega(S)} \int_{a\cdot\xi\geq 0} a \cdot \xi \, d\omega(\xi) \tag{2}$$

over the half-sphere

$$\{ \xi \in S \mid a \cdot \xi \geq 0 \}$$

with north-pole a is independent of a. To determine k_d we choose an orthonormal basis $e_1, \dots, e_d$ for E and use $a = e_d$. The coordinates of a vector $v \in E$ with respect to $e_1, \dots, e_d$ are denoted $(x_1, \dots, x_d)$ and the mapping $v \mapsto (x_1, \dots, x_d)$ is an isomorphism of E onto $\mathbb{R}^d$. We consider first the integral of $a \cdot v$ over the half-ball

$$B_a = \{v \in E \mid \|v\| \leq 1\,,\; a \cdot v \geq 0\}\,,$$

and using polar coordinates $v = r\xi$ we find

$$\int_{B_a} a \cdot v \, dm_E(v) = \int_{a\cdot\xi\geq 0} \int_0^1 r(a\cdot\xi) r^{d-1} \, dr \, d\omega(\xi) = \frac{k_d \omega(S)}{d+1}\,.$$

On the other hand this integral is also

$$\int_{B_+} x_d \, dx_1 \cdots dx_d\,,$$

where

$$B_+ = \{(x_1, \dots, x_d) \mid \|x\| \leq 1\,,\; x_d \geq 0\}\,,$$

and this integral can be calculated by Fubini. If we put $x' = (x_1, \dots x_{d-1})$ and suppose $\|x'\| \leq 1$, then the section $(B_+)_{x'}$ is given by

$$(B_+)_{x'} = \{x_d \mid (x', x_d) \in B_+\} = [0, (1 - \|x'\|^2)^{\frac{1}{2}}]$$

so

$$\int_{B_+} x_d \, dx_1 \cdots dx_d = \int_{\|x'\|\leq 1} \Big(\int_0^{(1-\|x'\|^2)^{\frac{1}{2}}} x_d \, dx_d \Big) dx'$$

$$= \frac{1}{2} \int_{\|x'\|\leq 1} (1 - \|x'\|^2) \, dx'\,,$$

and this integral over the unit ball in $\mathbb{R}^{d-1}$ can be calculated by polar coordinates in $\mathbb{R}^{d-1}$:

$$= \frac{1}{2} \int_0^1 (1-r^2) r^{d-2} \, dr \int_{S^{d-2}} d\omega_{d-1} = \frac{1}{2}\Big(\frac{1}{d-1} - \frac{1}{d+1}\Big) \omega_{d-1}(S^{d-2})\,.$$

We finally get

$$k_d = \frac{\omega_{d-1}(S^{d-2})}{(d-1)\omega_d(S^{d-1})} = \frac{V_{d-1}}{\omega_d(S^{d-1})} = \frac{V_{d-1}}{d V_d}$$

or using (1)

$$k_d = \frac{\Gamma(\frac{d}{2})}{2\sqrt{\pi}\,\Gamma(\frac{d+1}{2})}\,. \tag{3}$$

We have $k_1 = \frac{1}{2}$, $k_2 = \frac{1}{\pi}$, $k_3 = \frac{1}{4}$, $k_4 = \frac{2}{3\pi}$, ...

THEOREM 4.1. *Let E be a euclidean space of dimension d. From any finite family $v_1, \ldots, v_n$ of vectors from E it is possible to choose a subfamily $(v_i)_{i \in I}$ with $I \subseteq \{1, \ldots, n\}$ so that*

$$\|\sum_{i \in I} v_i\| \geq k_d \sum_{i=1}^{n} \|v_i\| . \tag{4}$$

The constant k_d is best possible.

PROOF: We define a continuous function $\varphi : S \to [0, \infty[$ by

$$\varphi(\xi) = \sum_{i=1}^{n} (\xi \cdot v_i)^+ ,$$

and have

$$\begin{aligned}\max_{\xi \in S} \varphi(\xi) &\geq \frac{1}{\omega(S)} \int \varphi(\xi) d\omega(\xi) \\ &= \sum_{i=1}^{n} \frac{1}{\omega(S)} \int (\xi \cdot v_i)^+ d\omega(\xi) = k_d \sum_{i=1}^{n} \|v_i\| .\end{aligned}$$

The function φ assumes its maximum at a point $\xi_0 \in S$, and defining $I = \{i \in \{1, \ldots, n\} \mid \xi_0 \cdot v_i \geq 0\}$ we have

$$\varphi(\xi_0) = (\sum_{i \in I} v_i) \cdot \xi_0 \leq \|\sum_{i \in I} v_i\| ,$$

which shows (4). □

In order to show that k_d is the best possible constant in (4), we need the following:

LEMMA 4.2. *Assume that (4) holds with some constant $k > 0$ instead of k_d. For any E-valued measure $\mu : (X, \mathbb{E}) \to E$ we then have*

$$\sup\{\|\mu(A)\| \mid A \in \mathbb{E}\} \geq k \|\mu\| . \tag{5}$$

PROOF: Let $\varepsilon > 0$ be given. In analogy with Corollary 1.6 we have

$$\|\mu\| = \sup\{\sum_{i=1}^{n} \|\mu(E_i)\| \mid \{E_1, \cdots, E_n\} \text{ partition of } X\} ,$$

so there exists a partition $\{E_1, \ldots, E_n\}$ of X such that

$$\|\mu\| - \varepsilon \leq \sum_{i=1}^{n} \|\mu(E_i)\| \,.$$

Applying Theorem 4.1 to the family $v_i = \mu(E_i)$, $i = 1, \ldots, n$ (with k instead of k_d), we get $I \subseteq \{1, \ldots, n\}$ so that

$$\|\sum_{i \in I} \mu(E_i)\| \geq k \sum_{i=1}^{n} \|\mu(E_i)\| \,,$$

whence with $A = \cup_{i \in I} E_i$

$$\|\mu(A)\| \geq k(\|\mu\| - \varepsilon) \,,$$

and (5) follows. □

COROLLARY 4.3. *For any E-valued measure μ there exists $A \in \mathbb{E}$ so that*

$$\|\mu(A)\| \geq k_d \|\mu\| \,. \tag{6}$$

PROOF: By Liapounov's theorem $\mu(\mathbb{E})$ is compact so there exists $A \in \mathbb{E}$ such that

$$\|\mu(A)\| = \sup\{\|\mu(B)\| \mid B \in \mathbb{E}\} \,,$$

and the result follows from Lemma 4.2 with $k = k_d$. □

END OF PROOF OF THEOREM 4.1: Assume that (4) holds with some constant $k > 0$ instead of k_d. We shall prove that $k \leq k_d$. Let us consider the E-valued vector measure μ on $(S, \mathbb{B}(S))$ defined by

$$\mu(A) = \int_A \xi \, d\omega(\xi) \,,$$

i.e. the i'th component of μ with respect to an orthonormal basis $e_1, \ldots, e_d$ is

$$\mu_i(A) = \int_A \xi_i \, d\omega(\xi) \,,$$

where $\xi \in S$ has the coordinates $(\xi_1, \ldots, \xi_d)$.

By Lemma 4.2 and Liapounov's theorem there exists a Borel set $A_0 \subseteq S$ such that

$$\|\mu(A_0)\| = \sup\{\|\mu(A)\| \mid A \in \mathbb{B}(S)\} \geq k\|\mu\| \,. \tag{7}$$

The vector measure μ has the density $\xi \mapsto \xi$ with respect to ω, and like in Example 1.8 (b) the total variation $\|\mu\|$ has the density $\xi \mapsto \|\xi\| = 1$ with respect to ω, i.e. $\|\mu\| = \omega$, so the right-hand side of (7) is $k\omega(S)$.

To calculate the left-hand side of (7) we consider the unit vector

$$\xi_0 = \frac{\mu(A_0)}{\|\mu(A_0)\|} \in S$$

and the half-sphere $H = \{\xi \in S \mid \xi \cdot \xi_0 \geq 0\}$ and claim that $A_0 = H$ ω-a.e., i.e.

$$\omega(A_0 \setminus H) = \omega(H \setminus A_0) = 0 \,.$$

To see the first equation note that

$$\|\mu(A_0)\| = \mu(A_0) \cdot \xi_0 = \mu(A_0 \cap H) \cdot \xi_0 + \mu(A_0 \setminus H) \cdot \xi_0 \,,$$

so by the Cauchy-Schwarz inequality

$$\|\mu(A_0 \cap H)\| \geq \mu(A_0 \cap H) \cdot \xi_0 = \|\mu(A_0)\| - \mu(A_0 \setminus H) \cdot \xi_0 \,,$$

and by the maximality of A_0, cf. (7), we have $\mu(A_0 \setminus H) \cdot \xi_0 \geq 0$, but

$$\mu(A_0 \setminus H) \cdot \xi_0 = \int_{A_0 \setminus H} \xi \cdot \xi_0 \, d\omega(\xi) < 0$$

unless $\omega(A_0 \setminus H) = 0$, because we have $\xi \cdot \xi_0 < 0$ on $A_0 \setminus H$.

To see the second equation we write

$$\mu(H) = \mu(H \setminus A_0) + \mu(H \cap A_0) \,,$$

and denoting $\overset{\circ}{H} = \{\xi \in S \mid \xi \cdot \xi_0 > 0\}$, we find

$$\mu(H) = \mu(\overset{\circ}{H} \setminus A_0) + \mu(A_0)$$

since $\omega(H \setminus \overset{\circ}{H}) = \omega(A_0 \setminus H) = 0$. Using the Cauchy-Schwarz inequality we get

$$\|\mu(H)\| \geq \mu(H) \cdot \xi_0 = \mu(\overset{\circ}{H} \setminus A_0) \cdot \xi_0 + \|\mu(A_0)\| \,,$$

so by maximality of A_0 we have $\mu(\overset{\circ}{H} \setminus A_0) \cdot \xi_0 \leq 0$, hence $\omega(H \setminus A_0) = \omega(\overset{\circ}{H} \setminus A_0) = 0$.

Knowing that $A_0 = H$ ω-a.e. it is now easy to calculate $\|\mu(A_0)\|$ because

$$\|\mu(A_0)\| = \mu(A_0) \cdot \xi_0 = \int_H \xi \cdot \xi_0 \, d\omega(\xi) = \omega(S) k_d \,,$$

and (7) now gives $k_d \geq k$. □

REMARK. The present proof is a modification of the proof in Kaufman and Rickert: *An inequality concerning measures.* Bull. Amer. Math. Soc.72 (1966), 672–676.

§5. Functions of bounded variation

Borel measures of finite total mass on the real line are in one-to-one correspondence with distribution functions. More precisely we recall the following result from 2MA §II.5.3:

THEOREM 5.1. *Let μ be a finite Borel measure on $\mathbb{R}$. Then $\varphi : \mathbb{R} \to [0, \infty[$ defined by*

$$\varphi(x) = \mu(]-\infty, x]) \tag{1}$$

is

(i) *increasing,*

(ii) *right-continuous,*

(iii) $\varphi(-\infty) = 0$, $\varphi(\infty) < \infty$.

Conversely, if $\varphi : \mathbb{R} \to [0, \infty[$ satisfies (i)-(iii), then there is a unique finite Borel measure μ on $\mathbb{R}$ such that $\varphi(x) = \mu(]-\infty, x])$ for $x \in \mathbb{R}$.

It is a natural problem to determine the class of functions (1) corresponding to arbitrary complex measures on $(\mathbb{R}, \mathbb{B})$.

Let $I = [a, b]$ be a compact interval. For a function $f : I \to \mathbb{C}$ we introduce the *variation* of f over I as

$$v_I(f) = \sup \sum_{i=1}^{N} |f(t_i) - f(t_{i-1})| \,, \tag{2}$$

where the supremum is taken over all N and all choices of (t_i) such that

$$a = t_0 < t_1 < \cdots < t_N = b \,.$$

The functions in

$$BV(I, \mathbb{C}) = \{f : I \to \mathbb{C} \mid v_I(f) < \infty\} \tag{3}$$

are called of *bounded variation* over I.

We clearly have

$$v_I(f) = 0 \Leftrightarrow f \text{ is constant}$$

$$v_I(\lambda f) = |\lambda| v_I(f) \,, \; v_I(f+g) \le v_I(f) + v_I(g) \,,$$

which shows that $BV(I, \mathbb{C})$ is a complex vector space and v_I is a semi-norm, (but not a norm).

It is also easy to see that $f : I \to \mathbb{C}$ is of bounded variation if and only if $\operatorname{Re} f$, $\operatorname{Im} f$ are of bounded variation, so we can restrict the attention to the real subspace $BV(I, \mathbb{R})$ of real functions of bounded variation.

THEOREM 5.2. 1°. *Every monotonic function $f : I \to \mathbb{R}$ is of bounded variation and $v_I(f) = |f(b) - f(a)|$.*

2°. *Every function $f \in BV(I, \mathbb{R})$ can be written as difference of two increasing functions.*

3°. *Every C^1-function $f : I \to \mathbb{C}$ is of bounded variation.*

PROOF: 1°. If f is increasing and $a = t_0 < t_1 < \cdots < t_N = b$ we have

$$\sum_{i=1}^{N} |f(t_i) - f(t_{i-1})| = f(b) - f(a) ,$$

and the assertion $v_I(f) = |f(b) - f(a)|$ follows. If f is decreasing, then $-f$ is increasing so the assertion also holds in this case.

2°. Let $a \le x < y \le b$ and $a = t_0 < t_1 < \cdots < t_N = x$. Then we have

$$v_{[a,y]}(f) \ge |f(y) - f(x)| + \sum_{i=1}^{N} |f(t_i) - f(t_{i-1})| ,$$

and hence

$$v_{[a,y]}(f) \ge |f(y) - f(x)| + v_{[a,x]}(f) . \tag{4}$$

Defining $F(x) = v_{[a,x]}(f)$ we have in particular

$$\begin{aligned} &F(x) \le F(y) \\ &F(x) \pm (f(y) - f(x)) \le F(y) , \end{aligned}$$

which shows that F, $F + f$, $F - f$ are increasing functions.

Finally

$$f = \frac{1}{2}(F + f - (F - f))$$

is the difference of two increasing functions.

3°. It is enough to prove the assertion for real functions, and in this case the mean-value theorem gives

$$v_I(f) \le (b - a) \max_{\xi \in I} |f'(\xi)| < \infty .$$

□

COROLLARY 5.3. *For a function $f \in BV(I, \mathbb{C})$ the left-sided limit*

$$f(x-) := \lim_{y \to x^-} f(y) \quad \textit{exists for} \quad x \in]a, b] ,$$

and the right-sided limit

$$f(x+) := \lim_{y \to x+} f(y) \ \textit{ exists for } \ x \in [a, b[\ .$$

The set of points of discontinuity of f is at most countable.

PROOF: The result is well-known for increasing functions, and the general result follows from Theorem 5.2 2°. □

For a function $f : J \to \mathbb{C}$ defined on an arbitrary interval J we say that it is *locally of bounded variation* if $v_I(f) < \infty$ for every compact interval $I \subseteq J$, and we say that it is of *bounded variation over* J if

$$v_J(f) := \sup_I v_I(f) < \infty \ ,$$

where the supremum is taken over all compact intervals $I \subseteq J$. We clearly have

$$v_J(f) = \sup \sum_{i=1}^{N} |f(t_i) - f(t_{i-1})| \ ,$$

where the supremum is taken over all N and all choices of (t_i) from J so that $t_0 < t_1 < \cdots < t_N$.

We shall now in particular consider the case $J = \mathbb{R}$. With $f : \mathbb{R} \to \mathbb{C}$ we associate the *total variation function* $T_f : \mathbb{R} \to [0, \infty]$ defined by

$$T_f(x) = v_{]-\infty, x]}(f) = \sup \sum_{i=1}^{N} |f(t_i) - f(t_{i-1})| \ , \tag{5}$$

where the supremum is taken over all N and all choices $-\infty < t_0 < t_1 < \cdots < t_N = x$.

The function T_f is clearly increasing and $v_{\mathbb{R}}(f) = \lim_{x \to \infty} T_f(x)$. The set

$$BV = BV(\mathbb{R}, \mathbb{C}) = \{ f : \mathbb{R} \to \mathbb{C} \mid v_{\mathbb{R}}(f) < \infty \}$$

is a subspace of the vector space of bounded functions and $v_{\mathbb{R}}$ is a semi-norm.

We call a function $f \in BV(\mathbb{R}, \mathbb{C})$ *normalized* if

(i) $\lim_{x \to -\infty} f(x) = 0$,

(ii) f is right-continuous,

and the set of these functions is a subspace of BV denoted $BV_n = BV_n(\mathbb{R}, \mathbb{C})$. □

THEOREM 5.4. 1°. *If* $f \in BV(\mathbb{R}, \mathbb{C})$ *and* $x < y$ *then*

$$|f(y) - f(x)| \leq T_f(y) - T_f(x) .$$

2°. *For a monotonic function* $f : \mathbb{R} \to \mathbb{R}$ *we have*

$$v_{\mathbb{R}}(f) = |\lim_{x\to\infty} f(x) - \lim_{x\to-\infty} f(x)| ,$$

and f *is of bounded variation if and only if it is bounded.*

3°. *Every function* $f \in BV(\mathbb{R}, \mathbb{R})$ *can be written as difference of two increasing bounded functions.*

4°. *If* $f \in BV(\mathbb{R}, \mathbb{C})$ *then* $f(x-)$ *exists for every* $x \in]-\infty, \infty]$ *and* $f(x+)$ *exists for every* $x \in [-\infty, \infty[$. *The set of points of discontinuity of* f *is at most countable, and there exists a unique pair* (c, g) *with* $c \in \mathbb{C}$ *and* $g \in BV_n$ *such that*

$$f(x) = c + g(x) \tag{6}$$

at all points of continuity of f. *Also*

$$v_{\mathbb{R}}(g) \leq v_{\mathbb{R}}(f) .$$

5°. *If* $f \in BV$ *then* $T_f \in BV$ *and* $T_f(-\infty) = 0$.

6°. *If* $f \in BV_n$ *then* $T_f \in BV_n$.

PROOF: 1°. For $x < y$ and $\varepsilon > 0$ there are points $t_0 < t_1 < \cdots < t_N = x$ so that

$$\sum_{i=1}^{N} |f(t_i) - f(t_{i-1})| > T_f(x) - \varepsilon ,$$

and hence

$$T_f(y) \geq |f(y) - f(x)| + \sum_{i=1}^{N} |f(t_i) - f(t_{i-1})| > |f(y) - f(x)| + T_f(x) - \varepsilon ,$$

proving 1°. The assertions 2° and 3° and first half of 4° are proved like in Theorem 5.2 and Corollary 5.3.

Concerning 4° assume first that we have a representation (6). Choosing a decreasing sequence (x_n) of points of continuity of f tending to $-\infty$ (respectively to x) we get

$$f(-\infty) = \lim_{n\to\infty} f(x_n) = c + \lim_{n\to\infty} g(x_n) = c ,$$

respectively

$$f(x+) = c + g(x+) = f(-\infty) + g(x) ,$$

showing that the pair (c, g) is uniquely determined. For the existence of (6) we define

$$c := f(-\infty)\ ,\ g(x) := f(x+) - c\ ,$$

and g is clearly right-continuous with limit 0 for $x \to -\infty$. Furthermore $f(x) = g(x) + c$ for all points x where $f(x) = f(x+)$, in particular for all points of continuity. To see that g has bounded variation we consider $-\infty < t_0 < \cdots < t_N$ and find

$$\sum_{i=1}^{N} |g(t_i) - g(t_{i-1})| = \lim_{\delta \to 0+} \sum_{i=1}^{N} |f(t_i + \delta) - f(t_{i-1} + \delta)| \leq v_{\mathbf{R}}(f)\ ,$$

hence $v_{\mathbf{R}}(g) \leq v_{\mathbf{R}}(f) < \infty$.

5°. If $f \in BV$ and $\varepsilon > 0$ we choose

$$t_0 < t_1 < \cdots < t_N = 0$$

so that

$$T_f(0) - \varepsilon < \sum_{i=1}^{N} |f(t_i) - f(t_{i-1})|\ . \tag{7}$$

For points $y_0 < y_1 < \cdots < y_M = t_0$ we then have

$$T_f(0) \geq \sum_{i=1}^{M} |f(y_i) - f(y_{i-1})| + \sum_{i=1}^{N} |f(t_i) - f(t_{i-1})|\ ,$$

and combined with (7) we get

$$\sum_{i=1}^{M} |f(y_i) - f(y_{i-1})| < \varepsilon\ .$$

This shows that $T_f(t_0) \leq \varepsilon$ whence $T_f(-\infty) = 0$.

Since T_f is increasing and bounded it belongs to BV by 2°.

6°. For $f \in BV$ and $x < y$ we clearly have $T_f(x) + v_{[x,y]}(f) = T_f(y)$. This shows, that if $f \in BV_n$, then T_f is right-continuous at x if and only if

$$\lim_{y \to x+} v_{[x,y]}(f) = 0\ .$$

Let $\varepsilon > 0$ be given. Since $v_{[x,x+1]}(f) < \infty$ there exists $x = t_0 < t_1 < \cdots < t_N = x + 1$ so that

$$\sum_{j=1}^{N} |f(t_j) - f(t_{j-1})| > v_{[x,x+1]}(f) - \varepsilon\ ,$$

and for $x < y < t_1$ we then have

$$|f(y) - f(x)| + |f(t_1) - f(y)| + \sum_{j=2}^{N} |f(t_j) - f(t_{j-1})| > v_{[x,x+1]}(f) - \varepsilon \,. \quad (8)$$

Since f is right-continuous at x there exists $0 < \delta < t_1 - x$ so that

$$|f(y) - f(x)| < \varepsilon \text{ for } y \in]x, x+\delta[\,,$$

and (8) implies

$$v_{[y,x+1]}(f) > v_{[x,x+1]}(f) - 2\varepsilon \text{ for } y \in]x, x+\delta[\,,$$

hence

$$v_{[x,y]}(f) < 2\varepsilon \text{ for } y \in]x, x+\delta[\,.$$

□

We can now prove the main theorem.

THEOREM 5.5. *(a) If μ is a complex measure on $(\mathbb{R}, \mathbb{B})$ and if*

$$f(x) = \mu(]-\infty, x]) \,, \quad x \in \mathbb{R} \quad (9)$$

then $f \in BV_n$.

(b) Conversely, to every $f \in BV_n(\mathbb{R}, \mathbb{C})$ there corresponds a unique complex measure μ on $(\mathbb{R}, \mathbb{B})$ such that (9) holds. For this μ

$$T_f(x) = |\mu|(]-\infty, x]) \,, \quad x \in \mathbb{R} \quad (10)$$

and $v_{\mathbb{R}}(f) = \|\mu\|$.

PROOF: (a) For $t_0 < \cdots < t_N = x$ we have

$$\sum_{i=1}^{N} |f(t_i) - f(t_{i-1})| = \sum_{i=1}^{N} |\mu(]t_{i-1}, t_i])| \le |\mu|(]t_0, t_N])$$
$$\le |\mu|(]-\infty, x]) \,,$$

and hence

$$T_f(x) \le |\mu|(]-\infty, x]) \,, \quad x \in \mathbb{R} \quad (11)$$

which shows that f is of bounded variation. Splitting μ in real and imaginary part and using the Jordan decomposition (see p.1.11) we see that $f \in BV_n$.

(b) There is at most one complex measure μ satisfying (9) for given $f \in BV_n$. In fact, if two complex measures both satisfy (9) then they agree on all standard intervals $]\alpha, \beta]$, and hence on all open sets. For this we use that any open set in $\mathbb{R}$ is countable union of disjoint standard intervals, cf. 2MA. To see that the two complex measures finally agree on all Borel sets it is sufficient to remark that for any two complex measure μ and ν and any Borel set B, there exists a sequence of open sets (U_n), $U_n \supseteq B$ so that

$$\mu(B) = \lim_{n\to\infty} \mu(U_n)\,,\; \nu(B) = \lim_{n\to\infty} \nu(U_n)\,. \tag{12}$$

To see this we write $\mu = \mu_1 - \mu_2 + i(\mu_3 - \mu_4)$ and $\nu = \mu_5 - \mu_6 + i(\mu_7 - \mu_8)$ with $\mu_i \geq 0$, and using the outer regularity of each μ_i we see that given $B \in \mathbb{B}$ and $\varepsilon > 0$ there exists an open set G_i with $B \subseteq G_i$ and $\mu_i(G_i) \leq \mu_i(B) + \varepsilon$ for $i = 1, \cdots, 8$. The intersection $G(\varepsilon)$ of these eight open sets is open, $B \subseteq G(\varepsilon)$ and

$$\mu_i(B) \leq \mu_i(G(\varepsilon)) \leq \mu_i(G_i) \leq \mu_i(B) + \varepsilon\;,\; i = 1, \cdots, 8\,.$$

The sequence of sets $U_n = G(\frac{1}{n})$ has the property

$$\lim_{n\to\infty} \mu_i(U_n) = \mu_i(B)\;,\; i = 1, \cdots, 8\,,$$

and (12) follows.

Let $f \in BV_n$ be given. Since T_f has the properties in Theorem 5.1 (by Theorem 5.4), there exists a positive finite measure σ on $(\mathbb{R}, \mathbb{B})$ so that

$$T_f(x) = \sigma(]-\infty, x]) \;\text{ for }\; x \in \mathbb{R}\,.$$

If f is real-valued we know that $f = \frac{1}{2}(T_f + f - (T_f - f))$ is the difference of two increasing functions in BV_n. By Theorem 5.1 we can associate two positive measures μ_1, μ_2 to the functions $\frac{1}{2}(T_f \pm f)$ so that (1) holds. Then $\mu = \mu_1 - \mu_2$ is a real measure such that (9) holds. The complex case is easily reduced to the real case.

If μ is the complex measure so that (9) holds, then

$$|\mu(]\alpha, \beta])| \leq \sigma(]\alpha, \beta]) \;\text{ for }\; \alpha < \beta\,.$$

This follows immediately from the inequality in Theorem 5.4 1°. Any open set G is disjoint union of a sequence (I_n) of standard intervals, and hence

$$|\mu(G)| = |\sum \mu(I_n)| \leq \sum |\mu(I_n)| \leq \sum \sigma(I_n) = \sigma(G)\,.$$

Finally by (12) we get $|\mu(B)| \leq \sigma(B)$ for any Borel set B, but $|\mu|$ is the smallest positive measure with this property, hence $|\mu| \leq \sigma$ and in particular

$$|\mu|(]-\infty, x]) \leq \sigma(]-\infty, x]) = T_f(x)\,,$$

which together with (11) establishes (10). □

PROPOSITION 5.6. *Let μ be a complex measure on $(\mathbb{R}, \mathbb{B})$ and*

$$f(x) = \mu(]-\infty, x]) \ , \ x \in \mathbb{R} \ .$$

Then f is continuous if and only if $\mu(\{x\}) = 0$ for all $x \in \mathbb{R}$, and this is equivalent with μ being purely non-atomic.

PROOF: We have

$$\mu(\{x\}) = \lim_{n \to \infty} \mu(]x - \frac{1}{n}, x]) = f(x) - f(x-) \ ,$$

so $\mu(\{x\}) = 0$ if and only if f is left-continuous at x, but this is equivalent to the continuity of f at x.

If $\mu(\{x\}) \neq 0$ then $\{x\}$ is clearly an atom. Conversely if $B \in \mathbb{B}$ is an atom for μ then it is easy to see that there exists $x \in B$ so that $\mu(\{x\}) = \mu(B)$, cf. exercise 16. □

DEFINITION 5.7. A function $f : \mathbb{R} \to \mathbb{C}$ is called *absolutely continuous* if to every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\sum_{i=1}^{N} (\beta_i - \alpha_i) < \delta \Rightarrow \sum_{i=1}^{N} |f(\beta_i) - f(\alpha_i)| < \varepsilon \tag{13}$$

whenever $]\alpha_1, \beta_1[, \cdots,]\alpha_N, \beta_N[$ are disjoint intervals.

Every absolutely continuous function is uniformly continuous ($N = 1$) and $v_I(f) < \infty$ for any finite interval. In fact, if $\delta > 0$ corresponds to $\varepsilon = 1$ and I is of finite length we can divide $I = I_1 \cup \cdots \cup I_p$ in finitely many intervals of length $< \delta$, so $v_{I_j}(f) \leq 1$ by (13) and hence $v_I(f) \leq p$.

If f is a Lipschitz function then it is automatically absolutely continuous. In particular $f(x) = \sin x$ is absolutely continuous, but $f \notin BV$.

PROPOSITION 5.8. *Let μ be a complex measure on $(\mathbb{R}, \mathbb{B})$ and*

$$f(x) = \mu(]-\infty, x]) \ , \ x \in \mathbb{R} \ .$$

Then $\mu \ll m$ if and only if f is absolutely continuous, (where m is Lebesgue measure).

PROOF: Suppose first that f is absolutely continuous and let $E \in \mathbb{B}$ with $m(E) = 0$ and $\varepsilon > 0$ be given. Choose $\delta > 0$ according to (13). Using the same idea as in (12) we can find open sets $U_n \supseteq E$ so that

$$\mu(U_n) \to \mu(E) \ , \ m(U_n) \to m(E) = 0 \ . \tag{14}$$

We may therefore assume that $m(U_n) < \delta$ for all n. Every U_n is disjoint union of a sequence of standard intervals $I_j =]\alpha_j, \beta_j]$ for which $m(U_n) = \sum(\beta_j - \alpha_j) < \delta$. This is in particular valid for all finite partial sums, so by (13) we get

$$\sum |f(\beta_j) - f(\alpha_j)| \leq \varepsilon$$

(in the first place for all the finite partial sums and hence for the series itself), and consequently

$$|\mu(U_n)| \leq \sum |\mu(I_j)| = \sum |f(\beta_j) - f(\alpha_j)| \leq \varepsilon \, .$$

By (14) $|\mu(E)| \leq \varepsilon$, and since $\varepsilon > 0$ was arbitrary $\mu(E) = 0$.

If $\mu \ll m$ then Theorem 2.2 immediately gives (13). □

For $f : \mathbb{R} \to \mathbb{R}$ we introduce the *positive variation* P_f and the *negative variation* N_f of f as

$$P_f(x) = \sup \sum_{i=1}^{N} (f(t_i) - f(t_{i-1}))^+ \tag{15}$$

$$N_f(x) = \sup \sum_{i=1}^{N} (f(t_i) - f(t_{i-1}))^- \, , \tag{16}$$

where the suprema are taken for all $N \in \mathbb{N}$ and all choices $t_0 < \cdots < t_N = x$. These functions are clearly increasing, and they satisfy

$$T_f(x) = P_f(x) + N_f(x) \quad , \; x \in \mathbb{R} \, , \tag{17}$$

which is an easy consequence of $|a| = a^+ + a^-$ for $a \in \mathbb{R}$. (For the inequality $\geq$ one should merge two subdivisions associated with P_f and N_f).

From (17) follows that $f \in BV$ if and only if $P_f(\infty) < \infty$ and $N_f(\infty) < \infty$.

For $f \in BV$ we have the additional equation

$$f(x) - f(-\infty) = P_f(x) - N_f(x) \, , \; x \in \mathbb{R} \, , \tag{18}$$

which follows from $a = a^+ - a^-$ for $a \in \mathbb{R}$.

In fact, given $\varepsilon > 0$ there exists a subdivision $t_0 < \cdots < t_N = x$ such that $|f(t_0) - f(-\infty)| < \varepsilon/2$ and

$$P_f(x) - \varepsilon/2 < \sum_{i=1}^{N} (f(t_i) - f(t_{i-1}))^+ \, ,$$

and we get

$$P_f(x) - N_f(x) < \sum_{i=1}^{N}(f(t_i) - f(t_{i-1}))^+ - \sum_{i=1}^{N}(f(t_i) - f(t_{i-1}))^- + \varepsilon/2$$
$$= f(x) - f(t_0) + \varepsilon/2 < f(x) - f(-\infty) + \varepsilon\ .$$

Similarly, given $\varepsilon > 0$ there exists a subdivision $s_0 < \cdots < s_M = x$ such that $|f(s_0) - f(-\infty)| < \varepsilon/2$ and

$$N_f(x) - \varepsilon/2 < \sum_{i=1}^{M}(f(s_i) - f(s_{i-1}))^-\ ,$$

and we get

$$f(x) + N_f(x) < f(x) - f(s_0) + \sum_{i=1}^{M}(f(s_i) - f(s_{i-1}))^- + \varepsilon/2 + f(s_0)$$
$$= \sum_{i=1}^{M}(f(s_i) - f(s_{i-1}))^+ + \varepsilon/2 + f(s_0) < P_f(x) + \varepsilon + f(-\infty)\ .$$

PROPOSITION 5.9. *Let μ be a real measure on $(\mathbb{R}, \mathbb{B})$ with distribution function*

$$f(x) = \mu(]-\infty, x]) \quad , \ x \in \mathbb{R}\ ,$$

and Jordan decomposition $\mu = \mu^+ - \mu^-$. Then we have

$$P_f(x) = \mu^+(]-\infty, x])\ ,\ N_f(x) = \mu^-(]-\infty, x]) \quad , \ x \in \mathbb{R}\ .$$

PROOF: We know that $\mu^\pm = \frac{1}{2}(|\mu| \pm \mu)$ so the distribution functions of $\mu^\pm$ are given by

$$\mu^\pm(]-\infty, x]) = \frac{1}{2}(T_f(x) \pm f(x))\ ,$$

but these function are $P_f(x)$ and $N_f(x)$ by (17) and (18) because $f(-\infty) = 0$.

□

Exercises

1. Show that if an infinite series $\sum z_n$ of complex numbers is absolutely convergent, then it is unconditionally convergent.

2. Let $(B, \|\cdot\|)$ be a Banach space. For an infinite series $\sum z_n$ with elements $z_n \in B$ the concepts of convergence, absolute convergence and unconditional convergence are straightforward to define.

Prove:

(i) absolute convergence $\Rightarrow$ unconditional convergence.

(ii) Prove the converse implication in case $B = \mathbb{C}^k$ with the maximum norm.

(iii) Generalize (ii) to any finite dimensional Banach space. (Hint: Choose a basis and use that the given norm is equivalent with the maximum norm with respect to the basis).

(iv) Let B be an infinite dimensional Hilbert space and let $e_1, e_2, \cdots$ be an orthonormal sequence, i.e. $e_i \cdot e_j = \delta_{ij}$.

Show that the series $\sum_1^\infty \frac{1}{n} e_n$ is *unconditionally convergent but not absolutely convergent.*

3. A famous theorem due to Dvoretzky and Rogers states: *A Banach spaces B is of finite dimension if and only if every unconditionally convergent series is absolutely convergent.* Find the precise reference to the journal(s), where this was first published.

4. Let $(X, \mathbb{E})$ be a measurable space and $(B, \|\cdot\|)$ a Banach space. A function $\mu : \mathbb{E} \to B$ is called a B-valued measure if

$$\mu(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mu(E_i)$$

for any sequence $(E_i)_{i \geq 1}$ of pairwise disjoint sets from $\mathbb{E}$. We also call μ a *vector measure* (with values in B).

The series in question is unconditionally convergent. We can introduce the *absolute value* $|\mu|$ of such a vector measure:

$$|\mu|(E) = \sup\{\sum_{i=1}^{\infty} \|\mu(E_i)\| \mid (E_i) \text{ partition of } E\} .$$

Show that $|\mu| : \mathsf{E} \to [0, \infty]$ is a positive measure.

5. Let μ be a B-valued vector measure on (X, E), where B is a *finite* dimensional normed space. Show that $|\mu|(X) < \infty$.

6. Let B be an infinite dimensional Hilbert space and let (e_n) be an orthonormal sequence. Let μ be the B-valued vector measure on $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ defined by

$$\mu(\{n\}) = \frac{1}{n} e_n .$$

(This is a vector measure by 2(iv)). Find $|\mu|$ and show that $|\mu|(\mathbb{N}) = \infty$.

7. Let μ be a complex measure on (X, E). Show that

$$\left.\begin{array}{l} |\mathrm{Re}\ \mu| \\ |\mathrm{Im}\ \mu| \end{array}\right\} \leq |\mu| \leq |\mathrm{Re}\ \mu| + |\mathrm{Im}\ \mu| .$$

8. Let (X, E) and (Y, F) be two measurable spaces, and let $\phi : X \to Y$ be measurable. Show that if μ is a complex measure on (X, E) then $\phi(\mu)$ is a complex measure on (Y, F) called the image measure, if we define

$$\phi(\mu)(F) = \mu(\phi^{-1}(F)) \quad \text{for} \quad F \in \mathsf{F}.$$

Show that $|\phi(\mu)| \leq \phi(|\mu|)$, and give an example showing that equality need not occur.

9. Let μ be a real measure on (X, E). Show that for $E \in \mathsf{E}$

$\mu^+(E) = \sup\{\mu(A) \mid A \in \mathsf{E}, A \subseteq E\},$

$\mu^-(E) = -\inf\{\mu(A) \mid A \in \mathsf{E}, A \subseteq E\}.$

10. Let $M_{\mathbf{R}}(X, \mathsf{E})$ denote the real vector space of real measures on (X, E) and let $\mu, \nu \in M_{\mathbf{R}}(X, \mathsf{E})$. For $E \in \mathsf{E}$ we define

$$(\mu \vee \nu)(E) = \sup\{\mu(A) + \nu(B) \mid \{A, B\} \text{ partition of } E\} ,$$
$$(\mu \wedge \nu)(E) = \inf\{\mu(A) + \nu(B) \mid \{A, B\} \text{ partition of } E\} .$$

(i) Show that $\mu \vee \nu$ and $\mu \wedge \nu$ are real measures on (X, E) and that $\mu \vee \nu$ is the smallest real measure which majorizes μ and ν.
(ii) Show similarly that $\mu \wedge \nu$ is the biggest real measure minorizing μ and ν.
(iii) $\mu + \nu = (\mu \vee \nu) + (\mu \wedge \nu)$
(iv) $\mu^+ = \mu \vee 0$, $\mu^- = -(\mu \wedge 0)$.

11. Let $[0,1]$ be equipped with the smallest σ-algebra E on $[0,1]$ which contains the countable sets. Show that $\mathsf{E} \neq \mathsf{B}$. Let $\mu : \mathsf{E} \to [0, \infty]$ be counting measure. Show that

$$\int x f(x) d\mu(x)$$

is well-defined for $f \in \mathcal{L}_1(\mu)$ and that it defines a continuous linear functional Φ on $L_1(\mu)$. Show that Φ does not have the form T_φ for some $\varphi \in \mathcal{L}_\infty(\mu)$.

12. Suppose X consists of two points a and b and define a measure μ on X by $\mu(a) = 1$, $\mu(b) = \infty$. Show that any measure on X is absolutely continuous with respect to μ, but the Radon-Nikodym theorem does not hold. Describe $L_1(\mu)$ and $L_\infty(\mu)$. Is it true that $L_\infty(\mu)$ is the dual of $L_1(\mu)$ in this case?.

13. Let X denote a topological space and let $f : X \to [-\infty, \infty]$. We say that f is *lower semicontinuous* at $a \in X$, if for any $t < f(a)$ there exists a neighbourhood U of a so that $f(y) > t$ for all y in U. Note that f is automatically lower semicontinuous at any point where f is $-\infty$.

Define upper semicontinuity similarly. (It should be so that f is upper semicontinuous at a if and only if $-f$ is lower semicontinuous at a.)

Show that f is continuous at a if and only if it is both upper and lower semicontinuous at a.

Show that f is lower semicontinuous, i.e. at every point of X, if and only if $\{x \in X \mid f(x) > t\}$ is open in X for any $t \in \mathbb{R}$.

Show that if $(f_i)_{i \in I}$ is any family of lower semicontinuous functions, then $f := \sup_{i \in I} f_i$ is again lower semicontinuous.

Show that if $f + g$ is defined at all points of X (i.e. there are no points where $f(x) = \infty$ and $g(x) = -\infty$ or vice versa) and if f, g are lower semicontinuous, then $f + g$ is lower semicontinuous.

Show finally that the indicator function 1_A of a subset A of X is lower (resp. upper) semicontinuous if and only if A is open (resp. closed) in X.

14. Let $f \in \mathcal{L}_{loc}(\mathbb{R})$. Show that

$$|f(x)| \leq M(f)(x)$$

at every Lebesgue point of f.

15. *The density topology.* A set $E \in \mathbb{R}^k$ is called *approximatively open* if for every $x \in E$ there exists a Borel set B_x with $x \in B_x \subseteq E$ and so that B_x has the metric density 1 at x.

Show that the family $\mathcal{A}$ of approximatively open sets in $\mathbb{R}^k$ is a topology (called the density topology).

Show that this topology is strictly finer than the ordinary topology on $\mathbb{R}^k$.

For more information about this topology see J. Ridder, Fund. Math. 13 (1929), 201-209.

16. (i) Let $A \subseteq \mathbb{R}$ be a Borel set with finite Lebesgue measure. Show that the restriction m_A of Lebesgue measure to A has no atoms.

(ii) Let $f \in \mathcal{L}_1(\mathbb{R}, m)$ and define the real measure

$$\mu(A) = \int_A f\, dm\,.$$

Show that μ has no atoms.

(iii) Let μ be a positive finite measure on $(\mathbb{R}, \mathbb{B})$. Show that if μ has an atom A, then there exists $a \in A$ so that

$$\mu(\{a\}) = \mu(A)\,.$$

17. Let (X, E, μ) be a measure space with $\mu(X) < \infty$ and let $L_1(X, \mathsf{E}, \mu)$ be the Banach space of equivalence classes of integrable functions. For $A \in \mathsf{E}$ we denote by $[1_A]$ the equivalence class containing 1_A. Show that

$$\sigma(A) = [1_A]$$

is a $L_1(X, \mathsf{E}, \mu)$-valued vector measure and that the absolute value of σ is μ.

Show that $\sigma(\mathsf{E})$ is closed in $L_1(X, \mathsf{E}, \mu)$.

Show finally that if (X, E, μ) is Lebesgue measure on the Borel subsets of $[0, 1]$, then $\sigma(\mathsf{E})$ is not convex.

From Paul R. Halmos: I want to be a mathematician. An automathography, Springer 1985.

There was a "night shift" and a "weekend shift" at Eckhart Hall; the building was always alive. Nostalgia?—yes, maybe, but with top-quality students in such breath-taking quantity, there is something to be nostalgic about. Personal attitudes aside, the mathematics department at the University of Chicago in the late 40's and 50's was either the best in the world or close enough that I can be forgiven for regarding it so.

The beginning of Hilbert space

In the late 1940's I began to act on one of my beliefs: to stay young, you have to change fields every five years. Looking back on it I can now see a couple of aspects of that glib commandment that weren't always obvious. One: I didn't first discover it and then act on it, but, instead, noting that I did in fact seem to change directions every so often, I made a virtue out of a fact and formulated it as a piece of wisdom. Two: it works. A creative thinker is alive only so long as he grows; you have to keep learning new things to understand the old. You don't really have to change fields—but you must stoke the furnace, branch out, make a strenuous effort to keep from being locked in.

As my own focus on measure theory began to waver, I published a couple of comments on other people's measure theory. One was on Liapounov's theorem (to the effect that the ranges of well-behaved vector-valued measures are closed convex sets). Kai Rander Buch published a paper on closedness, and that paper made me angry: it struck me as wordy and pretentious and unnecessarily complicated. Surely one can do better than that, I said; I thought about the question, saw a way of doing much better, and dashed off a note to the Bulletin of the AMS. My proof was a lot slicker than Kai Rander Buch's and a lot shorter, but his was right, and, to my mortification, mine turned out to be wrong. Both Jessen and Dieudonné wrote and told me that my Lemma 5, the crucial lemma, was false. A pity; it was such a nice lemma. It says that the span of two compact topologies is compact (span, supremum, generated topology)—a statement for which it's not only easy to find counterexamples but it's hard to find any non-trivial instances where it is true. Being caught stumbling in public was all the motivation I needed to sit down and think matters through more deeply and more effectively. My second note came out a year after the first (1948), and it was twice as long (six pages), but it was elegant and correct, and has been quoted quite a bit since then. It is all superseded by now; in 1966 Lindenstrauss came out with the slickest proof to end all proofs (J. of Math. and Mech.).

In 1949 I published another little note precipitated by an emotional reaction. The irritant in that case was a paper by Shin-Ichi Izumi proving