Real Analysis
A.A 2013–14

Giuseppe De Marco

1. Prolegomena

1.1. The extended real line. The extended real line

$$\tilde{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\} = [-\infty, \infty]$$

is constantly used in this text. As an ordered set with $-\infty < x < \infty$ for every $x \in \mathbb{R}$, the extended line is *order complete*, meaning that every non–empty subset of $\tilde{\mathbb{R}}$ has a supremum and an infimum in $\tilde{\mathbb{R}}$. As a topological space $\tilde{\mathbb{R}}$ is a *compactification* of the usual line, a compact space containing $\mathbb{R}$ as a dense subspace, and it is homeomorphic to a bounded closed interval of $\mathbb{R}$; recall that a neighborhood of ∞ (resp $-\infty$) in $\tilde{\mathbb{R}}$ is any subset of $\tilde{\mathbb{R}}$ containing a half–line $]a, \infty]$ (resp. $[-\infty, a[$) for some $a \in \mathbb{R}$).

1.1.1. *Operations on the extended real line.* Addition $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$ has a continuous extension

$$(\#) + (\#) : \tilde{\mathbb{R}} \times \tilde{\mathbb{R}} \smallsetminus \{(\infty, -\infty), (-\infty, \infty)\} \to \tilde{\mathbb{R}}$$

by $\infty + r = \infty + \infty = \infty$ and $-\infty + r = -\infty + (-\infty) = -\infty$ for every $r \in \mathbb{R}$; $\infty - \infty$ and $-\infty + \infty$ are not defined (no possible definition of $\infty - \infty$ in $\tilde{\mathbb{R}}$ can make addition continuous on the point $(\infty, -\infty)$: for every $a \in \mathbb{R}$ the real sequences $n + a$ and $-n$ tend to ∞, $-\infty$ repectively, and $(n + a) + (-n) = a$ tends to a). Similarly, multiplication is extended by continuity to $\tilde{\mathbb{R}} \times \tilde{\mathbb{R}} \smallsetminus \{(0, \pm\infty), (\pm\infty, 0)\}$ in the way dictated by the theorem on limits, and $0(\pm\infty)$ and $(\pm\infty)0$ cannot be defined. We have the easy but useful result:

. *Let A, B be non–empty subset of $\tilde{\mathbb{R}}$; assume that $a + b$ is defined for every $a \in A$ and every $b \in B$, so that $A + B := \{a + b : a \in A, b \in B\}$ is also defined. Then:*

$$\inf(A + B) = \inf A + \inf B; \qquad \sup(A + B) = \sup A + \sup B$$

provided that also the right–hand sides are defined (in particular, this always holds when A and B are non–empty sets of real numbers).

Proof. Let's prove it for suprema. If $\sup A + \sup B = \infty$, then either $\sup A = \infty$ or $\sup B = \infty$, in any case either A or B has no upper bound in $\mathbb{R}$, and this immediately implies that $A + B$ has no upper bound in $\mathbb{R}$. If $\sup A + \sup B = -\infty$, then either $\sup A = -\infty$ and equivalently $A = \{-\infty\}$, or $\sup B = -\infty$, and equivalently $B = \{-\infty\}$, or both, but in any of these cases we get $A + B = \{-\infty\}$ and the conclusion follows. If $\sup A$ and $\sup B$ are both finite we first get $a + b \le \sup A + \sup B$ for every $a \in A$ and every $b \in B$ just summing up the two inequalities $a \le \sup A$ and $b \le \sup B$, so that $\sup(A + B) \le \sup A + \sup B$. And given $\varepsilon > 0$, pick $a \in A$ and $b \in B$ such that $a > \sup A - \varepsilon/2$ and $b > \sup B - \varepsilon/2$; then $a + b > \sup A + \sup B - \varepsilon$. This proves that $\sup(A + B) \ge \sup(A) + \sup B$. □

Remark. One can show that these formulae are in fact equivalent to the continuity of the extended addition map: $] - \infty, \infty]$ and $[-\infty, \infty[$ are *topological semigroups* under addition. Notice that $A + B$ is defined if and only if A and B are both subsets of one of these two semigroups.

There is an entirely analogous multiplicative statement for subsets of $[0, \infty]$, with 0 playing the role of $-\infty$; we leave it to the reader.

1.1.2. *Additivity of* inf *and* sup. By induction and 1.1.1 it is clear that we also have

. *If $A_1, \dots A_m$ are non empty subsets of $\tilde{\mathbb{R}}$ then we have*

$$\inf(A_1 + \cdots + A_m) = \inf(A_1) + \cdots + \inf(A_m); \quad \sup(A_1 + \cdots + A_m) = \sup(A_1) + \cdots + \sup(A_m),$$

provided that both sides of the equalities are meaningful.

1.2. Lattice operations on real valued functions. Two binary operations, $\vee$ and $\wedge$, are defined on all the extended real line:

$$x \vee y = \max\{x,\, y\}; \quad x \wedge y = \min\{x,\, y\}.$$

It is easy to see that $\vee$ and $\wedge$ are continuous as mappings of $\tilde{\mathbb{R}} \times \tilde{\mathbb{R}}$ to $\tilde{\mathbb{R}}$; as binary operations they are commutative and associative (with $\mp\infty$ as neutral elements for $\vee$ and $\wedge$, respectively).

If X is a set ad $f, g : X \to \tilde{\mathbb{R}}$ are extended real valued functions we define $f \vee g$, $f \wedge g : X \to \tilde{\mathbb{R}}$ as

$$f \vee g(x) = f(x) \vee g(x); \quad f \wedge g(x) = f(x) \wedge g(x);$$

in particular the *positive part* f^+ and *negative part* f^- of a function $f : X \to \tilde{\mathbb{R}}$ are defined as $f^+ = f \vee 0$ and $f^- = (-f) \vee 0$ (notice that also the negative part is a positive function!), The function f may the be written as as $f = f^+ - f^-$ (notice that if $f^-(x) > 0$ then $f^+(x) = 0$, so that $f^+(x) - f^-(x)$ is always defined). The *absolute value* $|f|$ is defined as $|f| = f \vee (-f)$; we also have $|f| = f^+ + f^-$.

1.2.1. *Minimality of the positive and negative parts.* For future use we note here the following fact: the decomposition of f as $f = f^+ - f^-$ is minimal in the following sense: if $f = u - v$, with $u \geq 0$, $v \geq 0$ positive functions, then $f^+ \leq u$ and $f^- \leq v$. In fact

$$f^+(x) = f(x) \vee 0 = (u(x) - v(x)) \vee 0 \leq u(x) \vee 0 = u(x),$$

and similarly for $f^-(x)$. In particular we have for real functions f and g that $(f+g)^+ \leq f^+ + g^+$ and $(f+g)^- \leq f^- + g^-$. Notice also that $f \leq g$ is equivalent to $f^+ \leq g^+$ and $f^- \geq g^-$: in fact for every $x \in X$ we have $f^+(x) = f(x) \vee 0 \leq g(x) \vee 0 = g^+(x)$, and similarly for f^-, g^-: $g^-(x) = (-g(x)) \vee 0 \leq (-f(x)) \vee 0 = f^-(x)$.

1.2.2. *Subsets and characteristic functions.* A subset A of a given ambient set X is often identified with its *characteristic function*, or *indicator function*, $\chi_A : X \to \mathbb{R}$ defined by $\chi_A(x) = 1$ for $x \in A$, and $\chi_A(x) = 0$ for $x \in X \setminus A$. Notice that if $A, B \subseteq X$ then

$$\chi_{A\cup B} = \chi_A \vee \chi_B; \quad \chi_{A\cap B} = \chi_A \wedge \chi_B = \chi_A\, \chi_B; \quad \chi_{A\Delta B} = |\chi_A - \chi_B| = \chi_A + \chi_B - 2\chi_A\, \chi_B.$$

1.2.3. The following fact is sometimes useful:

. *If a, b, c, d are real numbers then*

$$|a \wedge b - c \wedge d| \leq |a - c| \vee |b - d| (\leq |a - c| + |b - d|); \quad |a \vee b - c \vee d| \leq |a - c| \vee |b - d| (\leq |a - c| + |b - d|).$$

Proof. First formula: the only case that is not immediate is when $a \wedge b = a$ and $c \wedge d = d$; we may assume $d \geq a$ by exchanging $a \wedge b$ with $c \wedge d$ if needed. Then we have $a \leq d \leq c$, so that $c - a \geq d - a \geq 0$, and $|a - c| = c - a \geq d - a = |d - a|$, as required. The second formula has an analogous proof, or comes from the first by changing signs. $\square$

1.3. Some topological facts. Recall that a topological space is said to be *separable* if it has a countable dense subset. A topological space is said to be *second countable* when it satisfies the second axiom of countability, that is, it has a countable base for its open sets. Recall that if τ is a topology on the set X, a base for τ is a subset $\mathcal{E} \subseteq \tau$ such that every $A \in \tau$ is a union of members of $\mathcal{E}$, that is, $A = \bigcup\{B \in \mathcal{E} :\, B \subseteq A\}$ for every $A \in \tau$. The following statement is then trivial:

. *If (X, τ) is a topological space, a subset $\mathcal{E}$ of τ is a base for τ if and only if given $p \in X$ and $A \in \tau$ with $p \in A$ there is $B \in \mathcal{E}$ with $p \in B$ and $B \subseteq A$.*

Next we observe:

. *Every second countable space is separable.*

Proof. Given a countable base $\mathcal{E}$, pick $x_B \in B$ for every non–empty $B \in \mathcal{E}$; the subset D of X so obtained is countable and dense: given an open non–empty subset A of X, A contains some $B \in \mathcal{E}$, hence $x_B \in A$, so that $A \cap D$ is non–empty. $\square$

1.3.1. *Metrizable and separable spaces.* Separability does not imply second countability: the Sorgenfrey line S (see 3.0.15) is an example of a separable not second countable space. In metrizable spaces this cannot happen:

Proposition. *A metrizable space is second countable if and only if it is separable.*

Proof. Let D be a countable dense subset of the metrizable space X, and let d be a topology determining metric on X. Let $\mathcal{E}$ be the set of all open balls with center at points $x \in D$ and rational strictly positive radii,

$$\mathcal{E} = \{B(x, r[:\ x \in D,\ r \in \mathbb{Q}^{>}\}.$$

Clearly $\mathcal{E}$ is countable, and we prove that it is a base for τ. If $p \in X$ and $A \in \tau$ with $p \in A$, pick first $r > 0$ such that $B(p, r[\subseteq A$. Next, pick $x \in D$ such that $d(p, x) < r/3$ (possible because D is dense in X) and a rational ρ with $r/3 < \rho < 2r/3$. Then $p \in B(x, \rho[$, and $B(x, \rho[\subseteq B(p, r[\subseteq A$. In fact, if $y \in B(x, \rho[$ we have

$$d(y, p) \le d(y, x) + d(x, p) < \rho + \frac{r}{3} < \frac{2r}{3} + \frac{r}{3} = r.$$

$\square$

Second countability is a hereditary property, which means that every subspace of a second countable space is also second countable (trivial: if $\mathcal{E}$ is a base for τ and $S \subseteq X$ then $\mathcal{E}_S = \{B \cap S :\ B \in \mathcal{E}\}$ is a base for the topology induced on S). Hence a second countable space is also *hereditarily separable*, that is, every subspace of a second countable space is separable.

1.3.2. *Disjoint families.* A family $(E_\lambda)_{\lambda\in\Lambda}$ of subsets of a set is called *disjoint* if its members are pairwise disjoint, that is, $\lambda, \mu \in \Lambda$ and $\lambda \ne \mu$ imply that $E_\lambda \cap E_\mu = \emptyset$. Then

. *In a separable space a disjoint family $(U_\lambda)_{\lambda\in\Lambda}$ of non–empty open subsets is at most countable.*

Proof. Let D be a countable dense subset of the space. Then $U_\lambda \cap D$ is non–empty for every $\lambda \in \Lambda$; picking a point $x(\lambda) \in U_\lambda \cap D$ we get a one–to–one map $\lambda \mapsto x(\lambda)$ of Λ into D. Then Λ is countable. $\square$

1.3.3. *Separating closed sets by continuous functions.* The following simple fact is often useful:

. *If X is a metrizable space and A, B are disjoint closed subsets of X there exists a continuous function $u : X \to [0, 1]$ such that $u(x) = 1$ for every $x \in A$ and $u(x) = 0$ for every $x \in B$.*

Proof. Let d be a topology determining metric on X; recall that the function $\rho_A(x) = \mathrm{dist}(x, A)(:= \inf\{d(x, y) : y \in A\})$ is continuous (it is even Lipschitz continuous) and $\rho_A(x) = 0$ iff $x \in A$; analogously for ρ_B. Then

$$u(x) = \frac{\rho_B(x)}{\rho_A(x) + \rho_B(x)}$$

is as required (notice that $\rho_A(x) + \rho_B(x) > 0$ for every $x \in X$, since ρ_A and ρ_B are non negative with no common zero). $\square$

Remark. If A and B are closed and disjoint in $\mathbb{R}^n$ it can be proved that u may also be taken smooth, that is, there is $u \in C^\infty(\mathbb{R}^n, \mathbb{R})$ such that $u(x) = 1$ on A, $u(x) = 0$ on B, and $0 \le u(x) \le 1$ for every $x \in \mathbb{R}^n$.

1.3.4. *Left and right limits.*

Proposition. *Let I be an interval of $\tilde{\mathbb{R}}$, let Y be a Hausdorff topological space, let $f : I \to Y$ be a function and let $\ell \in Y$. Then: if c is not the maximum (resp: not the minimum) of I the following are equivalent:*

(i) $\lim_{x\to c^+} f(x) = \ell$ *(resp:* $\lim_{x\to c^-} f(x) = \ell$*).*
(ii) *For every sequence $x_j \in I$ with $\lim_{j\to\infty} x_j = c$ and $x_j > c$ for every j (resp: $x_j < c$ for every j) we have $\lim_{j\to\infty} f(x_j) = \ell$.*
(iii) *For every strictly decreasing (resp: increasing) sequence $x_j \in I$ with $\lim_{j\to\infty} x_j = c$ we have $\lim_{j\to\infty} f(x_j) = \ell$.*

Proof. (i) implies (ii) implies (iii) are trivial. Let's prove that (iii) implies (i). Arguing by contradiction, if (i) is false then there is a neighborhood V of ℓ in Y such that for no $b \in I$ with $b > c$ we have $f(]c, b[) \subseteq V$. Define x_j in the following way: $x_0 > c$, $x_0 \in I$ is such that $f(x_0) \notin V$; assuming $x_0 > x_1 > \cdots > x_{j-1}$ have been defined, pick $x_j \in]c, x_{j-1}[\cap]c, c + 1/(j + 1)[$ such that $f(x_j) \notin V$. Then x_j is strictly decreasing and converges to c, but $f(x_j) \notin V$ for every j, so that $f(x_j)$ cannot converge to ℓ. $\square$

1.3.5. *Continuity of monotone functions.* We recall some easy results on monotone functions, that ought to be well–known.

PROPOSITION. *Let I be an interval of $\tilde{\mathbb{R}}$, and let $f : I \to \tilde{\mathbb{R}}$ be increasing (resp: decreasing). Then*

(i) *If $c \in I$ is not the maximum of I we have $\lim_{x\to c^+} f(x) = \inf f(\{x \in I : x > c\})$ (resp: $\lim_{x\to c^+} f(x) = \sup f(\{x \in I : x > c\})$.*

(ii) *If $c \in I$ is not the minimum of I we have $\lim_{x\to c^+} f(x) = \sup f(\{x \in I : x < c\})$ (resp: $\lim_{x\to c^+} f(x) = \inf f(\{x \in I : x < c\})$.*

(iii) *The set of points of discontinuity of f is at most countable.*

(iv) *If I is a closed interval then f is continuous if and only if for every non–empty subset A of I we have $f(\inf A) = \inf f(A)$ and $f(\sup A) = \sup f(A)$ (resp: $f(\inf A) = \sup f(A)$ and $f(\inf A) = \sup f(A)$).*

Proof. (i) Let $\ell = \inf f(\{x \in I : x > c\})$ and let V be an open interval of $\tilde{\mathbb{R}}$ containing ℓ, so that $\beta = \sup V > \ell$. We then have $f(b) < \beta$ for some $b \in I$ with $b > c$; by monotonicity of f we get $f(x) \le f(b) < \beta$ for every x such that $c < x \le b$. On the other hand $\ell \le f(x)$ for every $x \in]c, b]$, so $f(x) \in [\ell, \beta[\subseteq V$ for every $x \in]c, b]$, proving that $\lim_{x\to c^+} f(x) = \ell$.

(ii) has a proof entirely analogous to (i).

(iii) Assuming f increasing, by what just proved f is discontinuous at c in the interior of I iff $f(c^-) < f(c^+)$. Now if $c < d$ and $c, d \in \mathrm{int}(I)$ we have $f(c^+) \le f(d^-)$, so that the intervals $]f(c^-), f(c^+)[$ and $]f(d^-), f(d^+)[$ are disjoint; by 1.3.2, these intervals form a countable set.

(iv) That the condition is sufficient for continuity is immediate, by (i) and (ii). The necessity is a consequence of (i),(ii) and 1.3.4, recalling that if a set A has no minimum then there is a strictly decreasing sequence in A converging to $\inf A$, and analogously for $\max A$ and $\sup A$. □

1.3.6. *Increasing functions and left or right continuity.*

PROPOSITION. *Let I be an open interval of $\mathbb{R}$, and let $f,\ g : I \to \tilde{\mathbb{R}}$ be increasing functions; let $C(f)$ and $C(g)$ be the sets of points of continuity of f and g, respectively. The following are equivalent:*

(i) *$f(x) = g(x)$ for all x of a dense subset of I.*

(ii) *For every $c \in I$ we have $f(c^-) = g(c^-)$ and $f(c^+) = g(c^+)$.*

(iii) *$C(f) = C(g)$ and $f(x) = g(x)$ for every $x \in C(f) = C(g)$.*

Proof. (i) implies (ii). Let D be the dense subset of I on which f and g agree. Then $f(c^-) = \lim_{x\in D,\, x<c} f(x) = \lim_{x\in D,\, x<c} g(x) = g(c^-)$ and similarly for c^+.

(ii) implies (iii). We have $c \in C(f)$ if and only if $f(c^-) = f(c^+)$, and by (ii) this is equivalent to $g(c^-) = g(c^+)$, that is, to $c \in C(g)$.

(iii) implies (i). Since $I \smallsetminus C(f)$ is at most countable, $C(f)$ is dense in I. □

Then, if we alter an increasing function f on its points of discontinuity (which may be an infinite and even dense subset of I, see 2.2.2) in such a way that the function g so obtained is still increasing, the functions f and g have still the same right and left limits. The *right continuous modification of f* is the function $f_+ : I \to \tilde{\mathbb{R}}$ defined by $f_+(x) = f(x^+)$, while the *left continuous modification* is $f_-(x) = f(x^-)$. Sometimes (e.g. in Fourier analysis) it is useful to consider the *emivalue modification* f_0, defined by

$$f_0(x) = \frac{f(x^-) + f(x^+)}{2},$$

whose value is the center of the interval $[f(x^-), f(x^+)]$.

1.4. Boolean algebras of sets.

DEFINITION. Given a set X, a non–empty subset $\mathcal{A}$ of $\mathcal{P}(X)$ is said to be a (boolean) algebra of parts of X if given $A, B \in \mathcal{A}$ then $A \cup B$ and $X \smallsetminus A$ belong to $\mathcal{A}$ (briefly: an algebra of parts of X is a non–empty subset of $\mathcal{P}(X)$ closed under union and complementation).

Since a complementation closed subset of $\mathcal{P}(X)$ is closed under union iff it is closed under intersection (De Morgan's formulae: $A \cap B = X \smallsetminus ((X \smallsetminus A) \cup (X \smallsetminus B))$ and $A \cup B = X \smallsetminus ((X \smallsetminus A) \cap (X \smallsetminus B))$, we immediately get that an algebra is also closed under intersection, difference ($A \smallsetminus B := A \cap (X \smallsetminus B)$) and symmetric difference ($A \mathbin{\triangle} B := (A \smallsetminus B) \cup (B \smallsetminus A)$).

Recall that $(\mathcal{P}(X), \triangle, \cap)$ is a commutative ring, with $\triangle$ as addition, and $\cap$ as multiplication (neutral elements are $\emptyset$, additive, and X, multiplicative). Algebras of parts of X are exactly the subrings of $\mathcal{P}(X)$

containing the multiplicative identity X. In fact, if $\mathcal{A} \subseteq \mathcal{P}(X)$ contains X and is closed under symmetric difference, then it is closed under complementation, since $X \smallsetminus A = X \vartriangle A$ if $A \subseteq X$, and if $\mathcal{A}$ is also closed under intersection then it is closed under union, as remarked above, so that it is an algebra. Since the intersection of a set of subalgebras of $\mathcal{P}(X)$ is a subalgebra of $\mathcal{P}(X)$, given a subset $\mathcal{E}$ of $\mathcal{P}(X)$ we can speak of the algebra generated by $\mathcal{E}$ in $\mathcal{P}(X)$, the intersection of all subalgebras of $\mathcal{P}(X)$ containing $\mathcal{E}$.

1.4.1. *Finitely generated subalgebras of* $\mathcal{P}(X)$. We want to describe the algebra $\mathcal{A} = \mathcal{A}(\mathcal{E})$ generated by a finite subset $\mathcal{E}$ of $\mathcal{P}(X)$; we shall see that in this case $\mathcal{A}$ is also finite, and is generated by a finite partition of X. Let $\mathcal{G} = \mathcal{E} \cup \{X \smallsetminus E : E \in \mathcal{E}\}$ (that is, $\mathcal{G}$ is obtained by adding to $\mathcal{E}$ the complement in X of each of its elements). For every $x \in X$ we set $\mathcal{G}(x) = \{G \in \mathcal{G} : x \in G\}$ and let $G(x) = \bigcap_{G \in \mathcal{G}(x)} G$. Clearly $G(x)$ belongs to $\mathcal{A}$ (it is a finite intersection of elements of $\mathcal{A}$); moreover $x \in G(x)$, so that $G(x)$ is non–empty. And plainly $G = \bigcup_{x \in G} G(x)$, for every $G \in \mathcal{G}$. We claim that if $G(x) \neq G(y)$ then $G(x) \cap G(y) = \emptyset$, so that $\mathcal{B} = \{G(x) : x \in X\}$ is indeed a partition of X. In fact, $G(x) \neq G(y)$ implies $\mathcal{G}(x) \neq \mathcal{G}(y)$ (otherwise $G(x) = G(y)$, being intersections of the same set of subsets of X), and if a set E belongs to one but not to the other, say $x \in E$ but $y \notin E$, then $X \smallsetminus E \in \mathcal{G}(y)$, so that $G(x) \subseteq E$ but $G(y) \subseteq X \smallsetminus E$, and $G(x)$, $G(y)$ are actually disjoint. Clearly $\mathcal{B}$ is finite, it has at most $2^{2|\mathcal{E}|}$ elements. We claim now that:

. *Every element $A \in \mathcal{A}(\mathcal{E})$ can be written in a unique way as the union of a subset of $\mathcal{B}$.*

in fact, it is readily verified that the set of all these unions is an algebra, clearly contained in $\mathcal{A}(\mathcal{E})$; and this algebra contains $\mathcal{G} \supseteq \mathcal{E}$, since, as above seen, $G = \bigcup_{x \in G} G(x)$ for every $G \in \mathcal{G}$, so that it coincides with $\mathcal{A}(\mathcal{E})$.

Thus every finite algebra $\mathcal{A}$ is generated by a finite partition $\mathcal{B}$ of X; we call $\mathcal{B}$ the natural base of $\mathcal{A}$. We can describe $\mathcal{B}$ as the finest partition of X into elements of $\mathcal{A}$; we have $|\mathcal{A}| = 2^{|\mathcal{B}|}$.

1.4.2. *Semialgebras.* By a semialgebra of parts of a given set X we mean a subset $\mathcal{E} \subseteq \mathcal{P}(X)$ which is closed under intersection (A, $B \in \mathcal{E}$ implies $A \cap B \in \mathcal{E}$) and is such that for every $A \in \mathcal{E}$ the complement $X \smallsetminus A$ can be written as a finite disjoint union of elements of $\mathcal{E}$; we also assume that $\emptyset \in \mathcal{E}$. It is easy to prove that

. *If $\mathcal{E}$ is a semialgebra then the set $\mathcal{A}$ of all finite disjoint unions of elements of $\mathcal{E}$ is an algebra.*

Proof. Clearly $\mathcal{A}$ is closed under intersection:

$$(E_1 \cup \dots \cup E_m) \cap (F_1 \cup \dots \cup F_n) = \bigcup_{\{1 \le j \le m,\, 1 \le k \le n\}} E_j \cap F_k,$$

and hence it is closed also under complementation: for each j we have $X \smallsetminus E_j \in \mathcal{A}$ by hypothesis, and

$$X \smallsetminus (E_1 \cup \dots \cup E_m) = \bigcap_{j=1}^{m} (X \smallsetminus E_j).$$

□

One of the most important examples of a semialgebra is the semialgebra of all intervals of $\mathbb{R}$. Remember that an interval of $\mathbb{R}$ is subset of $\mathbb{R}$ which is order–convex, that is, $I \subseteq \mathbb{R}$ is an interval iff x_1, $x_2 \in I$ and $x_1 < x < x_2$ imply $x \in I$. The associated algebra is that of *plurintervals*, finite disjoint unions of intervals, the *interval algebra* for short. These intervals may be open, closed, half–open, reduced to singletons, bounded or unbounded, etc..

Exercise 1.4.1. Prove that the set of *left half–open* intervals $\mathcal{E} = \{]a, b] : a, b \in [-\infty, +\infty[\} \cup \{]a, +\infty[: a \in [-\infty, +\infty[\}$ is a semialgebra of subsets of $\mathbb{R}$. Same for the right half–open intervals.

Exercise 1.4.2. (*Important*) Let X, Y be sets, and let $\mathcal{E}$ and $\mathcal{F}$ be semialgebras of parts of X, Y respectively. Then $\mathcal{G} = \{E \times F : E \in \mathcal{E}, F \in \mathcal{F}\}$ is a semialgebra of parts of $X \times Y$.

Solution. Since $(E_1 \times F_1) \cap (E_2 \times F_2) = (E_1 \cap E_2) \times (F_1 \cap F_2)$ the set $\mathcal{G}$ is closed under intersection. And if $E \times F \in \mathcal{G}$ then $X \times Y \smallsetminus (E \times F) = ((X \smallsetminus E) \times Y) \cup (E \times (Y \smallsetminus F))$, disjoint union; if $X \smallsetminus E = E_1 \cup \dots \cup E_m$ and $Y \smallsetminus F = F_1 \cup \dots \cup F_n$, both disjoint unions, then

$$X \times Y \smallsetminus (E \times F) = \left(\left(\bigcup_{j=1}^{m} E_j \right) \times (F \cup F_1 \cup \dots \cup F_n) \right) \cup \left(\bigcup_{k=1}^{n} E \times F_k \right) =$$

$$\left(\bigcup_{j=1}^{m} E_j \times F\right) \cup \left(\bigcup_{1\le j\le m, 1\le k\le n} E_j \times F_k\right) \cup \left(\bigcup_{k=1}^{n} E \times F_k\right),$$

a finite disjoint union of elements of $\mathcal{G}$. □

By induction it is easy to generalize this fact to any finite family of factors: if $(X_k)_{1\le k\le m}$ is a finite family of sets, and $\mathcal{E}_k$ is a semialgebra on X_k, for $k \in \{1, \dots, m\}$, then $\mathcal{G} = \{E_1 \times \cdots \times E_m : E_k \in \mathcal{E}_k\}$ is a semialgebra on $\prod_{k=1}^{m} X_k$. If we call $n-$dimensional interval any subset of $\mathbb{R}^n$ which is of the form $I_1 \times \cdots \times I_n$, where each I_k is an interval of $\mathbb{R}$, then $n-$dimensional intervals in $\mathbb{R}^n$ are a semialgebra.

1.5. Ideals and filters. In a commutative ring $(R, +, \cdot)$ an ideal is a subset I of R which is additively a subgroup and is such that $a \in R$ and $b \in I$ imply $a\,b \in I$. In an algebra $\mathcal{A}$ of parts of X an ideal is then a subset $\mathcal{I}$ which is closed under symmetric difference and such that $A \in \mathcal{A}$ and $B \in \mathcal{I}$ imply $A \cap B \in \mathcal{I}$. It is easy to see that $\mathcal{I} \subseteq \mathcal{A}$ is an ideal if and only if it is closed under union and contains the subsets of its elements which are elements of $\mathcal{A}$. An ideal $\mathcal{I}$ of the algebra $\mathcal{A}$ is of course said to be proper if it is properly contained in $\mathcal{A}$, equivalently iff $X \notin \mathcal{I}$.

Exercise 1.5.1. Let X be a set. Prove that for a subset $\mathcal{R}$ of $\mathcal{P}(X)$ the following are equivalent:

(i) $\mathcal{R}$ is closed under symmetric difference and intersection.
(ii) $\mathcal{R}$ is closed under union and (set–theoretic) difference.

A non–empty subset $\mathcal{R}$ of $\mathcal{P}(X)$ closed under $\triangle$ and $\cap$ is called a (boolean) *ring* of parts of X. Prove that if $\mathcal{R}$ is a ring then

$$\mathcal{A} = \mathcal{R} \cup \{X \setminus A : A \in \mathcal{R}\}$$

is an algebra, and that $\mathcal{R}$ is an ideal of $\mathcal{A}$.

Solution. (i) implies (ii) Given A, $B \in \mathcal{R}$ we have $A \cup B = (A \triangle B) \triangle (A \cap B)$, and $A \setminus B = A \triangle (A \cap B)$.

(ii) implies (i) Given A, $B \in \mathcal{R}$ we have $A \triangle B = (A \setminus B) \cup (B \setminus A)$ and $A \cap B = A \setminus (A \setminus B)$.

Next we prove that $\mathcal{A}$ is an algebra. Clearly $\mathcal{A}$ is complementation closed. We prove that it is closed under intersection. Given $A, B \in \mathcal{R}$ we know that $A \cap B \in \mathcal{R}$; if $A \in \mathcal{R}$ and $X \setminus B \in \mathcal{R}$ then $A \cap B = A \setminus (X \setminus B) \in \mathcal{R}$; finally, if $X \setminus A$ and $X \setminus B \in \mathcal{R}$ then $(X \setminus A) \cap (X \setminus B) \in \mathcal{R}$, so that the complement $A \cap B$ of this set is in $\mathcal{A}$. In the course of the proof we have also proved that $\mathcal{R}$ is an ideal of $\mathcal{A}$. □

1.5.1. *Dual ideals, or filters.* If $\mathcal{I}$ is a proper ideal of the algebra $\mathcal{P}(X)$ of all subsets of X, then the set

$$\mathcal{F}_{\mathcal{I}} = \{X \setminus A : A \in \mathcal{I}\}$$

consisting of all complements of elements of $\mathcal{I}$ is called *filter* on X. Directly stated:

Definition. If X is a set, a filter on X is subset $\mathcal{F}$ of $\mathcal{P}(X)$ such that:

(i) If U, $V \in \mathcal{F}$ then $U \cap V \in \mathcal{F}$ (a filter is closed under intersection).
(ii) If $U \in \mathcal{F}$, $V \in \mathcal{P}(X)$ and $V \supseteq U$, then $V \in \mathcal{F}$ (a filter is closed under the formation of supersets).
(iii) The emptyset does not belong to $\mathcal{F}$ (equivalently under (ii): $\mathcal{F}$ is a proper subset of $\mathcal{P}(X)$).

It is easy to see that $\mathcal{F} \subseteq \mathcal{P}(X)$ is a filter if and only if $\mathcal{I} = \{X \setminus U : U \in \mathcal{F}\}$ is a proper ideal of $\mathcal{P}(X)$.

A base of the filter $\mathcal{F}$ on X is a subset $\mathcal{B}$ of $\mathcal{F}$ such that for every $F \in \mathcal{F}$ there is $B \in \mathcal{B}$ contained in F, $F \supseteq B$. Any subset $\mathcal{B}$ of $\mathcal{P}(X)$ which is the base of a filter is called *filterbase.* It is easy to see that

Proposition. *Let X be a set, and let $\mathcal{B}$ be a subset of $\mathcal{P}(X)$, Then $\mathcal{B}$ is a filterbase if and only if $\emptyset \notin \mathcal{B}$, and given U, $V \in \mathcal{B}$ there exists $W \in \mathcal{B}$ such that $W \subseteq U \cap V$.*

Of course a filterbase is base of a unique filter on X, which is

$$\mathcal{F}(\mathcal{B}) = \{F \subseteq X : \text{there is } U \in \mathcal{B} \text{ such that } F \supseteq U\}.$$

Given a filter $\mathcal{F}$ on X, consider the dual ideal $\mathcal{I} = \{X \setminus U : U \in \mathcal{F}\}$. Then $\mathcal{A} = \mathcal{I} \cup \mathcal{F}$ is an algebra of parts of X, in which $\mathcal{I}$ is a proper ideal (the verification is immediate: clearly $\mathcal{A}$ is complementation closed, and if A, B are both in $\mathcal{I}$ then $A \cup B = X \setminus ((X \setminus A) \cap (X \setminus B))$ is also in I (since $X \setminus A$ and $X \setminus B$ are in $\mathcal{F}$ their intersection is in $\mathcal{F}$, hence the complement of this intersection is in $\mathcal{I}$); and if either $A \in \mathcal{F}$ or $B \in \mathcal{F}$ then $A \cup B \in \mathcal{F}$ as superset of a set in $\mathcal{F}$, so that $\mathcal{A}$ is an algebra).

1.5.2. *Examples of filters.* Given a topological space (X, τ) and a point $p \in X$, recall that a neighborhood of p in X is any subset U of X containing an open set A which contains p as an element, $p \in A \subseteq U$. It is plain that the set $\mathcal{I}_p$ of all neighborhoods of p in X is a filter; this filter has its open elements as a base; in a metrizable space, also the balls centered at p with strictly positive radius are a base.

In a set X, call a subset U *cofinite* if its complement $X \setminus U$ is finite; it is clear that if X is infinite the cofinite sets are a filter; so that the set $\mathcal{A}$ consisting of all finite and cofinite subsets of X is an algebra (the cofinite algebra).

We call *countable* a set that is finite, or countably infinite (i.e. of the same cardinality as $\mathbb{N}$), and we call co–countable in X any subset of X whose complement in X is countable; in an uncountable set the co–countable subsets are a filter, and countable and co–countable subsets of X together make an algebra (the co–countable algebra).

1.6. Generalized sequences (nets). Sequences are very useful in analysis, and they suffice to describe the topology of every metrizable space. But the need, or at least the convenience of use, for a more general notion often arises, in various circumstances. Let's give some definitions. Recall that a *preorder* on a set is a binary relation $\preceq$ that is reflexive and transitive; a preordered set is a pair $(D, \preceq)$ formed by a set D and a preorder $\preceq$ on D. A *directed set* is a preordered set such that for any pair $\alpha, \beta \in D$ there is $\gamma \in D$ such that $\alpha \preceq \gamma$ and $\beta \preceq \gamma$. Given a set X, and a directed set $(D, \preceq)$ a *generalized sequence*, or *net* in X, indexed by D, is a function $x : D \to X$, $\alpha \mapsto x_\alpha = x(\alpha)$.

1.6.1. *Limit of a net in a topological space.* If X has a topology, and $\ell \in X$, we say that ℓ is a limit of the net $x : D \to X$ if for every neighborhhood V of ℓ in X there is $\alpha(V) \in D$ such that $x_\alpha \in V$ for every $\alpha \succeq \alpha(V)$ (this fact is also expressed by saying that the net x_α is *eventually in* V). It is easy to see that if X is an Hausdorff space then a net has at most one limit: if ℓ_1, $\ell_2 \in X$, with $\ell_1 \neq \ell_2$ then the net x cannot converge to both ℓ_1 and ℓ_2: let V_1, V_2 be disjoint nbhds of ℓ_1 and ℓ_2 respectively; then there is α_1 such that $x_\alpha \in V_1$ for $\alpha \succeq \alpha_1$ and there is α_2 such that $x_\alpha \in V_2$ for $\alpha \geq \alpha_2$; since D is directed there is $\alpha \succeq \alpha_1, \alpha_2$, and for such an α we have $x_\alpha \in V_1 \cap V_2 = \emptyset$, absurd.

A subset E of the directed set $(D, \preceq)$ is said to be *cofinal* in D if for every $\alpha \in D$ there is $\beta \in D$ such that $\beta \succeq \alpha$. Observe that a net $x : D \to X$ in the topological space X does not have the point $p \in X$ as limit if and only if there is a neighborhood V of p in X such that the set $x^{\leftarrow}(X \setminus V) = \{\beta \in D : x(\beta) \notin V\}$ is cofinal in D.

1.6.2. *The directed set associated to a filterbase.* Every filterbase $\mathcal{B}$, partially ordered by reverse inclusion, is a directed set (given $U, V \in \mathcal{B}$ the set $U \cap V$ contains some $W \in \mathcal{B}$). However the directed set associated to $\mathcal{B}$ is not $\mathcal{B}$ itself, but the following

$$D = \{(x, U) : U \in \mathcal{B}, x \in U\}; \quad \text{with the preorder} \quad (x, U) \preceq (y, V) \overset{\text{def}}{\iff} U \supseteq V.$$

1.6.3. *Nets and topology.* The notion of limit of a net is not really new, it can be interpreted also within the usual notion of limit in general topology. If $(D, \preceq)$ is a directed set, take an object $\infty(D) \notin D$, and consider $E = D \cup \{\infty(D)\}$; put on E the topology for which all subsets of D are open (equivalently, all points of D are isolated in E), while a subset of E that contains $\infty(D)$ is declared open iff it contains a subset like $T_\alpha = \{\beta \in D : \beta \succeq \alpha\}$, for some $\alpha \in D$ (the $\alpha-$tail of D). It is easy to see that this is a topology on E; directedness of the preorder is needed to ensure that a finite intersection of open sets containing $\infty(D)$ is still open. In this topology $\infty(D)$ is the only non–isolated point of E; and a net $x : D \to X$, with X a topological space, has $\ell \in X$ as limit if and only if $x : D \to X$ has limit ℓ as α tends to $\infty(D)$ in the topological space E. In this way the notion of limit of net is reduced to the topological notion.

Conversely, assume that T is a topological space and that c is a non–isolated point of T; let $S = T \setminus \{c\}$. Then the set

$$\mathcal{I}_c = \{U = Z \setminus \{c\} : \ Z \text{ a neighborhood of } c \text{ in } T\}$$

is a filter of subsets of S. Take any base $\mathcal{B}$ for this filter and consider the directed set $(D, \preceq)$ associated to $\mathcal{B}$ as above (1.6.2); if Y is another topological space and $f : S \to Y$ is a function we have that $\lim_{x \to c} f(x) = \ell \in Y$ if and only if the net $y_f : D \to Y$ defined by $y_f(x, U) = f(x)$ has limit ℓ; the proof is immediate. In this way a topological limit may be interpreted as the limit of a net.

1.6.4. *Limits of monotone nets.* As with sequences, monotone nets with values in the extended real line $\tilde{\mathbb{R}}$ always have limits. Of course a net is called increasing if $\alpha \preceq \beta$ implies $x(\alpha) \leq x(\beta)$, and decreasing if it implies $x(\alpha) \geq x(\beta)$.

PROPOSITION. *Let $(D, \preceq)$ be a directed set, and suppose that $x : D \mapsto \tilde{\mathbb{R}}$ is a monotone net. Then $\lim_{\alpha \in D} x(\alpha)$ always exists in the extended real line $\tilde{\mathbb{R}}$; it is $\sup x(D)$ if x is increasing, and $\inf x(D)$ if x is decreasing.*

The proof is identical to the one for sequences, and is left to the reader.

1.7. Lower and upper semicontinuous functions.

1.7.1. *Definition.* The *lower topology* on $\tilde{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ is the topology whose open sets are, besides $\emptyset$ and the whole space $\tilde{\mathbb{R}}$, the open half–lines with a lower bound, $V_a = \{t \in \tilde{\mathbb{R}} : t > a\}$, $a \in \mathbb{R}$, or $a = -\infty$; it is a non–Hausdorff topology. If (X, τ) is a topological space, a function $f : X \to \tilde{\mathbb{R}}$ is said to be *lower semicontinuous*, lsc for short, if it is continuous when the target space $\tilde{\mathbb{R}}$ has this topology, equivalently: for every $a \in \mathbb{R}$ the set $\{f > a\} := \{x \in X : f(x) > a\}$ is open in X. Clearly continuous functions, those that are continuous when $\tilde{\mathbb{R}}$ has the usual topology, are also lsc functions.

1.7.2. *Pointwise suprema of lsc functions are lsc.*

. *Let X be a topological space, and let $f_\lambda : X \to \tilde{\mathbb{R}}$ be a family of lsc functions. Then the function $f : X \to \tilde{\mathbb{R}}$ defined by $f(x) = \sup\{f_\lambda(x) : \lambda \in \Lambda\}$ is also lsc.*

Proof. For every $a \in \mathbb{R}$ we have

$$\{f > a\} = \bigcup_{\lambda \in \Lambda} \{f_\lambda > a\}.$$

□

In particular, a supremum of a family of continuous functions is lower semicontinuous, although in general not continuous if the family is infinite.

EXERCISE 1.7.1. Consider $\tilde{\mathbb{R}}$ with the lower topology. Prove that a non–empty subset $C \subseteq \tilde{\mathbb{R}}$ is compact in the induced topology if and only if has a minimum. Deduce from it that if X is compact and $f : X \to \tilde{\mathbb{R}}$ is l.s.c., then f has a minimum.

Solution. If C does not have a minimum in $\tilde{\mathbb{R}}$, then $\{]a, \infty] : a \in C\}$ is a cover of C (because for every $x \in C$ there is $a \in C$ with $a < x$) by sets open in the lower topology, which has no finite subcover (for every finite subset $\{a_1, \dots, a_m\}$ of C we have $\bigcup_{k=1}^{m}]a_k, \infty] =]a, \infty]$ where $a = \min\{a_1, \dots, a_m\} \in C$, so that $a \in C \setminus (\bigcup_{k=1}^{m}]a_k, \infty])$, and $\{]a_k, \infty] : k = 1, \dots, m\}$ is not a subcover). So if $C \neq \emptyset$, and $\min C$ does not exist, then C is not compact. And if $\min C = a$ exists, this minimum is either $-\infty$ or not; if it is $-\infty$ then the only open set containing it is the entire space; otherwise, given an open cover $(]a_\lambda, \infty])_{\lambda \in \Lambda}$ of C we must have $a \in]a_\lambda, \infty]$ for some $\lambda \in \Lambda$, and then $C \subseteq]a_\lambda, \infty]$ for this λ.

Since the continuous image of a compact space is compact, $f(X)$ must be compact in the lower topology, hence $\min f(X)$ must exist. □

1.7.3. *Upper topology.* To complete the picture we have to define also the *upper topology* of $\tilde{\mathbb{R}}$, the topology whose non trivial open sets are the upper bounded half–lines $\{t < a\}$, with $a \in \mathbb{R}$ or $a = \infty$. Given a topological space X, the *upper semicontinuous functions*, usc for short, are the $f : X \to \tilde{\mathbb{R}}$ continuous when the range space has this topology. Clearly, if X is a topological space and $f : X \to \tilde{\mathbb{R}}$ is a function then f is if and only if $\{f < a\}$ is open in X, for every $a \in \mathbb{R}$.

And a pointwise infimum of usc functions is still usc, etc. A function f is usc iff its opposite $-f$ is lsc.

1.7.4. *The usual topology.* The usual topology of $\tilde{\mathbb{R}}$ has the open half–lines $]b, \infty]$ and $[-\infty, a[$, with $a, b \in \mathbb{R}$ as a *subbase* for the open sets, in the sense that finite intersections of thes half–lines, the open intervals $]a, b[$ and the half–lines themselves are a base (see 1.3). It follows that a function $f : X \to \tilde{\mathbb{R}}$ from the topological space X the $\tilde{\mathbb{R}}$ with the usual topology is continuous if and ony if it is both lower semicontinuous and upper semicontinuous. Observe that the characteristic function of an open set is lsc, the characteristic function of a closed set is usc.

1.8. Limsup and liminf. There are various situation in which the use of the notion of limit is precluded from the fact that a limit does not always exist. For real valued sequences, or more generally for real valued functions, the presence of a complete order on $\tilde{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ allows us to define two notions weaker than the limit notion, but always existent; these notions are fruitful in many situations, and are used to formulate various important theorems of mathematical analysis.

1.8.1. *Definitions.* Let $(D, \preceq)$ be a directed set, and let $x : D \to \tilde{\mathbb{R}}$ be a generalized sequence with values in the extended real line $\tilde{\mathbb{R}}$. For each $\alpha \in D$ we set

$$x_*(\alpha) = \inf\{x(\beta);\, \beta \succeq \alpha\}; \quad x^*(\alpha) = \sup\{x(\beta) : \beta \succeq \alpha\}.$$

Observe that x_* is an increasing net, while x^* is decreasing: this is obvious, since if $\beta \succeq \alpha$ the set $\{x_\gamma : \gamma \succeq \beta\}$ is a subset of the set $\{x_\gamma : \gamma \succeq \alpha\}$, by transitivity of the preorder, so that it has a larger infimum, and a smaller supremum. Then these nets have a limit in the space $\tilde{\mathbb{R}}$ (see 1.6.4). We set

Definition.

$$\liminf_{\alpha \in D} x(\alpha) = \lim_{\alpha \in D} x_*(\alpha) \left(= \sup_{\alpha \in D} x_*(\alpha) = \sup_{\alpha \in D} \{\inf\{x(\beta) : \beta \succeq \alpha\}\} \right);$$

$$\limsup_{\alpha \in D} x(\alpha) = \lim_{\alpha \in D} x^*(\alpha) \left(= \inf_{\alpha \in D} x^*(\alpha) = \inf_{\alpha \in D} \{\sup\{x(\beta);\, \beta \succeq \alpha\}\} \right).$$

1.8.2. *First properties; limits.* Notice that if $a \in \tilde{\mathbb{R}}$, and $a > \liminf_{\alpha \in D} x(\alpha)$ then a cannot be the limit of the net (not even in the lower topology): if $a_1 < a$, but $a_1 > \liminf_{\alpha \in D} x(\alpha)$ then for every $\alpha \in D$ there is $\beta \succeq \alpha$ such that $x(\beta) < a_1$, and for no $\alpha \in D$ the set $\{x(\beta) : \beta \succeq \alpha\}$ is contained in the neighborhood $]a_1, +\infty]$ of a. Similarly, if $a < \limsup_{\alpha \in D} x(\alpha)$ then a is not a limit of the net x (not even in the upper topology).

. *If $x : D \to \tilde{\mathbb{R}}$ is a net then*

$$\liminf_{\alpha \in D} x(\alpha) \le \limsup_{\alpha \in D} x(\alpha). \tag{*}$$

Moreover the net has a limit in $\tilde{\mathbb{R}}$ if and only if $\limsup_{\alpha \in D} x(\alpha) \le \liminf_{\alpha \in D} x(\alpha)$; in this case we have $\limsup_{\alpha \in D} x(\alpha) = \liminf_{\alpha \in D} x(\alpha)$, and this common value is the limit of the net.

Proof. For every $\alpha \in D$ we have $x_*(\alpha) \le x(\alpha) \le x^*(\alpha)$; passing to the limit in the inequality $x_*(\alpha) \le x^*(\alpha)$ we get $\liminf_{\alpha \in D} x(\alpha) \le \limsup_{\alpha \in D} x(\alpha)$. Then $\limsup_{\alpha \in D} x(\alpha) \le \liminf_{\alpha \in D} x(\alpha)$ is equivalent to $\liminf_{\alpha \in D} x(\alpha) = \limsup_{\alpha \in D} x(\alpha)$; if this happens then also $\lim_{\alpha \in D} x(\alpha)$ is this common value (simply apply the three functions theorem (it. teorema dei carabinieri) to the inequality $x_*(\alpha) \le x(\alpha) \le x^*(\alpha)$). And if $\liminf_{\alpha \in D} x(\alpha) < \limsup_{\alpha \in D} x(\alpha)$, then the limit of the net cannot exist, by what observed just before the statement (a limit has to be not strictly larger than lim inf and not strictly smaller that lim sup). □

1.8.3. *Composition with monotone functions.*

. *Let $(D, \preceq)$ be a directed set, let $x : D \to I$ be a net, where I is a closed interval of $\tilde{\mathbb{R}}$, and let $\varphi : I \to \tilde{\mathbb{R}}$ be continuous and monotone. Then*

$$\liminf_{\alpha \in D} \varphi \circ x(\alpha) = \varphi(\liminf_{\alpha \in D} x(\alpha)); \quad \limsup_{\alpha \in D} \varphi \circ x(\alpha) = \varphi(\limsup_{\alpha \in D} x(\alpha)) \quad \text{if } \varphi \text{ is increasing}$$

$$\liminf_{\alpha \in D} \varphi \circ x(\alpha) = \varphi(\limsup_{\alpha \in D} x(\alpha)); \quad \limsup_{\alpha \in D} \varphi \circ x(\alpha) = \varphi(\liminf_{\alpha \in D} x(\alpha)) \quad \text{if } \varphi \text{ is decreasing.}$$

Proof. The easy proof is left to the reader: it uses the fact that if φ is continuous and increasing then $\varphi(\inf A) = \inf \varphi(A)$ and $\sup \varphi(A) = \varphi(\sup A)$ for every $A \subseteq I$, while continuous decreasing functions exchange inf and sup (see 1.3.5). Notice in particular that $\liminf_{\alpha \in D}(-x(\alpha)) = -\limsup_{\alpha \in D} x(\alpha)$ and $\limsup_{\alpha \in D}(-x(\alpha)) = -\liminf_{\alpha \in D} x(\alpha)$. □

Example 1.8.1. For $D = \mathbb{R}$, $c = +\infty$ we have $\liminf_{x \to +\infty} \sin x = -1$, $\limsup_{x \to +\infty} \sin x = 1$ (trivially: for every punctured nbhd U of $+\infty$ in $\tilde{\mathbb{R}}$ we have $\inf \sin U = -1$, $\sup \sin U = 1$). And $\liminf_{x \to +\infty} \tanh x \cos^2 x = 0$, $\limsup_{x \to +\infty} \tanh x \cos^2 x = 1$: also in this case, if $f(x) = \tanh x \cos^2 x$ we have $\inf f(U) = 0$ and $\sup f(U) = 1$ for every punctured nbhd U of $+\infty$ in $\tilde{\mathbb{R}}$ (for a more detailed analysis see at the end).

1.8.4. *Limsup and liminf for sequences.* Particularly important is the case of sequences ($D = \mathbb{N}$, $c = \infty$). Let us review the definition in this case, in detail. Given a sequence $n \mapsto x_n$ of real numbers, for every $m \in \mathbb{N}$ we set

$$x_{*m} = \inf\{x_n : n \ge m\}; \quad x^*_m = \sup\{x_n : n \ge m\},$$

so that x_{*m} is an increasing sequence of extended real numbers, and x^*_m a decreasing one; moreover $x_{*m} \le x_n \le x^*_m$ for every $n, m \in \mathbb{N}$ with $n \ge m$. As monotone sequences in the order–complete set $\tilde{\mathbb{R}}$ these sequences have a limit:

$$\lim_{m\to\infty} x_{*m} = \sup\{x_{*m} : m \in \mathbb{N}\} = \liminf_{n\to\infty} x_n;$$
$$\lim_{m\to\infty} x^*_m = \inf\{x^*_m : m \in \mathbb{N}\} = \limsup_{n\to\infty} x_n.$$

EXERCISE 1.8.2. For a sequence $(x_n)_{n\in\mathbb{N}}$ of real numbers the following are equivalent:

(i) The sequence has no lower bound (resp: no upper bound) in $\mathbb{R}$.
(ii) For some $m \in \mathbb{N}$ we have $x_{*m} = -\infty$ (resp: $x^*_m = +\infty$).
(iii) For every $m \in \mathbb{N}$ we have $x_{*m} = -\infty$ (resp: $x^*_m = +\infty$).

EXERCISE 1.8.3. Let $(s_n)_{n\ge 1}$ be the sequence inductively defined by

$$\begin{aligned} s_1 &= 1;\\ s_{2n} &= s_{2n-1} - 1 - \frac{(-1)^{n+1}}{n+1} \qquad (n \ge 1);\\ s_{2n+1} &= 1 + s_{2n} \qquad (n \ge 1)\,. \end{aligned}$$

Find

$$\limsup_{n\to+\infty} s_n, \qquad \liminf_{n\to+\infty} s_n\ .$$

EXERCISE 1.8.4. If $x_n = (-1)^n(n+1)/(n-1/2)$, determine $\liminf_{n\to\infty} x_n$ and $\limsup_{n\to\infty} x_n$.

Solution. Notice that $x \mapsto (x+1)/(x-1/2)$ is strictly decreasing for $x > 1/2$ (derivative $-3/(2(x-1/2)^2)$), and $\lim_{x\to+\infty}(x+1)/(x-1/2) = 1$. Excluding $n = 0$, the subsequence of terms of even index is decreasing to 1, while that of terms of odd index is increasing to -1. This clearly shows that

$$\liminf_{n\to\infty} x_n = -1; \qquad \limsup_{n\to\infty} x_n = 1.$$

For a more formal solution: given $m \ge 1$ we have:

$$\text{for } m \text{ odd}: \quad x_{*m} = -\frac{m+1}{m-1/2}; \quad x^*_m = \frac{(m+1)+1}{(m+1)-1/2}$$
$$\text{for } m \text{ even}: \quad x_{*m} = -\frac{(m+1)+1}{(m+1)-1/2}; \quad x^*_m = \frac{m+1}{m-1/2}.$$

□

Very often one writes maxlim or minlim in place of lim sup and lim inf. This may look surprising: in elementary calculus one is always at pains to distinguish between the quite different notions of maximum and supremum, or minimum and infimum! A justification is given by the following proposition. Given a sequence $(x_n)_{n\in\mathbb{N}}$ (in a metrizable space X) all the points of the space X which are limits of some subsequence $(x_{\mu(k)})_{k\in\mathbb{N}}$ are called *cluster points* (ital: valori di aderenza), sometimes also *limit points*, of the original sequence $(x_n)_{n\in\mathbb{N}}$ in X. Then:

PROPOSITION. *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of real numbers. The set of cluster points in $\tilde{\mathbb{R}}$ of the sequence has a minimum, which is $\liminf_{n\to\infty} x_n$, and a maximum, which is $\limsup_{n\to\infty} x_n$.*

Proof. The proof is at the end of the section. □

1.8.5. *Operations with limsup and liminf.*

. *Let $(D, \preceq)$ be a directed set, and let $x, y : D \to \tilde{\mathbb{R}}$ be generalized sequences.*

(0) *if $x(\alpha) \le y(\alpha)$ for every $\alpha \in D$, then*

$$\liminf_{\alpha\in D} x(\alpha) \le \liminf_{\alpha\in D} y(\alpha); \qquad \limsup_{\alpha\in D} x(\alpha) \le \limsup_{\alpha\in D} y(\alpha)$$

Assume now that $(x+y)(\alpha) = x(\alpha) + y(\alpha)$ is defined for every $\alpha \in D$. Then

(i) *Liminf is superadditive and limsup is subadditive, that is*

$$\liminf_{\alpha\in D}(x(\alpha)+y(\alpha)) \geq \liminf_{\alpha\in D} x(\alpha) + \liminf_{\alpha\in D} y(\alpha);$$
$$\limsup_{\alpha\in D}(x(\alpha)+y(\alpha)) \leq \limsup_{\alpha\in D} x(\alpha) + \limsup_{\alpha\in D} y(\alpha)$$

provided that the right-hand sides are meaningful, i.e no $+\infty-\infty$ *or* $-\infty+\infty$ *is encountered.*

(ii) *If* $\lim_{\alpha\in D} y(\alpha)$ *exists in* $\tilde{\mathbb{R}}$, *then additivity holds. That is, we have*

$$\liminf_{\alpha\in D}(x(\alpha)+y(\alpha)) = \liminf_{\alpha\in D} x(\alpha) + \lim_{\alpha\in D} y(\alpha)$$
$$\limsup_{\alpha\in D}(x(\alpha)+y(\alpha)) = \limsup_{\alpha\in D} x(\alpha) + \lim_{\alpha\in D} y(\alpha)$$

again provided that the right-hand sides are meaningful.

Proof. (0) is trivial. (i) We prove only the assertion on $\limsup$. Given $\alpha\in D$ we have for every $\beta\succeq\alpha$:

$$(x+y)(\beta) = x(\beta)+y(\beta) \leq x^*(\alpha)+y^*(\alpha),$$

so that, taking suprema on the left–hand side for $\beta\succeq\alpha$ we get

$$(x+y)^*(\alpha) = \sup\{x(\beta)+y(\beta) : \beta\succeq\alpha\} \leq x^*(\alpha)+y^*(\alpha);$$

taking limits as α varies in D we get the asserted inequality.

Remark. The inequality is in general strict. For instance if $f(x)=\sin x$ and $g(x)=-\sin x$ we have $\limsup_{x\to+\infty} f(x) = \limsup_{x\to+\infty} g(x) = 1$, so that the sum of the limsups is 2, but $f+g=0$ has limit 0.

We prove (ii) again for $\limsup$ only. Writing $a=\limsup_{\alpha\in D} x(\alpha)$ and $b=\lim_{\alpha\in D} y(\alpha)$ we need only to prove that

$$\limsup_{\alpha\in D}(x+y)(\alpha) \geq a+b.$$

We can assume $a+b>-\infty$, since otherwise the above inequality is trivial. Then $a>-\infty$ and $b>-\infty$. Given a real number $c<a+b$ we can pick real numbers $u<a$ and $v<b$ such that $u+v>c$; there is now $\bar\alpha\in D$ such that $y(\alpha)>v$ for $\alpha\succeq\bar\alpha$; and since $u<a\leq x^*(\alpha)$ we can pick $\beta\succeq\alpha$ such that $x(\beta)>u$. It follows that $(x+y)(\beta)=x(\beta)+y(\beta)>u+v$ so that $(x+y)^*(\alpha)>u+v>c$ for every $\alpha\succeq\bar\alpha$. This implies $\limsup_{\alpha\in D}(x+y)(\alpha)\geq c$, and since c is an arbitrary number smaller that $a+b$, we obtain $\limsup_{\alpha\in D}(x+y)(\alpha)\geq a+b$, as required. □

Exercise 1.8.5. Let a_n and b_n be real sequences, and let $a,b\in\mathbb{R}$. Assume that

$$\liminf_{n\to\infty} a_n \geq a, \quad \liminf_{n\to\infty} b_n \geq b; \quad \limsup_{n\to\infty}(a_n+b_n) \leq a+b.$$

Prove that then $\lim_{n\to\infty} a_n = a$ and $\lim_{n\to\infty} b_n = b$.

1.8.6. *Multiplicative versions.* For positive real valued functions there of course is a multiplicative analogue of the previous result, whose proof is entirely left to the reader:

. *Let* $(D,\preceq)$ *be a directed set, and let* $x,y: D\to[0,\infty]$ *be generalized sequences. Assume that for every* $\alpha\in D$ *the product* $x(\alpha)\,y(\alpha)$ *is defined. Then:*

(i) *Liminf is supermultiplicative and limsup is submultiplicative, that is*

$$\liminf_{\alpha\in D}(x(\alpha)\,y(\alpha)) \geq (\liminf_{\alpha\in D} x(\alpha))\,(\liminf_{\alpha\in D} y(\alpha));$$
$$\limsup_{\alpha\in D}(x(\alpha)\,y(\alpha)) \leq (\limsup_{\alpha\in D} x(\alpha))\,(\limsup_{\alpha\in D} y(\alpha))$$

provided that the right-hand sides are meaningful, i.e no $0\,(\pm\infty)$ *or* $(\pm\infty)\,0$ *is encountered.*

(ii) *If* $\lim_{\alpha\in D} y(\alpha)$ *exists in* $[0,\infty]$, *then multiplicativity holds. That is, we have*

$$\liminf_{\alpha\in D}(x(\alpha)\,y(\alpha)) = (\liminf_{\alpha\in D} x(\alpha))\,(\lim_{\alpha\in D} y(\alpha))$$
$$\limsup_{\alpha\in D}(x(\alpha)\,y(\alpha)) = (\limsup_{\alpha\in D} x(\alpha))\,(\lim_{\alpha\in D} y(\alpha))$$

again provided that the right-hand sides are meaningful.

1.8.7. *Some applications.* As a first application we give the true statement of the root test for the absolute convergence of a numerical series.

. Root test *Let $\sum_{n=0}^{\infty} a_n$ be a series of complex numbers, and let $\alpha = \limsup_{n\to\infty} |a_n|^{1/n}$. If $\alpha < 1$ then the series is absolutely convergent, if $\alpha > 1$ then the series does not converge.*

Proof. To say that $\limsup_{n\to\infty} |a_n|^{1/n} = \alpha < 1$ implies that if $\alpha < \rho < 1$ there is $m \in \mathbb{N}$ such that $|a_n|^{1/n} \le \rho$ for $n \ge m$. Then $|a_n| \le \rho^n$ for $n \ge m$, and the series converges since its $m-$tail is dominated by a geometric series of ratio $\rho < 1$. And if $\alpha > 1$, given ρ with $1 < \rho < \alpha$ we have $|a_n|^{1/n} > \rho \iff |a_n| > \rho^n$ for infinitely many $n \in \mathbb{N}$, so that $\limsup_{n\to\infty} |a_n| = +\infty$, and the series cannot converge, since its general term does not tend to 0.

$\square$

We also immediately get

. Cauchy–Hadamard criterion for the radius of convergence *Given a power series*

$$\sum_{n=0}^{\infty} a_n\, z^n$$

let $R = 1/\limsup_{n\to\infty} |a_n|^{1/n}$ (understanding that $R = 0$ if the denominator is ∞, $R = \infty$ if the denominator is 0). Then the power series is absolutely convergent if $|z| < R$, and is not convergent for $|z| > R$.

Proof. Set $\alpha = \limsup_{n\to\infty} |a_n|^{1/n}$. Then, assuming $z \neq 0$ we have $\limsup_{n\to\infty} |a_n\, z^n|^{1/n} = |z| \limsup_{n\to\infty} |a_n|^{1/n} = |z|\,\alpha$, so that by the root test the series converges if $|z|\,\alpha < 1$, does not converge if $|z|\,\alpha > 1$ $\square$

The quantity R so obtained is for obvious reasons called *radius of convergence* of the given power series.

Given a power series $\sum_{n=0}^{\infty} a_n\, z^n$, is *derived series* is the series $\sum_{n=1}^{\infty} n a_n\, z^{n-1}$. An immediate application of 1.8.6, (ii) is

. *A power series and its derived series have the same radius of convergence.*

Proof. Recall that $\lim_{n\to\infty} n^{1/n} = 1$ ($n^{1/n} = \exp(\log n/n)$ and $\lim_{n\to\infty} \log n/n = 0$). Then

$$\limsup_{n\to\infty} (n|a_n|)^{1/n} = \lim_{n\to\infty} n^{1/n} \limsup_{n\to\infty} |a_n|^{1/n} = \limsup_{n\to\infty} |a_n|^{1/n}.$$

$\square$

1.8.8. *Another application: a limit related to subadditive sequences.* A sequence $n \mapsto x(n)$ of real numbers is said to be *subadditive* if $x(m+n) \le x(m) + x(n)$ for every $m, n \in \mathbb{N}$. An interesting and useful result on subadditive sequences is the:

. Fekete lemma. *If $(x(n))_{n\in\mathbb{N}}$ is a subadditive sequence of real numbers then*

$$\lim_{n\to\infty} \frac{x(n)}{n} = \inf\left\{\frac{x(n)}{n} : n \ge 1\right\}$$

(finite or $-\infty$).

Proof. Plainly

$$\liminf_{n\to\infty} \frac{x(n)}{n} \ge \inf\left\{\frac{x(n)}{n} : n \ge 1\right\}.$$

If we prove that for every $m \ge 1$ we have $\limsup_{n\to\infty} x(n)/n \le x(m)/m$ we conclude that

$$\limsup_{n\to\infty} \frac{x(n)}{n} \le \inf\left\{\frac{x(m)}{m} : m \ge 1\right\} \le \liminf_{n\to\infty} \frac{x(n)}{n},$$

and our claim is proved. Observe first that by subadditivity we have $x(m\,n) = x(n + \cdots + n) \le x(n) + \cdots + x(n) = m\,x(n)$ for every $m,\, n \ge 1$. We can write $n = q(n)\,m + r(n)$, with $q(n) = [n/m]$ (integer part of n/m) and $r(n) \in \{0, \ldots, m-1\}$. By subadditivity we get

$$x(n) = x(q(n)\,m + r(n)) \le x(q(n)\,m) + x(r(n)) \le q(n)\,x(m) + x(r(n)) \Longrightarrow \frac{x(n)}{n} \le \frac{q(n)}{n}\,x(m) + \frac{x(r(n))}{n},$$

which implies

$$\frac{x(n)}{n} \le \frac{1}{m + r(n)/q(n)}\, x(m) + \frac{x(r(n))}{n} \le \frac{1}{1 + r(n)/(n - r(n))}\, \frac{x(m)}{m} + \frac{\mu(m)}{n},$$

where $\mu(m) = \max\{x(r) : r = 0, \dots, m-1\}$. As n tends to infinity, $n - r(n) \ge n - m + 1$ also tends to infinity and the rightmost hand side tends to $x(m)/m$; taking lim sup of both sides then gives:

$$\limsup_{n\to\infty} \frac{x(n)}{n} \le \limsup_{n\to\infty} \left(\frac{1}{1 + (r(n)/(n - r(n))}\, \frac{x(m)}{m} + \frac{\mu(m)}{n} \right) = \frac{x(m)}{m}.$$

□

Subadditivity is often checked with the following fact:

. *Let $q : [0, +\infty[\to \mathbb{R}$ be concave and such that $q(0) = 0$. Then*

$$q(x + y) \le q(x) + q(y) \qquad \textit{for every} \quad x, y \in [0, +\infty[.$$

Proof. Assume that $x, y > 0$. Then the line connecting $(0, q(0) = 0)$ with $((x + y), q(x + y))$ has equation $Y = (q(x + y)/(x + y))\, X$; by concavity we get

$$q(x) \ge \frac{q(x + y)}{x + y}\, x; \quad q(y) \ge \frac{q(x + y)}{x + y}\, y \Longrightarrow q(x) + q(y) \ge q(x + y).$$

□

But in this case the result of Fekete's lemma is immediate: $q(x)/x$ is decreasing, by concavity of q, hence

$$\lim_{x\to+\infty} \frac{q(x)}{x} = \inf \left\{ \frac{q(t)}{t} : t > 0 \right\}.$$

A function $q : [0, +\infty[\to \mathbb{R}$ may however be subadditive and zero at 0 without being concave, e.g. the function $q(x) = |\sin x|$: we have

$$q(x + y) = |\sin(x + y)| = |\sin x \,\cos y + \cos x \,\sin y| \le |\sin x \,\cos y| + |\cos x \,\sin y| =$$
$$|\sin x|\,|\cos y| + |\cos x|\,|\sin y| \le |\sin x| + |\sin y| = q(x) + q(y).$$

1.8.9. *Some proofs and more observations.*

Solution. Detailed solution of Exercise 1.8.1. Every nbhd of $+\infty$ contains a half line as $[a, +\infty[$, in particular points as $k\pi$ and $\pi/2 + k\pi$ for $k \in \mathbb{N}$ large. Then $0 \le f(x) < 1$ for $x \in [a, +\infty[$ (if $a > 0$), in particular $0 \le f_*([a, +\infty[) \le f^*([a, +\infty[) \le 1$. But $f(k\pi) = \tanh(k\pi)$ and $f(\pi/2 + k\pi) = 0$ so that

$$f_*([a, +\infty[) = \min f([a, +\infty[) = 0; \; f^*([a, +\infty[) = \sup f([a, +\infty[) \ge \sup\{f(k\pi) : k \ge a\} = 1.$$

□

Proof. (of 1.8.4). Recall that a subsequence of the sequence $(x_n)_{n\in\mathbb{N}}$ is a sequence $y_n = x_{\nu(n)}$, where $\mu : \mathbb{N} \to \mathbb{N}$ is a strictly increasing function. Given $m \in \mathbb{N}$ we then have

$$y_{*m} = \inf\{y_n : n \ge m\} = \inf\{x_{\nu(n)} : n \ge m\} \ge \inf\{x_k : k \ge \nu(m)\} = x_{*\nu(m)};$$

the inequality is due to the fact that $\{x_{\nu(n)} : n \ge m\} \subseteq \{x_k : k \ge \nu(m)\}$. Passing to the limit in the inequality $y_{*m} \ge x_{*\nu(m)}$ we get

$$\liminf_{n\to\infty} y_n = \lim_{m\to\infty} y_{*m} \ge \lim_{m\to\infty} x_{*\nu(m)} = \lim_{m\to\infty} x_{*m} = \liminf_{n\to\infty} x_n.$$

We have proved that every cluster point of a sequence is not smaller that the limit inferior of the sequence. It remains to prove that the limit inferior α is a cluster point, that is, the limit of a subsequence. We consider this obvious if $\alpha = -\infty$; we assume then $\alpha \in \mathbb{R}$ or $\alpha = +\infty$; we may skip the second case, since then the sequence has $+\infty$ as a limit, and hence as only cluster point. For $m \in \mathbb{N}$ let

$$A(m) = \{n \in \mathbb{N} : x_n < \alpha + 1/(m + 1)\};$$

clearly these sets are all infinite sets. Define now inductively $\nu : \mathbb{N} \to \mathbb{N}$ as $\nu(0) = \min A(0)$ and $\nu(m) = \min\{n \in A(m), n > \nu(m-1)\}$; then μ is strictly increasing and $x_{*\nu(m)} \le x_{\nu(m)} < \alpha + 1/(m+1)\}$. Similarly for limsup. □

Solution. (of 1.8.3) We get $s_2 = -1/2$; $s_3 = 1 + (-1/2) = 1/2$; observe that

$$s_{2n+2} = s_{2(n+1)} = s_{2n+1} - 1 - \frac{(-1)^{n+2}}{n+2} = s_{2n} + \frac{(-1)^{n+1}}{n+2}$$

so that

$$s_{2n+2} = -\frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{(-1)^{n+1}}{n+2}\,,$$

and

$$s_{2n+3} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{(-1)^{n+1}}{n+2}\,.$$

Recalling that $\log 2 = \lim_{m\to\infty} \sum_{n=1}^{m} (-1)^{n+1}/n$ we get

$$\liminf_{n\to\infty} s_n = \log 2 - 1; \qquad \limsup_{n\to\infty} s_n = \log 2\,. \quad \square$$

$\square$

Solution. (*of exercise 1.8.5*) Write $a_n = (a_n + b_n) - b_n$. Then

$$\limsup_{n\to\infty} a_n = \limsup_{n\to\infty}((a_n + b_n) - b_n) \le \limsup_{n\to\infty}(a_n + b_n) + \limsup_{n\to\infty}(-b_n) \le a + b - \liminf_{n\to\infty} b_n \le$$
$$\le a + b - b = a.$$

Then $\limsup_{n\to\infty} a_n \le a \le \liminf_{n\to\infty} a_n$, equivalently $\lim_{n\to\infty} a_n = a$. In the same way one sees that $\lim_{n\to\infty} b_n = b$. $\square$

1.9. Limsup and liminf of sequences of real valued functions. Sequences of real valued functions, or even of extended real valued functions have naturally pointwise suprema and infima, and hence also pointwise limsup and liminf: the notion presents no difficulty. Given a sequence $(f_n)_{n\in\mathbb{N}}$ of functions $f_n : X \to \tilde{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$, where X is any set, we consider for every $m \in \mathbb{N}$:

$$f_{*m} = \bigwedge_{n\ge m} f_n : X \to \tilde{\mathbb{R}} \quad \text{defined by} \quad f_{*m}(x) := \inf\{f_n(x) :\ n \ge m\},$$

and

$$f^*_m = \bigvee_{n\ge m} f_n : X \to \tilde{\mathbb{R}} \quad \text{defined by} \quad f^*_m(x) := \sup\{f_n(x) :\ n \ge m\},$$

the sequence (f_{*m} is increasing and f^*_m is decreasing; the limits f_* and f^* of these sequences are, by definition, the functions $\liminf_{n\to\infty} f_n$ and $\limsup_{n\to\infty} f_n$, so that we have, for every $x \in X$:

$$(\liminf_{n\to\infty} f_n)(x) = \liminf_{n\to\infty} f_n(x); \quad (\limsup_{n\to\infty} f_n)(x) = \limsup_{n\to\infty} f_n(x).$$

Exercise 1.9.1. Let $f_n : \mathbb{R} \to \mathbb{R}$ be given by $f_n(x) = (-1)^n\, |x|^{1/n}$. Find $\liminf_{n\to\infty} f_n$ and $\limsup_{n\to\infty} f_n$.

(We get $\liminf_{n\to\infty} f_n(x) = -1$ and $\limsup_{n\to\infty} f_n(x) = 1$ if $x \in \mathbb{R} \setminus \{0\}$; and $\liminf_{n\to\infty} f_n(0) = \limsup_{n\to\infty} f_n(0) = 0$).

1.9.1. *Limsup and liminf of sequences of sets.* Very important in measure theory is the case of liminf and limsup of sequences of subsets $(A_n)_{n\in\mathbb{N}}$ of a given set X: subsets of X are identified with their characteristic functions, and obviously both limsups and liminfs of characteristic functions are still caracteristic functions, in fact we have that $\bigwedge_{n\ge m} \chi_{A_n}$ is the characteristic function of the intersection $\bigcap_{n\ge m} A_n$, while $\bigvee_{n\ge m} \chi_{A_n}$ is the characteristic function of the union $\bigcup_{n\ge m} A_n$, so that

$$\liminf_{n\to\infty} A_n = \bigcup_{m=0}^{\infty} \left(\bigcap_{n\ge m} A_n \right); \quad \limsup_{n\to\infty} A_n = \bigcap_{m=0}^{\infty} \left(\bigcup_{n\ge m} A_n \right).$$

Notice that $\liminf_{n\to\infty} A_n$ may be characterized as the set of all $x \in X$ which are *eventually (it: definitivamente)* in the sequence, while $\limsup_{n\to\infty} A_n$ is the set of all $x \in X$ which are *frequently*, or *infinitely often*, in the sequence, that is, belong to A_n for infinitely many indices $n \in \mathbb{N}$.

Exercise 1.9.2. In $X = \mathbb{R}^2$ for every $n \in \mathbb{N}$ let $A_n = \mathrm{Epi}(p_n)$ be the epigraphic of the power function $p_n(x) = x^n$, that is $A_n = \{(x, y) :\ x \in \mathbb{R},\ y \ge x^n\}$. Describe $\liminf_{n\to\infty} A_n$ and $\limsup_{n\to\infty} A_n$.

(Results:
$$\liminf_{n\to\infty} A_n = (]-1,1[\times]0,\infty[) \cup (\{-1,1\}\times[1,\infty[) \cup \{(0,0)\}$$
and
$$\limsup_{n\to\infty} A_n =$$
$$(]-\infty,-1[\times\mathbb{R}) \cup (\{-1\}\times[-1,\infty[) \cup (]-1,0]\times[0,\infty[) \cup (]0,1[\times]0,\infty[) \cup (\{1\}\times[1,\infty[).)$$

EXERCISE 1.9.3. We say that a sequence $(A_n)_n$ of subsets of a set X has a limit if $\liminf_n A_n = \limsup A_n$ (this common value is the limit of the sequence). Prove that the sequence has a limit if and only if for each $x \in X$ there is $n(x) \in \mathbb{N}$ such that we either have $x \in A_n$ for every $n \geq n(x)$, or $x \notin A_n$ for every $n \geq n(x)$.

EXERCISE 1.9.4. For the sequence $x_n = n^2 \sin(n\pi/2)/(1+n^2)$ compute lim inf and lim sup.

EXERCISE 1.9.5. State the theorem on the behavior of $\liminf \varphi \circ f$ and $\limsup \varphi \circ f$ when f is post–composed with a monotone function φ.

For the sequence $x_n = \exp(3n/(n+1) + \cos(n\pi))$ compute lim inf and lim sup. Compute also $\limsup_{x\to+\infty} \exp(3x/(x+1) + \cos(\pi x))$.

EXERCISE 1.9.6. (i) Prove the

. RATIO TEST FOR THE CONVERGENCE OF A SERIES. *Let $(c_n)_{n\in\mathbb{N}}$ be a sequence of non–zero complex numbers. If $\limsup_{n\to\infty}(|c_{n+1}|/|c_n|) < 1$ then the series $\sum_{n=0}^{\infty} c_n$ is absolutely convergent. If $\liminf_{n\to\infty}(|c_{n+1}|/|c_n|) > 1$ then the series does not converge.*

(ii) Find a sequence $a_n > 0$ of strictly positive real numbers such that the series $\sum_{n=0}^{\infty} a_n$ is convergent, but $\limsup_{n\to\infty}(a_{n+1}/a_n) = +\infty$.

From now on $a = (a_n)_{n\in\mathbb{N}}$ is a sequence of strictly positive real numbers, $a_n > 0$; we consider the power series $\sum_{n=0}^{\infty} a_n z^n$; let $R = R_a$ be its radius of convergence.

(iii) Prove that $R = \liminf_{n\to\infty} a_n^{-1/n}$.

(iv) *(Requires some labor)* Prove that
$$\liminf_{n\to\infty} \frac{a_{n+1}}{a_n} \leq \liminf_{n\to\infty} a_n^{1/n} \leq \limsup_{n\to\infty} a_n^{1/n} \leq \limsup_{n\to\infty} \frac{a_{n+1}}{a_n},$$
and deduce from this that if $\lim_{n\to\infty} a_{n+1}/a_n$ exists in $\tilde{\mathbb{R}}$, then also $\lim_{n\to\infty} a_n^{1/n}$ exists and coincides with the previous limit.

(v) Compute $\lim_{n\to\infty}(n!)^{1/n}/n$.

Solution. (of Ex1.9.4) The function $x \mapsto x^2/(1+x^2)$ is increasing on $[0,+\infty[$(it has derivative $2x/(1+x^2)^2$) and has limit 1 as $x \to +\infty$; notice that $|x_n| < 1$ for every n. We have $\sin(n\pi/2) = 0$ if $n \in 2\mathbb{N}$ is even, $\sin(n\pi/2) = 1$ if $n = 4k+1$, $\sin(n\pi/2) = -1$ if $n = 4k+3$. For every $m \in \mathbb{N}$ there are numbers of the form $4k+1$ and of the form $4k+3$ larger than m, then the sequence has terms of the form $x_n = n^2/(1+n^2)$ and of the form $x_n = -n^2/(n^2+1)$ for infinitely many $n > m$; and no term is in absolute value larger than 1. It follows that for every $m \in \mathbb{N}$
$$\inf\{x_n : n \geq m\} = -1; \qquad \sup\{x_n : n \geq m\} = 1;$$
hence $\liminf_{n\to-\infty} x_n = -1$ and $\limsup_{n\to\infty} x_n = 1$.

Alternatively: every subsequence of the given sequence must have a subsequence in common with either the subsequence x_{2k} or with the subsequence x_{4k+1} or x_{4k+3}, since $\mathbb{N} = \{2k : k \in \mathbb{N}\} \cup \{4k+1 : k \in \mathbb{N}\} \cup \{4k+3 : k \in \mathbb{N}\}$. These subsequences converge to 0, 1 and -1 respectively. Then the cluster points of the sequence are $\{-1, 0, 1\}$ and the minimum is -1, the maximum is 1.

□

Solution. (of Ex 1.9.5) For the statement see 1.8.5. Since exp is continuous and strictly increasing we have
$$\liminf_{n\to\infty} \exp(3n/(n+1) + \cos(n\pi)) = \exp(\liminf_{n\to\infty}(3n/(n+1) + \cos(n\pi)));$$
$$\limsup_{n\to\infty} \exp(3n/(n+1) + \cos(n\pi)) = \exp(\limsup_{n\to\infty}(3n/(n+1) + \cos(n\pi)));$$
since $\cos(n\pi) = (-1)^n$ we have $\liminf_{n\to\infty} \cos(n\pi) = -1$ and $\limsup_{n\to\infty} \cos(n\pi) = 1$; and
$$\lim_{n\to\infty} (3n/(n+1) = 3,$$

so that:

$$\liminf_{n\to\infty}\left(\frac{3n}{n+1}+\cos(n\pi)\right)=3-1=2;\qquad \limsup_{n\to\infty}\left(\frac{3n}{n+1}+\cos(n\pi)\right)=3+1=4;$$

and finally

$$\liminf_{n\to\infty} x_n = e^2;\qquad \limsup_{n\to\infty} x_n = e^4.$$

A discussion identical to the preceding one proves that also:

$$\limsup_{x\to+\infty}\exp\left(\frac{3x}{x+1}+\cos(\pi x)\right)=e^4.$$

□

Solution. (of Ex 1.9.6). (i) Given ρ with $\limsup_{n\to\infty}(|c_{n+1}|/|c_n|)<\rho<1$ we have that there is $m\in\mathbb{N}$ such that $\sup\{|c_{n+1}|/|c_n| : n\geq m\}<\rho$, equivalently $|c_{n+1}|/|c_n|<\rho$ for $n\geq m$. Then we get $|c_{m+1}|<\rho\,|c_m|$, $|c_{m+2}|<\rho\,|c_{m+1}|<\rho^2\,|c_m|$, and by induction we easily see that $|c_n|<\rho^{n-m}\,|c_m|$ for every $n\geq m$. Then the series $\sum_{n=m}^\infty |c_n|$ is dominated by the series $\sum_{n=m}^\infty \rho^n\,(|c_m|/\rho^m)$, which converges since $0<\rho<1$, and the series $\sum_{n=0}^\infty c_n$ is then absolutely convergent. In the same way, if $\liminf_{n\to\infty}|c_{n+1}|/|c_n|>1$, given ρ with $1<\rho<\liminf_{n\to\infty}|c_{n+1}|/|c_n|$ we have that there exists $m\in\mathbb{N}$ such that $|c_{n+1}|/|c_n|>\rho$ for $n\geq m$, and this implies, in analogy to what above proved, that $|c_n|>\rho^{n-m}\,|c_m|$ for $n\geq m$; this implies clearly $\lim_{n\to\infty}|c_n|=+\infty$, forbidding the convergence of the series $\sum_{n=0}^\infty c_n$.

(ii) Simply take $a_n=1/4^n$ for n even, and $a_n=1/2^n$ for n odd. The series converges (we can also easily compute the sum) but $a_{2n+1}/a_{2n}=(1/2^{2n+1})/(1/4^{2n})=2^{4n}/2^{2n+1}=2^{2n-1}$ tends to $+\infty$.

Remark. The root test proves convergence of this series, in fact we have $\limsup_{n\to\infty} a_n^{1/n}=1/2<1$. This shows that the root test is more sensitive than the ratio test in determining the convergence of a series.

(iii) From the Cauchy–Hadamard criterion we know that $R=1/\limsup_{n\to\infty} a_n^{1/n}$; by the theorem on composition with decreasing maps, applied to the map $\varphi(x)=1/x$, decreasing homeomorphism of $[0,+\infty]$ onto itself, we get that $R=\varphi(\limsup_{n\to\infty} a_n^{1/n})=\liminf_{n\to\infty}\varphi(a_n^{1/n})=\liminf_{n\to\infty} a_n^{-1/n}$.

(iv) Let us first prove the leftmost inequality. If $\liminf_{n\to\infty} a_{n+1}/a_n=0$, there is nothing to prove. If not, we have that $\rho_m=\inf\{a_{n+1}/a_n : n\geq m\}>0$ for m large enough. As seen above in the proof of (i), second statement, this implies $a_n\geq\rho_m^{n-m}\,a_m$, for every $n\geq m$. Then $a_n^{1/n}\geq\rho_m(a_m/\rho_m^m)^{1/n}$; taking lim inf of both sides we get $\liminf_{n\to\infty} a_n^{1/n}\geq\rho_m$, for every $m\in\mathbb{N}$ (we have $\lim_{n\to\infty}(a_m/\rho_m^m)^{1/n}=1$). Since $\liminf_{n\to\infty} a_{n+1}/a_n=\sup\{\rho_m : m\in\mathbb{N}\}$ we conclude that $\liminf_{n\to\infty} a_n^{1/n}\geq\liminf_{n\to\infty} a_{n+1}/a_n$, as desired.

It remains to prove the rightmost inequality. This can be done with an argument as above, or as follows: let $\beta=\limsup_{n\to\infty} a_{n+1}/a_n$. For every $z\in\mathbb{C}$ we have

$$\limsup_{n\to\infty}|a_{n+1}\,z^{n+1}|/|a_n\,z^n|=|z|\,\limsup_{n\to\infty} a_{n+1}/a_n=|z|\,\beta;$$

then, if $|z|\,\beta<1$, equivalently if $|z|<1/\beta$, the series converges, by the ratio test, so that $|z|<1/\beta$ implies $|z|\leq R$. This of course implies $1/\beta\leq R=1/\limsup_{n\to\infty} a_n^{1/n}$, so that $\beta\geq\limsup_{n\to\infty} a_n^{1/n}$.

(v) Setting $a_n=n!/n^n$ we get

$$\frac{a_{n+1}}{a_n}=\frac{(n+1)!}{(n+1)^{n+1}}\,\frac{n^n}{n!}=\frac{n^n}{(n+1)^n}=\frac{1}{(1+1/n)^n},$$

so that $\lim_{n\to\infty} a_{n+1}/a_n=1/e$. It follows that $\lim_{n\to\infty}(n!)^{1/n}/n=1/e$.

Remark. We have obtained that the sequence $(n!)^{1/n}$ is asymptotic, as $n\to\infty$, to the sequence n/e, a fact often useful in estimating limits of sequences.

□

1.9.2. *de l'Hôpital's rule.* As a further important application of the notions of $\liminf$ and $\limsup$ we present the following result, which gives an immediate proof of the de l'Hôpital's rule:

. Let I be an interval of $\mathbb{R}$, and let $c \in \tilde{\mathbb{R}}$ be an accumulation point of I; assume that $f, g : I \setminus \{c\} \to \mathbb{R}$ are differentiable functions, with $g'(x) \neq 0$ for every $x \in I \setminus \{c\}$, and that either $\lim_{x\to c} f(x) = \lim_{x\to c} g(x) = 0$, or $\lim_{x\to c} |g(x)| = \infty$. Then

$$\liminf_{x\to c} \frac{f'(x)}{g'(x)} \le \liminf_{x\to c} \frac{f(x)}{g(x)} \le \limsup_{x\to c} \frac{f(x)}{g(x)} \le \limsup_{x\to c} \frac{f'(x)}{g'(x)}.$$

Proof. We assume that $c = \sup I$ so that all the limits are really for $x \to c^-$; an analogous proof will evidently work for $x \to c^+$. Also, we prove only the third inequality, the first is entirely analogous (or can be obtained from the third by considering $(-f)/g$ instead of f/g). If $\limsup_{x\to c} f'(x)/g'(x) = +\infty$, there is nothing to prove. Otherwise, given a real number $\beta > \limsup_{x\to c} f'(x)/g'(x)$ we prove that $\limsup_{x\to c} \dfrac{f(x)}{g(x)} \le \beta$, which clearly implies the desired inequality. We have a left punctured nbhd $[\gamma, c[$ of c in I such that $f'(x)/g'(x) \le \beta$ for every $x \in [\gamma, c[$. Given two points $x, u \in [\gamma, c[$, with $x \neq u$, the Cauchy theorem of "finite increments" says that

$$\frac{f(x)-f(u)}{g(x)-g(u)} = \frac{f'(\xi)}{g'(\xi)} \quad \text{for some } \xi \text{ in the open interval with extremes } x, u;$$

since clearly $\xi \in [\gamma, c[$ we have proved that

$$(*) \qquad \frac{f(x)-f(u)}{g(x)-g(u)} \le \beta \quad \text{for every } x, u \in [\gamma, c[,\ x \neq u.$$

Now, if $\lim_{u\to c} f(u) = \lim_{u\to c} g(u) = 0$ we keep $x \in [\gamma, c[$ fixed and take the limit in (*) for $u \to c$, obtaining $f(x)/g(x) \le \beta$ for every $x \in [\gamma, c[$, so that $\limsup_{x\to c} f(x)/g(x) \le \beta$, as desired. In the other case, of a diverging g, we write

$$\frac{f(x)-f(u)}{g(x)-g(u)} = \frac{f(x)}{g(x)-g(u)} - \frac{f(u)}{g(x)-g(u)} = \frac{f(x)}{g(x)}\,\frac{1}{1-g(u)/g(x)} - \frac{f(u)}{g(x)-g(u)},$$

so that from(*) we get

$$\frac{f(x)}{g(x)}\,\frac{1}{1-g(u)/g(x)} \le \beta + \frac{f(u)}{g(x)-g(u)} \qquad x, u \in [\gamma, c[,\ x \neq u.$$

We now keep $u \in [\gamma, c[$ fixed; since $\lim_{x\to c}(1 - g(u)/g(x)) = 1$, for x close enough to c, say $x \in [\gamma', c[$ with $\gamma' \ge \gamma$ we have $(1 - g(u)/g(x)) > 0$ so that we can multiply both sides of the preceding inequality by it obtaining

$$\frac{f(x)}{g(x)} \le \left(\beta + \frac{f(u)}{g(x)-g(u)}\right)\left(1 - \frac{g(u)}{g(x)}\right), \quad \text{for every } x \in [\gamma', c[;$$

taking $\limsup$ in both sides:

$$\limsup_{x\to c} \frac{f(x)}{g(x)} \le \limsup_{x\to c} \left(\beta + \frac{f(u)}{g(x)-g(u)}\right)\left(1 - \frac{g(u)}{g(x)}\right) = \lim_{x\to c} \left(\beta + \frac{f(u)}{g(x)-g(u)}\right)\left(1 - \frac{g(u)}{g(x)}\right) = \beta,$$

as desired. $\square$

Exercise 1.9.7. Using the above result prove that if $f\ :]0, \infty[\to \mathbb{R}$ is bounded and everywhere differentiable then

$$\liminf_{x\to\infty} f'(x) \le 0 \le \limsup_{x\to\infty} f'(x).$$

Deduce from it what in Italy is called *teorema dell'asintoto*: if f is bounded and $\lim_{x\to\infty} f'(x)$ exists in $\tilde{\mathbb{R}}$, then this limit is zero (consider $f(x)/x \dots$). Give an example of an f with $\lim_{x\to\infty} f(x) = 0$, but

$$\liminf_{x\to\infty} f'(x) < 0 < \limsup_{x\to\infty} f'(x),$$

and an example of an f with $\lim_{x\to\infty} f(x) = \infty$, but $\lim_{x\to\infty} f'(x) = 0$

Exercise 1.9.8. Compute $\limsup_{x\to\infty}(\cos x + \sin x)$, and observe that

$$\limsup_{x\to\infty}(\cos x + \sin x) < \limsup_{x\to\infty} \cos x + \limsup_{x\to\infty} \sin x.$$

Exercise 1.9.9. Lower and upper topology on $\tilde{\mathbb{R}}$ have been defined in 1.7.1. These topologies are both non–Hausdorff, so that uniqueness of limits for $\tilde{\mathbb{R}}-$valued nets in these topologies may fail, as in fact it does. Let $(D, \preceq)$ be a directed set and let $x : D \to \tilde{\mathbb{R}}$ be a generalized sequence.

(a) Prove that if $\ell \in \tilde{\mathbb{R}}$ is a limit of the net x, with respect to the upper (resp: lower) topology on the range space $\tilde{\mathbb{R}}$ then every $a > \ell$ (resp.: $a < \ell$) is also a limit.
(b) Prove that the set of limits of x in the upper topology is the closed interval $[\limsup_{\alpha\in D} x(\alpha), \infty]$, while the set of limits in the lower topology is $[-\infty, \liminf_{\alpha\in D} x(\alpha)]$.

1.9.3. *Upper and lower semicontinuous approximations or real valued functions.* The following result has some intrinsic interest and will be used when comparing Riemann and Lebesgue integration (4.6.2).

Proposition. *Let X be a topological space, and let $f : X \to \mathbb{R}$ be a bounded function. Then there exist functions f_*, $f^* : X \to \mathbb{R}$ such that f_* is lower semicontinuous, f^* is upper semicontinuous, $f_* \le f \le f^*$ and moreover:*

(i) *If $u : X \to \mathbb{R}$ is lower semicontinuous and $u \le f$, then $u \le f_*$;*
(ii) *If $v : X \to \mathbb{R}$ is upper semicontinuous and $v \ge f$, then $v \ge f^*$.*

(In other words, f_* is the largest lsc function smaller that f, and f^* is the smallest usc function larger that f)

(iii) *f is continuous at x if and only if $f_*(x) = f^*(x)$.*

.

Proof. We will see that

$$f_*(x) = \sup\{\inf f(U) : U \text{ a neighborhood of } x \text{ in } X\};$$
$$f^*(x) = \inf\{\sup f(U) : U \text{ a neighborhood of } x \text{ in } X\}.$$

We now prove (ii). That f^*, as defined, is usc, is immediate: if $a > f^*(x)$ there is a nbhd U of x, which we may assume open, such that $\sup f(U) < a$; it follows that $f^*(\xi) < a$ for every $\xi \in U$. And clearly if $v \ge f$ and v is usc, then $v(x) \ge f^*(x)$ for every $x \in X$: if for some $c \in X$ we have $v(c) < f^*(c)$, given $a \in \mathbb{R}$ with $v(c) < a < f^*(c)$ there is an open nbhd U of c such that $v(x) < a$ for every $x \in U$; since $f(x) \le v(x)$ for every $x \in X$ this implies $\sup f(U) \le a < f^*(c)$; hence $f^*(c) \le a < f^*(c)$, a contradiction. In the same way (i) can be proved; (iii) is trivial. □

Remark. When x is non isolated in X we have

$$f_*(x) = f(x) \wedge \liminf_{t\to x} f(t); \qquad f^*(x) = f(x) \vee \limsup_{t\to x} f(t),$$

as is easy to see. The function $f^* - f_*$ is the *oscillation function* of f.

Exercise 1.9.10. Let X be a topological space, and let A be a subset of X. If $f = \chi_A$ then $f_* = \chi_{\mathrm{int}(A)}$, and $f^* = \chi_{\mathrm{cl}(A)}$ (easy).

1.10. Infinite sums of positive functions. Let X be a set, and let $w : X \to [0, \infty]$ be a positive extended real-valued function. For every finite subset F of X the *sum of w over F* is defined: $\sum_\emptyset w = 0$ by definition, $\sum_F w = \sum_{x\in F} w(x)$ for F finite non–empty. The sum is defined over arbitrary subsets A of X by the formula:

$$\sum_A w = \sum_{x\in A} w(x) := \sup\left\{\sum_F w : \ F \subseteq A,\, F \text{ finite}\right\}.$$

Notice that this immediately implies that if $A \subseteq B \subseteq X$ then:

$$\sum_A w \le \sum_B w,$$

with equality iff $w(x) = 0$ for every $x \in B \setminus A$.

Remark. For every set X, the set $\Phi(X)$ of all finite subsets of X, partially ordered by inclusion, is a directed set (the union of two finite subsets is still finite), and given $w : X \to [0, \infty]$ we can define a net $s_w : \Phi(X) \to [0, \infty]$ by $s_w(F) = \sum_F w$, for every $F \in \Phi(X)$. Since $w(x) \ge 0$ for every $x \in X$, this net is increasing and $\sup\{s_w(F) : F \in \Phi(X)\}$ is exactly its limit in $\tilde{\mathbb{R}}$. So we have defined $\sum_X w$ as the limit of this net, the limit of the sums of w on finite subsets, as the set gets larger and larger ....

1.10.1. *Associativity.* Associativity of the sum in the commutative semigroup $([0,\infty],+)$ may be expressed by saying that if $A_1,\dots,A_m$ is a finite disjoint family of finite subsets of X, and $A=\bigcup_{k=1}^m A_k$ is its union, then

$$\sum_A w=\sum_{k=1}^m\left(\sum_{A_k} w\right).$$

This holds also for infinite sums:

. UNRESTRICTED ASSOCIATIVITY FOR SUMS OF POSITIVE FUNCTIONS. *If $w:X\to[0,\infty]$ is a positive function, $(A_\lambda)_{\lambda\in\Lambda}$ is a disjoint family of subsets of X and $A=\bigcup_{\lambda\in\Lambda}A_\lambda$ then:*

$$\sum_{x\in A}w(x)=\sum_{\lambda\in\Lambda}\left(\sum_{x\in A_\lambda}w(x)\right).$$

Proof. If $F\subseteq A$ is finite, then $F=\bigcup_{\lambda\in\Lambda}(F\cap A_\lambda)$, disjoint union, where each $F\cap A_\lambda$ is finite; and $F\cap A_\lambda$ is non–empty only for a finite subset M of Λ, so that actually $F=\bigcup_{\lambda\in M}(F\cap A_\lambda)$ a finite disjoint union of finite sets. By usual (finite) associativity we have

$$\sum_F w=\sum_{\lambda\in M}\left(\sum_{x\in F\cap A_\lambda}w(x)\right);$$

since $\sum_{F\cap A_\lambda}w\le\sum_{A_\lambda}w$ for every λ we have

$$\sum_F w=\sum_{\lambda\in M}\left(\sum_{F\cap A_\lambda}w\right)\le\sum_{\lambda\in M}\left(\sum_{A_\lambda}w\right)\le\sum_{\lambda\in\Lambda}\left(\sum_{A_\lambda}w\right);$$

we have proved that for every finite subset F of A we have

$$\sum_F w\le\sum_{\lambda\in\Lambda}\left(\sum_{A_\lambda}w\right)$$

equivalently

$$\sum_A w\le\sum_{\lambda\in\Lambda}\left(\sum_{A_\lambda}w\right).$$

Conversely, take a finite subset M of Λ and for each $\lambda\in M$ a finite subset $F_\lambda\subseteq A_\lambda$; by finite associativity we have, setting $F=\bigcup_{\lambda\in M}F_\lambda$:

$$\sum_{\lambda\in M}\left(\sum_{F_\lambda}w\right)=\sum_F w\le\sum_A w,$$

so that

$$\sum_{\lambda\in M}\left(\sum_{F_\lambda}w\right)\le\sum_A w;$$

Taking suprema as F_λ varies among the finite subsets of A_λ, and using 1.1.2, we get

$$\sum_{\lambda\in M}\left(\sum_{A_\lambda}w\right)\le\sum_A w$$

for every finite subset M of Λ; equivalently

$$\sum_{\lambda\in\Lambda}\left(\sum_{A_\lambda}w\right)\le\sum_A w.$$

□

1.10.2. *Isotony of infinite sums of positive numbers.*

. *If X ia a set, $v, w : X \to [0,\infty]$ are positive functions, and $v(x) \le w(x)$ for every $x \in X$, then*

$$\sum_A v \le \sum_A w \quad \text{for every } A \subseteq X.$$

Proof. For every finite subset F of A we have

$$\sum_F v \le \sum_F w,$$

then

$$\sum_F v \le \sum_F w \le \sum_A w,$$

and the inequality $\sum_F v \le \sum_A w$ for every finite subset F of A is equivalent to $\sum_A v \le \sum_A w$. □

1.10.3. *Finiteness of the sum.* The following elementary fact is frequently encountered:

Proposition. *Let $w : X \to [0,\infty[$ be a finite valued positive function. If $\sum_X w < \infty$, then w is different from 0 on a set which is at most countable (i.e. $\{w \ne 0\} := \{x \in X : w(x) > 0\}$ has cardinality not larger than $\aleph_0 = |\mathbb{N}|$).*

Proof. Let $s = \sum_X w$; then s is a positive real number. Given $\alpha > 0$, let

$$E(\alpha) = \{w \ge \alpha\} := \{x \in X : \, w(x) \ge \alpha\}.$$

We claim that $E(\alpha)$ is finite and that $|E(\alpha)| \le s/\alpha$. In fact, if $F \subseteq E(\alpha)$ is finite we have $\sum_F w \le s$, and also $\sum_{x\in F} w(x) \ge \sum_{x\in F} \alpha = \alpha\,|F|$, so that $|F| \le s/\alpha$ for every finite subset F contained in $E(\alpha)$. But then $E(\alpha)$ itself is finite and $|E(\alpha)| \le s/\alpha$. Since we have

$$\{w > 0\} = \bigcup_{n=1}^{\infty} \{w \ge 1/n\},$$

the set $\{w > 0\}$ is countable, as a countable union of finite sets. □

2. Measures and measurable spaces

2.1. Additive set functions. Areas, volumes, masses, electric charges are concepts which lead to the definition of additive set functions (a set function is a function whose domain has sets as elements). Given a set X and a subset $\mathcal{A} \subseteq \mathcal{P}(X)$, a function $\phi : \mathcal{A} \to \mathbb{R}$ is said to be additive if given $A, B \in \mathcal{A}$, with $A \cup B \in \mathcal{A}$ and A, B disjoint we have $\phi(A \cup B) = \phi(A) + \phi(B)$. We shall start the study of additive set functions with positive values (allowing also ∞ as a value); additive set functions of an arbitrary sign, and even complex or vector–valued, are also important, but shall be studied later.

2.1.1. *Positive set functions finitely and countably additive.*

Definition. Let $\mathcal{A}$ be a subset of $\mathcal{P}(X)$, with $\emptyset \in \mathcal{A}$, and let $\phi : \mathcal{A} \to [0, +\infty]$ be a positive function. We say that ϕ is *finitely additive* if $\phi(\emptyset) = 0$, and for every finite family $(A_1, \dots, A_m)$ of pairwise disjoint sets in $\mathcal{A}$ whose union $\bigcup_{n=1}^m A_n$ belongs to $\mathcal{A}$ we have

$$\phi\left(\bigcup_{n=1}^{m} A_n\right) = \sum_{n=1}^{m} \phi(A_n);$$

and ϕ is said to be *countably additive*, or also $\sigma-$additive, if for every sequence $(A_n)_{n\in\mathbb{N}}$ of pairwise disjoint elements of $\mathcal{A}$ whose union belongs to $\mathcal{A}$ we have

$$\phi\left(\bigcup_{n=0}^{\infty} A_n\right) = \sum_{n=0}^{\infty} \phi(A_n).$$

Countable additivity implies finite additivity (since $\emptyset \in \mathcal{A}$ the finite sequence $(A_1, \dots, A_m)$ can be extended to $(A_1, \dots, A_m, \emptyset, \emptyset, \dots)$)

2.1.2. *Algebras and $\sigma-$algebras of sets.* To be able to work comfortably with additivity, it is indispensable that the class $\mathcal{A} \subseteq \mathcal{P}(X)$ of subsets of X on which the additive function under examination is defined be closed under all elementary set theoretic operations: that is, if A, $B \in \mathcal{A}$ then $A \cup B$, $A \cap B$, the complement $X \setminus A$, etc. must all belong to $\mathcal{A}$: in other words $\mathcal{A}$ must be an *algebra* of subsets of X (see 1.4). And countable additivity can be exploited successfully if we have closure under countable union as well.

Definition. A $\sigma-$algebra, or *tribe*, of parts of X is an algebra which is also closed under countable union, that is, if $(A_n)_{n\in\mathbb{N}}$ is a sequence of elements of $\mathcal{A}$, then also $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

And a complementation closed subset of $\mathcal{P}(X)$ is a $\sigma-$algebra iff it is closed under countable intersection, always by De Morgan's formulae.

Example 2.1.1. For any set X there is a largest algebra of parts of X, namely $\mathcal{P}(X)$ itself, and a smallest algebra, namely $\{\emptyset,\, X\}$; these are also $\sigma-$algebras.

Example 2.1.2. Given any set X, the subset $\mathcal{A}$ of $\mathcal{P}(X)$ consisting of all finite and cofinite subsets of X is an algebra, which is not a tribe unless X is finite. And the subset $\mathcal{B}$ of all countable or co–countable parts of X is a tribe, which coincides with $\mathcal{P}(X)$ iff X itself is countable.

Example 2.1.3. An algebra $\mathcal{A}$ of parts of X is finite when it has a finite set of elements. Of course a finite algebra is also a $\sigma-$algebra. In 1.4.1 we proved that

. Finite boolean algebras *Let $\mathcal{A}$ be a finite subalgebra of $\mathcal{P}(X)$. Then there is a finite partition $\{A_1, \dots A_m\}$ of X, with $A_k \in \mathcal{A}$ for every $k = 1, \dots, m$, such that every $A \in \mathcal{A}$ is representable, in a unique way, as*

$$A = \bigcup_{k\in S} A_k \qquad \textit{with } S \textit{ a subset of } \{1, \dots, m\}.$$

This partition is the natural basis *of the algebra $\mathcal{A}$.*

2.1.3. *Finitely additive measures, premeasures, measures.*

Definition. Let $\mathcal{A}$ be an algebra of parts of X, and let $\mu : \mathcal{A} \to [0, +\infty]$ be a positive extended real valued function. We say that μ is a *finitely additive measure* if μ is finitely additive. We say that μ is a *premeasure* if it is also *countably additive*. A *measure* is a premeasure whose domain $\mathcal{A}$ is a $\sigma-$algebra.

Recall that by definition of additivity $\mu(\emptyset) = 0$. We observed in 2.1.1 that countable additivity is stronger than finite additivity: that is, every premeasure is also finitely additive; we note here that it is strictly stronger: on an infinite set X let $\mathcal{A}$ be the algebra of finite or cofinite parts of X (see 2.1.2);

define $\mu : \mathcal{A} \to [0, +\infty[$ by $\mu(A) = 0$ if A is finite, and $\mu(A) = 1$ if A is cofinite. It is easy to check that μ is finitely additive (two cofinite subsets of $\mathbb{N}$ are never disjoint), but if X is countable then μ is not countably additive: $\mu(X) = 1 \neq \sum_{x\in X} \mu(\{x\}) = \sum_{x\in X} 0 = 0$, so that μ is not countably additive. However if X is uncountable then μ is a premeasure, i.e. countably additive, as it follows from the next exercise.

EXERCISE 2.1.4. Let X be an uncountable set, and let $\mathcal{A}$ be the $\sigma-$algebra of countable or co–countable subsets of X (2.1.2). Define $\mu : \mathcal{A} \to [0, +\infty[$ by $\mu(A) = 0$ if A is countable, $\mu(A) = 1$ if A is co–countable. Prove that μ is a measure (called the co–countable measure).

Solution. Let $(A_0, A_1, \dots)$ be a disjoint sequence of elements of $\mathcal{A}$. Let $A = \bigcup_{n=0}^{\infty} A_n$; if A is countable, then all A_n are countable, and $\mu(A) = \mu(A_n) = 0$ for every n. If A is cocountable then one of the A_n's must be cocountable, since a countable union of countable sets is countable, and in an uncountable set no countable subset is also cocountable. Say that A_0 is cocountable: then A_n is countable for every $n \neq 0$, being contained in $X \setminus A_0$. Then $1 = \mu(A) = \mu(A_0) + 0 + 0 + 0 + \dots$ holds true. □

The simplest example of a non entirely trivial measure is the following: given a set X and a positive function $w : X \to [0, +\infty]$ we define $\mu = \mu_w : \mathcal{P}(X) \to [0, +\infty]$ by setting

$$\mu(A) = \sum_A w \left(:= \sup \left\{ \sum_{x\in F} w(x) : F \subseteq A,\ F \text{ finite} \right\} \right) \text{ (finite or } +\infty).$$

Countable additivity has been proved in 1.10.1, in a stronger form, for any family of disjoint sets, of arbitrary cardinality. We can think of X as the universe, and at every point of X there is a mass $w(x)$; when $w(x) > 0$ we can think that $\{x\}$ is an *atom*; $\mu(A)$ is the sum of the masses of all atoms contained in A. When $w(x) = 1$ for every $x \in X$ we get the *counting measure* on X: the measure of a subset is its cardinality when the subset is finite, and is otherwise $+\infty$. When w is the characteristic function of a singleton $\{c\} \subseteq X$ we get the so called *Dirac measure* at c, or *unit mass* at c, often denoted δ_c: we have only one atom, $\{c\}$, so $\delta_c(A) = 1$ if $c \in A$, $\delta_c(A) = 0$ otherwise. Construction of non-atomic measures, like Lebesgue measure, is a non trivial matter that requires some effort.

2.1.4. *First properties of measures.*

. *Let $\mathcal{A}$ be an algebra of subsets of X, and let $\mu : \mathcal{A} \to [0, +\infty]$ be a finitely additive positive measure. Then*

(i) *μ is monotone, that is, if $A \subseteq B$ and $A, B \in \mathcal{A}$, then $\mu(A) \leq \mu(B)$.*
(ii) *μ is finitely subtractive, that is, if $A \subseteq B$, $A, B \in \mathcal{A}$, and $\mu(A) < \infty$, then $\mu(B \setminus A) = \mu(B) - \mu(A)$.*

Proof. (i) We have $B = A \cup (B \setminus A)$, disjoint union, so that $\mu(B) = \mu(A) + \mu(B \setminus A)$ by finite additivity; since $\mu(B \setminus A) \geq 0$, we have $\mu(B) \geq \mu(A)$.

(ii) Since $\mu(A)$ is finite, we can add $-\mu(A)$ to both sides of the equality $\mu(B) = \mu(A) + \mu(B \setminus A)$ obtaining $\mu(B) - \mu(A) = \mu(B \setminus A)$; if $\mu(B)$ is finite this is an equality in $\mathbb{R}$, otherwise both sides are ∞. □

When $A \subseteq B$ and $\mu(A) = \infty$, then of course by monotonicity also $\mu(B) = \infty$, and $\mu(B \setminus A)$ can be anything, finite or infinite: e.g take the counting measure on $\mathbb{N}$, $B = \mathbb{N}$ and $A = \{n \in \mathbb{N} : n \geq m\}$; we have $B \setminus A = \{0, \dots, m-1\}$ so that $\mu(B \setminus A) = m$; or take $B = \mathbb{N}$ and $A = 2\mathbb{N}$, subset of even numbers: then $\mu(B \setminus A) = \infty$.

2.1.5. *Countable subadditivity.* Given a positive set function $\mu : \mathcal{A} \to [0, +\infty]$, where $\mathcal{A}$ is an algebra of subsets of X, we say that μ is *countably subadditive* if whenever $(A_n)_{n\in\mathbb{N}}$ is a sequence of elements of $\mathcal{A}$ whose union $\bigcup_{n\in\mathbb{N}} A_n$ belongs to $\mathcal{A}$ we have

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leq \sum_{n=0}^{\infty} \mu(A_n).$$

We have the following:

PROPOSITION. *Let $\mu : \mathcal{A} \to [0, +\infty]$ be a positive finitely additive measure, where $\mathcal{A}$ is an algebra of subsets of X. Then μ is countably additive if and only if it is countably subadditive.*

Proof. There is a standard trick, worth remembering, to write a countable union of elements of an algebra as a disjoint countable union of elements of the same algebra. Given the sequence A_0, $A_1, \dots$ of elements of $\mathcal{A}$ we set

$$B_0 = A_0;\ B_1 = A_1 \setminus A_0, \dots, \text{ in general } \quad B_n = A_n \setminus \left(\bigcup_{k=0}^{n-1} A_k\right);$$

it is immediate to prove that the B_n's are pairwise disjoint and that, for every $m \in \mathbb{N}$:

$$\bigcup_{n=0}^{m} A_n = \bigcup_{n=0}^{m} B_n \quad \text{so that also} \quad \bigcup_{n=0}^{\infty} A_n = \bigcup_{n=0}^{\infty} B_n,$$

moreover $B_n \subseteq A_n$ for every $n \in \mathbb{N}$. This immediately shows how countable subadditivity follows from countable additivity, for positive measures:

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \mu\left(\bigcup_{n\in\mathbb{N}} B_n\right) = \sum_{n=0}^{\infty} \mu(B_n);$$

and since $B_n \subseteq A_n$ for every n, we have $\mu(B_n) \le \mu(A_n)$ for every n, so that the last sum is dominated by $\sum_{n=0}^{\infty} \mu(A_n)$.

And it is easy to show that finite additivity and countable subadditivity together imply countable additivity; if $A = \bigcup_{n\in\mathbb{N}} A_n$, disjoint union, then for every m we get

$$\sum_{n=0}^{m} \mu(A_n) = \mu\left(\bigcup_{n=0}^{m} A_n\right) \le \mu(A),$$

the last inequality due to monotonicity. Then, letting m tend to infinity we get

$$\sum_{n=0}^{\infty} \mu(A_n) \le \mu(A),$$

and the reverse inequality is provided by countable subadditivity. □

Exercise 2.1.5. Define finite subadditivity and prove that a positive finitely additive measure is finitely subadditive.

Exercise 2.1.6. Let $\mathcal{A}$ be an algebra of parts of X and let $\mu : \mathcal{A} \to [0, \infty]$ be a premeasure. We say that two sets $A, B \in \mathcal{A}$ are *almost disjoint* if $\mu(A \cap B) = 0$. Prove that if $(A_k)_{k\in\mathbb{N}}$ is a sequence of pairwise almost disjoint sets in $\mathcal{A}$, with union $A = \bigcup_{k\in\mathbb{N}} A_k \in \mathcal{A}$ then

$$\mu(A) = \sum_{k=0}^{\infty} \mu(A_k).$$

Solution. Let's apply the usual trick for making a disjoint union, $B_k = A_k \setminus \left(\bigcup_{j=0}^{k-1} A_j\right)$. We have $B_k \subseteq A_k$, and if the A_k's are pairwise almost disjoint then $\mu(B_k) = \mu(A_k)$: in fact $A_k \setminus B_k = A_k \cap \left(\bigcup_{j=0}^{k-1} A_j\right) = \bigcup_{j=0}^{k-1} A_k \cap A_j$ is a finite union of sets of measure zero, and has then measure zero. Then, since $A = \bigcup_{k\in\mathbb{N}} A_k = \bigcup_{k\in\mathbb{N}} B_k$ we have

$$\mu(A) = \sum_{k=0}^{\infty} \mu(B_k) = \sum_{k=0}^{\infty} \mu(A_k)$$

(the first equality due to the fact that the B_k's are pairwise disjoint), as desired. □

2.1.6. *Continuity for increasing sequences.*

. Continuity from below *Let $\mu : \mathcal{A} \to [0, +\infty]$ be a positive finitely additive measure, where $\mathcal{A}$ is an algebra of subsets of X. Then μ is countably additive if and only if for every increasing sequence $A_0 \subseteq A_1 \subseteq \dots$ of sets in $\mathcal{A}$ whose union $A = \bigcup_{n\in\mathbb{N}} A_n$ belongs to $\mathcal{A}$ we have $\mu(A) = \lim_{n\to\infty} \mu(A_n)(= \sup\{\mu(A_n) : n \in \mathbb{N}\})$.*

Proof. Assume μ countably additive and let A_n be as in the statement. If $B_n = A_n \setminus A_{n-1}$ for $n \geq 1$, and $B_0 = A_0$ we have that A is the disjoint union of the B_n's, so that

$$\mu(A) = \sum_{n=0}^{\infty} \mu(B_n) = \lim_{m\to\infty} \left(\sum_{n=0}^{m} \mu(B_n)\right) = \lim_{m\to\infty} \mu\left(\bigcup_{n=0}^{m} B_n\right) = \lim_{m\to\infty} \mu(A_m).$$

We have proved that countable additivity implies continuity for increasing sequences. We leave to the reader the easy converse, that finite additivity and continuity for increasing sequences together imply countable additivity. □

2.1.7. *Decreasing sequences.* What about decreasing sequences? In general the analogous statement is false: given $\mathbb{N}$ with the counting measure, and $A_n = \{k \in \mathbb{N} : k \geq n\}$ we have that $A_0 \supseteq A_1 \supseteq \dots$, and $\mu(A_n) = \infty$ for each $n \in \mathbb{N}$, so that $\lim_{n\to\infty} \mu(A_n) = \infty$, but $\bigcap_{n=0}^{\infty} A_n = \emptyset$, and $\mu(\emptyset) = 0 \neq \infty$. However:

. Continuity from above on sets of finite measure *Let $\mu : \mathcal{A} \to [0, +\infty]$ be a positive countably additive premeasure, where $\mathcal{A}$ is an algebra of subsets of X. Then for every decreasing sequence $A_0 \supseteq A_1 \supseteq \dots$ of elements of $\mathcal{A}$ whose intersection $A = \bigcap_{n\in\mathbb{N}} A_n$ belongs to $\mathcal{A}$, and is such that $\mu(A_m) < \infty$ for some $m \in \mathbb{N}$, we have $\mu(A) = \lim_{n\to\infty} \mu(A_n)(= \inf\{\mu(A_n) : n \in \mathbb{N}\})$.*

Proof. By discarding the first terms and reindexing the sequence if necessary we may assume that $\mu(A_0)$ is finite, so that all $\mu(A_n)$ and $\mu(A)$ are finite; and we have that $A_0 \setminus A_n$ is an increasing sequence in $\mathcal{A}$ whose union is $A_0 \setminus A \in \mathcal{A}$; by 2.1.6 we have $\mu(A_0 \setminus A) = \lim_{n\to\infty} \mu(A_0 \setminus A_n)$; but by finite subtractivity:

$$\mu(A_0) - \mu(A) = \mu(A_0 \setminus A) = \lim_{n\to\infty} (\mu(A_0 \setminus A_n)) = \lim_{n\to\infty} (\mu(A_0) - \mu(A_n)) = \mu(A_0) - \lim_{n\to\infty} \mu(A_n),$$

which clearly implies $\mu(A) = \lim_{n\to\infty} \mu(A_n)$. □

Exercise 2.1.7. Let $\mu : \mathcal{A} \to [0, +\infty]$ be a positive finitely additive measure.

(i) Assume that for every decreasing sequence $A_0 \supseteq A_1 \supseteq \dots$ of elements of $\mathcal{A}$ with empty intersection we have $\lim_{n\to\infty} \mu(A_n) = 0$. Then μ is countably additive.

(ii) Assume that μ verifies the thesis of the above proposition, that is: for every decreasing sequence $A_0 \supseteq A_1 \supseteq \dots$ whose intersection $A = \bigcap_{n\in\mathbb{N}} A_n$ belongs to $\mathcal{A}$, and is such that $\mu(A_m) < \infty$ for some $m \in \mathbb{N}$, we have $\mu(A) = \lim_{n\to\infty} \mu(A_n)(= \inf\{\mu(A_n) : n \in \mathbb{N}\})$. Prove that $\mu_{|\mathcal{F}(\mu)}$ is countably additive; here $\mathcal{F}(\mu) = \{A \in \mathcal{A} : \mu(A) < \infty\}$ is the ideal of sets of finite measure.

The next problem requires the notions of lim inf and lim sup for a a sequence $(A_n)_{n\in\mathbb{N}}$ of subsets of a set X given in 1.9; recall that from the definitions we have

$$\liminf_{n\to\infty} A_n = \bigcup_{m=0}^{\infty} \left(\bigcap_{n=m}^{\infty} A_n\right); \qquad \limsup_{n\to\infty} A_n = \bigcap_{m=0}^{\infty} \left(\bigcup_{n=m}^{\infty} A_n\right).$$

Exercise 2.1.8. Let $\mu : \mathcal{A} \to [0, +\infty]$ be a measure (that is, $\mu(\emptyset) = 0$, μ is countably additive, and $\mathcal{A}$ is a $\sigma-$algebra) and let $(A_n)_{n\in\mathbb{N}}$ be a sequence of sets in $\mathcal{A}$.

(i) Prove that $\mu(\liminf_{n\to\infty} A_n) \leq \liminf_{n\to\infty} \mu(A_n)$.

(ii) Assume that for some $m \in \mathbb{N}$ we have $\mu\left(\bigcup_{n=m}^{\infty} A_n\right) < \infty$. Then

$$\limsup_{n\to\infty} \mu(A_n) \leq \mu(\limsup_{n\to\infty} A_n).$$

(iii) Prove that if $\sum_{n=0}^{\infty} \mu(A_n) < \infty$, then $\mu(\limsup_{n\to\infty} A_n) = 0$.

Solution. (i) For simplicity let $B_m = \bigcap_{n=m}^{\infty} A_n$; for every m, and every $n \geq m$ we have $B_m \subseteq A_n$ so that, by monotonicity, $\mu(B_m) \leq \mu(A_n)$ for every $n \geq m$ and hence

$$(*) \qquad \mu(B_m) \leq \inf\{\mu(A_n) : n \geq m\};$$

now B_m is an increasing sequence of sets of $\mathcal{A}$ whose union is $\liminf_{n\to\infty} A_n$; and $a_m = \inf\{\mu(A_n) : n \geq m\}$ is an increasing sequence in $\tilde{\mathbb{R}}$ which by definition has $\liminf_{n\to\infty} \mu(A_n)$ as limit; so we get the required inequality simply by passing to the limit in (*).

(ii) Repeat, mutatis mutandis, the proof of (i), this time using continuity from above instead of continuity from below.

(iii) By countable subadditivity we get $\mu\left(\bigcup_{n=m}^{\infty} A_n\right) \leq \sum_{n=m}^{\infty} \mu(A_n)$; and if the series converges then $\lim_{m\to\infty} \left(\sum_{n=m}^{\infty} \mu(A_n)\right) = 0$. □

Exercise 2.1.9. Let $\mathcal{A}$ be an algebra of subsets of X. Then $\mathcal{A}$ is also a $\sigma-$algebra if and only if

(i) $\mathcal{A}$ is closed under countable disjoint union, that is, for every disjoint sequence $(A_n)_{n\in\mathbb{N}}$ of elements of $\mathcal{A}$ the union $\bigcup_{n=0}^{\infty} A_n$ belongs to $\mathcal{A}$.
(ii) $\mathcal{A}$ is closed under countable increasing union, that is, for every increasing sequence $A_0 \subseteq A_1 \subseteq \dots$ of elements of $\mathcal{A}$ the union $\bigcup_{n=0}^{\infty} A_n$ belongs to $\mathcal{A}$.
(iii) $\mathcal{A}$ is closed under countable decreasing intersection, that is, for every decreasing sequence $A_0 \supseteq A_1 \supseteq \dots$ of elements of $\mathcal{A}$ the intersection $\bigcap_{n=0}^{\infty} A_n$ belongs to $\mathcal{A}$.

2.1.8. *Operations and order on premeasures.* If $\mu, \nu : \mathcal{A} \to [0, +\infty]$ are premeasures on the same algebra of parts of subsets of X, then $\mu + \nu : \mathcal{A} \to [0, +\infty]$ defined by $(\mu + \nu)(A) = \mu(A) + \nu(A)$ is also a premeasure; and if $\alpha > 0$ then $\alpha\,\mu : \mathcal{A} \to [0, +\infty]$ defined by $(\alpha\,\mu)(A) = \alpha\,\mu(A)$ is a premeasure. Also, if $(\mu_\lambda)_{\lambda\in\Lambda}$ is any family of premeasures, then $\mu = \sum_{\lambda\in\Lambda} \mu_\lambda$, the function $\mu : \mathcal{A} \to [0, \infty]$ defined by:

$$\mu(A) := \sum_{\lambda\in\Lambda} \mu_\lambda(A) \left(:= \sup \left\{ \sum_{\lambda\in F} \mu_\lambda(A);\ F \subseteq \Lambda,\ F \,\text{finite} \right\} \right),$$

is a premeasure on $\mathcal{A}$ (countable additivity is a consequence of unrestricted associativity of sums of positive numbers). Given a function $w : X \to [0, \infty]$ the measure μ_w defined in 2.1.3 is often denoted $\sum_{x\in X} w(x)\,\delta_x$, an infinite positive combination of Dirac measures. There is also a natural partial order on premeasures: $\mu \le \nu$ means that $\mu(A) \le \nu(A)$ for every $A \in \mathcal{A}$.

Exercise 2.1.10. Assume that $\mu_0 \le \mu_1 \le \dots$ is an increasing sequence of premeasures on an algebra $\mathcal{A}$. Prove that then $\mu(A) = \lim_{n\to\infty} \mu_n(A)\,(= \sup\{\mu_n(A);\ n \in \mathbb{N}\})$ defines a premeasure μ on $\mathcal{A}$.

2.2. Radon–Stjelties premeasures on $\mathbb{R}$ and $\mathbb{R}^n$. There is a quite natural correspondence between positive premeasures on the algebra $\mathcal{A}$ of plurintervals of $\mathbb{R}$ which are finite on compact intervals (called here *Radon–Stieltjes premeasures*; clearly they are finite also on bounded intervals) and increasing functions $f : \mathbb{R} \to \mathbb{R}$; this correspondence is also a bijection if we conventionally restrict the admissible class of increasing functions. We now describe this correspondence. For basic facts concerning increasing functions see 1.3.5.

2.2.1. *Premeasures to functions.* Let $\mu : \mathcal{A} \to [0, +\infty]$ be a Radon–Stieltjes premeasure on the interval algebra $\mathcal{A}$ of $\mathbb{R}$. Define $f = f_\mu : \mathbb{R} \to \mathbb{R}$ by setting

$$f(x) = \mu(]0, x]) \quad \text{for } x \ge 0; \quad f(x) = -\mu(]x, 0]) \quad \text{for } x < 0.$$

It is immediate to check that f is increasing, and $f(0) = 0$. Moreover, f is right–continuous: if $x \ge 0$, and $x_j \downarrow x$, then $]0, x] = \bigcap_{j=0}^{\infty}]0, x_j]$ and all $]0, x_j]$ have finite measure, so that $\mu(]0, x]) = \lim_{j\to\infty} \mu(]0, x_j])$ by continuity from above of premeasures; and if $x < 0$ and $x_j \downarrow x$, then $]x, 0] = \bigcup_{j=0}^{\infty}]x_j, 0]$ and $\mu(]x, 0]) = \lim_{j\to\infty} \mu(]x_j, 0])$ by continuity from below of premeasures. We leave it to the reader to check that for $a, b \in \mathbb{R}$

$$\mu(]a, b[) = f(b^-) - f(a^+);\ \mu([a, b]) = f(b^+) - f(a^-);\ \mu([a, b[) = f(b^-) - f(a^-), \dots$$

etc; in particular for every $c \in \mathbb{R}$ we have, since $\{c\} = [c, c]$:

$$\mu(\{c\}) = f(c^+) - f(c^-) = \sigma_f(c) \quad \text{the jump of } f \text{ at } c,$$

so that f is continuous at a point c if and only if the singleton $\{c\}$ has measure 0.

The function f is the *distribution function (it. funzione di ripartizione)* of the premeasure μ, with *origin* 0.

Remark. When the premeasure μ is finite, e.g. in probability theory, another distribution function is preferred, the one with origin $-\infty$:

$$F(x) = F_\mu(x) = \mu(] - \infty, x]), \quad \text{for every } x \in \mathbb{R}.$$

Clearly F differs from f by an additive constant: we have $f(x) = F(x) - F(0) = F(x) - \mu(] - \infty, 0])$ (prove it), so that F has the same properties as f, in particular it is right continuous, but now $F(0)$ is not necessarily 0, and $F(-\infty) := \lim_{x\to-\infty} F(x) = 0$ (why?).

Exercise 2.2.1. Consider the premeasure $\mu = (1/2)\delta_{-1} + (1/3)\delta_0 + (1/\pi)\delta_{1/2} + \delta_1$; plot the graphs of F_μ and f_μ.

Exercise 2.2.2. Prove that there is an increasing right continuous function $F : \mathbb{R} \to \mathbb{R}$ such that $F(-\infty) = 0$, $F(\infty) = 1$, and the set of discontinuities of F is exactly the set $\mathbb{Q}$ of all rationals. Prove that every such function is necessarily strictly increasing.

Solution. We prove the last assertion first. Given $x < y$, there is a rational $r \in]x, y[$, hence

$$F(x) \le F(r^-) < F(r^+) = F(r) \le F(y), \quad \text{so that} \quad F(x) < F(y).$$

Choose a bijective indexing $n \mapsto r_n$ of the rationals. Define $w : \mathbb{R} \to [0, +\infty[$ by $w(x) = 0$ if $x \in \mathbb{R} \smallsetminus \mathbb{Q}$, and $w(x) = 1/2^{n+1}$ if $x = r_n \in \mathbb{Q}$; w gives a measure $\mu = \mu_w$ on $\mathcal{P}(\mathbb{R})$ by the formula $\mu(E) = \sum_{x\in E} w(x)$; clearly this measure is finite ($\mu(\mathbb{R}) = 1$), in particular it is a Radon measure. Define $F : \mathbb{R} \to \mathbb{R}$ by $F(x) = \mu(]-\infty, x]) = \sum_{t\le x} w(t)$ (as in the Remark above). Then F is right-continuous and $F(x^-) = \mu(]-\infty, x[) = \mu(]-\infty, x] - \{x\}) = F(x) - \mu(\{x\}) = F(x) - w(x)$, so that F is discontinuous at x iff $w(x) > 0$, i.e. iff $x \in \mathbb{Q}$; the jump at x is exactly $w(x) = \mu(\{x\})$. And $F(\infty) = \sum_{t\in\mathbb{R}} w(t) = \sum_{n=0}^{\infty} 1/2^{n+1} = 1$. □

2.2.2. *Functions to premeasures.* Assume now that we have an increasing function $f : \mathbb{R} \to \mathbb{R}$. Inspired by the above construction we define a set function $\lambda = \lambda_f : \mathcal{I} \to [0, +\infty]$ on the semialgebra $\mathcal{I}$ of all intervals of $\mathbb{R}$ by setting

$$\lambda(]a, b[) = f(b^-) - f(a^+);\ \lambda([a, b]) = f(b^+) - f(a^-);\ \lambda([a, b[) = f(b^-) - f(a^-), \dots$$

etc. (by definition $f(+\infty) = \sup f(\mathbb{R})$ and $f(-\infty) = \inf f(\mathbb{R})$). It can be proved that this set function is finitely additive on $\mathcal{I}$ (try to prove it; the proof is in 2.2.4), so that we can extend λ to a finitely additive measure on the interval algebra $\mathcal{A}$. We prove that λ is also countably additive by proving that it is countably subadditive (recall 2.1.5).

We need first the following fact, whose proof is easy (a trivial application of left and right limits for f) and left as an exercise:

. *For every interval I of $\mathbb{R}$ we have*

$$\lambda(I) = \sup\{\lambda(K) : K \subseteq I,\ K \text{ a compact interval}\}$$
$$\lambda(I) = \inf\{\lambda(J) : I \subseteq J,\ J \text{ an open interval}\}.$$

Finally we prove

Proposition. *With the above notations and terminology, if I and I_n are intervals, and $I \subseteq \bigcup_{n=1}^{\infty} I_n$, then $\lambda(I) \le \sum_{n=1}^{\infty} \lambda(I_n)$.*

Proof. If $\sum_{n=1}^{\infty} \lambda(I_n) = \infty$ there is nothing to prove. If not, then $\lambda(I_n) < \infty$ for every $n \ge 1$; given $\varepsilon > 0$ pick an open interval J_n containing I_n such that $\lambda(J_n) \le \lambda(I_n) + \varepsilon/2^n$. If K is a compact interval contained in I, then $K \subseteq I \subseteq \bigcup_{n=1}^{\infty} I_n \subseteq \bigcup_{n=1}^{\infty} J_n$. Then $\{J_n : n \ge 1\}$ is an open cover of the compact interval K, which must have a finite subcover. In other words there is $m \ge 1$ such that $K \subseteq \bigcup_{n=1}^{m} J_n$, so that

$$\lambda(K) \le \lambda\left(\bigcup_{n=1}^{m} J_n\right)$$

By finite subadditivity we get

$$\lambda\left(\bigcup_{n=1}^{m} J_n\right) \le \sum_{n=1}^{m} \lambda(J_n) \le \sum_{n=1}^{m} (\lambda(I_n) + \varepsilon/2^n) \le \sum_{n=1}^{\infty} \lambda(I_n) + \varepsilon,$$

so that, for every compact interval K contained in I and every $\varepsilon > 0$:

$$\lambda(K) \le \sum_{n=1}^{\infty} \lambda(I_n) + \varepsilon;$$

since $\varepsilon > 0$ is arbitrary we get

$$\lambda(K) \le \sum_{n=1}^{\infty} \lambda(I_n),$$

and taking suprema as K varies in the compact subintervals of I we get $\lambda(I) \le \sum_{n=1}^{\infty} \lambda(I_n)$, as desired. □

The premeasure λ_1 associated to the identity function $f(x) = x$ of $\mathbb{R}$ is the *length*, or Lebesgue premeasure, of the intervals.

2.2.3. *Conclusion.* We have seen how to associate to every Radon–Stieltjes premeasure μ on the algebra of intervals a right–continuous increasing function f_μ, its distribution function, which is 0 at 0, and conversely to every increasing $f : \mathbb{R} \to \mathbb{R}$ a Radon–Stieltjes premeasure λ_f on intervals. Of course the same measure λ_f is obtained if we consider $f + k$, with k a real constant, in place of f; and if f, g are increasing functions that coincide where they are continuous, then again the measures λ_f and λ_g coincide. We standardize a choice by taking right continuous increasing functions which are 0 at 0; in this way we have a bijective correspondence between premeasures and functions.

Remark. For finite premeasures μ one often standardizes by taking distribution functions F_μ still right continuous, but with $F_\mu(-\infty) = 0$, so that $f_\mu = F_\mu - F_\mu(0)$, as seen above.

Every increasing function $f : \mathbb{R} \to \mathbb{R}$, even not normalized, defines anyway a Radon–Stieltjes premeasure on intervals, often written df; this notation is due to the fact that if $f \in C^1(\mathbb{R})$, or even if f is continuous and only piecewise C^1, we have for every interval $[a, b]$ (or $]a, b]$ or $[a, b[$, or $]a, b[$, with $a < b$ and $a, b \in \mathbb{R}$):

$$f(b) - f(a) = \lambda_f([a, b]) := \int_a^b df(t) \quad \text{and} \quad f(b) - f(a) = \int_a^b f'(t)\, dt,$$

the last equality being the fundamental theorem of calculus. The premeasure is in some sense the *derivative* of its distribution function.

Finally, we have considered only measures and functions on the entire real line $\mathbb{R}$, but it is quite clear that the same constructions, with obvious modifications, can be repeated for increasing functions and measures on any interval I of $\mathbb{R}$: boundedness will now have to be replaced with "having a compact closure in I". For instance, the function log defines a measure on $]0, \infty[$, finite on compact subintervals, but infinite on the bounded subinterval $]0, 1]$, whose compact closure is not contained in $]0, \infty[$.

2.2.4. *Proof of finite additivity.* We have to prove that if $I = \bigcup_{n=1}^m I_n$, where I is an interval and I_n are pairwise disjoint intervals, then $\lambda_f(I) = \sum_{n=1}^m \lambda_f(I_n)$ (see 2.2.2). Observe that if I_j and I_k are disjoint intervals, then either $I_j < I_k$ (in the sense that $x < y$ if $x \in I_j$ and $y \in I_k$) or $I_k < I_j$, so that we may assume that $I_1 < I_2 < \cdots < I_m$. Since then $I_1 \cup \cdots \cup I_{m-1}$ is also an interval, we are reduced to proving the case $m = 2$; this is boring (various cases have to be considered, e.g. $I = [a, b]$ and $[a, b] = [a, c] \cup]c, b]$ or $[a, b] = [a, c[\cup [c, b]$ etc.) but easy.

2.2.5. *Radon premeasures on intervals of* $\mathbb{R}^n$. A *Radon premeasure* on $\mathbb{R}^n$ is a premeasure defined on the algebra $\mathcal{A}_n$ of $n-$dimensional intervals of $\mathbb{R}^n$, $\mu : \mathcal{A}_n \to [0, \infty]$, that is finite on all compact intervals (hence also on all bounded intervals). A standard way of obtaining Radon premeasures on $\mathbb{R}^n$ is the following: if $\mu_k : \mathcal{A}_1 \to [0, \infty]$, $k = 1, \ldots, n$ is an $n-$tuple of Radon premeasures on $\mathbb{R}^1$ we define $\mu : \mathcal{A}_n \to [0, \infty]$ by $\mu(I_1 \times \cdots \times I_n) = \prod_{k=1}^n \mu_k(I_k)$ (making use of $\infty \cdot 0 = 0 \cdot \infty = 0$); this is the (tensor) *product premeasure* of the premeasures μ_k. One can verify that μ is indeed a premeasure on $\mathcal{A}_n$ (it is not a completely straightforward proof, even finite additivity is not at all immediate). The most important case is that of the $n-$dimensional Lebesgue premeasure $\lambda_n = \lambda_1^{\otimes n}$, product of n one dimensional measures.

Exercise 2.2.3. *(Important)*

(i) Given $a, b \in \mathbb{R}$ with $a < b$ write an explicit decreasing sequence J_k of open intervals with intersection $I =]a, b]$ and an explicit increasing sequence K_j of compact intervals with I as union.

One can then prove easily (accept this fact) that every interval I of $\mathbb{R}$ is the intersection of a decreasing sequence of open intervals, and the union of an increasing sequence of compact intervals; moreover, if the interval is bounded the open intervals may also be taken bounded.

(ii) Prove that the same is true for $n-$dimensional intervals.

(iii) Assume that $\mu : \mathcal{A}_n \to [0, +\infty]$ is premeasure finite on bounded intervals. Then we have, for every interval I:

$$\mu(I) = \sup\{\mu(K) : K \subseteq I,\ K \text{ a compact interval}\},$$

and if I is bounded then also

$$\mu(I) = \inf\{\mu(J) : J \supseteq I,\ J \text{ a bounded open interval}\}; \tag{*}$$

(iv) Let $\mu : \mathcal{A}_n \to [0, +\infty]$ be a finitely additive measure on the $n-$dimensional interval algebra $\mathcal{A}_n$, that is finite on bounded intervals. Assume that for every interval I we have:

$$\mu(I) = \sup\{\mu(K) : K \subseteq I,\ K \text{ a compact interval}\},$$

and that for every bounded interval I:

$$\mu(I) = \inf\{\mu(J) : J \supseteq I, J \text{ a bounded open interval}\};$$

then μ is a premeasure (i.e it is countably additive).

Solution. (i) Take $J_k =]a, b + 1/k[$ and $K_j = [a + (b-a)/j, b]$, with $k, j \geq 1$.

(ii) consider the interval $I = \prod_{k=1}^n I_k$. For every $k \in \{1, \dots, n\}$ let K_r^k, $r = 1, 2, \dots$ be an increasing sequence of compact intervals with union I_k, and J_r^k, $r = 1, 2, \dots$, a decreasing sequence of open intervals with intersection I_k. Then $L_r = \prod_{k=1}^n K_r^k$, $r \geq 1$ is an increasing sequence of compact $n-$dimensional intervals with union I, and $U_r = \prod_{k=1}^n J_r^k$, $r \geq 1$ is a decreasing sequence of open $n-$dimensional intervals with intersection I. Since a product of bounded one–dimensional intervals is bounded, the open intervals U_r are bounded if I is bounded.

(iii) The fact that $\mu(I)$ is the supremum of the measures of compact intervals contained in (i) is an immediate consequence of (ii) and continuity from below of premeasures. For $\mu(I)$ the infimum of measures of open intervals containing it, we can again use (ii) and continuity from above for decreasing sequences of sets of finite measure: a bounded interval is the intersection of a decreasing sequence of bounded open intervals.

(iv) We need to prove countable subadditivity, that if an interval I is the union of a sequence I_k of intervals then $\mu(I) \leq \sum_{k=1}^\infty \mu(I_k)$. If the intervals I_k are bounded we can repeat verbatim the argument given in the proposition 2.2.2 to prove the assertion. If some I_k is unbounded, take a compact interval $K \subseteq I$; we then have, by the bounded case:

$$\mu(K) \leq \sum_{k=1}^\infty \mu(I_k \cap K);$$

since $\mu(I_k \cap K) \leq \mu(I_k)$ we get

$$\mu(K) \leq \sum_{k=1}^\infty \mu(I_k),$$

for every compact subinterval K of I; taking suprema as K varies among compact subintervals of I we conclude. □

Remark. Unlike the one–dimensional case, it is in general not true that the measure of an unbounded $n-$interval is the infimum of the measures of the open *intervals* containing it, if $n \geq 2$. For instance, the $y-$axis $\{0\} \times \mathbb{R}$ has $2-$dimensional Lebesgue measure 0, but every open interval containing it has measure ∞.

2.3. Borel sets and $\sigma-$algebras generated by a class of sets. Given a set X, we may consider the set of all $\sigma-$subalgebras of $\mathcal{P}(X)$, called $\sigma-$algebras on X; this set is partially ordered by inclusion, and it is clear that the intersection of a set of $\sigma-$subalgebras of $\mathcal{P}(X)$ is a $\sigma-$subalgebra of $\mathcal{P}(X)$. Given an arbitrary subset $\mathcal{E}$ of $\mathcal{P}(X)$, we may consider the set of all the $\sigma-$algebras on X containing $\mathcal{E}$ (there is always one, at least all of $\mathcal{P}(X)$); the intersection of this set is the smallest $\sigma-$algebra on X containing $\mathcal{E}$, is denoted by $\mathcal{M}(\mathcal{E})$, and called the *$\sigma-$algebra generated by $\mathcal{E}$*. Unless $\mathcal{E}$ is finite (a very peculiar case, discussed in 1.4.1) there is no simple way of describing all the elements of $\mathcal{M}(\mathcal{E})$. Of course, if $\mathcal{E}$ and $\mathcal{F}$ are classes of subsets of X, and $\mathcal{E} \subseteq \mathcal{F}$ then we have $\mathcal{M}(\mathcal{E}) \subseteq \mathcal{M}(\mathcal{F})$: every $\sigma-$algebra containing $\mathcal{F}$ contains also $\mathcal{E}$, so that $\mathcal{M}(\mathcal{F})$ is the intersection of a smaller set of $\sigma-$algebras than $\mathcal{M}(\mathcal{E})$, and hence is a larger $\sigma-$algebra.

Definition. If (X, τ) is a topological space, the Borel $\sigma-$algebra on X is the $\sigma-$subalgebra $\mathcal{B}(X)$ of $\mathcal{P}(X)$ generated by the topology τ.

The elements of $\mathcal{B}(X)$ are called Borel sets of X (it. boreliani). Trivially the closed sets generate also $\mathcal{B}(X)$. We focus now on $\mathcal{B}(\mathbb{R})$. It is generated by various interesting classes of subsets of $\mathbb{R}$: all intervals; the open intervals; the closed intervals; the open (closed) left rays $]-\infty, a[$ ($a \in \mathbb{R}$) or the open (closed) right rays $]a, +\infty[$ with $a \in \mathbb{R}$; and we can even let a vary on a dense subset C of $\mathbb{R}$, instead of all of $\mathbb{R}$. We give only one of these proofs; from this all the others should be clear:

. *Let C be a dense subset of $\mathbb{R}$. Then the family $\mathcal{E}_C = \{]a, +\infty[: a \in C\}$ generates $\mathcal{B}(\mathbb{R})$ as a $\sigma-$algebra.*

Proof. Since $\mathcal{E}_C$ is contained in the topology of $\mathbb{R}$, the generated $\sigma-$algebra is smaller than $\mathcal{B}(\mathbb{R})$. Let's prove that all open sets are contained in $\mathcal{M}(\mathcal{E}_C)$. First observe that all open right rays are contained in $\mathcal{M}(\mathcal{E}_C)$: given $a \in \mathbb{R}$, pick a sequence $a_n \in C$ such that $a_n \to a$ and $a_n > a$; then $]a, +\infty[= \bigcup_{n\in\mathbb{N}}]a_n, +\infty[$. By complementation, $\mathcal{M}(\mathcal{E}_C)$ contains every closed left ray, $]-\infty, b]$, and by intersection every left half–open interval $]a, b]$. But since every open interval is a countable union of half open intervals, in fact $]a,\ b[= \bigcup_{n=1}^{\infty}]a, b-(b-a)/2^n]$, $\mathcal{M}(\mathcal{E}_C)$ contains all open intervals, hence all open sets, because every open set A is a countable union of open intervals (e.g, those contained in A with both extremes in $\mathbb{Q}$). □

Similarly, $\mathcal{B}(\mathbb{R}^n)$ is generated by all intervals, or all open intervals; this is easy, since every open subset of $\mathbb{R}^n$ is a countable union of cubes of rational side–length, and center in $\mathbb{Q}^n$.

2.3.1. *The tribe induced on a subset.* If $\mathcal{M}$ is a $\sigma-$algebra of parts of the set X, and S is a subset of X, then

$$\mathcal{M}_S = \{A \cap S : A \in \mathcal{M}\}$$

is a $\sigma-$algebra of parts of S, as is easy to see; it is the $\sigma-$algebra induced, or traced, by $\mathcal{M}$ on S. We have

PROPOSITION. *Let X be a set, S subset of X, $\mathcal{E}$ a subset of $\mathcal{P}(X)$, $\mathcal{M} = \mathcal{M}(\mathcal{E})$ the $\sigma-$subalgebra of $\mathcal{P}(X)$ generated by $\mathcal{E}$. If $\mathcal{M}_S = \{A \cap S : A \in \mathcal{M}\}$ is the $\sigma-$subalgebra of $\mathcal{P}(S)$ traced by $\mathcal{M}$ on S, then $\mathcal{M}_S$ is the $\sigma-$subalgebra of $\mathcal{P}(S)$ generated by $\{E \cap S : S \in \mathcal{E}\}$.*

Proof. Let $\mathcal{N}$ be the $\sigma-$ subalgebra of $\mathcal{P}(S)$ generated by $\{E \cap S : S \in \mathcal{E}\}$. Clearly $\mathcal{M}_S \supseteq \{E \cap S : S \in \mathcal{E}\}$, so that $\mathcal{M}_S \supseteq \mathcal{N}$. On the other hand the set $\{A \subseteq X : A \cap S \in \mathcal{N}\}$ is a $\sigma-$subalgebra of $\mathcal{P}(X)$; this $\sigma-$algebra contains $\mathcal{E}$, hence also $\mathcal{M} = \mathcal{M}(\mathcal{E})$, in other words, $A \cap S \in \mathcal{N}$ for every $A \in \mathcal{M}$, that is, $\mathcal{M}_S \subseteq \mathcal{N}$. □

COROLLARY. *If (X, τ) is a topological space, and $S \subseteq X$ has the induced topology, then $\mathcal{B}(S)$ is the $\sigma-$algebra induced on S by the Borel $\sigma-$algebra $\mathcal{B}(X)$.*

EXERCISE 2.3.1. Let $\mathcal{M}$ be an infinite $\sigma-$algebra of parts of X. Prove that there is a countably infinite partition $(A_n)_{n\in\mathbb{N}}$ of X consisting of elements of $\mathcal{M}$. Deduce that every infinite $\sigma-$algebra has cardinality not less than the continuum $\mathfrak{c} = |\mathcal{P}(\mathbb{N})|$.

Solution. We prove that if $\mathcal{M}$ is an infinite algebra of parts of X then X admits a countably infinite partition $(A_n)_{n\in\mathbb{N}}$ consisting of elements of $\mathcal{M}$; when this is done, if the algebra is a $\sigma-$algebra then we have an injective mapping of $\mathcal{P}(\mathbb{N})$ into $\mathcal{M}$ defined by $S \mapsto \bigcup_{n\in S} A_n$.

The proof is by induction, based on the following trivial observation: if $\{X_1, \dots, X_m\}$ is a finite partition of X by elements of $\mathcal{M}$ then at least one of the X_k may be split as $X_k = A \cup B$, disjoint union of two non empty sets $A, B \in \mathcal{M}$: if not, $\mathcal{M}$ has only 2^m elements (those obtained from unions of the X_k's, see 2.1.3).

Starting with the trivial partition $X_{0,1} = X$, we split X as the disjoint union of $X_{1,1},\ X_{1,2} \in \mathcal{M}$, both non–empty; and one of these sets may be further split giving the partition $\{X_{2,1},\ X_{2,2}, X_{2,3}\}$ which refines the preceding one; by what above observed the process never stops, and yields the required countable partition.

REMARK. In fact we have proved that $\mathcal{M}$ contains a copy of the $\sigma-$algebra $\mathcal{P}(\mathbb{N})$ of all subsets of $\mathbb{N}$.

□

2.4. Measurable spaces; measure spaces; complete measure spaces.

DEFINITION. We call *measurable space* an ordered pair $(X, \mathcal{M})$ consisting of a set X and a $\sigma-$algebra $\mathcal{M}$ of parts of X. We call *measure space* a triple $(X, \mathcal{M}, \mu)$, where $(X, \mathcal{M})$ is a measurable space, and $\mu : \mathcal{M} \to [0, +\infty]$ is a measure on $\mathcal{M}$.

In a measurable space $(X, \mathcal{M})$ the elements of $\mathcal{M}$ are often called *measurable sets*. The $\sigma-$algebra of a measure space has two noteworthy *ideals* (see 1.5): the ideal $\mathcal{F}(\mu) = \{A \in \mathcal{M} : \mu(A) < \infty\}$ of sets of finite measure, and the $\sigma-$ideal (see the remark below) $\mathcal{N}(\mu) = \{A \in \mathcal{M} : \mu(A) = 0\}$ of sets of 0 measure: notice in fact that, by countable subadditivity, every countable union of sets of zero measure is still a set of measure zero.

REMARK. In a $\sigma-$algebra the $\sigma-$ideals are ideals closed under countable union: in a measure space the sets of zero measure, and the sets of $\sigma-$finite measure (see next section) are both $\sigma-$ideals. The notion of ideal is however not terribly important in measure theory, and we use only, from time to time, the name.

2.4.1. *Complete measure spaces and completions.* A measure space $(X, \mathcal{M}, \mu)$ is called *complete* if every subset of a set of $\mu-$measure zero is still measurable, i.e. $N \subseteq M$ and $\mu(M) = 0$ imply $N \in \mathcal{M}$ (and $\mu(N) = 0$ too, of course). We can always enlarge the $\sigma-$algebra of all measurable subsets of a measure space $(X, \mathcal{M}, \mu)$ to get a complete measure space $(X, \overline{\mathcal{M}}, \bar{\mu})$: define $\overline{\mathcal{M}}$ as the set of all $E \subseteq X$ such that there are $A, B \in \mathcal{M}$ with $A \subseteq E \subseteq B$ and $\mu(B \setminus A) = 0$ (observe that it is equivalent to define $\overline{\mathcal{M}}$ as the set of all $S \in \mathcal{P}(X)$ of the form $A \cup N$, where N is subset of a set $M \in \mathcal{M}$ with $\mu(M) = 0$). It is easy to see that $\overline{\mathcal{M}}$ is a $\sigma-$algebra containing $\mathcal{M}$, and that the formula $\bar{\mu}(E) = \mu(A)(= \mu(B))$ defines a measure $\bar{\mu} : \overline{\mathcal{M}} \to [0, \infty]$ which makes $(X, \overline{\mathcal{M}}, \bar{\mu})$ a complete space (see below).

Proof. $\overline{\mathcal{M}}$ is a $\sigma-$algebra: if $A \subseteq E \subseteq B$ and $\mu(B \setminus A) = 0$ we have $X \setminus B \subseteq X \setminus E \subseteq X \setminus A$ and $(X \setminus A) \setminus (X \setminus B) = B \setminus A$, so that $\overline{\mathcal{M}}$ is complementation closed. If $(E_n)_{n\in\mathbb{N}}$ is a disjoint sequence in $\overline{\mathcal{M}}$, and $A_n \subseteq E_n \subseteq B_n$, then $\bigcup_n A_n \subseteq \bigcup_n E_n \subseteq \bigcup_n B_n$ and $\bigcup_n B_n \setminus \bigcup_n A_n \subseteq \bigcup_n (B_n \setminus A_n)$, with $\bigcup_n (B_n \setminus A_n)$ of measure zero, so that $\bigcup_n E_n \in \overline{\mathcal{M}}$. And if the E_n's are pairwise disjoint, then so are the A_n's, so that

$$\bar{\mu}\left(\bigcup_n E_n\right) := \mu\left(\bigcup_n A_n\right) = \sum_n \mu(A_n) = \sum_n \bar{\mu}(E_n).$$

□

2.4.2. *Subspaces of a measure space.* If $(X, \mathcal{M}, \mu)$ is a measure space and $S \in \mathcal{M}$, we have the trace $\sigma-$ algebra $\mathcal{M}_S = \{A \cap S : A \in \mathcal{M}\}$; since $\mathcal{M}_S \subseteq \mathcal{M}$, the measure μ restricts to a measure $\mu \mid S : \mathcal{M}_S \to [0, \infty]$; the triple $(S, \mathcal{M}_S, \mu \mid S)$ is the measure space induced on S by the original one.

Remark. Unlike the notion of measurable subspace, where measurability of S was not assumed, we suppose here that $S \in \mathcal{M}$. It is possible to give a notion of induced measure on the subspace $(X, \mathcal{M}_S)$ even with $S \notin \mathcal{M}$ (and thus $\mathcal{M}_S \nsubseteq \mathcal{M}$), but the generality thus gained seems to be not worth the trouble.

2.4.3. *Finite, $\sigma-$finite and semifinite measure spaces.*

Definition. A measure space $(X, \mathcal{M}, \mu)$ is said to be:

(i) *Finite* if $\mu(X) < \infty$, in particular a *probability space* if $\mu(X) = 1$;
(ii) *$\sigma-$finite* if X can be covered by countably many sets in $\mathcal{M}$ of finite measure;
(iii) *semifinite* if for every $E \in \mathcal{M}$ with $\mu(E) > 0$ there is $F \subseteq E$, $F \in \mathcal{M}$, with $0 < \mu(F) < \infty$.

We have that finite implies $\sigma-$finite, and $\sigma-$finite implies semifinite. The first implication is trivial; for the second, if $A_0 \subseteq A_1 \subseteq A_2 \subseteq \dots$ is a sequence in $\mathcal{M}$ such that $A_n \uparrow X$, with each A_n of finite measure, for every $E \in \mathcal{M}$ we have that $E \cap A_n \uparrow E$, and each $E \cap A_n$ is of finite measure, so that $\mu(E) = \lim_n \mu(E \cap A_n) = \sup\{\mu(E \cap A_n) : n \in \mathbb{N}\}$, clearly proving semifiniteness of the measure. Most spaces of interest in analysis are $\sigma-$finite. Non semifinite spaces are rather pathological: by definition they contain measurable sets on which the induced measure has the values 0, ∞ and no other value (atoms of infinite measure, see below, 2.4.4). Often what is needed in a proof is only the fact that every set of nonzero measure contains a subset of finite nonzero measure, that is semifiniteness, much less than $\sigma-$finiteness. But $\sigma-$finite measure spaces are the only ones well–behaved under products.

Example 2.4.1. Let $(X, \mathcal{P}(X), \mu)$ with μ the counting measure; μ is $\sigma-$finite iff X is countable, and is always semifinite. For a non semifinite measure consider a non–empty set X, with the trivial $\sigma-$algebra $\{\emptyset, X\}$ and the measure which is zero on the emptyset and ∞ on X.

2.4.4. *Atoms; purely atomic and atomless measure spaces.*

Definitions. An *atom* in a measure space $(X, \mathcal{M}, \mu)$ is a set $A \in \mathcal{M}$ such that $0 < \mu(A)$, and for every measurable $E \subseteq A$ we have either $\mu(E) = 0$ or $\mu(A \setminus E) = 0$.

A measure space is said to be *atomless* if it has no atoms. It is called *purely atomic* if every $A \in \mathcal{M}$ of strictly positive measure contains an atom.

A tipical example of a purely atomic measure space is the space $(X, \mathcal{P}(X), \mu_w)$ where $w : X \to [0, \infty]$ is any function (see 2.1.3). All purely atomic measure spaces are essentially of this sort. An atomless measure space will be the real line with Lebesgue measure.

Of course, measure subspaces of atomless measure spaces are also atomless, and measure subspaces of purely atomic measure spaces are purely atomic.

2.4.5. *Direct sums of measure spaces.* Let X, Y be disjoint sets and let $\mathcal{A}$ and $\mathcal{B}$ be $\sigma-$algebras over X and Y respectively. The direct sum, or coproduct, of the two $\sigma-$algebras $\mathcal{A}$ and $\mathcal{B}$ is the $\sigma-$algebra $\mathcal{A}\oplus\mathcal{B}$ over $X\cup Y$ defined by $\{A\cup B:\ A\in\mathcal{A},\ B\in\mathcal{B}\}$ (verify that it is a $s-$algebra). If $\mu:\mathcal{A}\to[0,\infty]$ and $\nu:\mathcal{B}\to[0,\infty]$ are measures, then $\mu\oplus\nu:\mathcal{A}\oplus\mathcal{B}\to[0,+\infty]$ defined by $\mu\oplus\nu(A\cup B)=\mu(A)+\nu(B)$ is a measure on $\mathcal{A}\oplus\mathcal{B}$, as it is immediately checked. The measure space $(X\cup Y,\mathcal{A}\oplus\mathcal{B},\ \mu\oplus\nu)$ is the coproduct of the two measure spaces $(X,\mathcal{A},\mu)$ and $(Y,\mathcal{B},\nu)$, and of course has these both as measure subspaces. The construction may be extended to an arbitrary family $(X_\lambda,\mathcal{A}_\lambda,\mu_\lambda)_{\lambda\in\Lambda}$ of measure spaces: assuming the X_λ pairwise disjoint we consider their union X, and on it the $\sigma-$algebra $\mathcal{A}=\bigoplus_{\lambda\in\Lambda}\mathcal{A}_\lambda$ consisting of all subsets A of X such that $A\cap X_\lambda\in\mathcal{A}_\lambda$ for every $\lambda\in\Lambda$; the measure is

$$\mu(A):=\sum_{\lambda\in\Lambda}\mu_\lambda(A\cap A_\lambda).$$

2.4.6. *Exercises.*

Exercise 2.4.2. Let $(X,\mathcal{M},\mu)$ be a measure space. Assume that $E\in\mathcal{M}$ has $\sigma-$finite measure, and that $(A_\lambda)_{\lambda\in\Lambda}$ is an almost disjoint family of measurable subsets of E of strictly positive measure. Prove that then Λ is countable (hint: prove it first with $\mu(E)<\infty$). Deduce from it that in a purely atomic space every set $E\in\mathcal{M}$ of $\sigma-$finite measure is a countable disjoint union of atoms of finite measure.

Solution. Assume first that $0<\mu(E)<\infty$. For every finite subset F of Λ we have

$$\sum_{\lambda\in F}\mu(A_\lambda)=\mu\left(\bigcup_{\lambda\in F}A_\lambda\right)\le\mu(E),$$

(the equality because the A_λ are almost disjoint, the inequality by isotony of μ), so that, denoting by $\Phi(\Lambda)$ the set of all finite subsets of Λ:

$$\sum_{\lambda\in\Lambda}\mu(A_\lambda)\left(:=\sup\left\{\sum_{\lambda\in F}\mu(A_\lambda):\ F\in\Phi(\Lambda)\right\}\right)\le\mu(E)<\infty;$$

by 1.10.3 the set Λ is countable. If E has $\sigma-$finite measure we can write $E=\bigcup_{k\in\mathbb{N}}E(k)$, where each $E(k)$ has finite measure. Let $\Lambda(k)=\{\lambda\in\Lambda:\ \mu(A_\lambda\cap E(k))>0\}$. Then each $\Lambda(k)$ is countable, by what just proved; and $\Lambda=\bigcup_{k\in\mathbb{N}}\Lambda(k)$ (if $\alpha\in\Lambda$ belongs to no $\Lambda(k)$, we have $\mu(A_\alpha\cap E(k))=0$ for every $k\in\mathbb{N}$, but then $A_\alpha=\bigcup_{k\in\mathbb{N}}A_\alpha\cap E(k)$ has measure zero, being a countable union of sets of measure zero).

Then, if E has $\sigma-$finite measure it contains only a countable set of pairwise almost disjoint atoms (all of finite measure, of course: an atom of infinite measure is almost disjoint from every set of $\sigma-$finite measure). If F is the union of this set, then $E\setminus F$ contains no atom; and if the measure is purely atomic, this means that $\mu(E\setminus F)=0$. □

Exercise 2.4.3. Let $(X,\mathcal{M},\mu)$ be a measure space. Prove that the following are equivalent:

(i) $\inf\{\mu(E):\ E\in\mathcal{M},\ \mu(E)>0\}=\alpha>0$.

(ii) Every $E\in\mathcal{M}$ of finite nonzero measure is a finite disjoint union of atoms.

And if (i) holds then the space is purely atomic.

Solution. We prove the last assertion first, that if $\mu(E)>0$ then some atom A is contained in E. If $\mu(E)=\infty$, and E contains no measurable set of finite nonzero measure, then E itself is an atom of infinite measure. If not, E contains some measurable subset E_0 with $0<\mu(E_0)<\infty$. If E_0 is not an atom, then it contains a measurable subset E_1 with $0<\mu(E_1)<\mu(E_0)$; inductively, unless we find an atom, we get a sequence $E_0\supseteq E_2\supseteq\dots$ of measurable sets with strictly decreasing measures, $\mu(E_0)>\mu(E_1)>\dots$; this clearly contradicts the hypothesis, since then $\sum_{n=0}^{\infty}\mu(E_n\setminus E_{n+1})\le\mu(E_0)<\infty$, and hence $\lim_{n\to\infty}\mu(E_n\setminus E_{n+1})=0$, but $\mu(E_n\setminus E_{n+1})>0$ for every n. We are then bound to find an atom contained in E.

(i) implies (ii): if $0<\mu(E)<\infty$, since the space is purely atomic E is a countable disjoint union of atoms (see 2.4.2), necessarily of finite measure; and since these atoms have measures bounded away from zero this union must actually be finite.

(ii) implies (i): if (i) is false there is a sequence $E(n)\in\mathcal{M}$ such that $0<\mu(E(n))<\infty$ and $\lim_{n\to\infty}\mu(E(n))=0$; taking a subsequence if necessary we can assume $\mu(E(n))\le 1/2^n$. Then the set $E=\bigcup_{n=0}^{\infty}E(n)$ has finite nonzero measure, and contains all atoms contained in any $E(n)$, which cannot be a finite set, since their measures have 0 as infimum not a minimum. □

2.5. Outer measures and extension of premeasures. To be able to construct a significant theory we need to have measures defined on $\sigma-$algebras of subsets of a set. Given a premeasure μ defined on an algebra $\mathcal{A}$ we prove that μ can always be extended to a measure on the $\sigma-$algebra $\mathcal{M}(\mathcal{A})$ generated by $\mathcal{A}$. There are various ways of doing that. We describe the one that seems to be the most commonly used nowadays, through *outer measures.*

2.5.1. *Outer measures: definition.*

Definition. An *outer measure* on a set X is a function $\phi : \mathcal{P}(X) \to [0, +\infty]$ such that

(i) $\phi(\emptyset) = 0$;
(ii) ϕ is monotone, that is $A \subseteq B$ implies $\phi(A) \le \phi(B)$;
(iii) ϕ is countably subadditive, that is, for every sequence $(A_n)_{n\in\mathbb{N}}$ of subsets of X we have

$$\phi\left(\bigcup_{n\in\mathbb{N}} A_n\right) \le \sum_{n=0}^{\infty} \phi(A_n).$$

2.5.2. *Constructing outer measures.* A standard way of constructing outer measures is as follows: we are given a set $\mathcal{E} \subseteq \mathcal{P}(X)$ and a function $\rho : \mathcal{E} \to [0, +\infty]$. We define $\phi = \phi_\rho : \mathcal{P}(X) \to [0, +\infty]$ by $\phi(\emptyset) = 0$ and for every $A \subseteq X$:

$$\phi(A) = \inf\left\{\sum_{n=0}^{\infty} \rho(E_n) : \ A \subseteq \bigcup_{n\in\mathbb{N}} E_n,\ E_n \in \mathcal{E}\right\}$$

(the infimum is taken on the set of sums obtained as $(E_n)_{n\in\mathbb{N}}$ varies on all countable covers of A by elements of $\mathcal{E}$; if no such cover exists, then $\phi(A) = +\infty$). Trivially ϕ is monotone. To check countable subadditivity assume that $A = \bigcup_{n\in\mathbb{N}} A_n$; we prove that

$$\phi(A) \le \sum_{n=0}^{\infty} \phi(A_n).$$

If the right hand side is $+\infty$ there is nothing to prove. If it is finite, given $\varepsilon > 0$ pick for every n a countable cover $(E(n,\, k) : \, k \in \mathbb{N}\}$ of A_n by elements of $\mathcal{E}$ such that

$$\sum_{k=0}^{\infty} \rho(E(n,\, k)) \le \phi(A_n) + \frac{\varepsilon}{2^{n+1}};$$

Then $\{E(n,\, k) : \, (n,\, k) \in \mathbb{N}\times\mathbb{N}\}$ is a countable cover of $A = \bigcup_{n\in\mathbb{N}} A_n$, so that

$$\phi(A) \le \sum_{(n\, k)\in\mathbb{N}\times\mathbb{N}} \rho(E(n,\, k)) = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{\infty} \rho(E(n,\, k))\right) \le \sum_{n=0}^{\infty}\left(\phi(A_n) + \frac{\varepsilon}{2^{n+1}}\right) = \sum_{n=0}^{\infty} \phi(A_n) + \varepsilon;$$

since $\varepsilon > 0$ is arbitrary we conclude.

Remark. Notice that $\phi = \phi_\rho$ is unchanged if we remove from $\mathcal{E}$ the sets with $\rho(E) = \infty$ (if all countable covers of $A \subseteq X$ by elements of $\mathcal{E}$ must contain elements $E \in \mathcal{E}$ with $\rho(E) = \infty$, then $\phi_\rho(A) = \infty$, anyway).

2.5.3. *The outer measure associated to a premeasure.* If $\mathcal{A}$ is an algebra of parts of X and $\mu : \mathcal{A} \to [0, \infty]$ is a premeasure, the outer measure defined on $\mathcal{P}(X)$ as in 2.5.2 is denoted by μ^* and it is the outer measure naturally associated to the premeasure μ.

. *For every $E \subseteq X$ we have*

$$\mu^*(E) = \inf\left\{\sum_{n=0}^{\infty} \mu(A_n) : \ A_n \in \mathcal{A} \text{ a disjoint sequence, } \mu(A_n) < \infty,\ A \subseteq \bigcup_{n=0}^{\infty} A_n\right\}.$$

(with the usual proviso that if no such cover exists, then $\mu^*(E) = \infty$).

The proof is left as an exercise, using 2.1.5. Notice that if $\mathcal{A}$ is a $\sigma-$algebra, i.e. μ is a measure, then we have, for every $E \subseteq X$:

$$\mu^*(E) = \inf\{\mu(A) : \, E \subseteq A,\ A \in \mathcal{A}\}.$$

Notice also that, as observed above, μ^* depends only on the ideal $\mathcal{R}$ of $\mathcal{A}$ consisting of the sets of finite measure; $\mathcal{R}$ is a ring of sets, closed under union and difference.

2.5.4. *Measurability with respect to an outer measure.* If an outer measure ϕ on X is not countably additive, then it cannot be finitely additive (2.1.5). This means that there are pairs E, F of disjoint subsets of X such that $\phi(E \cup F) < \phi(E) + \phi(F)$ (subadditivity of course implies that $\phi(E \cup F) \leq \phi(E) + \phi(F)$). This is a partial motivation for the following:

Definition. If $\phi : \mathcal{P}(X) \to [0, +\infty]$ is an outer measure, a subset A of X is said to be *measurable* (with respect to ϕ) if for every E in $\mathcal{P}(X)$ we have

$$\phi(E) = \phi(E \cap A) + \phi(E \setminus A)(= (\phi(E \cap A) + \phi(E \cap (X \setminus A))).$$

Briefly: a set is measurable if it splits additively every other set. To check measurability we need only to prove that $\phi(E) \geq \phi(E \cap A) + \phi(E \setminus A)$, since subadditivity gives the other inequality; hence we can also assume $\phi(E) < \infty$. Notice that all sets of outer measure 0 are measurable: if $\phi(A) = 0$ then $\phi(E \cap A) \leq \phi(A) = 0$ for every E, so that $\phi(E \cap A) + \phi(E \setminus A) = \phi(E \setminus A) \leq \phi(E)$ by monotonicity. One of the most important results of measure theory is

. The Carathèodory theorem *Let $\phi : \mathcal{P}(X) \to [0, +\infty]$ be an outer measure. Then the set $\mathcal{M}$ of all measurable subsets of X is a $\sigma-$algebra on X, and the restriction of ϕ to $\mathcal{M}$ is a complete measure on $\mathcal{M}$.*

We divide the proof in some steps.

. *Step 1 $\mathcal{M}$ is an algebra.*

Proof. First notice that $\mathcal{M}$ is closed under complementation, since the definition of measurability of A is symmetric in A and $X \setminus A$. To prove that it is an algebra we prove that it is closed under union.

Let $A, B \in \mathcal{M}$. We have to prove that $\phi(E)$ coincides with:

$$\phi(E \cap (A \cup B)) + \phi(E \setminus (A \cup B)) \tag{*}$$

for every $E \subseteq X$. By measurability of A we have:

$$\phi(E \cap (A \cup B)) = \phi(E \cap (A \cup B) \cap A) + \phi(E \cap (A \cup B) \setminus A);$$

now $E \cap (A \cup B) \cap A = E \cap A$, and $E \cap (A \cup B) \setminus A = (E \setminus A) \cap B$ (just use distributivity of $\cap$ with respect to $\cup$), so that (*) becomes (noticing also that $E \setminus (A \cup B) = (E \setminus A) \setminus B$):

$$\phi(E \cap A) + \phi((E \setminus A) \cap B) + \phi((E \setminus A) \setminus B); \tag{**}$$

by measurability of B we have $\phi((E \setminus A) \cap B) + \phi((E \setminus A) \setminus B) = \phi(E \setminus A)$, so that (**) coincides with

$$\phi(E \cap A) + \phi(E \setminus A), \tag{***}$$

which is $\phi(E)$, by measurability of A. □

We have proved that $\mathcal{M}$ is an algebra. It remains to prove that it is a $\sigma-$algebra, and that ϕ is countably additive on it. Let us prove

. *Step 2 If $A_1, \dots, A_m$ are measurable and pairwise disjoint, then for every $E \subseteq X$ we have*

$$\phi(E \cap (A_1 \cup \dots \cup A_m)) = \sum_{k=1}^{m} \phi(E \cap A_k).$$

Proof. This is easy by induction, we need to prove it for $m = 2$. In fact, if A is measurable, and E, B are subsets of X, with B disjoint from A, we have, by measurability of A:

$$\phi(E \cap (A \cup B)) = \phi(E \cap (A \cup B) \cap A)) + \phi(E \cap (A \cup B) \cap (X \setminus A)) = \phi(E \cap A) + \phi(E \cap B).$$

□

. *Step 3 $\mathcal{M}$ is a tribe, and ϕ induces a measure on $\mathcal{M}$.*

Proof. Since we know that $\mathcal{M}$ is an algebra, we need only to prove that $\mathcal{M}$ is closed under countable disjoint union. Let $(A_n)_{n \geq 1}$ be a disjoint sequence of measurable sets. Let E be a subset of X. For every $m \geq 1$ measurability of $\bigcup_{n=1}^{m} A_n$ implies

$$\phi(E) = \phi(E \cap (A_1 \cup \dots \cup A_m)) + \phi(E \cap (X \setminus (A_1 \cup \dots \cup A_m)),$$

and by Step 2 we get

$$\phi(E \cap (A_1 \cup \dots \cup A_m)) = \sum_{k=1}^{m} \phi(E \cap A_k).$$

If $A = \bigcup_{n=1}^{\infty} A_n$ we have $X \smallsetminus A \subseteq X \smallsetminus (A_1 \cup \cdots \cup A_m)$, so that, by monotonicity

$$\phi(E \cap (X \smallsetminus (A_1 \cup \cdots \cup A_m))) \geq \phi(E \cap (X \smallsetminus A)) = \phi(E \smallsetminus A);$$

hence

$$\phi(E) \geq \sum_{k=1}^{m} \phi(E \cap A_k) + \phi(E \smallsetminus A).$$

Taking the limit on the right–hand side as m tends to infinity we get

$$(*) \qquad \phi(E) \geq \sum_{k=1}^{\infty} \phi(E \cap A_k) + \phi(E \smallsetminus A) \geq \phi(E \cap A) + \phi(E \smallsetminus A),$$

the last inequality being due to countable subadditivity of ϕ. Then A is measurable. From (*) we also get, for every $E \subseteq X$:

$$\phi(E) = \sum_{k=1}^{\infty} \phi(E \cap A_k) + \phi(E \smallsetminus A),$$

and putting in this equality $E = A$ we get countable additivity of ϕ on $\mathcal{M}$. □

Exercise 2.5.1. Let X be a non–empty set, and define $\phi : \mathcal{P}(X) \to [0, \infty]$ by $\phi(\emptyset) = 0$, $\phi(X) = 1$, and $\phi(E) = 1/2$ for $\emptyset \subsetneqq E \subsetneqq X$. Prove that ϕ is an outer measure. Assuming $|X| \geq 3$, prove that the only $\phi-$measurable sets are $\emptyset$ and X. What can be said for $|X| \leq 2$?

Solution. (*of Exercise 2.5.1*) Monotonicity and countable subadditivity are immediate. Assume that a proper non–empty subset A of X is $\phi-$measurable (the existence of such a set implies that $|X| > 1$). Pick E such that $E \cap A$ and $E \smallsetminus A$ are both non–empty, say both consisting of one point. Then $\phi(E) = \phi(E \cap A) + \phi(E \smallsetminus A) = 1/2 + 1/2 = 1$, which is possible iff $E = X$, and implies that X has exactly two elements. In this case clearly ϕ is a measure on $\mathcal{P}(X)$. □

2.5.5. *Extension of a premeasure.* Assume now that we have a premeasure $\mu : \mathcal{A} \to [0, +\infty]$ on an algebra $\mathcal{A}$ of parts of X. In 2.5.3 we have defined the outer measure μ^* associated to μ. Then:

Lemma. *With μ, μ^* and $\mathcal{A}$ as above, if $A \in \mathcal{A}$ then: (i) $\mu^*(A) = \mu(A)$; (ii) A is μ^*-measurable.*

Proof. (i) Clearly $\mu^*(A) \leq \mu(A)$ (consider the cover $(A, \emptyset, \emptyset, \dots)$). And if $(A_n)_{n\in\mathbb{N}}$ is a countable cover of A by elements of $\mathcal{A}$ we have $A = \bigcup_{n=0}^{\infty}(A \cap A_n)$, so that, by countable subadditivity of μ on $\mathcal{A}$, we have $\mu(A) \leq \sum_{n=0}^{\infty} \mu(A \cap A_n)$; since $\mu(A \cap A_n) \leq \mu(A_n)$ for every $n \in \mathbb{N}$ we get

$$\mu(A) \leq \sum_{n=0}^{\infty} \mu(A \cap A_n) \leq \sum_{n=0}^{\infty} \mu(A_n),$$

for every countable cover $(A_n)_{n\in\mathbb{N}}$ of A by elements of $\mathcal{A}$, proving that $\mu(A) \leq \mu^*(A)$.

(ii) Given A in $\mathcal{A}$ an $E \in \mathcal{P}(X)$ of finite μ^* measure we have to prove that

$$\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \smallsetminus A).$$

Given $\varepsilon > 0$ take a countable cover $(A_n)_{n\in\mathbb{N}}$ of E by elements of $\mathcal{A}$ such that $\sum_{n=0}^{\infty} \mu(A_n) \leq \mu^*(E) + \varepsilon$. Since $\mu(A_n) = \mu(A_n \cap A) + \mu(A_n \smallsetminus A)$ we have

$$\mu^*(E) + \varepsilon \geq \sum_{n=0}^{\infty} \mu(A_n) = \sum_{n=0}^{\infty} \mu(A_n \cap A) + \sum_{n=0}^{\infty} \mu(A_n \smallsetminus A);$$

since $(A_n \cap A)_{n\in\mathbb{N}}$ and $(A_n \smallsetminus A)_{n\in\mathbb{N}}$ are countable covers of $E \cap A$ and $E \smallsetminus A$, respectively, we get

$$\sum_{n=0}^{\infty} \mu(A_n \cap A) \geq \mu^*(E \cap A); \quad \sum_{n=0}^{\infty} \mu(A_n \smallsetminus A) \geq \mu^*(E \smallsetminus A),$$

so that

$$\mu^*(E) + \varepsilon \geq \mu^*(E \cap A) + \mu^*(E \smallsetminus A);$$

since $\varepsilon > 0$ is arbitrary this implies

$$\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \smallsetminus A),$$

thus proving μ^*-measurability of A. □

Remark. The preceding lemma remains valid if we assume only that $\mathcal{A}$ is a *ring* of parts of X, with $\mu : \mathcal{A} \to [0, \infty]$ countably additive and zero on the emptyset: we do not need $X \in \mathcal{A}$, we only need to know that $\mathcal{A}$ is closed under union and difference, as is apparent from the proof. This fact is often useful. Countable additivity of μ on $\mathcal{A}$ of course cannot be dispensed with.

2.5.6. *Uniqueness of the extension.* Let $\mathcal{A}$ be an algebra of parts of a set X, and let $\mu : \mathcal{A} \to [0, +\infty]$ be a premeasure on $\mathcal{A}$, and let μ^* be the outer measure induced from μ. Caratheòdory's theorem shows that μ may be extended to a complete measure $\bar{\mu}$ on the $\sigma-$algebra $\mathcal{M}_\mu$ of μ^*-measurable sets, a $\sigma-$algebra of parts of X that contains $\mathcal{A}$, and hence contains also the $\sigma-$algebra $\mathcal{M}(\mathcal{A})$ generated by $\mathcal{A}$. This is the *Caratheòdory's extension* of the measure, $\bar{\mu} : \mathcal{M}_\mu \to [0, +\infty]$. It is a maximal extension in this sense: if $\nu : \mathcal{S} \to [0, \infty]$ is another measure on a $\sigma - algebra$ $\mathcal{S}$ containing $\mathcal{A}$, and ν coincides with μ on $\mathcal{A}$, then $\nu(E) \le \bar{\mu}(E)$ for every $E \in \mathcal{M}_\mu \cap \mathcal{S}$. In fact, if $E \in \mathcal{M}_\mu \cap \mathcal{S}$ is covered by the sequence $(A_n)_{n\in\mathbb{N}}$ of $\mathcal{A}$ we have by subadditivity and monotonicity of ν:

$$\nu(E) \le \sum_{n=0}^{\infty} \nu(A_n) = \sum_{n=0}^{\infty} \mu(A_n) \Longrightarrow \nu(E) \le \mu^*(E) = \bar{\mu}(E).$$

And we have equality for all sets of $\sigma-$finite $\bar{\mu}$ measure:

If $\bar{\mu}(E) < \infty$, given $\varepsilon > 0$ we pick a sequence $(A_n)_{n\in\mathbb{N}}$ of $\mathcal{A}$ which covers E and is such that $\sum_{n=0}^{\infty} \mu(A_n) \le \bar{\mu}(E) + \varepsilon$; if $A = \bigcup_{n=0}^{\infty} A_n$ we have

$$\bar{\mu}(A) = \lim_{m\to\infty} \bar{\mu}\left(\bigcup_{n=0}^{m} A_n\right) = \lim_{m\to\infty} \nu\left(\bigcup_{n=0}^{m} A_n\right) = \nu(A);$$

and we have

$$\bar{\mu}(E) \le \bar{\mu}(A) = \nu(A) = \nu(E) + \nu(A \setminus E) \le \nu(E) + \bar{\mu}(A \setminus E) \le \nu(E) + \varepsilon,$$

so that $\bar{\mu}(E) \le \nu(E) + \varepsilon$ for every $\varepsilon > 0$, and hence $\bar{\mu}(E) \le \nu(E)$; notice that $\bar{\mu}(A) \le \sum_{n=0}^{\infty} \mu(A_n) \le \bar{\mu}(E) + \varepsilon$. Since every set of $\sigma-$finite $\bar{\mu}-$measure is a countable disjoint union of sets of finite measure, the proof is achieved.

The following proposition contains the relevant part of the preceding discussion.

Proposition. *Let $\mathcal{A}$ be an algebra of parts of a set X, and let $\mu : \mathcal{A} \to [0, +\infty]$ be a premeasure on $\mathcal{A}$. The Caratheòdory's extension extends μ to a complete measure $\bar{\mu}$ on the $\sigma-$algebra $\mathcal{M}_\mu$, a $\sigma-$algebra containing $\mathcal{M}(\mathcal{A})$; if $\nu : \mathcal{S} \to [0, \infty]$ is another measure on a $\sigma-$algebra $\mathcal{S}$ containing $\mathcal{A}$, which coincides with μ on $\mathcal{A}$, then we have $\nu(E) \le \bar{\mu}(E)$ for every $E \in \mathcal{M}_\mu \cap \mathcal{S}$, with equality for all sets E of $\sigma-$finite $\bar{\mu}-$measure.*

Proof. See above. □

2.5.7. *Approximation of sets of finite measure.* The following simple fact will be very important in integration theory:

. *If $\mu : \mathcal{A} \to [0, +\infty]$ is a premeasure and $\bar{\mu} : \mathcal{M}_\mu \to [0, +\infty]$ its Caratheòdory extension, for every $E \in \mathcal{M}_\mu$ of finite $\bar{\mu}-$measure and every $\varepsilon > 0$ there exists $A \in \mathcal{A}$ such that $\bar{\mu}(E \triangle A) < \varepsilon$.*

Proof. Find a countable cover $(A_n)_{n\in\mathbb{N}}$ of E by elements of $\mathcal{A}$ such that $\sum_{n=0}^{\infty} \mu(A_n) \le \bar{\mu}(E) + \varepsilon$; if $B = \bigcup_{n=0}^{\infty} A_n$we then have $B \in \mathcal{M}_\mu$ and $\bar{\mu}(B) \le \bar{\mu}(E) + \varepsilon$, equivalently $\bar{\mu}(B \setminus E) \le \varepsilon$ (because $\bar{\mu}(E) < \infty$). Pick $m \in \mathbb{N}$ such that $\sum_{n=m+1} \mu(A_n) < \varepsilon$, and set $A = \bigcup_{n=0}^{m} A_n$. Then, noticing that $E \setminus A \subseteq \bigcup_{n=m+1}^{\infty} A_n$:

$$\bar{\mu}(E \triangle A) = \bar{\mu}(E \setminus A) + \bar{\mu}(A \setminus E) \le \bar{\mu}\left(\bigcup_{n=m+1}^{\infty} A_n\right) + \bar{\mu}(B \setminus E) \le \sum_{n=m+1}^{\infty} \mu(A_n) + \varepsilon < 2\varepsilon.$$

□

2.5.8. *An example.*

Exercise 2.5.2. We have seen (2.1.3) that if X is a set and $w : X \to [0, \infty]$ is a positive function then the formula $\mu_w(A) = \sum_{x\in A} w(x)$ for $A \subseteq X$ defines a measure $\mu_w : \mathcal{P}(X) \to [0, \infty]$. We assume that X is an infinite set, that w is finite–valued, and that $\mu_w(X) = \infty$. Let $\mathcal{A}$ be the subalgebra of $\mathcal{P}(X)$ consisting of finite and cofinite subsets; define the set function $\mu : \mathcal{A} \to [0, \infty]$ by the formula $\mu(A) = \mu_w(A)$ if A is finite, and $\mu(A) = \infty$ if A is cofinite.

(i) Prove that $\mu(A) = \mu_w(A)$ for every $A \in \mathcal{A}$ and that μ is a premeasure

Let now $\mu^* : \mathcal{P}(X) \to [0, \infty]$ be the outer measure associated to μ by the usual procedure.

(ii) Prove that $\mu_w(E) \le \mu^*(E)$ for every $E \subseteq X$, with equality for every countable set.
(iii) Prove that $\mu^*(E) = \infty$ for every uncountable subset of X.
(iv) Prove that every $E \in \mathcal{P}(X)$ is μ^*-measurable.

We now specialize by taking $X = \mathbb{R}$ and $w : \mathbb{R} \to [0, \infty[$ defined by $w(x) = 0$ if $x \in \mathbb{R} \smallsetminus \mathbb{Z}$, and $w(x) = 2^{-x} \wedge 1$ for $x \in \mathbb{N}$.

(v) Observe that on every uncountable subset E of $\mathbb{R} \smallsetminus \mathbb{Z}$ we have $0 = \mu_w(E) < \mu^*(E) = \infty$.

Solution. (i) Notice that since w is finite–valued we have $\mu_w(E) < \infty$ for every finite subset E of X. For every $A \subseteq X$ we have $\mu_w(X) = \mu_w(A) + \mu_w(X \smallsetminus A)$; if A is cofinite then $\mu_w(X \smallsetminus A) < \infty$, and by the hypothesis $\mu_w(X) = \infty$ we get $\mu_w(A) = \infty$, so that $\mu_w(A) = \mu(A)$ for cofinite sets. Countable additivity of μ on $\mathcal{A}$ is now obvious, since $\mu(A) = \mu_w(A)$ for every $A \in \mathcal{A}$ and μ_w is countably additive on $\mathcal{P}(X)$. Then μ is a premeasure.

(ii) Given $E \subseteq X$, assume that $(A_n)_{n\in\mathbb{N}}$ is a countable cover of E by elements of $\mathcal{A}$. Then, first by monotonicity and then by countable subadditivity of μ_w:

$$\mu_w(E) \le \mu_w\left(\bigcup_{n\in\mathbb{N}} A_n\right) \le \sum_{n\in\mathbb{N}} \mu_w(A_n) = \sum_{n\in\mathbb{N}} \mu(A_n),$$

the last equality due to the fact that $\mu_w(A_n) = \mu(A_n)$ for every $n \in \mathbb{N}$ since $A_n \in \mathcal{A}$. We have proved that

$$\mu_w(E) \le \sum_{n\in\mathbb{N}} \mu(A_n) \quad \text{for every countable cover of } E \text{ by elements of } \mathcal{A},$$

which implies $\mu_w(E) \le \mu^*(E)$. If E is countable then $\{\{x\} : x \in E\}$ is a countable cover of E by elements of $\mathcal{A}$ and $\mu_w(E) = \sum_{x\in E} w(x) = \sum_{x\in E} \mu(\{x\})$, so that $\mu^*(E) \le \mu_w(E)$, and equality $\mu^*(E) = \mu_w(E)$ then holds.

(iii) If E is uncountable then every countable cover $(A_n)_{n\in\mathbb{N}}$ by elements of $\mathcal{A}$ must contain a cofinite set, since a countable union of finite sets is at most countable and cannot contain the uncountable set E. Then $\sum_{n\in\mathbb{N}} \mu(A_n) = \infty$, because one element of the sum is infinite. We have seen that for every countable cover $(A_n)_{n\in\mathbb{N}}$ of E by elements of $\mathcal{A}$ we have $\sum_{n\in\mathbb{N}} \mu(A_n) = \infty$: this is equivalent to say that $\mu^*(E) = \infty$.

(iv) We know that a set $A \subseteq X$ is μ^*-measurable if and only if $\mu^*(E) = \mu^*(A \cap E) + \mu^*(E \smallsetminus A)$ for every $E \subseteq X$ with $\mu^*(E) < \infty$. This means that E is countable; then also $E \cap A$ and $E \smallsetminus A$ are countable, being subsets of E. But then $\mu_w(E) = \mu_w(E \cap A) + \mu_w(E \smallsetminus A)$ by additivity of μ_w, and on each of these sets μ_w and μ^* coincide, as seen in (ii); in other words the preceding equality is exactly $\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \smallsetminus A)$.

(v) Trivial. □

This example shows that the Carathèodory extension may in fact be not unique on the generated $\sigma-$algebra: μ_w and μ^* are both defined on the same tribe $\mathcal{P}(\mathbb{R})$, they coincide on $\mathcal{A}$, but differ on co–countable sets, which are in the $\sigma-$algebra generated by $\mathcal{A}$, and shows also that the $\sigma-$algebra of μ^*-measurable sets, in this case $\mathcal{P}(\mathbb{R})$, may be much larger than the $\sigma-$algebra generated by $\mathcal{A}$.

2.6. Uniqueness of measures. Assume that μ, $\nu : \mathcal{M} \to [0, \infty]$ are measures on the same measurable space $(X, \mathcal{M})$. The coincidence set $\mathcal{C} = \{E \in \mathcal{M} : \mu(E) = \nu(E)\}$ contains of course the emptyset, and is closed under countable disjoint union and countable increasing union. But this set is in general not an algebra, even assuming finite measures and $\mu(X) = \nu(X)$:

Example 2.6.1. Let $X = \{1, 2, 3, 4\}$, $\mathcal{M} = \mathcal{P}(X)$; define measures $\mu, \nu : \mathcal{M} \to [0, 1]$ by $\mu(\{a\}) = 1/4$ for every $a \in X$, and $\nu(\{2\}) = \nu(\{4\}) = 1/2$, $\nu(\{1\}) = \nu(\{3\}) = 0$. Then:

$$\mathcal{C} := \{E \in \mathcal{M} : \mu(E) = \nu(E)\} = \{\emptyset, \{1, 2, 3, 4\}, \{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 1\}\},$$

as is easy to check, and $\mu(X) = \nu(X) = 1$; $\mathcal{C}$ is not an algebra, the algebra generated by $\mathcal{C}$ is $\mathcal{M} = \mathcal{P}(X)$ (every singleton is an intersection of two of the pairs in $\mathcal{C}$). So two finite measures can coincide on a set of generators of the $\sigma-$algebra of measurable sets, and still be different.

2.6.1. *Dynkin classes.*

Definition. A Dynkin class of parts of X is a subset $\mathcal{D}$ of $\mathcal{P}(X)$ which contains the emptyset (as an element), is complementation closed, and closed under countable disjoint union.

Every $\sigma-$algebra is clearly a Dynkin class; and $\{\emptyset, \{1,2,3,4\}, \{1,2\}, \{2,3\}, \{3,4\}, \{4,1\}\}$ is a Dynkin class that is not a $\sigma-$algebra. Every intersection of a family of Dynkin classes is of course a Dynkin class, so that we may speak of the Dynkin class $\mathcal{D}(\mathcal{E})$ generated by the subset $\mathcal{E} \subseteq \mathcal{P}(X)$: it is the intersection of all Dynkin subclasses of $\mathcal{P}(X)$ containing $\mathcal{E}$; among these are all $\sigma-$subalgebras of $\mathcal{P}(X)$ containing $\mathcal{E}$, so that

$$\mathcal{D}(\mathcal{E}) \subseteq \mathcal{M}(\mathcal{E}).$$

Recall that an algebra of parts of X is a $\sigma-$algebra iff it is closed under countable disjoint union (see 2.1.8) so that a Dynkin class is a $\sigma-$algebra iff it is an algebra; and since it is complementation closed, this happens iff the Dynkin class is closed under intersection (a $\pi-$system, as some authors say).

Lemma. *Let $\mathcal{D}$ be a Dynkin class. Given $A \in \mathcal{D}$ let $\mathcal{D}_A = \{E \in \mathcal{D} : A \cap E \in \mathcal{D}\}$. Then $\mathcal{D}_A$ is a Dynkin class containing A.*

Proof. Plainly $X \in \mathcal{D}_A$ and $\mathcal{D}_A$ is closed under countable disjoint union (if E_n is a disjoint sequence in $\mathcal{D}_A$ with union $E \in \mathcal{D}$ then $E \cap A$ is the disjoint union of the sequence $E_n \cap A \in \mathcal{D}$, hence $E \cap A \in \mathcal{D}$). It remains to prove that $E \in \mathcal{D}_A$ implies $X \setminus E \in \mathcal{D}_A$, i.e. $(X \setminus E) \cap A \in \mathcal{D}$; equivalently, we prove that $X \setminus ((X \setminus E) \cap A) \in \mathcal{D}$, true because this set is the disjoint union of $E \cap A \in \mathcal{D}$ and $X \setminus A \in \mathcal{D}$. □

Let's now prove the

. Dynkin's theorem *If $\mathcal{E}$ is a subset of $\mathcal{P}(X)$ closed under intersection, then the Dynkin class generated by $\mathcal{E}$ coincides with the $\sigma-$algebra generated by $\mathcal{E}$:*

$$\mathcal{D}(\mathcal{E}) = \mathcal{M}(\mathcal{E}).$$

Proof. We know that $\mathcal{D}(\mathcal{E}) \subseteq \mathcal{M}(\mathcal{E})$; to prove the reverse inclusion we prove that $\mathcal{D}(\mathcal{E})$ is a $\sigma-$algebra, proving that it is closed under intersection. Given $E \in \mathcal{E}$ we consider, as in the lemma, the set $\mathcal{D}_E = \{D \in \mathcal{D}(\mathcal{E}) : D \cap E \in \mathcal{D}(\mathcal{E})\}$; this set contains $\mathcal{E}$, because $E \cap D \in \mathcal{E} \subseteq \mathcal{D}(\mathcal{E})$ by the hypothesis that $\mathcal{E}$ is closed under intersection, and is a Dynkin class by the lemma; then $\mathcal{D}(\mathcal{E}) \subseteq \mathcal{D}_E$, and hence $\mathcal{D}(\mathcal{E}) = \mathcal{D}_E$. We have proved: for every $A \in \mathcal{D}(\mathcal{E})$ and every $E \in \mathcal{E}$ we have that $A \cap E \in \mathcal{D}(\mathcal{E})$. Now, given $A \in \mathcal{D}(\mathcal{E})$ consider $\mathcal{D}_A = \{E \in \mathcal{D}(\mathcal{E}) : A \cap E \in \mathcal{D}(\mathcal{E})\}$. By the lemma this is a Dynkin class, and by what just proved $\mathcal{E} \subseteq \mathcal{D}_A$, so that $\mathcal{D}(\mathcal{E}) \subseteq \mathcal{D}_A$, hence actually $\mathcal{D}(\mathcal{E}) = \mathcal{D}_A$. We have proved: for every $A \in \mathcal{D}(\mathcal{E})$ and every $B \in \mathcal{D}(\mathcal{E})$ we have $A \cap B \in \mathcal{D}(\mathcal{E})$. Then $\mathcal{D}(\mathcal{E})$ is a $\sigma-$algebra. □

2.6.2. *Uniqueness of finite measures.*

Proposition. *Let μ, ν be finite measures on the same measurable space $(X, \mathcal{M})$; assume $\mu(X) = \nu(X)$, and that the set $\mathcal{E} \subseteq \mathcal{C} = \{E \in \mathcal{M} : \mu(E) = \nu(E)\}$ is closed under intersection. Then $\mathcal{C}$ contains the $\sigma-$ algebra $\mathcal{M}(\mathcal{E})$ generated by $\mathcal{E}$.*

Proof. It is easily checked that in the stated hypotheses $\mathcal{C} = \{E \in \mathcal{M} : \mu(E) = \nu(E)\}$ is a Dynkin class; then $\mathcal{C} \supseteq \mathcal{D}(\mathcal{E})$, and $\mathcal{D}(\mathcal{E}) = \mathcal{M}(\mathcal{E})$ by Dynkin's theorem. □

In particular, two probability measures that coincide on a family of sets closed under intersection coincide on the $\sigma-$algebra generated by this family.

2.6.3. *Coincidence sets of arbitrary positive measures.*

Proposition. *Let μ, ν be measures on the same measurable space $(X, \mathcal{M})$. Assume that the two measures coincide and are finite valued on a set $\mathcal{E} \subseteq \mathcal{M}$ closed under intersection. Then the two measures coincide on all the sets of the $\sigma-$algebra $\mathcal{M}(\mathcal{E})$ generated by $\mathcal{E}$ which can be covered by countably many sets of $\mathcal{E}$. In particular, if X has a countable cover by sets in $\mathcal{E}$, then the two measures coincide on all $\mathcal{M}(\mathcal{E})$.*

Proof. For every $E \in \mathcal{E}$ we set $\mathcal{E}_E = \{F \cap E : F \in \mathcal{E}\}$ and $\mathcal{M}_E = \{A \cap E : A \in \mathcal{M}(\mathcal{E})\}$. Then $\mathcal{M}_E$ is a $\sigma-$algebra of parts of E, and we know that $\mathcal{M}_E$ is exactly the $\sigma-$algebra of parts of E generated by $\mathcal{E}_E$ (see 2.3.1). By 2.6.2 we have $\mu(G) = \nu(G)$ for every $G \in \mathcal{M}_E$. That is, we have $\mu(G) = \nu(G)$ for every $G \in \mathcal{M}(\mathcal{E})$ that is contained in some member of $\mathcal{E}$. If $A \in \mathcal{M}(\mathcal{E})$ and $A = \bigcup_{k\in\mathbb{N}} A_k$, with $A_k \subseteq E_k \in \mathcal{E}$, then with the usual technique A can be written as a disjoint union of sets $F_k \subseteq E_k$, with $F_k = A_k \setminus \left(\bigcup_{j=0}^{k-1} A_j\right) \in \mathcal{M}(\mathcal{E})$, so that $\mu(F_k) = \nu(F_k)$ and hence also $\mu(A) = \nu(A)$. □

The following corollary is then immediate:

Corollary. *Let Ω be an open subset of $\mathbb{R}^n$. Two measures on $\mathcal{B}(\Omega)$ that coincide and are finite on compact intervals contained in Ω coincide on all Borel subsets of Ω.*

Of course we could get the corollary also from the uniqueness part of the Carathèodory extension; but the above proof is much neater and more direct.

2.7. Lebesgue measure and the topology of $\mathbb{R}^n$. There is an intimate connection between Lebesgue measure on $\mathbb{R}^n$ and the topology of $\mathbb{R}^n$. This is true not only for Lebesgue measure, but more generally for Radon measures (see also 2.2.5). We call Radon measure on $\mathbb{R}^n$ any measure defined on Borel subsets of $\mathbb{R}^n$ that is finite on compact subsets of $\mathbb{R}^n$. Since $\mathbb{R}^n$ is a countable union of compact intervals (e.g $\mathbb{R}^n = \bigcup_{m=1}^{\infty}[-m,m]^n$) any Radon measure is the (restriction to Borel sets of the) Carathèodory extension of a Radon premeasure on intervals (see 2.5.5). We call $\mu-$measurable any set measurable in the Carathèodory extension of μ, and we still denote μ this extension. Remember that all sets of outer measure 0 are measurable. And remember that the Lebesgue measure on $\mathbb{R}^n$ is the measure obtained by Carathèodory extension of the premeasure on intervals defined by $\lambda_n(I_1\times\cdots\times I_n)=\prod_{k=1}^n\lambda_1(I_k)$; the $\sigma-$algebra $\mathcal{L}_n$ of Lebesgue measurable subsets is then the set of all $A\subseteq\mathbb{R}^n$ which split additively every other set. All singletons and all countable subsets of $\mathbb{R}^n$ have Lebesgue measure 0. In $\mathbb{R}^n$ every coordinate hyperplane has Lebesgue measure 0 (the set $\{x\in\mathbb{R}^n : x_k=c\}$ is a degenerate interval, product of copies of $\mathbb{R}$ on every place, with the degenerate interval $\{c\}$ in the k^{th} place).

Exercise 2.7.1. Compute the measure $\lambda_n(A)$ of the set

$$A=\{x\in\mathbb{R}^n : \text{ at least one coordinate of } x \text{ is a rational number}\}.$$

Solution. The set $A_1=\mathbb{Q}\times\mathbb{R}^{n-1}$ has measure 0, being a countable union of the sets $\{q\}\times\mathbb{R}^{n-1}$, $q\in\mathbb{Q}$. So the set A_1 of all points in $\mathbb{R}^n$ whose first coordinate is rational has measure 0. Similarly, if A_k is the set of all points whose $k-$th coordinate is rational we have $\lambda_n(A_k)=0$. Since $A=\bigcup_{k=1}^n A_k$ is a finite union of sets of measure 0, we also have $\lambda_n(A)=0$. □

2.7.1. *Outer regularity.*

. *If μ is Radon measure on $\mathbb{R}^n$, then for every measurable set A we have*

$$\mu(A)=\inf\{\mu(U) : U\supseteq A,\ U \text{ open in } \mathbb{R}^n\}.$$

Proof. Clearly only the case $\mu(A)<\infty$ needs proof. Given $\varepsilon>0$ we can find a countable cover $(I_k)_{k\in\mathbb{N}}$ of A by intervals such that $\sum_{k=0}^{\infty}\mu(I_k)\le\mu(A)+\varepsilon/2$; we can assume these intervals to be bounded (every interval is a countable disjoint union of bounded intervals); then for every k we can find a bounded open interval J_k containing I_k such that $\mu(J_k)\le\mu(I_k)+\varepsilon/2^{k+2}$ (2.2.5). If $U=\bigcup_{k=0}^{\infty}J_k$ we have that U is open and

$$\mu(U)\le\sum_{k=0}^{\infty}\mu(J_k)\le\sum_{k=0}^{\infty}(\mu(I_k)+\varepsilon/2^{k+2})=\sum_{k=0}^{\infty}\mu(I_k)+\frac{\varepsilon}{2}\le\mu(A)+\varepsilon.$$

Remark. We have really proved that $\mu^*(A)=\inf\{\mu^*(U) : U\supseteq A,\ U \text{ open in } \mathbb{R}^n\}$ for every subset A of $\mathbb{R}^n$, if $\mu^*:\mathcal{P}(\mathbb{R}^n)\to[0,\infty]$ is the outer measure associated to μ.

□

2.7.2. *Regularity.* Notice that since $A\subseteq U$, when $\mu(A)<\infty$ we have proved that given $\varepsilon>0$ we have an open set U containing A such that $\mu(U\setminus A)\le\varepsilon$. Since the measure μ is $\sigma-$finite this limitation can be removed, that is: given any measurable $A\subseteq\mathbb{R}^n$ and $\varepsilon>0$ there is an open set $U\supseteq A$ such that $\mu(U\setminus A)\le\varepsilon$: simply write A as a disjoint union $\bigcup_{k=0}^{\infty}A_k$ of measurable sets of finite measure, and for each k pick an open $U_k\supseteq A_k$ such that $\mu(U_k\setminus A_k)\le\varepsilon/2^{k+1}$. Then $U=\bigcup_{k=0}^{\infty}U_k$ is open, $U\supseteq A$ and $U\setminus A\subseteq\bigcup_{k=0}^{\infty}(U_k\setminus A_k)$ so that $\mu(U\setminus A)\le\sum_{k=0}^{\infty}\mu(U_k\setminus A_k)\le\sum_{k=0}^{\infty}\varepsilon/2^{k+1}=\varepsilon$.

From the inside we can use closed sets:

Lemma. *If μ is Radon measure on $\mathbb{R}^n$, then for every measurable set A and $\varepsilon>0$ there exist an open set U containing A and a closed set F contained in A such that $\mu(U\setminus F)\le\varepsilon$.*

Proof. Pick first an open set U containing A with $\mu(U\setminus A)\le\varepsilon/2$; next an open V containing $B=\mathbb{R}^n\setminus A$ and $\mu(V\setminus B)\le\varepsilon/2$. Then $F=\mathbb{R}^n\setminus V\subseteq A$ and $U\setminus F\subseteq(U\setminus A)\cup(V\setminus B)$ so that by subadditivity we have $\mu(U\setminus F)\le\mu(U\setminus A)+\mu(V\setminus B)\le\varepsilon$. □

In a topological space a subset G is said to be a $G_\delta-$set if it is a countable intersection of open sets, a subset E is an $F_\sigma-$set if it is a countable union of closed sets, equivalently its complement is a $G_\delta-$set. Clearly G_δ and F_σ sets are all Borel sets.

THEOREM. *Let μ be a Radon measure on $\mathbb{R}^n$. For every $\mu-$measurable set A there are a $G_\delta-$set G containing A and an $F_\sigma-$set E contained in A such that $\mu(G \setminus E) = 0$.*

Proof. For each k we pick open sets U_k and closed sets F_k such that $F_k \subseteq A \subseteq U_k$ and $\mu(U_k \setminus F_k) \le 1/(k+1)$. If $U = \bigcap_{k=0}^{\infty} U_k$ and $E = \bigcup_{k=0}^{\infty} F_k$ we have $E \subseteq A \subseteq G$ and $G \setminus E \subseteq U_k \setminus F_k$ for every k so that $\mu(G \setminus E) \le \mu(U_k \setminus F_k) \le 1/(k+1)$ for every k, and hence $\mu(G \setminus E) = 0$. □

The preceding proposition implies the important fact that $\mu-$measurable sets are all of the form $E \cup N$, where E is a Borel set, and N has (outer) $\mu-$measure zero, in particular Lebesgue measurable sets are of this form. And it is important to notice explicitly that

. *A subset N of $\mathbb{R}^n$ has Lebesgue measure zero if and only if for every $\varepsilon > 0$ there exists a countable cover $(I_k)_k$ of N by intervals ($N \subseteq \bigcup_{k=0}^{\infty} I_k$) such that $\sum_{k=0}^{\infty} \lambda_n(I_k) \le \varepsilon$.*

These intervals may also be assumed open and bounded.

2.7.3. *Inner regularity.*

. *Every Radon measure μ on $\mathbb{R}^n$ is inner regular. That is, for every $\mu-$measurable set A we have*

$$\mu(A) = \sup\{\mu(K) : K \subseteq A, K \text{ compact}\}.$$

Proof. From the preceding theorem we clearly have $\mu(A) = \sup\{\mu(F) : F \subseteq A, F \text{ closed in } \mathbb{R}^n\}$. And for every closed set F we have $\mu(F) = \sup\{\mu(F \cap [-k,k]^n) : k = 1,2,3,\dots\}$; the sets $F \cap [-k,k]^n$ are clearly compact. □

2.7.4. *A compact set of positive Lebesgue measure with empty interior.* We have seen that measurable sets can be approximated from the outside with open sets, from the inside with closed or even compact sets. The opposite is not in general possible : there are sets of positive Lebesgue measure which have empty interior, as we now see. Order the rationals of $[0,1]$ in a sequence $(x_k)_k$. Fix $\varepsilon > 0$; for each k consider the open interval $I_k =]x_k - \varepsilon/2^{k+2}, x_k + \varepsilon/2^{k+2}[$, and set

$$K = [0,1] \setminus \bigcup_{k=0}^{\infty} I_k \cap [0,1].$$

Then K is compact (it is closed and contained in $[0,1]$). And

$$\lambda_1(K) = 1 - \lambda_1\left(\bigcup_{k=0}^{\infty} I_k \cap [0,1]\right) \ge 1 - \sum_{k=0}^{\infty} \lambda_1(I_k) \ge 1 - \sum_{k=0}^{\infty} \frac{\varepsilon}{2^{k+1}} = 1 - \varepsilon$$

Thus with ε small we can make the measure of K very close to 1. It is clear that the interior of K is empty, since K contains no rational number, and rationals are dense in $\mathbb{R}$.

Of course analogous constructions can be carried out in $\mathbb{R}^n$. And one cannot approximate from the outside with closed sets: e.g. the rationals have measure 0 but are dense, so the only closed set containing them is $\mathbb{R}$.

2.7.5. *Elementary measure and Lebesgue measure.* For definiteness we discuss area, two–dimensional measure in $\mathbb{R}^2$; it will be clear that analogous things can be said for $n-$dimensional measure. Our definition of area is intuitively much less appealing than the one given by, say, Archimedes, and preserved more or less unchanged until the beginning of 1900, when Lebesgue and others created the modern theory of measure.

We consider a bounded subset E of $\mathbb{R}^2$; consider the class $\mathcal{A}_*(E)$ of all elements of $A \in \mathcal{A}_2$ (two–dimensional plurintervals) contained in E, and the class $\mathcal{A}^*(E)$ of all $B \in \mathcal{A}_2$ containing E. Clearly we have $\lambda_2(A) \le \lambda_2(B)$ for every $A \in \mathcal{A}_*(E)$ and every $B \in \mathcal{A}^*(E)$; we say that E is elementarily measurable (or measurable according to Peano–Jordan) if

$$\sup\{\lambda_2(A) : A \in \mathcal{A}_*(E)\} = \inf\{\lambda_2(B) : B \in \mathcal{A}^*(E)\},$$

and of course this common value is, by definition, the *area* of E. It is not difficult to prove that if E is elementarily measurable then it is Lebesgue measurable and its area is $\lambda_2(E)$ (we can find sequences $A_j \in \mathcal{A}_*(E)$ and $B_j \in \mathcal{A}^*(E)$ such that A_j is increasing, B_j is decreasing and $\lambda_2(B_j \setminus A_j) \le 1/2^j$; then if $A = \bigcup_j A_j$, $B = \bigcap_j B_j$ we have $A \subseteq E \subseteq B$ and $\lambda_n(B \setminus A) = 0$, with A and B both Borel sets). But every set with empty interior will have $\sup\{\lambda_2(A) : A \in \mathcal{A}_*(E)\} = 0$ (a rectangle with empty interior has area zero), and we know that there are compact sets with empty interior and strictly positive measure (see 2.7.4), so that there are Lebesgue measurable sets not elementarily measurable. If we want to measure all open and all closed subsets, which seems to be highly desirable, we are forced to abandon approximation from the inside by rectangles (we still have approximation from inside by *compact sets*).

2.7.6. *Vitali's example of a non–measurable subset of* $\mathbb{R}$. From a practical point of view all subsets of $\mathbb{R}^n$ are Lebesgue measurable: all ordinary constructions for building new sets out of old ones do not make us leave the class of measurable sets. But in the most common version of set theory used in Mathematics today nonmeasurable sets do exist. Using the axiom of choice in an essential way, the italian mathematician Vitali was the first to give an example of a non measurable subset of $\mathbb{R}$ (example easily adapted to $\mathbb{R}^n$). The example is also based on the important fact that Lebesgue measure on $\mathbb{R}^n$ is *translation invariant*: that is, for every $a \in \mathbb{R}^n$ the translation $x \mapsto a + x$ transforms measurable sets into measurable sets of the same measure (this is immediate, from the fact that the statement is true for intervals).

Consider in $\mathbb{R}$ the congruence modulo the additive subgroup $\mathbb{Q}$ of the rational numbers; its classes are $x + \mathbb{Q}$, translates of $\mathbb{Q}$, and so they are all dense in $\mathbb{R}$, so that they have non–empty intersection with every nondegenerate interval of $\mathbb{R}$. We may then find a set E, subset of $[0, 1[$, such that $E \cap (x + \mathbb{Q})$ consists of exactly one element, for every class $x + \mathbb{Q}$: the existence of such a set is ensured by the axiom of choice. Observe now that if $r, s \in \mathbb{Q}$ are distinct, then $r + E$ and $s + E$ are disjoint ($r + x = s + y \iff x - y = s - r \in \mathbb{Q}$, which is possible iff $x = y$, but then $r = s$). If we consider the set A of all rational numbers contained in $]-1, 1[$ we have

$$[0, 1[\subseteq \bigcup_{r\in A}(r + E) \subseteq]-1, 2[$$

The second inclusion is immediate (if $-1 < r < 1$ and $0 \le x < 1$ then $-1 < r + x < 2$); for the first, given $a \in [0, 1[$, we know that there is exactly one $x \in E$ such that $r = a - x \in \mathbb{Q}$; since $0 \le a < 1$ and $0 \le x < 1$ we get $-1 < a - x < 1$. Then E cannot have measure 0: by monotonicity $1 = \lambda_1([0, 1[) \le \lambda_1\left(\bigcup_{r\in A}(r + E)\right)$ and by countable additivity $\lambda_1\left(\bigcup_{r\in A}(r + E)\right) = \sum_{r\in A}\lambda_1(r + E) = \sum_{r\in A}\lambda_1(E)$, the last by translation invariance. Then $\lambda_1(E) > 0$, which implies $\lambda_1\left(\bigcup_{r\in A}(r + E)\right) = \infty$. On the other hand $\lambda_1\left(\bigcup_{r\in A}(r + E)\right) \le \lambda_1(]-1, 2[) = 3$, so that $\bigcup_{r\in A}(r + E)$ has finite measure. The contradiction proves non–measurability of E.

Exercise 2.7.2. With E as above, prove that all Lebesgue measurable subsets of E have measure zero.

Exercise 2.7.3. Prove the:

. Difference theorem *If $A \subseteq \mathbb{R}^n$ is Lebesgue measurable and $\lambda_n(A) > 0$, then*

$$A - A := \{x - y : x \in A, y \in A\}$$

contains a neighborhood of 0 *in* $\mathbb{R}^n$.

Proof. Follow these hints:

(i) Prove that there is a compact $K \subseteq A$ such that $\lambda_n(K) > 0$ (we shall prove that $K - K$ contains a nbhd of 0).
(ii) Prove that there is an open subset U of $\mathbb{R}^n$ such that $K \subseteq U$ and $\lambda_n(U) < 2\lambda_n(K)$.
(iii) It is well–known that the function $x \mapsto \operatorname{dist}(x, \mathbb{R}^n \setminus U)$ is continuous; let $\delta = \min\{\operatorname{dist}(x, \mathbb{R}^n \setminus U) : x \in K\}$; then $\delta > 0$, and if $|x| < \delta$ then $x + K \subseteq U$.
(iv) Observe that $K - K = \{x \in \mathbb{R}^n : (x + K) \cap K \ne \emptyset\}$ and conclude.

□

Solution. (i) Trivial, by inner regularity. (ii) Since $0 < \lambda_n(K) < \infty$ we have $\lambda_n(K) < 2\lambda_n(K)$; then U exists by outer regularity of K.

(iii) That $\delta > 0$ is clear: the minimum of the distance function on K exists because this function is continuous and K is compact; and the minimum cannot be 0, because the zero–set of the function is exactly $\mathbb{R}^n \setminus U$, disjoint from K. Then $x + K \subseteq U$ if $|x| < \delta$: if $y = x + a$ with $a \in K$, then $|y - a| = |x| < \delta$, so that $y \notin \mathbb{R}^n \setminus U$.

(iv) We have: $(x + K) \cap K \ne \emptyset \iff$ (there exist $a, b \in K$ such that $x + a = b \iff x = b - a$), exactly the definition of $K - K$. And if for some x we have $(x + K) \cap K = \emptyset$ then we have $\lambda_n((x + K) \cup K) = \lambda_n(x + K) + \lambda_n(K) = \lambda_n(K) + \lambda_n(K) = 2\lambda_n(K) > \lambda_n(U)$. Then it is not possible that $(x + K) \cup K \subseteq U$; since $K \subseteq U$ we have that $x + K \subseteq U$ implies $(x + K) \cap K \ne \emptyset$; since in (iii) we proved that $x + K \subseteq U$ if $|x| < \delta$ we get that the open ball of center the origin and radius δ is contained in $K - K$. □

Exercise 2.7.4. We want to prove (generalizing Vitali's example) that in $\mathbb{R}^n$ there is a subset A such that every Lebesgue measurable subset of A has measure 0, and every measurable subset of the complement $B = \mathbb{R}^n \setminus A$ has also measure 0 (prove that necessarily such a set cannot be measurable, and that the outer measures of A and B are both strictly positive). Let $a \in \mathbb{R}^n$ be a vector with at least one irrational coordinate. Set $G = \mathbb{Q}^n + \mathbb{Z}\, a$, $G_0 = \mathbb{Q}^n + 2\mathbb{Z}\, a$, $G_1 = G_0 + a$.

(i) G and G_0 are additive subgroups of $\mathbb{R}^n$, $G = G_0 \cup G_1$ (disjoint union), and G_1 is dense in $\mathbb{R}^n$.

Let $E \subseteq \mathbb{R}^n$ be a set of representatives of $\mathbb{R}^n/G$ ((that is, $E \cap (x + G)$ contains exactly one element out of each coset $x + G$ of $\mathbb{R}^n/G$); set $A = E + G_0$.

(ii) Prove that $B = \mathbb{R}^n \setminus A = E + G_1$.
(iii) Prove that $A - A := \{x - y : x, y \in A\}$ and $B - B = \{x - y : x, y \in B\}$ do not intersect G_1.
(iv) Using the difference theorem (2.7.3) prove that neither A nor B contain measurable sets of nonzero measure.
(v) Prove that every $E \subseteq \mathbb{R}^n$ of strictly positive outer measure contains a non–measurable subset.

3. Measurable functions

Recall that a measurable space is a pair $(X, \mathcal{M})$ formed by a set X and $\sigma-$algebra of parts of X. The relevant maps between such spaces are the measurable maps, which we now define.

DEFINITION. Let $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ be measurable spaces. A function $f : X \to Y$ is said to be *measurable* if $f^{\leftarrow}(V) \in \mathcal{M}$ for every $V \in \mathcal{N}$ (that is, the inverse image of every measurable set in the range is measurable in the domain).

The theory of measurable functions has several analogies (and differences ...) with the theory of continuous functions; recall that a map $f : X \to Y$ of topological spaces (X, τ_X) and (Y, τ_Y) is continuous iff $f^{\leftarrow}(V) \in \tau_X$ for every $V \in \tau_Y$.

3.0.7. *First results.* Some easy remarks:

. *Constant maps are measurable. Compositions of measurable maps are measurable. Restrictions of measurable maps are measurable.*

Proof. Assume that $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ are measurable spaces, and $f : X \to Y$ is a function. If f is constant the inverse image of a subset of Y is either empty or X. And if $(Z, \mathcal{P})$ is a third measure space and $f : X \to Y$ and $g : Y \to Z$ are measurable then $(g \circ f)^{\leftarrow}(W) = f^{\leftarrow}(g^{\leftarrow}(W)) \in \mathcal{M}$ for every $W \in \mathcal{P}$ since $g^{\leftarrow}(W) \in \mathcal{N}$, by measurability of g. And if $S \subseteq X$, then $f_{|S} : S \to Y$ is measurable with respect to the induced $\sigma-$algebra $\mathcal{M}_S = \{B \cap S : B \in \mathcal{M}\}$. □

EXERCISE 3.0.5. Let $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ be measurable spaces, $f : X \to Y$ a function. Let $(X_k)_{k \in \mathbb{N}}$ be a countable cover of X by measurable sets (i.e. $X_k \in \mathcal{M}$ for every $k \in \mathbb{N}$ and $X = \bigcup_{k \in \mathbb{N}} X_k$). Then f is measurable if and only if $f_{|X_k}$ is measurable, for every $k \in \mathbb{N}$.

3.0.8. *Changing $\sigma-$algebras.* As with continuity in topological spaces, measurability of $f : X \to Y$, with $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ measurable spaces is the more significant the largest is the $\sigma-$algebra on the range space, and the smaller is the $\sigma-$algebra on the domain space. In other words, if $f : X \to Y$ is measurable, it remains measurable when the $\sigma-$algebra $\mathcal{M}$ on the domain is enlarged to a bigger $\sigma-$algebra, and also if $\mathcal{N}$ on the target space is replaced by a smaller $\sigma-$algebra; at the extreme, if $\mathcal{M} = \mathcal{P}(X)$ then every f with domain X is measurable, whatever the range space $(Y, \mathcal{N})$; and symmetrically, if $\mathcal{N} = \{\emptyset, Y\}$ is the trivial $\sigma-$algebra, then every $f : X \to Y$ is measurable, whatever the domain $(X, \mathcal{M})$. This elementary observation motivates some of the subsequent definitions.

3.0.9. *Final $\sigma-$algebra.* Assume that $(X, \mathcal{M})$ is a measurable space, Y is a set, and $f : X \to Y$ a function. It is immediate to see that

$$\mathcal{Q}_f = \mathcal{Q} = \{V \subseteq Y : f^{\leftarrow}(V) \in \mathcal{M}\},$$

is a $\sigma-$algebra of parts of Y. It is the largest possible $\sigma-$algebra on Y which makes f measurable, and is called the final, or quotient $\sigma-$algebra of $\mathcal{M}$ *via* f. It has the following property:

. *With notations and terminology as above, let $g : Y \to Z$ be a function, where $(Z, \mathcal{P})$ is a measurable space. If $g \circ f : X \to Z$ is measurable, then g is $(Y, \mathcal{Q}_f)-$measurable.*

Proof. For every $W \in \mathcal{P}$ we have by measurability of $g \circ f$ that $(g \circ f)^{\leftarrow}(W) = f^{\leftarrow}(g^{\leftarrow}(W)) \in \mathcal{M}$, and this is equivalent to say that $g^{\leftarrow}(W) \in \mathcal{Q}_f$. □

3.0.10. *A sufficient condition for measurability.* Notice that $f : (X, \mathcal{M}) \to (Y, \mathcal{N})$ is measurable if and only if $\mathcal{Q}_f \supseteq \mathcal{N}$. Then

PROPOSITION. *Assume that $f : (X, \mathcal{M}) \to (Y, \mathcal{N})$ is a map, and that $\mathcal{E}$ is a set of generators for $\mathcal{N}$. Then f is measurable if and only if $f^{\leftarrow}(E) \in \mathcal{M}$ for every $E \in \mathcal{E}$.*

Proof. The condition says that $\mathcal{E} \subseteq \mathcal{Q}_f$, where $\mathcal{Q}_f$ is the final $\sigma-$algebra of f. Then $\mathcal{N} = \mathcal{M}(\mathcal{E}) \subseteq \mathcal{Q}_f$, which implies measurability of the original f. □

The preceding proposition is constantly used. A very important class of measurable functions will be that of real valued functions from a measurable space $(X, \mathcal{M})$ into $\mathbb{R}$, where, unless otherwise specified, it is understood that the $\sigma-$algebra is that $\mathcal{B}(\mathbb{R})$ of Borel subsets of $\mathbb{R}$. To check that $f : X \to \mathbb{R}$ is measurable, it is enough to check that strict upper Lebesgue sets as $\{f > \alpha\} := \{x \in X : f(x) > \alpha\}$ belong to $\mathcal{M}$, for every $\alpha \in \mathbb{R}$ (or even only for every $\alpha \in \mathbb{Q}$); analogously for strict lower Lebesgue sets $\{f < \alpha\}$, etc. (2.3). Same for extended real valued functions $f : X \to \tilde{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$: recall that the Borel sets of $\tilde{\mathbb{R}}$ are those of $\mathbb{R}$, united with a subset of $\{-\infty, \infty\}$. A very important corollary is:

Corollary. *Let (X, τ_X) and (Y, τ_Y) be topological spaces, and let $f : X \to Y$ be continuous. Then $f : (X, \mathcal{B}(X)) \to (Y, \mathcal{B}(Y))$ is measurable.*

Proof. For every $V \in \tau_Y$ we have that $f^{\leftarrow}(V) \in \tau_X \subseteq \mathcal{B}(X)$, and τ_Y generates $\mathcal{B}(Y)$. □

Of course there are many more Borel measurable functions than continuous functions: observe that a subset A of a measurable space X is measurable if and only if its characteristic function $\chi_A : X \to \mathbb{R}$ is measurable, but in a topological space a characteristic function is continuous iff it is the characteristic function of an open–and–closed set. As an important example we single out the following one:

Exercise 3.0.6. The sign function $\mathrm{sgn} : \mathbb{C} \to \mathbb{C}$ is defined by $\mathrm{sgn}(0) = 0$, and $\mathrm{sgn}\, z = z/|z|$ for $z \neq 0$. Prove that sgn is measurable (no specification of $\sigma-$algebras means that Borel $\sigma-$algebras are used, both in domain and in range).

Solution. Let V be an open subset of $\mathbb{C}$. If $0 \notin V$ then $\mathrm{sgn}^{\leftarrow}(V) = (\mathrm{sgn}_{|\mathbb{C}_*})^{\leftarrow}(V)$ is an open subset of $\mathbb{C}_* = \mathbb{C} \setminus \{0\}$, the punctured plane, since $\mathrm{sgn}_{|\mathbb{C}_*}$ is a continuous self–map of $\mathbb{C}_*$. And if $0 \in V$, then $\mathrm{sgn}^{\leftarrow}(V) = \{0\} \cup \mathrm{sgn}^{\leftarrow}(V \setminus \{0\})$ is Borel subset of $\mathbb{C}$ (clearly sgn is not continuous at 0: the inverse image of the open unit disk $\Delta = \{w \in \mathbb{C} : |w| < 1\}$ is $\{0\}$, not a neighborhood of 0!) □

3.0.11. *Initial $\sigma-$algebras.* Assume now that X is a set, that $(Y_\lambda, \mathcal{N}_\lambda)_{\lambda \in \Lambda}$ is a family of measurable spaces, and that for each $\lambda \in \Lambda$ we have a function $f_\lambda : X \to Y_\lambda$. There is a smallest $\sigma-$algebra $\mathcal{M}$ on X among those that make all functions f_λ measurable. This $\sigma-$algebra may be described in the following way: for each $\lambda \in \Lambda$ consider a set $\mathcal{B}_\lambda$ of generators for $\mathcal{N}_\lambda$ (which might be of course all of $\mathcal{N}_\lambda$). Then $\mathcal{M}$ is the $\sigma-$algebra on X generated by all sets of the form $f_\lambda^{\leftarrow}(V_\lambda)$, with $V_\lambda \in \mathcal{B}_\lambda$ and $\lambda \in \Lambda$. The most important application will be to the case of product spaces. But even the case of a single map $f : X \to Y$ has interest. For instance we can have $X = S \subseteq Y$ and $f = j_S : S \to Y$ the canonical injection. In this case we have of course $j_S^{\leftarrow}(V) = V \cap S$, and the initial $\sigma-$algebra is the $\sigma-$algebra $\mathcal{N}_S = \{V \cap S : V \in \mathcal{N}\}$, already discussed in 3.0.7.

3.0.12. *Product of a finite family of measurable spaces.* Given two measurable spaces $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ we consider the product set $X \times Y$, and we put on $X \times Y$ the initial $\sigma-$algebra determined by the canonical projections $p_X : X \times Y \to X$ given by $p_X(x, y) = x$ and $p_Y : X \times Y \to Y$ given by $p_Y(x, y) = y$. This $\sigma-$algebra is denoted by $\mathcal{M} \otimes \mathcal{N}$ (read em tensor en), and is generated by all *measurable cylinders* $A \times Y = p_X^{\leftarrow}(A)$ and $X \times B = p_Y^{\leftarrow}(B)$, with $A \in \mathcal{M}$ and $B \in \mathcal{N}$ (or A varying in system of generators $\mathcal{E}$ for $\mathcal{M}$, same for B, $\mathcal{F}$ and $\mathcal{N}$. Of course finite intersections of such sets are also a generating system; and if $A \subseteq X$ and $B \subseteq Y$ we have

$$A \times B = (A \times Y) \cap (X \times B) = p_X^{\leftarrow}(A) \cap p_Y^{\leftarrow}(B).$$

If $\mathcal{E}$ and $\mathcal{F}$ are semialgebras which generate $\mathcal{M}$ and $\mathcal{N}$ respectively we have seen that the set of all products $\mathcal{G} = \{A \times B : A \in \mathcal{E}, B \in \mathcal{F}\}$ is a semialgebra of parts of $X \times Y$ (see 1.4.2); this semialgebra generates $\mathcal{M} \otimes \mathcal{N}$. All this can be repeated with obvious changes for a product of a finite family $((X_k, \mathcal{M}_k))_{k \in \{1,\dots,m\}}$ of measurable spaces: the product $X = \prod_{k=1}^m X_k$, set of all $m-$tuples $(x_1, \dots, x_m)$ with $x_k \in X_k$ is equipped with the $\sigma-$algebra generated by the sets $A_1 \times \cdots \times A_m = \bigcap_{k=1}^m p_k^{\leftarrow}(A_k)$, with A_k varying in a system of generators for $\mathcal{M}_k$

To specify a function $f : T \to X = \prod_{k=1}^m X_k$ arriving to a product, we can specify its *component functions* $f_k = p_k \circ f$. Products are characterized by the following property: a function which arrives to a product is measurable if and only if its component functions are measurable:

. *Let $(T, \mathcal{T})$ be a measurable space, let $((X_k, \mathcal{M}_k))_{k \in \{1,\dots,m\}}$ be a family of measurable spaces, and let $f : T \to X = \prod_{k=1}^m X_k$ be a function. Then f is measurable if and only if all its component functions $f_k = p_k \circ f$, $k = 1, \dots, m$ are measurable.*

Proof. Necessity is obvious, since compositions of measurable maps are measurable. And sets like $p_k^{\leftarrow}(A_k)$ are a system of generators for $\bigotimes_{k=1}^m \mathcal{M}_k$; since for every such set $f^{\leftarrow}(p_k^{\leftarrow}(A_k)) = (p_k \circ f)^{\leftarrow}(A_k)$ is $\mathcal{T}-$measurable, by 3.0.10 f is measurable. □

3.0.13. *Sections.* To simplify notations we limit ourselves to the product of two measurable spaces $(X, \mathcal{M})$ and $(Y, \mathcal{N})$. Having a third space $(Z, \mathcal{P})$ and a function $f : X \times Y \to Z$ we have the sections of f, functions obtained by keeping one variable fixed and letting the other change. Given $y \in Y$, the $y-$section of f is the function $f(\#, y) : X \to Z$ given by $x \mapsto f(x, y)$; similarly, given $x \in X$, the $x-$section $f(x, \#) : Y \to Z$ is defined by $y \mapsto f(x, y)$. The next result is very important for the theory of multiple integration.

PROPOSITION. *Let $(X, \mathcal{M})$,$(Y, \mathcal{N})$ and $(Z, \mathcal{P})$ be measurable spaces, and assume that*

$$f : (X \times Y, \mathcal{M} \otimes \mathcal{N}) \to (Z, \mathcal{P}) \quad \textit{is measurable.}$$

Then all sections $f(x, \#)$ and $f(\#, y)$ of f are measurable.

Proof. Given $x \in X$ we have the injection $j^x : Y \to X \times Y$ given by $j^x(y) = (x, y)$. This map is $(\mathcal{N}, \mathcal{M} \otimes \mathcal{N})$ measurable (its components are the maps constantly x and the identity of Y, both measurable). Then the composition $f \circ j^x : Y \to Z$ is measurable; this composition is exactly the $x-$section $f(x, \#)$. Similarly for the $y-$section. □

Given $a \in X$, by $a-$section of a subset $S \subseteq X \times Y$ we mean the set $S(a) = \{y \in Y : (a, y) \in S\}$, of which the function $y \mapsto \chi_S(a, y)$ is the characteristic function; similarly for $b \in Y$.

EXERCISE 3.0.7. Let $(X \times Y, \mathcal{M} \otimes \mathcal{N})$ be the product of the measurable spaces $(X, \mathcal{M})$ and $(Y, \mathcal{N})$. Let A, B be subsets of X and Y respectively, both non empty. Prove that $A \times B$ is $\mathcal{M} \otimes \mathcal{N}-$measurable if and only if $A \in \mathcal{M}$ and $B \in \mathcal{N}$. Prove that $\mathcal{M}$ and $\mathcal{N}$ are the final $\sigma-$algebras of p_X and p_Y.

Solution. If $A \times B \in \mathcal{M} \otimes \mathcal{N}$ then for every $a \in A$ the $a-$section of $A \times B$, namely B, belongs to $\mathcal{N}$; similarly for $b \in B$ the $b-$section A of $A \times B$ belongs to $\mathcal{M}$. Then, if $p_X^{\leftarrow}(A) = A \times Y \in \mathcal{M} \otimes \mathcal{N}$ we have $A \in \mathcal{M}$, proving that $\mathcal{M}$ is the final $\sigma-$algebra of $\mathcal{M} \otimes \mathcal{N}$ by p_X; similarly for p_Y. □

3.0.14. *Borel sets in a product of topological spaces.* Recall that if (X, τ_X) is a topological space a base for the open sets of X is subset $\mathcal{E}$ of open sets such that every open set of the space is a union of elements of $\mathcal{E}$. Since $\mathcal{E} \subseteq \tau_X$ we have $\mathcal{M}(\mathcal{E}) \subseteq \mathcal{B}(X) = \mathcal{M}(\tau_X)$. If $\mathcal{E}$ is a countable set then of course we have $\tau_X \subseteq \mathcal{M}(\mathcal{E})$ and hence $\mathcal{M}(\mathcal{E}) = \mathcal{B}(X)$; but in general the inclusion is proper. If (X, τ_X) and (Y, τ_Y) are topological spaces, then the product topology on $X \times Y$ has the open rectangles $U \times V$ as a base for the open sets. It follows that $\mathcal{B}(X) \otimes \mathcal{B}(Y) \subseteq \mathcal{B}(X \times Y)$, and the inclusion might be proper (examples exist, see 3.0.15). If however τ_X and τ_Y have countable bases, say $\mathcal{E}_X$ and $\mathcal{E}_Y$, then $\{U \times V : U \in \mathcal{E}_X,\ V \in \mathcal{E}_Y\}$ is a base for the product topology of $X \times Y$, and a generating system for the $\sigma-$algebra $\mathcal{B}(X) \otimes \mathcal{B}(Y)$ as well, so that $\mathcal{B}(X) \otimes \mathcal{B}(Y) = \mathcal{B}(X \times Y)$. A topological space is said to be *second countable* if it has a countable base for the open sets; metrizable and separable spaces are second countable (a countable base for the open sets are the open balls of rational radii centered at the points of a countable dense set, see 1.3). In particular $\mathcal{B}(\mathbb{R}^2) = \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$, etc.

PROPOSITION. *If (X, τ_X), (Y, τ_Y) are second countable topological spaces and (Z, τ_Z) is a topological space, then every continuous function $f : X \times Y \to Z$ is $(\mathcal{B}(X) \otimes \mathcal{B}(Y), \mathcal{B}(Z))-$measurable.*

Proof. Because $\mathcal{B}(X) \otimes \mathcal{B}(Y) = \mathcal{B}(X \times Y)$. □

EXAMPLE 3.0.8. The functions $(x, y) \mapsto x \vee y = \max\{x, y\}$ and $(x, y) \mapsto x \wedge y = \min\{x, y\}$ are continuous from $\tilde{\mathbb{R}} \times \tilde{\mathbb{R}} \to \tilde{\mathbb{R}}$, hence measurable. Similarly the addition from $\mathbb{R} \times \mathbb{R}$ to $\mathbb{R}$, or $\mathbb{C} \times \mathbb{C}$ to $\mathbb{C}$, or $[0, \infty] \times [0, \infty] \to [0, \infty]$, or even $\tilde{\mathbb{R}} \times \tilde{\mathbb{R}} \setminus \{(\infty, -\infty), (-\infty, \infty)\}$ to $\tilde{\mathbb{R}}$ is continuous and hence measurable. Multiplication $\tilde{\mathbb{R}} \times \tilde{\mathbb{R}} \to \tilde{\mathbb{R}}$, extended by declaring that $0(\pm\infty) = (\pm\infty)0 = 0$, is not continuous, but it is Borel measurable.

3.0.15. *The Sorgenfrey plane.* The Sorgenfrey line S defined in Analisi Due, 2.4.15, is a space which is not second countable: if $\mathcal{S}$ is a base for the open sets of S, for every real number a there exists $U(a) \in \mathcal{S}$ such that $a \in U(a) \subseteq [a, a+1[$, in particular $\min U(a) = a$. If we consider the subset $\mathcal{U}$ of $\mathcal{S}$ consisting of all $U \in \mathcal{S}$ that have a minimum, we get a map $U \mapsto \min U$ from $\mathcal{U}$ to $\mathbb{R}$, which is onto $\mathbb{R}$. Then $|\mathbb{R}| \le |\mathcal{U}| \le |\mathcal{S}|$, proving that any base for the topology of S has at least cardinality $\mathfrak{c}$, the continuum. But one can prove that $\mathcal{B}(S) = \mathcal{B}(\mathbb{R})$: with a little effort one can in fact prove that every $S-$open set is a countable union of right half–open intervals. Then $\mathcal{B}(S) \otimes \mathcal{B}(S) = \mathcal{B}(\mathbb{R}^2)$. But the Borel tribe of the Sorgenfrey plane $S \times S$ is much larger than $\mathcal{B}(\mathbb{R}^2)$: the topology induced by $S \times S$ on the second diagonal $D = \{(x, -x) : x \in S\}$ is the discrete topology, so every subset of D is in $\mathcal{B}(S \times S)$, whereas $\mathcal{B}(\mathbb{R}^2)$ traces on D the Borel tribe (observe that $x \mapsto (x, -x)$ is a homeomorphism from $\mathbb{R}$ onto D, when D has the topology induced on it by $\mathbb{R}^2$ (usual).

3.1. Measurable real functions. We now consider the particularly important special case of extended real valued functions from a measurable space $(X, \mathcal{M})$ to the extended real line $\tilde{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$, equipped with the Borel $\sigma-$algebra. As observed in 3.0.10, such a function $f : X \to \tilde{\mathbb{R}}$ is measurable iff $\{f > \alpha\}$ (or $\{f < \alpha\}$, or $\{f \ge \alpha\}$ or $\{f \le \alpha\}$) is measurable for each $\alpha \in \mathbb{R}$ (or even only every $\alpha \in \mathbb{Q}$).

3.1.1. *Basic properties.*

. *Let $(X, \mathcal{M})$ be a measurable space.*

(i) *if $(f_n)_{n\in\mathbb{N}}$ is a sequence of extended real valued measurable functions, then the functions $\bigvee_{n\in\mathbb{N}} f_n$ and $\bigwedge_{n\in\mathbb{N}} f_n$, defined by*

$$\left(\bigvee_{n\in\mathbb{N}} f_n\right)(x) := \sup\{f_n(x) :\, n \in \mathbb{N}\}, \qquad \left(\bigwedge_{n\in\mathbb{N}} f_n\right)(x) = \inf\{f_n(x) :\, n \in \mathbb{N}\},$$

are measurable; and the functions f^*, f_*, defined by

$$f^*(x) = \limsup_{n\to\infty} f_n(x) \quad \textit{that is} \quad f^* = \bigwedge_{m\in\mathbb{N}}\left(\bigvee_{n\geq m} f_n\right)$$

$$f_*(x) = \liminf_{n\to\infty} f_n(x) \quad \textit{that is} \quad f_* = \bigvee_{m\in\mathbb{N}}\left(\bigwedge_{n\geq m} f_n\right)$$

are measurable; in particular, the limit of a pointwise convergent sequence of measurable functions is measurable.

(ii) *If $f, g : X \to \tilde{\mathbb{R}}$ are measurable, then $f \vee g$, $f \wedge g$, f^+, f^-, $|f|$ are measurable.*

(iii) *If $f, g : X \to \mathbb{R}$ are measurable, then $f + g$ and $f\,g$ are measurable.*

(iv) *$f : X \to \mathbb{C}$ is measurable iff $\operatorname{Re} f$ and $\operatorname{Im} f$ are measurable; and $|f|$, $\operatorname{sgn} f$ are measurable.*

Proof. (i) Writing for simplicity f in place of $\bigvee_n f_n$ we have, for every $\alpha \in \mathbb{R}$:

$$\{f > \alpha\} = \bigcup_{n\in\mathbb{N}} \{f_n > \alpha\}$$

(the supremum of a set of extended real numbers is strictly larger than some real number α if and only if some member of the set is strictly larger than α). Similarly

$$\left\{\bigwedge_n f_n < \alpha\right\} = \bigcup_{n\in\mathbb{N}} \{f_n < \alpha\}.$$

Measurability of liminf and limsup follows easily iterating this fact; and from measurability of these follows measurability of the limit, when the limit exists.

(ii) is a particular case of the first statement, for a two element set $\{f,\, g\}$ in place of a sequence. Notice that $|f| = f \vee (-f)$.

(iii) Consider the function $f \triangledown g : x \mapsto (f(x), g(x))$ from $X \to \mathbb{R}^2$; its component are measurable so that $f \triangledown g$ is $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ measurable; addition and multiplication, being continuous, are $(\mathcal{B}(\mathbb{R}\times\mathbb{R}), \mathcal{B}(\mathbb{R}))$ measurable; but since $\mathbb{R}$ is second countable we have $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R} \times \mathbb{R})$. Then $f + g$ and $f\,g$ are measurable, as compositions of measurable functions.

(iv) In the usual identification of $\mathbb{C}$ with $\mathbb{R}^2$ we have $\mathcal{B}(\mathbb{C}) = \mathcal{B}(\mathbb{R}^2) = \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$, Re and Im are the projections, so that the statement is just the proposition in 3.0.12. And if $\phi : \mathbb{C} \to \mathbb{C}$ is Borel measurable, then $\phi \circ f$ is measurable: the sign function is Borel measurable (see 3.0.6), and the absolute value function is continuous.

□

3.1.2. *Linear spaces of measurable functions.* Given a measurable space $(X, \mathcal{M})$ we denote by $\mathcal{L}^+_{\mathcal{M}}(X)$ or simply $\mathcal{L}^+(X)$ or even only $\mathcal{L}^+$ the set of all extended real valued positive measurable functions. We denote $\mathcal{L}_{\mathcal{M}}(X, \mathbb{K})$ (or simply $\mathcal{L}(X, \mathbb{K})$ or even only $\mathcal{L}(X)$ when the tribe $\mathcal{M}$ is understood) the set of all $\mathbb{K}-$valued measurable functions defined on X; this set is a vector space and an algebra of functions, and $\mathcal{L}(X, \mathbb{R})$ is also a lattice. We have seen that $\mathcal{L}(X)$ is closed under post-composition with Borel measurable functions, and under pointwise convergence of sequences,

Exercise 3.1.1. If $f : X \to \tilde{\mathbb{R}}$ is measurable, and $\phi : \tilde{\mathbb{R}} \to \tilde{\mathbb{R}}$ is Borel measurable (in particular continuous) then $\phi \circ f$ is measurable. Prove that if ϕ is a homeomorphism then measurability of $\phi \circ f$ implies measurability of f.

Remark. If ϕ is not a homeomorphism of $\tilde{\mathbb{R}}$ then $\phi \circ f$ can be measurable without f being measurable. For instance it can happen that $|f|$ is measurable and f is not measurable: e.g. take a non measurable set $A \subseteq X$ and define $f(x) = 1$ for $x \in A$, $f(x) = -1$ for $x \in X \setminus A$. Then f is non measurable but $|f|$ is the constant 1 and is then measurable.

EXERCISE 3.1.2. Let $(X, \mathcal{M})$ be a measurable space, and let $f,\, g : X \to \tilde{\mathbb{R}}$ be measurable functions. Prove that the set $\{g > f\} := \{x \in X : g(x) > f(x)\}$ is measurable.

Given a sequence $f_n : X \to \tilde{\mathbb{R}}$ of measurable functions, prove that the set

$$\{x \in X : \lim_{n\to\infty} f_n(x) \text{ exists in } \tilde{\mathbb{R}}\}$$

is measurable.

Solution. If the functions were real valued, then we would have $\{g > f\} = \{g - f > 0\}$, with the latter set measurable since $g - f$ is measurable. In the general case we can consider $T = \{(u, v) \in \tilde{\mathbb{R}}^2 : u < v\}$; this is a Borel set, in fact an open subset of $\tilde{\mathbb{R}}^2$; the map $f \triangledown g : X \to \tilde{\mathbb{R}}^2$ that has f and g as components is measurable and hence $(f \triangledown g)^{\leftarrow}(T) = \{g > f\}$ is measurable.

For the second question simply note that if $f_* = \liminf_n f_n$ and $f^* = \limsup_n f_n$, then the set where the limit exists is $X \setminus \{f_* < f^*\} = \{f^* \le f_*\}$, hence is measurable.

Of course the limit function is measurable on the measurable set where it exists (by (i) of 3.1.1). □

EXERCISE 3.1.3. Let $f : \mathbb{R} \to \mathbb{K}$ be a function. Given $a \in \mathbb{R}$, the *translate of f by a* is the function $\mathrm{tr}_a f : \mathbb{R} \to \mathbb{R}$ defined by $\mathrm{tr}_a f(x) = f(x - a)$; prove that if f is Borel measurable then its translates are Borel measurable.

Assuming f everywhere differentiable, prove that the derivative f' is measurable.

3.1.3. *Some exercises and some solutions.*

Solution. (*of exercise 3.0.5*) Since restrictions of measurable functions are measurable, necessity is trivial. And if $f_k = f_{|X_k}$ is measurable for every $k \in \mathbb{N}$ and $V \in \mathcal{N}$ we have

$$f^{\leftarrow}(V) = \bigcup_{k\in\mathbb{N}} X_k \cap f^{\leftarrow}(V) = \bigcup_{k\in\mathbb{N}} f_k^{\leftarrow}(V);$$

by the measurabilty of f_k each $f_k^{\leftarrow}(V)$ belongs to the $\sigma-$algebra induced by $\mathcal{M}$ on X_k. in particular belongs to $\mathcal{M}$; then $f^{\leftarrow}(V)$ is a countable union of elements of $\mathcal{M}$. □

EXERCISE 3.1.4. Let X, Y be Hausdorff topological spaces, and let $A \subseteq X$ be a finite or countably infinite subset. Let $f : X \to Y$ be continuous on $X \setminus A$. Prove that $f : X \to Y$ is $(\mathcal{B}(X), \mathcal{B}(Y))-$measurable.

Addition from $(\#)+(\#) : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ and multiplication $(\#)(\#) : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ may be extended by continuity to

$$(\#) + (\#) : \tilde{\mathbb{R}} \times \tilde{\mathbb{R}} \setminus \{(\infty, -\infty), (-\infty, \infty)\} \to \tilde{\mathbb{R}}; \quad (\#)(\#) : \tilde{\mathbb{R}} \times \tilde{\mathbb{R}} \setminus \{(0, \pm\infty), (\pm\infty, 0)\} \to \tilde{\mathbb{R}}$$

prove that any further extension to all of $\tilde{\mathbb{R}} \times \tilde{\mathbb{R}}$ of the these maps is measurable (e.g one can define $\infty + (-\infty) = (-\infty) + \infty = 0$, or any other arbitrary value, and similarly for multiplication one can e.g. define $0 \cdot (\pm\infty) = (\pm\infty) \cdot 0 = 0$ and addition and multiplication remain measurable).

Solution. Since X is Hausdorff, singletons of X are closed sets, in particular Borel sets, and hence all countable subsets of X are Borel sets, in particular the $\sigma-$algebra induced on A by $\mathcal{B}(X)$ is $\mathcal{P}(A)$, and $X \setminus A$ is a Borel set. If $f : X \to Y$ is continuous on $X \setminus A$, then for every open set V of Y the inverse image $f^{\leftarrow}(V)$ is of the form $(U \cap (X \setminus A)) \cup f_{|A}^{\leftarrow}(V)$, with U open in X, since $f_{|X\setminus A}$ is continuous; the second set is a countable set, hence a Borel set, and $U \cap (X \setminus A)$ is also Borel, intersection of the open U with $X \setminus A$, co–countable, hence Borel.

All the rest of the exercise is a trivial application of this result. □

3.1.4. *Almost everywhere, almost every.* Given a measure space $(X, \mathcal{M}, \mu)$, a subset $N \subseteq X$ is said to be *negligible* (it: trascurabile), $\mu-$negligible if the measure has to be specified, if it is contained in a set $M \in \mathcal{M}$ of measure $\mu(M) = 0$; being a *null* set is synonymous to having zero measure: when the space is complete negligible and null sets coincide. A property $P(x)$ applicable to points $x \in X$ of a measure space is said to be true *almost everywhere*, abbreviated a.e. (it: quasi ovunque, abbr. q.o.) if the set $\{x \in X : P(x) \text{ is false}\}$ is negligible; equivalently, the property $P(x)$ is true, except at most on a null set (which contains the set where $P(x)$ fails to be true). A function $f : X \to \tilde{\mathbb{R}}$, or $f : X \to \mathbb{C}$ whose cozero–set is negligible will be called negligible; a null function is a measurable negligible function. But we shall more frequently speak of a function which is a.e zero. These distinctions disappear on complete measure spaces. The set of all equivalence classes of measurable functions modulo null functions will be denoted $L_\mu(X, \mathbb{K})$ or simply $L(\mu)$: its elements are the cosets $f + \mathcal{N}(\mu)$, if $\mathcal{N}(\mu)$ denotes the set of all null functions.

3.1.5. *Measurable functions defined only almost everywhere.* Very often it is the case that on a given measure space $(X, \mathcal{M}, \mu)$ a numerical valued function f is defined not on the whole space X, but only on a subset D of X whose complement $X \setminus D$ is negligible, i.e. contained in a null set $M \in \mathcal{M}$. If f is measurable (with respect to the induced $\sigma-$algebra $\mathcal{M}_D = \{A \cap D : A \in \mathcal{M}\}$) then there is one and only one class of $L(\mu)$ associated to f, in the following sense:

. *In the above hypotheses there exists a measurable $g : X \to \mathbb{K}$ such that the set $\{x \in D : f(x) \neq g(x)\}$ is negligible; and if $h : X \to \mathbb{K}$ is another such function, then $g = h$ a.e. on X.*

Proof. Let M be a null set containing $X \setminus D$. Define $g(x) = f(x)$ for $x \in X \setminus M$, and $g(x) = 0$ for $x \in M$; then g is measurable and $\{x \in D;\ f(x) \neq g(x)\}$ is contained in M. Clearly if $h : X \to \mathbb{K}$ is measurable and $\{x \in D : h(x) \neq f(x)\}$ is negligible, say contained in the null set N, then $\{x \in X : g(x) \neq h(x)\} \subseteq M \cup N$. □

3.1.6. *A. e. equal with respect to Lebesgue measure.* The following exercise highlights a very important aspect of Lebesgue measure.

Exercise 3.1.5. We consider $\mathbb{R}^n$ with Lebesgue measure.

(i) Prove that every non–empty open subset $U \subseteq \mathbb{R}^n$ has strictly positive measure, $\lambda_n(U) > 0$.

(ii) Let $D \subseteq \mathbb{R}^n$ be open, let $f, g : D \to \mathbb{K}$ be functions, and let c be a point of the closure of D in $\mathbb{R}^n$. Assume that f and g are a.e. equal in D and that $\lim_{x\to c, x\in D} f(x) = a \in \mathbb{K}$ and $\lim_{x\to c, x\in D} g(x) = b \in \mathbb{K}$ exist in $\mathbb{K}$. Prove that then $a = b$. In particular, if f and g are both continuous at some point $c \in D$, and are a.e. equal, then $f(c) = g(c)$, so that:

two continuous a.e. equal functions $f, g : D \to \mathbb{K}$ are equal everywhere on D

(instead of $\mathbb{K}$ we might have any Hausdorff topological space as target space).

(iii) If I is an open interval of $\mathbb{R}$ and $f,\ g : I \to \mathbb{R}$ are monotone and a.e. equal, then they coincide on the complement of a countable set, and have the same left and right limits, differing at most at jump points (see also 1.3.6).

Solution. (i) Immediate: given $c \in U$, for $\delta > 0$ small enough the open cube Q centered at c of half–side δ, namely $Q = \{x \in \mathbb{R}^n : \|x - c\|_\infty < \delta\}$ is contained in U, so that $\lambda_n(U) \geq \lambda_n(Q) = 2^n\delta^n > 0$.

(ii) If $a \neq b$ we may take disjoint nbhds V_1 and V_2 of a and b respectively. There is then an open nbhd W of $c \in \mathbb{R}^n$ such that if $x \in W \cap D \setminus \{c\}$ we have $f(x) \in V_1$ and $g(x) \in V_2$; in particular then $f(x) \neq g(x)$ for every $x \in W \cap D \setminus \{c\}$, so that $W \cap D \setminus \{c\} \subseteq \{f \neq g\}$; but $W \cap D \setminus \{c\}$ is a non–empty open subset of $\mathbb{R}^n$, so that $\lambda_n(W \cap D \setminus \{c\}) > 0$, contradicting the hypothesis $f = g$ a.e.

All remaining questions now follow easily.

Remark. Notice the important corollary: If in the class $[g]$ of functions $f \in \mathbb{K}^D$, with D open in $\mathbb{R}^n$, that are a.e. equal to a given function $g \in \mathbb{K}^D$, there is a continuous representative h, then this is the only continuous representative in the class. This will naturally (in general) be the chosen representative of the class.

□

3.2. Measurable simple functions. Given an algebra $\mathcal{A}$ of parts of a set X, a function $f : X \to \mathbb{K}$ is called $\mathcal{A}-$*simple*, just simple when $\mathcal{A}$ is understood, if it has the following form: there is a finite partition $\{E(1), \dots, E(n)\}$ of X into elements $E(k) \in \mathcal{A}$ such that f is constant on each $E(k)$. Such a partition is said to be *admissible* for f, or associated to f; if $f(E(k)) = \{\alpha_k\}$, with $\alpha_k \in \mathbb{K}$, then we have

$$f = \sum_{k=1}^{n} \alpha_k\, \chi_{E(k)}.$$

The range of a simple function is of course finite, $f(X) = \{\alpha_1, \dots, \alpha_n\}$; if $\{y_1, \dots, y_m\}$ is a bijective indexing of the range (i.e. no y_j is repeated) then the set $f^{\leftarrow}(y_j) = \bigcup\{E(k) : \alpha_k = y_j\}$ belongs to $\mathcal{A}$, and we have the *standard representation* of f, $f = \sum_{j=1}^{m} y_j\, \chi_{f^{\leftarrow}(y_j)}$, the one with the coarsest partition. Given two $\mathcal{A}-$simple functions, there is a partition associated to both: if $\{A(1), \dots, A(m)\}$ is associated to f and $\{B(1), \dots, B(n)\}$ to g, then $\{A(j) \cap B(k)\}_{(j,k)}$ (discarding the emptyset if encountered) is admissible for both; then $f + g$ and $f\,g$ have this as associated partition, as well as $f \vee g$ and $f \wedge g$ when $\mathbb{K} = \mathbb{R}$. So the $\mathcal{A}-$simple $\mathbb{K}-$valued functions are a $\mathbb{K}-$algebra of functions, and the real valued simple functions are also a lattice.

EXAMPLE 3.2.1. If $X = \mathbb{R}^n$, and $\mathcal{A}_n$ is the interval algebra, the $\mathcal{A}_n-$simple functions are called *step functions*. If $\mathcal{A}$ is the algebra of finite or cofinite subsets of X then $\mathcal{A}-$simple functions are all functions which are constant outside some finite subset of X.

EXERCISE 3.2.2. If $\mathcal{A}$ is an algebra of parts of X, the $\mathbb{K}-$valued $\mathcal{A}-$simple functions are the functions of the vector subspace of $\mathbb{K}^X$ generated by the characteristic functions of sets $A \in \mathcal{A}$.

Observe that if $(X, \mathcal{M})$ is a measurable space, then the $\mathcal{M}-$simple functions are exactly the measurable functions with finite range. Remember that simple functions are always finite valued: we do not allow ∞ as a value for a simple function.

3.2.1. *Approximation by simple functions.* The next result is of fundamental importance. Assume that $(X, \mathcal{M})$ is a measurable space, and that $f : X \to [0, \infty]$ is a measurable extended real valued positive function. A *subdivision* of the positive half–line $[0, \infty[$ is a finite set $\alpha = \{a_0, \dots, a_m\}$ of positive real numbers, with $a_0 = 0 < a_1 < \cdots < a_m$; the *mesh* of α is $\max\{a_j - a_{j-1} : j = 1, \dots, m\}$. If α, β are subdivisions of $[0, \infty[$ we say that β is *finer* than α if it has more points, i.e. $\beta \supseteq \alpha$. To each subdivision α of $[0, \infty[$ and each measurable $f : X \to [0, \infty]$ we associate a measurable positive simple function $f_\alpha : X \to [0, \infty[$ such that $f_\alpha \le f$ in the following way:

$$f_\alpha = \sum_{j=1}^{m} a_{j-1}\, \chi_{f^{\leftarrow}(]a_{j-1}, a_j])} + a_m\, \chi_{f^{\leftarrow}(]a_m, \infty])}.$$

Notice that $f_\alpha(x) = 0$ if $f(x) = 0$, $f_\alpha(x) = a_m$ if $f(x) = \infty$, and $f_\alpha(x) = a_{j-1}$ if $a_{j-1} < f(x) \le a_j$ for a (necessarily unique) $j \in \{1, \dots, m\}$. If β is another subdivision finer than α then we have $f_\alpha \le f_\beta$, as is easy to see.

PROPOSITION. *Let $\alpha(k)$ be a sequence of subdivisions of $[0, \infty[$ such that:*

(i) $\alpha(k) \subseteq \alpha(k + 1)$.
(ii) $\lim_{k\to\infty} \max \alpha(k) = \infty$
(iii) $\lim_{k\to\infty} \operatorname{mesh} \alpha(k) = 0$.

Then the sequence $f_k = f_{\alpha(k)}$ is increasing and converges pointwise to f; and the convergence is uniform on every set on which f is bounded.

Proof. Condition (i) implies that the sequence is increasing. Put $a(k) = \max \alpha(k)$ and $\delta(k) = \operatorname{mesh} \alpha(k)$. Given $x \in X$ we have to prove that $\lim_{k\to\infty} f_k(x) = f(x)$. If $f(x) = 0$, then $f_k(x) = 0$ for every k; if $f(x) = \infty$, then $f_k(x) = a(k)$ for every k, and by (ii) we have $\lim_{k\to\infty} a(k) = \infty$. Finally, if $0 < f(x) \le a < \infty$, and $a(k) \ge a$ we have:

$$0 < f(x) - f_k(x) \le \delta(k),$$

and since by (iii) we have $\lim_{k\to\infty} \delta(k) = 0$ the conclusion is reached. □

A sequence $\alpha(k)$ verifying (i), (ii) and (iii) is for instance

$$\alpha(k) = \left\{0, \frac{1}{2^k}, \dots, \frac{j}{2^k}, \dots, k\right\} = \{j/2^k : j = 0, \dots, k\, 2^k\},$$

with $\max \alpha(k) = k$ and $\operatorname{mesh} \alpha(k) = 1/2^k$.

3.2.2. We immediately get:

COROLLARY. *Let $(X, \mathcal{M})$ be a measurable space, and let $f : X \to \mathbb{C}$ be measurable. Then there is a sequence of measurable simple functions φ_n which converges pointwise to f and such that $|\varphi_n| \uparrow |f|$; the convergence is uniform on sets on which f is bounded.*

Proof. Let $u = \operatorname{Re} f$ and $v = \operatorname{Im} f$; choose sequences $u_n^\pm$ e $v_n^\pm$ of positive functions converging upwards to $u^\pm$ e $v^\pm$. Then $\varphi_n = (u_n^+ - u_n^-) + i\,(v_n^+ - v_n^-)$ is as required (details of the proof are left as an exercise, see below). □

Details of the proof Let's prove first the real valued case, $v = 0$. We have to prove that $u_n^+ + u_n^-$ is increasing. Given $x \in X$, if $u(x) = 0$ then $u^+(x) = u^-(x) = 0$ and also $u_n^+(x) = u_n^-(x) = 0$ for every n; if $u^+(x) > 0$ then $u^-(x) = 0$ and $u_n^-(x) = 0$ for every n, etc. Next, if $|u_n| \le |u_{n+1}|$ and $|v_n| \le |v_{n+1}|$ we have, trivially, $|u_n + i\, v_n| \le |u_{n+1} + i\, v_{n+1}|$, since $\sqrt{u^2 + v^2}$ increases if $|u|$ and $|v|$ both increase.

3.2.3. *Measurable functions on completions.* When completing a measure space, the set of (equivalence classes of) measurable functions modulo null functions does not change (in other words $L(\mu) = L(\bar{\mu})$):

PROPOSITION. *Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $(X, \overline{\mathcal{M}}, \bar{\mu})$ be its completion. For every $\overline{\mathcal{M}}-$measurable function $f : X \to \mathbb{K}$ there is a $\mathcal{M}-$measurable function $g : X \to \mathbb{K}$ such that the set $\{f \neq g\}$ is contained in a set $M \in \mathcal{M}$ with $\mu(M) = 0$.*

Proof. For the characteristic function of a set in $\overline{\mathcal{M}}$ this is true (2.4.1). Then it is true for simple functions; and since every measurable function is a limit of a sequence of simple functions, it is true for every measurable function as well. $\square$

4. Integration

We begin now the theory of abstract integration. Integrals will be first defined for measurable positive simple functions, then for arbitrary measurable positive functions, then for some measurable functions; these will have an integral only if they are not too large.

4.1. Integrals of measurable functions.

4.1.1. *Integrals of positive simple functions.* We start by defining the integral of a positive $\mathcal{A}$−simple function, with respect to a positive finitely additive measure $\mu : \mathcal{A} \to [0, \infty]$, $\mathcal{A}$ an algebra of parts of X. Assume that $f = \sum_{j=1}^m y_j\, \chi_{f^{\leftarrow}(y_j)}$ is the standard representation of f (i.e. $f(X) = \{y_1, \dots, y_m\}$, with $y_j \neq y_k$ if $j \neq k$); then, by definition

$$\int_X f\, d\mu = \int_X f = \int f := \sum_{j=1}^m y_j\, \mu(f^{\leftarrow}(y_j))$$

(as usual in measure theory, $0\,\infty = 0$ is used); this integral is a non negative real number, or ∞ (this last if and only if $y_j > 0$ and $\mu(f^{\leftarrow}(y_j)) = \infty$ for at least one $j \in \{1, \dots, m\}$). We can immediately prove that if $\{E(1), \dots, E(n)\}$ is an admissible partition for $f = \sum_{k=1}^n \alpha_k\, \chi_{E(k)}$ then

$$\int_X f = \sum_{k=1}^n \alpha_k\, \mu(E(k));$$

in fact, collecting all the $E(k)$ such that $\alpha_k = y_j$:

$$\sum_{k=1}^n \alpha_k\, \mu(E(k)) = \sum_{j=1}^m y_j \left(\sum \{\mu(E(k)) : \alpha_k = y_j\}\right) = \sum_{j=1}^m y_j\, \mu(f^{\leftarrow}(y_j))$$

by finite additivity of μ. We then get additivity of the integral: if f and g are positive and $\mathcal{A}$−simple there is a partition admissible for both f and g, so that $f = \sum_{k=1}^n \alpha_k\, \chi_{E(k)}$ and $g = \sum_{k=1}^n \beta_k\, \chi_{E(k)}$ and $f + g = \sum_{k=1}^n (\alpha_k + \beta_k)\, \chi_{E(k)}$, whence

$$\int_X (f+g) = \sum_{k=1}^n (\alpha_k + \beta_k)\, \mu(E(k)) = \sum_{k=1}^n \alpha_k\, \mu(E(k)) + \sum_{k=1}^n \beta_k\, \mu(E(k)) = \int_X f + \int_X g.$$

That $\int_X (\lambda f) = \lambda \int_X f$ for each $\lambda > 0$ and each positive simple f is trivial. This implies that if a function is represented as a positive linear combination of characteristic functions of sets in $\mathcal{A}$, $f = \sum_{k=1}^p \alpha_k \chi_{E(k)}$, with $\alpha_k \geq 0$, we have

$$\int_X f = \sum_{k=1}^p \alpha_k\, \mu(E(k))$$

even if the $E(k)$'s are not pairwise disjoint. Notice that the integral is zero if and only if $\alpha_k > 0$ implies $\mu(E(k)) = 0$, and that the integral is finite if and only if $\mu(E(k)) = \infty$ implies $\alpha_k = 0$. Note that the constant 1 has a finite integral if and only if $\mu(X) < \infty$: by definition $\int_X 1\, d\mu = \mu(X)$. Remember that

. *A simple function f has finite integral if and only if its cozero set* $\mathrm{Coz}(f) = \{f \neq 0\}$ *has finite measure.*

Finally, the integral is isotone: if $f \leq g$ with f, g positive simple then also $g = f + h$ with h positive and simple so that $\int_X g = \int_X f + \int_X h \geq \int_X f$.

4.1.2. *Integral of a positive measurable function.* Let $(X, \mathcal{M}, \mu)$ be a measure space. For $f \in \mathcal{L}^+$, i.e. f positive measurable, we define its integral as:

$$\int_X f\, d\mu = \int_X f = \int f := \sup\left\{\int_X \varphi\, d\mu : 0 \leq \varphi \leq f,\ \varphi \text{ simple measurable}\right\}.$$

It is non negative real number, or ∞; clearly for positive simple functions this integral coincides with the one already defined. Trivially this integral is isotone:

$$f, g \in \mathcal{L}^+,\ f \leq g \implies \int_X f \leq \int_X g,$$

and positively homogeneous: $f \in \mathcal{L}^+$ and $\lambda \geq 0$ imply $\int_X \lambda f = \lambda \int_X f$. We also define the *integral of f extended to A* as $\int_A f := \int_X f\, \chi_A$ for $f \in \mathcal{L}^+$ and $A \in \mathcal{M}$.

4.1.3. *Positive functions with zero integral.* Let us prove that a positive measurable function has integral zero if and only if it is zero almost everywhere, that is, its cozero–set $\mathrm{Coz}(f) = \{f \neq 0\}$ has measure zero

. *If $f \in \mathcal{L}^+$ then*

$$\int_X f\,d\mu = 0 \iff \mu(\{f > 0\}) = 0.$$

Proof. We have

$$\{f > 0\} = \bigcup_{n=1}^{\infty} \{f \geq 1/n\},$$

and clearly $\varphi_n = (1/n)\,\chi_{\{f \geq 1/n\}}$ is a simple function dominated by f, with integral $(1/n)\mu(\{f \geq 1/n\})$. Now, $\int_X f = 0$ means that $\int_X \varphi = 0$ for every positive simple function under f, in particular

$$\int_X \varphi_n = \frac{1}{n}\,\mu(\{f \geq 1/n\}) = 0,$$

equivalently $\mu(\{f \geq 1/n\}) = 0$, for every $n \geq 1$. Then $\{f > 0\}$, a countable union of sets of measure zero, has also measure 0. Conversely, if $\mu(\{f > 0\}) = 0$ and $\varphi = \sum_{k=1}^{m} \alpha_k\,\chi_{E(k)}$ is a positive simple function with $\alpha_j > 0$ for every $j \in \{1, \dots, m\}$, we have $\varphi(x) \geq \alpha_k > 0$ for every $x \in E(k)$, so that $E(k) \subseteq \{f > 0\}$, and $\mu(E(k)) \leq \mu(\{f > 0\}) = 0$ for every $k \in \{1, \dots, m\}$. Then

$$\int_X \varphi = \sum_{k=1}^{m} \alpha_k\,\mu(E(k)) = 0;$$

this is equivalent to say that the integral of f is zero. □

4.1.4. *Monotone convergence.* The most important theorem of abstract integration is the:

. Monotone convergence theorem. *Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $f_0 \leq f_1 \leq f_2 \leq \dots$ be an increasing sequence of functions in $\mathcal{L}^+(X)$, such that $f_n \uparrow f$ pointwise. Then*

$$\int_X f = \lim_{n\to\infty} \int_X f_n \left(= \sup\left\{ \int_X f_n \,:\, n \in \mathbb{N} \right\} \right).$$

Proof. By isotony we have $\int_X f_n \leq \int_X f$ for every n; we need to prove that $\int_X f \leq \sup\left\{\int_X f_n \,:\, n \in \mathbb{N}\right\}$. Equivalently, given a simple measurable φ, with $0 \leq \varphi \leq f$, we have to prove that $\int_X \varphi \leq \lim_{n\to\infty} \int_X f_n$. We shall prove that for every λ with $0 < \lambda < 1$ we have

$$\lim_{n\to\infty} \int_X f_n \geq \lambda \int_X \varphi,$$

and we get the conclusion taking the limit in this formula as $\lambda \to 1^-$. Set $E(n) = \{f_n \geq \lambda\,\varphi\}$, i.e. $E(n) = \{x \in X \,:\, f_n(x) \geq \lambda\,\varphi(x)\}$; then $E(n) \subseteq E(n+1)$ (because $f_n \leq f_{n+1}$) and $\bigcup_{n=0}^{\infty} E(n) = X$ (if $\varphi(x) = 0$ then $x \in E(n)$ for every n; and if $\varphi(x) > 0$ then $\lambda\,\varphi(x) < \varphi(x) \leq f(x) = \lim_{n\to\infty} f_n(x)$, so that eventually $f_n(x) > \lambda\,\varphi(x)$). Then

$$(*) \qquad \int_X f_n \geq \int_{E(n)} f_n \geq \int_{E(n)} \lambda\,\varphi = \lambda \int_X \varphi\,\chi_{E(n)}.$$

If $\varphi = \sum_{k=1}^{m} \alpha_k\,\chi_{A(k)}$, with $\alpha_k > 0$ and $A(k) \in \mathcal{M}$ for $k = 1, \dots, m$ we have $\varphi\,\chi_{E(n)} = \sum_{k=1}^{m} \alpha_k\,\chi_{A(k)\cap E(n)}$, so that

$$\int_X \varphi\,\chi_{E(n)} = \sum_{k=1}^{m} \alpha_k\,\mu(A_k \cap E(n)),$$

and from (*) we get

$$(**) \qquad \int_X f_n \geq \lambda \sum_{k=1}^{m} \alpha_k\,\mu(A_k \cap E(n)),$$

Since $E(n) \uparrow X$ we have $A(k) \cap E(n) \uparrow_n A(k)$; by continuity from below of measures we have $\mu(A(k)) = \lim_{n\to\infty} \mu(A(k) \cap E(n))$, for every $k \in \{1, \dots, m\}$. Then, passing to the limit in (**) as n tends to ∞ we get

$$\lim_{n\to\infty} \int_X f_n \geq \lambda \sum_{k=1}^{m} \alpha_k\,\mu(A(k)) = \lambda \int_X \varphi,$$

as desired. □

4.1.5. *Additivity of the integral.*

. *Let $(X, \mathcal{M}, \mu)$ be a measure space. If $f, g \in \mathcal{L}^+(X)$ then*

$$\int_X (f+g) = \int_X f + \int_X g.$$

Termwise integration of series of positive functions *(or countable additivity of the integral of positive functions). If $\sum_{n=0}^{\infty} f_n$ is a series of functions in $\mathcal{L}^+(X)$, with pointwise sum f, then*

$$\int_X f = \sum_{n=0}^{\infty} \int_X f_n.$$

Proof. Take sequences u_n and v_n of simple functions in $\mathcal{L}^+(X)$ such that $u_n \uparrow f$ and $v_n \uparrow g$; u_n and v_n exist by 3.2.2. Then $u_n + v_n \uparrow f + g$, so that by the monotone convergence theorem we have

$$\int_X (f+g) = \lim_{n\to\infty} \int_X (u_n+v_n);\ \text{but}\ \int_X (u_n+v_n) = \int_X u_n + \int_X v_n,\ \lim_{n\to\infty}\int_X u_n = \int_X f,\ \lim_{n\to\infty}\int_X v_n = \int_X g,$$

so that

$$\int_X (f+g) = \int_X f + \int_X g.$$

For the series theorem: if $g_m = \sum_{n=0}^{m} f_n$, then $g_m \uparrow f$ and by monotone convergence the integrals of g_m converge to the integral of f as $m \to \infty$, but by finite additivity we have:

$$\int_X g_m = \int_X \left(\sum_{n=0}^{m} f_n\right) = \sum_{n=0}^{m} \int_X f_n; \quad \text{passing to the limit with } m \uparrow \infty\text{:} \quad \int_X \left(\sum_{n=0}^{\infty} f_n\right) = \sum_{n=0}^{\infty} \int_X f_n.$$

□

4.1.6. *Čebičeff inequality.* The following inequality is almost trivial but quite useful (some people call it Markov inequality):

. *Given a measure space $(X, \mathcal{M}, \mu)$ and $f \in \mathcal{L}^+(X)$ for every $\alpha > 0$ we have*

$$\mu(\{f \geq \alpha\}) \leq \frac{1}{\alpha} \int_X f\,d\mu.$$

Proof. We have

$$\int_X f\,d\mu \geq \int_{\{f\geq\alpha\}} f\,d\mu \geq \int_{\{f\geq\alpha\}} \alpha\,d\mu = \alpha\,\mu(\{f \geq \alpha\}) \quad \text{which implies} \quad \mu(\{f \geq \alpha\}) \leq \frac{1}{\alpha}\int_X f\,d\mu.$$

□

Remember that the cozero–set of a real or complex valued function $f : X \to \mathbb{K}$ is the set $\mathrm{Coz}(f) = \{x \in X : f(x) \neq 0\}$.

Corollary. *If $f \in \mathcal{L}^+$ and $\int_X f\,d\mu < \infty$ then $\{f = \infty\}$ has measure zero and* $\mathrm{Coz}(f) = \{f > 0\}$ *has $\sigma-$finite measure.*

Proof. We have $\mu(\{f \geq 1/n\} \leq n \int_X f\,d\mu$, so that $\{f \geq 1/n\}$ has finite measure and $\mathrm{Coz}(f) = \bigcup_{n=1}^{\infty}\{f \geq 1/n\}$ has $\sigma-$finite measure. And $\{f = \infty\} \subseteq \{f \geq n\}$ for every n so that $\mu(\{f = \infty\}) \leq \mu(\{f \geq n\}) \leq (1/n)\int_X f$ for every n implies $\mu(\{f = \infty\}) = 0$. □

Thus a positive function with finite integral can assume the value ∞, but only on a null set, it is a.e. finite–valued .

Notice also that for every $f \in \mathcal{L}^+$ and every $M \in \mathcal{M}$ with $\mu(M) = 0$ we have $\int_M f = 0$, since $f\,\chi_M$ can be nonzero only on M.

4.1.7. *A.e. equal functions have the same integral.* We have seen in 4.1.3 that a function $f \in \mathcal{L}^+$ has zero integral iff it is a null function. Let us see that

. *If $f, g \in \mathcal{L}^+$ are equal a.e., then $\int_X f = \int_X g$.*

Proof. Let M be a null set containing the set $\{f \neq g\}$, and let $A = X \setminus M$. Then $f\,\chi_A = g\,\chi_A$, so that $\int_A f = \int_A g$; and $\int_M f = \int_M g = 0$. Then

$$\int_X f = \int_M f + \int_A f = \int_A f = \int_A g = \int_A g + \int_M g = \int_X g.$$

□

EXERCISE 4.1.1. (solution in 4.1.11) Prove that for every $f \in \mathcal{L}^+$:

$$\int_X f = \lim_{n\to\infty} \int_{\{f>1/n\}} f.$$

Prove that if $\int_X f < \infty$, then for every $\varepsilon > 0$ there is $E \in \mathcal{M}$ with $\mu(E) < \infty$ such that $\int_E f > \int_X f - \varepsilon$.

4.1.8. *Decreasing sequences.* Given the monotone convergence theorem, one might wonder if there an analogous theorem for decreasing sequences: there is, but with an added hypothesis on the finiteness of integrals. The situation exactly parallels that of measures, which are continuous from below; continuity from above requires a finiteness assumption. These situations will be handled in the sequel by the more general dominated convergence theorem, 4.2.5; see also 4.4.2.

4.1.9. *Indefinite integral and density.* (*Important*) Let $(X, \mathcal{M}, \mu)$ be a measure space. Let $\rho : X \to [0, +\infty]$ be measurable. The *indefinite integral* of ρ is the set function $\nu = \nu_\rho : \mathcal{M} \to [0, +\infty]$ defined for $A \in \mathcal{M}$ by

$$\nu(A) = \int_A \rho\, d\mu.$$

PROPOSITION. *The set function ν is a measure, and for every $f \in \mathcal{L}^+$ we have*

$$\int_X f\, d\nu = \int_X f\, \rho\, d\mu.$$

Proof. For every $E \in \mathcal{M}$ such that $\mu(E) = 0$ we have that $\int_E \rho = \int_X \rho\,\chi_E = 0$ because $\mathrm{Coz}(\rho\,\chi_E) \subseteq E$ has measure 0; in particular this holds for $E = \emptyset$. And if $(A(n))_{n\in\mathbb{N}}$ is a disjoint sequence in $\mathcal{M}$ with union A, we have

$$f\,\chi_A = \sum_{n=0}^{\infty} f\,\chi_{A(n)} \Longrightarrow \int_X f\,\chi_A = \sum_{n=0}^{\infty} \int_X f\,\chi_{A(n)} \iff \nu(A) = \sum_{n=0}^{\infty} \nu(A_n),$$

by the theorem on termwise integration of series of positive functions. Next, if $f = \sum_{k=1}^m \alpha_k \chi_{E(k)}$ is measurable we have by linearity of the integral:

$$\int_X f\, d\nu = \sum_{k=1}^m \alpha_k \int_X \chi_{E(k)}\, d\nu = \sum_{k=1}^m \alpha_k\, \nu(E(k)) = \sum_{k=1}^m \alpha_k \int_{E(k)} \rho\, d\mu = \sum_{k=1}^m \alpha_k \int_X \chi_{E(k)}\, \rho\, d\mu =$$
$$\int_X \left(\sum_{k=1}^m \alpha_k\, \chi_{E(k)}\right) \rho\, d\mu = \int_X f\, \rho\, d\mu.$$

Given $f \in \mathcal{L}^+$ we then pick an increasing sequence f_n of positive measurable simple functions converging to f; then $f_n\,\rho$ is an increasing sequence in $\mathcal{L}^+$ converging to $f\,\rho$, and for every n we have, by what just proved:

$$\int_X f_n\, d\nu = \int_X f_n\, \rho\, d\mu;$$

passing to the limit in this formula as $n \to \infty$ the monotone convergence theorem implies that

$$\int_X f\, d\nu = \int_X f\, \rho\, d\mu.$$

□

Remark. Observe that $\mu(E)=0$ implies $\nu(E)=0$. One writes $d\nu=\rho\,d\mu$; ρ is the *density function* of the measure ν with respect to the measure μ; one also writes

$$\frac{d\nu}{d\mu}=\rho$$

and thinks of ρ as the *derivative* of the measure ν with respect to the measure μ; we shall study this situation in the sequel (Radon–Nikodym theorem, 7.1.7).

4.1.10. *Fatou's lemma.* One of the basic tenets of integration theory, Fatou's lemma deals with arbitrary sequences of positive functions.

. *Let $(X,\mathcal{M},\mu)$ be a measure space, and let f_n be a sequence of positive measurable extended real valued functions on X ($f_n\in\mathcal{L}^+(X)$). Then*

$$\int_X \liminf_{n\to\infty} f_n\,d\mu \le \liminf_{n\to\infty}\int_X f_n\,d\mu.$$

Proof. For every $m\in\mathbb{N}$ set $f_{*m}(x)=\inf\{f_n(x)\,:\,n\ge m\}$. Then $f_{*m}\le f_n$ for every $n\ge m$, hence $\int_X f_{*m}\le\int_X f_n$ for every $n\ge m$ so that

$$\int_X f_{*m}\le\inf\left\{\int_X f_n\,:\,n\ge m\right\};$$

taking limits in this inequality as $m\uparrow\infty$ we get $\int_X \liminf_{n\to\infty} f_n$ on the left–hand side, by monotone convergence, and $\liminf_{n\to\infty}\int_X f_n$, by definition of $\liminf$, on the right–hand side. $\square$

Remark. One might say that Fatou's lemma expresses the sequential lower semicontinuity of the integral on the set of positive measurable functions.

4.1.11. *Solutions of the previous exercises.*

Solution. (*of exercise 4.1.1*). If $E(n)=\{f>1/n\}$ and $f_n=f\,\chi_{E(n)}$ we have $f_n\uparrow f$ (easy: $E(n)\subseteq E(n{+}1)$ so that $f\,\chi_{E(n)}\le f\,\chi_{E(n+1)}$; and if $x\in X$ and $f(x)>0$ then $f_n(x)=f(x)$ as soon as $1/n<f(x)$, while if $f(x)=0$ then $f_n(x)=f(x)$ for every $n\in\mathbb{N}$). By monotone convergence we then have

$$\int_{E(n)} f\,d\mu=\int_X f_n\uparrow\int_X f.$$

If $\int_X f<\infty$, then for every $\varepsilon>0$ we have $\int_X f-\varepsilon<\int_X f$, so that there is $m\in\mathbb{N}$ such that $\int_{E(n)} f>\int_X f-\varepsilon$ for $n\ge m$; by Čebičeff inequality we have that $\mu(E(m))\le m\int_X f$ is finite. $\square$

4.1.12. *Image measure.* If $(X,\mathcal{M},\mu)$ is a measure space, $(Y,\mathcal{N})$ is a measurable space and $\phi:X\to Y$ is a measurable map there is a natural way of transporting the measure μ on Y to make $(Y,\mathcal{N})$ into a measure space.

. *The set function $\phi_\#\mu=\mu\phi^{\leftarrow}:\mathcal{N}\to[0,\infty]$ defined for every $B\in\mathcal{N}$ by $\mu\phi^{\leftarrow}(B)=\mu(\phi^{\leftarrow}(B))$ is a measure on $\mathcal{N}$. For every $f\in\mathcal{L}^+_{\mathcal{N}}(Y)$ we have*

$$\int_Y f(y)\,d\mu\phi^{\leftarrow}(y)=\int_X f\circ\phi(x)\,d\mu(x).$$

The measure $\phi_\#\mu$ is the *image measure* of μ by means of ϕ; notice that $\mu\phi^{\leftarrow}$ is naturally defined also on the final $\sigma-$algebra $\mathcal{Q}_\phi\supseteq\mathcal{N}$ of ϕ.

Proof. Clearly $\mu\phi^{\leftarrow}(\emptyset)=0$. And if $(A_n)_{n\in\mathbb{N}}$ is a disjoint sequence of elements of $\mathcal{N}$, or of $\mathcal{Q}_\varphi$, with union A, then $(\phi^{\leftarrow}(A_n))_{n\in\mathbb{N}}$ is a disjoint sequence of elements of $\mathcal{M}$, with union $B=\phi^{\leftarrow}(A)$. Then

$$\mu\phi^{\leftarrow}(A):=\mu(B)=\sum_{n=0}^{\infty}\mu(\phi^{\leftarrow}(A_n))=\sum_{n=0}^{\infty}\mu\phi^{\leftarrow}(A_n),$$

proving countable additivity of $\mu\phi^{\leftarrow}$. Notice now that if $A\subseteq Y$ then $\chi_A\circ\phi=\chi_{\phi^{\leftarrow}(A)}$, so that if $u=\sum_{j=1}^m\alpha_j\,\chi_{A(j)}$ is an $\mathcal{N}-$simple function then $u\circ\phi=\sum_{j=1}^m\alpha_j\,\chi_{\phi^{\leftarrow}(A(j))}$ is $\mathcal{M}-$simple and for $u\ge 0$:

$$\int_Y u\,d\mu\phi^{\leftarrow}=\sum_{j=1}^m\alpha_j\,\mu\phi^{\leftarrow}(A(j))=\sum_{j=1}^m\alpha_j\,\mu(\phi^{\leftarrow}(A(j))=\int_X u\circ\phi\,d\mu.$$

For $f \in \mathcal{L}^+(Y)$ pick a sequence $u_n \uparrow f$ of positive simple functions and pass to the limit in the formula $\int_X u_n \, d\mu\phi^{\leftarrow} = \int_X u_n \circ \phi \, d\mu$, using the monotone convergence theorem (if u_n is increasing then clearly also $u_n \circ \phi$ is increasing). $\square$

The *Lebesgue measure on the unit circle* $\mathbb{U} = \{z \in \mathbb{C} : |z| = 1\}$ is the image measure μ, by the canonical winding map $t \mapsto e^{it}$, of the measure induced by Lebesgue measure on the interval $]-\pi, \pi]$ (or on any interval of length 2π): it is of course the ordinary arclength measure. This measure is invariant by rotation, that is $\mu(a\,E) = \mu(E)$ for every measurable E and every $a \in \mathbb{U}$; this rotational invariance of the arclength measure on the circle is a consequence of translational invariance of Lebesgue measure on the line, as we now see. If $\phi(t) = e^{it}$, and $a = e^{i\alpha}$ with $\alpha \in]-\pi, \pi]$ then $\phi^{\leftarrow}(aE)$ is the translate by α of $\phi^{\leftarrow}(E)$, that is $T = \alpha + \phi^{\leftarrow}(E)$, provided that the piece of T which might fall out of $]-\pi, \pi]$ is translated back to this interval by a translation of 2π if $\alpha < 0$, of -2π if $\alpha > 0$. So the Lebesgue measure of this set is $\lambda_1(\phi^{\leftarrow}(E)) = \lambda_1(T)$.

4.1.13. *Preimage measure and variable change.* We now address the following question: having a set X, a measure space $(Y, \mathcal{N}, \nu)$ and a function $\phi : X \to Y$, is there a measure λ on the initial $\sigma-$algebra $\mathcal{M} = \{\phi^{\leftarrow}(V) : \mathcal{N} \in \mathcal{N}\}$, such that ν is the image measure of λ by ϕ, $\nu = \phi_{\#}\lambda = \lambda\phi^{\leftarrow}$? There is only one way to define λ, and this is $\lambda(\phi^{\leftarrow}(V)) = \nu(V)$; this set function must be well–defined on $\mathcal{M}$, and be a measure.

Leaving the general case to exercise 4.1.2 we confine ourselves to the easier and important case in which ϕ is a bijection; in this case of course the measure λ is the image measure by ϕ^{-1} of the measure ν, that is $\lambda = \phi^{-1}_{\#}\nu$: for every $E \in \phi^{\leftarrow}(\mathcal{N})$ we set (notice that $E = \phi^{\leftarrow}(\phi(E))$)

$$\lambda(E) = \nu(\phi(E)).$$

It follows that $f \in \mathcal{L}^+_{\mathcal{N}}(Y)$ iff $f \circ \phi \in \mathcal{L}^+_{\phi^{\leftarrow}(\mathcal{N})}(X)$ and

$$\int_Y f(y)\, d\nu(y) = \int_X f \circ \phi(x)\, d\lambda(x).$$

The measure λ is the *preimage measure* of ν by ϕ. Most interesting is the case in which X has already a measure μ defined on a $\sigma-$algebra $\mathcal{M}$ containing $\phi^{\leftarrow}(\mathcal{N})$ and λ has a density with respect to μ, that is there is $\rho \in \mathcal{L}^+_{\phi^{\leftarrow}(\mathcal{N})}(X)$ such that

$$\lambda(E) = \int_E \rho(x)\, d\mu(x) \qquad \text{for every } E \in \phi^{\leftarrow}(\mathcal{N});$$

then, for every $f \in \mathcal{L}^+_{\mathcal{N}}(Y)$ we have

$$\int_Y f(y)\, d\nu(y) = \int_X f \circ \phi(x)\, d\lambda(x) = \int_X f \circ \phi(x)\, \rho(x)\, d\mu(x);$$

In Calculus the following fact is used: we have open sets $U, V \subseteq \mathbb{R}^n$ and $\phi : U \to V$ a C^1 diffeomorphism. We consider the Borel $\sigma-$algebras $\mathcal{B}(U)$ and $\mathcal{B}(V)$ on U and V, so that ϕ and ϕ^{-1} are both measurable, and as measures μ and ν on $\mathcal{B}(U)$ and $\mathcal{B}(V)$ respectively we take the (induced) Lebesgue $n-$dimensional measure m. Now the preimage measure of m by ϕ has density $|\det \phi'(x)|$, that is for every Borel measurable $E \subseteq U$ we have

$$m(\phi(E)) = \int_E |\det \phi'(x)|\, dm(x).$$

It is clearly enough to prove it for compact intervals (remember: if the measures $\lambda = \phi^{-1}m$ and $|\det \phi'|\, dm$ coincide on compact subintervals of U they coincide on $\mathcal{B}(U)$ (2.6.3)); the proof for general n will be given in 8.8; we give here the proof for the case $n = 1$. Let $Q = [a, b]$ $(a < b)$ be a compact interval contained in U; since ϕ is a homeomorphism, it is strictly monotone on each connected component of the open set $U \subseteq \mathbb{R}^1$, so that $\phi([a, b])$ is the interval $[\phi(a), \phi(b)]$ if ϕ is increasing on $[a, b]$, and is the interval $[\phi(b), \phi(a)]$ if ϕ is decreasing on $[a, b]$. In the first case $\phi'(x) > 0$ and

$$m(\phi([a, b])) = \phi(b) - \phi(a) = \int_a^b \phi'(x)\, dx = \int_{[a,b]} |\phi'(x)|\, dx;$$

in the second case $\phi'(x) < 0$ for every $x \in [a, b]$ and

$$m(\phi([a, b])) = \phi(a) - \phi(b) = \int_b^a \phi'(x)\, dx = \int_{[a,b]} (-\phi'(x))\, dx = \int_{[a,b]} |\phi'(x)|\, dx,$$

and the theorem is proved for $n = 1$. Notice that then for every positive measurable f on V we have

$$\int_V f(y)\,dy = \int_U f(\phi(x))\,|\phi'(x)|\,dx.$$

EXERCISE 4.1.2. In the situation described above (X a set $(Y, \mathcal{N}, \nu)$ a measure space, $\phi : X \to Y$ a function), prove that the formula:

$$\lambda(\phi^{\leftarrow}(V)) = \nu(V) \qquad (V \in \mathcal{N})$$

is a good definition of a set function λ on the initial σ−algebra $\phi^{\leftarrow}(\mathcal{N}) = \{\phi^{\leftarrow}(V)) : V \in \mathcal{N}\}$ if and only if $\phi(X)$ has non–empty intersection with every set $B \in \mathcal{N}$ with $\nu(B) > 0$, and that if this holds, then λ is a measure on $\phi^{\leftarrow}(\mathcal{N})$ such that $\phi_{\#}\lambda = \nu$.

Solution. Observe that $\phi^{\leftarrow}(V) = \phi^{\leftarrow}(W)$ if and only if $V \cap \phi(X) = W \cap \phi(X)$, equivalently if $V \vartriangle W \subseteq Y \smallsetminus \phi(X)$, so that $\nu(V \vartriangle W) = 0$, and this clearly implies $\nu(V) = \nu(W)$. Then the condition on $\phi(X)$ implies that λ is well defined. Moreover if $\phi^{\leftarrow}(V) = \emptyset$ then $\nu(V) = 0$. so that $\lambda(\emptyset) = 0$. And $\phi^{\leftarrow}(V_n)$ pairwise disjoint is equivalent to assert that the sets V_n are pairwise almost disjoint; if $V = \bigcup_{n\in\mathbb{N}} V_n$ is the union of the V_n's this implies $\nu(V) = \sum_{n\in\mathbb{N}} \nu(V_n)$ (see exercise 2.1.6), hence also $\lambda(\phi^{\leftarrow}(V)) = \sum_{n\in\mathbb{N}} \lambda(\phi^{\leftarrow}(V_n))$, so that λ is a measure

□

4.1.14. *Decreasing distribution function.* In the following exercise we present an alternative approach to the definition of integral of measurable positive function.

EXERCISE 4.1.3. Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $f : X \to \tilde{\mathbb{R}}$ be measurable. We define the *decreasing distribution function of* f as the function $\rho_f : \tilde{\mathbb{R}} \to \tilde{\mathbb{R}}$ given by $\rho_f(t) = \mu(\{f > t\})$.

(i) Prove that ρ_f is decreasing and right–continuous.
(ii) Prove that $f \le g$ implies $\rho_f(t) \le \rho_g(t)$, for every $t \in \tilde{\mathbb{R}}$.
(iii) Let f_n be an increasing sequence of measurable functions, converging pointwise to f. Prove that then $\rho_{f_n}(t) \uparrow \rho_f(t)$, for every $t \in \tilde{\mathbb{R}}$.

From now on we consider only functions $f : X \to [0, \infty]$, so that ρ_f is relevant only for $t \ge 0$.

(iv) Let $\varphi = \sum_{j=0}^m a_j\,\chi_{A(j)}$ be a positive simple function, where we assume $0 = a_0 < a_1 \cdots < a_m$. Plot the graph of ρ_φ, assuming that all $A(j)$ have finite measure if $j > 0$, and are pairwise disjoint; in particular, note that ρ_φ is a step function with compact support in $[0, \infty[$, and prove that
$$\int_X \varphi\,d\mu = \int_0^\infty \rho_\varphi(t)\,dt.$$
(v) Prove that for every $f \in \mathcal{L}^+$ we have
$$\int_X f\,d\mu = \int_0^\infty \rho_f(t)\,dt$$
(the latter integral may be intended as a Lebesgue integral, or a generalized Riemann integral of the monotone function ρ_f; supposing this known, we could use it to define the integral of f; monotone convergence would in this way be very easy to prove).

Solution. (i) Is $s < t$ then clearly $\{f > s\} \supseteq \{f > t\}$ so that $\mu(\{f > s\}) \ge \mu(\{f > t\})$. And if $t_n \downarrow t$ in $\tilde{\mathbb{R}}$, then $\{f > t_n\} \uparrow \{f > t\}$ so that $\mu(\{f > t_n\}) \uparrow \mu(\{f > t\}$ by continuity from below of the measure μ.

(ii) If $f \le g$ and $f(x) > t$ then also $g(x) \ge f(x) > t$, so that $\{f > t\} \subseteq \{g > t\}$, and monotonicity of the measure implies $\mu(\{f > t\}) \le \mu(\{g > t\})$.

(iii) We have $\{f > t\} \supseteq \{f_{n+1} > t\} \supseteq \{f_n > t\}$ for every n, as observed in (ii); and if $f(x) > t$ then since $f(x) = \sup_n\{f_n(x)\}$ we have $x \in \{f_n > t\}$ for n large; in other words, $\{f_n > t\} \uparrow \{f > t\}$, so that $\rho_{f_n}(t) \uparrow \rho_f(t)$.

(iv) We have $\rho_\varphi(t) = \mu(X)$ if $t < 0$, For $t \in [0 = a_0, a_1[$ we have $\rho_\varphi(t) = \sum_{k=1}^m \mu(A(k))$. For $t \in [a_1, a_2[$ we have $\rho_\varphi(t) = \sum_{k=2}^m \mu(A(k))$. In general for $t \in [a_{j-1}, a_j[$, $j \in \{1, \dots, m\}$ we have $\rho_\varphi(t) = \sum_{k=j}^m \mu(A(k))$, and if $t \in [a_m, \infty]$ then $\rho_\varphi(t) = 0$. Thus

$$\rho_\varphi = \sum_{j=1}^m \left(\sum_{k=j}^m \mu(A(k))\right) \chi_{[a_{j-1},a_j[}.$$

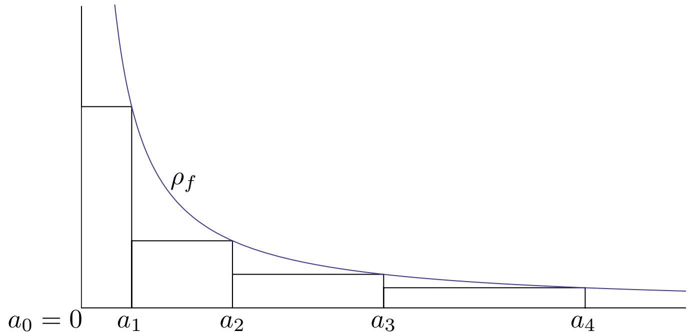

FIGURE 1. A decreasing distribution function and the step function of an approximating simple function.

Then, assuming first that all measures $\mu(A(j))$ are finite for $j = 1, \dots, m$:

$$\int_0^\infty \rho_\varphi(t)\,dt = \sum_{j=1}^m \left(\sum_{k=j}^p \mu(A(k))\right)(a_j - a_{j-1}) = \sum_{j=1}^m \left(\sum_{k=j}^m \mu(A(k))\right) a_j - \sum_{j=1}^m \left(\sum_{k=j}^m \mu(A(k))\right) a_{j-1};$$

in the second sum we set $j - 1 = l$ so that this sum becomes, recalling that $a_0 = 0$

$$\sum_{l=0}^{m-1} \left(\sum_{k=l+1}^m \mu(A(k))\right) a_l = \sum_{l=1}^{m-1} \left(\sum_{k=l+1}^m \mu(A(k))\right) a_l;$$

putting again j in place of l in this sum we get:

$$= \sum_{j=1}^m \left(\sum_{k=j}^m \mu(A(k))\right) a_j - \sum_{j=1}^{m-1} \left(\sum_{k=j+1}^m \mu(A(k))\right) a_j = \mu(A(m))\,a_m + \sum_{j=1}^{m-1} \mu(A(j))\,a_j$$

which is of course $\int_X \varphi\,d\mu$. If for some $j > 0$ we have $\mu(A(j)) = \infty$ then $\int_X \varphi = \infty$, and trivially also $\int_0^\infty \rho_\varphi(t)\,dt = \infty$.

REMARK. Notice that if $f : X \to \infty]$ is positive measurable and $\alpha = \{a_0 = 0 < a_1 < \cdots < a_m\}$ is a subdivision of $[0, \infty[$ (see 3.2.2), letting $A(j) = f^{\leftarrow}(]a_j, a_{j+1}])$ for $j < m$, and $A(m)) = f^{\leftarrow}(]a_m, \infty])$ the function $f_\alpha = \sum_{j=0}^m a_j\,\chi_{A(j)}$ is exactly the one considered in 3.2.2, and the integral of f_α coincides with the integral of the step function $\rho_\alpha = \rho_{f_\alpha} : [0, \infty[\to [0, \infty]$:

$$\rho_\alpha(t) = \sum_{j=1}^m \rho_f(a_j)\,\chi_{[a_{j-1}, a_j[},$$

approximating ρ_f from below.

(v) is now immediate: if u_n is an increasing sequence of positive simple functions that increases to f, by (iii) ρ_{u_n} is a sequence of positive step functions which increases to ρ_f; we apply the monotone convergence theorem. $\square$

EXERCISE 4.1.4. Let $(X, \mathcal{M}, \mu)$ be a measure space. Prove, using the previous exercise, that for every $f \in \mathcal{L}^+$

$$\int_X f = \lim_{m\to\infty} \frac{1}{2^m} \sum_{j=1}^{m2^m} \mu(\{x \in X : f(x) > j/2^m\}).$$

EXERCISE 4.1.5. (Assumes that 4.1.14 is known) Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $f : X \to]0, \infty[$ be measurable and strictly positive. Assume that the decreasing distribution function ρ_f of f is finite valued for $t > 0$. The function $-\rho_f : [0, +\infty[\to \tilde{\mathbb{R}}$ is increasing and hence it has an associated Radon–Stieltjes measure $-d\rho_f$, finite on compact subintervals of $]0, \infty[$. Prove that this measure is the image measure $\mu f^{\leftarrow}$, and deduce that

$$\int_X f\,d\mu = \int_{[0,\infty[} t\,(-d\rho_f).$$

EXERCISE 4.1.6. Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $f \in L^+(X)$.

(i) Prove that if $\sum_{n=1}^{\infty} n\,\mu(f^{\leftarrow}(]n-1,n])) < \infty$, then $\int_X f < \infty$.
(ii) Find $f \in L^+(\mathbb{R})$ such that $\int_{\mathbb{R}} f < \infty$, but $\sum_{n=1}^{\infty} n\,\mu(f^{\leftarrow}(]n-1,n])) = \infty$, with $\mu = \lambda_1$, Lebesgue measure.
(iii) Find $f \in L^+(\mathbb{R})$ with $\int_{\mathbb{R}} f = \infty$ but

$$\sum_{n=1}^{\infty}(n-1)\,\mu(f^{\leftarrow}(]n-1,n])) < \infty \quad \text{(as above, } \mu = \text{Lebesgue measure).}$$

(iv) Assuming $\mu(X) < \infty$, prove that $\int_X f < \infty$ if and only if $\sum_{n=1}^{\infty}(n-1)\,\mu(f^{\leftarrow}(]n-1,n])) < \infty$.
(v) Always assuming $\mu(X) < \infty$ prove that if $f : X \to \mathbb{R}$ is measurable then $f^{\pm}$ have both finite integrals if and only if

$$\sum_{n\in\mathbb{Z}} |n-1|\,\mu(f^{\leftarrow}(]n-1,n])) < \infty.$$

EXERCISE 4.1.7. Compute ρ_f (see 4.1.14) for the measure space $(\mathbb{R}, \mathcal{L}_1, \lambda)$ and

$$f(x) = \sin x;\; f(x) = e^x;\; f(x) = e^{-|x|};\; f(x) = |x|^{-1/2}\,\chi_{]0,\infty[};\; f(x) = \frac{1}{1+x^2};$$

plot ρ_f whenever possible, and compute $\int_0^{\infty} \rho_f(t)\,dt$. Now consider the subspace of $\mathbb{R}^n$ given by $B = \{x \in \mathbb{R}^n : \|x\|_\infty = \max\{|x_k| : k = 1, \dots, n\} \le 1\}$, with measure induced by the $n-$dimensional Lebesgue measure, and let $f(x) = f_\alpha(x) = 1/\|x\|_\infty^\alpha$, where $\alpha > 0$. Find ρ_f, plot it, and use it to compute $\int_B f$.

EXERCISE 4.1.8. (*A look into countable subadditivity*) Let $(X, \mathcal{M}, \mu)$ be a measure space. Let $(A_n)_{n\in\mathbb{N}}$ be a sequence of sets in $\mathcal{M}$; for every $n \in \mathbb{N}$ let $\chi_n = \chi_{A_n}$ be the characteristic function of A_n; for every $x \in X$ let $\nu(x) = \sum_{n=0}^{\infty} \chi_n(x)$. Notice that $\nu(x)$ is the cardinality of the set $\{n \in \mathbb{N} : x \in A_n\}$ when this set is finite, and is ∞ otherwise; in particular, $\limsup_{n\to\infty} A_n = \{\nu = \infty\}$. Let $A = \bigcup_{n=0}^{\infty} A_n$. Prove that for every $f \in L^+$ we have

$$\int_A f \le \int_A \nu\, f = \sum_{n=0}^{\infty} \int_{A_n} f,$$

and prove that if ν is bounded, $\nu(x) \le m \in \mathbb{N}^{>}$ for every $x \in X$, then

$$\sum_{n=0}^{\infty} \int_{A_n} f \le m \int_A f \quad \text{in particular, with } f = 1 \text{ we get} \quad \sum_{n=0}^{\infty} \mu(A_n) \le m\,\mu\left(\bigcup_{n=0}^{\infty} A_n\right);$$

(if every x in the union belongs to at most m sets of the sequence then the sum of the measures of all the sets is at most m times the measure of the union).

4.2. The space $L^1(\mu)$ of summable functions. We now extend the integral to some not necessarily positive measurable functions.

DEFINITION. Given a measure space $(X, \mathcal{M}, \mu)$ we define $\mathcal{L}^1_\mu(X, \mathbb{K})$ as the set of all measurable functions $f : X \to \mathbb{K}$ such that $\|f\|_1 := \int_X |f| < \infty$.

Functions in $\mathcal{L}^1_\mu(X, \mathbb{K})$ are called integrable, or summable. We have $\|f\|_1 = 0$ if and only if $\mathrm{Coz}(|f|) = \mathrm{Coz}(f)$ has measure zero, that is, f is a null function (4.1.3). And for every $f \in \mathcal{L}^1(\mu)$ the cozero–set $\mathrm{Coz}(f)$ has $\sigma-$finite measure. Notice also that, by definition, functions in $\mathcal{L}^1_\mu$ are finite valued; then a function in $\mathcal{L}^+$ with finite integral is not necessarily in $\mathcal{L}^1(\mu)$; but it is always a.e. equal to a function in $\mathcal{L}^1(\mu)$, since it assumes the value ∞ only on a set of measure zero (4.1.6).

Remember: *nonzero constant functions belong to $\mathcal{L}^1_\mu(X, \mathbb{K})$ if and only if $\mu(X)$ is finite.*

4.2.1. *Integral of functions in $\mathcal{L}^1(\mu)$.* It is immediate to see that $\mathcal{L}^1_\mu(X, \mathbb{K})$ is a $\mathbb{K}-$vector subspace of $\mathbb{K}^X$: if f, g are measurable then $f+g$ is measurable and $|f+g| \le |f|+|g|$ so that $\|f+g\|_1 \le \|f\|_1 + \|g\|_1$ is also finite; and if $\alpha \in \mathbb{K}$ we have that αf is measurable and $|\alpha f| = |\alpha|\,|f|$ so that $\|\alpha f\|_1 = |\alpha|\,\|f\|_1$. There is a unique way of extending to $\mathcal{L}^1$ the integral already defined on $\mathcal{L}^+$ if we want it to be $\mathbb{K}-$linear: for a real valued f we set:

$$\int_X f := \int_X f^+ - \int_X f^-;$$

this is a real number: by hypothesis, $|f| = f^+ + f^-$ has a finite integral, and since $f^{\pm} \le |f|$ the integrals $\int_X f^{\pm}$ are both finite. Homogeneity is trivial, $\int_X (\alpha f) = \alpha \int_X f$ for real α and $f \in \mathcal{L}^1_\mu(X, \mathbb{R})$. Additivity:

if $h = f + g$ then we have $h^+ - h^- = f^+ - f^- + g^+ - g^- \iff h^+ + f^- + g^- = f^+ + g^+ + h^-$, whence
$$\int_X h^+ + \int_X f^- + \int_X g^- = \int_X f^+ + \int_X g^+ + \int_X h^- \iff \int_X h^+ - \int_X h^- = \int_X f^+ - \int_X f^- + \int_X g^+ - \int_X g^-.$$
We have seen that the integral on $\mathcal{L}^1_\mu(X, \mathbb{R})$ is a linear form. It is also isotone:

. *Assume that $f, g \in \mathcal{L}^1_\mu(X, \mathbb{R})$ and that $f \le g$. Then $\int_X f \le \int_X g$, and equality $\int_X f = \int_X g$ holds if and only if $\{f < g\}$ is a null set.*

Proof. $\int_X g - \int_X f = \int_X (g - f)$ and $g - f \in \mathcal{L}^+(X)$, with $\operatorname{Coz}(g - f) = \{f < g\}$; apply 4.1.3. □

4.2.2. *Integral of complex functions.* For a complex valued $f : X \to \mathbb{C}$ we have
$$|\operatorname{Re} f| \vee |\operatorname{Im} f| \le |f| \le |\operatorname{Re} f| + |\operatorname{Im} f|,$$
so that $f \in \mathcal{L}^1_\mu(X, \mathbb{C})$ if and only if $\operatorname{Re} f$ and $\operatorname{Im} f$ are both in $\mathcal{L}^1_\mu(X, \mathbb{R})$. We define, for $f \in \mathcal{L}^1_\mu(X, \mathbb{C})$:
$$\int_X f := \int_X \operatorname{Re} f + i \int_X \operatorname{Im} f;$$
straightforward calculations prove $\mathbb{C}-$linearity of this integral. We now prove:

. Fundamental inequality *For every $f \in \mathcal{L}^1_\mu(X, \mathbb{K})$ we have*
$$\left| \int_X f \right| \le \int_X |f| (= \|f\|_1).$$

Proof. For f real we have $-|f| \le f \le |f|$ so that by isotony
$$-\int_X |f| \le \int_X f \le \int_X |f| \iff \max\left\{ \int_X f, -\int_X f \right\} \le \int_X |f|.$$
For f complex–valued, notice first that the inequality is trivial if the integral of f is zero. Assume then $\int_X f = a \ne 0$ and let $\alpha = \overline{\operatorname{sgn}}(a)$. Then $|a| = \alpha\, a$ so that:
$$\left| \int_X f \right| = \alpha \int_X f = \int_X (\alpha\, f) = \int_X \operatorname{Re}(\alpha\, f),$$
the last equality because the integral of $\alpha\, f$ is real, so that $\int_X \operatorname{Im}(\alpha\, f) = 0$; then, by the fundamental inequality for real functions
$$\int_X \operatorname{Re}(\alpha\, f) \le \left| \int_X \operatorname{Re}(\alpha\, f) \right| \le \int_X |\operatorname{Re}(\alpha\, f)| \le \int_X |\alpha\, f| = \int_X |f|,$$
recalling that $|\alpha| = 1$. □

Exercise 4.2.1. Prove that the fundamental inequality is an equality if and only if f has sign almost everywhere constant on X.

Solution. (synthetic) In the real case the fundamental inequality is an equality iff either $\int_X f^- = 0$ or $\int_X f^+ = 0$, that is iff either $\{f < 0\}$ or $\{f > 0\}$ have measure 0; in the complex case the chain of inequalities used in the proof has to be a chain of equalities; we must have $\int_X (|\operatorname{Re}(\alpha\, f)| - \operatorname{Re}(\alpha\, f)) = 0$, and $\int_X (|\alpha\, f| - |\operatorname{Re}(\alpha\, f)|) = 0$, which is true iff $|\alpha\, f| = |\operatorname{Re}(\alpha\, f)| = \operatorname{Re}(\alpha\, f)$ a.e. on X. In other words $\alpha\, f(x) \ge 0$ for almost every $x \in X$, that is $f(x) \in \mathbb{R}_+ \operatorname{sgn}(a)$ for a.e. $x \in X$, with $a = \int_X f$.
□

4.2.3. The following will be used in 7.1.7

Proposition. *Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $f, g \in \mathcal{L}^1_\mu(X, \mathbb{C})$. Then $f = g$ a.e. if and only if $\int_E f = \int_E g$ for every $E \in \mathcal{M}$.*

Proof. Clearly $f = g$ a.e. implies $f\, \chi_E = g\, \chi_E$ a.e. for every $E \in \mathcal{M}$ and so also
$$\int_E f = \int_X f\, \chi_E = \int_X g\, \chi_E = \int_E g, \quad \text{for every} \quad E \in \mathcal{M}.$$
Conversely, assume $\int_E f = \int_E g$ for every $E \in \mathcal{M}$; then $\int_E \operatorname{Re} f = \int_E \operatorname{Re} g$ and $\int_E \operatorname{Im} f = \int_E \operatorname{Im} g$ for every $E \in \mathcal{M}$ and we are reduced to the case of real f, g. Setting $E = \{f > g\}$ we have $\int_E (f - g) \ge 0$, and since $f(x) > g(x)$ for every $x \in E$ we can have $\int_E f = \int_E g$ only if $\mu(E) = 0$ (4.2.1). Similarly, if $F = \{f < g\}$ we get $\mu(F) = 0$. Then $\{f \ne g\} = \{f > g\} \cup \{f < g\} = E \cup F$ has measure 0. □

REMARK. The statement:

. *If $f, g \in \mathcal{L}^+(X)$ are such that $\int_E f = \int_E g$ for every $E \in \mathcal{M}$, then $f = g$ a.e.*

is FALSE unless some additional hypothesis is made on μ: take an uncountable set X with the $\sigma-$algebra of countable or co–countable subsets, and the measure μ that is ∞ for co–countable, and 0 for countable sets: the constants 1 and 2 have integral 0 on countable and ∞ on co–countable sets, but are never equal. We can prove (but the proof is much more complicated than (ii) above, owing to possibly infinite integrals):

. *If μ is semifinite, and $f, g \in \mathcal{L}^+(X)$ are such that $\int_E f = \int_E g$ for every $E \in \mathcal{M}$, then $f = g$ a.e.*

Proof. Let $A = \{f < g\}$; it is enough to prove that $\mu(A) = 0$ (an analogous proof will work for $B = \{g < f\}$). Given $n \in \mathbb{N}$, let $E(n) = \{g \le n\} \cap A$. Then $\mu(E(n)) = 0$; in fact, if not, by semifiniteness we get $E \subseteq E(n)$ with $0 < \mu(E) < \infty$. Then $\int_E f = \int_E g \le \int_E n = n\,\mu(E) < \infty$; it follows that $\int_E (g - f) = 0$, but $g(x) - f(x) > 0$ for every $x \in E$, impossible if $\mu(E) > 0$. Then $\mu(E(n)) = 0$ for every n, so that $\mu(\{g < \infty\} \cap A) = 0$ (since $\{g < \infty\} = \bigcup_{n=1}^\infty \{g \le n\}$). If $\mu(\{g = \infty\} \cap A) > 0$ we still get a contradiction: notice that since $f(x) < \infty$ for every $x \in A$ we still have $\{g = \infty\} \cap A = \bigcup_{n=1}^\infty \{g = \infty\} \cap A \cap \{f \le n\}$; unless these sets have all measure zero we can get $E \subseteq \{g = \infty\} \cap A \cap \{f \le n\}$ with $0 < \mu(E) < \infty$; then $\int_E f \le n\,\mu(E) < \infty$, but $\int_E g = \infty$. Then $\mu(A) = \mu(A \cap \{g < \infty\}) + \mu(A \cap \{g = \infty\}) = 0$. □

4.2.4. *The normed space $L^1(\mu)$.* We have seen that on the vector space $\mathcal{L}^1_\mu(X, \mathbb{K})$ the function $f \mapsto \|f\|_1$ has all the properties of a norm, except that $\|f\|_1 = 0$ is equivalent to f being a null function, that is, $f = 0$ a.e. and not to f identically zero (it is only a *seminorm*, not a norm). The set $\mathcal{N}_\mu(X, \mathbb{K})$ of null functions is clearly a vector subspace of $\mathcal{L}^1_\mu(X, \mathbb{K})$; the quotient space $\mathcal{L}^1_\mu(X, \mathbb{K})/\mathcal{N}_\mu(X, \mathbb{K})$ is denoted by $L^1_\mu(X, \mathbb{K})$ or simply by $L^1(\mu)$ when the scalar field has been specified somehow. Its elements are the cosets $f + \mathcal{N}_\mu(X, \mathbb{K})$ of the subspace $\mathcal{N}_\mu(X, \mathbb{K})$: for a given f the coset $f + \mathcal{N}_\mu(X, \mathbb{K})$ is the set of all measurable functions a.e. equal to f. Clearly $\|f\|_1 = \|g\|_1$ for every pair of functions f and g equal a.e.; on the quotient space the function $f + \mathcal{N}_\mu(X, \mathbb{K}) \to \|f\|_1$ is well defined, and is a norm. But we often deliberately confuse $L^1(\mu)$ and $\mathcal{L}^1_\mu$, and we speak of a function $f \in L^1(\mu)$: it is a convenient abuse of language of which it is indispensable to be aware, and it is in general harmless. Sometimes we speak of the integral of a function f only defined a.e.: by 3.1.5, if a measurable $g : X \to \mathbb{K}$ that coincides with f a.e on the domain of f is in $L^1(\mu)$, then all such functions h are a.e. equal to each other, hence with the same integral. Convergence in $L^1(\mu)$ of a sequence $f_n \in L^1(\mu)$ to $f \in L^1(\mu)$ means of course convergence in this normed space, that is $\lim_n \|f - f_n\|_1 = 0$.

REMARK. If $f_n \in L^1(\mu)$ converges to $f \in L^1(\mu)$ in the L^1-norm, then clearly we also have

$$\lim_{n\to\infty} \int_X f_n = \int_X \lim_{n\to\infty} f_n, \quad \text{in fact:} \quad \left|\int_X f - \int_X f_n\right| = \left|\int_X (f - f_n)\right| \le \int_X |f - f_n| = \|f - f_n\|_1,$$

In other words, the integral is a continuous linear form on $L^1(\mu)$, i.e. convergence in L^1 implies that the limit may be carried under the integral sign, the integral of the L^1-limit is the limit of the integrals. But the fact that $\lim_{n\to\infty} \int_X f_n = \int_X f$, even if coupled with pointwise convergence of f_n to f, does *not* ensure convergence of f_n to f in the L^1-norm. An example is the following: $X = [0, 1]$, $m =$ Lebesgue measure, $f_n = n(\chi_{]0,1/(2n)]} - \chi_{]1/(2n),1/n]})$; we have that $f_n(x) \to 0$ for every $x \in [0, 1]$ (prove it), so that $f = 0$; and clearly we have

$$\int_{[0,1]} f_n = 0 \quad \text{but} \quad \int_{[0,1]} |f_n| = \int_0^1 n\chi_{]0,1/n]} = 1, \quad \text{for every } n \ge 1,$$

so that the integral of the limit is the limit of the integrals, but f_n does not converge to 0 in L^1, since $\|f_n\|_1 = 1$ for every $n \ge 1$. But there is a particular case in which this is true, which is of interest in probability theory:

. SCHEFFÉ'S THEOREM *Let $(X, \mathcal{M}, \mu)$ be a measure space; assume that f_n is sequence of positive functions in $L^1(\mu)$ converging pointwise a.e to f, and assume that*

$$\lim_{n\to\infty} \int_X f_n = \int_X f < \infty.$$

Then f_n converges to f in $L^1(\mu)$.

For the proof see exercise 4.2.5.

EXERCISE 4.2.2. Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $f \in L^1_\mu(X, \mathbb{K})$. Prove that if $g : \mathbb{K} \to \mathbb{K}$ is Lipschitz continuous, and $g(0) = 0$, then $g \circ f \in L^1_\mu(X, \mathbb{K})$. Find a function $f \in L^1_m([0, 1], \mathbb{R})$ such that $f^2 \notin L^1$; in particular, $L^1(\mu)$ in general is not closed under multiplication (m Lebesgue measure).

Solution. If $L > 0$ is a Lipschitz constant for g we have $|g(y)| = |g(y) - g(0)| \le L\,|y - 0| = L\,|y|$ for every $y \in \mathbb{R}$, so that $|g(f(x))| \le L\,|f(x)|$ for every $x \in X$; $g \circ f$ is measurable since so is f and g is continuous, and $L\,|f| \in L^1(\mu)$ since $f \in L^1(\mu)$. An example is $f(x) = 1/\sqrt{x}$ for $x \in]0,1]$, and $f(0) = 0$; clearly $f \in L^1_m([0,1])$, but $(f(x))^2 = 1/x \notin L^1_m([0,1])$.

□

4.2.5. *The dominated convergence theorem.* Monotone sequences are handled very well by the integral we are discussing. The Lebesgue dominated convergence theorem deals with not necessarily monotone sequences; we deduce it from Fatou's lemma.

Theorem. *Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions in $L^1(\mu)$ which converges a.e. to the function f; assume that there is $g \in L^1(\mu)$ such that $|f_n| \le g$, a.e. in X, for every $n \in \mathbb{N}$. Then $f \in L^1(\mu)$ and f_n converges to f in $L^1(\mu)$, so that we also have*

$$\lim_n \int_X f_n = \int_X f.$$

Proof. We can redefine f, g and all the f_n's if necessary on a null set so that convergence is everywhere and $|f_n(x)| \le g(x)$ holds everywhere (reset all these functions to zero on the union of all the countably many null sets on which these requirements are violated). Then $|f| \le g$ and f measurable imply $f \in L^1(\mu)$.

Notice that $|f - f_n| \le |f| + |f_n| \le g + g$, so that $h_n = 2g - |f - f_n| \ge 0$, and the sequence h_n converges pointwise to $2g$. We apply Fatou's lemma to h_n, obtaining

$$\int_X 2g \le \liminf_{n\to\infty} \int_X h_n = \liminf_{n\to\infty} \left(\int_X 2g - \int_X |f - f_n| \right) = \int_X 2g + \liminf_{n\to\infty} \left(- \int_X |f - f_n| \right) =$$
$$= \int_X 2g - \limsup_{n\to\infty} \int_X |f - f_n|,$$

which immediately implies

$$\limsup_{n\to\infty} \int_X |f - f_n| \le 0, \quad \text{equivalently} \quad \lim_{n\to\infty} \int_X |f - f_n| = 0,$$

which is the desired conclusion. □

4.2.6. *Normally convergent series in $L^1(\mu)$.* In a normed space $(E, \|\#\|)$ one can consider series of vectors, $\sum_{n=0}^{\infty} u_n$, with $u_n \in E$; as usual, we say that the series converges to $s \in E$ if the sequence $(s_m)_{m\in\mathbb{N}}$ of partial sums converges to s; by definition $s_m := \sum_{n=0}^{m} u_n$. In the normed space E we say that the series $\sum_{n=0}^{\infty} u_n$ is *normally convergent* if the series of the norms is convergent, i.e. $\sum_{n=0}^{\infty} \|u_n\| < \infty$: this name is perhaps badly chosen, since not necessarily a normally convergent series is convergent, see 4.2.10; but in $L^1(\mu)$ we have

. Theorem on normally convergent series in $L^1(\mu)$. *Let $\sum_{n=0}^{\infty} f_n$ be a series in $L^1(\mu)$ that is normally convergent, i.e. $\sum_{n=0}^{\infty} \|f_n\|_1 < \infty$. Then $\sum_{n=0}^{\infty} f_n$ converges a.e. and in $L^1(\mu)$ to a function $f \in L^1(\mu)$; and $\int_X f = \sum_{n=0}^{\infty} \int_X f_n$.*

Proof. For every n we have $|f_n| \in L^+$; if for every $x \in X$ we set $g(x) = \sum_{n=0}^{\infty} |f_n(x)|$, by the countable additivity of integrals of positive functions 4.1.5 we have $\int_X g = \sum_{n=0}^{\infty} \int_X |f_n| < \infty$. Then $\mu(\{g = \infty\}) = 0$ (4.1.6); we have that the series $\sum_{n=0}^{\infty} |f_n(x)|$ is convergent for every $x \in X \setminus \{g = \infty\}$; since absolute convergence implies convergence the series $\sum_{n=0}^{\infty} f_n(x)$ is convergent on $X \setminus \{g = \infty\}$; and

$$\left| \sum_{n=0}^{m} f_n(x) \right| \le \sum_{n=0}^{m} |f_n(x)| \le g(x) \quad \text{for every } x \in X \setminus \{g = \infty\},$$

so that dominated convergence theorem allows us to conclude. □

4.2.7. *Convergence in $L^1(\mu)$ and convergence almost everywhere.* Convergence in $L^1(\mu)$ does *not* in general imply a.e. convergence. The following is an example of a sequence converging to 0 in $L^1([0,1])$ with Lebesgue measure, which converges at no point of $[0,1]$. Start with $g_0 = 1$, the constant 1. Next let $g_{1,1}$ be the characteristic function of $[0,1/2]$, $g_{1,2}$ the characteristic function of $]1/2,1]$; in general $g_{n,1}$ is the characteristic function of $[0,1/2^n]$, while $g_{n,k}$, with $k = 2, \dots, 2^n$ is the characteristic function of $](k-1)/2^n, k/2^n]$. Order all the $g_{n,k}$ in a sequence $(f_m)_{m\in\mathbb{N}}$; it is clear that $f_m(x)$ does not converge for any $x \in [0,1]$ (we have $f_m(x) = 0$ and $f_n(x) = 1$ for infinitely many $m, n \in \mathbb{N}$), but since $\int_{[0,1]} g_{n,k} = 1/2^n$ clearly $\|f_m\|_1 \to 0$ as $m \to \infty$. However we have the following very important:

Proposition. *If $(f_n)_{n\in\mathbb{N}}$ converges to f in $L^1(\mu)$ then some subsequence $(f_{\nu(k)})_{k\in\mathbb{N}}$ of $(f_n)_{n\in\mathbb{N}}$ converges pointwise a.e. to f.*

Proof. Simply recall that in a metric space every Cauchy sequence contains a subsequence of bounded variation, i.e. a subsequence $(f_{\nu(k)})_{k\in\mathbb{N}}$ whose *total variation*

$$\sum_{k=1}^{\infty} \|f_{\nu(k)} - f_{\nu(k-1)}\|_1 < \infty$$

is finite (see e.g. Analisi Due, 2.11.7). Then the series

$$f_{\nu(0)} + \sum_{k=1}^{\infty}(f_{\nu(k)} - f_{\nu(k-1)})$$

is normally convergent in $L^1(\mu)$, so that its sequence of partial sums, which is $f_{\nu(k)}$, converges a.e. to its sum f (4.2.6). □

Exercise 4.2.3. The space $L^1(\varkappa)$ of the discrete measure space $(X, \mathcal{P}(X), \varkappa)$, where $\varkappa$ is the counting measure, is generally denoted $\ell^1(X) = \ell^1$ (see 4.3) Prove that convergence in ℓ^1 implies uniform convergence (in particular then also pointwise convergence).

4.2.8. *On dominated convergence.* If a sequence in $L^1(\mu)$ converges pointwise to a function and is dominated by a given function in $L^1(\mu)$ then it also converges in $L^1(\mu)$, as proved by Lebesgue's theorem. One might wonder if convergence in $L^1(\mu)$ coupled with pointwise convergence implies dominated convergence, that is: if a sequence $f_n \in L^1(\mu)$ converges to f, both in $L^1(\mu)$ and a.e., does there exist a function $g \in L^1(\mu)$ such that $|f_n| \le g$ for every $n \in \mathbb{N}$? The answer is negative:

Example 4.2.4. In the measure space $[0,1]$ with Lebesgue measure, for $n \ge 1$ and $k = 1, \dots, n$ let $f_{n,k} = n\,\chi_{](k-1)/n^2, k/n^2]}$. Then $\|f_{n,k}\|_1 = 1/n$, so that arranging all f_{nk} in a sequence f_m this converges to 0 both a.e. and in L^1. But $g = \bigvee_{n,k} f_{n,k} \notin L^1$: in fact $g(x) = \sum_{n=1}^{\infty} n\,\chi_{]1/(n+1),1/n]}$ so that

$$\int_{[0,1]} g = \sum_{n=1}^{\infty} n\left(\frac{1}{n} - \frac{1}{n+1}\right) = \sum_{n=1}^{\infty} \frac{1}{n+1} = \infty.$$

Of course, as seen above, a sequence of bounded variation in $L^1(\mu)$ of functions in $L^1(\mu)$ is dominated by a function of $L^1(\mu)$.

Exercise 4.2.5. Prove Scheffé's theorem (4.2.4) in the following way: prove first that $(f - f_n)^+ \le f$; apply dominated convergence to prove that $\lim_{n\to\infty} \int_X (f - f_n)^+ = 0$; the rest is easy ...

Solution. Since $f_n(x) \ge 0$ we have $f(x) - f_n(x) \le f(x)$; since the positive part $t \mapsto t^+ = \max\{t, 0\}$ is an increasing function we have $(f(x) - f_n(x))^+ \le (f(x))^+ = f(x)$, so that $(f - f_n)^+ \le f$ for every n. Clearly $f \in L^1(\mu)$ (it is a positive function with a finite integral); by continuity of the positive part, we have $\lim_{n\to\infty}(f - f_n)^+ = 0$ a.e., so that by dominated convergence $\lim_{n\to\infty} \int_X (f - f_n)^+ = 0$. Now we have $|f - f_n| = (f - f_n)^+ + (f - f_n)^-$ and $f - f_n = (f - f_n)^+ - (f - f_n)^-$, so that

$$|f - f_n| = 2(f - f_n)^+ - (f - f_n) \Longrightarrow \int_X |f - f_n| = 2\int_X (f - f_n)^+ - \int_X (f - f_n),$$

and taking limits as $n \to \infty$ we get zero on the right–hand side. □

A simpler proof may be obtained using the generalized dominated convergence theorem 4.4.7

4.2.9. *Approximation of L^1 functions.* We denote by $S(\mu)$ the subspace of $L^1(\mu)$ consisting of simple functions, so that $S(\mu)$ is the space of integrable simple functions: it consists of all measurable simple functions which are 0 outside a set of finite measure. It is clear that

Lemma. *$S(\mu)$ is dense in $L^1(\mu)$*

Proof. If $(\varphi_n)_{n\in\mathbb{N}}$ is a sequence of measurable simple functions which converges pointwise to $f \in L^1(\mu)$ and such that $|\varphi_n| \uparrow |f|$ (3.2.2) we have $\varphi_n \in S(\mu)$ (since $|\varphi_n| \le |f| \in L^1(\mu)$) and by dominated convergence $\lim_n \|f - \varphi_n\|_1 = 0$. □

But measurable sets are often quite complicated, so simple functions are not really simple. We are interested in approximating functions in $L^1(\mu)$ with functions we know better about.

Proposition. *Let $(X, \mathcal{M}, \mu)$ be the Carathèodory extension of a premeasure on an algebra of parts of X. Then the space of $\mathcal{A}-$simple integrable functions is dense in $L^1(\mu)$. If $X = \mathbb{R}^n$ and μ is a Radon measure, then the integrable step functions are dense in $L^1(\mu)$; and continuous functions with compact support are also dense in $L^1(\mu)$.*

Proof. We need to prove that every function $\varphi \in S(\mu)$ is approximable with $\mathcal{A}-$simple functions in the $L^1(\mu)$ norm. Let $\varphi = \sum_{j=1}^m a_j\, \chi_{E(j)}$, with $a_j \neq 0$ and $0 < \mu(E(j)) < \infty$ for every $j = 1, \dots, m$; set $a = \sum_{j=1}^m |a_j|$. Given $\varepsilon > 0$, by 2.5.7 there is $A(j) \in \mathcal{A}$ such that $\mu(E(j) \vartriangle A(j)) \leq \varepsilon/a$, for $j = 1, \dots, m$. Then the function $\psi = \sum_{j=1}^m a_j\, \chi_{A(j)}$ is $\mathcal{A}-$simple, and:

$$\|\varphi - \psi\|_1 = \int_X \left| \sum_{j=1}^m a_j (\chi_{E(j)} - \chi_{A(j)}) \right| \leq \int_X \sum_{j=1}^m |a_j|\, |\chi_{E(j)} - \chi_{A(j)}| = \sum_{j=1}^m |a_j| \int_X |\chi_{E(j)} - \chi_{A(j)}| =$$
$$\sum_{j=1}^m |a_j|\, \mu((E(j) \vartriangle A(j))) \leq \varepsilon.$$

If the algebra $\mathcal{A}$ is the interval algebra then every $A(j)$ is a finite disjoint union of bounded intervals. To conclude that continuous functions with compact support are dense in $L^1_\mu(\mathbb{R}^n)$ it is clearly enough to prove that for every bounded interval I and every $\varepsilon > 0$ there is a continuous function u with compact support such that $\|\chi_I - u\|_1 \leq \varepsilon$. Remember that we can find a compact interval $K \subseteq I$ and a bounded open interval $V \supseteq I$ such that $\mu(V \smallsetminus K) \leq \varepsilon$ (see 2.2.5). By 1.3.3 there exists a continuous function $u : \mathbb{R}^n \to [0, 1]$ such that $u(x) = 0$ for $x \notin V$ (so that $\mathrm{Supp}(u) \subseteq \bar{V}$ is compact) and $u(x) = 1$ for $x \in K$; accepting this result we have

$$|\chi_I - u| \leq \chi_{V \smallsetminus K} \quad \text{so that} \quad \|\chi_I - u\|_1 \leq \int_{\mathbb{R}^n} \chi_{V \smallsetminus K}\, d\mu = \mu(V \smallsetminus K) \leq \varepsilon.$$

Remark. Using the remark in 1.3.3 we see that u may even be supposed C^∞.

□

4.2.10. *Completeness of* $L^1(\mu)$. The space $L^1(\mu)$ is a Banach space, i.e. it is a complete normed space, as we now see. Remember that a normed space $(E, \|\#\|)$ is called complete when every Cauchy sequence of E has a limit in E; and in a normed space this is equivalent to assert that every normally convergent series of E is convergent in E (see e.g. Analisi Due, 2.12.4; let's quickly review the argument: a Cauchy sequence with a converging subsequence is convergent; every Cauchy sequence has a subsequence of bounded variation, and convergence of sequences of bounded variation is trivially equivalent to convergence of normally convergent series). Thus the theorem on normally convergent series expresses completeness of $L^1(\mu)$.

Completeness is perhaps the main profit obtained from this more general theory of integration.

4.3. The space $\ell^1(X)$. We now take a look at the discrete measure space $(X, \mathcal{P}(X), \varkappa)$, where $\varkappa$ is the counting measure. All functions are of course measurable. The only set of measure 0 is the empty set, and the sets of finite measure are exactly the finite subsets, so that simple functions with finite integral are exactly the functions which are finite–valued and nonzero only on a finite set. It is quite obvious that for every positive function $f : X \to [0, \infty]$ we have:

$$\int_X f(x)\, d\varkappa(x) = \sup \left\{ \sum_{x \in F} f(x) : \; F \subseteq X,\, F \text{ finite} \right\}$$

(by definition, the integral is the supremum of the integrals of positive simple functions dominated by f; every simple function dominated by f with finite integral is then of the form $\sum_{x \in F} \varphi(x)\, \chi_{\{x\}}$, with F a finite subset of X, and $0 \leq \varphi(x) \leq f(x)$, $\varphi(x) < \infty$, for every $x \in F$ etc, …). In other words, the integral in the counting measure is exactly the infinite sum $\sum_X f$ described in 1.10. Notice that the necessary condition for finiteness of the sum given in 1.10.3, that $\{f \neq 0\}$ be countable, is exactly the $\sigma-$finiteness of the measure of $\mathrm{Coz}(f)$, if f has a finite integral. We have then that $f \in L^1(\varkappa)$ if and only if f is finite valued and representable as $f = (u^+ - u^-) + i\,(v^+ - v^-)$, with all the four positive functions $u^\pm$, $v^\pm$ summable; of course $\mathrm{Coz}(f)$ is at most countable, for every f with finite sum, positive or not.

It is customary to write $\ell^1(X)$ instead of $L^1(\varkappa)$; the integral of f over $A \subseteq X$ will be denoted $\sum_A f$. Instead of saying that $f \in \ell^1(X)$ we also say that f is *summable* over X.

4.3.1. *Unrestricted additivity.* Unrestricted associativity holds also for sums of functions in $\ell^1(X)$, and it may also be thought of as *unrestricted additivity* of infinite sums; this is trivial, since it is true for positive functions

. Unrestricted additivity for summable functions. *Let X be a set, and let $f : X \to \mathbb{K}$ be a function; let $(A(\lambda))_{\lambda\in\Lambda}$ be a partition of X. Then*

(i) *If $f \in \ell^1(X)$ then the function $s : \Lambda \to \mathbb{K}$ defined by $s(\lambda) = \sum_{A(\lambda)} f$ is in $\ell^1(\Lambda)$ and*

$$\sum_X f = \sum_\Lambda s \left(= \sum_{\lambda\in\Lambda} \left(\sum_{x\in A(\lambda)} f(x) \right) \right). \tag{*}$$

(ii) *If $f \in \ell^1(A(\lambda))$ for every $\lambda \in \Lambda$, and*

$$\sum_{\lambda\in\Lambda} \left(\sum_{A(\lambda)} |f| \right) < \infty, \tag{**}$$

then $f \in \ell^1(X)$ (consequently, (i) holds and the sum on X of f may be obtained as in (*), by removing the absolute values).

Proof. (i) has been observed above: it is a consequence of 1.10.1 and of the linearity of sums.
(ii) 1.10.1 says that the sum in (**) is exactly $\|f\|_1 = \sum_X |f|$. □

We present some important applications of this fact.
4.3.2. *Fubini–Tonelli, the discrete case.*

Proposition. *Let X, Y be sets, and let $f : X \times Y \to \mathbb{K}$ be a function.*

(i) (Fubini) *Assume that $f \in \ell^1(X \times Y, \mathbb{K})$. Then for every $x \in X$ the function $y \mapsto f(x, y)$ is in $\ell^1(Y)$, the function $x \mapsto \sum_{y\in Y} f(x, y)$ is in $\ell^1(X)$ (the same is true exchanging X and Y) and*

$$\sum_{(x,y)\in X\times Y} f(x, y) = \sum_{x\in X} \left(\sum_{y\in Y} f(x, y) \right) = \sum_{y\in Y} \left(\sum_{x\in X} f(x, y) \right).$$

(ii) (Tonelli) *If*

$$\sum_{x\in X} \left(\sum_{y\in Y} |f(x, y)| \right) < \infty, \quad \text{or if} \quad \sum_{y\in Y} \left(\sum_{x\in X} |f(x, y)| \right) < \infty,$$

then $f \in \ell^1(X \times Y, \mathbb{K})$ (so that the sum of f on $X \times Y$ may be computed as in (i)).

Proof. (i) is simply unrestricted additivity 4.3.1 with the partitions $(\{x\} \times Y)_{x\in X}$ or $(X \times \{y\})_{y\in Y}$; (ii) is the second part of 4.3.1 □

As an important consequence of the above proposition we prove the homomorphism property of the complex exponential. Recall that by definition $\exp z = \sum_{n=0}^\infty z^n/n!$, where the series is absolutely convergent for every $z \in \mathbb{C}$. We have

$$\exp(w + z) = \sum_{n=0}^\infty \frac{(w + z)^n}{n!} = \sum_{n=0}^\infty \left(\sum_{k=0}^n \binom{n}{k} \frac{w^{n-k}\, z^k}{n!} \right) = \sum_{n=0}^\infty \left(\sum_{k=0}^n \frac{w^{n-k}\, z^k}{(n - k)!\, k!} \right).$$

On $\mathbb{N} \times \mathbb{N}$ consider the function $(j, k) \mapsto w^j\, z^k/(j!\, k!)$; the preceding sum is the sum of this function on the partition given by $M(n) = \{(j, k) \in \mathbb{N} \times \mathbb{N} : j + k = n\}$; moreover we certainly have summability, since by absolute convergence w and z may be replaced by their absolute values, leaving the sum finite with value $\exp(|w| + |z|)$; thus (Fubini):

$$\exp(w + z) = \sum_{n=0}^\infty \left(\sum_{k=0}^n \frac{w^{n-k}\, z^k}{(n - k)!\, k!} \right) = \sum_{(j,k)\in\mathbb{N}\times\mathbb{N}} \frac{w^j\, z^k}{j!\, k!} = \sum_{j\in\mathbb{N}} \left(\sum_{k\in\mathbb{N}} \frac{w^j\, z^k}{j!\, k!} \right) =$$

$$\sum_{j\in\mathbb{N}} \frac{w^j}{j!} \left(\sum_{k\in\mathbb{N}} \frac{z^k}{k!} \right) = \sum_{j\in\mathbb{N}} \frac{w^j}{j!} \exp z = \exp w \, \exp z.$$

EXERCISE 4.3.1. If $f, g \in \mathbb{K}^{\mathbb{Z}}$ their *convolution* is, by definition, the function $f * g : \mathbb{Z} \to \mathbb{K}$ defined by

$$f * g(x) = \sum_{y \in \mathbb{Z}} f(x-y)\, g(y)$$

whenever this formula is meaningful, that is, the right hand side is defined somehow.

(i) Prove that if $f, g \in \ell^1(\mathbb{Z})$ then $f * g$ is defined on $\mathbb{Z}$, that $f * g \in \ell^1(\mathbb{Z})$ and that

$$\|f * g\|_1 \le \||f| * |g|\|_1 = \|f\|_1\, \|g\|_1; \quad \sum_{\mathbb{Z}} f * g = \sum_{\mathbb{Z}} f \sum_{\mathbb{Z}} g.$$

If $a : n \mapsto a_n$ and $b : n \mapsto b_n$ are two sequences ($n \in \mathbb{N}$) they may be considered as two–sided sequences $a, b \in \mathbb{K}^{\mathbb{Z}}$ by assuming $a_n = b_n = 0$ for $n < 0$; their convolution $a * b$ is then always defined, and

$$(a * b)_n = \sum_{k=0}^{n} a_{n-k}\, b_k \quad \text{for } n \ge 0, \text{ while } (a * b)_n = 0 \text{ for } n < 0.$$

Given the two series $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$, their *Cauchy product* is, by definition, the series

$$\sum_{n=0}^{\infty} (a * b)_n = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_{n-k}\, b_k \right).$$

(ii) Prove that if a and b are absolutely convergent then their Cauchy product is also absolutely convergent and its sum is the product of the sums $(\sum_{\mathbb{N}} a)(\sum_{\mathbb{N}} b)$ (this is of course a particular case of (i), and includes the preceding result on the homomorphism formula of exp as a particular case).

EXERCISE 4.3.2. If $f \in \mathbb{K}^X$ and $g \in \mathbb{K}^Y$, their *tensor product* $f \otimes g : X \times Y \to \mathbb{K}$ is defined by $f \otimes g(x, y) = f(x)\, g(y)$. Prove that if $f \in \ell^1(X, \mathbb{K})$ and $g \in \ell^1(Y, \mathbb{K})$ then $f \otimes g \in \ell^1(X \times Y, \mathbb{K})$ and $\|f \otimes g\|_1 = \|f\|_1\, \|g\|_1$.

4.3.3. *Euler's product formula for the ζ function.* As a final result we present the famous Euler's product formula for the zeta function: for $s > 1$ we set

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s};$$

(we know that $n \mapsto 1/n^s$ is summable over $\mathbb{N}^>$ if $s > 1$). Given a prime p, every $n \in \mathbb{N}^>$ may be written uniquely as $n = p^{\nu(n)}\, m(n)$, where $m(n)$ is prime to p, and $\nu(n) \ge 0$; we partition $\mathbb{N}^>$ in sets $M(m) = \{p^\nu\, m : \nu \in \mathbb{N}\}$, where m ranges in the set $N(p)$ of all integers in $\mathbb{N}^>$ prime to p; then

$$\zeta(s) = \sum_{m \in N(p)} \left(\sum_{\nu=0}^{\infty} \frac{1}{(p^\nu\, m)^s} \right) = \sum_{m \in N(p)} \frac{1}{m^s} \left(\sum_{\nu=0}^{\infty} \frac{1}{(p^s)^\nu} \right) = \frac{1}{1 - p^{-s}} \sum_{m \in N(p)} \frac{1}{m^s}.$$

Given another prime q, the argument may be repeated to show that

$$\sum_{m \in N(p)} \frac{1}{m^s} = \frac{1}{1 - q^{-s}} \sum_{m \in N(p,q)} \frac{1}{m^s},$$

where now $N(p, q)$ is the set of all strictly positive integers prime to both p and q. If $p_1, \dots, p_j, \dots$ is a bijective elencation of all primes (in any order) we then have, for every integer r:

$$(*) \qquad \zeta(s) = \prod_{j=1}^{r} \frac{1}{1 - p_j^{-s}} \sum_{m \in N(P_r)} \frac{1}{m^s},$$

where $N(P_r)$ is the set of all strictly positive integers whose factorization does not contain any of the primes $p_1, \dots, p_r$. Notice that $N(P_1) \supseteq N(P_2) \supseteq N(P_3) \supseteq \dots$, and moreover $\bigcap_{r=1}^{\infty} N(P_r) = \{1\}$: every integer strictly larger than 1 has a prime divisor. From this one easily sees that

$$\lim_{r \to \infty} \left(\sum_{m \in N(P_r)} \frac{1}{m^s} \right) = 1,$$

so that, passing to the limit in (*) as $r \to \infty$ we get

$$\zeta(s) = \prod_{j=1}^{\infty} \frac{1}{1 - p_j^{-s}} \left(:= \lim_{r \to \infty} \prod_{j=1}^{r} \frac{1}{1 - p_j^{-s}} \right)$$

one of the most beautiful results of Euler.

Exercise 4.3.3. Prove that the sum, where $p \in \mathbb{R}$:

$$\sum_{(m,n)\in\mathbb{N}\times\mathbb{N}\smallsetminus\{(0,0)\}} \frac{1}{(m+n)^p}$$

is finite if and only if $p > 2$, and in this case its value is $\zeta(p-1)+\zeta(p)$. (hint: sum on the blocks $M(k) = \{(m,n)\in\mathbb{N}\times\mathbb{N} : m+n = k\}$, with $k = 1,2,3,\dots$).

4.4. Uniform convergence almost everywhere and $L^\infty(\mu)$. Let $(X,\mathcal{M},\mu)$ be a measure space, and let $g \in L(X,\mathbb{R})$ be measurable. In exercise 4.1.14 we defined the decreasing distribution function ρ_g of g as $\rho_g(t) = \mu(\{g > t\})$; this function is positive, decreasing and right–continuous, so that if $\rho_g(t) = 0$ for some $t > 0$ we have $\rho_g(s) = 0$ for all $s > t$; the set $\{t \in \mathbb{R}_+ : \rho_g(t) = 0\}$, if non–empty, is then a closed half–line $[a,+\infty[$ (closed because of right continuity). This number a is called the *essential supremum* of g; in other words

$$\mathrm{essup}(g) = \inf\{t \in \mathbb{R} : \mu(\{g > t\}) = 0\}$$

(when no such t exists, then the essential supremum of g is ∞; when the set is non–empty, the infimum is actually a minimum, as seen above). For every measurable $f \in L(X,\mathbb{K})$ we define

$$\|f\|_\infty = \mathrm{essup}(|f|) = \inf\{t \in \mathbb{R} : \mu(\{|f| > t\}) = 0\};$$

notice that $\|f\|_\infty = \sup\{|f(x)| : x \in X \smallsetminus \{|f| > \|f\|_\infty\}$. The essentially bounded functions are those for which this quantity is finite, they form a linear subspace $\mathcal{L}^\infty_\mu(X,\mathbb{K}) \subseteq \mathcal{L}(X,\mathbb{K})$; then $\|\#\|_\infty$ is a norm on $L^\infty(\mu) = L^\infty_\mu(X,\mathbb{K}) = \mathcal{L}^\infty_\mu(X,\mathbb{K})/\mathcal{N}_\mu(X,\mathbb{K})$, the essential supremum norm. Of course if the measure is such that the only null set is the empty set (e.g. the counting measure, where L^∞ is denoted ℓ^∞), then the essential supremum norm is exactly the sup–norm. We leave to the reader the proof of the following fact:

. *A sequence $f_n \in \mathcal{L}^\infty(\mu)$ converges to $f \in \mathcal{L}^\infty(\mu)$ in the essential supremum norm if and only if there is a null set $N \subseteq X$ such that on $X \smallsetminus N$ the sequence f_n converges uniformly to f. The space $(L^\infty(\mu), \|\#\|_\infty)$ is a Banach space; it is also closed under (pointwise) multiplication, in fact*

$$\|f\,g\|_\infty \le \|f\|_\infty\,\|g\|_\infty,$$

so $L^\infty(\mu)$ is also a Banach algebra.

Exercise 4.4.1. (*Easy but important*) Let $(X,\mathcal{M},\mu)$ be a measure space. Prove that $L^\infty(\mu) \subseteq L^1(\mu)$ if and only if $\mu(X) < \infty$ (bounded measurable functions are summable on finite measure spaces) and in this case $\|f\|_1 \le \mu(X)\,\|f\|_\infty$. Deduce that on a finite measure space convergence in L^∞ implies convergence in L^1. The sequence $f_n = \chi_{[0,n]}/n$ in $L^1_\lambda(\mathbb{R})$ proves that this is not true in absence of finiteness of the measure space.

Exercise 4.4.2. We have seen that in general $L^1(\mu)$ is not closed under multiplication (4.2.2). However: if $f \in L^1(\mu)$ and $g \in L^\infty(\mu)$ then $g\,f \in L^1(\mu)$, and $\|g\,f\|_1 \le \|g\|_\infty\,\|f\|_1$ (easy). If μ is semifinite, then we have also:

. *Let $g \in L(X)$ be such that $g\,f \in L^1(\mu)$ for every $f \in L^1(\mu)$. Then $g \in L^\infty(\mu)$.*

(requires some ingenuity ...).

Exercise 4.4.3. Let X be an uncountable set, $\mathcal{M}$ the $\sigma-$algebra of countable or co–countable subsets.

(i) Prove that $f : X \to \mathbb{K}$ is measurable iff it is constant on some co–countable subset of X (hint: find the simple functions first and recall that every measurable function is the pointwise limit of a sequence of simple functions).
(ii) Let $\mu : \mathcal{M} \to [0,\infty]$ be the measure which is zero on countable sets, ∞ on the others. Prove that $L^1(\mu) = \{0\}$.
(iii) With μ as in (ii) describe $\mathcal{L}^\infty_\mu(X,\mathbb{K})$ and prove that $L^\infty(\mu)$ is isomorphic to $\mathbb{K}$ (as a vector space and as a $\mathbb{K}-$algebra).

Exercise 4.4.4. Let X be a set. Prove that if $f \in \ell^1(X)$ then there is $c \in X$ such that $|f(c)| = \max|f|(X)$; deduce that $\ell^1 \subseteq \ell^\infty$ and that $\|f\|_\infty \le \|f\|_1$, for every $f \in \ell^1$; equality $\|f\|_\infty = \|f\|_1$ holds only for functions whose support is a singleton, or empty.

Solution. Assume that $f \in \ell^1$ is not identically zero; then there is $a \in X$ such that $|f(a)| > 0$; the set $E = \{|f| \geq |f(a)|\}$ is then finite (by 1.10.3) and non–empty because $a \in E$; then there is $c \in E$ such that $|f(c)| = \max\{|f(x)| : x \in E\}$; we also have $|f(c)| = \|f\|_\infty$, because if $x \notin E$ we have $|f(x)| < |f(a)| \leq |f(c)|$. Clearly $\|f\|_1 = \sum_{x \in X} |f(x)| \geq |f(c)|$, and if there is $a \in X \smallsetminus \{c\}$ such that $f(a) \neq 0$ we have

$$\|f\|_\infty = |f(c)| < |f(c)| + |f(a)| \leq \|f\|_1.$$

$\square$

EXERCISE 4.4.5. Let $(X, \mathcal{M}, \mu)$ be a measure space and let $F \subseteq \mathbb{K}$ be closed in $\mathbb{K}$. Denote by $L^1_\mu(X, F)$ the set of all $f \in L^1(\mu)$ such that $f(x) \in F$ for a.e. $x \in X$. Prove that $L^1_\mu(X, F)$ is a closed subset of $L^1(\mu)$. Identifying two measurable sets $A, B \in \mathcal{M}$ such that $\mu(A \vartriangle B) = 0$, prove that the ideal $\mathcal{F}(\mu) = \{A \in \mathcal{M} : \mu(A) < \infty\}$ is a complete metric space in the metric $\rho(A, B) = \mu(A \vartriangle B)$ (consider $L^1_\mu(X, \{0, 1\}) \dots$).

EXERCISE 4.4.6. (*One–sided convergence*) Let $(X, \mathcal{M}, \mu)$ be a measure space. Let $(f_n)_n$ be a sequence in $L^1_\mu(X, \mathbb{R})$.

(i) Assume that $0 \leq f_n \leq f$ for every n, and that $f_n \to f$ a.e. Then $\int_X f_n \to \int_X f$ (this last may be finite or $+\infty$).
(ii) The conclusion of (i) remains true assuming $g \leq f_n \leq f$ in place of $0 \leq f_n \leq f$, provided that $g \in L^1_\mu(X, \mathbb{R})$.

Solution. (i) Fatou's lemma says that $\int_X f \leq \liminf_n \int_X f_n$; and since $\int_X f_n \leq \int_X f$ for every n we get $\limsup_n \int_X f_n \leq \int_X f$ and we conclude; (ii) simply take $f_n - g$, $f - g$ .... $\square$

EXERCISE 4.4.7. *A generalized dominated convergence theorem*

. *Let $(X, \mathcal{M}, \mu)$ be a measure space. Assume that we have two sequences $f_n, g_n \in L^1(\mu)$, a.e. converging to f and g respectively, that $|f_n| \leq g_n$ a.e. for every $n \in \mathbb{N}$, and $\lim_n \int_X g_n = \int_X g < \infty$. Then $f \in L^1(\mu)$, we have*

$$\lim_n \int_X f_n = \int_X f,$$

and f_n converges to f in $L^1(\mu)$.

COROLLARY. *If f_n is a sequence in $L^1(\mu)$ converging a.e. to f and $\lim_n \|f_n\|_1 = \|f\|_1 < \infty$, then f_n converges to f in $L^1(\mu)$.*

Solution. We have $|f - f_n| \leq |f| + |f_n| \leq g + g_n$, and $\lim_{n\to\infty}(g + g_n) = 2g$ a.e. on X. We apply Fatou's lemma to the sequence $(g + g_n) - |f - f_n| \geq 0$, obtaining

$$\int_X 2g \leq \liminf_{n\to\infty} \left(\int_X g + \int_X g_n - \int_X |f - f_n| \right) = \int_X 2g - \limsup_{n\to\infty} \int_X |f - f_n|,$$

and cancelling $\int_X 2g$ we get

$$\limsup_{n\to\infty} \int_X |f - f_n| \leq 0.$$

For the corollary, apply the above with $g_n = |f_n|$. $\square$

REMARK. Notice that Scheffé's theorem (4.2.4) is a particular case of the corollary, with $f_n = |f_n|$.

EXERCISE 4.4.8. Compute the limits:

$$\lim_n \int_0^1 \frac{1 - \sin(x/n)}{\sqrt{x^2 + 1/n}}\, dx; \quad \lim_n \int_1^{2011} \frac{1 - \sin(x/n)}{\sqrt{x^2 + 1/n}}\, dx;$$

$$\lim_n \int_1^n \frac{1 - \sin^4(x/n)}{x + 1/n}\, dx; \quad \lim_n \int_0^\infty \left(\frac{\sin(x/n)}{x} \right)^3 dx.$$

(you may use the convergence theorems, but for some of the above also trivial estimates work; for the last limit use $\int_0^\infty = \int_0^1 + \int_1^\infty$). Same questions for:

$$\lim_n \int_0^{2011} \frac{n\,\cos(x/n)}{nx + 1}\, dx; \quad \lim_n \int_1^n \frac{\cos^2(x/n)}{x}\, dx; \quad \lim_n \int_0^\infty \frac{\sin^2(x/n)}{x^{5/2}}\, dx.$$

EXERCISE 4.4.9. Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $f \in L^+$, with $0 < \int_X f < \infty$. Prove that for every $\alpha > 0$ and $n \in \mathbb{N}^>$ the function $g(x) = g_{n,\alpha}(x) = n \sin(f(x)/n^\alpha)$ is in $L^1(\mu)$, and compute, expressing the answer in terms of $\int_X f$:

$$A(\alpha) = \lim_{n\to\infty} \int_X g_{n,\alpha}(x)\, d\mu(x)$$

(hint: get to a situation in which you can use dominated convergence; recall that $|\sin t/t| \le 1 \dots$).

EXERCISE 4.4.10. (*assumes 4.1.14 and 4.1.5*) Let $(X, \mathcal{M}, \mu)$ be a finite measure space, $0 < \mu(X) < \infty$, and let $f : X \to \mathbb{R}$ be measurable. The (increasing) distribution function of f is the function $F_f : \mathbb{R} \to \mathbb{R}$ defined by $F(t) = F_f(t) = \mu(\{f \le t\}) = \mu(f^{\leftarrow}(]-\infty, t]))$.

(i) Prove that F_f is increasing and right continuous, and express it by means of the decreasing distribution function $\rho_f(t) = \mu(\{f > t\})$.

(ii) Prove that:

For $t < 0$: $F_f(t) = \rho_{f^-}((-t)^-)$ (left limit of ρ_{f^-} in $-t > 0$).

For $t \ge 0$: $F_f(t) = \mu(X) - \rho_{f^+}(t)$.

and deduce from that and what proved in Exercise 4.1.14 that $f \in L^1(\mu)$ if and only if

$$\int_{-\infty}^0 F_f(t)\, dt \quad \text{and} \quad \int_0^\infty (\mu(X) - F_f(t))\, dt$$

are both finite, and in that case:

$$\int_X f = \int_0^\infty (\mu(X) - F_f(t))\, dt - \int_{-\infty}^0 F_f(t)\, dt.$$

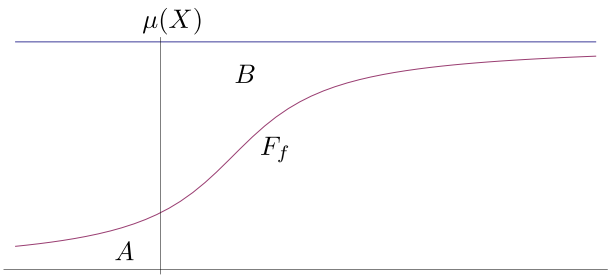

FIGURE 2. If F_f is as plotted, we have $f \in L^1(\mu)$ iff the two regions A and B have finite area, and in that case $\int_X f = \text{area}(B) - \text{area}(A)$.

(iv) Prove that the image measure $\mu F^{\leftarrow}$ (see 4.1.5) is the Radon–Stieltjes measure dF_f, and deduce from it that if $f \in L^1(\mu)$ then

$$\int_X f\, d\mu = \int_{\mathbb{R}} t\, dF_f(t).$$

EXERCISE 4.4.11. Let $(X, \mathcal{M}, \mu)$ be a measure space. If $E \in \mathcal{M}$ is such that $0 < \mu(E) < \infty$, and $f : X \to \mathbb{K}$ is such that $f_{|E} \in L^1_\mu(E)$, we define the *average* of f over E as

$$⨍_E f := \frac{1}{\mu(E)} \int_E f \left(= \int_E f \frac{d\mu}{\mu(E)} \right)$$

(it may be interpreted as the integral on E with respect to the "rescaled" measure $\mu/\mu(E)$). Assume now that $C \subseteq \mathbb{K}$ is a closed subset of $\mathbb{K}$ such that $⨍_E f \in C$ for every E of finite nonzero measure, and that the measure is semifinite. Prove that then $f(x) \in C$ for a.e. $x \in X$ (hint: prove that for every open disk $B(c, r[\subseteq \mathbb{K} \setminus C$ we have $\mu(f^{\leftarrow}(B(c, r[)) = 0$; given a countable base for the open sets of $\mathbb{K}$ consisting of open disks $\dots$).

EXERCISE 4.4.12. Using the power series expansion of $\sin x/x$ prove that if $a > 1$:

$$\int_0^\infty e^{-ax} \frac{\sin x}{x}\, dx = \arctan\left(\frac{1}{a}\right).$$

Solution. We have for every $x \in \mathbb{R}$:

$$\frac{\sin x}{x} = \sum_{n=0}^{\infty}(-1)^n \frac{x^{2n}}{(2n+1)!}, \quad \text{so that} \quad \int_0^{\infty} e^{-ax}\frac{\sin x}{x}\,dx = \int_0^{\infty} e^{-ax}\left(\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n}}{(2n+1)!}\right)dx;$$

we want to prove that if $a > 1$ then

$$\int_0^{\infty} e^{-ax}\left(\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n}}{(2n+1)!}\right)dx = \sum_{n=0}^{\infty}(-1)^n\int_0^{\infty} e^{-ax}\frac{x^{2n}}{(2n+1)!}\,dx.$$

A sufficient condition for this is the convergence of the series of the L^1norms on $[1,\infty[$ (theorem on normal convergence of a series in L^1), that is the series

$$\sum_{n=0}^{\infty}\int_0^{\infty} e^{-ax}\frac{x^{2n}}{(2n+1)!}\,dx$$

(this says that the sum of the series of absolute values of the functions is in $L^1([0,\infty[)$, and this sum dominates the partial sums of the original series). The integrals are easily computed:

$$\int_0^{\infty} e^{-ax}\frac{x^{2n}}{(2n+1)!}\,dx = \frac{1}{(2n+1)!}\int_0^{\infty} e^{-t}\,\frac{t^{2n}}{a^{2n}}\,\frac{dt}{a} = \frac{1}{a^{2n+1}(2n+1)!}\int_0^{\infty} t^{(2n+1)-1}\,e^{-t}\,dt = \frac{\Gamma(2n+1)}{a^{2n+1}(2n+1)!} = \frac{1}{a^{2n+1}(2n+1)},$$

and clearly the series $\sum_{n=0}^{\infty} 1/(a^{2n+1}(2n+1))$ converges iff $a > 1$. Then, if $a > 1$ we also have, by the theorem on normal convergence

$$\int_0^{\infty} e^{-ax}\frac{\sin x}{x}\,dx = \sum_{n=0}^{\infty}(-1)^n\int_0^{\infty} e^{-ax}\frac{x^{2n}}{(2n+1)!}\,dx = \sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}\,\frac{1}{a^{2n+1}} = \arctan\left(\frac{1}{a}\right),$$

recalling the ciclometric series $\arctan x = \sum_{n=0}^{\infty}(-1)^n x^{2n+1}/(2n+1)$, for $|x| \le 1$.

Remark. For $0 < a < 1$ the series of integrals does not converge. However we have

$$\varphi(a) = \int_0^{\infty} e^{-ax}(\sin x/x)\,dx = \arctan(1/a) \quad \text{for every } a > 0.$$

In fact, the integral converges for $a > 0$, and the theorem on differentiation of parameter depending integrals says that

$$\varphi'(a) = -\int_0^{\infty} e^{-ax}\,\sin x\,dx = -\left[\frac{e^{-ax}}{1+a^2}(-a\,\sin x - \cos x)\right]_{x=0}^{x=\infty} = \frac{-1}{1+a^2};$$

and since the derivative of $\arctan(1/a)$ is $-1/(1+a^2)$ for every $a \neq 0$, and

$$\lim_{a\to\infty}\varphi(a) = \lim_{a\to\infty}\arctan(1/a) = 0,$$

we conclude.

$\square$

4.4.1. *Infinite integrals.* If $(X, \mathcal{M}, \mu)$ is a measure space and $f : X \to \tilde{\mathbb{R}}$ is measurable, the formula

$$\int_X f = \int_X f^+ - \int_X f^-$$

is meaningless only when $\int_X f^+ = \int_X f^- = \infty$. If at least one of the integrals $\int_X f^+$ and $\int_X f^-$ is finite, we say that f *is integrable in the extended sense*, and we define its integral, which may then be ∞ or $-\infty$, by this formula. If f is integrable in the extended sense, then so is $f\,\chi_E$ for every $E \in \mathcal{M}$, and the indefinite integral $\nu = \nu_f : \mathcal{M} \to [-\infty,\infty]$ defined by $\nu(E) = \int_E f\ \left(:= \int_X f\,\chi_E\right)$ is a countably additive function: the easy proof is left to the reader (notice that actually $\nu(E) = \nu_{f^+}(E) - \nu_{f^-}(E)$ is the difference of two positive measures, of which one at most can be infinite valued).

We note the following facts:

. *Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $f, g : X \to \mathbb{R}$ be real valued measurable functions. Then:*

(i) *If $f \le g$ and f and g are integrable in the extended sense, then $\int_X f \le \int_X g$.*

(ii) *If $f \le g$, f is integrable in the extended sense, and $\int_X f > -\infty$, then g is integrable in the extended sense (similarly, if g is integrable in the extended sense and $\int_X g < \infty$, then f is integrable in the extended sense).*

(iii) *If f and g are integrable in the extended sense, and $\int_X f + \int_X g$ is defined, then $f+g$ is integrable in the extended sense, and $\int_X (f+g) = \int_X f + \int_X g$.*

Proof. Exercise. □

Solution. (of the previous exercise) (i) As seen in 1.2.1, $f \le g$ is equivalent to $f^+ \le g^+$ and $f^- \ge g^- \iff -f^- \le -g^-$; summing the inequalities obtained by integrating we have

$$\int_X f = \int_X f^+ - \int_X f^- \le \int_X g^+ - \int_X g^- = \int_X g.$$

(ii) As above we get $f^+ \le g^+$ and $f^- \ge g^-$; if $\int_X f$ exists and is not $-\infty$, we have that $\int_X f^-$ is finite, and then also $\int_X g^-$ is finite.

(iii) If both integrals are finite the proof is that given for $L^1_\mu(X, \mathbb{R})$. We can then reduce to the case in which one, or both, integrals are infinite; exchanging f and g we may assume that $\int_X f$ is infinite; by changing signs we may also assume $\int_X f = \infty$. Then $\int_X f^-$ is finite, and $\int_X g$ cannot be $-\infty$, so that $\int_X g^-$ is also finite. Recall that, by 1.2.1, $(f+g)^+ \le f^+ + g^+$ and $(f+g)^- \le f^- + g^-$. In our case we then get $\int_X (f+g)^- < \infty$, so that $\int_X (f+g)$ is defined and is not $-\infty$. Moreover $\int_X (f+g) = \infty$, since otherwise $f^+ + g^+ = (f+g) + (f^- + g^-)$ has also a finite integral, but $\int_X f^+ = \infty$ by assumption. Then $\int_X (f+g) = \infty$. □

Sometimes we say integrable without adding "in the extended sense" unless for some reason this has to be emphasized. Of course, a function integrable in the extended sense with a finite integral differs from a function in $L^1_\mu(X, \mathbb{R})$ only on a null set, and may be considered belonging to $L^1(\mu)$.

Exercise 4.4.13. Let X be a set, and let $f : X \to \mathbb{R}$ be a real valued function. The set $\Phi(X)$ of all finite subsets of X, ordered by inclusion, is a directed set and we can define a net $\sum f : \Phi(X) \to \mathbb{R}$ by the formula $A \mapsto \sum_A f \left(:= \sum_{x \in A} f(x)\right)$. Prove that this net has a limit in $\tilde{\mathbb{R}}$ if and only if f is integrable in the extended sense with respect to the counting measure $\varkappa$ on X, and that

$$\lim_{A \in \Phi(X)} \sum_A f = \sum_X f \left(:= \int_X f \, d\varkappa \right).$$

Solution. For positive functions the assertion is trivial: the net $A \mapsto \sum_A f$ is increasing, it always has its supremum $\sup\{\sum_A f : A \in \Phi(X)\}$, clearly coinciding with $\int_X f \, d\varkappa$, as limit. Writing $f = f^+ - f^-$ we have $\sum_A f = \sum_A f^+ - \sum_A f^-$ and continuity of addition in $]-\infty, \infty]$ or in $[-\infty, \infty[$ ensures that if f is integrable then the net has the integral as its limit. It only remains to prove that if $\sum_X f^+ = \sum_X f^- = \infty$ then the net has no limit in $\tilde{\mathbb{R}}$. Let $P = \{f > 0\}$ and $Q = \{f < 0\}$. Given any $A \in \Phi(X)$, since $\sum_{P \setminus A} f = \infty$ we can add to A a finite set $B \subseteq P \setminus A$ such that $\sum_{A \cup B} f$ is large as we please, and since $\sum_{Q \setminus A} f = -\infty$ we can also add to A a finite subset C of Q to make $\sum_{A \cup C} f$ smaller than any negative number we want. □

4.4.2. *Monotone convergence theorem, extended version.*

. *Let $(X, \mathcal{M}, \mu)$ be a measure space, and let f_n be a monotone sequence of measurable extended real valued functions, with pointwise limit f; assume that f_0 is integrable (in the extended sense). If the sequence is increasing (resp: decreasing) and $\int_X f_0 > -\infty$ (resp: $\int_X f_0 < \infty$) then f and all f_n are integrable and*

$$\lim_{n \to \infty} \int_X f_n = \int_X f.$$

Proof. Assume that f_n is increasing. By (ii) of 4.4.1 we have that all f_n are integrable and the integrals are an increasing sequence in $\tilde{\mathbb{R}}$; moreover f_n^+ is increasing and f_n^- is a decreasing sequence of positive functions with a finite integral. By the monotone convergence theorem we have

$$\lim_{n \to \infty} \int_X f_n^+ = \int_X f^+,$$

and by dominated convergence we have also

$$\lim_{n \to \infty} \int_X f_n^- = \int_X f^- (< \infty);$$

the conclusion is now immediate. For the decreasing case, just change signs. □

We also easily get as a corollary

. Fatou's lemma, extended version *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of extended real valued measurable functions on the measure space $(X, \mathcal{M}, \mu)$.*

(i) *Assume that there is a function $u : X \to \tilde{\mathbb{R}}$ integrable, with $\int_X u > -\infty$, such that $u \le f_n$ a.e., for every $n \in \mathbb{N}$. Then every f_n is integrable and*

$$\int_X \liminf_{n\to\infty} f_n \le \liminf_{n\to\infty} \int_X f_n.$$

(ii) *Assume that there is a function $v : X \to \tilde{\mathbb{R}}$ integrable, with $\int_X v < \infty$, such that $v \ge f_n$ a.e., for every $n \in \mathbb{N}$. Then every f_n is integrable and*

$$\limsup_{n\to\infty} \int_X f_n \le \int_X \limsup_{n\to\infty} f_n.$$

Proof. In both cases integrability of the f_n's follows from (ii) of 4.4.1. The proof of (i) is then identical to the proof of the original Fatou's lemma, by using the above extended version of monotone convergence, and is left to the reader. Proof of (ii) may follow similar lines, or is simply deduced from (i) by changing the sign. □

4.4.3. *Some solutions.*

Solution. (*of Exercise 4.1.5*) For simplicity, put $\alpha(t) = -\rho_f(t) = -\mu\{(f > t\})$. If $s < t$ we have $]s, t] =]s, \infty[-]t, \infty[$ so that

$$\mu(f^{\leftarrow}(]s, t]) = \mu(f^{\leftarrow}(]s, \infty[\setminus f^{\leftarrow}(]t, \infty[)) = \mu(f^{\leftarrow}(]s, \infty[)) - \mu(f^{\leftarrow}(]t, \infty[) = \rho_f(s) - \rho_f(t) = \alpha(t) - \alpha(s)$$

(finiteness of ρ_f has been used in an essential way for subtractivity). The image measure $\mu f^{\leftarrow}$ and the Radon–Stieltjes measure λ_α of α then coincide and are finite on the right half–open intervals of $]0, \infty[$ with lower bound strictly positive; then they coincide on the Borel subsets of $]0, \infty[$, by $\sigma-$finiteness and 2.5.5. For every positive measurable $g :]0, \infty[\to [0, \infty]$ we have (see 4.1.12)

$$\int_{[0,\infty[} g(t)\, d\mu f^{\leftarrow}(t) = \int_X g(f(x))\, d\mu(x),$$

and we get the result with $g(t) = t$. □

Solution. (*of Exercise 4.1.6*) (i) We clearly have $f \le g$ if $g := \sum_{n=1}^{\infty} n\, \chi_{f^{\leftarrow}(]n-1,n])}$: given $x \in X$, if $f(x) > 0$ then $x \in f^{\leftarrow}(]n-1, n])$ for a unique $n \in \mathbb{N}^{>}$, so that $f(x) \le n = g(x)$. Since $\int_X g = \sum_{n=1}^{\infty} n\, \mu(f^{\leftarrow}(]n-1, n])) < \infty$, we also have $\int_X f < \infty$.

(ii) If $f(x) = 1/(1 + x^2)$ it is well–known that $\int_{\mathbb{R}} f < \infty$; but $g = \chi_{\mathbb{R}}$ has infinite integral.

(iii) Consider $f(x) = 1/(1 + |x|)$; then $\int_X f = \infty$, but $h := \sum_{n=1}^{\infty} (n-1)\, \chi_{f^{\leftarrow}(]n-1,n])} = 0$.

(iv) We have, with h as in (iii) and g as in (i) and (ii):

$$g = \sum_{n=1}^{\infty} n\, \chi_{f^{\leftarrow}(]n-1,n])} = \sum_{n=1}^{\infty} ((n-1) + 1)\, \chi_{f^{\leftarrow}(]n-1,n])} =$$
$$\sum_{n=1}^{\infty} (n-1)\, \chi_{f^{\leftarrow}(]n-1,n])} + \sum_{n=1}^{\infty} \chi_{f^{\leftarrow}(]n-1,n])} = h + \chi_{\{f>0\}};$$

since X has finite measure we have that $\chi_{\{f>0\}}$ has finite integral $(\mu(\chi_{\{f>0\}}))$, hence $\int_X g$ is finite iff $\int_X h$ is finite; and f is sandwiched between h and g, $h \le f \le g$.

(v) Put $\varphi = \sum_{n\in\mathbb{Z}} |n-1|\, \chi_{]n-1,n]}$. If $x \in f^{\leftarrow}(]n-1, n])$ we have $n - 1 < f(x) \le n$, and if $n \le 0$ then $|f(x)| < |n-1|$ and $f(x) \le 0$, i.e. $f(x) = f^{-}(x)$, while if $n > 0$ then $|f(x)| \le \varphi(x) + 1$, and $f(x) \ge 0$ so that $f(x) = f^{+}(x)$. We then have $|f(x)| \le \varphi(x) + \chi_{\{f>0\}}(x)$, so that if φ is summable f is also summable. And we also have $0 \le \varphi - \chi_{\{f<0\}} \le |f|$, so that summability of f implies summability of φ, always taking account of the fact that $\chi_{\{f>0\}}$ and $\chi_{\{f<0\}}$ are summable because $\mu(X) < \infty$. □

Solution. (*of exercise 4.1.7*). Clearly $\rho_{\sin}(t) = \infty$ for $t < 1$, $\rho_{\sin}(t) = 0$ for $t \ge 1$. Clearly $\rho_f(t) = \infty$ for every $t \in \mathbb{R}$ if $f(x) = e^x$. If $f(x) = e^{-|x|}$ we have clearly $\rho_f(t) = \infty$ if $t \le 0$; if $t \ge 1$ then $\rho_f(t) = 0$ (1 is the absolute maximum of f), and if $0 < t < 1$ then

$$\{x \in \mathbb{R} :\, e^{-|x|} > t\} = \{x \in \mathbb{R} :\, -|x| > \log t\} = \{x \in \mathbb{R} : |x| \le \log(1/t)\}, \quad \text{so that } \rho_f(t) = 2\log(1/t);$$

hence $\rho_f(t) = 2(\log(1/t))^+$ for $t > 0$. The integral is easily computed:

$$\int_0^\infty \rho_f(t)\,dt = \int_0^1 2\log(1/t)\,dt = [-2t\log t]_0^1 + \int_0^1 2\,dt = 2.$$

Now $f(x) = |x|^{-1/2}\,\chi_{]0,\infty[}$; clearly $\rho_f(t) = \infty$ it $t \le 0$. If $t > 0$ then

$$\{x \in \mathbb{R} : f(x) > t\} = \{x > 0 :\ x^{-1/2} > t\} = \{x < 1/t^2\} =]0, 1/t^2[.$$

Then $\rho_f(t) = 1/t^2$ for $t > 0$, and hence $\int_0^\infty \rho_f(t)\,dt = \infty$.

$f(x) = 1/(1+x^2)$; $\rho_f = \infty$ for $t \le 0$, and if $t > 0$ we have $\{f > t\} = \{x \in \mathbb{R} :\ 1 + x^2 > 1/t\}$; this set is empty for $t \ge 1$ (clear: 1 is the absolute maximum of f), and is $]-\sqrt{1/t-1}, \sqrt{1/t-1}[$ if $0 < t < 1$. Then $\rho_f(t) = 2\sqrt{1/t-1}$ if $t \in]0,1[$ and $\rho_f(t) = 0$ if $t \ge 1$. Integral:

$$\int_0^\infty \rho_f(t)\,dt = \int_0^1 2\sqrt{\frac{1}{t} - 1}\,dt = 4\int_0^1 \sqrt{1-t}\,\frac{dt}{2\sqrt{t}} = 4\int_0^1 \sqrt{1-\theta^2}\,d\theta = \pi.$$

For the last: $\rho_f(t) = \lambda_n(B) = 2^n$ if $t \le 1$. It $t > 1$ then

$$\{x \in B :\ \|x\|_\infty^{-\alpha} > t\} = \{x \in B :\ \|x\|_\infty < t^{-1/\alpha}\} \quad \text{so that} \quad \rho_f(t) = 2^n\,t^{-n/\alpha}\ (t > 1).$$

We then get

$$\int_0^\infty \rho_f(t)\,dt = \int_0^1 2^n\,dt + 2^n\int_1^\infty t^{-n/\alpha}\,dt;$$

we know that $\int_1^\infty dt/t^\beta = \infty$ if $\beta \le 1$ and $1/(\beta-1)$ if $\beta > 1$; then:

$$\int_B f_\alpha\,d\lambda_n = \int_0^\infty \rho_{f_\alpha}(t)\,dt = \begin{cases} \infty & \text{if} \quad \alpha \ge n \\ 2^n\left(1 + \dfrac{\alpha}{n-\alpha}\right) & \text{if} \quad \alpha < n \end{cases}.$$

□

Solution. (*of Exercise 4.4.2*) Clearly $g\,f$ is measurable and $|g\,f| = |g|\,|f| \le \|g\|_\infty\,|f|$, and moreover $\|g\|_\infty\,|f| \in L^1(\mu)$ since $f \in L^1(\mu)$; thus $g\,f \in L^1(\mu)$.

Assume now that g is measurable but not in $L^\infty(\mu)$. Then infinitely many of the sets $\{n < |g| \le n+1\}$ have nonzero measure; there is then an increasing sequence $n_k \to \infty$ of integers such that $\mu(\{n_k < |g| \le n_k + 1\}) > 0$, and we may also assume that $n_k \ge 2^k$ (taking a subsequence). By semifiniteness we have a measurable $E(k) \subseteq \{n_k < |g| \le n_k + 1\}$ such that $0 < \mu(E(k)) < \infty$. Then:

$$f = \sum_{k=0}^\infty \frac{1}{n_k\,\mu(E(k))}\,\chi_{E(k)} \in L^1(\mu), \quad \text{in fact} \quad \int_X f = \sum_{k=0}^\infty \frac{\mu(E(k))}{n_k\,\mu(E(k))} = \sum_{k=0}^\infty \frac{1}{n_k} \le \sum_{k=0}^\infty \frac{1}{2^k} < \infty.$$

But $g\,f \notin L^1(\mu)$, in fact:

$$\int_X |g\,f| = \sum_{k=0}^\infty \frac{1}{n_k\mu(E(k)}\int_{E(k)} |g| \ge \sum_{k=0}^\infty \frac{1}{n_k\mu(E(k)}\int_{E(k)} n_k = \sum_{k=0}^\infty 1 = \infty.$$

□

Solution. (*of Exercise 4.4.3*) (i) If $f : X \to \mathbb{K}$ is a simple function with range $f(X) = \{y_1, \dots, y_p\}$ then $A_j = f^{\leftarrow}(y_j)$ is a partition of X into measurable sets; exactly one of these sets is co–countable, the others all countable; so the simple functions are those that are constant on a co-countable set. Then the same is true for measurable functions: if f_n is a sequence of simple functions converging pointwise to f, and each f_n is constantly a_n on the co–countable set A_n, then f is constantly $a = \lim_{n\to\infty} a_n$ on the set $A = \bigcap_{n=0}^\infty A_n$, a co–countable set. Then every measurable function is constant outside some countable subset of X; and it is clear that every such function is measurable.

(ii) Notice that the null functions are exactly those that are zero outside a countable subset of X. Every positive measurable function has integral 0 (when it is zero outside a countable set, hence a.e. 0) or $+\infty$ (when it is constantly $a > 0$ outside a countable set).

(iii) Modulo null functions all measurable functions are constant; so $\mathcal{L}^\infty_\mu$ coincides with the set of all measurable functions, and $L^\infty(\mu) = \mathcal{L}^\infty_\mu/\mathcal{N}_\mu$ can be identified with the space of constant functions. □

Solution. (*of exercise 4.4.5*) Given a sequence $f_n \in L^1_\mu(X,F)$ converging to $f \in L^1(\mu)$ we prove that $f(x) \in F$ for a.e. $x \in X$. In fact we have a subsequence $f_{\nu(k)}$ of f_n that converges to f also a.e.; the set $N(k) = \{x \in X : f_{\nu(k}(x) \notin F\}$ has measure 0, hence also $N = \bigcup_{k\in\mathbb{N}} N(k)$ has measure 0. For $x \in X \setminus N$ we have $f_{\nu(k)}(x) \in F$ for every k, so that $f(x) = \lim_{k\to\infty} f_{\nu(k)}(x) \in F$, since F is closed.

Then $L^1_\mu(X,F)$ is a complete metric subspace of $L^1(\mu)$, for every closed $F \subseteq \mathbb{K}$, in particular $L^1_\mu(X,\{0,1\})$ is complete. And clearly this metric space is isometric to $(\mathcal{F}(\mu),\rho)$ the isometry being $A \mapsto \chi_A$; in fact, for $A,B \in \mathcal{F}(\mu)$:

$$\rho(A,B) = \mu(A \vartriangle B) = \int_{A\Delta B} d\mu = \int_X \chi_{A\Delta B} = \int_X |\chi_A - \chi_B| = \|\chi_A - \chi_B\|_1,$$

recalling that $|\chi_A - \chi_B| = \chi_{A\Delta B}$. □

Solution. (*of Exercise 4.4.9*) Since $|g_{n,\alpha}(x)| \le n^{1-\alpha}\,|f(x)|$ and $f \in L^1(\mu)$ we have $g_{n,\alpha} \in L^1(\mu)$. We restrict the integral to the set $C = \{f > 0\}$, and we write

$$\int_X g_{n,\alpha} = n^{1-\alpha} \int_C \frac{\sin(f(x)/n^\alpha)}{f(x)/n^\alpha}\, f(x)\, d\mu(x).$$

Now we have:

$$\left|\frac{\sin(f(x)/n^\alpha)}{f(x)/n^\alpha}\, f(x)\right| \le f(x),$$

so that dominated convergence says that

$$\lim_{n\to\infty} \int_C \frac{\sin(f(x)/n^\alpha)}{f(x)/n^\alpha}\, f(x)\, d\mu(x) = \int_C f,$$

so that, if $\alpha < 1$

$$\lim_{n\to\infty} \int_X g_{n,\alpha} = \lim_{n\to\infty} n^{1-\alpha} \int_C \frac{\sin(f(x)/n^\alpha)}{f(x)/n^\alpha}\, f(x)\, d\mu(x) = \infty.$$

while for $\alpha = 1$ this limit is $\int_X f$, and for $\alpha > 1$ the limit is 0. □

Solution. (*of Exercise 4.4.10*) (i) If $s < t$ then $]-\infty,s] \subseteq]-\infty,t]$, then $f^{\leftarrow}(]-\infty,s] \subseteq f^{\leftarrow}(]-\infty,t])$, and by monotonicity of the measure we have $F_f(s) = \mu(f^{\leftarrow}(]-\infty,s]) \le \mu(f^{\leftarrow}(]-\infty,t])) = F_f(t)$. Right–continuity is continuity from above of measures on set of finite measure: if $t_n \downarrow t$ then $]-\infty,t_n] \downarrow]-\infty,t]$ so that $F_f(t_n) = \mu(f^{\leftarrow}(]-\infty,t_n])) \downarrow \mu(f^{\leftarrow}(]-\infty,t])) = F_f(t)$. Since $f^{\leftarrow}(]-\infty,t]) = X \setminus f^{\leftarrow}(]t,\infty[)$ and the measure is finite, by subtractivity we get

$$\mu(f^{\leftarrow}(]-\infty,t])) = \mu(X \setminus f^{\leftarrow}(]t,\infty[)) = \mu(X) - \mu(f^{\leftarrow}(]t,\infty[)) = \mu(X) - \rho_f(t).$$

(ii) If $t < 0$ then $\{x \in X : f(x) \le t\} = \{x \in X : -f^-(x) \le t\} = \{x \in X : f^-(x) \ge -t\}$ (if $f(x) \le t < 0$ then $f(x) = -f^-(x)$). Now, if $a > 0$ and $g : X \to [0,\infty[$ is positive measurable, then $\mu(\{g \ge a\}) = \lim_{s\to a^-} \rho_g(s) (:= \rho_g(a^-)$, as is easy to see, if the measures are finite (if $s_n \uparrow a$, then $\{g > s_n\} \downarrow \{g \ge a\}$). Then, if $t < 0$, from (i) we have $F_f(t) = \rho_{f^-}(-t^-)$.

If $t \ge 0$ then easily $\{f > t\} = \{f^+ > t\}$ so that as seen in (i) we get $F_f(t) = \mu(X) - \rho_{f^+}(t)$.

(iii) We have that $f \in L^1(\mu)$ if and only if $f^\pm \in L^1(\mu)$; by 4.1.14 this happens if $\int_0^\infty \rho_{f^\pm}(t)\,dt < \infty$; and

$$\int_0^\infty \rho_{f^+}(t)\,dt = \int_0^\infty (\mu(X) - F_f(t))\,dt < \infty,$$
$$\int_0^\infty \rho_{f^-}(s)\,ds = \int_0^\infty \rho_{f^-}(s^-)\,ds = \int_{-\infty}^0 \rho_{f^-}(-t^-)\,dt = \int_{-\infty}^0 F_f(t)\,dt$$

(If $\rho : [0,\infty] \to [0,\infty]$ is decreasing, then $\int_0^\infty \rho(t^-)\,dt = \int_0^\infty \rho(t^+)\,dt$: left and right–continuous modifications of ρ differ on an at most countable set, which has Lebesgue measure zero).

(it is clear that $\int_0^\infty (\mu(X) - F_f(t))\,dt$ is area A, while $\int_{-\infty}^0 F_f(t)\,dt$ is area B).

(iv) Is easy, using what done in 4.1.5. □

Solution. (*of Exercise 4.4.11*) Recall that semifinite measure means that every set of nonzero measure contains a set of finite nonzero measure. If for some disk as in the hint we have $\mu(f^{\leftarrow}(B(c,r[)) > 0$ then we can find a measurable $E \subseteq f^{\leftarrow}(B(c,r[)$ with $0 < \mu(E) < \infty$. Then, since $|f(x) - c| < r$ for every $x \in E$ we have

$$\left|⨍_E f - c\right| = \left|⨍_E f - ⨍_E c\right| = \left|⨍_E (f - c)\right| \le ⨍_E |f - c| < ⨍_E r = r,$$

so that $⨍_E f \notin C$, contrary to the assumption.

Then $f^{\leftarrow}(B(c,r[)$ has measure zero for every disk disjoint from C; and since $\mathbb{K} \setminus C$ is a countable union of such disks we have that $\mu(f^{\leftarrow}(\mathbb{K} \setminus C)) = 0$. □

4.5. Continuity and differentiability of parameter depending integrals.

Theorem. *Let D be a metrizable space; let $(T, \mathcal{M}, \mu)$ be a measure space, and let $f : D \times T \to \mathbb{K}$ be a function. Assume that for every $x \in D$ the function $T \mapsto f(x,t)$ is in $L^1_\mu(T, \mathbb{K})$, so that the formula:*

$$F(x) = \int_T f(x,t)\, d\mu(t)$$

defines a function $F : D \to \mathbb{K}$. Let $x \in D$ be given.

(i) *Assume that there are a nbhd. U of x in D, and a function $\gamma_x \in L^1_\mu(T)$ such that for every $y \in U$ the inequality $|f(y,t)| \leq \gamma_x(t)$ holds for $\mu-$a.e. $t \in T$. Assume also that for $\mu-$a.e. $t \in T$ the function $y \mapsto f(y,t)$ is continuous at x. Then F is continuous at x.*

(ii) *Assume now that D is an open subset of the normed $\mathbb{K}-$space Y, that u is a vector of Y, that for every $y \in D$ the derivative $\partial_u f(y,t)$ exists for almost every $t \in T$, and that moreover there are a nbhd. U of x in D, and a function $\gamma_x \in L^1_\mu(T)$ such that for every $y \in U$ the inequality $|\partial_u f(y,t)| \leq \gamma_x(t)$ holds for a.e. $t \in T$. Then $\partial_u F(x)$ exists and*

$$\partial_u F(x) = \int_T \partial_u f(x,t)\, d\mu(t).$$

The proof is a simple application of dominated convergence:

Proof. (i) Given $x \in D$ let $(x_n)_n$ be a sequence in D converging to x in D. We have to prove that $\lim_{n\to\infty} F(x_n) = F(x)$, that is, defining f_n, $g : T \to \mathbb{K}$ by $f_n(t) = f(x_n, t)$, and $g(t) = f(x,t)$ we claim that

$$(*) \qquad \lim_{n\to\infty} \int_T f_n(t)\, d\mu(t) = \int_T g(t)\, d\mu(t).$$

In fact, by the hypotheses $f_n(t) \to g(t)$ for $\mu-$a.e. $t \in T$ (continuity of f at x in the first variable); and given a nbhd U of x and $\gamma_x \in L^1_\mu(T)$ as in the hypotheses, we have $x_n \in U$ eventually, say for $n \geq N$, so that $|f_n(t)| \leq \gamma_x$ for $\mu-$a.e. $t \in T$, as $n \geq N$; by dominated convergence (*) holds.

(ii) Given $x \in D$, there is $\delta > 0$ such that $x + \lambda\, u \in D$ for every $\lambda \in \mathbb{K}$ with $\lambda \in \delta\, B = \{\xi \in \mathbb{K} : |\xi| \leq \delta\}$; apply (i) to the function $g : \delta\, B \times T \to \mathbb{K}$ given by:

$$g(0,t) = \partial_u f(x,t); \quad g(\lambda, t) = \frac{f(x + \lambda\, u, t) - f(x,t)}{\lambda} \qquad \lambda \in \delta\, B \setminus \{0\},\, t \in T;$$

for $|\lambda|$ small enough (such that $x + \lambda\, u \in U$) we have, by the mean value theorem:

$$|g(\lambda, t)| = \left| \frac{f(x + \lambda\, u, t) - f(x,t)}{\lambda} \right| \leq \sup\{|\partial_u f(x + \xi\, u, t)| : |\xi| \leq |\lambda|\} \leq \gamma_x(t).$$

□

Exercise 4.5.1. (see also 4.4.12) Prove that the formula:

$$(*) \qquad \varphi(x) = \int_0^\infty e^{-xt} \frac{\sin t}{t}\, dt$$

defines a function $\varphi \in C^1(]0, \infty[, \mathbb{R})$. Give an esplicit formula for $\varphi'(x)$, not containing integrals, and deduce from it an analogous expression for $\varphi(x)$.

Solution. The derivative with respect to x of the integrand is $-e^{-xt} \sin t$. Given $x > 0$, let $a = x/2$ (or simply pick any a with $0 < a < x$), and let $U = [a, \infty[$. For $y \in U$ we have

$$|-e^{-yt} \sin t| = e^{-yt}\, |\sin t| \leq e^{-yt} \leq e^{-at};$$

of course $t \mapsto e^{-at}$ belongs to $L^1([0, \infty[)$, since $a > 0$. Then $\varphi \in C^1(]0, \infty[$, and (see formula for the primitive of $e^{-xt} \sin t$):

$$\varphi'(x) = \int_0^\infty (-e^{-xt} \sin t)\, dt = \left[\frac{e^{-xt}}{1 + x^2} (\sin t + \cos t) \right]_{t=0}^{t=\infty} = \frac{-1}{1 + x^2}.$$

Then we get

$$\varphi(x) = \operatorname{arccotan} x + k \qquad (x > 0);$$

but one easily sees that $\lim_{x\to\infty} \varphi(x) = 0$ (e.g., by dominated convergence; or simply because $|\varphi(x)| \leq \int_0^\infty e^{-xt}\, dt = 1/x$), so that

$$\varphi(x) = \operatorname{arccotan} x \qquad (x > 0).$$

□

4.5.1. *Continuity and differentiability of sums of power series.* The following is well known; but we want to get it as a corollary of the previous theorem. If $a = (a_n)_{n\in\mathbb{N}}$ is a sequence of complex numbers, and c is a complex numbers, we have the *power series with coefficient sequence $a = (a_n)_{n\in\mathbb{N}}$ and initial point c*,

$$\sum_{n=0}^{\infty} a_n\,(z-c)^n.$$

It is clear that, by a simple translation, we may confine ourselves to the case $c = 0$. We know (see e.g. 1.8.7) that if $R = R_a = 1/\limsup_{n\to\infty} |a_n|^{1/n}$ is the radius of convergence of the series $\sum_{n=0}^{\infty} a_n\,z^n$, then the convergence set C_a of the series is sandwiched like that:

$$\{z \in \mathbb{C} :\ 0 \le |z| < R\} \subseteq C_a \subseteq \{z \in \mathbb{C} :\ 0 \le |z| \le R\}.$$

If $R > 0$ then C_a has a non–empty interior, the open disc of convergence, all of $\mathbb{C}$ when $R = \infty$, otherwise the interior of C_a is $R\,B = \{0 < |z| < R\}$. Recall that the derived power series $\sum_{n=1}^{\infty} n\,a_n\,z^{n-1}$ has the same radius of convergence R_a. On C_a we may define the function $f_a(z) = \sum_{n=0}^{\infty} a_n\,z^n$, so that on the open disc of convergence the sum $g_a(z) = \sum_{n=1}^{\infty} n\,a_n\,z^{n-1}$ is also defined.

Let us prove

. Continuity and complex differentiability of power series sums *In the interior of C_a the function f_a is continuous, and differentiable in the complex sense, with derivative $f_a'(z) = g_a(z)$.*

Proof. The space $(T, \mathcal{M}, \mu)$ is $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \varkappa)$ with $\varkappa$ the counting measure. That$(x =)z \in \mathrm{int}(C_a)$ means $|z| < R$; pick r with $|z| < r < R$; the nbhd U of z is $r\,B = \{w \in \mathbb{C} :\ |w| < r\}$. The function $(y, t) \mapsto f(y, t)$ is here $(w, n) \mapsto a_n\,w^n$; the function $\gamma_x = \gamma_z : \mathbb{N} \to [0, \infty[$ is $n \mapsto |a_n|\,r^n$; we have $\gamma_z \in L^1(T) = \ell^1(\mathbb{N})$ because every power series is absolutely convergent at every point in the interior of C_a, in particular at $w = r$, that is $\sum_{n=0}^{\infty} |a_n\,r^n| < \infty$). This proves continuity of f_a. For differentiability we only have to change γ_z, that is now $\gamma_z(n) = n\,|a_n|\,r^{n-1}$.

Remark. We can apply the differentiation theorem as stated also to complex derivatives essentially because a mean value theorem holds for complex derivatives: if $D \subseteq \mathbb{C}$ is an open set containing the segment $[a, b] = \{a + t(b-a) :\ t \in [0, 1]\}$, and $f : D \to \mathbb{C}$ has a complex derivative $f'(z)$ at every $z \in [a, b]$, then $|f(b) - f(a)| \le \|f'\|_{[a,b]}\,|b - a|$, where $\|f'\|_{[a,b]} = \sup\{|f'(z)| :\ z \in [a, b]\}$.

□

4.6. Riemann integral and Lebesgue integral. The Riemann integral is defined with a particular kind of simple functions, the step functions with compact support (recall that the support $\mathrm{Supp}(f)$ of a function $f : X \to \mathbb{K}$, where X is a topological space, is the closure of the cozero–set of f, $\mathrm{Supp}(f) = \mathrm{cl}_X\{f \neq 0\}$). These are the vector subspace $S_c = S_c(\mathbb{R}^n)$ of the space of all functions from $\mathbb{R}^n$ to $\mathbb{K}$ generated by the characteristic functions of bounded intervals of $\mathbb{R}^n$. Every nonzero $\varphi \in S_c$ is representable as $\varphi = \sum_{k=1}^{m} \alpha_k\,\chi_{E(k)}$, where $\alpha_k \in \mathbb{K}$, and $\{E(1), \ldots, E(m)\}$ is a finite set of pairwise disjoint bounded intervals; the integral of φ is, by definition,

$$\int_{\mathbb{R}^n} \varphi := \sum_{k=1}^{m} \alpha_k\,\lambda^n(E(k)).$$

Clearly $S_c(\mathbb{R}^n)$ is contained in $L^1(\lambda_n)$. We confine our attention to real valued functions.

Definition. A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be Riemann integrable (R–integrable for short) if for every $\varepsilon > 0$ there are $u, v \in S_c$ such that $u \le f \le v$, and $\int_{\mathbb{R}^n}(v - u) \le \varepsilon$.

This condition is clearly equivalent to the assertion:

$$\sup\left\{\int_{\mathbb{R}^n} u :\ u \le f,\, u \in S_c\right\} = \inf\left\{\int_{\mathbb{R}^n} v :\ v \in S_c,\, v \ge f\right\};$$

by definition, this separator is the Riemann integral of f. From the definition it immediately follows that an R–integrable function is bounded and has compact support (since there exist $u, v \in S_c$ with $u \le f \le v$ we have that $\mathrm{Supp}(f) \subseteq \mathrm{Supp}(u) \cup \mathrm{Supp}(v)$, and $\min u \le f \le \max v$). And it is also immediate to verify that

Proposition. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. The following are equivalent:*

(i) *f is R–integrable.*
(ii) *For every sequence $\varepsilon_k > 0$ with $\lim_{k\to\infty} \varepsilon_k = 0$ there are sequences u_k, $v_k \in S_c$ such that:*
(1) *$u_k \le f \le v_k$ for every $k \in \mathbb{N}$;*
(2) *u_k is increasing and v_k is decreasing;*
(3) *$\int_{\mathbb{R}^n}(v_k - u_k) \le \varepsilon_k$, for every $k \in \mathbb{N}$.*

Proof. (i) implies (ii) By the definition, we find $u_0, v_0 \in S_c$ such that $u_0 \le f \le v_0$ and $\int_{\mathbb{R}^n}(v_0 - u_0) \le \varepsilon_0$. By induction, assuming that

$$u_0 \le u_1 \le \cdots \le u_{k-1} \le f \le v_{k-1} \le \cdots \le v_1 \le v_0$$

have been defined, we find φ_k, $\psi_k \in S_c$ with $\varphi_k \le f \le \psi_k$ and $\int_{\mathbb{R}^n}(\psi_k - \varphi_k) \le \varepsilon_k$. Setting $u_k = u_{k-1} \vee \varphi_k$ and $v_k = \psi_k \wedge v_{k-1}$ we conclude the induction (note that $v_k - u_k \le \psi_k - \varphi_k$). That (ii) implies (i) is trivial. □

Assuming f Riemann integrable and u_k, v_k as above we now set:

$$u(x) = \lim_{k\to\infty} u_k(x) = \sup\{u_k(x) : k \in \mathbb{N}\}; \qquad v(x) = \lim_{k\to\infty} v_k(x) = \inf\{v_k(x) : k \in \mathbb{N}\};$$

then $u \le f \le v$, the functions u and v are Borel measurable, and all three functions are Riemann integrable with the same integral. By dominated convergence $\int_{\mathbb{R}^n} u_k$ converges to $\int_{\mathbb{R}^n} u$ and $\int_{\mathbb{R}^n} v_k$ converges to $\int_{\mathbb{R}^n} v$ in $L^1(\lambda_n)$. Then $\int_{\mathbb{R}^n}(v - u) = 0$ both in the Lebesgue and the Riemann sense, so that the Lebesgue measure of the set $\{x \in \mathbb{R}^n : v(x) - u(x) > 0\}$ is zero. This implies that f is Lebesgue measurable (it is a.e. equal to the Borel measurable functions u and v), and $\int_{\mathbb{R}^n} f = \int_{\mathbb{R}^n} u = \int_{\mathbb{R}^n} v$, both in the Lebesgue and the Riemann sense. We have proved that if f is R-integrable then it is also in $L^1(\lambda^n)$, and the two integrals coincide.

4.6.1. *Lebesgue integral and generalized Riemann integral.* Assume now that I is an open interval of $\mathbb{R}$, and that $f : I \to \mathbb{K}$ is R-integrable on every compact subinterval of I. Let $[a_k, b_k]$ be an increasing sequence of compact subintervals of I whose union is I (such a sequence always exists, pick a_k, $b_k \in I$ such that $a_k \downarrow \inf I$ and $b_k \uparrow \sup I$). If $\chi_k = \chi_{[a_k,b_k]}$ then $|f|\,\chi_k \uparrow |f|$ on I; by monotone convergence we then have that the Lebesgue integral of $|f|$ on I is

$$\int_I |f| = \lim_{k\to\infty} \int_I |f|\,\chi_k = \lim_{k\to\infty} \int_{a_k}^{b_k} |f(x)|\,dx;$$

but the last limit is also the generalized integral, in the Riemann sense, of $|f|$; we have proved that a locally R-integrable function is Lebesgue summable on the open interval I if and only if it absolutely R–integrable in the generalized sense. Of course, if this is the case, dominated convergence implies also that

$$\int_I f(x)\,d\lambda(x) = \lim_{k\to\infty} \int_I f\,\chi_k\,d\lambda = \lim_{k\to\infty} \int_{a_k}^{b_k} f(x)\,dx,$$

so that, in this case, the Lebesgue integral coincides with the generalized Riemann integral; but without absolute convergence the function is not Lebesgue integrable.

4.6.2. *Theorem of Lebesgue–Vitali.* Step functions are not pliable enough to adapt closely to wildly oscillating functions, and R-integrability is equivalent to continuity almost everywhere, as the following theorem shows.

. Theorem of Lebesgue–Vitali. *A function $f : \mathbb{R}^n \to \mathbb{K}$ is R–integrable if and only if it is bounded with compact support, and its set of points of discontinuity has Lebesgue measure 0.*

Proof. We assume that f is real valued. By definition the lower Riemann integral of f is

$$\underline{I}(f) = \sup\left\{\int_{\mathbb{R}^n} u : u \in S_c,\ u \le f\right\},$$

while the upper Riemann integral of f is

$$\overline{I}(f) = \inf\left\{\int_{\mathbb{R}^n} v : v \in S_c,\ v \ge f\right\},$$

so that f is Riemann integrable if and only if $\underline{I}(f) = \overline{I}(f)$. Recall the lsc and usc approximations f_* and f^* of f discussed in 1.9.3; the functions f_* and f^* are Borel measurable, bounded and with compact support (their support is contained in $\mathrm{Supp}(f)$), so they both belong to $L^1(\lambda_n)$. We shall prove that

$$(*) \qquad \underline{I}(f) = \int_{\mathbb{R}^n} f_*\,d\lambda_n; \qquad \overline{I}(f) = \int_{\mathbb{R}^n} f^*\,d\lambda_n,$$

(both integrals intended as Lebesgue integrals). This concludes the proof because:

$$f \text{ is Riemann integrable } \iff \int_{\mathbb{R}^n} (f^* - f_*)\, d\lambda_n = 0 \iff f^*(x) = f_*(x) \quad \text{for almost every } x \in \mathbb{R}^n$$
$$\iff f \text{ is continuous almost everywhere.}$$

First we see that

. *If $u \in S_c$ and $u \le f$, then $u(x) \le f_*(x)$ for a.e. $x \in \mathbb{R}^n$; and similarly if $v \in S_c$ and $v \ge f$ then $v(x) \ge f^*(x)$ for a.e. $x \in \mathbb{R}^n$; then $\underline{I}(f) \le \int_{\mathbb{R}^n} f_*$ and $\overline{I}(f) \ge \int_{\mathbb{R}^n} f^*$.*

In fact, if $u = \sum_{k=1}^m \alpha_k\, \chi_{E(k)}$ with the $E(k)$ pairwise disjoint we have $u(x) \le f_*(x)$ for every $x \notin \bigcup_{k=1}^m \partial E(k)$: for, if x belongs to the interior of $E(k)$ then we have $u(t) = \alpha_k \le f(t)$ for every $t \in \operatorname{int}(E(k))$, so that $\alpha_k \le \inf f(\operatorname{int}(E(k))) \le f_*(x)$, hence $u(x) = \alpha_k \le f_*(x)$ (by the same argument we also have $u(x) = 0 \le f_*(x)$ for $x \notin \bigcup_{k=1}^m \operatorname{cl}(E(k))$). Since $\bigcup_{k=1}^m \partial E(k)$ has $n-$dimensional measure 0 we conclude. The proof for v is entirely analogous. It remains to prove equality in (*). We choose a compact interval Q containing the support of f in its interior; by *partition of* Q we mean, in this section, a finite disjoint family $\mathcal{P} = \{A(1), \dots, A(m)\}$ of intervals whose union is Q; the *mesh* of the partition is $\operatorname{mesh}(\mathcal{P}) = \max\{\operatorname{diam}(A(j)) : j \in \{1, \dots, m\}\}$. To every partition $\mathcal{P} = \{A(1), \dots, A(m)\}$ of Q there are a lower step function $u_{\mathcal{P}} \le f$ and an upper step function $v_{\mathcal{P}} \ge f$ associated to it, in the following way:

$$u_{\mathcal{P}} = \sum_{j=1}^m \inf f(A(j))\, \chi_{A(j)};$$
$$v_{\mathcal{P}} = \sum_{j=1}^m \sup f(A(j))\, \chi_{A(j)} :$$

in other words $u_{\mathcal{P}}$ is the largest step function smaller than f with $\mathcal{P}$ as an associated partition, while $v_{\mathcal{P}}$ is the smallest step function larger than f with $\mathcal{P}$ as associated partition. Assume now that $\mathcal{P}(k)$ is a sequence of partitions of Q such that $\lim_{k\to\infty} \operatorname{mesh}(\mathcal{P}(k)) = 0$, and for every k set $u_k = u_{\mathcal{P}(k)}$ and $v_k = v_{\mathcal{P}(k)}$. Let $F(k)$ be the union of the boundaries of all intervals in $\mathcal{P}(k)$; then $F(k)$ has $n-$dimensional measure zero, and hence the same holds for $F = \bigcup_{k=0}^\infty F(k)$. We shall prove:

. *For every $x \in \mathbb{R}^n \setminus F$ we have $f_*(x) = \lim_{k\to\infty} u_k(x)$ and $f^*(x) = \lim_{k\to\infty} v_k(x)$.*

Proof. We prove that $f_*(x) = \lim_{k\to\infty} u_k(x)$ if $x \notin F$, analogous proof for v_k and f^*. Given $x \in Q$ and $\varepsilon > 0$, we find an open ball $B(x, \delta[$ centered at x such that $\inf f(B(x, \delta[) \ge f_*(x) - \varepsilon$. For k large enough, say $k \ge k_\varepsilon$ we have $\operatorname{mesh}(\mathcal{P}(k)) \le \delta$, so that every interval of $\mathcal{P}(k)$ containing x is contained in $B(x, \delta[$. Since $x \notin F$, for every k the point x is in the interior of some interval $A \in \mathcal{P}(k)$, and if $A \subseteq B(x, \delta[$ we have $u_k(x) = \inf f(A)$ and $f_*(x) \ge \inf f(A) \ge \inf f(B(x, \delta[) \ge f_*(x) - \varepsilon$. $\square$

Since $\lambda_n(F) = 0$ we have that u_k converges a.e, to f_* and v_k converges a.e to f^*; both sequences are clearly dominated by $\|f\|_\infty\, \chi_Q \in L^1(\lambda_n)$, so that by dominated convergence

$$\int_{\mathbb{R}^n} f_* = \lim_{k\to\infty} \int_{\mathbb{R}^n} u_k, \qquad \int_{\mathbb{R}^n} f^* = \lim_{k\to\infty} \int_{\mathbb{R}^n} v_k.$$

$\square$

5. Product measures

Given two measurable spaces $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ we have defined a product $(X \times Y, \mathcal{M} \otimes \mathcal{N})$ of these spaces (3.0.12). If we have two measure spaces $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ we plan to define a measure on $\mathcal{M} \otimes \mathcal{N}$, the *product measure*, which will be denoted $\mu \times \nu$ or $\mu \otimes \nu$, and to relate integrals with respect to this measure to integrals on the factor spaces.

5.1. Product measure. The set $\mathcal{G} = \{E \times F : E \in \mathcal{M}, F \in \mathcal{N}\}$ of all measurable rectangles is a semialgebra of parts in $X \times Y$ (see 1.4.2); let's call $\mathcal{A}$ the generated algebra, consisting of all finite (disjoint) unions of elements of $\mathcal{G}$; the $\sigma-$algebra generated by $\mathcal{A}$ is of course $\mathcal{M} \otimes \mathcal{N}$ (3.0.12). Guided by our geometrical intuition we define $\mu \otimes \nu(E \times F) := \mu(E)\, \nu(F)$ (as usual the convention $0 \cdot \infty = \infty \cdot 0 = 0$ is used), and we prove (see 5.1.1) that on $\mathcal{G}$ this set function is countably additive: this implies that we can extend it to a premeasure on the generated rectangle algebra $\mathcal{A}$. We can then give the:

Definition. Given two measure spaces $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ their *product measure space* is the triple

$$(X \times Y, \mathcal{M} \otimes \mathcal{N}, \mu \otimes \nu),$$

where $\mu \otimes \nu$ is the restriction to $\mathcal{M} \otimes \mathcal{N}$ of the Carathèodory extension of the premeasure defined on the rectangle algebra by $\mu \otimes \nu(E \times F) = \mu(E)\, \nu(F)$, for every rectangle $E \times F$.

5.1.1. *Product premeasure.* A function $f \in L^+(X \times Y) (= L^+_{\mathcal{M} \otimes \mathcal{N}}(X \times Y))$ will be called a function of class $\mathcal{I}$ if the functions obtained integrating the sections (recall that all sections of a measurable function are measurable, see 3.0.12):

$$\varphi(x) = \int_Y f(x, y)\, d\nu(y); \quad \psi(y) = \int_X f(x, y)\, d\mu(x)$$

belong to $L^+(X)$ and $L^+(Y)$ respectively, and moreover

$$\int_X \varphi(x)\, d\mu(x) = \int_Y \psi(y)\, d\nu(y).$$

In other words, function of class $\mathcal{I}$ are those positive measurable function on the product for which the *iterated integrals* both exist and coincide:

$$\int_X \left(\int_Y f(x, y)\, d\nu(y) \right) d\mu(x) = \int_Y \left(\int_X f(x, y)\, d\mu(x) \right) d\nu(y);$$

this common value (which will be $\int_{X \times Y} f\, d(\mu \otimes \nu)$ in the right cases) is provisionally denoted $I(f)$.

Proposition. *If $f, g \in L^+(X \times Y)$ are of class $\mathcal{I}$ then $f + g$ and λf, for every $\lambda \geq 0$, are of class $\mathcal{I}$ and $I(f + g) = I(f) + I(g)$, $I(\lambda f) = \lambda I(f)$. If f_n is an increasing sequence of functions of class $\mathcal{I}$ and $f_n \uparrow f$, then f is of class $\mathcal{I}$ and $I(f_n) \uparrow I(f)$. And if $f_n \in L^+(X \times Y)$ is a sequence of functions of class $\mathcal{I}$ then $\sum_{n=0}^{\infty} f_n$ is of class $\mathcal{I}$ and:*

$$I\left(\sum_{n=0}^{\infty} f_n \right) = \sum_{n=0}^{\infty} I(f_n). \tag{*}$$

Proof. Left to the reader, we only prove the monotone convergence: for every $x \in X$ the $x-$section $f_n(x, \#)$ is an increasing sequence in $L^+(Y)$, converging to the $x-$section $f(x, \#)$ of f, so that $\varphi_n(x) = \int_Y f_n(x, y)\, d\nu(y) \uparrow \int_Y f(x, y)\, d\nu(y) = \varphi(x)$; then $\varphi \in L^+(X)$, and since $\varphi_n \uparrow \varphi$ we have also

$$I(f_n) = \int_X \varphi_n(x)\, d\mu(x) \uparrow \int_X \left(\int_Y f(x, y)\, d\nu(y) \right) d\mu(x);$$

exchanging the variables we get

$$I(f_n) = \int_Y \psi_n(y)\, d\nu(y) \uparrow \int_Y \left(\int_X f(x, y)\, d\mu(x) \right) d\nu(y),$$

and the iterated integrals of f must then coincide, being limits of the same sequence $I(f_n)$. □

Notice now that if $f \in L^+(X)$ and $g \in L^+(Y)$ then $f \otimes g \in L^+(X \times Y)$ is of class $\mathcal{I}$ and $I(f \otimes g) = \int_X f \int_Y g$ (trivial):

$$\int_X \left(\int_Y f(x)\, g(y)\, d\nu(y) \right) d\mu(x) = \int_X f(x) \left(\int_Y g(y)\, d\nu(y) \right) d\mu(x) = \left(\int_Y g(y)\, d\nu(y) \right) \int_X f(x)\, d\mu(x),$$

and exchanging x and y we get the same result. If $E \in \mathcal{M}$ and $F \in \mathcal{N}$ then $\chi_{E\times F} = \chi_E \otimes \chi_F$, so that all characteristic functions of measurable rectangles, and hence also all positive $\mathcal{A}-$simple functions are of class $\mathcal{I}$; formula (*) of the preceding proposition says that the set function

$$\mu \otimes \nu(E \times F) = I(\chi_E \otimes \chi_F)$$

is well–defined (because it depends only on the characteristic function $\chi_{E\times F} = \chi_E \otimes \chi_F$ and not on the representation of $E \times F$ as a countable disjoint union of rectangles) and is countably additive on the set $\mathcal{G}$ of all measurable rectangles, and hence is a premeasure on the generated algebra $\mathcal{A}$. Then we can define the product measure on $\mathcal{M} \otimes \mathcal{N}$ as in 5.1. From now on we call *Tonelli function* every function f of the class $\mathcal{I}$ such that $I(f) = \int_{X\times Y} f\, d\mu \otimes \nu$; a Tonelli set is a subset of $X \times Y$ whose characteristic function is a Tonelli function; thus all measurable rectangles and their finite unions are Tonelli sets; from (*) we get that the class $\mathcal{T}$ of all Tonelli sets is also closed under countable disjoint union. But as the following example shows, in absence of $\sigma-$finiteness *not all measurable sets are Tonelli sets.*

5.1.2. *An example.* First we get some practice with the product measure:

Exercise 5.1.1. Prove that a subset E of $X \times Y$ has $\sigma-$finite $\mu \otimes \nu-$outer measure if and only if

$$E \subseteq (P \times Y) \cup (X \times Q) \cup (A \times B),$$

where P, Q are null sets, and A, B are sets of $\sigma-$finite measure, in X and Y respectively.

Solution. Since $\mu \otimes \nu(P \times Y) = \mu \otimes \nu(X \times Q) = 0$, it is clear that a set verifying this condition has $\sigma-$finite outer measure; let's prove the converse. Clearly E has $\sigma-$finite $\mu \otimes \nu-$outer measure iff there exists a sequence $(A_k \times B_k)_{k\in\mathbb{N}}$ of rectangles of finite measure ($\mu \otimes \nu(A_k \times B_k) = \mu(A_k)\,\nu(B_k) < \infty$, for every $k \in \mathbb{N}$) which covers E. Let

$$I = \{k \in \mathbb{N} : \mu(A_k) = 0\};\ J = \{k \in \mathbb{N} : \nu(B_k) = 0\}; K = \mathbb{N} \smallsetminus (I \cup J).$$

If $P = \bigcup_{k\in I} A_k$, $Q = \bigcup_{k\in J} B_k$, $A = \bigcup_{k\in K} A_k$, $B = \bigcup_{k\in K} B_k$ it is clear that these sets are as required. $\square$

Now an example to show that not all $\mathcal{M} \otimes \mathcal{N}-$measurable sets are Tonelli sets. Consider $X = Y = [0, 1]$ with $\mathcal{M} = \mathcal{N} = \mathcal{B}([0, 1])$, Borel subsets of $[0, 1]$, μ counting measure, ν Lebesgue measure. Let $D = \{(x, x) : x \in [0, 1]\}$ be the diagonal; clearly D is $\mathcal{M} \otimes \mathcal{N}-$measurable ($D$ is closed in the square $[0, 1]^2$, hence it is a Borel subset of the square, and $\mathcal{B}([0, 1]^2) = \mathcal{B}([0, 1]) \otimes \mathcal{B}([0, 1]) = \mathcal{M} \otimes \mathcal{N}$, by second countability of $[0, 1]$). Clearly

$$\int_X \left(\int_Y \chi_D(x, y)\, d\nu(y)\right) d\mu(x) = \int_X 0\, d\mu(x) = 0; \quad \int_Y \left(\int_X \chi_D(x, y)\, d\mu(x)\right) d\nu(y) = \int_Y 1\, d\nu(y) = 1,$$

so that χ_D is not a Tonelli function. Let us prove that $\mu \otimes \nu(D) = \infty$; even more, every subset E of D of $\sigma-$finite $\mu \otimes \nu-$measure is contained in a rectangle such as $X \times N$, with N of Lebesgue measure 0, and hence has $\mu \otimes \nu-$measure zero. In fact the previous exercise shows that $E \subseteq (X \times Q) \cup (A \times B)$, where Q has Lebesgue measure zero, and A and B are $\sigma-$finite (in the counting measure the only null set is the empty set); then A is countable, so that $F = D \cap (A \times B) = \{(x, x) : x \in A \cap B\}$ is countable. It follows that $p_2(E) \subseteq Q \cup p_2(F)$; if N is this set, then $\nu(N) = 0$, and $E \subseteq X \times N$. Thus there are situations in which integrals of measures of sections have no relation to the measure of the set; absence of $\sigma-$finiteness is at the root of this bad behavior.

5.1.3. *Fubini–Tonelli's theorem.* In all the following we assume that $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are fixed measure spaces. A rectangle with $\sigma-$finite sides in $X \times Y$ is a measurable rectangle $A \times B$ whose sides A, B are both $\sigma-$finite in X and Y respectively; if X and Y are both $\sigma-$finite, then clearly the entire space $X \times Y$ is a rectangle with $\sigma-$finite sides.

Lemma. *Every measurable subset of $X \times Y$ contained in a rectangle with $\sigma-$finite sides is a Tonelli set.*

Proof. First we prove that every measurable subset E of a measurable rectangle $A \times B$ whose sides A, B have both finite measure is a Tonelli set. Consider the class of sets:

$$\mathcal{D} = \{E \subseteq A \times B : E \text{ is a Tonelli set}\}.$$

Clearly $A \times B \in \mathcal{D}$; by (*) of 5.1 $\mathcal{D}$ is closed under countable disjoint union; if we prove that $E \in \mathcal{D}$ implies $A \times B \smallsetminus E \in \mathcal{D}$, then $\mathcal{D}$ is a Dynkin class of parts of $A \times B$, which contains all measurable rectangles contained in $A \times B$; since these rectangles are closed under intersection, this Dynkin class is a $\sigma-$algebra of parts of $A \times B$ (see 2.6.1) which contains all rectangles contained in $A \times B$; thus $\mathcal{D}$ contains all $\mathcal{M} \otimes \mathcal{N}-$measurable subsets of $A \times B$.

We have, assuming $E \in \mathcal{D}$:

$$\int_X \left(\int_Y (\chi_{A\times B}(x,y) - \chi_E(x,y))\, d\nu(y)\right) d\mu(x) = \int_X \left(\chi_A(x) \int_Y \chi_B(y) - \int_Y \chi_E(x,y)\, d\nu(y)\right) d\mu(x) =$$
$$\int_X \left(\chi_A(x)\, \nu(B) - \int_Y \chi_E(x,y)\, d\nu(y)\right) d\mu(x) = \mu(A)\, \nu(B) - I(\chi_E) = \mu\otimes\nu(A\times B) - \mu\otimes\nu(E) =$$
$$\mu\otimes\nu(A\times B \setminus E),$$

and clearly the same holds exchanging x and y, so that $A\times B\setminus E \in \mathcal{D}$. Finally, every rectangle $A\times B$ with $\sigma-$finite sides can be written as a countable disjoint union $\bigcup_{k=0}^\infty A_k\times B_k$ of rectangles whose sides have finite measure; if $E\subseteq A\times B$ is measurable, and $E_k = E\cap(A_k\times B_k)$, then every E_k is a Tonelli set, hence so is E, by (*) of 5.1. □

. The Fubini–Tonelli theorem *Let $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$ be $\sigma-$finite measure spaces.*

(i) (Tonelli) *If $f\in L^+(X\times Y)$ then*

$$\int_{X\times Y} f(x,y)\, d\mu\otimes\nu(x,y) = \int_X \left(\int_Y f(x,y)\, d\nu(y)\right) d\mu(x) = \int_Y \left(\int_X f(x,y)\, d\mu(x)\right) d\nu(y).$$

(ii) (Fubini) *If $f\in L^1(\mu\otimes\nu)$ then the $x-$section $f(x,\#)$ of f is in $L^1(\nu)$ for $\mu-$a.e. $x\in X$, the a.e. defined function $\varphi(x) = \int_Y f(x,y)\, d\nu(y)$ is in $L^1(\mu)$ and we have $\int_X \varphi(x)\, d\mu(x) = \int_{X\times Y} f(x,y)\, d\mu\otimes\nu(x,y)$, in other words*

$$\int_{X\times Y} f(x,y)\, d\mu\otimes\nu(x,y) = \int_X \left(\int_Y f(x,y)\, d\nu(y)\right) d\mu(x).$$

Of course an entirely analogous statement holds with x and y exchanged.

Proof. (i) By the preceding lemma and $\sigma-$finiteness of the measures the characteristic functions of all measurable sets in the product are Tonelli functions, so that every positive simple function is Tonelli, and hence every function in $L^+(X\times Y)$ is Tonelli.

(ii) With the usual decomposition of a measurable function we can assume $f\in L^+(X\times Y)$ and the theorem is already proved by (i). We only observe that for f positive the integral of the $x-$section $\varphi(x) = \int_Y f(x,y)\, d\nu(y)$ is defined for every $x\in X$, possibly being ∞, and that $\varphi\in L^+(X)$; since $\int_X \varphi\, d\mu < \infty$, the set $\{\varphi = \infty\}$ has $\mu-$measure zero: not necessarily all sections are in L^1, in general only almost all. □

Remark. We assume $\sigma-$finiteness to make the statement simpler; in fact, for the validity of (i) it is enough to assume that $\mathrm{Coz}(f)$ is contained in a rectangle with $\sigma-$ finite sides. And since cozero sets of functions in L^1 always have $\sigma-$finite measure, Fubini's theorem is always applicable.

5.1.4. *Completion of a product measure.* Even if the factor spaces $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$ are complete, the product measure is almost never complete: simply note that if $P\in\mathcal{M}$ is a non–empty $\mu-$null set then $P\times Y$ has $\mu\otimes\nu-$measure zero, but if there is $S\subseteq Y$ with $S\notin\mathcal{N}$ then $P\times S\notin\mathcal{M}\otimes\mathcal{N}$. But we do not discuss this situation in depth. If $(X\times Y,\mathcal{L},\lambda)$ is the completion of $\mu\otimes\nu$, then for every $\mathcal{L}-$measurable function $f: X\times Y\to\mathbb{K}$ there is a $g: X\times Y\to\mathbb{K}$ which is $\mu\otimes\nu-$measurable and such that $\{f\neq g\}$ is a $\mu\otimes\nu-$null set. So we may always replace f with such a g and reason on g, assuming of course $\sigma-$finite factor spaces.

5.1.5. *Finite products.* There is no difficulty in treating products of finite families of measure spaces by induction on the number of factors. There is an obvious associativity of products, in the sense that $(X\times Y)\times Z$ is naturally identified with $X\times(Y\times Z)$, and both are identified with the *set of ordered triples* $X\times Y\times Z$, analogous identifications with $\sigma-$algebras and measures.

The Lebesgue measure on $\mathbb{R}^n$ was defined as the complete Carathèodory extension of the premeasure defined on intervals by $\lambda_n\left(\prod_{k=1}^n I_k\right) = \prod_{k=1}^n \lambda_1(I_k)$; it is then also the completion of the product measure $\lambda_1^{\otimes n} = \lambda_1\otimes\cdots\otimes\lambda_1$ on Borel subsets of $\mathbb{R}^n$ (these measures coincide and are finite on compact intervals): Lebesgue measurable sets are obtained adding to Borel sets sets of outer measure zero (or removing from them sets of outer measure 0).

Exercise 5.1.2. Consider the map $\delta:\mathbb{R}^{2n} = \mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}^n$ given by $\delta(x,y) = x-y$ (difference). Prove that if $N\subseteq\mathbb{R}^n$ is λ_n-null, then $\delta^{\leftarrow}(N)$ is $\lambda_{2n}-$null. Deduce from this that if $f:\mathbb{R}^n\to\mathbb{K}$ is Lebesgue measurable then $f\circ\delta:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{K}$ (i.e the map $(x,y)\mapsto f(x-y)$) is also Lebesgue measurable.

Solution. If N has Lebesgue measure zero, it is contained in a Borel set $B \subseteq \mathbb{R}^n$ of λ_n-measure zero. Then $\delta^{\leftarrow}(B) = C$ is a Borel subset of $\mathbb{R}^n \times \mathbb{R}^n$. On Borel sets we have

$$\lambda_{2n}(C) = \int_{\mathbb{R}^n} \lambda_n(C_x)\, d\lambda_n(x) \quad \text{where } C_x = \{y \in \mathbb{R}^n : (x,y) \in C\},$$

equivalently $C_x = \{y \in \mathbb{R}^n : x - y \in B\} = x - B$; then $\lambda_n(C_x) = \lambda_n(x - B) = \lambda_n(B) = 0$ for every $x \in \mathbb{R}^n$, so that $\lambda_{2n}(C) = 0$. Since $\delta^{\leftarrow}(N) \subseteq C$, this implies $\lambda_{2n}(\delta^{\leftarrow}(N)) = 0$.

Given a Borel subset V of $\mathbb{K}$ the set $f^{\leftarrow}(V)$ is Lebesgue measurable in $\mathbb{R}^n$, hence it has the form $A \cup N$, with $A \in \mathcal{B}(\mathbb{R}^n)$ and N a Lebesgue null set; then $(f \circ \delta)^{\leftarrow}(V) = \delta^{\leftarrow}(A) \cup \delta^{\leftarrow}(N)$, with $\delta^{\leftarrow}(A) \in \mathcal{B}(\mathbb{R}^{2n})$ and $\lambda_{2n}(\delta^{\leftarrow}(N)) = 0$.

□

Exercise 5.1.3. Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $f \in \mathcal{L}^+(X)$.

(i) Prove that all these sets are measurable in $X \times [0, \infty]$ with $\mathcal{M} \otimes \mathcal{B}([0,\infty])$: the trapezoid $\mathrm{Trap}(f) = \{(x,y) \in X \times [0,\infty] : 0 \le y \le f(x)\}$; the epigraphic $\mathrm{Epi}(f) = \{(x,y) \in X \times [0,\infty] : f(x) \le y\}$ and the graphic $\mathrm{Graph}(f) = \{(x,y) \in X \times [0,\infty] : y = f(x)\}$.

Assume now that the measure space is $\sigma-$finite.

(ii) Prove that (λ is one dimensional Lebesgue measure on $[0,\infty]$)

$$\int_X f = \mu \otimes \lambda(\mathrm{Trap}(f)); \qquad \int_X f = \int_0^\infty \mu(\{f > t\})\, dt$$

(the last formula is already known).

(iii) Prove that the graph of f has measure zero.

Solution. (i) The function $g : X \times [0,\infty] \to [0,\infty] \times [0,\infty]$ defined by $g(x,y) = (f(x), y)$ is measurable (its composition with the first projection is $f \circ p_X$, with the second projection the composition is $p_{[0,\infty]}$). Then

$$\mathrm{Trap}(f) = g^{\leftarrow}(\{(s,t) \in [0,\infty]^2 : s \le t\}); \quad \mathrm{Epi}(f) = g^{\leftarrow}(\{(s,t) \in [0,\infty]^2 : s \ge t\});$$
$$\mathrm{Graph}(f) = g^{\leftarrow}(\{(s,t) \in [0,\infty]^2 : s = t\}).$$

(ii) We get, by Tonelli's thorem:

$$\mu \otimes \lambda(\mathrm{Trap}(f)) = \int_X \left(\int_{[0,+\infty]} \chi_{\mathrm{Trap}(f)}(x,y)\, d\lambda(y) \right) d\mu(x) = \int_X f\, d\mu.$$

(note that the $x-$section of $\chi_{\mathrm{Trap}(f)}$ is $\chi_{[0,f(x)]}$). Integrating in the other way:

$$\mu \otimes \lambda(\mathrm{Trap}(f)) = \int_{[0,\infty]} \left(\int_X \chi_{\mathrm{Trap}(f)}(x,y)\, d\mu(x) \right) d\lambda(y),$$

and $\{x \in X : (x,y) \in \mathrm{Trap}(f)\} = \{x \in X : y \le f(x)\}$, so that

$$\int_X \chi_{\mathrm{Trap}(f)}(x,y)\, d\mu(x) = \mu(\{f \ge y\}) = \rho_f(y^-),$$

if $\rho_f(y) := \mu(\{f > y\}$ (see 4.1.14). Then

$$\mu \otimes \lambda(\mathrm{Trap}(f)) = \int_{[0,\infty]} \rho_f(y^-)\, dy \left(= \int_{[0,\infty]} \rho_f(y)\, dy \right),$$

the last equality due to the fact that $\rho_f(y) = \rho_f(y^-)$ for all $y \in [0,\infty]$ but a countable subset.

(iii) The $x-$sections ot the graph are singletons, of zero Lebesgue measure in $[0,\infty]$. □

Exercise 5.1.4. For every $a > 0$ let $E(a) = [0,a] \times [0,\infty[$ and let $E = [0,\infty[^2$. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x,y) = e^{-xy} \sin x$.

(i) Prove that $f \in L^1(E(a))$ for every $a > 0$ and express the integral of f on $E(a)$ as an iterated integral, reducing it to integrals in one variable, in both ways.

(ii) Prove that $f \notin L^1(E)$.

(iii) In spite of (ii), $\lim_{a\to\infty}\int_{E(a)} f(x,y)\,dxdy$ exist and is finite; compute it, and use it to evaluate the generalized integral

$$\int_0^{\uparrow\infty}\frac{\sin x}{x}\,dx := \lim_{a\to\infty}\int_0^a \frac{\sin x}{x}\,dx.$$

Solution. (i) The function f is continuous and hence measurable, and all sets are measurable. We apply Tonelli's theorem to $|f|$, and check finiteness of the iterated integral, integrating in y first:

$$\int_{E(a)} |f(x,y)|\,dxdy = \int_0^a \left(\int_{y=0}^{y=\infty} e^{-xy}\,dy\right)\sin x\,dx = \int_0^a \frac{|\sin x|}{x}\,dx < \infty,$$

since clearly $x\mapsto |\sin x|/x$ is bounded measurable on the set $[0,a]$, of finite measure. We may as well answer now to (ii): the same reduction brings us to

$$\int_E |f(x,y)|\,dxdy = \int_0^\infty \frac{|\sin x|}{x}\,dx.$$

It is well known, and worth remembering, that this last integral is ∞:

$$\int_0^\infty \frac{|\sin x|}{x}\,dx = \sum_{k=1}^\infty \int_{(k-1)\pi}^{k\pi}\frac{|\sin x|}{x}\,dx \ge \sum_{k=1}^\infty \int_{(k-1)\pi}^{k\pi}\frac{|\sin x|}{k\pi}\,dx = \frac{2}{\pi}\sum_{k=1}^\infty \frac{1}{k} = \infty.$$

Let's complete the answer to (i): removing the absolute value clearly brings

$$\int_{E(a)} f(x,y)\,dxdy = \int_0^a \frac{\sin x}{x}\,dx;$$

integrating first in the variable x we get

$$\int_{E(a)} f(x,y)\,dxdy = \int_{y=0}^{y=\infty}\left(\int_{x=0}^{x=a} e^{-xy}\sin x\,dx\right)dy = \int_{y=0}^{y=\infty}\left[\frac{-e^{-xy}}{1+y^2}(\cos x + y\,\sin x)\right]_{x=0}^{x=a} dy =$$
$$\int_0^\infty \frac{dy}{1+y^2} - \int_0^\infty \frac{e^{-ay}}{1+y^2}(\cos a + y\sin a)\,dy.$$

We have obtained that for every $a>0$:

$$\int_{E(a)} f(x,y)\,dxdy = \int_0^a \frac{\sin x}{x}\,dx = \frac{\pi}{2} - \int_0^\infty \frac{e^{-ay}}{1+y^2}(\cos a + y\sin a)\,dy.$$

Now we have, for $y \ge 0$:

$$\frac{|\cos a + y\sin a|}{1+y^2} \le \frac{1+y}{1+y^2};$$

the function $y\mapsto (1+y)/(1+y^2)$ has a finite maximum μ in $[0,\infty]$ (it is continuous, positive, and tends to 0 at infinity; we could easily compute μ, but it is not relevant), so that

$$\left|\int_0^\infty \frac{e^{-ay}}{1+y^2}(\cos a + y\sin a)\,dy\right| \le \int_0^\infty \left|\frac{e^{-ay}}{1+y^2}(\cos a + y\sin a)\right|dy \le \mu\int_0^\infty e^{-ay}\,dy = \frac{\mu}{a},$$

which tends to 0 as $a\to\infty$. We have obtained the important result (Dirichlet's integral):

$$\int_0^{\uparrow\infty}\frac{\sin x}{x}\,dx := \lim_{a\to\infty}\int_0^a \frac{\sin x}{x}\,dx = \frac{\pi}{2}.$$

We reiterate the fact that *this is not a Lebesgue integral*: as seen above $\sin x/x \notin L^1_m([0,\infty[)$ □

5.1.6. *Partial integration for Radon–Stieltjes measures.*

. *Let $F, G:\mathbb{R}\to\mathbb{R}$ be increasing and right–continuous. Denote the respective measures by $\int_E dF$ and $\int_E dG$ respectively. Prove that if $a,b\in\mathbb{R}$ and $a<b$ then*

$$\int_{]a,b]} F(x^-)\,dG(x) = F(b)G(b) - F(a)G(a) - \int_{]a,b]} G(x)\,dF(x)$$

(consider the triangle $T=\{(x,y)\in\mathbb{R}^2 : a<y\le x\le b\}$; compute the product measure of this triangle with Fubini's theorem …). Prove also that if F and G do not have a common point of discontinuity (in particular, if one of them is continuous) then the formula can also be written as:

$$\int_{]a,b]} F(x)\,dG(x) = F(b)G(b) - F(a)G(a) - \int_{]a,b]} G(x)\,dF(x).$$

Proof. Observe first that all measures considered are $\sigma-$finite (actually even finite on the sets considered) so that Fubini's theorem is applicable. Integrating first with respect to y, then to x

$$\int_T (dF \otimes dG) = \int_{x\in]a,b]} \left(\int_{y\in T(x)} dG(y) \right) dF(x) = \int_{x\in]a,b]} (G(x) - G(a))\, dF(x) =$$

$(T(x) = \{y :\ a < y \le x\} =]a,x])$

$$\int_{]a,b]} G(x)\, dF(x) - G(a)(F(b) - F(a));$$

reversing the order of integration:

$$\int_T (dF \otimes dG) = \int_{y\in]a,b]} \left(\int_{x\in T(y)} dF(x) \right) dG(y) = \int_{y\in]a,b]} (F(b) - F(y^-))\, dG(y) =$$

$(T(y) = \{x :\ y \le x \le b\} = [y,b])$

$$F(b)(G(b) - G(a)) - \int_{]a,b]} F(y^-)\, dG(y);$$

comparison of results yields the formula.

If F and G do not have a common point of discontinuity then the set of discontinuities of F has dG measure zero, hence the functions F and $x \mapsto F(x^-)$ are $dG-$a.e. equal, so that

$$\int_{]a,b]} F(x^-)\, dG(x) = \int_{]a,b]} F(x)\, dG(x).$$

□

Remark. We can drop the requirement of right continuity; the same proof leads then to the formula

$$\int_{]a,b]} F(x^-)\, dG(x) = F(b^+)G(b^+) - F(a^+)G(a^+) - \int_{]a,b]} G(x^+)\, dF(x),$$

which simplifies to

$$\int_{]a,b]} F(x)\, dG(x) = F(b^+)G(b^+) - F(a^+)G(a^+) - \int_{]a,b]} G(x)\, dF(x)$$

if F and G have no common point of discontinuity.

Exercise 5.1.5. (*Lebesgue measure of the euclidean unit ball in* $\mathbb{R}^n$) By Fubini's theorem we have, if $f(x) = e^{-|x|^2}$:

$$\int_{\mathbb{R}^n} f(x)\, d\lambda_n(x) = \int_{\mathbb{R}^n} e^{-x_1^2} \dots e^{-x_n^2}\, d\lambda_1^{\otimes n} = \prod_{k=1}^n \int_{\mathbb{R}} e^{-x^2}\, dx = \pi^{n/2}.$$

Compute the function $\rho_f(t) = \lambda_n(\{f > t\})$ and $\int_0^\infty \rho_f(t)\, dt$, and deduce that, if $B = \{x \in \mathbb{R}^n :\ |x| \le 1\}$ we have

$$\lambda_n(B) = \frac{\pi^{n/2}}{\Gamma(n/2+1)} = \frac{2\,\pi^{n/2}}{n\,\Gamma(n/2)}.$$

Solution. Clearly $\rho_f(t) = 0$ for $t \ge 1$. If $0 < t < 1$ we get $\{f > t\} = \{x \in \mathbb{R}^n :\ |x| < (\log(1/t))^{1/2}\}$, the open ball of center 0 and radius $(\log(1/t))^{1/2}$, with measure $(\log(1/t))^{n/2}\, \lambda_n(B)$, so that (using also the change of variable $t = e^{-x}$):

$$\int_0^\infty \rho_f(t)\, dt = \lambda_n(B) \int_0^1 (\log(1/t))^{n/2}\, dt = \lambda_n(B) \int_0^\infty x^{n/2}\, e^{-x}\, dx = \lambda_n(B)\, \Gamma(n/2+1).$$

Observe that $\lambda_1(B) = 2$, $\lambda_2(B) = \pi$, $\lambda_3(B) = 4\pi/3$, $\lambda_4(B) = \pi^2/2$, ecc. □

Exercise 5.1.6. Prove the formula

$$\Gamma(x+y)\, B(x,y) = \Gamma(x)\, \Gamma(y) \qquad \left(\mathrm{Re}\, x,\ \mathrm{Re}\, y > 0,\ B(x,y) := \int_0^1 (1-t)^{x-1}\, t^{y-1}\, dt \right);$$

(you may find it, e.g, in Analisi Due, 9.24.7).

6. L^p SPACES

There are various situations in which one needs to consider integrals of a measurable function raised to some power; the most common situation is integrals of squares of functions, but other exponents are also sometimes useful. The theory is very important in modern analysis.

6.1. Basics.

DEFINITION. Given a measure space $(X, \mathcal{M}, \mu)$, a measurable function $f \in \mathcal{L}(X, \mathbb{K})$, and a real number $p > 0$ we set

$$\|f\|_p = \left(\int_X |f|^p\right)^{1/p}.$$

This is a non–negative real number, or ∞ ($\infty^{1/p} := \infty$), and is zero iff $f = 0$ a.e.. We denote $\mathcal{L}^p_\mu(X, \mathbb{K})$ the set of all measurable $\mathbb{K}-$valued functions for which $\|f\|_p < \infty$.

Thus $\mathcal{L}^p_\mu(X, \mathbb{K})$ may also be described as the set of all measurable functions such that $|f|^p \in \mathcal{L}^1(\mu)$. For $p = 1$ we naturally re–obtain the already familiar $\mathcal{L}^1_\mu(X, \mathbb{K})$. We easily see that

. *For every $p > 0$ the set $\mathcal{L}^p_\mu(X, \mathbb{K})$ is a vector subspace of $L(X, \mathbb{K})$.*

Proof. If $f, g \in \mathcal{L}^p_\mu(X, \mathbb{K})$ then

$$|f + g|^p \le (|f| + |g|)^p \le (2|f| \vee |g|)^p = 2^p|f|^p \vee |g|^p \le 2^p(|f|^p + |g|^p),$$

and the last function is clearly in $L^1(\mu)$, since $|f|^p$, $|g|^p \in L^1(\mu)$. For $\lambda \in \mathbb{K}$ we get $\|\lambda f\|_p = |\lambda| \, \|f\|_p$ □

The quotient space $\mathcal{L}^p_\mu(X, \mathbb{K})/\mathcal{N}_\mu(X, \mathbb{K})$ where a.e. equal functions are identified will be denoted $L^p_\mu(X, \mathbb{K})$ or simply $L^p(\mu)$. Caution: $f \mapsto \|f\|_p$ will be a norm on $L^p(\mu)$ *only for* $p \ge 1$; if $0 < p < 1$ the function $\|\#\|_p$ is absolutely homogeneous, as observed in the above proof, but it is not subadditive, barring extremely trivial cases (6.4.2).

6.1.1. *$L^\infty(\mu)$ and $L^p(\mu)$.* The space $L^\infty(\mu)$ of essentially bounded functions has ben defined in 4.4; let us relate it to $L^p(\mu)$. On infinite measure spaces constant (non–zero) functions are not in L^p, for no $p > 0$, and in general there are bounded functions not in any L^p; for these functions their finite L^∞ norm is not related to their infinite L^p norms. But we have

. *If $\|f\|_\infty = \infty$, or if $f \in L^p(\mu)$ for some $p > 0$ then $\|f\|_\infty = \lim_{q\to\infty} \|f\|_q$.*

Proof. We can assume $f \ne 0$. Let's prove first that $\liminf_{q\to\infty} \|f\|_q \ge \|f\|_\infty$. Take $0 < \alpha < \|f\|_\infty$; then $\{|f| \ge \alpha\}$ has strictly positive measure and for every $q > 0$:

$$\|f\|_q = \left(\int_X |f|^q\right)^{1/q} \ge \left(\int_{\{|f|\ge\alpha\}} |f|^q\right)^{1/q} \ge \left(\int_{\{|f|\ge\alpha\}} |\alpha|^q\right)^{1/q} = \alpha \, (\mu(\{|f| \ge \alpha\}))^{1/q};$$

so that, for every $\alpha > 0$:

$$\|f\|_q \ge \alpha \, (\mu(\{|f| \ge \alpha\}))^{1/q};$$

taking liminf in the preceding inequality we obtain $\liminf_{q\to\infty} \|f\|_q \ge \alpha$ (if $\mu(\{f \ge \alpha\}) = \infty$ then both sides are ∞ for every q; otherwise $\lim_{q\to\infty} \alpha \, (\mu(\{|f| \ge \alpha\}))^{1/q} = \alpha$). Since α is an arbitrary positive number strictly less than $\|f\|_\infty$, we conclude. If $\|f\|_\infty = \infty$ the proof is concluded. Assume then $\|f\|_\infty < \infty$ and that $f \in L^p(\mu)$ for some $p > 0$, and let's prove that $\limsup_{q\to\infty} \|f\|_q \le \|f\|_\infty$; we have, for $q > p$: $|f|^q = |f|^{q-p} \, |f|^p \le \|f\|_\infty^{q-p} \, |f|^p$, so that

$$\int_X |f|^q \le \|f\|_\infty^{q-p} \int_X |f|^p \Longrightarrow \|f\|_q \le \|f\|_\infty^{1-p/q} \, \|f\|_p^{p/q},$$

and taking limsups as $q \to \infty$ we get the desired inequality.

□

REMARK. Observe that if $\mu(X) < \infty$ then $\|f\|_\infty = \lim_{q\to\infty} \|f\|_q$ for every measurable function $f \in L(X)$.

In the course of the proof of the preceding proposition we have proved the:

. ČEBIČEFF'S INEQUALITY FOR L^p. *For every measurable f and every α, $p > 0$ we have*

$$\|f\|_p \ge \alpha \, (\mu(\{|f| \ge \alpha\}))^{1/p}.$$

6.1.2. *Conjugate exponents and Hölder inequality.* Given $p>1$ there is one and only one number $\tilde p>1$ such that $(1/p)+(1/\tilde p)=1$, namely $\tilde p=p/(p-1)$; $\tilde p$ is the *conjugate exponent* of p, and p, $\tilde p$ are a pair of conjugate exponents (clearly p is the exponent conjugate to $\tilde p$). To complete the definition, we say that $1,\infty$ are also a pair of conjugate exponents; notice that 2 is a self–conjugate exponent. Recall the familiar arithmetico–geometric inequality: if $u,\, v>0$ then we have $\sqrt{u\,v}\le(u+v)/2$: for two positive real numbers the geometric mean $\sqrt{u\,v}$ is smaller than the arithmetic mean $(u+v)/2$. This inequality generalizes to weighted averages: if $\alpha,\beta>0$ and $\alpha+\beta=1$, then, for every pair of real numbers $u,\, v>0$:

$$u^\alpha\, v^\beta\le\alpha\, u+\beta\, v\quad\text{(with equality iff } u=v).$$

The proof is an immediate consequence of the (strict) concavity of the logarithm function on $]0,\infty[$; if $u,\, v$ are distinct strictly positive real numbers, and $\alpha+\beta=1$ with $\alpha,\beta>0$, then the *convex combination* $\alpha\, u+\beta\, v$ belongs to the open interval of extremes u and v, so that

$$\log(\alpha\, u+\beta\, v)>\alpha\,\log u+\beta\,\log v=\log(u^\alpha\, v^\beta)\iff u^\alpha\, v^\beta<\alpha\, u+\beta\, v.$$

We need the following corollary:

. Young's inequality *If $p,q>1$ are conjugate exponents (i.e. $1/p+1/q=1$) and $a,\, b\ge 0$ then*

$$ab\le\frac{a^p}{p}+\frac{b^q}{q},$$

and equality holds if and only if $a^p=b^q$.

Proof. In the above inequality put $\alpha=1/p$, $\beta=1/q$, $u=a^p$, $v=b^q$. □

. Hölder's inequality *Let $(X,\mathcal{M},\mu)$ be a measure space, let f,g be measurable, and let p,q be conjugate exponents. Then*

$$\|f\,g\|_1\le\|f\|_p\,\|g\|_q.$$

Proof. The case $p=1$, $q=\infty$ is trivial, and has been proved in 4.4.2. Assume then $p,\, q>1$. If f, or g, is a null function, then both sides of the inequality are zero. Excluding this case, if $\|f\|_p=\infty$ or if $\|g\|_q=\infty$ the right hand–side is ∞. We are then reduced to the case in which $\|f\|_p$ and $\|g\|_q$ are both finite and nonzero; putting $|f(x)|/\|f\|_p$ in place of a and $|g(x)|/\|g\|_q$ in place of b in the above inequality we get

$$\frac{|f(x)|\,|g(x)|}{\|f\|_p\,\|g\|_q}\le\frac{|f(x)|^p}{p\,\|f\|_p^p}+\frac{|g(x)|^q}{q\,\|g\|_q^q};$$

integrating both sides of this inequality we get

$$\frac{1}{\|f\|_p\,\|g\|_q}\,\|f\,g\|_1\le\frac{\|f\|_p^p}{p\,\|f\|_p^p}+\frac{\|g\|_q^q}{q\,\|g\|_q^q}=\frac1p+\frac1q=1,$$

and we conclude. □

When $p=q=2$ Hölder's inequality is also known as the *Cauchy–Schwarz inequality for integrals.*

Exercise 6.1.1. Assuming $p>1$ and $q=\tilde p$, prove that Hölder's inequality is an equality if and only if $|f|^p$ and $|g|^q$ are linearly dependent in $L(X)$.

One often uses the fact that if $f\in L^p(\mu)$ then $|f|^{p-1}\in L^{\tilde p}(\mu)$, in fact $(p-1)\tilde p=p$; the space $L^{\tilde p}(\mu)$ is called *the conjugate space* of $L^p(\mu)$.

6.1.3. *Triangular, or Minkowski's, inequality.* We now prove that on $L^p(\mu)$, for $p\ge 1$ the map $f\mapsto\|f\|_p$ is indeed a norm.

. *Let $(X,\mathcal{M},\mu)$ be a measure space, and let $f,g\in L(X)$ be measurable. If $p\ge 1$ we have*

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

Proof. For $p=1$ and $p=\infty$ the inequality is known, assume $1<p<\infty$. If f or g are not in $L^p(\mu)$ then the right–hand side is ∞. Assuming $f,\, g\in L^p$ we have $f+g\in L^p(\mu)$ and:

$$|f+g|^p=|f+g|\,|f+g|^{p-1}\le|f|\,|f+g|^{p-1}+|g|\,|f+g|^{p-1};$$

we integrate both sides and apply Hölder's inequality, putting for simplicity $q=\tilde p$:

$$\|f+g\|_p^p\le\|f\|_p\,\|f+g\|_p^{p/q}+\|g\|_p\,\|f+g\|_p^{p/q}\Longrightarrow\|f+g\|_p^{p-p/q}\le\|f\|_p+\|g\|_p,$$

and since $p-p/q=1$ we conclude. □

6.2. Completeness of $L^p(\mu)$. We have proved that $L^1(\mu)$ is Banach space. A similar proof yields the same result for $L^p(\mu)$. Remember that convergence of f_n to f in $L^p(\mu)$ is equivalent to assert that $|f - f_n|^p$ converges to 0 in $L^1(\mu)$.

. *Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $\sum_{n=0}^{\infty} f_n$ be a normally convergent series in $L^p(\mu)$. Then the series converges a.e. and in $L^p(\mu)$ to a function $f \in L^p(\mu)$.*

Proof. Let $g_m = \sum_{k=0}^{m} |f_k|$; then g_m is an increasing sequence and

$$\|g_m\|_p \le \sum_{k=0}^{m} \|f_k\|_p \le \sum_{k=0}^{\infty} \|f_k\|_p = S,$$

so that g_m^p is an increasing sequence in $L^1(\mu)$ whose integrals $\int_X g_m^p \le S^p$ have a finite upper bound; then g_m^p converges a.e. to $g^p \in L^1(\mu)$. It follows that the series $\sum_{k=0}^{\infty} f_k(x)$ converges for a.e. $x \in X$ to a function which we call $f(x)$. Let's prove that the series converges to f in $L^p(\mu)$; in fact

$$\left| f(x) - \sum_{k=0}^{m} f_k(x) \right| = \left| \sum_{k=m+1}^{\infty} f_k(x) \right| \le \sum_{k=m+1}^{\infty} |f_k(x)| = g(x) - g_m(x) \le g(x),$$

for a.e. $x \in X$. Then $|f - \sum_{k=0}^{m} f_k|^p \le g^p \in L^1(\mu)$; by dominated convergence, we have

$$\lim_{m\to\infty} \int_X |f - \sum_{k=0}^{m} f_k|^p = 0, \text{ but } \int_X |f - \sum_{k=0}^{m} f_k|^p = \left\| f - \sum_{k=0}^{m} f_k \right\|_p^p$$

and the proof ends.

□

6.2.1. *Convergence in $L^p(\mu)$ and pointwise convergence.* The sequence of functions considered in 4.2.7 converges to the zero function in $L^p([0,1])$ for every $p \in]0, \infty[$, but nowhere pointwise, so that convergence in $L^p(\mu)$ does not in general imply a.e. convergence if $p < \infty$. However

Proposition. *For $0 < p < \infty$, if a sequence $f_n \in L^p(\mu)$ converges to f in $L^p(\mu)$, then some subsequence converges to f also a.e.*

Proof. Pick a subsequence $f_{\nu(k)}$ of bounded variation, i.e. such that $\sum_{k=0}^{\infty} \|f_{\nu(k+1)} - f_{\nu(k)}\|_p < \infty$, and apply the theorem on normal convergence of series in $L^p(\mu)$. □

Of course, convergence of a sequence in $L^\infty(\mu)$ is uniform convergence on the complement of some set of measure zero, and implies convergence a.e.

6.2.2. *Approximation of L^p functions.* Recall that $S(\mu)$ is the set of simple functions that are in $L^1(\mu)$ (4.2.9). Trivially, $S(\mu) \subseteq L^p(\mu)$ for every $p \ge 0$ (notice that if $\varphi = \sum_{j=1}^{m} \alpha_j \chi_{E(j)}$, standard representation with the value 0 omitted if in the range of φ, then $|\varphi|^p = \sum_{j=1}^{m} |\alpha_j|^p \chi_{E(j)} \in S(\mu)$, since all $E(j)$ have finite measure if $\varphi \in S(\mu)$). It is easy to see that $S(\mu)$ is dense in $L^p(\mu)$: if $f \in L^p(\mu)$ and φ_n is a sequence of simple functions converging pointwise to f, and such that $|\varphi_n| \uparrow |f|$ (3.2.2), then $\varphi_n \in L^p(\mu)$ and $\|f - \varphi_n\|_p \to 0$ as $n \to \infty$ ($|f - \varphi_n|^p \le 2^p |f|^p \in L^1(\mu)$). Now, if E and F are sets of finite measure, and $p > \infty$ then

$$\|\chi_E - \chi_F\|_p = \left(\int_X |\chi_E - \chi_F|^p \right)^{1/p} = \left(\int_X \chi_{E \triangle F}^p \right)^{1/p} = \mu(E \triangle F)^{1/p} = \|\chi_E - \chi_F\|_1^{1/p},$$

so that if the measure μ is the Carathèodory extension of a premeasure on an algebra $\mathcal{A}$, for every $f \in L^p(\mu)$ and every $\varepsilon > 0$ there is an $\mathcal{A}$−simple function g which is also in $S(\mu)$ such that $\|f - g\|_p \le \varepsilon$. In particular, if μ is a Radon measure on $\mathbb{R}^n$, for every $f \in L^p_\mu(\mathbb{R}^n)$ and $\varepsilon > 0$ there is a step function φ with compact support such that $\|f - \varphi\|_p \le \varepsilon$. We leave to the reader the proof of the fact that continuous compactly supported functions can also approximate functions in $L^p_\mu(\mathbb{R}^n)$ if $p < \infty$: the proof in 4.2.9 is easily adapted to this case. For $p = \infty$ and $\mu = m$, Lebesgue measure, the closure of $C_c(\mathbb{R}^n)$ in $L^\infty_m(\mathbb{R}^n)$ is the space $C_0(\mathbb{R}^n)$ of continuous functions in $\mathbb{R}^n$ which are 0 at infinity, i.e. such that $\lim_{x\to\infty} f(x) = 0$; the proof is not difficult (but perhaps non–trivial for beginners).

6.3. Variation of L^p with p. As a paradigmatic easy example consider the function $f(x) = 1/x$ for $x > 0$: we have that $f \in L^p([0, 1])$ for $0 < p < 1$, while $f(x) \in L^p([1, \infty[)$ for $p > 1$; there is no $p > 0$ such that $f \in L^p(]0, \infty[)$. We have

Proposition. *Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $0 < p < q < r \le \infty$. Then $L^q(\mu) \subseteq L^p(\mu) + L^r(\mu)$, and $L^p(\mu) \cap L^r(\mu) \subseteq L^q(\mu)$; moreover, if $1/q = \alpha/p + \beta/r$, a convex combination ($\alpha, \beta > 0$, $\alpha + \beta = 1$) then we have*

$$\|f\|_q \le \|f\|_p^{\alpha}\, \|f\|_r^{\beta}.$$

Proof. We can write a measurable $f : X \to \mathbb{K}$ as the sum of a function whose non zero values are of absolute value larger than 1, and a function bounded by 1: simply take $E = \{|f| > 1\}$, $F = X \smallsetminus E = \{|f| \le 1\}$ and write $f = f\,\chi_E + f\,\chi_F$. We have $|f|^q = |f|^q\,\chi_E + |f|^q\,\chi_F$, for every $q > 0$, so that, if we assume $f \in L^q(\mu)$, we have $|f|^q\,\chi_E$, $|f|^q\,\chi_F \in L^1(\mu)$. Notice that if $p < q$ then $|f|^p\,\chi_E \le |f|^q\,\chi_E$, so that $|f|^p\,\chi_E \in L^1$, i.e. $f\,\chi_E \in L^p$; and if $r > q$ then $|f|^r\,\chi_F \le |f|^q\,\chi_F$, so that $|f|^r\,\chi_F \in L^1$, in other words $f\,\chi_F \in L^r$. This proves that $L^q \subseteq L^p + L^r$. For the second part assume first $r = \infty$; then $\alpha = p/q$, and $\beta = 1 - p/q$, and

$$|f|^q = |f|^p\,|f|^{q-p} \le |f|^p\,\|f\|_\infty^{q-p} \Longrightarrow \int_X |f|^q \le \int_X |f|^p\,\|f\|_\infty^{q-p} \Longrightarrow \|f\|_q \le \|f\|_p^{p/q}\,\|f\|_\infty^{1-p/q}.$$

If $r < \infty$ we have $1 = \alpha\,q/p + \beta\,q/r = 1/(p/(\alpha\,q)) + 1/(r/(\beta\,q))$ and we apply Hölder's inequality with conjugate exponents $p/(\alpha\,q)$ and $r/(\beta\,q)$ to the pair of functions $|f|^{\alpha q}$ and $|f|^{\beta\,q}$:

$$\int_X |f|^q = \int_X |f|^{\alpha q}\,|f|^{\beta\,q} \le \left(\int_X |f|^p\right)^{\alpha\,q/p} \left(\int_X |f|^r\right)^{\beta\,q/r} \Longrightarrow \|f\|_q \le \|f\|_p^{\alpha}\,\|f\|_r^{\beta}.$$

□

Exercise 6.3.1. Strictly connected to the interpolation inequality is the following

. Generalization of Hölder's inequality *If $p, q > 1$ and $1/r = 1/p + 1/q$ then, for any measurable f, g we have*

$$\|f\,g\|_r \le \|f\|_p\,\|g\|_q.$$

Hint: apply Hölder's inequality to $|f|^r$ and $|g|^r$, with conjugate exponents p/r and q/r.

Solution. Following the hint we get

$$\||f|^r |g|^r\|_1 \le \||f|^r\|_{p/r}\,\||g|^r\|_{q/r} = \left(\int_X |f|^p\right)^{r/p} \left(\int_X |g|^q\right)^{r/q} = \|f\|_p^r\,\|g\|_q^r,$$

while the left-hand side is $\|f\,g\|_r^r$.

□

6.3.1. *L^p spaces if $\mu(X) < \infty$.* The following result is very important:

Proposition. *Let $(X, \mathcal{M}, \mu)$ be a finite measure space, $\mu(X) < \infty$. Then $L^p(\mu)$ decreases as p increases, and if $p < q \le \infty$ then*

$$\|f\|_p \le \mu(X)^{1/p - 1/q}\|f\|_q \quad \text{in particular, if } q = \infty \text{ we get } \|f\|_p \le \mu(X)^{1/p}\,\|f\|_\infty.$$

Proof. On finite measure spaces bounded measurable functions are in L^1; then $L^\infty(\mu) \subseteq L^p(\mu)$ for every $p > 0$, if $\mu(X) < \infty$. The preceding proposition 6.3 then implies that if $p < q < \infty$ we have $L^q \subseteq L^p + L^\infty = L^p$, and also the inequality with the norms (which then is obtained by applying Hölder's inequality to $|f|^p$ and $|f|^0 = 1$, with conjugate exponents q/p and $q/(q - p)$). □

Remark. The situation is particularly neat in the case of probability spaces, $\mu(X) = 1$: then $p < q$ implies $\|f\|_p \le \|f\|_q$.

Exercise 6.3.2. Find $f \in \bigcap_{p>0} L^p([0, 1]) \smallsetminus L^\infty([0, 1])$.

6.3.2. *ℓ^p spaces.* The opposite behavior is encountered in the case of ℓ^p spaces, spaces $L^p(X,\mathcal{P}(X),\varkappa)$ where $\varkappa$ is the counting measure. Here $\ell^\infty \supseteq \ell^p$ for all $p>0$: in fact $|f(x)|^p \le \sum_{t\in X}|f(t)|^p = \|f\|_p^p$ for every $x\in X$, so that $|f(x)|\le \|f\|_p$ for every $x\in X$, equivalently $\|f\|_\infty \le \|f\|_p$. Then spaces ℓ^p increase with p, while $p\mapsto \|f\|_p$ decreases. In fact by proposition 6.3 if if $p<q<\infty$ we have $\ell^p\cap\ell^\infty\subseteq \ell^q$, that is $\ell^p\subseteq\ell^q$; and the inequality in that same proposition gives ($\alpha=p/q$, $\beta=1-p/q$)

$$\|f\|_q \le \|f\|_p^{p/q}\,\|f\|_\infty^{1-p/q} \le \|f\|_p^{p/q}\,\|f\|_p^{1-p/q} = \|f\|_p.$$

We have proved:

PROPOSITION. *Let X be a set, and put $\ell^p=\ell^p(X)$. If $0<p<q\le\infty$ then $\ell^p\subseteq\ell^q$, and $\|f\|_p\ge\|f\|_q$.*

Proof. See above. □

The inequality on the norms means that the norm topologies get stronger as p decreases. Uniform convergence on these spaces is weaker than any ℓ^p convergence. On the contrary, on finite measure spaces L^∞ (i.e. uniform a.e.) convergence is stronger than any L^p convergence, and these convergences get weaker as p decreases, as implied by the inequality $\|f\|_p\le \mu(X)^{1/p-1/q}\|f\|_q$ for $p<q$.

EXERCISE 6.3.3. Assume $0<p<\infty$, and $f\in\ell^p$. Prove that for every $\varepsilon>0$ there is a finite set $F_\varepsilon\subseteq X$ such that $|f(x)|<\varepsilon$ for $x\in X\smallsetminus F_\varepsilon$, and deduce from this that there is $c\in X$ such that $|f(c)|=\|f\|_\infty$, and that $\|f\|_p=\|f\|_\infty$ if and only if the cozero set of f contains at most one element.

EXERCISE 6.3.4. Observe that if $\alpha>0$ then $n\mapsto 1/n^\alpha$ is in $\ell^p(\mathbb{N}^>)$ iff $p>1/\alpha$. Deduce from this that if there exist p,q with $0<p<q\le\infty$ such that $\ell^p(X)=\ell^q(X)$ then X is a finite set.

6.3.3. *Exercises on the variation of L^p with p.*

EXERCISE 6.3.5. Let $(X,\mathcal{M},\mu)$ be a measure space. Assume that $(A(n))_{n\in\mathbb{N}}$ is a disjoint sequence of sets in $\mathcal{M}$ of finite strictly positive measure. Let $0<p<q<\infty$.

(i) Assuming that for some subsequence $B(k)=A(n(k))$ we have $\lim_{k\to\infty}\mu(A(n(k))=0$, prove that there is $f\in L^p(\mu)\smallsetminus L^q(\mu)$ (hint: it is not restrictive to assume $\mu(B(k))\le 1/2^k$; set $b_k=\mu(B_k)$, and for $\alpha>0$ define the measurable function $g_\alpha: X\to\mathbb{R}$ by $g_\alpha=\sum_{k=0}^\infty b_k^{-\alpha}\,\chi_{B(k)}$. Given $0<p<q<\infty$, prove that if $1/q<\alpha<1/p$ we have $g_\alpha\in L^p(\mu)\smallsetminus L^q(\mu)$).

(ii) Prove that if for some subsequence $B_k=A(n(k))$ we have $b_k=\mu(B(k))\ge a>0$, then there is $f\in L^q(\mu)\smallsetminus L^p(\mu)$ (consider separately the cases of b_k bounded and $\lim_{k\to\infty}b_k=\infty$; in the second case we may assume $b_k\ge 2^k\ \dots$).

Solution. (i) We have, for $r>0$:

$$\int_X g_\alpha^r = \sum_{k=0}^\infty b_k^{-\alpha r}\,b_k = \sum_{k=0}^\infty b_k^{1-\alpha r};$$

Since $\lim_{k\to\infty}b_k=0$, if $1-\alpha r<0$ we have $\lim_{k\to\infty}b_k^{1-\alpha r}=\infty$ and the integral is infinite; if $1-\alpha r>0$ we have

$$b_k^{1-\alpha r}\le \frac{1}{2^{k(1-\alpha r)}}=\left(\frac{1}{2^{1-\alpha r}}\right)^k,$$

so that $\sum_{k=0}^\infty b_k^{1-\alpha r}<\infty$; the integral is finite.

(ii) Assume b_k bounded, then we have $a\le b_k\le b$ for every $k\in\mathbb{N}$. Given $\alpha>0$ we consider $g_\alpha=\sum_{k=0}^\infty \chi_{B(k)}/(k+1)^\alpha$; for $r>0$ we have

$$\int_X g_\alpha^r\,d\mu=\sum_{k=0}^\infty \frac{1}{(k+1)^{\alpha r}}\,b_k,\quad\text{so that}\quad \sum_{k=0}^\infty\frac{a}{(k+1)^{\alpha r}}\le\int_X g_\alpha^r\,d\mu\le\sum_{k=0}^\infty\frac{b}{(k+1)^{\alpha r}}$$

then $\int_X g_\alpha^r\,d\mu<\infty$ if and only if $\alpha\,r>1$; if $1/q<\alpha<1/p$ we have $g_\alpha\in L^q(\mu)$ but $g_\alpha\notin L^p(\mu)$.

If b_k has no bounded subsequence then $\lim_{k\to\infty}b_k=\infty$, and passing to a subsequence we may assume $b_k\ge 2^k$. Given $\alpha>0$ we consider $g_\alpha=\sum_{k=0}^\infty b_k^{-\alpha}\,\chi_{B(k)}$, as in (i); then again

$$\int_X g_\alpha^r=\sum_{k=0}^\infty b_k^{-\alpha r}\,b_k=\sum_{k=0}^\infty b_k^{1-\alpha r},$$

and we conclude as in (i), noting that now $b_k\ge 2^k$. □

EXERCISE 6.3.6. Let $(X,\mathcal{M},\mu)$ be a measure space. The following are equivalent:

(a) There exists a sequence $E_n \in \mathcal{M}$ with $\lim_{n\to\infty} \mu(E_n) = 0$ and $0 < \mu(E_n)$ for every n.
(b) There is a sequence $A_k \in \mathcal{M}$ with $0 < \mu(A_k) \le 1/2^k$ for every k.
(c) There is a function $f \in L^1(\mu) \smallsetminus L^\infty(\mu)$.
(d) There is a disjoint sequence $B_k \in \mathcal{M}$ with $0 < \mu(B_k) \le 1/2^k$ for every k.

((a) implies (b) easy; for (b) implies (c) prove that the formula $f(x) = \sum_{k=0}^\infty k\,\chi_{A_k}$ defines a.e. a function $f \in L^1(\mu) \smallsetminus L^\infty(\mu)$; for (c) implies (d) consider a suitable subsequence of the sequence $E_n = \{n < |f| \le n+1\}$, with $f \in L^1(\mu) \smallsetminus L^\infty(\mu) \ldots$). Other equivalences are:

(e) For every p, q with $0 < p < q$ there is $f \in L^p(\mu) \smallsetminus L^q(\mu)$.
(f) There exist p, q in $\mathbb{R}$ with $0 < p < q$ such that $L^p(\mu) \subsetneq L^q(\mu)$.

Solution. That (a) implies (b) is trivial: if a sequence of strictly positive numbers tends to 0, then there is a subsequence $(\mu(E_{n(k)}))_{k\in\mathbb{N}}$ such that $\mu(E_{n(k)}) \le 1/2^k$; simply set $A_k = E_{n(k)}$.

(b) implies (c) The series $\sum_{k=0}^\infty k\,\chi_{A_k}$ is a series of positive measurable functions, so that we have

$$\int_X f = \sum_{k=0}^\infty k\,\mu(A_k) \le \sum_{k=0}^\infty \frac{k}{2^k} < \infty.$$

Then $\{f = \infty\}$ has measure 0, and $f \in L^1(\mu)$ (to be more precise for the punctilious: f coincides a.e. with a function in $L^1(\mu)$, which we still call f). And $f \notin L^\infty(\mu)$: since all terms are positive, we have $f \ge k\,\chi_{A_k}$, so that $\{f \ge k\} \supseteq A_k$, hence $\mu(\{f \ge k\}) \ge \mu(A_k) > 0$, for every $k \in \mathbb{N}$, consequently $\|f\|_\infty = \infty$.

(c) implies (d) Since $f \notin L^\infty(\mu)$, infinitely many E_n have strictly positive measure. Moreover $\lim_{n\to\infty} \mu(E_n) = 0$, since by Čebičeff's inequality we have $\mu(E_n) \le (1/n)\|f\|_1$; and the E_n are pairwise disjoint. Some subsequence $B_k = E_{n(k)}$ will then be such that $\mu(B_k) \le 1/2^k$.

That (d) implies (a) is trivial.

For the other equivalences: (d) implies (e) is exercise 6.3.5; (e) implies (f) is trivial. Finally, (f) implies (c): if $f \in L^p(\mu) \smallsetminus L^q(\mu)$ then $|f|^p \in L^1(\mu)$ by definition, but $|f|^p \notin L^\infty(\mu)$ (by Proposition 6.3, $L^1(\mu) \cap L^\infty(\mu) \subseteq L^q(\mu)$ for every $q > p$).

□

Exercise 6.3.7. (uses the previous exercises 6.3.5 and 6.3.6) Let $(X, \mathcal{M}, \mu)$ be a measure space. Prove that the following are equivalent:

(i) There exist p, q, with $0 < p < q < \infty$, such that $L^p(\mu) \supseteq L^q(\mu)$.
(ii) For every disjoint family $E(n) \in \mathcal{M}$ of sets of finite nonzero measure we have $\lim_{n\to\infty} \mu(E(n)) = 0$.
(iii) For every disjoint sequence $E(n) \in \mathcal{M}$ of sets of finite nonzero measure we have that $\sup\{\mu(E(n) : n \in \mathbb{N}\} < \infty$.
(iv) There are disjoint sets $A, B \in \mathcal{M}$, with A of finite measure, and B either empty or an atom of infinite measure, such that $X = A \cup B$.
(v) If $0 < p < q < \infty$ we have $L^p(\mu) \supseteq L^q(\mu)$.

And if (ii) holds, then $L^\infty(\mu) \subseteq L^p(\mu)$ for every $p > 0$ if and only if B is empty.

Solution. (i) implies (ii): exercise 6.3.5; (ii) implies (iii) is trivial. (iii) implies (iv) Let $s = \sup\{\mu(E) : E \in \mathcal{M},\ \mu(E) < \infty\}$; we claim that s is finite, and in fact $s = \max\{\mu(E) : E \in \mathcal{M},\ \mu(E) < \infty\}$: picking a sequence $E(n)$ of sets of finite measure such that $\lim_{n\to\infty} \mu(E(n)) = s$, we consider $F(n) = \bigcup_{k=0}^n E(k)$; then $F(n)$ is increasing, and $s \ge \mu(F(n)) \ge \mu(E(n))$ so that, setting $F = \bigcup_{n=0}^\infty F(n)$ we have

$$s \ge \mu(F) = \lim_{n\to\infty} \mu(F(n)) \ge \lim_{n\to\infty} \mu(E(n)) = s,$$

i.e., $\mu(F) = s$. If $s = \infty$ then for some subsequence $F(n(k))$ of $F(n)$ we have $\lim_{k\to\infty} \mu(F(n(k+1)) \smallsetminus F(n(k))) = \infty$, contradicting hypothesis (i). Then s is finite, and F has maximum measure among the sets of finite measure. Consequently $X \smallsetminus F$ cannot contain sets of finite strictly positive measure: if $\mu(X \smallsetminus F) = 0$ we set $A = X$ and $B = \emptyset$; if $\mu(X \smallsetminus F) > 0$, then $X \smallsetminus F$ is an atom of infinite measure; we set $A = F$ and $B = X \smallsetminus F$.

(iv) implies (v). On an atom of infinite measure every summable function is almost everywhere 0; then $L^p_\mu(X, \mathbb{C})$ can be identified with $L^p_\mu(A, \mathbb{C})$; then recall 6.3.1. And (v) trivially implies (i) □

Exercise 6.3.8. Let $(X, \mathcal{M}, \mu)$ be a measure space. The following are equivalent:

(i) There are $p, q \in \mathbb{R}$, with $0 < p < q$ such that $L^p(\mu) \subseteq L^q(\mu)$.

(ii) For every disjoint sequence $E(n) \in \mathcal{M}$ of sets of finite nonzero measure we have that $\inf\{\mu(E(n) : n \in \mathbb{N}\} = a > 0$.
(iii) $\inf\{\mu(E) : E \in \mathcal{M},\ \mu(E) > 0\} = \alpha > 0$.
(iv) Every $E \in \mathcal{M}$ of finite nonzero measure is a finite disjoint union of atoms.
(v) For every real number $p > 0$ we have $L^p(\mu) \subseteq L^\infty(\mu)$.
(vi) For every pair of real numbers p, q with $0 < p < q$ we have $L^p(\mu) \subseteq L^q(\mu)$.

Solution. (i) implies (ii): see 6.3.5; (ii) implies (iii): see 6.3.6; (iii) is equivalent to (iv) was proved in 2.4.3.

(iii), (iv) imply (v): if f is in $L^p(\mu)$ then $\mathrm{Coz}(f) = \{f \neq 0\}$ has $\sigma-$finite measure; by (iv) every set $E \in \mathcal{M}$ of $\sigma-$finite measure is a countable disjoint union of atoms of finite measure, so that $\mathrm{Coz}(f) = \bigcup_{n=0}^\infty A(n)$, where each $A(n)$ is an atom of finite measure, and the $A(n)$ are pairwise disjoint, so that we can write $f = \sum_{n=0}^\infty y_n\,\chi_n$, where $\chi_n = \chi_{A(n)}$ and $y_n \in \mathbb{K}$. By (iii) the measures of these atoms are bounded way from zero, $a_n = \mu(A(n)) \geq \alpha > 0$ for every n. Then, for every $m \in \mathbb{N}$:

$$\|f\|_p^p = \sum_{n=0}^\infty |y_n|^p\,\mu(A(n)) \geq \sum_{n=0}^\infty |y_n|^p\,\alpha \geq |y_m|^p\,\alpha$$

which implies $|y_m| \leq \|f\|_p/\alpha^{1/p}$ for every m, hence

$$\|f\|_\infty \leq \frac{\|f\|_p}{\alpha^{1/p}}.$$

(v) implies (vi): Proposition 6.3; (iv) implies (i) trivially. □

6.4. Convexity and L^p spaces.

6.4.1. *Convexity combinations of arbitrary length.* In a real vector space V, given two vectors a, b the set $\{\alpha\,a + \beta\,b : \alpha, \beta \geq 0,\ \alpha + \beta = 1\} = \{(1-t)a + t\,b : t \in [0,1]\}$ of all convex combinations of a, b is the *convex hull* of the set $\{a, b\}$, the smallest convex subset of V containing the two points a, b, namely the segment of extremes a, b. Given a set of n vectors $\{a_1, \dots, a_n\} \subseteq V$ a *convex combination* of these vectors is

$$\sum_{k=1}^n \alpha_k\,a_k, \quad \text{provided that} \quad \alpha_k \geq 0, \sum_{k=1}^n \alpha_k = 1.$$

By induction on the cardinality of the subset one proves easily:

. *A convex set contains all convex combinations of its finite subsets.*

Proof. For singletons it is trivial. Assuming that all convex combinations of subsets with n elements of the convex set C are in C, let $\{a_1, \dots, a_n,\ a_{n+1}\} \subseteq C$, and let $\alpha_k > 0$ be such that $\sum_{k=1}^{n+1} \alpha_k = 1$; we want to prove that $\sum_{k=1}^{n+1} \alpha_k\,a_k \in C$. In fact, setting $\alpha = \sum_{k=1}^n \alpha_k$, and $a = \sum_{k=1}^n (\alpha_k/\alpha)\,a_k$, a convex combination of the first n vectors:

$$\sum_{k=1}^{n+1} \alpha_k\,a_k = \sum_{k=1}^{n} \alpha_k\,a_k + \alpha_{n+1}\,a_{n+1} = \alpha \sum_{k=1}^{m} \frac{\alpha_k}{\alpha}\,a_k + \alpha_{n+1}\,a_{n+1} = \alpha\,a + \alpha_{n+1}\,a_{n+1}.$$

By the inductive hypothesis $a \in C$; then $\alpha\,a + \alpha_{n+1}\,a_{n+1} \in C$ as the convex combinations of two vectors, a and a_{n+1}, belonging to C. □

Corollary. *Let I be an interval of $\mathbb{R}$, and let $f : I \to \mathbb{R}$ be convex. Then, for every convex combination $\sum_{k=1}^n \alpha_k\,x_k$ of a finite subset $\{x_1, \dots, x_n\}$ of I we have*

$$f(\alpha_1\,x_1 + \cdots + \alpha_n\,x_n) \leq \alpha_1\,f(x_1) + \cdots + \alpha_n\,f(x_n) \quad \text{or} \quad f\left(\sum_{k=1}^n \alpha_k\,x_k\right) \leq \sum_{k=1}^n \alpha_k\,f(x_k)$$

(with strict inequality if f is strictly convex, all α_k are strictly positive, and the x_k are distinct).

Proof. $\{(x_k, f(x_k)) : k = 1, \dots, n\}$ is a finite subset of the convex set $\mathrm{Epi}(f) = \{(x, y) : x \in I,\ y \geq f(x)\}$, which contains every convex combination of its finite subsets. □

Exercise 6.4.1. Prove

. Jensen's inequality for positive functions *Let $(X, \mathcal{M}, \mu)$ be a probability space ($\mu(X) = 1$), and let $\omega : [0, \infty[\to [0, \infty[$ be continuous, convex and increasing. Then for every $f \in \mathcal{L}^+$ we have*

$$\omega\left(\int_X f\right) \le \int_X \omega \circ f.$$

Hint: pick a sequence φ_n of positive simple functions such that $\varphi_n \uparrow f$ ...

Solution. If $\varphi_n = \sum_{k=1}^{m(n)} y_k \,\chi_{E(k)}$ is the standard representation of φ_n, a representation of $\omega \circ \varphi_n$ is $\omega \circ \varphi_n = \sum_{k=1}^{m(n)} \omega(y_k)\,\chi_{E(k)}$; observe that

$$\int_X \varphi_n = \sum_{k=1}^{m(n)} y_k\,\mu(E(k)) \quad \text{and} \quad \int_X \omega \circ \varphi_n = \sum_{k=1}^{m(n)} \omega(y_k)\,\mu(E(k)),$$

are convex combinations, with $\mu(E(k))$, $k = 1, \dots, m(n)$ as coefficients (in fact $X = \bigcup_{k=1}^{m(n)} E(k)$, disjoint union, so that $1 = \sum_{k=1}^{m(n)} \mu(E(k))$), and convexity of ω says that

$$\omega\left(\sum_{k=1}^{m(n)} y_k\,\mu(E(k))\right) \le \sum_{k=1}^{m(n)} \omega(y_k)\,\mu(E(k)),$$

in other words

$$\omega\left(\int_X \varphi_n\right) \le \int_X \omega \circ \varphi_n \quad \text{for every } n \in \mathbb{N}.$$

Since ω is increasing $\omega \circ \varphi_n$ is also increasing and by continuity of ω this sequence converges pointwise to $\omega \circ f$, i.e $\omega \circ \varphi_n \uparrow \omega \circ f$. By monotone convergence the right–hand side has $\int_X \omega \circ f$ as limit, while by continuity of ω the left–hand side tends to $\omega\left(\int_X f\right)$. $\square$

Exercise 6.4.2. In a probability space $(X, \mathcal{M}, \mu)$ we have $\|f\|_p \le \|f\|_q$ for $p < q$ and every measurable $f : X \to \mathbb{K}$ (6.3.1). Prove it by Jensen inequality; then get from this the inequality $\|f\|_p \le \mu(X)^{1/p - 1/q}\,\|f\|_q$ valid when $\mu(X)$ is finite but not necessarily 1.

Exercise 6.4.3. (*A generalized Hölder's inequality*) Let $(p(k)_{1 \le k \le n}$ be an $n-$tuple of positive numbers such that $1 = \sum_{k=1}^m 1/p(k)$. Prove that for every $n-$tuple $a_1, \dots, a_n$ of positive numbers we have

$$(*) \qquad \prod_{k=1}^n a_k \le \sum_{k=1}^n \frac{a_k^{p(k)}}{p(k)},$$

and that for every $n-$tuple $f_1, \dots, f_n$ of measurable functions we have, denoting by $f = \prod_{k=1}^n f_k$ their product:

$$\|f\|_1 \le \prod_{k=1}^n \|f_k\|_{p(k)}.$$

Assume now that $\sum_{k=1}^m 1/p(k) \le 1$, and that $1/p = \sum_{k=1}^m 1/p(k)$. Then, with f_k and f as above:

$$\|f\|_p \le \prod_{k=1}^n \|f_k\|_{p(k)}.$$

Solution. Taking logarithms of both sides the given numerical inequality is equivalent to

$$\sum_{k=1}^n \log a_k \le \log\left(\sum_{k=1}^n \frac{a_k^{p(k)}}{p(k)}\right),$$

which, setting $x_k = a_k^{p(k)}$, equivalently $a_k = x_k^{1/p(k)}$, becomes

$$\sum_{k=1}^n \frac{\log x_k}{p(k)} \le \log\left(\sum_{k=1}^n \frac{x_k}{p(k)}\right),$$

which, since $1 = \sum_{k=1}^n 1/p(k)$, is consequence of the concavity of the logarithm function, and hence true. Then we simply put in inequality (*) $a_k = |f_k(x)|/\|f_k\|_{p(k)}$, and integrate both sides.

For the second part: we have $1 = \sum_{k=1}^{n} 1/(p(k)/p)$; apply the first part with $p(k)/p$ in place of $p(k)$ and with f_k replaced by $|f_k|^p$ to get:

$$\||f|^p\|_1 \le \prod_{k=1}^{n} \||f_k|^p\|_{p(k)/p} = \prod_{k=1}^{n} \|f_k\|_{p(k)}^p,$$

and raising both sides to the power $1/p$ we conclude. □

6.4.2. *Subadditivity, concavity, convexity.* As observed in 1.8.8 a concave function $q : [0, \infty[\to \mathbb{R}$ which is 0 at 0 is subadditive: given $a, b > 0$ we get a, b as convex combinations of 0 and $a + b$, that is $a = (a/(a+b))\,(a+b) + (1 - a/(a+b))\,0$ so that $q(a) \ge (a/(a+b))\,q(a+b)$ and analogously $q(b) \ge (b/(a+b))\,q(a+b)$ so that, adding these inequalities we get $q(a) + q(b) \ge q(a+b)$, with strict inequality if a, $b > 0$ and q is strictly concave. If $0 < p < 1$ the function $q(x) = x^p$ is strictly concave (in fact $q''(x) = p(p-1)\,x^{p-2} < 0$ if $x > 0$), so that we have $(a+b)^p < a^p + b^p$ if $a, b > 0$.

Conversely a convex function $\rho : [0, \infty[\to \mathbb{R}$ such that $\rho(0) = 0$ is *superadditive*, that is $\rho(a+b) \ge \rho(a) + \rho(b)$ if a, $b > 0$, with strict inequality if ρ is strictly convex. This shows why $\|\#\|_p$ is not subadditive for $0 < p < 1$: if E, F are disjoint sets of finite strictly positive measure then

$$\|\chi_E + \chi_F\|_p = \left(\int_X \chi_E + \int_X \chi_F\right)^{1/p} = (\mu(E) + \mu(F))^{1/p} > \mu(E)^{1/p} + \mu(F)^{1/p} = \|\chi_E\|_p + \|\chi_F\|_p;$$

at the same time subadditivity of $x \mapsto x^p$ for $0 < p < 1$ proves that the map $f \mapsto [f]_p = \int_X |f|^p = \|f\|_p^p$ is subadditive. Then this map defines a *metric* on $L^p(\mu)$, by $\mathrm{dist}(f, g) = [f - g]_p$, translation invariant (i.e. $\mathrm{dist}(f + h, g + h) = \mathrm{dist}(f, g)$ for any $f, g, h \in L^p$) (but we have $[\alpha\, f]_p = |\alpha|^p\,[f]_p$, so that $[f]_p$ is not a norm) this metric makes L^p a metrizable complete topological vector space, completeness is still the theorem on normal convergence. But these spaces have a limited interest in analysis.

We have seen that if $p > 1$ and $a, b > 0$ then $(a+b)^p \le 2^p(a^p + b^p)$; but a better constant for it is 2^{p-1}, as we see in the following exercise.

EXERCISE 6.4.4. Given $a_1, \dots, a_n > 0$ we have

$$\begin{aligned} &\text{If} \quad p > 1: && a_1^p + \cdots + a_n^p < (a_1 + \cdots + a_n)^p < n^{p-1}(a_1^p + \cdots + a_n^p)\\ &\text{if} \quad 0 < p < 1: && n^{p-1}(a_1^p + \cdots + a_n^p) < (a_1 + \cdots + a_n)^p < a_1^p + \cdots + a_n^p. \end{aligned}$$

Solution. Superadditivity of the convex function $x \mapsto x^p$ has been proved above. And we have

$$(a_1 + \cdots + a_n)^p = n^p \left(\frac{a_1}{n} + \cdots + \frac{a_n}{n}\right)^p \le n^p \left(\frac{a_1^p}{n} + \cdots + \frac{a_n^p}{n}\right) = n^{p-1}\,(a_1^p + \cdots + a_n^p),$$

the inequality due to the convexity of x^p. □

6.4.3. *Some problems.*

EXERCISE 6.4.5. Let $(X, \mathcal{M}, \mu)$ be a finite measure space. Let f be a measurable non zero function belonging to L^p for every $p < \infty$. Prove that

$$\lim_{p \to \infty} \frac{\|f\|_{p+1}^{p+1}}{\|f\|_p^p} = \|f\|_\infty.$$

Solution. We have $|f|^{p+1} = |f|\,|f|^p \le \|f\|_\infty\,|f|^p$; integrating both sides we get $\|f\|_{p+1}^{p+1} \le \|f\|_\infty\,\|f\|_p^p$ so that

$$\frac{\|f\|_{p+1}^{p+1}}{\|f\|_p^p} \le \|f\|_\infty \Longrightarrow \limsup_{p \to \infty} \frac{\|f\|_{p+1}^{p+1}}{\|f\|_p^p} \le \|f\|_\infty.$$

Recall that if $p < q$ then $\|f\|_p \le \mu(X)^{1/p - 1/q}\|f\|_q$; applying this formula with $q = p + 1$ we get:

$$\|f\|_{p+1} \ge \frac{\|f\|_p}{\mu(X)^{1/p - 1/(p+1)}} \Longrightarrow \|f\|_{p+1}^{p+1} \ge \frac{\|f\|_p^{p+1}}{\mu(X)^{1/p}},$$

so that

$$\frac{\|f\|_{p+1}^{p+1}}{\|f\|_p^p} \ge \frac{\|f\|_p}{\mu(X)^{1/p}};$$

the right–hand side tends to $\|f\|_\infty$ as p tends to ∞ so that the liminf of the left–hand side is larger than $\|f\|_\infty$. □

Exercise 6.4.6. Let $f : [0,\infty[\to \mathbb{R}$ be defined by $f(x) = x\,e^{-x}$. Compute $\|f\|_p$ for every $p > 0$. Find the limit

$$\lim_{n\to\infty} \frac{(n!)^{1/n}}{n}.$$

Solution. Since f is positive, continuous, and $0 = f(0) = f(\infty)$ we have $\|f\|_\infty = \max f([0,\infty])$; now $f'(x) = (1-x)\,e^{-x}$ is 0 for $x = 1$, so that $\|f\|_\infty = f(1) = 1/e$. And:

$$\int_0^\infty |f(x)|^p\,dx = \int_0^\infty x^p\,e^{-px}\,dx = \int_0^\infty \frac{t^p}{p^p}\,e^{-t}\,\frac{dt}{p} = \frac{\Gamma(p+1)}{p^{p+1}};$$

then

$$\|f\|_p = \frac{(\Gamma(p+1))^{1/p}}{p\,p^{1/p}};\text{ since}\quad \lim_{p\to\infty}\|f\|_p = \|f\|_\infty = \frac{1}{e},$$

and since $\lim_{p\to\infty} p^{1/p} = 1$, we get

$$\lim_{p\to\infty}\frac{(\Gamma(p+1))^{1/p}}{p} = \frac{1}{e} \Longrightarrow \lim_{n\to\infty}\frac{(n!)^{1/n}}{n} = \frac{1}{e}.$$

$\square$

Exercise 6.4.7. Consider $f_\alpha :]0,\infty[\to \mathbb{R}$ defined by $f_\alpha(x) = (1/(x(1+\log^2 x)))^\alpha$; $\alpha > 0$ is constant.
(i) Find $\{p > 0 : f_\alpha \in L^p([0,1])\}$.
(ii) Find $\{p > 0 :\ f_\alpha \in L^p([1,\infty[)\}$.
(iii) Find $\{p > 0 :\ f_\alpha \in L^p([0,\infty[)\}$.

Do the same for $g(x) = 1/x^\alpha$. Given $0 < p < r < \infty$ find functions $f, g \in L([0,\infty[)$ such that $f \in L^q([0,\infty[)$ iff $q \in [p,r]$ and iff $q \in]p,r[$, respectively.

(iv) Given $\alpha,\beta > 0$ let $f(= f_{\alpha,\beta}) : \mathbb{R}^n \to \mathbb{R}$ be defined by

$$f(x) = \frac{1}{(1+|x|)^\alpha\,(1+|\log|x||^\beta)}\qquad |x| = \left(\sum_{k=1}^n x_k^2\right)^{1/2}.$$

Given $p > 0$ find $\alpha, \beta > 0$ such that $f \in L^p(\mathbb{R}^n)$ (use integration on spheres, Analisi Due 9.27).

Exercise 6.4.8. Let $(X,\mathcal{M},\mu)$ be a measure space, and let $1 \le p < \infty$.
(i) Prove the dominated convergence theorem for L^p:

. *If f_n, f, g are measurable functions, $f_n \to f$ a.e., $|f_n| \le g$ for every $n \in \mathbb{N}$, and $g \in L^p$ then $f_n, f \in L^p$ and f_n converges to f in L^p.*

(ii) Generalize the dominated convergence theorem for L^p: if f_n, f, g_n, g are measurable, $f_n \to f$ a.e., $g_n \to g$ a.e., $|f_n| \le g_n$, and $g_n, g \in L^p$ are such that $\int_X g_n^p \to \int_X g^p$, then f_n converges to f in L^p.
(iii) If $f_n \to f$ a.e., and f_n, $f \in L^p$, then f_n converges to f in L^p if and only if $\|f_n\|_p \to \|f\|_p$.

Solution. (i) Since $|f_n|^p \le g^p \in L^1$, clearly also $|f|^p \le g^p \in L^1$, so that f_n, $f \in L^p$. And we have $|f - f_n|^p \le (|f| + |f_n|)^p \le 2^{p-1}(|f|^p + |f_n|^p) \le 2^{p-1}(g^p + g^p) \le 2^p\,g^p \in L^1$; clearly $|f - f_n|^p \to 0$ a.e.. The dominated convergence theorem says that $\int_X |f - f_n|^p \to 0$, that is $\|f - f_n\|_p^p \to 0$.

(ii) We have, again:

$$|f - f_n|^p \le (|f| + |f_n|)^p \le 2^{p-1}(|f|^p + |f_n|^p) \le 2^{p-1}(g^p + g_n^p),$$

and we can apply the generalized dominated convergence theorem to conclude that $\int_X |f - f_n|^p \to 0$.

(iii) If f_n converges to f in L^p, then $\|f_n\|_p \to \|f\|_p$, as in any normed space: recall that by the triangle inequality we get $|\|f\|_p - \|f_n\|_p| \le \|f - f_n\|_p$. The converse is the preceding result, with $|f_n|$ as g_n, $|f|$ as g. $\square$

Exercise 6.4.9. Let $(X,\mathcal{M},\mu)$ be measure space. Let $1 \le p < \infty$ and $1 \le r \le \infty$.
(i) Prove that $L^1 \cap L^\infty$ is dense in L^p.
(ii) Prove that the set $\{f \in L^p \cap L^r : \|f\|_r \le 1\}$ is closed in L^p (hint: Fatou's lemma).
(iii) Let f_n be a sequence in $L^p \cap L^r$, and let $f \in L^r$. Assume that f_n converges to f in L^p and that there is $a > 0$ such that for every n we have $\|f_n\|_r \le a$. Prove that $f \in L^q$, and that f_n converges to f in L^q for every $q \ne r$ in the closed interval of extremes p, r.

Solution. (i) $L^1 \cap L^\infty$ contains $S(\mu)$, space of integrable simple functions, which is dense in every L^p with $p < \infty$.

(ii) Assume that f_n is a sequence in the given set, converging to f in L^p. We have to prove that f is still in the set, i.e. that $\|f\|_r \leq 1$. There is a subsequence that converges to f also a.e., and we may as well assume that f_n itself converges a.e to f. If $r < \infty$ then, by Fatou's lemma:

$$\int_X |f|^r \leq \liminf_{n\to\infty} \int_X |f_n|^r \leq 1;$$

if $r = \infty$ then $|f_n(x)| \leq 1$ for a.e. $x \in X$ clearly implies $|f(x)| \leq 1$ for a.e $x \in X$, that is, $\|f\|_\infty \leq 1$.

(iii) If $q \in [p, r[$ there is α, with $0 < \alpha \leq 1$, such that $1/q = \alpha/p + (1-\alpha)/r$. The interpolation inequality then says that for every $n \in \mathbb{N}$:

$$\|f - f_n\|_q \leq \|f - f_n\|_p^\alpha \, \|f - f_n\|_r^{1-\alpha}.$$

Moreover from (ii), with a in place of 1, we know that $\|f\|_r \leq a$, so that $\|f - f_n\|_r \leq \|f\|_r + \|f_n\|_r \leq 2a$, so that

$$\|f - f_n\|_q \leq (2a)^{1-\alpha} \, \|f - f_n\|_p^\alpha;$$

the proof is completed. □

Exercise 6.4.10. (*A more general Young's inequality*) Let $\varphi : [0, \infty[\to [0, \infty[$ be a self–homeomorphism of $[0, \infty[$.

(i) Prove that φ is strictly increasing, that $\varphi(0) = 0$ and $\lim_{x\to\infty} \varphi(x) = \infty$.

Given now$a, b > 0$ we consider subsets $S(a)$ and $T(b)$ $E(a, b)$ of $\mathbb{R}^2$ so obtained:
$S(a) = \{(x, y) \in \mathbb{R}^2 : x \in [0, a],\ 0 \leq y \leq \varphi(x)\}$, trapezoid of φ over $[0, a]$.
$T(b) = \{(x, y) \in \mathbb{R}^2 : y \in [0, b],\ 0 \leq x \leq \varphi^{-1}(y)\}$, trapezoid of φ^{-1} over $[0, b]$.

(ii) Prove that

$$\lambda_2(S(a)) = \int_0^a \varphi(x)\, dx, \quad \lambda_2(T(b)) = \int_0^b \varphi^{-1}(y)\, dy,$$

and that, if $E(a, b) = S(a) \cup T(b)$ then

$$\lambda_2(E(a, b)) = \lambda_2(S(a)) + \lambda_2(T(b)) = \int_0^a \varphi(x)\, dx + \int_0^b \varphi^{-1}(y)\, dy.$$

(iii) Let $R(a, b) = [0, a] \times [0, b]$ be the rectangle; prove that $R(a, b) \subseteq E(a, b)$, and that unless $b = \varphi(a)$ then

$$\lambda_2(R(a, b)) = ab < \lambda_2(E(a, b)) \quad \text{that is} \quad ab < \int_0^a \varphi(x)\, dx + \int_0^b \varphi^{-1}(y)\, dy.$$

(iv) Let $p > 1$; using $\varphi(x) = x^{p-1}$ prove that $ab < a^p/p + b^q/q$ if $1/p + 1/q = 1$ and $a^p \neq b^q$.

Solution. Schematic: (i) easy (an injective continuous function on an interval is strictly monotone; since 0 is the minimum of $[0, \infty[$ and $0 = \varphi(a)$ for some $a \geq 0$, φ cannot be decreasing, otherwise $\varphi(c) < 0$ if $c > a$, impossible. Then φ is increasing; by surjectivity $\varphi(0) = 0$ and $\lim_{x\to\infty} \varphi(x) = \infty$).

(ii) Fubini's theorem; note that $S(a) \cap T(b) \subseteq \mathrm{Graph}(\varphi)$ and that $\lambda_2(\mathrm{Graph}(\varphi)) = 0$, so that $\lambda_2(E(a, b)) = \lambda_2(S(a)) + \lambda_2(T(b))$.

(iii) Look at the figures; if $b < \varphi(a)$ then $E(a, b) \setminus R(a, b) = \{(x, y) : \varphi^{-1}(b) < x \leq a,\ b < y \leq \varphi(x)\}$, a set with area $\int_{\varphi^{-1}(b)}^a (\varphi(x) - b)\, dx > 0$. If $b > \varphi(a)$ then $E(a, b) \setminus R(a, b) = \{(x, y) : \varphi(a) < y \leq b,\ a < x \leq \varphi^{-1}(y)\}$, a set with area $\int_{\varphi(a)}^b (\varphi^{-1}(y) - a)\, dy > 0$. If $b = \varphi(a)$ then $E(a, b) = R(a, b)$.

(iv) Compute. □

Exercise 6.4.11. Let $(X, \mathcal{M}, \mu)$ be a measure space.

(i) Let $f, g \in L^1(\mu)$. Prove that $\|f + g\|_1 = \|f\|_1 + \|g\|_1$ if and only if $f = 0$, or there exists $\lambda : X \to [0, +\infty[$ measurable and such that $g(x) = \lambda(x)\, f(x)$ for a.e. $x \in \mathrm{Coz}(f)$.

(ii) Let $p \in]1, \infty[$ and assume that $f, g \in L^p(\mu)$. Prove that $\|f + g\|_p = \|f\|_p + \|g\|_p$ if and only if $f = 0$, or there exists a constant $\lambda \geq 0$ such that $g(x) = \lambda\, f(x)$ for a.e. $x \in X$ (in other words, L^p spaces are *strictly convex* if $1 < p < \infty$).

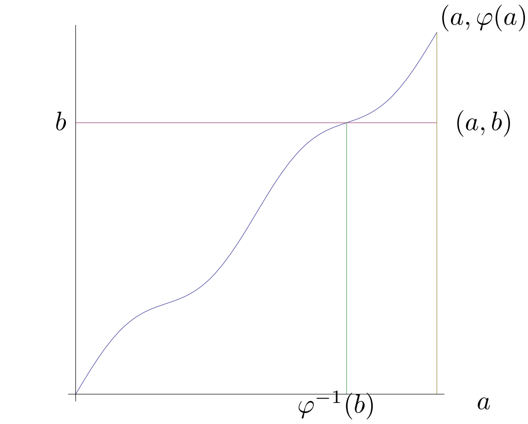

FIGURE 3. Case $b < \varphi(a)$.

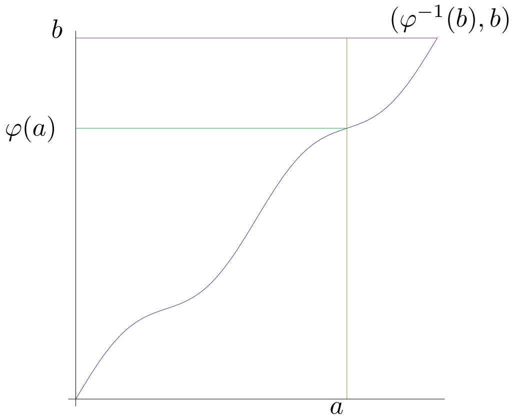

FIGURE 4. Case $b > \varphi(a)$.

Solution. Recall first that the absolute value on $\mathbb{K}$ is a strictly convex norm: that is, we have $|a+b| = |a|+|b|$ for $a, b \in \mathbb{K}$ iff $a = 0$ or $b = \lambda\, a$ with $\lambda \geq 0$.

(i) The inequality $\|f+g\|_1 \leq \|f\|_1+\|g\|_1$ is obtained by integrating the inequality $|f(x)+g(x)| \leq |f(x)|+|g(x)|$, and the resulting inequality between integrals is an equality if and only if $|f(x)+g(x)| = |f(x)|+|g(x)|$ for a.e. $x \in X$; in turn this happens iff $f(x) = 0$ or $g(x) = \lambda(x)\, f(x)$, with $\lambda(x) \geq 0$, for a.e. $x \in X$. If $f(x) = 0$ a.e., then clearly we have equality; if not, setting $\lambda(x) = g(x)/f(x)$ for every $x \in \mathrm{Coz}(f)$ we have equality iff $\lambda(x) \geq 0$ for almost all $x \in \mathrm{Coz}(f)$.

(ii) The conditions are clearly sufficient for equality, as in every normed space. Assume then f, g both nonzero; we also assume $f+g$ nonzero (otherwise the inequality is assuredly strict) Repeating the proof of Minkowski inequality we see that it is satisfied as an equality iff

(a) We have

$$|f(x)+g(x)|\,|f(x)+g(x)|^{p-1} = |f(x)|\,|f(x)+g(x)|^{p-1} + |g(x)|\,|f(x)+g(x)|^{p-1} \quad \text{for almost every } x \in X,$$

equivalently $|f(x)+g(x)| = |f(x)|+|g(x)|$ for a.e. $x \in \mathrm{Coz}(f+g)$;

(b) Hölder's inequality with functions $|f|$, $|f+g|^{p-1}$ and $|g|$, $|f+g|^{p-1}$ are verified as equalities.

This last condition, since $f+g \neq 0$, holds iff $|f|^p = a\,|f+g|^p$ and $|g|^p = b\,|f+g|^p$ for some constants $a, b > 0$ (6.1.1). Then $|g|^p = k\,|f|^p$ for some $k > 0$, so that $|g(x)| = c\,|f(x)|$ for some $c > 0$; on $\mathrm{Coz}(f+g)$ we must, by (a), have $g(x)/f(x) = \lambda(x) \geq 0$, so that $c = |\lambda(x)| = \lambda(x)$ for a.e. $x \in \mathrm{Coz}(f)$. □

6.4.4. *Jensen inequality.*

. *Let $(X, \mathcal{M}, \mu)$ be a probability space (i.e. $\mu(X) = 1$). Let I be an open interval of $\mathbb{R}$, and let $\omega : I \to \mathbb{R}$ be a convex function. Let $f : X \to I$ belong to $L^1(\mu)$. Then $\int_X f \in I$, the function $\omega \circ f$ is integrable in the extended sense, and we have the*

JENSEN INEQUALITY
$$\omega\left(\int_X f\right) \leq \int_X \omega \circ f.$$

If ω is strictly convex and f is non–constant, then the inequality is strict.

Proof. Let $c = \int_X f$; then $c \in I$; in fact, by the hypothesis $f \in L^1(\mu)$ we have that $c \in \mathbb{R}$, and if $\inf I = a$ is finite then we have $a < f(x)$ for every $x \in X$, so that integrating we get $a = \int_X a < \int_X f = c$, i.e. $a < c$; similarly we get $c < b$ if $b = \sup I < \infty$. Since c is in the interior of I and ω is convex there is a real m such that $\omega(t) \geq \omega(c) + m\,(t - c)$ for every $t \in I$ (simply take any m between the left and right derivatives of ω at c). Then for every $x \in X$ we have

$$\omega(f(x)) \geq \omega(c) + m\,(f(x) - c);$$

the function on the right–hand side of this inequality is in $L^1(\mu)$; then the function on the left–hand side, which is measurable because ω is continuous, is integrable in the extended sense (4.4.1) and

$$\int_X \omega \circ f \geq \omega(c) + m\left(\int_X f - c\right) = \omega(c) + m\,(c - c) = \omega\left(\int_X f\right).$$

If ω is strictly convex we have $\omega(t) > \omega(c) + m\,(t - c)$ for every $t \in I \setminus \{c\}$, so that integrating we get a strict inequality, unless $f(x) = c$ for a.e. $x \in X$. □

Remark. It may very well happen that $\omega \circ f$ has ∞ as integral. If $X = [0, 1]$ with Lebesgue measure, $I = \mathbb{R}$, $\omega(t) = t^2$ and $f(x) = 1/\sqrt{x}$, then $\int_X f = 2$, but $\omega \circ f(x) = 1/x$ has infinite integral over X.

Of course we also have

. Let $(X, \mathcal{M}, \mu)$ be a probability space (i.e. $\mu(X) = 1$). Let I be an open interval of $\mathbb{R}$, and let $\phi : I \to \mathbb{R}$ be a concave function. Let $f : X \to I$ belong to $L^1(\mu)$. Then $\int_X f \in I$, the function $\phi \circ f$ is integrable in the extended sense, and we have the

Jensen inequality for concave functions
$$\phi\left(\int_X f\right) \geq \int_X \phi \circ f.$$

If ϕ is strictly convex and f is non–constant, then the inequality is strict.

Proof. Apply Jensen inequality to the convex function $\omega = -\phi$. □

Exercise 6.4.12. Recall that on a finite measure space $(X, \mathcal{M}, \mu)$ for every measurable f we have $\|f\|_\infty = \lim_{p\to\infty} \|f\|_p$. Given a real valued measurable $f : X \to \mathbb{R}$, compute:

$$\lim_{p\to\infty} \frac{1}{p} \log\left(\int_X e^{p\,f(x)}\,d\mu(x)\right); \quad \lim_{p\to\infty} \frac{1}{p} \log\left(\int_X \cosh(p\,f(x))\,d\mu(x)\right)$$

(the second may be a little difficult ...).

Exercise 6.4.13. Let $(X, \mathcal{M}, \mu)$ be a measure space. Given a measurable not a.e. zero $f \in L(X)$ we consider the function $\varphi :]0, \infty] \to]0, \infty]$ given by

$$\varphi(p) = \int_X |f|^p = \|f\|_p^p$$

Prove that this function is finite–valued on an interval I of $]0, \infty]$ (which may be empty or reduced to a single point) and that in this interval $\log\varphi$ is a convex function.

Solution. (*of Exercise 6.4.2*) If $q > p$ the function $\omega(t) = t^{q/p}$ is continuous, positive, increasing and strictly convex on $[0, \infty[$; given a measurable $f : X \to \mathbb{K}$ we apply Jensen inequality for positive functions (6.4.1) to $|f|^p$, obtaining

$$\omega\left(\int_X |f|^p\right) \leq \int_X \omega \circ |f|^p \quad \text{that is} \quad \left(\int_X |f|^p\right)^{q/p} \leq \int_X |f|^q \iff \|f\|_p \leq \|f\|_q.$$

This inequality can be applied to finite measure spaces with the measure $\nu = \mu/\mu(X)$ obtaining:

$$\left(\int_X |f|^p\,d\nu\right)^{1/p} \leq \left(\int_X |f|^q\,d\nu\right)^{1/q} \iff \left(\int_X |f|^p \frac{d\mu}{\mu(X)}\right)^{1/p} \leq \left(\int_X |f|^q \frac{d\mu}{\mu(X)}\right)^{1/q},$$

which immediately implies the desired result. □

Solution. (*of Exercise 6.4.12*) If $g(x) = e^{f(x)}$ then $\int_X e^{p\,f(x)}\,d\mu(x) = \int_x (g(x))^p\,d\mu(x)$ so that

$$\frac{1}{p}\log\left(\int_X e^{p\,f(x)}\,d\mu(x)\right) = \log\left(\int_X e^{p\,f(x)}\,d\mu(x)\right)^{1/p} = \log\|g\|_p \to \log\|g\|_\infty \text{ as } p \to \infty.$$

Since $g \geq 0$ we have $\|g\|_\infty = \operatorname{essup} g$, and since exp is increasing and continuous on the reals $\operatorname{essup} g = \exp(\operatorname{essup} f)$ so that the required limit is $\operatorname{essup} f$.

For the second limit first notice that $\cosh(p\,f(x)) = \cosh(p\,|f(x)|)$, since cosh is an even function. It is not restrictive then to assume $f(x) = |f(x)| \geq 0$. As p goes to ∞ the term $e^{-p\,f(x)}$ tends to 0 if $f(x) > 0$, and is constantly 1 if $f(x) = 0$, so that it should be ininfluent; in other words we guess that (assuming $f(x) \geq 0$) the second limit coincides with the first. In fact:

$$\frac{1}{p}\log\left(\int_X \cosh(p\,f(x))\,d\mu(x)\right) - \frac{1}{p}\log\left(\int_X e^{p\,f(x)}\,d\mu(x)\right) = \frac{1}{p}\log\left(\frac{\int_X \cosh(pf)}{\int_X e^{p\,f}}\right);$$

for $t \geq 0$ we have $e^t/2 \leq \cosh t \leq e^t$ so that the argument of the logarithm is between $1/2$ and 1, hence the preceding is between $-\log 2/p$ and 0, and tends to 0 as $p \to \infty$. Then the second limit is exactly $\operatorname{essup}|f| = \|f\|_\infty$. $\square$

Solution. (*of Exercise 6.4.13*) We know that if $0 < p < q < r < \infty$ and $1/q = \alpha/p + \beta/r$, with α, $\beta > 0$ and $\alpha + \beta = 1$ then $\|f\|_q \leq \|f\|_p^\alpha\,\|f\|_r^\beta$, so that

$$\varphi(q) = \|f\|_q^q \leq \|f\|_p^{\alpha\,q}\,\|f\|_r^{\beta\,q} = \varphi(p)^{\alpha\,q/p}\,\varphi(r)^{\beta\,q/r};$$

taking logarithms we get

$$\log\varphi(q) \leq \frac{\alpha\,q}{p}\,\log\varphi(p) + \frac{\alpha\,q}{r}\,\log\varphi(r),$$

we now have $\alpha\,q/p + \beta\,q/r = 1$, so that if $\gamma = \alpha\,q/p$ and $\delta = \beta\,q/r$ we have $q = \gamma\,p + \delta\,r$, a convex combination, and from above we get

$$\log\varphi(\gamma\,p + \delta\,r) \leq \gamma\,\log\varphi(p) + \delta\,\log\varphi(r),$$

thus concluding the proof.

$\square$

7. Signed and complex measures

As remarked at the beginning, additive functions appear also with arbitrary real values, not only positive values, e.g the electric charge contained in a given portion of space. The fundamental requirement is that of countable additivity.

7.1. Preliminaries. Assume that $(X, \mathcal{M})$ is a measurable space; we have seen (2.1.8) that any linear combination with positive coefficients of positive measures defined on $\mathcal{M}$ is still a positive measure on $\mathcal{M}$. We then try to define a signed measure in the following way: assume that $\mu, \nu : \mathcal{M} \to [0, \infty]$ are positive measures defined on $\mathcal{M}$; define $\lambda : \mathcal{M} \to [-\infty, \infty]$ by $\lambda(E) = \mu(E) - \nu(E)$, for every $E \in \mathcal{M}$; of course, since $\infty - \infty$ is not defined, one of the two measures has to be finite for this to make sense; assume for definiteness $\mu(X) < \infty$; then λ is defined, and λ does not assume the value ∞. It is easy to check that λ is countably additive.

Definition. Let $(X, \mathcal{M})$ be a measurable space. A function $\nu : \mathcal{M} \to \tilde{\mathbb{R}}$ is said to be a *signed measure* if $\nu(\emptyset) = 0$ and ν is countably additive, that is, for every disjoint sequence $(A_n)_{n\in\mathbb{N}}$ of $\mathcal{M}$ we have

$$\nu\left(\bigcup_{n=0}^{\infty} A_n\right) = \sum_{n=0}^{\infty} \nu(A_n).$$

Implicit in this formula is the fact that $\sum_{n=0}^{\infty} \nu(A_n)$ is meaningful, whatever the disjoint sequence $(A_n)_{n\in\mathbb{N}}$ is, in particular, no $\infty - \infty$ or $-\infty + \infty$ is ever encountered; and we have

$$\sum_{n=0}^{\infty} \nu(A_n) := \lim_{m\to\infty} \sum_{n=0}^{m} \nu(A_n),$$

but the convergence is necessarily also absolute, as is easy to see: the union $\bigcup_{n=0}^{\infty} A_n$ does not depend on the order of the terms, so that the series is commutatively convergent, hence absolutely convergent; this fact will be confirmed also indirectly. A signed measure can assume at most one of the two values $\{-\infty, \infty\}$: if $A \in \mathcal{M}$ has measure $\nu(A) = \pm\infty$ then every $B \in \mathcal{M}$ containing A has the same infinite measure: in fact $\nu(B) = \nu(A) + \nu(B \setminus A) = \pm\infty + \nu(B \setminus A)$; then $\nu(B \setminus A)$ cannot be $\mp\infty$, and $\nu(B) = \pm\infty$. In particular, $\nu(X)$ must be ∞ $(-\infty)$ if some set $A \in \mathcal{M}$ has measure ∞ $(-\infty)$, so if $\infty \in \nu(\mathcal{M})$ then $-\infty \notin \nu(\mathcal{M})$.

As seen above, the difference of two positive measures, one of which finite, is a signed measure. If $\mu : \mathcal{M} \to [0, \infty]$ is a positive measure, and f is an extended real valued measurable function, $\mu-$integrable in the extended sense (4.4.1) the indefinite integral of f , $\nu_f = \nu : \mathcal{M} \to \tilde{\mathbb{R}}$ defined by

$$\nu(E) = \int_E f\,d\mu = \int_E f^+\,d\mu - \int_E f^-\,d\mu$$

is then a signed measure (each of $\nu^{\pm}(E) = \int_E f^{\pm}$ is a measure, and at least one of them is finite). We shall prove that every signed measure arises in this way, for a convenient positive measure μ.

7.1.1. *Continuity on monotone sequences.* Exactly as in the positive case we get:

Proposition. *Let $(X, \mathcal{M}, \nu)$ be a signed measure space. Then :*

- (i) Continuity from below *If $A_0 \subseteq A_1 \subseteq \dots$ is an increasing sequence in $\mathcal{M}$ with union A, then $\nu(A) = \lim_{n\to\infty} \nu(A_n)$.*
- (ii) Continuity from above on sets of finite measure *$A_0 \supseteq A_1 \supseteq \dots$ is a decreasing sequence in $\mathcal{M}$ with intersection A, and $\nu(A_m)$ is finite for some $m \in \mathbb{N}$, then $\nu(A) = \lim_{n\to\infty} \nu(A_n)$.*

Proof. Imitate the proof given in 2.1.6 and 2.1.7 for positive measures. □

7.1.2. *Terminology.* If $(X, \mathcal{M}, \nu)$ is signed measure space, given a subset $A \in \mathcal{M}$ we say that:

- A is *positive* if $\nu(B) \geq 0$ for every $B \subseteq A$, $B \in \mathcal{M}$.
- A is *negative* if $\nu(B) \leq 0$ for every $B \subseteq A$, $B \in \mathcal{M}$.
- A is *null* if $\nu(B) = 0$ for every $B \subseteq A$, $B \in \mathcal{M}$.

Then a set is null if and only if it is simultaneously positive and negative. If $\nu(E) = \int_E f\,d\mu$ for a function $f : X \to \tilde{\mathbb{R}}$ integrable in the extended sense on the measure space $(X, \mathcal{M}, \mu)$, then every measurable subset of $P = \{f \geq 0\}$ is positive, every measurable subset of $Q = \{f \leq 0\}$ is negative, and every measurable subset of $Z(f) = \{f = 0\}$ is null. Clearly:

. *The positive sets $\mathcal{P}$, the negative sets $\mathcal{Q}$, and the null sets $\mathcal{N}$ are all $\sigma-$ideals of the tribe $\mathcal{M}$, that is, they are closed under countable union and measurable subsets.*

Proof. Easy exercise. □

Of course a set of strictly positive/negative measure is not necessarily a positive/negative set!

7.1.3. *The Hahn decomposition.* The following lemma is the key result in our approach to the theory of signed measures:

LEMMA. *Let $\nu : \mathcal{M} \to [-\infty, \infty[$ be a signed measure which does not assume the value ∞. Then $s = \sup\{\nu(E) : E \in \mathcal{M}\}$ is finite, and actually there is $P \in \mathcal{M}$ such that $\nu(P) = s$ (i.e., the supremum is a maximum).*

Proof. Clearly we have $0 \leq s \leq \infty$. Pick a sequence $s_n = \nu(A_n) \geq 0$, with $A_n \in \mathcal{M}$, such that $\lim_{n\to\infty} s_n = s$. For every $n \in \mathbb{N}$ let $\mathcal{A}_n$ be the subalgebra of $\mathcal{M}$ generated by the subset $\{A_0, \ldots, A_n\}$: it is a finite algebra of parts of X, whose elements can be written as disjoint unions of intersections of the A_k and their complements (see 1.4.1). Let B_n be the union of all the elements G of the basis of $\mathcal{A}_n$ with $\nu(G) \geq 0$: then $\nu(B_n) = \max\{\nu(E) : E \in \mathcal{A}_n\}$, in particular $\nu(B_n) \geq \nu(A_n)$.

We prove that $\nu(\limsup_{n\to\infty} B_n) = s$, so that we conclude, with $P = \limsup_{n\to\infty} B_n$. For, consider the sequence $k \mapsto \nu(B_n \cup \cdots \cup B_{n+k})$; we claim that this is an increasing sequence of real numbers. In fact, since $\mathcal{A}_{n+k+1} \supseteq \mathcal{A}_{n+k}$, the set $B_{n+k+1} \smallsetminus (B_n \cup \cdots \cup B_{n+k})$ is the union of those elements G of the basis of $\mathcal{A}_{n+k+1}$ not already contained in $B_n \cup \cdots \cup B_{n+k}$, with $\nu(G) \geq 0$, so that we have $\nu(B_{n+k+1} \smallsetminus (B_n \cup \cdots \cup B_{n+k})) \geq 0$. By continuity from below, if $C_n = \bigcup_{k=0}^{\infty} B_{n+k}$ we get $\nu(C_n) = \lim_{k\to\infty} \nu(B_n \cup \cdots \cup B_{n+k})$, with $\nu(C_n) \geq \nu(B_n) \geq \nu(A_n)$. Now $C_0 \supseteq C_1 \supseteq C_2 \supseteq \ldots$ is a decreasing sequence of sets of finite positive measure, so that if $P = \bigcap_{n=0}^{\infty} C_n (= \limsup B_n)$ we get $\nu(P) = \lim_{n\to\infty} \nu(C_n) \geq \lim_{n\to\infty} \nu(A_n) = s$ by continuity from above. □

. THE HAHN DECOMPOSITION THEOREM *If $(X, \mathcal{M}, \nu)$ is a signed measure space then there is a positive set $P \in \mathcal{M}$ whose complement $Q = X \smallsetminus P$ is a negative set. Such a decomposition is unique modulo null sets: if P', Q' is another, then $P \vartriangle P' = Q \vartriangle Q'$ is a null set.*

Proof. Assume that the measure does not assume the value ∞. By the lemma there is $P \in \mathcal{M}$ such that $\nu(P) = s = \max\{\nu(A) : A \in \mathcal{M}\}$ is the largest possible measure of a measurable set. If $A \subseteq P$ is measurable with $\nu(A) < 0$ then $\nu(A)$ is finite (otherwise, as seen, $\nu(P) = -\infty$, too) and then $\nu(P \smallsetminus A) = \nu(P) - \nu(A) > \nu(P)$, contradicting maximality of $\nu(P)$, so that P is a positive set. And $Q = X \smallsetminus P$ cannot contain a set B of strictly positive measure, otherwise $P \cup B$ has measure $\nu(P) + \nu(B) > \nu(P)$, again contradicting maximality of $\nu(P)$. Hence Q is a negative set. If P' is another positive set with negative complement then $P \smallsetminus P'$ is negative being disjoint from P', and positive being contained in P, hence it is null; same for $P' \smallsetminus P$. □

Given a signed measure ν and a Hahn decomposition $X = P \cup Q$ for ν, we set for $E \in \mathcal{M}$:

$$\nu^+(E) = \nu(E \cap P), \quad \nu^-(E) = -\nu(E \cap Q); \quad |\nu|(E) = \nu^+(E) + \nu^-(E);$$

then $\nu = \nu^+ - \nu^-$, and ν^+, ν^-, $|\nu|$ are positive measures, called respectively the *positive part*, the *negative part*, and the *total variation* of ν. If $\sigma : X \to \{-1, 1\}$ is defined by $\sigma(x) = \chi_P - \chi_Q$ then

$$\nu(E) = \nu^+(E) - \nu^-(E) = \int_E \sigma(x)\, d|\nu|(x), \quad \text{for every } E \in \mathcal{M}.$$

We have done what we set out to do with signed measures.

The Hahn decomposition has certain minimality properties that we explore in the next exercise.

EXERCISE 7.1.1. Let $(X, \mathcal{M}, \nu)$ be a signed measure space; let $\mu : \mathcal{M} \to [0, \infty]$ be a positive measure that dominates ν, in the sense that $|\nu(E)| \leq \mu(E)$ for every $E \in \mathcal{M}$. Prove that then $|\nu|(E) \leq \mu(E)$ for every $E \in \mathcal{M}$.

Thus the total variation $|\nu|$ of the signed measure ν may be described as the smallest positive measure that dominates ν.

Solution. Immediate: if $X = P \cup Q$ is a Hahn decomposition for ν we have, for every $E \in \mathcal{M}$:

$$\nu^+(E) = \nu(E \cap P) \leq \mu(E \cap P); \quad \nu^-(E) = -\nu(E \cap Q) = |\nu(E \cap Q)| \leq \mu(E \cap Q),$$

so that

$$|\nu|(E) = \nu^+(E) + \nu^-(E) \leq \mu(E \cap P) + \mu(E \cap Q) = \mu(E).$$

□

We stress the fact that a signed measure which does not assume the value ∞ has actually a finite maximum value ($\nu(P) = \nu^+(X)$), and if it does not assume the value $-\infty$ it has a finite minimum value ($\nu(Q) = -\nu^-(X)$). It can be proved that the set of values of a finite signed measure is always a compact subset of $\mathbb{R}$.

7.1.4. *Mutual singularity of measures.* Let $(X, \mathcal{M})$ be a measurable space, and let $\nu : \mathcal{M} \to \tilde{\mathbb{R}}$ be a signed measure. We say that ν is *supported* by a measurable $M \in \mathcal{M}$ if $X \setminus M$ is a $\nu-$null set. Clearly the set of all $M \in \mathcal{M}$ which support ν is closed under countable intersection (a countable union of $\nu-$null sets is a $\nu-$null set), and the formation of supersets (it is the set of all complements of the $\sigma-$ideal of $\nu-$null sets). To give some examples: Lebesgue measure is supported by $\mathbb{R} \setminus \mathbb{Q}$, a counting measure is supported only by the entire space, the Dirac measure δ_0 on $\mathcal{P}(\mathbb{R})$ is supported by the singleton $\{0\}$, or by any set containing 0. The identically zero measure is supported by any set in $\mathcal{M}$, $\emptyset$ included.

Definition. Let ν and μ be signed measures on the measurable space $(X\mathcal{M})$. We say that ν is singular with respect to μ, and write $\nu \perp \mu$, if ν is supported by a set $B \in \mathcal{M}$ that is null for μ.

This relation is clearly symmetric: if B is null for μ then μ is supported by $A = X \setminus B$, a set null for ν. Then we often say that μ and ν are *mutually singular*, instead of saying that ν is singular with respect to μ, or vice–versa. In other words:

Two signed measures μ, ν on the same measurable space $(X, \mathcal{M})$ are said to be *mutually singular* if they are supported by disjoint sets, equivalently there is a partition $X = A \cup B$ of X into disjoint sets $A, B \in \mathcal{M}$ such that B is a null set for μ and A is null for ν. If ν is a signed measure then ν^+ and ν^-, the positive and negative parts of ν, are mutually singular, being supported on the two complementary pieces of a Hahn decomposition for ν. The Lebesgue measure and Dirac's measure are mutually singular.

7.1.5. *Absolute continuity.*

Definition. Let ν be a signed measure, μ a positive measure on the measurable space $(X, \mathcal{M})$. We say that ν is absolutely continuous with respect to μ, and write $\nu \ll \mu$, if $E \in \mathcal{M}$ and $\mu(E) = 0$ imply $\nu(E) = 0$.

Typically, if $f : X \to \tilde{\mathbb{R}}$ is $\mu-$integrable in the extended sense, then $\nu(E) := \int_E f\, d\mu$ is absolutely continuous with respect to μ. And any measure is absolutely continuous with respect to the counting measure, for the simple fact that the only set with counting measure zero is the empty set, which by definition has $\nu-$measure zero for every measure ν. Clearly $\nu \ll \mu$ if and only if ν^+, $\nu^- \ll \mu$, as is immediate to see. Absolute continuity and mutual singularity exclude each other, in this sense:

. *If $\mu, \nu : \mathcal{M} \to \tilde{\mathbb{R}}$ are measures, μ positive and ν signed, $\nu \perp \mu$ and $\nu \ll \mu$ together imply $\nu = 0$.*

Proof. By mutual singularity we can write $X = P \cup Q$, disjoint union, with $\mu(Q) = 0$ and P null for ν. For every $E \in \mathcal{M}$ contained in Q we have $\mu(E) = 0$, so that $\nu(E) = 0$ by absolute continuity. Then Q is a null set for ν, so that X, union of two sets null for ν, is null for ν. □

If $F : \mathbb{R} \to \mathbb{R}$ is increasing, then clearly the Radon–Stieltjes measure $dF = \mu_F$ is not absolutely continuous with respect to Lebesgue measure when F is not continuous (singletons have Lebesgue measure zero). Surprisingly, there are continuous increasing functions F such that not only μ_F is *not* absolutely continuous with respect to λ, it is even singular, $\mu_F \perp \lambda$! One such example is the *devil's staircase*, or Cantor's function, which we shall describe later (8.5).

Exercise 7.1.2. On a measurable space $(X, \mathcal{M})$ let μ be a positive measure, and let λ, ν be signed measures. Assume that $\lambda + \nu$ is also a signed measure. Prove that:

(i) If $\lambda \perp \mu$ and $\nu \perp \mu$ then $(\lambda + \nu) \perp \mu$.
(ii) If $\lambda \ll \mu$ and $\nu \ll \mu$ then $\lambda + \nu \ll \mu$.

Solution. (i) If $X = P_1 \cup Q_1$ and $X = P_2 \cup Q_2$ are measurable partitions of X such that μ is supported by Q_1 and Q_2, and λ is supported by P_1, ν by P_2, then μ is supported by $Q_1 \cap Q_2$ and $\lambda + \nu$ by $P_1 \cup P_2 = X \setminus (Q_1 \cap Q_2)$.

(ii) Trivial.

□

7.1.6. As said above, the typical example of the situation $\nu \ll \mu$ is when $\nu(E) = \int_E f\, d\mu$ for some f integrable in the extended sense. But it may happen that $\nu \ll \mu$, yet there is no f such that $d\nu = f\, d\mu$: in $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, with λ Lebegue measure and $\mu = \varkappa$ counting measure we have $\lambda \ll \mu$, but there is no $f : \mathbb{R} \to \mathbb{R}$ such that $\lambda(E) = \int_E f\, d\mu = \sum_{x \in E} f(x)$ for every $E \in \mathcal{B}(\mathbb{R})$: necessarily such an f, having finite sum on any bounded interval, would have a countable cozero set, thus of Lebesgue measure zero, and for every $E \subseteq \mathbb{R} \setminus \mathrm{Coz}(f)$ we would have $\int_E f\,d\mu = 0$; in fact we have $\nu \perp \lambda$, if $\nu(E) = \int_E f\,d\mu$. As it frequently happens, this situation is due to lack of $\sigma-$finiteness of the measure μ.

We turn to the proof of the most important result in this context, the Radon–Nikodym theorem. First a lemma:

Lemma. *Let μ, ν be finite positive measures on the measurable space $(X, \mathcal{M})$. Then, unless $\nu \perp \mu$ there is $\varepsilon > 0$ and $E \in \mathcal{M}$ such that $\mu(E) > 0$ and E is a positive set for $\nu - \varepsilon\,\mu$.*

Proof. Consider the signed measure $\nu - (1/n)\mu$, for $n = 1, 2, \dots$; for every $n \geq 1$ let $P_n \cup Q_n$ be a Hahn decomposition for it. If for some $n \geq 1$ we have $\mu(P_n) > 0$, then we have our $E = P_n$ and our $\varepsilon = 1/n$. If not, then $\mu(P_n) = 0$ for every n implies $\mu(P) = 0$, with $P = \bigcup_{n=1}^{\infty} P_n$. Then $\mu \perp \nu$; in fact, if $Q = X \setminus P = \bigcap_{n=1}^{\infty} Q_n$ we have $(\nu - (1/n)\mu)(Q) \leq 0$ for every $n \geq 1$ (since $Q \subseteq Q_n$, a negative set for $\nu - (1/n)\mu$), hence $\nu(Q) \leq (1/n)\mu(Q)$ for every $n \geq 1$, that is, $\nu(Q) = 0$. □

7.1.7. *The Radon–Nikodym theorem.*

Theorem. *Let $(X, \mathcal{M}, \mu)$ be a $\sigma-$finite measure space, and let $\nu : \mathcal{M} \to \tilde{\mathbb{R}}$ be a $\sigma-$finite signed measure. Then there exist unique signed measures $\nu_{\mathrm{s}}, \nu_{\mathrm{a}} : \mathcal{M} \to \tilde{\mathbb{R}}$ such that $\nu_{\mathrm{s}} \perp \mu$, $\nu_{\mathrm{a}} \ll \mu$ and $\nu = \nu_{\mathrm{s}} + \nu_{\mathrm{a}}$. Moreover there is, unique up to equality $\mu-$almost everywhere, a function $\rho : X \to \mathbb{R}$, integrable in the extended sense, such that $\nu_{\mathrm{a}}(E) = \int_E \rho\,d\mu$, for every $E \in \mathcal{M}$.*

Proof. First, we split ν into its negative and its positive part, and prove the theorem separately for $\nu^{\pm}$; in other words we assume ν positive. We can also assume that both measures are finite: since the measures are both $\sigma-$finite by hypothesis we can consider a countable partition $X = \bigcup_{n=0}^{\infty} X_k$ into disjoint sets X_k such that $\mu(X_k) < \infty$ and $\nu(X_k) < \infty$, and prove the theorem on each piece X_k separately.

So we assume that μ and ν are finite positive measures on $(X, \mathcal{M})$. Let (all integrals are with respect to the measure μ):

$$V = \{f \in L^+(X) : \int_E f\,d\mu \leq \nu(E) \quad \text{for every } E \in \mathcal{M}\}.$$

Observe that V is closed under $\vee$, that is, if $f, g \in V$ then $f \vee g \in V$. In fact, for every $E \in \mathcal{M}$:

$$\int_E f \vee g = \int_{E \cap \{f \leq g\}} f \vee g + \int_{E \cap \{f > g\}} f \vee g = \int_{E \cap \{f \leq g\}} g + \int_{E \cap \{f > g\}} f \leq$$
$$\leq \nu(E \cap \{f \leq g\}) + \nu(E \cap \{f > g\}) = \nu(E).$$

Let now:

$$a = \sup\left\{\int_X f \,:\, f \in V\right\}.$$

We prove that there exists $\rho \in V$ such that $\int_X \rho = a$. In fact there is a sequence $f_n \in V$ such that $a_n = \int_X f_n \uparrow a$; if $g_n = f_0 \vee f_1 \cdots \vee f_n$ then g_n is an increasing sequence in V, hence $\int_E g_n \leq \nu(E)$ for every $E \in \mathcal{M}$ and every $n \in \mathbb{N}$, and $\int_X g_n \geq a_n$, so that, if $\rho \in L^+(X)$ is the limit of g_n, by monotone convergence we have $\int_E \rho = \lim_n \int_E g_n \leq \nu(E)$ for every $E \in \mathcal{M}$, in particular for $E = X$ we get $\int_X \rho = \lim_n \int_X g_n \geq \lim_n a_n = a$. Let $\nu_{\mathrm{a}}(E) = \int_E \rho\,d\mu$; then $\nu_{\mathrm{a}} \ll \mu$, and we claim that $(\nu - \nu_{\mathrm{a}}) \perp \mu$. If not, the lemma just proved implies that there are $\varepsilon > 0$ and $E \in \mathcal{M}$ such that $\mu(E) > 0$ and $\nu(A) - \nu_{\mathrm{a}}(A) \geq \varepsilon\,\mu(A)$ per ogni $A \in \mathcal{M}$ contained in E; then $\rho + \varepsilon\,\chi_E \in V$, but $\int_X (\rho + \varepsilon\,\chi_E) = a + \varepsilon\,\mu(E) > a$, a contradiction.

Uniqueness of the singular part and the absolutely continuous part: if $\nu = \nu_{\mathrm{s}} + \nu_{\mathrm{a}} = \lambda_{\mathrm{s}} + \lambda_{\mathrm{a}}$, with ν_{s}, $\lambda_{\mathrm{s}} \perp \mu$ and $\nu_{\mathrm{a}} \ll \mu$, $\nu_{\mathrm{a}} \ll \mu$ then

$$\alpha := \nu_{\mathrm{s}} - \lambda_{\mathrm{s}} = \lambda_{\mathrm{a}} - \nu_{\mathrm{a}},$$

where the left–hand side is singular with respect to μ, and the right–hand side is absolutely continuous. Then the measure α is identically zero (7.1.5). Uniqueness of the density function is in 4.2.3. □

Remark. The most interesting and important part of this theorem is the existence of the density ρ; that is, observe that

. *If $(X, \mathcal{M}, \mu)$ is a $\sigma-$finite measure space, $\nu : \mathcal{M} \to \tilde{\mathbb{R}}$ is a $\sigma-$finite signed measure, and $\nu \ll \mu$, then there is, unique up to equality $\mu-$almost everywhere, a function $\rho : X \to \mathbb{R}$, integrable in the extended sense, such that $\nu(E) = \int_E \rho\,d\mu$, for every $E \in \mathcal{M}$.*

7.1.8. L^1 *for a sum of measures.* Observe that if μ, ν are positive measures on the measurable space $(X,\mathcal{M})$, then $\mathcal{L}^1_{\mu+\nu}=\mathcal{L}^1_\mu\cap\mathcal{L}^1_\nu$, and that for $f\in L^1(\mu+\nu)$ we have

$$\int_X f\,d(\mu+\nu)=\int_X f\,d\mu+\int_X f\,d\nu.$$

As usual the proof is carried out first for positive simple functions, then for functions in L^+, etc.; do it. Moreover we trivially have $\mu\ll\mu+\nu$ and $\nu\ll\mu+\nu$. If μ and ν are $\sigma-$finite we then have $d\mu=\alpha\,d(\mu+\nu)$ and $d\nu=\beta\,d(\mu+\nu)$ for some positive functions α,β; since $d(\mu+\nu)=(\alpha+\beta)\,d(\mu+\nu)$ we clearly have $\alpha+\beta=1$ (a.e).

7.1.9. *The* $(\varepsilon,\delta)-$*condition for absolute continuity.* For future use we give the following:

Definition. Let $(X,\mathcal{M})$ be a measurable space, let $\mu:\mathcal{M}\to[0,\infty]$ be a positive measure and $\nu:\mathcal{M}\to\tilde{\mathbb{R}}$ a signed measure. We say that ν is $(\varepsilon,\delta)-$absolutely continuous with respect to μ if for every $\varepsilon>0$ there exists $\delta=\delta_\varepsilon>0$ such that $|\nu(E)|\le\varepsilon$ for every $E\in\mathcal{M}$ with $\mu(E)\le\delta$.

Trivially this form of absolute continuity is stronger than the previous one: if $\mu(E)=0$ then $|\nu(E)|\le\varepsilon$ for every $\varepsilon>0$, so that $\nu(E)=0$. And it is strictly stronger, in general: if $X=[0,1]$, $\mathcal{M}=\mathcal{B}([0,1])$, $\mu=m=dx=$Lebesgue measure, and $d\nu=dx/x$ (that is, $\nu(E)=\int_E dx/x$) then $\nu\ll\mu$, but for any given $\delta>0$ we have $\nu(]0,\delta])=\int_{]0,\delta]}dx/x=\infty$. However:

Exercise 7.1.3. With μ and ν as above, if ν is a finite measure then $\nu\ll\mu$ implies that ν is $(\varepsilon,\delta)-$absolutely continuous with respect to μ (reduce to ν positive; see also 7.2.9).

7.2. Complex measures.

Definition. If $(X,\mathcal{M})$ is a measurable space, a countably additive function $\mu:\mathcal{M}\to\mathbb{C}$ is said to be a *complex measure*.

Notice that $\mu(\emptyset)=0$: in fact $\mu(\emptyset)=\mu(\emptyset\cup\emptyset)=\mu(\emptyset)+\mu(\emptyset)$, and since $\mu(\emptyset)$ is finite this implies $\mu(\emptyset)=0$. Of course, if $\mu:\mathcal{M}\to\mathbb{C}$ is a complex measure then $\mu_r=\operatorname{Re}\mu$ and $\mu_\iota=\operatorname{Im}\mu$ are finite signed measures, and as such they have finite total variations; if $\nu=|\mu_r|+|\mu_\iota|$, then $\mu_r\ll\nu$ and $\mu_\iota\ll\nu$; since all measures are finite by Radon–Nikodym theorem there are real functions $u,v\in L^1(\nu)$ such that $d\mu_r=u\,d\nu$ and $d\mu_\iota=v\,d\nu$, so that $d\mu=(u+i\,v)\,d\nu$, meaning that

$$\mu(E)=\int_E(u+i\,v)\,d\nu\quad\text{for every } E\in\mathcal{M};$$

Let's define the *total variation* of the complex measure μ, a positive measure $|\mu|$ that dominates μ, in the sense that $|\mu(E)|\le|\mu|(E)$, and $|\mu|$ is the smallest positive measures ν that verifies this condition; it is of course to be proved that such a minimal measure exists.

Assume that ν is any positive measure on $\mathcal{M}$ such that $d\mu=\rho\,d\nu$ with $\rho\in L^1_\nu(X,\mathbb{C})$ (we have seen that such measures exist, e.g. $\nu=|\mu_r|+|\mu_\iota|$). We shall prove that then $|\rho|\,d\nu$ is the required total variation.

For every $E\in\mathcal{M}$ define

(Total Variation) $$|\mu|(E)=\sup\left\{\sum_{k=1}^m|\mu(E_k)|:\,E_1,\dots,E_m\in\mathcal{M},\,E=\bigcup_{k=1}^m E_k,\quad\text{disjoint union}\right\}.$$

We immediately have $|\mu|(E)\le\lambda(E)$ for any positive measure λ that dominates μ:

$$\sum_{k=1}^m|\mu(E_k)|\le\sum_{k=1}^m\lambda(E_k)=\lambda(E);$$

if we prove that $|\mu|(E)=\int_E|\rho|\,d\nu$ for every $E\in\mathcal{M}$ we are done: this proves also that $|\mu|$ is a measure. The proof, which requires a little work, is in the next number. However, the real case is very easy:

Exercise 7.2.1. Prove that if $\mu:\mathcal{M}\to\mathbb{R}$ is a signed measure and $d\mu=\rho\,d\nu$ for some positive measure ν on $\mathcal{M}$ and some $\rho:X\to\mathbb{R}$, with ρ $\nu-$integrable in the extended sense, then $|\mu|$, as defined by the above formula, is exactly $\mu^+ + \mu^-$, and we also have $d|\mu|=|\rho|\,d\nu$.

Exercise 7.2.2. Prove directly that the Total Variation formula defines $|\mu|$ as a positive measure.

Exercise 7.2.3. If $\mu,\,\nu:\mathcal{M}\to\mathbb{C}$ are complex measures on the same measurable space $(X,\mathcal{M})$, then $|\mu+\nu|\le|\mu|+|\nu|$, and for every $\alpha\in\mathbb{C}$ $|\alpha\,\mu|=|\alpha|\,|\mu|$.

Solution. The second statement is trivial. For the first, for every $E \in \mathcal{M}$:

$$|(\mu+\nu)E| = |\mu(E)+\nu(E)| \le |\mu(E)| + |\nu(E)| \le |\mu|(E) + |\nu|(E),$$

so that $|\mu| + |\nu|$ is a measure that dominates $\mu+\nu$, and is hence larger than $|\mu+\nu|$, smallest measure that does it. □

7.2.1. *The total variation as an integral.* Given an interval I of $\mathbb{R}$, we consider the *angle* (with vertex at the origin) given by $A(I) = \{z \in \mathbb{C} : z = r\, e^{i\theta},\ \theta \in I,\ r \ge 0\}$. If $0 < \varepsilon < \pi$, since cos is decreasing in $[0, \varepsilon]$ we have $\cos\vartheta \ge \cos\varepsilon$ for $\vartheta \in [-\varepsilon, \varepsilon]$, so that

$$A[-\varepsilon,\varepsilon] = \{r\,(\cos\vartheta + i\sin\vartheta) : r \ge 0,\ \vartheta \in [-\varepsilon,\varepsilon]\} = \{z \in \mathbb{C} : \operatorname{Re} z \ge |z|\, \cos\varepsilon\};$$

then, if $\alpha, \beta \in \mathbb{R}$, $(\beta-\alpha)/2 = \varepsilon < \pi$ and $\gamma = (\alpha+\beta)/2$ we have, since $e^{-i\gamma}\, A[\alpha,\beta] = A[-\varepsilon,\varepsilon]$ that

$$A[\alpha,\beta] = \{z \in \mathbb{C} : \operatorname{Re}(e^{-i\gamma}\, z) \ge \cos\varepsilon\, |z|\}.$$

Proposition. *Let $\mu : \mathcal{M} \to \mathbb{C}$ be a complex measure on the measurable space $(X, \mathcal{M})$, and assume that $\nu : \mathcal{M} \to [0,\infty]$ is a positive measure such that $d\mu = \rho\, d\nu$, for a function $\rho \in L^1_\nu(X, \mathbb{C})$. Then $|\mu|(E) = \int_E |\rho|\, d\nu$ for every $E \in \mathcal{M}$, if $|\mu|$ is defined as in the above formula.*

Proof. It is clearly enough to prove that

$$|\mu|(X) = \int_X |\rho|\, d\nu$$

since the same proof will work for every $E \in \mathcal{M}$. Let $Z = Z(\rho)$ be the zero–set of ρ. For $n = 3, 4, \ldots$ we partition the complex plane into n angles $A\,]\theta_{n,(k-1)}, \theta_{n,k}]$ where $\theta_{n,k} = -\pi + 2k\pi/n$, $k = 0, \ldots, n$, and consider $E(nk) = \rho^{\leftarrow}(A\,]\theta_{n,(k-1)}, \theta_{n,k}] \setminus \{0\})$. Then, for every $n \ge 3$ the family $\{Z, E(n1), \ldots, E(nn)\}$ is a partition of X into measurable sets; we write $E(k)$ in place of $E(n,k)$ and $\gamma(k)$ in place of $\gamma(nk) = (\theta_{n,(k-1)} + \theta_{n,k})/2$ in the proof that follows. For $x \in E(k)$ we have $\cos(\pi/n)|\rho(x)| \le \operatorname{Re}(e^{-i\gamma(k)}\, \rho(x))$ so that:

$$\cos(\pi/n) \int_X |\rho(x)|\, d\mu(x) = \cos(\pi/n) \int_{X\setminus Z} |\rho(x)|\, d\mu(x) = \sum_{k=1}^{n} \int_{E(k)} \cos(\pi/n)\, |\rho(x)|\, d\mu(x) \le$$

$$\le \sum_{k=1}^{n} \int_{E(k)} \operatorname{Re}(e^{-i\gamma(k)}\, \rho(x))\, d\mu(x) = \sum_{k=1}^{n} \operatorname{Re}\left(e^{-i\gamma(k)} \int_{E(k)} \rho(x)\, d\mu(x)\right) \le \sum_{k=1}^{n} \left| e^{-i\gamma(k)} \int_{E(k)} \rho(x)\, d\mu(x)\right| =$$

$$= \sum_{k=1}^{n} |\nu(E(k))| \le |\nu|(X).$$

We have obtained

$$\cos(\pi/n) \int_X |\rho(x)|\, d\mu(x) \le |\nu|(X),$$

for every $n \ge 3$; letting n tend to ∞ we conclude. □

Notice that with notations as in the preceding proposition we have for every $E \in \mathcal{M}$:

$$\mu(E) = \int_E \operatorname{sgn}\rho(x)\, |\rho|\, d\nu = \int_E \sigma(x)\, d|\mu|,$$

where $\sigma(x) = \operatorname{sgn}\rho(x)$ is a measurable function of absolute value 1 for $|\mu|-$almost every $x \in X$. Of course if $\mu = \mu_r + i\,\mu_\iota$ then $\mu_r \ll |\mu|$ and $\mu_\iota \ll |\mu|$ so that $d\mu_r = u\, d|\mu|$ and $d\mu_\iota = v\, d|\mu|$ with u, v positive functions in $L^1(|\mu|)$; then from $\mu(E) = \int_E \sigma(x)\, d|\mu|$ we get

$$\mu_r(E) = \int_E \operatorname{Re}\sigma\, d|\mu| = \int_E u\, d|\mu| \quad \text{and} \quad \mu_\iota(E) = \int_E \operatorname{Im}\sigma\, d|\mu| = \int_E v\, d|\mu|,$$

so that $u = \operatorname{Re}\sigma$ and $v = \operatorname{Im}\sigma$ ($|\mu|-$a.e. in X).

7.2.2. *Total variation of a premeasure.* The following result has some intrinsic interest, and will be used in the future:

. *Let $(X, \mathcal{M})$ be a measurable space, and let $\nu : \mathcal{M} \to \mathbb{C}$ be a complex measure. Assume that $\mathcal{A}$ is an algebra of parts of X which generates $\mathcal{M}$ as a $\sigma-$algebra. Then for every $A \in \mathcal{A}$ we have*

$$|\nu|(A) = \sup\left\{\sum_{k=1}^{m} |\nu(A_k)| :\ A_k \in \mathcal{A},\ \{A_1, \ldots, A_m\} \text{ pairwise disjoint subsets of } A\right\}.$$

Proof. It is clearly enough to prove it for $A = X$. Given a disjoint finite family $E_1, \dots, E_m \in \mathcal{M}$, and $\varepsilon > 0$, for every $k \in \{1, \dots, m\}$ we pick $A_k \in \mathcal{A}$ such that $|\nu|(E_k \triangle A_k) \le \varepsilon$ (being finite, the measure $|\nu|$ on $\mathcal{M}$ is the Carathèodory extension of the restriction of $|\nu|$ to $\mathcal{A}$, so we can apply 2.5.7). By the usual trick we make $\{A_1, \dots, A_m\}$ into a disjoint family, $B_k = A_k \smallsetminus \left(\bigcup_{j=1}^{k-1} A_j\right)$. We claim that

$$\sum_{k=1}^{m} |\nu(B_k)| \ge \sum_{k=1}^{m} |\nu(E_k)| - \alpha(\varepsilon),$$

where $\lim_{\varepsilon\to 0^+} \alpha(\varepsilon) = 0$. This clearly concludes the proof. Notice that

$$|\nu(B_k)| = \left|\nu(A_k) - \nu\left(\bigcup_{j=1}^{k-1} A_k \cap A_j\right)\right| \ge |\nu(A_k)| - \left|\nu\left(\bigcup_{j=1}^{k-1} A_k \cap A_j\right)\right| \ge$$
$$\ge |\nu(A_k)| - |\nu|\left(\bigcup_{j=1}^{k-1} A_k \cap A_j\right) \ge |\nu(A_k)| - \sum_{j=1}^{k-1} |\nu|(A_j \cap A_k).$$

Now we have, for $j \ne k$, since $E_j \cap E_k = \emptyset$:

$$|\chi_{A(k)} \wedge \chi_{A(j)}| = |\chi_{A(k)} \wedge \chi_{A(j)} - \chi_{E(k)} \wedge \chi_{E(j)}| \le |\chi_{A(k)} - \chi_{E(k)}| \vee |\chi_{A(j)} - \chi_{E(j)}|,$$

(formula 1.2.3) so that

$$|\nu|(A_k \cap A_j) = \int_X \chi_{A(k)} \wedge \chi_{A(j)}\, d|\nu| \le \int_X |\chi_{A(k)} - \chi_{E(k)}| \vee |\chi_{A(j)} - \chi_{E(j)}|\, d|\nu| \le$$
$$\le \int_X (|\chi_{A(k)} - \chi_{E(k)}| + |\chi_{A(j)} - \chi_{E(j)}|)\, d|\nu| = |\nu|(A_k \triangle E_k) + |\nu|(A_j \triangle E_j) \le 2\varepsilon.$$

We then have

$$|\nu(B_k)| \ge |\nu(A_k)| - 2(k-1)\varepsilon \ge |\nu(A_k)| - 2m\varepsilon,$$

so that

$$\sum_{k=1}^{m} |\nu(B_k)| \ge \sum_{k=1}^{m} |\nu(A_k)| - 2m^2\varepsilon;$$

moreover

$$|\nu(A_k)| = |\nu(A_k) - \nu(E_k) + \nu(E_k)| \ge |\nu(E_k)| - |\nu(A_k) - \nu(E_k)|,$$

and

$$|\nu(A_k) - \nu(E_k)| = |\nu(A_k \smallsetminus E_k) - \nu(E_k \smallsetminus A_k)| \le |\nu(A_k \smallsetminus E_k)| + |\nu(E_k \smallsetminus A_k)| \le |\nu|(A_k \triangle E_k) \le \varepsilon,$$

so that $|\nu(A_k)| \ge |\nu(E_k)| - \varepsilon$; finally we get

$$\sum_{k=1}^{m} |\nu(B_k)| \ge \sum_{k=1}^{m} |\nu(E_k)| - (2m^2 + m)\varepsilon.$$

□

EXERCISE 7.2.4. Prove that if μ is a complex measure and $\mu(X) = |\mu|(X)$ then μ is a positive measure.

Solution. Recall that for a complex function the modulus of the integral (with respect to a positive measure) is the integral of the modulus if and only if there is $\alpha \in \mathbb{R}$ such that for almost all points in the domain the values of the function are in the ray $\{re^{i\alpha} : r \ge 0\}$ (4.2.1). Then, if

$$\mu(X) = |\mu|(X) = \int_X \sigma(x)\, d|\mu| = \int_X d|\mu|(> 0),$$

we have that $\sigma(x) = 1$ for $|\mu|$–a.e. $x \in X$, so that $\mu(E) = |\mu|(E)$ for every measurable E. □

7.2.3. $L^1(\mu)$ *for signed or a complex measure.* Having a measurable space $(X,\mathcal{M})$ and a signed or a complex measure μ on it, we define $L^1(\mu)$ as $L^1(|\mu|)$, with $|\mu|$ the total variation of μ. By definition the integral is

$$\int_X f\,d\mu := \int_X f\sigma\,d|\mu| \left(= \int_X f\,\mathrm{Re}\,\sigma\,d|\mu| + i\int_X f\,\mathrm{Im}\,\sigma\,d|\mu| = \int_X f\,d\mu_r + i\int_X f\,d\mu_\iota\right).$$

where σ is the density of μ with respect to $|\mu|$, a measurable function with values in the unit circle $\mathbb{U}$ for $|\mu|-$almost every $x\in X$. We have the fundamental inequality:

$$\left|\int_X f\,d\mu\right| \le \int_X |f|\,d|\mu|.$$

If μ is a signed measure, $\mu = \mu^+ - \mu^-$, with $d\mu^+ = \chi_P\,d|\mu|$ and $d\mu^- = \chi_Q\,d|\mu|$, with P, Q a Hahn decomposition for μ we evidently have

$$\int_X f\,d\mu := \int_X f(\chi_P - \chi_Q)\,d|\mu| = \int_X f\,\chi_P\,d|\mu| - \int_X f\,\chi_Q\,d|\mu| = \int_X f\,d\mu^+ - \int_X f\,d\mu^-,$$

in other words

$$\int_X f\,d\mu = \int_X f\,d\mu^+ - \int_X f\,d\mu^-.$$

From this it is easy to see, arguing as we did for proving linearity of the integral on $L^1_\mu(X,\mathbb{R})$ that if μ and ν are real valued signed measures, and $f\in\mathcal{L}^1(\mu)\cap\mathcal{L}^1(\nu)$, then (assuming that $\mu+\nu$ is defined):

$$\int_X f\,d(\mu+\nu) = \int_X f\,d\mu + \int_X f\,d\nu.$$

The same is true for a sum of complex measures, as is easy to check. In other words, the symbol $\int_X f\,d\mu$ is bilinear, linear in f when μ is given, but also in μ for f fixed

Exercise 7.2.5. Let μ be a complex measure on the measurable space $(X,\mathcal{M})$. Prove that

$$|\mu|(X) = \sup\left\{\left|\int_X f\,d\mu\right| : f\,\text{measurable},\ \ |f|\le 1\right\};$$

and that the sup is actually a max.

Solution. First observe that since $|\mu|$ is a finite measure every bounded measurable function is in $L^1(\mu)$, so the integrals in the formula exist. By the fundamental inequality

$$\left|\int_X f\,d\mu\right| \le \int_X |f|\,d|\mu| \le \int_X |d\mu| = |\mu|(X);$$

and since $d\mu = \sigma\,d|\mu|$ with $|\sigma(x)| = 1$ for $|\mu|-$a.e. $x\in X$ we have

$$\left|\int_X \bar\sigma\,d\mu\right| = \int_X d|\mu| = |\mu|(X).$$

□

7.2.4. *The chain rule.* If $(X,\mathcal{M})$ is a measurable space, $\nu:\mathcal{M}\to[0,\infty]$ a positive measure and μ a signed or a complex measure defined on $\mathcal{M}$, we have used the notation $d\mu = \rho\,d\nu$ to mean that for every $E\in\mathcal{M}$ we have $\mu(E) = \int_E \rho\,d\nu$, with ρ in $L^1(\nu)$ if μ is a finite measure, an extended real valued function integrable in the extended sense if ν is a not finite signed measure. In this case of course $\mu \ll \nu$; and the Radon–Nikodym theorem says that if both measures are $\sigma-$finite then ρ exists. That $d\mu = \rho\,d\nu$ may also be denoted as

$$\frac{d\mu}{d\nu} = \rho;$$

in fact the function ρ, if it exists, is uniquely determined up to $\nu-$a.e. equality (4.2.3).

Proposition. *(i) In the above hypotheses a measurable function $f: X\to\mathbb{C}$ is in $L^1(\mu)$ if and only if $f\rho\in L^1(\nu)$, and*

$$\int_X f\,d\mu = \int_X f\,\rho\,d\nu.$$

(ii) Assume that μ is also a positive measure, that λ is a signed or complex measure defined on $\mathcal{M}$, and that $d\lambda = \tau\, d\mu$. Then $d\lambda = \tau\,\rho\, d\nu$; in the other notation

$$\frac{d\lambda}{d\nu} = \frac{d\lambda}{d\mu}\,\frac{d\mu}{d\nu}.$$

Proof. (i) By definition $f \in L^1(\mu)$ iff $f \in L^1(|\mu|)$ and $\int_X f\, d\mu = \int_X f\,\sigma\, d|\mu|$, and from 7.2.1 we have $d|\mu| = |\rho|\, d\nu$, and $\sigma = \operatorname{sgn}\rho$. Since $\sigma(x)$ has absolute value 1 for $|\mu|$−a.e. $x \in X$, we have $f\,\sigma \in L^1(|\mu|)$ iff $f \in L^1(|\mu|)$; with the usual decomposition into real and imaginary parts, and into positive and negative parts, we have to prove that if $f \in \mathcal{L}^+(X)$ then

$$\int_X f\, d|\mu| = \int_X f\,|\rho|\, d\nu,$$

knowing that for every $E \in \mathcal{M}$ we have $|\mu|(E) = \int_E |\rho|\, d\nu$. This was proposition 4.1.9.

(ii) Since $\int_X f\, d\lambda = \int_X f \operatorname{sgn}\tau(x)\, d|\lambda|$ we can consider only positive measures. That is, we are reduced to prove the chain rule for positive measures. We have to prove that $\lambda(E) = \int_E \tau\,\rho\, d\nu$ for every $E \in \mathcal{M}$. From (i) we know that

$$\int_X (\chi_E\,\tau)\,\rho\, d\nu = \int_X (\chi_E \tau)\, d\mu := \int_E \tau\, d\mu = \lambda(E);$$

the last equality being $d\lambda = \tau\, d\mu$. The proof is concluded. □

EXERCISE 7.2.6. Prove that if μ is a complex measure on the measurable space $(X, \mathcal{M})$ then, for every $f \in L^1(|\mu|)$ we have

$$\int_X f\, d\mu = \int_X f\, d\mu_r + i \int_X f\, d\mu_\iota.$$

Prove that if μ, ν are complex measures on $\mathcal{M}$ and $f : X \to \mathbb{C}$ is a bounded measurable function then

$$\int_X f\, d(\mu + \nu) = \int_X f\, d\mu + \int_X f\, d\nu.$$

Solution. Since $d\mu = \sigma\, d|\mu|$ we have $d\mu_r = \operatorname{Re}(\sigma)\, d|\mu|$ and $d\mu_\iota = \operatorname{Im}(\sigma)\, d|\mu|$; if $f \in L^1(|\mu|)$:

$$\int_X f\, d\mu := \int_X f\,\sigma\, d|\mu| = \int_X f\,(\operatorname{Re}(\sigma) + i\, \operatorname{Im}\sigma)\, d|\mu| = \int_X f\, \operatorname{Re}(\sigma)\, d|\mu| + i \int_X f\, \operatorname{Im}\sigma\, d|\mu| =$$
$$= \int_X f\, d\mu_r + i \int_X f\, d\mu_\iota,$$

the last equality being due to (i) of the above proposition.

For the last assertion: the hypothesis that f is bounded measurable implies that $f \in L^1(\lambda)$ for every finite positive measure λ on $(X, \mathcal{M})$, in particular then $f \in L^1(|\mu + \nu|)$. We clearly have

$$(\mu + \nu)_r = \mu_r + \nu_r;\ (\mu + \nu)_\iota = \mu_\iota + \nu_\iota,$$

so that, by what just proved:

$$\int_X f\, d(\mu + \nu) = \int_X f\, d(\mu_r + \nu_r) + i \int_X f\, d(\mu_\iota + \nu_\iota),$$

and we are reduced to the case in which μ and ν are real, i.e. signed measures. We know by direct proof (see 7.1.8) that if μ, ν are positive measures then $\int_X f\, d(\mu + \nu) = \int_x f\, d\mu + \int_X f\, d\nu$, so that we have to prove that for $f \in L^1(\mu) \cap L^1(\nu)$, with μ and ν positive measures we have

$$\int_X f\, d(\mu - \nu) = \int_X f\, d\mu - \int_X f\, d\nu.$$

If $d\mu = u\, d(\mu + \nu)$ and $d\nu = v\, d(\mu + \nu)$ then $0 \le u, v \le 1$ so that $f\,u,\ f\,v \in L^1(\mu) \cap L^1(\nu)$ if $f \in L^1(\mu) \cap L^1(\nu)$, and since $d(\mu - \nu) = (u - v)\, d(\mu + \nu)$ we get

$$\int_X f\, d(\mu - \nu) = \int_X f\,(u - v)\, d(\mu + \nu) = \int_X f\,u\, d(\mu + \nu) - \int_X f\,v\, d(\mu + \nu) = \int_X f\, d\mu - \int_X f\, d\nu.$$

□

7.2.5. *The Banach space of finite* $\mathbb{K}-$*valued measures.* If $(X, \mathcal{M})$ is a measurable space a finite $\mathbb{K}-$valued measure on $\mathcal{M}$ is a countably additive function $\mu : \mathcal{M} \to \mathbb{K}$. We shall drop the adjective finite unless needed for emphasis or to avoid ambiguity. When $\mathbb{K} = \mathbb{R}$ one also speaks of a *finite signed measure, or charge* (as we have seen signed measures are allowed to take $-\infty$ or $+\infty$ as values, but we exclude this in what follows).

It is clear that any finite linear combination with coefficients in $\mathbb{K}$ of $\mathbb{K}-$valued measures on $\mathcal{M}$ is still a $\mathbb{K}-$valued measure on $\mathcal{M}$, so that these measure are naturally a $\mathbb{K}-$linear space. We have

. *For every measurable space* $(X, \mathcal{M})$ *the set* $M(X) = M_{\mathcal{M}}(X, \mathbb{K})$ *of all finite* $\mathbb{K}-$*valued measures defined on* $\mathcal{M}$ *is a Banach space under the norm* $\|\mu\| = |\mu|(X)$.

Proof. By exercise 7.2.5, if μ, ν are $\mathbb{K}-$valued measures on $(X, \mathcal{M})$ we have $|\mu + \nu| \leq |\mu| + |\nu|$, in particular $|\mu + \nu|(X) \leq |\mu|(X) + |\nu|(X)$, so that $\mu \mapsto |\mu|(X)$ is subadditive and hence a norm (the remaining conditions are trivial). Completeness: assume that $\sum_{n\in\mathbb{N}} \mu_n$ is a normally convergent series of $\mathbb{K}-$valued measures, that is, the series $\sum_{n\in\mathbb{N}} \|\mu_n\| = \sum_{n\in\mathbb{N}} |\mu_n|(X)$ is convergent. Given $E \in \mathcal{M}$ we set

$$\mu(E) := \sum_{n\in\mathbb{N}} \mu_n(E)$$

(note that this series is absolutely convergent, since $|\mu_n(E)| \leq |\mu_n|(E) \leq |\mu_n|(X) = \|\mu_n\|$, so that the definition makes sense), and we have to prove that μ is countably additive. If $E = \bigcup_{m\in\mathbb{N}} E_m$, with the $E_m \in \mathcal{M}$ pairwise disjoint, we get

$$\mu(E) = \sum_{n\in\mathbb{N}} \mu_n(E) = \sum_{n\in\mathbb{N}} \left(\sum_{m\in\mathbb{N}} \mu_n(E_m) \right) =$$

(we interchange the sums, proving later the admissibility of this action)

$$\sum_{m\in\mathbb{N}} \left(\sum_{n\in N} \mu_n(E_m) \right) = \sum_{m\in\mathbb{N}} \mu(E_m),$$

as required. The interchange is admissible because the sum of absolute values:

$$\sum_{n\in\mathbb{N}} \left(\sum_{m\in\mathbb{N}} |\mu_n(E_m)| \right) \leq \sum_{n\in\mathbb{N}} \left(\sum_{m\in\mathbb{N}} |\mu_n|(E_m) \right) = \sum_{n\in\mathbb{N}} |\mu_n|(E) \leq \sum_{n\in\mathbb{N}} |\mu_n|(X) < \infty$$

is finite; we apply Tonelli–Fubini theorem in the space $\ell^1(\mathbb{N} \times \mathbb{N})$. Thus μ is a $\mathbb{K}-$valued measure and since:

$$\left| \mu(E) - \sum_{n=0}^{m} \mu_n(E) \right| \leq \sum_{n=m+1}^{\infty} |\mu_n(E)| \leq \sum_{n=m+1}^{\infty} |\mu_n|(E),$$

we have (remember that the sum of a series of positive measures is always a (not necessarily finite) positive measure):

$$\left| \mu - \sum_{n=0}^{m} \mu_n \right| (E) \leq \sum_{n=m+1}^{\infty} |\mu_n|(E),$$

in particular, for $E = X$, we get

$$\left\| \mu - \sum_{n=0}^{m} \mu_n \right\| \leq \sum_{n=m+1}^{\infty} \|\mu_n\|,$$

and the proof ends. □

Exercise 7.2.7. Let $(X, \mathcal{M})$ be a measurable space and let $M(X) = M_{\mathcal{M}}(X, \mathbb{K})$ be the Banach space of finite $\mathbb{K}-$valued measures. As usual, we say that $\mu, \nu \in \mathcal{M}(X)$ are *mutually singular*, and we write $\mu \perp \nu$ if there is a decomposition of X as $X = A \cup B$, disjoint union, with A, $B \in \mathcal{M}$ and B null for μ, A null for ν.

(i) Prove that if $\mu \perp \nu$ then $|\mu + \nu| = |\mu| + |\nu|$, so that $\|\mu + \nu\| = \|\mu\| + \|\nu\|$.

(ii) Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence in $M(X)$ such that $\sum_{n=0}^{\infty} \|\mu_n\| < \infty$, and let $\mu : \mathcal{M} \to [0, \infty]$ be a $\sigma-$finite measure. Let $\mu_j = \lambda_j + \nu_j = \rho_j \, d\mu + \nu_j$ be the Radon-Nikodym decomposition of μ_j with respect to μ, in other words, $\nu_j \perp \mu$. Prove that then

$$\sum_{n=0}^{\infty} \mu_n = \sum_{n=0}^{\infty} \lambda_n + \sum_{n=0}^{\infty} \nu_n = \left(\sum_{n=0}^{\infty} \rho_n \right) d\mu + \sum_{n=0}^{\infty} \nu_n,$$

with both series $\sum_{n=0}^{\infty} \lambda_n$, $\sum_{n=0}^{\infty} \nu_n$ normally converging in $M(X)$, and $\sum_{n=0}^{\infty} \rho_n$ normally converging in $L^1(\mu)$.

7.2.6. *Exercises.*

EXERCISE 7.2.8. Let $\mu, \nu : \mathcal{M} \to [0, \infty]$ be positive measures such that $\lambda = \mu - \nu$ is a signed measure (equivalently, one at least of them is a finite measure). Then $\lambda^+ \leq \mu$ and $\lambda^- \leq \nu$; and if $\mu \perp \nu$, then $\lambda^+ = \mu$ and $\lambda^- = \nu$

Solution. Let $X = P \cup Q$ be a Hahn decomposition for λ. Then, for every $E \in \mathcal{M}$:

$$\lambda^+(E) = \lambda(E \cap P) = \mu(E \cap P) - \nu(E \cap P) \leq \mu(E \cap P) \leq \mu(E),$$

the last inequality being monotonicity of the positive measure μ. Similarly:

$$\lambda^-(E) = -\lambda(E \cap Q) = \nu(E \cap Q) - \mu(E \cap Q) \leq \nu(E \cap Q) \leq \nu(E).$$

If $\mu \perp \nu$ and $A \cup B$ is a partition of X into measurable sets, with μ supported by A and ν supported by B, then, trivially, this is also a Hahn decomposition for λ, so that

$$\lambda^+(E) = \lambda(E \cap A) = \mu(E \cap A) - \nu(E \cap A) = \mu(E \cap A) = \mu(E), \quad \text{for every } E \in \mathcal{M},$$

and similarly for λ^- and ν.

□

EXERCISE 7.2.9. Let $(X, \mathcal{M}, \mu)$ be a measure space. Let $\nu : \mathcal{M} \to \mathbb{C}$ be a (finite) measure.

(i) Prove that $\nu \ll \mu$ iff $\nu_r^\pm$ and $\nu_\iota^\pm$ are all absolutely continuous with respect to μ, and also iff $|\nu| \ll \mu$.
(ii) Prove that if $\nu \ll \mu$, then for every $\varepsilon > 0$ there is $\delta > 0$ such that $\mu(E) \leq \delta$ implies $|\nu(E)| \leq \varepsilon$.
(iii) Assertion (ii) fails if ν is not finite: consider $d\nu = dx/x$ on $[0, 1]$, or $d\nu = |x|\, dx$ on $\mathbb{R}$

(see also 7.1.9)

Solution. (i) Assume $\nu \ll \mu$ and that $\mu(E) = 0$. Then $\nu(E) = \nu_r(E) + i\,\nu_\iota(E) = 0 \iff \nu_r(E) = \nu_\iota(E) = 0$, so that $\nu \ll \mu \iff \nu_r \ll \mu$ and also $\nu_\iota \ll \mu$. Assuming ν signed measure we prove that $\nu \ll \mu$ iff $\nu^\pm \ll \mu$. Let $P \cup Q$ be a Hahn decomposition of ν; if $\mu(E) = 0$ then $\mu(E \cap P) = \mu(E \cap Q) = 0$ by monotonicity of μ, so that $\nu(E \cap P) = \nu^+(E) = 0$ and also $-\nu(E \cap Q) = \nu^-(E) = 0$. Thus $\nu \ll \mu$ implies $\nu^\pm \ll \mu$, and clearly $\nu^\pm \ll \mu$ implies in turn $\nu = \nu^+ - \nu^- \ll \mu$ and $|\nu| = \nu^+ + \nu^- \ll \mu$. And if ν is a complex measure such that ν_r, $\nu_\iota \ll \mu$ we have $|\nu_r|$, $|\nu_\iota| \ll \mu$ hence also $|\nu| \leq |\nu_r| + |\nu_\iota| \ll \mu$.

(ii) By (i) we can assume that ν is a finite positive measure. We argue by contradiction: if not, there is $\varepsilon > 0$ such that for every $n \in \mathbb{N}$ there is $E_n \in \mathcal{M}$, with $\mu(E_n) \leq 1/2^{n+1}$, such that $\nu(E_n) > \varepsilon$. Consider $F_m = \bigcup_{n \geq m} E_n$. Then $\mu(F_m) \leq \sum_{n \geq m} \mu(E_n) \leq \sum_{n \geq m} 1/2^{n+1} = 1/2^m$. The sequence F_m is decreasing, $F_0 \supseteq F_1 \supseteq F_2 \supseteq \dots$; let $F = \bigcap_{m=0}^\infty F_m$; then $\mu(F) = \lim_{m\to\infty} \mu(F_m) = 0$, but $\nu(F_m) \geq \nu(E_m) > \varepsilon$ for every m, so that $\nu(F) = \lim_{m\to\infty} \nu(F_m) \geq \varepsilon$, contradicting $\nu \ll \mu$.

(iii) In fact $\int_{[a,a+\delta]} dx/x = \log(1 + \delta/a) \to \infty$ as $a \to 0^+$, and $\int_{[a,a+\delta]} |x|\, dx = a\,\delta + \delta^2/2 \to \infty$ as $a \to \infty$. □

EXERCISE 7.2.10. Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $\nu : \mathcal{M} \to [0, \infty]$ be another positive measure. We consider the ideal $\mathcal{F} = \mathcal{F}(\mu) = \{E \in \mathcal{M} : \mu(E) < \infty\}$ of the sets of finite $\mu-$measure as a semimetric(=pseudometric) space under the semimetric $\rho(E, F) = \mu(E \triangle F) = \|\chi_E - \chi_F\|_1$, and we assume that ν is finite on $\mathcal{F}$.

Prove that the following are equivalent:

(a) ν is uniformly continuous.
(b) ν is continuous.
(c) ν is continuous at $\emptyset$
(d) $\nu \ll \mu$.

If we consider the quotient $\mathcal{F}_*$ of $\mathcal{F}$ obtained by identifying E and F when $\rho(E, F) = 0$, observe that ν is well defined on this quotient iff $\nu \ll \mu$.

Solution. That (a) implies (b) and that (b) implies (c) is trivial. (c) implies (d): continuity of ν at the emptyset is exactly this: given $\varepsilon > 0$ there is $\delta > 0$ such that $\rho(E, \emptyset)(= \mu(E)) \leq \delta$ implies $|\nu(E) - \nu(\emptyset)| = \nu(E) \leq \varepsilon$. If $\mu(E) = 0$, then $\mu(E) \leq \delta$ holds for every $\delta > 0$, then $\nu(E) \leq \varepsilon$ for every $\varepsilon > 0$, i.e $\nu(E) = 0$. By exercise 7.2.9, we also have (d) implies (c); that (c) implies (a) is trivial, because of additivity of ν.

□

8. Differentiation of measures in euclidean spaces

Recall that by a Radon measure in $\mathbb{R}^n$ we mean a positive Borel measure $\mu : \mathcal{B}_n \to [0, \infty]$ that is finite on compact (hence also on bounded) sets ($\mathcal{B}_n$ is the $\sigma-$algebra of Borel sets of $\mathbb{R}^n$). Let $\mathcal{B}_n^*$ denote the ideal of $\mathcal{B}_n$ consisting of bounded sets. If, as usual, $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$, a locally finite Radon $\mathbb{K}-$valued measure is a countably additive function $\mu : \mathcal{B}_n^* \to \mathbb{K}$. Notice that strictly speaking μ is not a measure, since it is not defined on a $\sigma-$algebra. However μ induces a finite measure on the Borel sets of any compact subset of $\mathbb{R}^n$. And there is a unique positive Radon measure $|\mu| : \mathcal{B}_n \to [0, \infty]$ such that $|\mu(E)| \le |\mu|(E)$ for every bounded set, and $|\mu|$ is the smallest measure that does so: simply put together all total variations of measures induced on compact subsets; specifically: $\mathbb{R}^n = \bigcup_{k=0}^{\infty} Q_k$, where $Q_k = [-k, k]^n$ is the compact cube centered at the origin with side $2k$; if μ_k is the $\mathbb{K}-$valued finite measure induced by μ on Q_k, let ν_k be the measure $\nu_k(E) = |\mu_k|(E \cap (Q_k \smallsetminus Q_{k-1}))$, and set $|\mu|(E) = \sum_{k=1}^{\infty} \nu_k(E)$. It is obvious that $|\mu|$ has the required properties. Moreover μ extends to a finite $\mathbb{K}-$valued measure on all of $\mathcal{B}_n$ iff $|\mu|(\mathbb{R}^n) < \infty$: finite $\mathbb{K}-$valued Borel measures are of course Radon measures as well. For every locally finite Radon measure μ and every bounded open set $U \subseteq \mathbb{R}^n$ there is a finite Borel measure ν which coincides with μ on subsets of U and is zero outside U: simply put $\nu(E) = \mu(U \cap E)$. This explains why, when concerned with properties of a local nature, we may assume μ finite and defined on all $\mathcal{B}_n$.

8.0.7. *Regularity of Radon measures.*

Proposition. *Every positive Radon measure μ is regular.*

Proof. The proof was given in section 2.7. ☐

Let us recall that this means that for every Borel set E we have $\mu(E) = \sup\{\mu(K) : K \subseteq E,\ K \text{ compact}\}$, and $\mu(E) = \inf\{\mu(U) : U \supseteq E,\ U \text{ open}\}$.

8.1. Derivatives with respect to the Lebesgue measure. The most important Radon measure is of course Lebesgue measure; we shorten λ_n to m, dimension will be understood somehow. Every locally finite Radon measure μ that is absolutely continuous with respect to Lebesgue measure is of the form $f\,dm$, where $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n, \mathbb{K})$: by this symbol we denote all measurable $f : \mathbb{R}^n \to \mathbb{K}$ such that $f_{|K} \in L^1_m(K)$, for every compact subset $K \subseteq \mathbb{R}^n$. All continuous functions, and in general all measurable functions bounded on bounded sets are in $L^1_{\mathrm{loc}}(\mathbb{R}^n)$, as all functions in $L^1(\mathbb{R}^n)$; but there are others, e.g. if $f(x) = 1/|x|^\alpha$ with $0 < \alpha < n$ then $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n) \smallsetminus L^1(\mathbb{R}^n)$, and f is unbounded in every nbhd of 0. Since Lebesgue measure is $\sigma-$finite, it is easy to deduce from the Lebesgue–Radon–Nikodym theorem:

. *If $\mu : \mathcal{B}_n^* \to \mathbb{K}$ is a locally finite Radon measure on $\mathbb{R}^n$, there exist and are unique two locally finite Radon measures λ and ν, with $\lambda \ll m$ and $\nu \perp m$ such that $\mu = \lambda + \nu$. Moreover there is a function $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n, \mathbb{K})$, unique up to equality $m-$a.e., such that $d\lambda = f\,dm$.*

Proof. Apply the Lebesgue–Radon–Nikodym theorem to Q_k, then put together the results. ☐

The Radon–Nikodym derivative with respect to m assumes a much more analytical character, and can be related to the usual notion of derivative given in Calculus (in the one–dimensional case; in the $n-$dimensional case it becomes the notion of density, as defined and used by physicists). To be more precise, the following result can be proved:

. *With μ, λ, ν and f as in the preceding proposition, we have for a.e. $x \in \mathbb{R}^n$*

$$\lim_{r \to 0^+} \frac{\mu(B(x,r])}{m(B(x,r])} = \lim_{r \to 0^+} \frac{\lambda(B(x,r])}{m(B(x,r])} = f(x) \qquad \lim_{r \to 0^+} \frac{\nu(B(x,r])}{m(B(x,r])} = 0.$$

here $B(x,r] = \{\xi \in \mathbb{R}^n : |\xi - x| \le r\}$ is the closed euclidean ball centered at x of radius r, but any norm can be used, and even sets much more general than balls. Call *substantial family* any family of Borel subsets of $\mathbb{R}^n$, indexed by the points of $\mathbb{R}^n$ and the strictly positive reals, $\{E_r(x) : x \in \mathbb{R}^n,\ r > 0\}$ such that:

- For every $x \in \mathbb{R}^n$ and $r > 0$ we have $E_r(x) \subseteq B(x, r]$.
- For every $x \in \mathbb{R}^n$ there is $\alpha = \alpha(x) > 0$ such that $m(E_r(x)) \ge \alpha\, m(B(x,r])$, for every $r > 0$.

That is, the sets $E_r(x)$ must have a measure which is at least a fixed percentage of the measure of the ball they are in. They do not have to contain the point x; given $E \subseteq B(0,1]$ with $m(E) > 0$ the family $E_r(x) = \{x + rE : x \in \mathbb{R}^n,\ r > 0\}$ is substantial. A typical example are the half–open intervals

$\{]x, x+r] : x \in \mathbb{R},\, r > 0\}$ and $\{]x-r, x] : x \in \mathbb{R},\, r > 0\}$. Notice also that no relation is assumed between $E_r(x)$ and $E_r(y)$ for different $x,\, y \in \mathbb{R}^n$.

Then the above theorem is true with $E_r(x)$ replacing $B(x, r]$. Let us give an explicit statement:

. THE DIFFERENTIATION THEOREM *Let $\mu : \mathcal{B}_n^* \to \mathbb{K}$ be a locally finite Radon measure on $\mathbb{R}^n$; let $\mu = \lambda + \nu$, with $\lambda \ll m$, $d\lambda = f\, dm$ and $\nu \perp m$. If $(E_r(x))_{x\in\mathbb{R}^n,\, r>0}$ is a substantial family in $\mathbb{R}^n$, then for $m-$almost every $x \in \mathbb{R}^n$ we have*

$$\lim_{r\to 0^+} \frac{\mu(E_r(x))}{m(E_r(x))} = \lim_{r\to 0^+} \frac{\lambda(E_r(x))}{m(E_r(x))} = f(x); \quad \lim_{r\to 0^+} \frac{\nu(E_r(x))}{m(E_r(x))} = 0.$$

We prove now parts of this theorem; completion of the proof is in 8.6

8.1.1. *Derivative of the absolutely continuous part.* For $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n, \mathbb{K})$ and $r > 0$, $x \in \mathbb{R}^n$ we define the *average* of f over the ball $B(x, r]$:

$$A_r f(x) = \fint_{B(x,r]} f(\xi)\, dm(\xi) := \frac{1}{m(B(x, r])} \int_{B(x,r]} f(\xi)\, dm(\xi);$$

the function $(x, r) \mapsto A_r f(x)$ is continuous from $\mathbb{R}^n\times]0, \infty[$ to $\mathbb{K}$ (see exercise 8.1.1). For the time being, we accept without proof the following difficult:

THEOREM. *If $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ then $\lim_{r\to 0^+} A_r f(x) = f(x)$ for almost every $x \in \mathbb{R}^n$.*

For every constant k we have $k = \fint_{B(x,r]} k\, dm(\xi)$, in particular $f(x) = \fint_{B(x,r]} f(x)\, dm(\xi)$ so that the preceding theorem says that

$$\lim_{r\to 0^+} \fint_{B(x,r]} (f(\xi) - f(x))\, d\xi = 0 \quad \text{for almost every } x \in \mathbb{R}^n.$$

We can prove that the same is true with an absolute value in the integral. Specifically: we call a point $x \in \mathbb{R}^n$ a *Lebesgue point* for f if

$$\lim_{r\to 0^+} \fint_{B(x,r]} |f(\xi) - f(x)|\, d\xi = 0;$$

this clearly implies that $f(x) = \lim_{r\to 0^+} A_r f(x)$. Then

. *If $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ then almost all points of $\mathbb{R}^n$ are Lebesgue points for f.*

Proof. For every constant $c \in \mathbb{K}$ we consider the function $x \mapsto |f(x) - c|$; by the preceding theorem we have

$$|f(x) - c| = \lim_{r\to 0^+} \fint_{B(x,r]} |f(\xi) - c|\, d\xi, \quad \text{for every } x \in \mathbb{R}^n \smallsetminus E(c) \text{ with } m(E(c)) = 0.$$

Let D be a countable subset of $\mathbb{K}$ dense in $\mathbb{K}$, and let $E = \bigcup_{c\in D} E(c)$; then $m(E) = 0$. We prove that every $x \in \mathbb{R}^n \smallsetminus E$ is Lebesgue point for f. In fact, given $x \in \mathbb{R}^n \smallsetminus E$ and $\varepsilon > 0$ we may pick $c \in D$ such that $|f(x) - c| \le \varepsilon$; then, for every $r > 0$:

$$\fint_{B(x,r]} |f(\xi) - f(x)|\, d\xi \le \fint_{B(x,r]} (|f(\xi) - c| + |c - f(x)|)\, d\xi = \fint_{B(x,r]} |f(\xi) - c|\, d\xi + |f(x) - c|,$$

and taking $\limsup_{r\to 0^+}$ of both sides:

$$\limsup_{r\to 0^+} \fint_{B(x,r]} |f(\xi) - f(x)|\, d\xi \le |f(x) - c| + |f(x) - c| \le 2\varepsilon.$$

$\square$

We now immediately have:

. *Let $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, and le $E_r(x)$ be a substantial family in $\mathbb{R}^n$. Then for every x in the Lebesgue set of f, hence for almost every $x \in \mathbb{R}^n$, we have*

$$\lim_{r\to 0^+} \fint_{E_r(x)} |f(\xi) - f(x)|\, d\xi = 0; \quad \lim_{r\to 0^+} \fint_{E_r(x)} f(\xi)\, d\xi = f(x).$$

Proof. We have, for x in the Lebesgue set of f:

$$\fint_{E_r(x)} |f(\xi)-f(x)|\,d\xi = \frac{1}{m(E_r(x))}\int_{E_r(x)} |f(\xi)-f(x)|\,d\xi \le \frac{1}{m(E_r(x))}\int_{B(x,r]} |f(\xi)-f(x)|\,d\xi \le$$
$$\le \frac{1}{\alpha\, m(B(x,r])}\int_{B(x,r]} |f(\xi)-f(x)|\,d\xi = \frac{1}{\alpha}\fint_{B(x,r]} |f(\xi)-f(x)|\,d\xi \to 0,$$

as $r\to 0^+$. Since

$$\left|\fint_{E_r(x)} f(\xi)\,d\xi - f(x)\right| = \left|\fint_{E_r(x)} (f(\xi)-f(x))\,d\xi\right| \le \fint_{E_r(x)} |f(\xi)-f(x)|\,d\xi,$$

the first statement immediately implies the second. □

EXERCISE 8.1.1. In $\mathbb{R}^n$ let x_k be a sequence converging to $x\in\mathbb{R}^n$, and let $r_k>0$ converge to $r>0$. If $\chi_k = \chi_{B(x_k,r_k]}$ then χ_k converges a.e. to $\chi=\chi_{B(x,r]}$. Moreover, if $R=\sup_k\{r_k+|x-x_k|\}$ then $B(x_k,r_k]\subseteq B(x,R]$ for every $k\in\mathbb{N}$. Prove that if $f\in L^1_{\rm loc}(\mathbb{R}^n)$ then $A_rf(x)$ is continuous as a map of $\mathbb{R}^n\times]0,\infty[\to\mathbb{K}$.

Solution. Assume that $|x-y|<r$, and pick t with $|x-y|<t<r$ then, since $|x_k-y|\to|x-y|$, and $r_k\to r$ for k large enough, say $k\ge N$, we have $|x_k-y|<r$ and $r_k>t$; then $\chi_k(y)=1$ for $k\ge N$. Similarly one proves that $\chi_k(y)=0$ for k large enough if $|x-y|>r$. Thus $\chi_k(y)$ converges to $\chi(y)$ for every $y\in\mathbb{R}^n$ but at most when $|x-y|=r$, the $(n-1)-$dimensional sphere of center x and radius r, a set of $n-$dimensional measure zero. The rest is clear: $|f|\,\chi_{B(x,R)}$ dominates all $f\,\chi_k$, because $B(x_k,r_k]\subseteq B(x,R]$ for every k (if $|x_k-y|\le r_k$ then $|x-y|\le|x-x_k|+|x_k-y|\le|x-x_k|+r_k\le R$) so that

$$\lim_{k\to\infty}\int_{B(x_k,r_k]} f(\xi)\,d\xi = \int_{B(x,r]} f(\xi)\,d\xi.$$

□

8.1.2. *Monotone functions are derivable almost everywhere.* As a first application of the preceding powerful theorems we show:

THEOREM. *Let $F:\mathbb{R}\to\mathbb{R}$ be increasing. Then F is differentiable a.e., with derivative $F'(x)\ge 0$ a.e.; moreover $F'\in L^1_{\rm loc}(\mathbb{R})$, and for every interval $[a,b]$ we have*

$$F(b)-F(a)\ge\int_{[a,b]} F'(x)\,dx.$$

Proof. Let $\mu=\mu_F$ be the Radon–Stieltjes measure associated to F, defined by $\mu(]a,b])=F(b^+)-F(a^+)$, $\mu(]a,b[)=F(b^-)-F(a^+)$, etc., see 2.2. Let C be the set of points of continuity of F. The following four families are all substantial families in $\mathbb{R}$

$$[x,x+r];\quad [x,x+r[;\quad]x-r,x];\quad [x-r,x],$$

so that, if $f\in L^1_{\rm loc}(\mathbb{R})$ is a representative of the Radon–Nikodym derivative $d\lambda/dm$, where λ is the absolutely continuous part of μ, we have $f(x)\ge 0$, because λ is a positive measure, and by the differentiation theorem 8.1, for a.e $x\in C$:

$$f(x)=\lim_{r\to0^+}\frac{\mu([x,x+r])}{r}=\lim_{r\to0^+}\frac{\mu([x,x+r[)}{r}=\lim_{r\to0^+}\frac{\mu(]x-r,x])}{r}=\lim_{r\to0^+}\frac{\mu([x-r,x])}{r},$$

that is

$$f(x)=\lim_{r\to0^+}\frac{F((x+r)^+)-F(x)}{r}=\lim_{r\to0^+}\frac{F((x+r)^-)-F(x)}{r}=$$
$$\lim_{r\to0^+}\frac{F(x)-F((x-r)^+)}{r}=\lim_{r\to0^+}\frac{F(x)-F((x-r)^-)}{r}.$$

But $F(t^-)\le F(t)\le F(t^+)$ for every $t\in\mathbb{R}$ so that

$$\frac{F((x+r)^-)-F(x)}{r}\le\frac{F(x+r)-F(x)}{r}\le\frac{F((x+r)^+)-F(x)}{r};$$
$$\frac{F(x)-F((x-r)^+)}{r}\le\frac{F(x)-F(x-r)}{r}\le\frac{F(x)-F((x-r)^-)}{r},$$

and the three functions theorem (it: teorema dei carabinieri) allows us to conclude that $F'(x) = f(x)$ exists for a.e. $x \in C$; and since $\mathbb{R} \setminus C$ is countable, hence of zero Lebesgue measure, $F'(x)$ exists a.e. in $\mathbb{R}$. If $\mu = \lambda + \nu$, with $\lambda \ll m$ and $\nu \perp m$ we have $d\lambda = f(x)\,dx$ and if $a, b \in \mathbb{R}$ with $a < b$ we have

$$\mu(]a,b[) = F(b^-) - F(a^+) = \int_{]a,b[} f(x)\,dx + \nu(]a,b]) \geq \int_{[a,b]} f(x)\,dx,$$

and since $F(b) - F(a) \geq F(b^-) - F(a^+)$ the proof is concluded (notice that $\int_{]a,b]} f(x)\,dx = \int_{[a,b]} f(x)\,dx = \int_{]a,b[} f(x)\,dx$ etc).

□

8.2. Radon measures on the real line and functions of bounded variation. Locally finite Radon measures on $\mathbb{R}$ have a distribution function; in 2.2.1 we defined it for positive Radon measures, called there Radon–Stieltjes measures. The same definition can be given here:

Definition. If $\mu : \mathcal{B}_1^* \to \mathbb{K}$ is a locally finite Radon measure its distribution function with initial point 0 is $F = F_\mu : \mathbb{R} \to \mathbb{K}$ defined by

$$F(x) = \begin{cases} \mu(]0,x]) & x \geq 0 \\ -\mu(]x,0]) & x < 0 \end{cases}.$$

When μ is a finite Borel measure one often chooses the distribution function with $-\infty$ as origin, the function $G(x) = \mu([-\infty, x])$; clearly we have $G(x) = F(x) - F(-\infty) = F(x) + \mu([-\infty, 0])$. Notice that on left–open intervals we recover the measure from F as $\mu(]a,b]) = F(b) - F(a)$.

8.2.1. *First properties.* Continuity from below and above of measures implies:

. If μ is a locally finite Radon measure then its distribution function F is right continuous and has finite left limits at every point.

If the measure μ is real then F is the difference of two right continuous increasing functions, $F = A - B$, the distribution functions of μ^+ and μ^-; if the measure is complex then $F = (A - B) + i\,(C - D)$, where all four functions A, B, C, D are right–continuous and increasing.

Proof. For continuity, imitate the proof given for positive measures. For the rest: write $\mu = \mu^+ - \mu^-$ and let A, B be the distribution functions of μ^+, μ^- respectively. If μ is complex, write $\mu = (\mu_r^+ - \mu_r^-) + i\,(\mu_\iota^+ - \mu_\iota^-)$ and let A, B, C, D be the distribution functions of these four positive measures. □

8.2.2. *Functions of bounded variation.* The total variation of a locally finite Radon measure μ on $\mathbb{R}$ is a positive Radon measure, whose distribution function (with initial point 0) is a monotone right–continuous function S with $S(0) = 0$. The minimality property of $|\mu|$ implies that for every pair $a, b \in \mathbb{R}$ with $a < b$ we have

$$|\mu(]a,b])| = |F(b) - F(a)| \leq S(b) - S(a).$$

And this is the property we need to characterize functions which are the distribution functions of some locally finite Radon measure.

Definition. Let I be an interval of $\mathbb{R}$, and let $F : I \to \mathbb{K}$ be a function. We say that F has locally bounded variation on I, or that F is a function of locally bounded variation on I if there is a monotone function $S : I \to \mathbb{R}$ such that

$$|F(x_1) - F(x_2)| \leq |S(x_1) - S(x_2)| \quad \text{for every } x_1, x_2 \in I.$$

We say that F has bounded variation on I if the function S above may be taken bounded on I.

The definition usually given is another, and we shall give it in 8.3. In the situation of the previous definition we say that the monotone function S *controls, or dominates* the variation of F; we may always assume that S is increasing, changing the sign if necessary, so that the definition above might be given by saying that there is an increasing function S such that x_1, $x_2 \in I$ and $x_1 < x_2$ imply $|F(x_2) - F(x_1)| \leq S(x_2) - S(x_1)$. The space of all $\mathbb{K}$−valued functions on I of locally bounded, resp. bounded variation is denoted $BV_{\rm loc}(I, \mathbb{K})$, resp. $BV(I, \mathbb{K})$, with the usual shortenings. Lipschitz continuous functions have locally bounded variation, controlled by linear functions. Clearly monotone functions are in $BV_{\rm loc}(I)$, and in $BV(I)$ if they also are bounded; in fact $BV_{\rm loc}(I)$ is the vector space generated by monotone functions. We collect here some results on functions of bounded variation:

. Let I be an interval of $\mathbb{R}$. Then

(i) $BV_{\rm loc}(I,\mathbb{K})$ *and* $BV(I,\mathbb{K})$ *are vector spaces.*
(ii) *A complex function* $F : I \to \mathbb{C}$ *is in* $BV(I)$ *or* $BV_{\rm loc}(I)$ *iff so are* $\operatorname{Re} F$ *and* $\operatorname{Im} F$
(iii) *A real function* $F : I \to \mathbb{R}$ *is in* $BV_{\rm loc}(I)$ *if and only if* F *is the difference of two increasing functions.*
(iv) *A function* $F \in BV_{\rm loc}(I)$ *has finite left and right limits at every point* $c \in I$*; and has an at most countable set of points of discontinuity. Moreover* $F'(x)$ *exists for* $m-$*a.e.* $x \in I$.

Proof. (i) If $F, G \in BV_{\rm loc}(I,\mathbb{K})$ and the increments of F and G are controlled by the increasing functions S and T respectively, then, assuming $x_1 < x_2$:

$$|(F+G)(x_2) - (F+G)(x_1)| \le |F(x_2) - F(x_1)| + |G(x_2) - G(x_1)| \le$$
$$\le (S(x_2) - S(x_1)) + (T(x_2) - T(x_1)) = (S+T)(x_2) - (S+T)(x_1),$$

with $S+T$ increasing; and $|\alpha\, F(x_2) - \alpha\, F(x_1)| \le |\alpha| S(x_2) - |\alpha| S(x_1)$, with $|\alpha|\, S$ increasing; same for BV.

(ii) If F has (locally) bounded variation then

$$|\operatorname{Re}(F)(x_1) - \operatorname{Re} F(x_2)| \le |F(x_1) - F(x_2)| \le |S(x_1) - S(x_2)|$$

for $x_1, x_2 \in I$, and the same for $\operatorname{Im} F$. And if $A, B : I \to \mathbb{R}$ have (locally) bounded variation, by (i) so has $F = A + i\,B$.

(iii) If F is real with variation dominated by the increasing function $S : I \to \mathbb{R}$ then $S+F$ and $S-F$ are increasing functions; assume in fact $x_1 < x_2$, $x_1,\, x_2 \in I$:

$$(S+F)(x_1) \le (S+F)(x_2) \iff F(x_1) - F(x_2) \le S(x_2) - S(x_1);$$
$$(S-F)(x_1) \le (S-F)(x_2) \iff F(x_2) - F(x_1) \le S(x_2) - S(x_1),$$

and since $|F(x_2) - F(x_1)| = \max\{F(x_1) - F(x_2),\, F(x_2) - F(x_1)\} \le S(x_2) - S(x_1)$, both are true. We have $F = (S+F)/2 - (S-F)/2$

(iv) All this is true for increasing functions, hence for their linear combinations. □

8.2.3. *The Radon measure defined by a function* $BV_{\rm loc}(\mathbb{R})$. We prove that:

Proposition. *Let* $F : \mathbb{R} \to \mathbb{K}$ *be a function in* $BV_{\rm loc}(\mathbb{R})$. *There is a unique locally finite Radon measure* $\mu = \mu_F : \mathcal{B}_1^* \to \mathbb{K}$ *such that for every* $a, b \in \mathbb{R}$, $a < b$*:*

$$\mu(]a,b]) = F(b^+) - F(a^+); \quad \mu(]a,b[) = F(b^-) - F(a^+); \quad \mu([a,b]) = F(b^+) - F(a^-);$$
$$\mu([a,b[) = F(b^-) - F(a^-); \quad \mu(\{a\}) = F(a^+) - F(a^-).$$

Moreover the distribution function of μ_F *with initial point* 0 *is*

$$G(x) = F(x^+) - F(0^+).$$

And the measure is finite on all of $\mathcal{B}_1$ *if and only if* F *has globally bounded variation.*

Proof. Write $F = (A - B) + i\,(C - D)$ (with C, D missing if F is real valued); then $\mu_A, \mu_B, \mu_C, \mu_D$ are positive Radon measures defined on $\mathcal{B}_1$; and $\mu = \mu_A - \mu_B + i\,(\mu_C - \mu_D)$ is a locally finite Radon measure on $\mathcal{B}_1^*$, whose values on intervals are as described. This proves existence; for uniqueness, two locally finite $\mathbb{K}-$valued Radon measures μ, ν which coincide on compact intervals coincide on all of $\mathcal{B}_1^*$: in fact, given a bounded Borel E set pick a compact interval containing it, say $[-a,a] \supseteq E$. The set $\mathcal{D} = \{A \subseteq [-a,a] :\, A \in \mathcal{B}_1,\, \mu(A) = \nu(A)\}$ is a Dynkin class of parts of $[-a,a]$, as is immediate to prove; since $\mathcal{D}$ contains all subintervals of $[-a,a]$, it contains also all Borel subsets of $[-a,a]$, in particular E.

The assertions on global boundedness are easy. □

Notice that the values of F on its points of discontinuity are completely irrelevant for the measure μ_F.

8.3. The point variation of a function on an interval. We have defined functions of bounded variation, or locally bounded variation, but we haven't defined variation; we are going to do it now. Given a function $f : I \to \mathbb{K}$ (the range could be also any normed space) its *total variation* $Vf(I)$ on I is a positive real number, or $+\infty$, defined in the following way. A subdivision of the interval I is a finite subset $\{x_0, \dots, x_m\}$ of I containing the extremes of I that belong to I; it is understood that a subdivision has always a strictly increasing indexing, $x_0 < \cdots < x_m$. The variation of f on the subdivision $\{x_0, \dots, x_m\}$

is defined as $\sum_{k=1}^m |f(x_k) - f(x_{k-1})|$. The total variation of f on the interval I is the supremum of the variations of f over all subdivisions of I:

$$Vf(I) = \sup\left\{\sum_{k=1}^m |f(x_k) - f(x_{k-1})| : \{x_0, \dots, x_m\} \text{ a subdivision of } I\right\}.$$

Observe that $Vf(I) = 0$ if and only if f is constant on I. We immediately see that if f is of bounded variation according to the definition in 8.2.2, then $Vf(I) < \infty$; in fact, if $S : I \to \mathbb{R}$ is bounded increasing, and $|f(t) - f(s)| \le S(t) - S(s)$ for $s < t$ and $s, t \in I$, then

$$\sum_{k=1}^m |f(x_k) - f(x_{k-1})| \le \sum_{k=1}^m (S(x_k) - S(x_{k-1})) = S(x_m) - S(x_0) \le S(\sup I) - S(\inf I),$$

so that

$$(*) \qquad Vf(I) \le S(\sup I) - S(\inf I).$$

8.3.1. *The variation function.* To prove the converse, namely that $Vf(I) < \infty$ implies the existence of a bounded monotone function controlling the variation, we need to prove a sort of additivity of the variation with respect to the interval I. First observe that the finer the subdivision, the higher the variation: we say of course that a subdivision $\{y_0, \dots, y_n\}$ is finer than the subdivision $\{x_0, \dots, x_m\}$ when it has more points, that is $\{x_0, \dots, x_m\} \subseteq \{y_0, \dots, y_n\}$; then

$$\sum_{k=1}^m |f(x_k) - f(x_{k-1})| \le \sum_{l=1}^n |f(y_l) - f(y_{l-1})|;$$

it is in fact enough to prove that the variation on a subdivision increases when adding just one point, and this is trivial. Notice in particular that if $a, b \in I$ and $a < b$ then $|f(b) - f(a)| \le Vf([a, b])$. This monotonicity shows that the variation is attained also on subdivisions which are prescribed to contain a given finite subset.

LEMMA. *If A, B are subintervals of I, and $\max A = \min B$, then $A \cup B$ is an interval, and $Vf(A \cup B) = Vf(A) + Vf(B)$ in particular, if $a < c < b$, a, $b \in I$ then $Vf([a, b]) = Vf([a, c]) + Vf([c, b])$.*

Proof. Let $c = \max A = \min B$. To compute the variations we may work with subdivisions of $A \cup B$ that contain the point c; if $\{x_0, \dots, x_{m-1}, c\}$ is a subdivision of A and $\{c, y_1, \dots, y_n\}$ is a subdivision of B then the union $\{x_0, \dots, x_{m-1}, c, y_1, \dots, y_n\}$ is a subdivision of $A \cup B$ so that

$$\sum_{k=0}^{m-1} |f(x_k) - f(x_{k-1})| + |f(c) - f(x_{m-1})| + |f(y_1) - f(c)| + \sum_{l=2}^n |f(y_l - f(y_{l-1})| \le Vf(A \cup B),$$

and taking suprema of the left–hand side we get (1.1.1):

$$Vf(A) + Vf(B) \le Vf(A \cup B);$$

but we also have

$$\sum_{k=0}^{m-1} |f(x_k) - f(x_{k-1})| + |f(c) - f(x_{m-1})| + |f(y_1) - f(c)| + \sum_{l=2}^n |f(y_l - f(y_{l-1})| \le Vf(A) + Vf(B),$$

and again taking suprema of the left–hand side we get the reverse inequality. □

PROPOSITION. *Let I be an interval of $\mathbb{R}$, and let $f : I \to \mathbb{K}$ be a function. The following are equivalent:*

(i) *The function f is locally of bounded variation, that is, there is an increasing $S : I \to \mathbb{R}$ such that $|f(x_2) - f(x_1)| \le S(x_2) - S(x_1)$ for $x_1 < x_2$, $x_1, x_2 \in I$.*

(ii) *For every compact subinterval $[a, b]$ of I the total variation $Vf([a, b])$ of f on $[a, b]$ is finite.*

Moreover S may be taken bounded if an only if $Vf(I) < \infty$.

Proof. That (i) implies (ii) has been observed above: $Vf([a, b]) \le S(b) - S(a)$. (ii) implies(i). Fix a point $c \in I$ and define $T : I \to \mathbb{R}$ by $T(x) = Vf([c, x])$ for $x \ge c$, $T(x) = -Vf([x, c])$ for $x < c$ ($T = T_c f$ is the *total variation function* of f, with initial point c; in the case of continuous curves, T is exactly the curvilinear abscissa; one can also pick $c = \inf I$ for functions of globally bounded variation). It is easy to see that:

$$\text{if } x_1 < x_2, \text{ with } x_1, x_2 \in I, \text{ then } T(x_2) - T(x_1) = Vf([x_1, x_2]);$$

(one has to consider some cases; if $c \le x_1 < x_2$ then we have, by the lemma, $Vf([c,x_2]) = Vf([c,x_1]) + Vf(x_1,x_2))$; if $x_1 \le c \le x_2$ we have $Vf([x_1,x_2]) = Vf([x_1,c]) + Vf([c,x_2])$, etc). Then T controls the variation of f:

$$(**) \qquad T(x_2) - T(x_1) = Vf([x_1,x_2]) \ge |f(x_2) - f(x_1)|.$$

The assertion on globally bounded variation is easy. □

8.3.2. *Continuity of the variation function.* Notice that formulae (*) and (**) above imply that if S is any increasing function that controls the variation of f, then S controls also the variation of T, that is $T(x_2) - T(x_1) \le S(x_2) - S(x_1)$. Some other interesting properties of T are investigated in the following proposition.

Proposition. *Let $f : I \to \mathbb{K}$ be of locally bounded variation, and let $T = T_c f : I \to \mathbb{R}$ the variation function of f of initial point c. Then, if a is in the interior of I we have*

$$T(a^+) = T(a) + |f(a) - f(a^+)|; \quad T(a) = T(a^-) + |f(a) - f(a^-)|;$$

in particular T is left/right continuous at a iff f is left/right continuous at a.

Proof. Given $\varepsilon > 0$ there is $\delta > 0$ such that if $x \in]a, a+\delta]$ then $T(x) - T(a^+) \le \varepsilon$ and $|f(x) - f(a^+)| \le \varepsilon$. Pick next a subdivision $\{a, x_1, \dots, x_m = x\}$ of $[a,x]$, with $x \in]a, a+\delta]$, such that

$$T(x) - T(a) - \varepsilon = Vf([a,x]) - \varepsilon \le |f(x_1) - f(a)| + \sum_{k=2}^{m} |f(x_k) - f(x_{k-1})| \le$$
$$|f(x_1) - f(a)| + T(x) - T(x_1) \le |f(x_1) - f(a)| + T(x) - T(a^+) + T(a^+) - T(x_1) \le$$
$$|f(x_1) - f(a)| + \varepsilon \le |f(x_1) - f(a^+) + f(a^+) - f(a)| + \varepsilon \le |f(a^+) - f(a)| + 2\varepsilon;$$

the last two inequalities because $x_1 \in]a, a+\delta]$. We have proved that given $\varepsilon > 0$ there is $\delta > 0$ such that for every $x \in]a, a+\delta]$ we have $T(x) - T(a) \le |f(a^+) - f(a)| + 3\varepsilon$; letting $x \to a^+$ we get $T(a^+) - T(a) \le |f(a^+) - f(a)| + 3\varepsilon$, so that $T(a^+) - T(a) \le |f(a^+) - f(a)|$; since $|f(x) - f(a)| \le T(x) - T(a)$ we get also $|f(a^+) - f(a)| \le T(a^+) - T(a)$, and finally

$$T(a^+) = T(a) + |f(a^+) - f(a)|;$$

in much the same way we get $T(a) = T(a^-) + |f(a) - f(a^-)|$. □

8.3.3. *Essential variation and point variation.* Unless $f \in BV_{\rm loc}(I)$ is continuous, the variation is not finitely additive on the interval algebra of I: if $a < c < b$, a, $b \in I$ then $Vf([a,b]) = Vf([a,c]) + Vf([c,b])$, as seen in 8.3.1; now $Vf([a,c]) = Vf([a,c[) + |f(c) - f(c^-)|$ and $Vf([c,b]) = |f(c) - f(c^+)| + Vf(]c,b])$, as seen in the previous proposition, so that

$$Vf([a,b]) - (Vf([a,c[) + Vf([c,b])) = |f(c) - f(c^-)|; \; Vf([a,b]) - (Vf([a,c]) + Vf(]c,b])) = |f(c) - f(c^+)|.$$

But if f is continuous then the variation is a premeasure on the interval algebra, the Radon–Stieltjes premeasure of its total variation function $T = T_c f$, as is immediate to see. And the variation is also a premeasure on the algebra of finite disjoint union of subintervals of I whose extremes are in the set C of points of continuity of f. We call this measure the *point variation measure* of f, and we denote it by $p\mu_f$.

We remarked in 8.2.3 that every function $f : I \to \mathbb{K}$ of locally bounded variation gives a locally finite Radon measure $\mu_f = df$, and that this measure depends only on left and right limits of f, and is insensible to the values of f on its points of discontinuity. Accordingly, the total variation of this measure must depend only on left and right limits of f. On the other hand the variation $Vf([a,b])$ introduced in the previous sections depends strongly on all values of f, including those on discontinuity points, and we won't have in general $Vf(J) = |\mu_f|(J)$, not even for intervals with extremes at points of continuity of f, as the following trivial example shows:

Exercise 8.3.1. For $f : \mathbb{R} \to \mathbb{R}$ the characteristic function of $\{0\}$ find

$$T(x) = T_{-\infty}(x) = Vf(]-\infty, x]),$$

write f as the difference of two increasing functions A and B and describe μ_f, $|\mu_f|$ and $p\mu_f$.

Solution. $T(x) = 0$ for $x < 0$, $T(0) = 1$, $T(x) = 2$ for $x > 0$; we have $A = (T + f)/2 = \chi_{[0,\infty[}$, $B = (T - f)/2 = \chi_{]0,\infty[}$, $\mu_A = \mu_B = \delta_0$, unit mass at 0, and the measure is $\mu_f = \delta_0 - \delta_0 = 0 (= |\mu_f|)$. The point variation measure is $p\mu_f = \mu_A + \mu_B = \mu_T = 2\delta_0$. □

Also, $|\mu_f|$ is a measure, and Vf is not a measure on the algebra of plurintervals contained in I, unless f is continuous, as seen above. The measure $|\mu_f|$ is the *essential variation* of f.

We now compare the measures $p\mu_f$ and $|\mu_f|$:

PROPOSITION. *Let I be an interval of $\mathbb{R}$, and let $f \in BV_{\rm loc}(I)$. For every $x \in I$ let*

$$w(x) = |f(x^-) - f(x)| + |f(x) - f(x^+)| - |f(x^+) - f(x^-)|;$$

(notice that $w(x) \geq 0$ for every $x \in I$, and $w(x) = 0$ if and only if $f(x)$ belongs to the segment $[f(x^-), f(x^+)]$). Let μ_w be the purely atomic measure $\mu_w(E) = \sum_{x\in E} w(x)$. Then for every Borel subset E of I we have

$$p\mu_f(E) = \mu_w(E) + |\mu_f|(E)|.$$

Consequently the point variation and essential variation of f coincide if and only if $f(x) \in [f(x^-), f(x^+)]$ for every $x \in I$.

The proof is in exercise ??.??. Notice only that if f is right continuous, or left continuous, or midpoint continuous then the essential variation and the point variation coincide.

EXERCISE 8.3.2. Find a formula for $T(x) = T_0 f(x)$, where $f(x) = x - [x]$ is the fractional part of x, and plot the graph of T. Write the measure μ_f as a positive and negative part, find a Hahn decomposition for μ_f. Is the measure defined on all of $\mathcal{B}_1 = \mathcal{B}(\mathbb{R})$?

Solution. The function T will be right continuous, and with a jump of 1 at every integer, since f is right–continuous with a jump of -1 at every integer. For $x \in [0, 1[$ we have $f(x) = x$, and for $x \in [n, n+1[$, $n \in \mathbb{Z}$, we have of course $T(x) = T(n) + (x - n)$; and $T(n) = 2n$, as easily seen by induction. Then a formula which describes T is for instance $T(x) = x + [x]$ (the plot, very easy, is omitted). We have $A(x) = (T + f)(x)/2 = x$, $B(x) = (T - f)(x)/2 = [x]$. The measure $\mu_A = \mu^+$ is the Lebesgue measure, $\mu_B = \mu^- = \sum_{n\in\mathbb{Z}} \delta_n$, with δ_n the Dirac measure at n. A Hahn decomposition is $\mathbb{R} \smallsetminus \mathbb{Z}$, positive set, and $Q = \mathbb{Z}$, negative set. Since $\mu_A(\mathbb{R}) = \infty$ and $\mu_B(\mathbb{R}) = \infty$, μ cannot be defined on $\mathcal{B}_1$. We have

$$\mu(E) = \mu_A(E) - \mu_B(E) = m(E) - \mathrm{Card}(E \cap \mathbb{Z}) \quad \text{for every bounded Borel set } E \in \mathcal{B}_1^*.$$

$\square$

EXAMPLE 8.3.3. Since functions of bounded variation are differentiable a.e., a nowhere differentiable function will have infinite variation on every non degenerate interval; in particular, continuous nowhere differentiable functions are of this sort.

8.3.4. *Functions of class C^1.* If a function f is of class C^1 on a compact interval $[a, b]$, then we have the formula

$$Vf([a, b]) = \int_a^b |f'(t)|\, dt$$

(formula for the length of a curve, see e.g Analisi Due, 3.4.1). It follows that if $f \in C^1(I)$, with I not necessarily compact, then

$$Vf(I) = \int_I |f'(t)|\, dt \quad \text{(finite or infinite)}.$$

In particular an $f \in C^1(I)$ is always in $BV_{\rm loc}(I)$, and is in $BV(I)$ iff $f' \in L^1(I)$.

EXERCISE 8.3.4. Prove that $f(x) = \sin x/x$ is not in $BV(\mathbb{R})$ in two ways: by directly exhibiting a sequence $x_0 < x_1 < \dots$ such that $\sum_{k=1}^\infty |f(x_k) - f(x_{k-1})| = \infty$, and by proving that $f' \notin L^1(\mathbb{R})$.

Prove that $g(x) = \sin^2 x/x^2$ is in $BV(\mathbb{R})$ in two ways: by proving that $g' \in L^1(\mathbb{R})$, and also in the following way: between two consecutive zeros $a_k < a_{k+1}$ the function g has a unique local maximum $g(c_k)$; express $Vg(\mathbb{R})$ in terms of a series built with the $g(c_k)$, and prove that this series converges (do NOT try to compute explicitly neither c_k nor $g(c_k)$!).

Solution. The obvious attempt is to take $x_0 = 0$ and $x_k = \pi/2+k\pi$; we have $f(x_k) = (-1)^k/(\pi/2+k\pi)$ and

$$\sum_{k=1}^m |f(x_k) - f(x_{k-1})| \geq \sum_{k=2}^m \left| \frac{(-1)^k}{\pi/2 + k\pi} - \frac{(-1)^{k-1}}{\pi/2 + (k-1)\pi} \right| = \sum_{k=2}^m \left(\frac{1}{\pi/2 + k\pi} + \frac{1}{\pi/2 + (k-1)\pi} \right) \geq$$

$$\geq \sum_{k=2}^m \frac{1}{\pi/2 + k\pi};$$

clearly the last sum diverges to ∞ as $m \to \infty$.

For $x \neq 0$ we have

$$f'(x) = \frac{x\cos x - \sin x}{x^2} = \frac{\cos x}{x} - \frac{\sin x}{x^2};$$

and $\sin x/x^2 \in L^1([1,\infty[)$, but $\cos x/x \notin L^1([1,\infty[)$.

We have, for $x \neq 0$:

$$g'(x) = \frac{2\sin x\cos x\, x^2 - 2x\, \sin^2 x}{x^4} = \frac{\sin(2x)}{x^2} - 2\frac{\sin^2 x}{x^3},$$

a function that is clearly in $L^1(\mathbb{R} \smallsetminus [-1,1])$, hence in $L^1(\mathbb{R})$, being continuous.

The zeroes of g are clearly $a_k = k\pi$, $k \in \mathbb{Z} \smallsetminus \{0\}$. We consider the function on $[0,\infty[$ (it is an even function). Writing

$$g'(x) = 2\frac{\sin x}{x^3}(x\cos x - \sin x),$$

the sign of the derivative is given by $\sin x(x\cos x - \sin x)$; in $]k\pi,(k+1)\pi[$ the derivative is zero at only one point, the solution c_k in this interval of the equation $\operatorname{cotan} x = 1/x$ (see at the end for a more detailed proof); since $f(k\pi) = f((k+1)\pi) = 0$ and $f(x) > 0$ in the interior, c_k is the only point of absolute maximum of f in the interval, and f is increasing in $[k\pi, c_k]$ and decreasing in $[c_k,(k+1)\pi]$. In every interval $[k\pi,(k+1)\pi]$ we then have $Vg([k\pi,(k+1)\pi]) = 2g(c_k)$; and $Vg([0,\pi]) = 1$ so that

$$Vg([0,\infty[) = 1 + \sum_{k=1}^{\infty} 2g(c_k);$$

now we have $g(c_k) = \sin^2 c_k/(c_k)^2 \leq 1/(c_k)^2 \leq 1/(k^2\pi^2)$. Then

$$Vg(\mathbb{R}) = 2Vg([0,\infty[) = 2 + \sum_{k=1}^{\infty} 4\, g(c_k) \leq 2 + \frac{4}{\pi^2}\sum_{k=1}^{\infty}\frac{1}{k^2} = 2 + \frac{2}{3}.$$

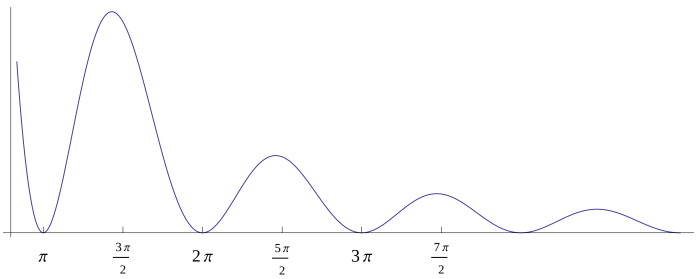

FIGURE 5. Plot of g (not on scale), with $5\pi/6 < x < 5\pi$.

REMARK. Let's give a more precise study of the sign of $\sin x(x\cos x - \sin x)$ in $]k\pi;(k+1)\pi[$; since $\sin^2 x > 0$ in this interval we have, dividing both sides by $x\,\sin^2 x$:

$$\sin x(x\cos x - \sin x) > 0 \iff \operatorname{cotan} x > 1/x \qquad x \in]k\pi,(k+1)\pi[;$$

it is plain that the set of these x is $]k\pi, c_k[$, with $k\pi < c_k < \pi/2 + k\pi$.

An ingenious application of methods of complex analysis, using fomulae related to the zeroes of $\operatorname{cotan} z - 1/z$, allows the exact computation of Vg, which if I remember correctly is $e^2 - 5$.

□

EXERCISE 8.3.5. (*Not entirely easy*) All functions considered here are real valued. Prove that a function f belongs to $L^2([0,1])$ if and only if $f \in L^1([0,1])$, and there is an increasing function $g : [0,1] \to \mathbb{R}$ such that, for $0 \leq x_1 < x_2 \leq 1$:

$$\left|\int_{x_1}^{x_2} f(t)\,dt\right|^2 \leq (g(x_2) - g(x_1))\,(x_2 - x_1)$$

(for the necessity part, use Cauchy–Schwarz inequality...).

Solution. Put $x_1 = x$ and $x_2 = x + h$ with $h > 0$; dividing both sides of the inequality by h^2 we get

$$\left|\frac{1}{h}\int_x^{x+h} f(t)\,dt\right|^2 \le \frac{g(x+h)-g(x)}{h};$$

as $h \to 0^+$ we have that $\int_x^{x+h} f(t)\,dt/h$ converges for a.e $x \in [0,1]$ to $f(x)$, so the left hand–side converges a.e. to $(f(x))^2$, while the right hand side tends a.e to the derivative $g'(x)$ of g. Then

$$(f(x))^2 \le g'(x) \quad \text{for a.e. } x \in [0,1];$$

integrating we get

$$\int_0^1 (f(x))^2\,dx \le \int_0^1 g'(x)\,dx \le g(1) - g(0).$$

Then $f \in L^2([0,1])$. Conversely, assuming $f \in L^2([0,1]$ a natural candidate for g is $g(x) = \int_0^x (f(t))^2\,dt$. To get the inequality we apply the Cauchy–Schwarz inequality to functions f and 1 on the interval $[x_1, x_2]$:

$$\left|\int_{x_1}^{x_2} f(t)\,dt\right| \le \int_{x_1}^{x_2} |f| \le \left(\int_{x_1}^{x_2} f^2\right)^{1/2} (x_2 - x_1)^{1/2};$$

squaring both sides we get the required inequality. □

Solution. (*Of exercise 7.2.7*) If X is the disjoint union of A, B with A null for ν and B null for μ then, given a partition $(A_j)_{1\le j\le m}$ if A into measurable sets, and a partition $(B_k)_{1\le k\le n}$ of B into measurable sets we have that the join of these partitions is a partition of X so that:

$$\sum_{j=1}^m |(\mu+\nu)(A_j)| + \sum_{k=1}^n |(\mu+\nu)(B_k)| \le |\mu+\nu|(X);$$

but since $\nu(A_j) = \mu(B_k) = 0$ the left–hand side is

$$\sum_{j=1}^m |\mu(A_j)| + \sum_{k=1}^n |\nu(B_k)|;$$

and the supremum of these sums as the partitions vary is $|\mu|(A) + |\nu|(B) = |\mu|(X) + |\nu|(X)$; we have proved that $|\mu|(X) + |\nu|(X) \le |\mu+\nu|(X)$, and since the reverse inequality is always true we get equality. The proof just given for X is clearly adaptable to any $E \in \mathcal{M}$.

(ii) Since $\lambda_j \ll \mu$ and $\nu_j \perp \mu$ we also have $\lambda_j \perp \nu_j$, so that $\|\mu_j\| = \|\lambda_j\| + \|\nu_j\|$, by what just proved. Then convergence of the series $\sum_{j=0}^\infty \|\mu_j\|$ is equivalent to the convergence of the two series $\sum_{j=0}^\infty \|\lambda_j\|$ and $\sum_{j=0}^\infty \|\nu_j\|$; and since $d|\lambda_j| = |\rho_j|\,d\mu$ and $\|\lambda_j\| = \|\rho_j\|_1$, everything follows easily. □

8.3.5. *More exercises.*

EXERCISE 8.3.6. Given $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ we have defined

$$A_r f(x) = \fint_{B(x,r]} f(y)\,dy,$$

Prove that if $f \in L^1(\mathbb{R}^n)$ then $A_r f \in L^1(\mathbb{R}^n)$, and compute $\int_{\mathbb{R}^n} A_r f(x)\,dx$ in terms of $\int_{\mathbb{R}^n} f$ (hint: write $A_r f(x) = \int_{\mathbb{R}^n} f(x+y)\,\chi_{rB}(y)\,dy/m(rB)$ and use Fubini–Tonelli's theorem. . .)

Solution. Consider the function $F : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{K}$ given by $(x, y) \mapsto f(x+y)\,\chi_{rB}(y)$; if f is Borel measurable this function is Borel measurable as a function of $\mathbb{R}^n \times \mathbb{R}^n$ into $\mathbb{K}$ ($(x, y) \mapsto x+y$ is continuous, hence measurable, and so $(x, y) \mapsto f(x+y)$ is measurable if f is Borel measurable; clearly $(x, y) \mapsto \chi_{rB}(y)$ is measurable; and even if f is only Lebesgue measurable still the function F is Lebesgue measurable, essentially because the inverse image under addition of a set of $n-$dimensional measure zero in $\mathbb{R}^n$ is a set of $2n-$dimensional measure zero in $\mathbb{R}^n \times \mathbb{R}^n$). We prove that $F \in L^1(\mathbb{R}^n \times \mathbb{R}^n)$. In fact, integrating first in the $x-$variable

$$\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} |F(x,y)|\,dx\right)dy = \int_{\mathbb{R}^n} \chi_{rB}(y)\left(\int_{\mathbb{R}^n} |F(x,y)|\,dx\right) = \int_{\mathbb{R}^n} \chi_{rB}(y)\left(\int_{\mathbb{R}^n} |f(x+y)|\,dx\right)dy =$$
$$\int_{\mathbb{R}^n} \|f\|_1 \chi_{rB}(y)\,dy = m(rB)\,\|f\|_1 < \infty.$$

By Tonelli's theorem F belongs to $L^1(\mathbb{R}^n \times \mathbb{R}^n)$. Then Fubini's theorem says that $x \mapsto \int_{\mathbb{R}^n} F(x,y)\,dy$, which is exactly $m(rB)\,A_r f(x)$, defines a.e (in our case everywhere) a function belonging to $L^1(\mathbb{R}^n)$ (thus proving that $A_r f \in L^1(\mathbb{R}^n)$) and that

$$\int_{\mathbb{R}^n\times\mathbb{R}^n} F(x,y)\,dxdy = \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} F(x,y)\,dy\right)dx \left(= m(rB)\int_{\mathbb{R}^n} A_r f(x)\,dx\right);$$

exchanging the order of integration in the integral of F we repeat the above calculation without the absolute value on f, and get

$$\int_{\mathbb{R}^n\times\mathbb{R}^n} F(x,y)\,dxdy = m(rB)\int_{\mathbb{R}^n} f(u)\,du;$$

since this equals $m(rB)\int_{\mathbb{R}^n} A_r f(x)\,dx$ we get

$$\int_{\mathbb{R}^n} A_r f(x)\,dx = \int_{\mathbb{R}^n} f(u)\,du.$$

□

Exercise 8.3.7. Let $F : \mathbb{R} \to \mathbb{R}$ be defined by

$$F(x) = \frac{1}{(1-x)^2} \quad \text{if } x < 0; \quad F(x) = \sin x \quad \text{if } 0 \le x < 2\pi; \quad F(x) = -(x-2\pi)^2 \quad \text{if } 2\pi \le x.$$

(i) Plot the graph of F.
(ii) Show that f has bounded variation on $]-\infty, x]$, for every $x \in \mathbb{R}$; describe $TF(x) = VF(]-\infty, x])$; plot it, and plot two increasing function A, B with $A - B = F$.
(iii) Prove that there is a unique signed measure $\nu : \mathcal{B}_1 \to \tilde{\mathbb{R}}$ such that $\nu(]a,b]) = F(b) - F(a)$ for every $a, b \in \mathbb{R}$, $a < b$.
(iv) Find a Hahn decomposition for ν.
(v) Find the Radon– Nikodym decomposition of ν with respect to Lebesgue measure m.
(vi) Compute $\int_{\mathbb{R}} u(x)\,d\nu(x)$, where $u(x)$ is so defined: $u(0) = 1/2$, $u(\pi) = 1$, $u(2\pi) = 3$; $u(x) = 1$ for $x < 0$, $u(x) = 3$ for $0 < x < \pi$, $u(x) = -1$ for $\pi < x < 2\pi$, $u(x) = 2$ for $2\pi < x < 3\pi$, $u(x) = 0$ for $x \ge 3\pi$.

Solution. (i) OK.

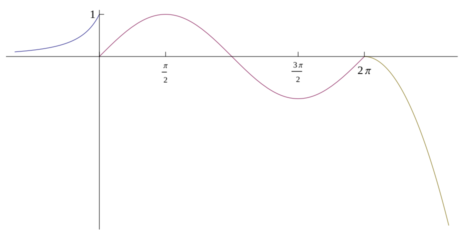

Figure 6. Graph of F.

(ii) First notice that since F is right continuous, and also left continuous but at $x = 0$, with a jump $\sigma_F(0) = F(0) - F(0^-) = -1$, T will be everywhere continuous but at 0, where will be right continuous with a jump of 1. We have $T(x) = 1/(1-x)^2$ for $x \le 0$ (for monotone functions, the variation on an interval is the absolute value of the increment), and hence $T(0^-) = 1$; then $T(0) = T(0^+) = 2$; in $[0, \pi/2]$ the function F is increasing, so that $T(x) = T(0) + \sin x - \sin 0 = 2 + \sin x$. In $[\pi/2, 3\pi/2]$ the function F is decreasing, and hence $T(x) = T(\pi/2) + F(\pi/2) - F(x) = 3 + 1 - \sin x = 4 - \sin x$; in $[3\pi/2, 2\pi]$ F increases again, and $T(x) = T(3\pi/2) + \sin x - \sin(3\pi/2) = 6 + \sin x$. Finally in $[2\pi, +\infty[$ the function F is decreasing and we then have $T(x) = T(2\pi) + F((2\pi) - F(x) = 6 + (x-2\pi)^2$. We now compute $A = (T+F)/2$:

$$A(x) = \frac{1}{(1-x)^2} \quad x \in]-\infty, 0[;\ A(x) = 1 + \sin x \quad x \in [0, \pi/2];\ A(x) = 2 \quad x \in [\pi/2, 3\pi/2];$$
$$A(x) = 3 + \sin x \quad x \in [3\pi/2, 2\pi];\ A(x) = 3 \quad x \in [2\pi, \infty[.$$

Computation of $B = (T-F)/2$:

$$B(x) = 0 \quad x \in]-\infty, 0[;\ B(x) = 1 \quad x \in [0, \pi/2];\ B(x) = 2 - \sin x \quad x \in [\pi/2, 3\pi/2];$$

$$B(x) = 3 \quad x \in [3\pi/2, 2\pi]; B(x) = 3 + (x - 2\pi)^2 \quad x \in [2\pi, \infty[.$$

(iii) We see that A is bounded; hence the measure μ_A will be finite, $\mu_A(\mathbb{R}) = A(\infty) - A(-\infty) = 3$, and $\nu = \mu_A - \mu_B$ can be defined.

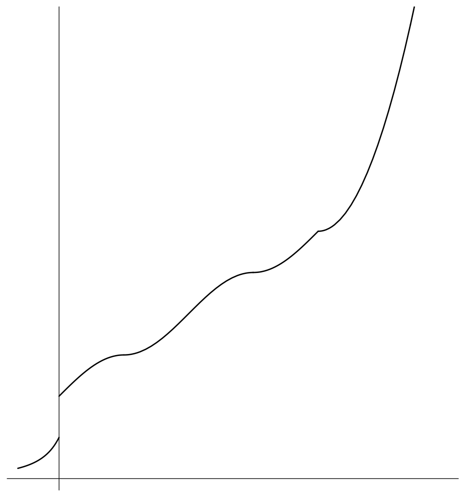

FIGURE 7. Graph of T.

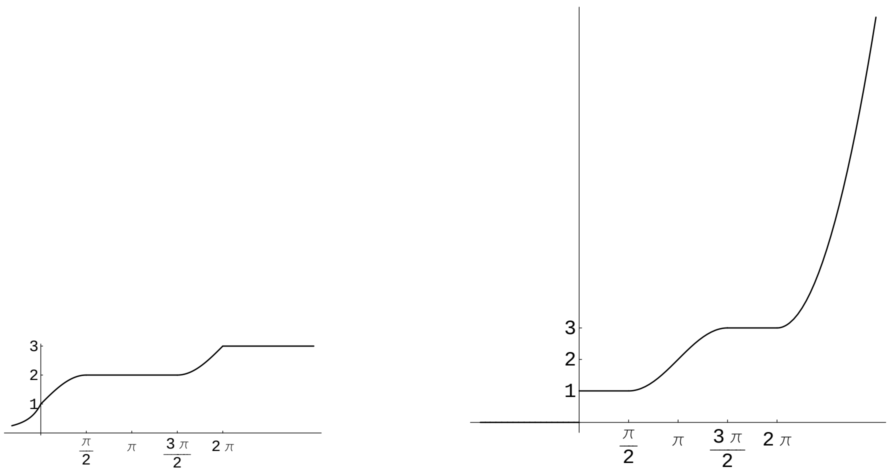

FIGURE 8. Graphs of A, B; B, on the right, has a jump of 1 at 0.

(iv) A set is positive for ν where A is increasing, negative where B is increasing; it is easy to see that $P =]-\infty, 0[\cup]0, \pi/2] \cup [3\pi/2, 2\pi[$,is a positive set for ν, its complement $Q = \mathbb{R} \setminus P = \{0\}\cup]\pi/2, 3\pi/2[\cup[2\pi, \infty[$ is a negative set.

(v) For the absolutely continuous part we compute the derivative F', which exists everywhere but at $x = 0, 2\pi$; this derivative is also continuous where it exists:

$$F'(x) = \frac{2}{(1-x)^3} \quad x \in]-\infty, 0[;\ F'(x) = \cos x \quad x \in]0, 2\pi[;\ F'(x) = -2(x - 2\pi) \quad x \in]2\pi, \infty[;$$

Then the absolutely continuous part is $d\lambda = F'(x)\, dm$; notice that $f(x) = F'(x)$ is integrable in the extended sense (the integral of the positive part is clearly finite). The singular part is $-\delta$. A pair of disjoint Borel sets which support these two measures is, for instance, $\mathbb{R} \setminus \{0\}$ for λ, and $\{0\}$ for the singular part.

(vi) Finally the integral is:

$$\int_{\mathbb{R}} u(x)\, d\lambda(x) + \int_{\mathbb{R}} u(x)\, d(-\delta(x)) = \int_{\mathbb{R}\setminus\{0\}} u(x)\, F'(x)\, dm(x) - u(0) =$$

$$-\frac{1}{2}+\int_{-\infty}^{0}\frac{2}{(1-x)^3}\,dx+\int_0^{\pi}3\cos x\,dx+\int_{\pi}^{2\pi}(-1)\cos x\,dx+\int_{2\pi}^{3\pi}2(-2(x-2\pi))\,dx=$$
$$-\frac{1}{2}+1+0+0-2\pi^2=\frac{1}{2}-2\pi^2.$$

□

8.4. Absolutely continuous functions. A function $F : I \to \mathbb{K}$ where I is an interval of $\mathbb{R}$ is said to be *locally absolutely continuous* in I if $F'(x)$ exists a.e. in I, it is in $L^1_{\text{loc}}(I)$, and for every pair $a, b \in I$ we have $F(b) - F(a) = \int_a^b F'(x)\,dx$; it is said to be *absolutely continuous* on I if it is locally absolutely continuous, and moreover $F' \in L^1(I)$. We denote by $AC_{\text{loc}}(I)$ the set of locally absolutely continuous functions on I, by $AC(I)$ the subset of globally absolutely continuous functions on I; plainly these are vector spaces of functions. As a reminder:

locally absolutely continuous functions are those a.e. differentiable functions which are the integral of their derivative.

They clearly are continuous functions, and may alternatively be described as those functions $F : I \to \mathbb{K}$ of the form
$$F(x) = F(c) + \int_c^x f(t)\,dt,$$
for some $c \in I$ and some $f \in L^1_{\text{loc}}(I)$ (or $f \in L^1(I)$ in the case of globally a.c. functions).

Remark. The space $AC(I)$ is, in Functional Analysis, denoted $W^{1,1}(I)$ and similarly $AC_{\text{loc}}(I)$ is denoted $W^{1,1}_{\text{loc}}(I)$ (Sobolev spaces).

Functions in $AC_{\text{loc}}(I)$ are those functions in $BV_{\text{loc}}(I)$ whose associated measure is absolutely continuous with respect to Lebesgue measure:

. *For every interval I we have $AC_{\text{loc}}(I) \subseteq BV_{\text{loc}}(I)$ and $AC(I) = BV(I) \cap AC_{\text{loc}}(I)$. Moreover for every function $F \in AC(I)$ we have*
$$VF(I) = \int_I |F'(x)|\,dx = \|F'\|_1.$$

Proof. Clearly the second statement implies the first. Given any subdivision $x_0 < \cdots < x_m$ of I we have
$$\sum_{k=0}^{m}|F(x_k)-F(x_{k-1})| \le \sum_{k=0}^{m}\left|\int_{x_{k-1}}^{x_k}F'(x)\,dx\right| \le \sum_{k=0}^{m}\int_{x_{k-1}}^{x_k}|F'(x)|\,dx = \int_{x_0}^{x_m}|F'(x)|\,dx \le \int_I |F'(x)|\,dx.$$
then:
$$VF(I) \le \int_I |F'(x)|\,dx;$$
this proves that F is of bounded variation. For the opposite inequality recall that if $d\lambda = F'(x)\,dx$ with $F' \in L^1(I)$ then $d|\lambda| = |F'(x)|\,dx$ (7.2.1); and by 8.3.3, since F is continuous its point variation and essential variation coincide. □

8.4.1. *The (ε, δ) condition for absolute continuity.* We have seen that if $\mu : \mathcal{M} \to [0, \infty]$ is a positive measure, and $\nu : \mathcal{M} \to \mathbb{K}$ is a finite measure, then the condition $\nu \ll \mu$, meaning that $\mu(E) = 0$ implies $\nu(E) = 0$, may also be expressed as: given $\varepsilon > 0$ there is $\delta > 0$ such that $\mu(E) \le \delta$ implies $|\nu(E)| \le \varepsilon$ (7.2.9 or 7.1.9). This condition may be used to characterize locally absolutely continuous functions among continuous functions, with no differentiability involved. A sequence $([a_j, b_j])_{1\le j\le n}$ of compact intervals is said to be non overlapping, or almost disjoint, if the interiors of two distinct intervals are disjoint, equivalently two distinct intervals intersect on a set of Lebesgue measure 0.

Definition. Let $f : I \to \mathbb{K}$ be a function, and let $[a, b]$ be a compact subinterval of I. We say that f satisfies the $\varepsilon - \delta$ condition for absolute continuity on $[a, b]$ (briefly: satisfies the (ε, δ)AC condition on $[a, b]$) if for every $\varepsilon > 0$ there is $\delta > 0$ such that for every sequence $([a_j, b_j])_{1\le j\le n}$ of non overlapping subintervals of $[a, b]$, with $\sum_{j=1}^n (b_j - a_j) \le \varepsilon$, we have $\sum_{j=1}^n |f(b_j) - f(a_j)| \le \varepsilon$.

It is clear that functions that satisy the (ε, δ)AC condition are continuous, and are a vector space of functions. Let's prove :

Lemma. *Let $f : [a, b] \to \mathbb{K}$ be a function. The following are equivalent:*

(i) *f satisfies the $(\varepsilon,\delta)AC$ condition on $[a,b]$.*
(ii) *For every $\varepsilon>0$ there is a $\delta>0$ such that for every sequence $([a_j,b_j])_{1\le j\le n}$ of non overlapping subintervals of $[a,b]$, with $\sum_{j=1}^n(b_j-a_j)\le\delta$, we have $\sum_{j=1}^n Vf([a_j,b_j])\le\varepsilon$.*

In other words, if (i) holds, then the total variation function $T(x)=Vf([a,x])$ of f satisfies the $(\varepsilon,\delta)AC$ condition; in particular this function is continuous, hence bounded on $[a,b]$.

Proof. Given $\varepsilon>0$ let $([a_j,b_j])_{1\le j\le n}$ be a sequence of non overlapping intervals such that $\sum_{j=1}^n(b_j-a_j)\le\delta$. For every $j\in\{1,\dots,n\}$ let $a_j=x(j,0)<x(j,1)<\cdots<x(j,n(j))=b_j$ a subdivision of $[a_j,b_j]$. Then $[x(j,k-1),x(j,k)]_{1\le j\le n,\,1\le k\le n(j)}$ is a sequence of non overlapping intervals with

$$\sum_{j=1}^n\left(\sum_{k=1}^{n(j)}(x(j,k)-x(j,k-1))\right)=\sum_{j=1}^n(b_j-a_j)\le\delta \text{ hence}\sum_{j=1}^n\left(\sum_{k=1}^{n(j)}|f(x(j,k))-f(x(j,k-1))|\right)\le\varepsilon.$$

Taking suprema as $x(j,k)_{0\le k\le n(j)}$ varies in the set of all subdivisions of $[a_j,b_j]$, for every j, we get

$$\sum_{j=1}^n Vf([a_j,b_j])\le\varepsilon.$$

Since $Vf([a_j,b_j])\ge|f(b_j)-f(a_j)|$, (ii) implies (i) is trivial. ☐

8.4.2. *Characterization of local absolute continuity.*

Proposition. *Let I be an interval of $\mathbb{R}$ and let $F:I\to\mathbb{K}$ be a function. The following are equivalent:*

(i) *On every compact subinterval of I the function F verifies the $(\varepsilon,\delta)AC$ condition.*
(ii) *F is locally absolutely continuous, that is F' exists a.e on I, and*

$$F(b)-F(a)=\int_a^b F'(x)\,dx\quad \textit{for every } a,b\in I.$$

Proof. We can of course assume that I is compact. (i) implies (ii): by lemma 8.4.1 the function F has locally bounded variation and its total variation function is continuous; we prove that if $M\subseteq I$ has zero Lebesgue measure then $\mu_T(M)(=|\mu_F|(M))=0$; this proves $\mu_F\ll m$. Given $\varepsilon>0$, let $\delta>0$ be as in the (ε,δ)AC definition for T. Since M has $m-$measure zero, it can be covered by a countable set $(I_j)_{j\in\mathbb{N}}$ of subintervals of I, which may be assumed pairwise disjoint, such that $\sum_{j=0}^\infty(b_j-a_j)\le\delta$, where $a_j=\inf I_j$ and $b_j=\sup I_j$. Then we have, for every $n\in\mathbb{N}$, $\sum_{j=0}^n(T(b_j)-T(a_j))\le\varepsilon$ because for every n the sequence $([a_j,b_j])_{0\le j\le n}$ consists of non overlapping intervals whose lengths have a sum less than δ, hence also $\sum_{j=0}^\infty(T(b_j)-T(a_j))\le\varepsilon$; this means that $\mu_T(M)=0$, as desired. (ii) implies (i): the formula $\nu(E)=\int_E F'(x)\,dx$ defines a finite measure on the Borel subsets of I; apply 7.2.9. ☐

Exercise 8.4.1. Prove that if $g:J\to\mathbb{K}$ is Lipschitz continuous, then $g\in AC_{\rm loc}(J)$, and $|g'(x)|\le k$, if k is a Lipschitz constant for g.

Let I be an interval of $\mathbb{R}$, let $F:I\to\mathbb{R}$ be in $AC_{\rm loc}(I)$, and let $g:F(I)\to\mathbb{K}$ be Lipschitz continuous. Prove that $g\circ F\in AC_{\rm loc}(I)$ (hint: use the (ε,δ)AC condition).

Exercise 8.4.2. The identity $F(x)=x$ of $\mathbb{R}$ verifies a global (ε,δ)AC condition (simply with $\delta=\varepsilon$) but its derivative, the constant 1, is not in $L^1(\mathbb{R})$; that is F is in $AC_{\rm loc}(\mathbb{R})\setminus AC(\mathbb{R})$. This does not happen on bounded intervals. Prove that if the interval I is bounded then $F:I\to\mathbb{K}$ is in $AC(I)$ if and only if it verifies a global (ε,δ)AC condition on I.

Solution. Clearly $F(x)=F(c)+\int_c^x f(t)\,dt$ with $f\in L^1_{\rm loc}(I)$ and we have to prove that if F verifies a global (ε,δ)AC condition on I then $f\in L^1(I)$. From lemma 8.4.1 we know that the total variation function $T_cF(x)=\int_c^x|f(t)|\,dt$ verifies the same condition, in particular then it is uniformly continuous on I. Then T_c is bounded on I and (theorem on extension of uniformly continuous functions) T_c extends to a continuous function on the compact closure $\bar I$ of I) and this clearly implies that $f\in L^1(I)$ ($\int_I|f(t)|\,dt=T_c(\sup I)-T_c(\inf I)$). ☐

8.5. The devil's staircase. One wonders if the following might be true: let F be continuous of (locally) bounded variation. Is F (locally) absolutely continuous? The simple minded examples of monotone non absolutely continuous functions have jump discontinuities. Nevertheless this conjecture is false. We describe here an example. Pick a strictly decreasing sequence $1 = \delta_0 > \delta_1 > \dots$, with limit 0. We construct a decreasing sequence $[0,1] = E_0 \supseteq E_1 \supseteq \dots$ of compact sets, with $m(E_n) = \delta_n$, in this way: remove from $[0,1] = E_0$ an open interval centered at $1/2$, in such a way that the two remaining intervals, $I(0) < I(1)$, have the same length $\delta_1/2$; this is clearly possible, since $\delta_1 < \delta_0$. If $E_1 = I(0) \cup I(1)$, then E_1 is compact, and $m(E_1) = \delta_1$; next, from each interval $I(0)$, $I(1)$ remove an open interval with the same center as $I(0)$ and $I(1)$, in such a way that calling $I(0,0)$ and $I(0,1)$ the intervals obtained from $I(0)$, and $I(1,0)$ and $I(1,1)$ those obtained from $I(1)$, each has length $\delta_2/4$; so that E_2, union of these four compact intervals, has length $m(E_2) = \delta_2$. It is clear how the induction proceeds: we have $E_n = \bigcup_{c \in \{0,1\}^n} I(c)$, union of 2^n disjoint compact intervals each of length $\delta_n/2^n$ so that $m(E_n) = \delta_n$; from each of the intervals $I(c)$ we remove an open interval with the same center as $I(c)$, of length $2r_n$ such that the two remaining intervals $I(c,0) < I(c,1)$ have both length $\delta_{n+1}/2^{n+1}$ (we take $r_n = (\delta_n - \delta_{n+1})/2^{n+1}$); and E_{n+1} is the union of all these 2^{n+1} intervals. Clearly $E = \bigcap_{n=0}^{\infty} E_n$ is non empty, as the intersection of a decreasing sequence of non empty compact sets, and $m(E) = \lim_{n\to\infty} m(E_n) = 0$; we might also prove that $|E| = \mathfrak{c}$ and that E is *perfect*, i.e it has no isolated point (see 8.5.1). But we are interested in another construction. For every n let $g_n = \chi_{E_n}/\delta_n$, and put

$$f_n(x) = \int_0^x g_n(t)\,dt.$$

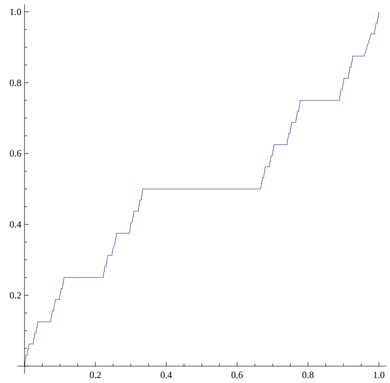

FIGURE 9. The devil's staircase (here $\delta_n = (2/3)^n$).

Then f_n is increasing, $f_n(x) = 0$ for $x \le 0$ and $f_n(x) = 1$ for $x \ge 1$. Notice that f_n is constant on each open interval that is a connected component of $\mathbb{R} \setminus E_n$; even more, if A is one such interval, then $f_n(x) = f_{n+1}(x)$ for every $x \in A$. The key observation is that for every $c \in \{0,1\}^n$ we have

$$\int_{I(c)} g_n(t)\,dt = \int_{I(c)} g_{n+1}(t)\,dt = \frac{1}{2^n} \quad \text{in fact} \quad \frac{1}{\delta_n}\frac{\delta_n}{2^n} = \frac{1}{\delta_{n+1}}\left(\frac{\delta_{n+1}}{2^{n+1}} + \frac{\delta_{n+1}}{2^{n+1}}\right) = \frac{1}{2^n},$$

so that if p is the cardinality of the set $J = \{c \in \{0,1\}^n : I(c) < A\}$ we have, for every $x \in A$:

$$f_n(x) = \sum_{c\in J} \int_{I(c)} g_n(x)\,dx = \frac{p}{2^n};\; f_{n+1}(x) = \sum_{c\in J}\int_{I(c)} g_{n+1}(x)\,dx = \frac{p}{2^n}.$$

And if $x \in I(c)$ for some $c \in \{0,1\}^n$ then f_n and f_{n+1} coincide on the interval of $\mathbb{R}\setminus E_n$ to the immediate left of $I(c)$, in particular $f_n(\min I(c)) = f_{n+1}(\min I(c))$, so that

$$|f_n(x) - f_{n+1}(x)| = \left|\int_{\min I(c)}^x (g_n(t) - g_{n+1}(t))\,dt\right| \le \int_{\min I(c)}^x |g_n(t) - g_{n+1}(t)|\,dt \le \int_{I(c)} |g_n(t) - g_{n+1}(t)|\,dt$$

for every $x \in I(c)$. The computation is easy:

$$\int_{I(c)} |g_n(t) - g_{n+1}(t)|\,dt = \int_{I(c,0)}\left(\frac{1}{\delta_{n+1}} - \frac{1}{\delta_n}\right)dt + \int_{I(c,1)}\left(\frac{1}{\delta_{n+1}} - \frac{1}{\delta_n}\right)dt + \int_{I(c)\setminus(I(c,0)\cup I(c,1))}\frac{dt}{\delta_n} =$$

$$\frac{2\delta_{n+1}}{2^{n+1}}\left(\frac{1}{\delta_{n+1}}-\frac{1}{\delta_n}\right)+\frac{1}{\delta_n}\left(\frac{\delta_n}{2^n}-\frac{2\delta_{n+1}}{2^{n+1}}\right)=\frac{1}{2^n}\left(2-2\frac{\delta_{n+1}}{\delta_n}\right)<\frac{1}{2^{n-1}}.$$

Then $\|f_{n+1}-f_n\|_\infty \le 1/2^{n-1}$ so that the sequence f_n is of finite variation in the uniform norm, and hence converges uniformly to a continuous increasing function f which is zero for $x\le 0$ and 1 for $x \ge 1$. On $\mathbb{R}\setminus E$ the derivative of this function is zero: if $x \notin E$, and n is the smallest integer for which $x\notin E_n$, then x belongs to an open interval A disjoint from E_k for all $k \ge n$, and all functions f_k, $k\ge n$ and f have the same constant value on A. Then $f'(x) = 0$ for every $x \in A$. Since $m(E) = 0$ we have $f'(x) = 0$ for a.e. $x\in\mathbb{R}$. The positive finite measure μ_f is supported by the set E, and we have $\mu_f \perp m$.

With the sequence $\delta_n = (2/3)^n$ the set E obtained is the celebrated *Cantor's ternary set*. The function f is the *Cantor's function*; its graph is called *devil's staircase*: f is a function that does its growth only on a set of length zero! Notice that $f([0,1]\setminus E)$ is a countable set: in fact $[0,1]\setminus E$ is the union of countably many open intervals, on each of which f is constant. It follows that the compact set $f(E)$ has Lebesgue measure 1: f is a continuous monotone function which maps a set of measure 0 onto a set of measure 1.

It can be proved that every set of strictly positive measure contains a non measurable subset. If $B\subseteq f(E)$ is non measurable, then $A = E\cap f^{\leftarrow}(B)$ is a Lebesgue measurable subset of $\mathbb{R}$ (it is a subset of a null set for m, so it is measurable) whose image B is non measurable.

EXERCISE 8.5.1. Prove that in fact $f(E) = [0,1]$. Prove that E has the continuum as cardinality, and that moreover no point of E is isolated.

Solution. Clearly $f([0,1]) = f(E)\cup f([0,1]\setminus E) = [0,1]$; if A is one of the open intervals that compose $[0,1]\setminus E$ its extremes are in E, and f is constant on A, so that $f(A)\subseteq f(E)$. Then $|E|\ge|f(E)| = \mathfrak{c}$, so that $|E| = \mathfrak{c}$. But a direct argument is perhaps more convincing: let $D = \{0,1\}^{\{1,2,3,\dots\}}$ be the set of all sequences of $\{0,1\}$; clearly $|D| = \mathfrak{c} = |\mathbb{R}|$. There is an obvious bijective map $\varphi : D\to E$: for every $c\in D$ consider the restriction $c|n$ of c to $\{1,\dots,n\}$; then $(I_{c|n})_{n\ge1}$ is a decreasing sequence of compact intervals, of lengths $m(I_{c|n}) = \delta_n/2^n$ with limit 0, so that their intersection contains exactly one point, which we call $\varphi(c)$. Clearly this map is injective: if $c, d\in D$ and $c\ne d$, let p be the smallest integer such that $c(p)\ne d(p)$, say $c(p) = 0$ and $d(p) = 1$; we have $\varphi(c)\in I_{c|p}$, $\varphi(d)\in I_{d|p}$ and the intervals $I_{c|p} < I_{d|p}$ have distance $2r_{p-1} = (\delta_{p-1}-\delta_p)/2^{p-1}$, so that also $\varphi(d)-\varphi(c)\ge((\delta_{p-1}-\delta_p)/2^{p-1}$. On the other hand, if $c|p = d|p$ then we have $|\varphi(d)-\varphi(c)|\le\delta_p/2^p$, since for every $n\ge p$ the intervals $I_{c|n}$ and $I_{d|n}$ are both contained in $I_{c|p} = I_{d|p}$, an interval of length $\delta_p/2^p$, and so also $\varphi(c),\varphi(d)$ are both in this interval. Then no point of E is isolated: given $\varepsilon > 0$, pick $p\ge1$ such that $\delta_p/2^p < \varepsilon$, and all infinitely many points $\varphi(d)$, with $d|p = c|p$ will be in $[\varphi(c)-\varepsilon, \varphi(c)+\varepsilon]$. □

REMARK. The set E has continuum cardinality and Lebesgue measure 0; all of its $2^{\mathfrak{c}}$ subsets have then measure 0. Knowing that $\mathcal{B}_1$ has cardinality $\mathfrak{c}$, this proves that there are Lebesgue measurable subsets of $\mathbb{R}$ that are not Borel sets.

EXERCISE 8.5.2. With f the Cantor's function, define $g : [0,1]\to\mathbb{R}$ by $g(x) = x + f(x)$. Then

(i) g induces a homeomorphism of $[0,1]$ onto $[0,2]$.
(ii) $m(g(E)) = m(g([0,1]\setminus E)) = 1$.

Let $h : [0,2]\to[0,1]$ be the inverse homeomorphism $g^{-1} : [0,2]\to[0,1]$; as every set of strictly positive measure, $g(E)$ contains a non Lebesgue measurable subset B; if $A = h(B)$, then $\chi_A\circ h : [0,2]\to\mathbb{R}$ is not Lebesgue measurable, although χ_A and h are both Lebesgue measurable (and h even continuous).

Having measure zero, clearly E contains no non–degenerate intervals. Since intervals are the connected subsets of $\mathbb{R}$, the connected components of E are the singletons. That is, E is a compact *totally disconnected* set, one that contains no connected subset of more than one point. This feature can be shared also by compact sets of positive measure, like e.g. the one constructed in 2.7.4. But we can repeat the preceding construction with a decreasing sequence $\delta_n > \delta_{n+1}$ with limit $\delta > 0$, since the proofs above given for the topological properties of E do not depend on $\delta = 0$, and prove that E is perfect and totally disconnected: given $\varphi(c) < \varphi(d)$, there is an interval of $\mathbb{R}\setminus E$ contained in $]\varphi(c),\varphi(d)[$.

8.5.1. *Continuous and discrete measures.* For functions of locally bounded variation on $\mathbb{R}$ (or on an interval I of $\mathbb{R}$), $F : I\to\mathbb{K}$ we denote by $dF = d\mu_F$ the measure μ_F; we have already used this symbol for positive measures when F is increasing. The only care to be observed is that we can write $dF = F'(x)\,dx$ *if and only if F is locally absolutely continuous.* In all other cases dF is of the form $dF = F'(x)\,dx + d\nu$, where ν is a measure singular with respect to Lebesgue measure dx. The function $F'(x)$ is in $L^1_{\rm loc}(I)$ and coincides a.e. with the classical derivative of F. The singular part can be further decomposed. Call *continuous* a locally finite Radon measure such that every singleton has zero measure, and call *discrete* a measure of this form: $\mu(E) = \sum_{x\in E} c(x)$, equivalently $\mu(E) = \sum_{x\in\mathbb{R}} c(x)\,\delta_x(E)$, where $c : \mathbb{R} \to \mathbb{K}$ is a function which is summable on every bounded subset E of $\mathbb{R}$. Such a measure is supported by the cozero–set of c, necessarily countable (its bounded parts are countable, since the sums over bounded sets are finite), and hence singular with respect to any continuous measure, for which all countable sets are null sets. For a function $F \in BV_{\rm loc}(\mathbb{R})$ the discrete part of $dF = \mu_F$ is of course

$$\sum_{x\in\mathbb{R}} (F(x^+) - F(x^-))\,\delta_x.$$

Sometimes the continuous part of a measure that is singular with respect to Lebesgue measure is called the *Cantor part* of the measure; some authors call *diffuse* these measures.

8.5.2. *Integration by parts.* For increasing functions $F, G : I \to \mathbb{R}$ we proved (see 5.1.6) that for $a < b$, $a, b \in I$ we have

$$\int_{]a,b]} F(x^-)\,dG(x) = F(b^+)G(b^+) - F(a^+)G(a^+) - \int_{]a,b]} G(x^+)\,dF(x);$$

the bilinearity of this formula in F and G immediately implies that it holds also when F and G are of locally bounded variation (for real F, G write $F = A - B$ and $G = C - B$ with A, B, C, D increasing right continuous). Let us see some applications. Unless explicitly stated otherwise, functions of locally bounded variation are taken right continuous.

Exercise 8.5.3. (*The Abel–Dirichlet's criterion for convergence of improper integrals*) Let $f \in L^1_{\rm loc}([0,\infty[)$, and let $g \in BV([0,\infty[)$. Assume that $F(x) = \int_0^x f(t)\,dt$ is bounded, and that $\lim_{x\to\infty} g(x) = 0$. Then

$$\lim_{x\to\infty}\int_0^x f(t)\,g(t)\,dt \quad \text{exists in } \mathbb{K}, \text{ in other words the generalized integral } \int_0^{\uparrow\infty} f(t)\,g(t)\,dt \text{ is finite.}$$

Solution. By the preceding formula (notice that F is continuous)

$$\int_0^x f(t)\,g(t)\,dt = F(x)g(x) - \int_0^x F(t)\,dg(t);$$

now F is continuous and bounded, and hence is in $L^1(d|g|)$, so that $\lim_{x\to\infty}\int_0^x F(t)\,dg(t) = \int_{[0,\infty[} F\,dg$ exists finite (essentially dominated convergence: if x_n is any sequence that diverges to ∞, the functions $F(t)\chi_{[0,x_n]}(t)\,\mathrm{sgn}\,g(t)$ are dominated by $\|F\|_\infty$, a constant which is in $L^1(d|g|)$ since $d|g|([0,\infty[) < \infty$). When $x \to \infty$, since F is bounded and g tends to 0 we get

$$\lim_{x\to\infty}\int_0^x f(t)\,g(t)\,dt = -\int_0^\infty F(t)\,dg(t).$$

□

Exercise 8.5.4. Let $f \in L^1([0,\infty[)$ be monotone. Prove that then $\lim_{x\to\infty} xf(x) = 0$.

Solution. Since f is monotone, $\lim_{x\to\infty} f(x)$ exists in $\tilde{\mathbb{R}}$; since $f \in L^1([0,\infty[)$, this limit must be 0. Then either f is positive and decreasing, or f is negative and increasing; by changing the sign we can always assume the first case. Then f is of bounded variation on $[0,\infty[$, with $Vf([0,\infty[) = f(0)$; and

$$\int_0^x f(t)\,dt = f(x)\,x - \int_0^x t\,df(t) = x\,f(x) + \int_0^x t\,(-df(t));$$

since $x\,f(x) \ge 0$ we have $\int_0^x t\,(-df(t)) \le \int_0^x f(t)\,dt$; moreover $x \mapsto \int_0^x t\,(-df(t)), x \mapsto \int_0^x f(t)\,dt$ are both increasing, and by the hypothesis $f \in L^1([0,\infty[)$ the limit $\lim_{x\to\infty}\int_0^x f(t)\,dt = \int_0^\infty f(t)\,dt$ is finite, so that both limits are finite. But then also $\lim_{x\to\infty} x\,f(x) \ge 0$ is finite, and this limit must be zero (otherwise $f \notin L^1([0,\infty[)$). Notice that we also get

$$\int_0^\infty f(x)\,dx = \int_0^\infty x\,(-df(x)) \quad \text{for every decreasing function in } L^1([0,\infty[).$$

□

Exercise 8.5.5. Assume that $F, G \in AC_{\rm loc}(I)$. Then for any $a, b \in I$ integration by parts formula gives

$$\int_a^b (F'G + FG')(x)\,dx = F(b)\,G(b) - F(a)\,G(a).$$

Deduce from it that the product of two locally AC functions is locally AC, and that

$$(F\,G)'(x) = F'(x)\,G(x) + F(x)\,G'(x) \quad \text{for a.e. } x \in I.$$

Conversely, prove directly that $FG \in AC_{\rm loc}(I)$ and deduce from this fact the integration by parts formula for locally AC functions.

Solution. First part: immediate: the very formula is the assert.

Second part: fix a compact subinterval J of I. Given $\varepsilon > 0$ find $\delta > 0$ such that for every non–overlapping family of subintervals $([a_j, b_j])_{1\le j\le n}$ of J such that $\sum_{j=1}^n (b_j - a_j) \le \delta$ we get that

$$\sum_{j=1}^n |F(b_j) - F(a_j)| \le \varepsilon; \quad \sum_{j=1}^n |G(b_j) - G(a_j)| \le \varepsilon;$$

Let us estimate with the usual "bilinear trick": for any $a, b \in J$ we have

$$\begin{aligned}|G(b)\,F(b) - G(a)\,F(a)| &= |G(b)\,F(b) - G(b)\,F(a) + G(b)\,F(a) - G(a)\,F(a)| \le \\ &\le |G(b)|\,|F(b) - F(a)| + |F(a)|\,|G(b) - G(a)| \le \\ &\le \|G\|_J\,|F(b) - F(a)| + \|F\|_J\,|G(b) - G(a)|,\end{aligned}$$

with $\|G\|_J = \max\{|G(x)|;\ x \in J\}$, $\|F\|_J = \max\{|F(x)|;\ x \in J\}$. Then

$$\sum_{j=1}^n |G(b_j)\,F(b_j) - G(a_j)\,F(a_j)| \le \sum_{j=1}^n (\|G\|_J\,|F(b_j) - F(a_j)| + \|F\|_J\,|G(b_j) - G(a_j)|) \le$$

$$\le \|G\|_J \sum_{j=1}^n |F(b_j) - F(a_j)| + \|F\|_J \sum_{j=1}^n |G(b_j) - G(a_j)| \le (\|G\|_J + \|F\|_J)\,\varepsilon.$$

Then $F\,G$ is locally absolutely continuous. The Leibniz rule for the derivative of a product of course holds wherever F and G are both differentiable, hence a.e.; then $(F\,G)'(x) = F'(x)\,G(x) + F(x)\,G'(x)$ for a.e. $x \in I$, and

$$\int_a^b (F'(x)\,G(x) + F(x)\,G'(x))\ dx = F(b)\,G(b) - F(a)\,G(a) \qquad a, b \in I,\ a < b.$$

Of course the left–hand side may also be written

$$\int_{[a,b]} G(x)\,dF(x) + \int_{[a,b]} F(x)\,dG(x).$$

$\square$

Exercise 8.5.6. (*Folland exercise 32*) If F_n, $F \in NBV(\mathbb{R})$ and $F_n \to F$ pointwise, then

$$T(x) \le \liminf_{n\to\infty} T_n(x),$$

for every $x \in \mathbb{R}$; here $T_n(x) = VF_n(]-\infty, x]$, $T(x) = VF(]-\infty, x])$.

Solution. Given $x \in \mathbb{R}$, let $x_0 < x_1 < \dots x_p = x$ be a subdivision of $]-\infty, x]$. For every n we have

$$\sum_{k=1}^p |F_n(x_k) - F_n(x_{k-1})| \le T_n(x).$$

Keeping the sudivision fixed take $\liminf_{n\to\infty}$ on both sides; on the left–hand side we have actually a limit, hence:

$$\sum_{k=1}^p |F(x_k) - F(x_{k-1})| \le \liminf_{n\to\infty} T_n(x);$$

since this holds for every subdivision $\{x_0, \dots, x_p\}$ of $]-\infty, x]$ we also have, taking suprema on the left–hand side:

$$VF(]-\infty, x]) = T(x) \le \liminf_{n\to\infty} T_n(x).$$

$\square$

Exercise 8.5.7. Observe that the sets E_n whose intersection is the Cantor set are all symmetric with respect to $x = 1/2$, that is, $\chi_{E_n}(t) = \chi_{E_n}(1-t)$ for every $t \in \mathbb{R}$; deduce from it that if f is the Cantor function then

$$f(1-x) = 1 - f(x) \quad \text{for every } x \in \mathbb{R} \text{ and hence} \quad \int_0^1 f(x)\,dx = \frac{1}{2}.$$

Solution. Clearly also $g_n(t) = g_n(1-t)$ if $g_n = \chi_{E_n}/\delta_n$. Then, by a change of variable, $t = 1-s$

$$f_n(x) = \int_0^x g_n(t)\,dt = \int_1^{1-x} g_n(1-s)\,(-ds) = \int_{1-x}^1 g_n(s)\,ds = \int_{1-x}^0 g_n(s)\,ds + \int_0^1 g_n(s)\,ds =$$
$$1 - \int_0^{1-x} g_n(s)\,ds = 1 - f_n(1-x).$$

Passing to the limit we get $f(x) = 1 - f(1-x)$. This functional relation (which geometrically means that the graph of f is self–symmetric in the plane symmetry of center $(1/2, 1/2)$) implies that the integral over $[0,1]$ is $1/2$:

$$\int_0^1 f(x)\,dx = \int_1^0 f(1-t)\,(-dt) = \int_0^1 (1 - f(t))\,dt = 1 - \int_0^1 f(t)\,dt,$$

so that

$$\int_0^1 f(x)\,dx = \frac{1}{2}.$$

□

Exercise 8.5.8. Let $\mu : \mathcal{B}_n \to [0, \infty]$ be a $\sigma-$finite positive measure on the Borel tribe $\mathcal{B}_n$ of $\mathbb{R}^n$. If μ is not absolutely continuous with respect to Lebesgue measure m, then there is a set $A \in \mathcal{B}_n$ with $m(A) = 0$ but $\mu(A) > 0$. Prove that nevertheless $m-$almost all translates of A have $\mu-$measure zero, i.e. $\mu(x + A) = 0$ for $m-$a.e. $x \in \mathbb{R}^n$ (let $B = \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n : x - y \in A\}$; compute $(\mu \otimes m)(B)$ by Fubini–Tonelli's theorem …).

Solution. Clearly B is Borel measurable, because $(x,y) \mapsto x - y$ is continuous. Both measures are $\sigma-$finite, so that Fubini–Tonelli's theorem is applicable and if $B_x = \{y \in \mathbb{R}^n : (x,y) \in B\} = x - A$ is the $x-$section of B we get

$$\mu \otimes m(B) = \int_{\mathbb{R}^n} \left(\int_{B_x} dm(y) \right) d\mu(x) = \int_{\mathbb{R}^n} m(x - A)\,d\mu(x) = \int_{\mathbb{R}^n} 0\,d\mu(x) = 0.$$

The $y-$section is $B^y = \{x \in \mathbb{R}^n : (x,y) \in B\} = y + A$, so that

$$\mu \otimes m(B) = \int_{\mathbb{R}^n} \left(\int_{B^y} d\mu(x) \right) dm(y) = \int_{\mathbb{R}^n} \mu(y + A)\,dm(y);$$

this must be zero; then the positive measurable function $y \mapsto \mu(y + A)$, having integral zero with respect to m, is zero $m-$a.e. □

Exercise 8.5.9. Let $I = [a,b]$ be a compact interval. Assume that $F : I \to \mathbb{K}$ is absolutely continuous, and that $F' \in L^p(I)$ with $1 < p \le \infty$; let q be the exponent conjugate to p, and let $\alpha = 1/q$. Prove that F satisfies a *Hölder condition of exponent* α, that is, there exists a constant $k \ge 0$ such that

$$|F(x_2) - F(x_1)| \le k\,|x_2 - x_1|^\alpha \quad \text{for every } x_1, x_2 \in I.$$

Solution. Immediate: apply Hölder's inequality with $|F'(x)|$ and 1 (assuming $x_1 < x_2$, $p < \infty$):

$$|F(x_2) - F(x_1)| = \left| \int_{x_1}^{x_2} F'(x)\,dx \right| \le \int_{[x_1,x_2]} |F'(x)|\,dx \le \left(\int_{[x_1,x_2]} |F'(x)|^p\,dx \right)^{1/p} \left(\int_{[x_1,x_2]} 1^q\,dx \right)^{1/q} \le$$
$$\le \left(\int_{[a,b]} |F'(x)|^p\,dx \right)^{1/p} |x_2 - x_1|^\alpha = \|F'\|_p\,|x_2 - x_1|^\alpha.$$

The case $p = \infty$ is the case of Lipschitz continuous functions, already discussed in a previous exercise.

□

Exercise 8.5.10. Let I be an interval of $\mathbb{R}$, and let $f : I \to \mathbb{K}$ be locally of bounded variation. Given $c \in I$ we define $T = T_c f$ as in 8.3, by $T(x) = Vf([a,x])$ for $x \ge a$, etc.

(i) Prove that for every $a \in I$ we have $T(a^+) - T(a^-) = |f(a^+) - f(a)| + |f(a - f(a^-)|$, so that the jump of T at a coincides with the absolute value of the jump of f at a iff $f(a) \in [f(a^-), f(a^+)]$ (see 8.3.2).

(ii) Given $a > c$, $a \in I$, observe that $T(a^-)(= \lim_{x \to a^-} Vf([c, x])) = Vf([c, a[)$.

(ii) Prove that the total variation is an additive function on the interval algebra of I if and only if f is continuous and that it is then countably additive, and the associated premeasure is then equal to $dT = \mu_T$.

Solution. (i) The result is contained in 8.3.2.

(ii) Given $\varepsilon > 0$, pick $a_\varepsilon \in [c, a[$ such that $T(a^-) - \varepsilon \le T(x)$ for every $x \in [a_\varepsilon, a[$. Then, if $x \in]a_\varepsilon, \varepsilon[$:

$$T(a^-) - \varepsilon \le Vf([c, a_\varepsilon]) \le Vf([c, x[) \le Vf([c, x] = T(x) \le T(a^-),$$

proving what claimed.

(ii) Assume that $a > c$ is a point of discontinuity for f. Then $T(a) = Vf([c, a]) = T(a^-) + |f(a) - f(a^-)|$ while $Vf([c, a[) = T(a^-)$, and $Vf(\{a\}) = 0$, so that $Vf([c, a]) > Vf([c, a[) + Vf(\{a\})$. Then, if f is discontinuous Vf is not additive. It is trivial to prove additivity of the variation with f continuous, and also that $Vf([a, b]) = T(b) - T(a)$; so that Vf is on intervals the premeasure dT. □

EXERCISE 8.5.11. A standard way of embedding isometrically $\ell^1(\mathbb{N})$ into $L^1([0, \infty[) \subseteq L^1(\mathbb{R})$ is by the following interpolation: the sequence $(a_n) = (a(n))_{n \in \mathbb{N}}$ is applied to the piecewise constant right–continuous function $F_a : [0, \infty[\to \mathbb{K}$ given by $F_a(x) = a_{[x]} = a([x])$, where of course $[x]$ is the integer part of x, largest integer not strictly larger than x. Often this function a is considered defined on $\mathbb{R}$, identically zero on $]-\infty, 0[$; the same trick is used to embed $\ell^1(\mathbb{Z})$ into $L^1(\mathbb{R})$, and also arbitrary sequences, not necessarily summable, two sided or not, can be identified in this way with piecewise constant functions. The embedding is isometrical because $a \in \ell^1(\mathbb{N})$ iff $F_a \in L^1([0, \infty[)$ and

$$\sum_{n \in \mathbb{N}} a_n = \int_0^\infty F_a(x)\, dx \quad \text{in particular, taking absolute values} \quad \|a\|_1 = \|F_a\|_1.$$

Notice that for any sequence we have

$$\sum_{k=0}^{n} a_k = \int_0^{n+1} F_a(x)\, dx$$

$$\left(= \int_{[0,n+1]} F_a(x)\, dx = \int_{[0,n+1[} F_a(x)\, dx = \int_{]0,n+1[} F_a(x)\, dx = \int_{]0,n+1]} F_a(x)\, dx \right);$$

and that, considering F_a as defined on $\mathbb{R}$ the measure dF_a is:

$$dF_a = \sum_{n=0}^{\infty} (a_n - a_{n-1})\, \delta_n \quad (a_{-1} = 0).$$

(i) For any sequence $a \in \mathbb{K}^{\mathbb{N}}$, prove that $VF_a([0, \infty[) = \sum_{n=1}^\infty |a_n - a_{n-1}|$ (this justifies the name: ”sequence of bounded variation” given to sequences $(a_n)_{n \in \mathbb{N}}$ such that $\sum_{n=1}^\infty |a_n - a_{n-1}| < \infty$).

(ii) Prove Abel's summation formula

$$\sum_{k=0}^{n} a_k\, b_k = b_n \sum_{k=0}^{n} a_k - \sum_{k=1}^{n} \left(\sum_{j=0}^{k-1} a_j \right) (b_k - b_{k-1}),$$

and deduce from it Abel–Dirichlet criterion for (nonabsolute) convergence of a series:

. *If a_n and b_n are sequences, a_n has bounded partial sums ($\sup_n \{|\sum_{k=0}^n a_k|\} < \infty$) and b_n is of bounded variation and with limit 0 at infinity, then the series $\sum_{k=0}^\infty a_k\, b_k$ is convergent, i.e. $\lim_{n \to \infty} \sum_{k=0}^n a_k\, b_k$ exists finite).*

Solution. (i) Easy (notice that pairs on the same interval $[n, n+1[$ give zero contribution to the variation).

(ii) With F_a and F_b as previously defined we have $\sum_{k=0}^n a_k\, b_k = \int_{[0,n+1[} F_a(x)\, F_b(x)\, dx$; by partial integration formula, taking $F_a(x)\, dx = dA(x)$, where $A(x) = \int_0^x F_a(t)\, dt$:

$$\int_{[0,n+1[} F_a(x)\, F_b(x)\, dx = A((n+1)^-)\, F_b((n+1)^-) - A(0^-) F_b(0^-) - \int_{[0,n+1[} A(x)\, dF_b(x) =$$

$$b_n \sum_{k=0}^{n} a_k - \int_{[0,n+1[} A(x) \left(\sum_{k=0}^{\infty} (b_k - b_{k-1})\, \delta_k \right) = b_n \sum_{k=0}^{n} a_k - \sum_{k=1}^{n} \left(\sum_{j=0}^{k-1} a_j \right) (b_k - b_{k-1})$$

(taking account of the fact that A is continuous, $A(0) = 0$, and in general $A(n) = \sum_{j<n} a_j$, for every integer n. The criterion is now immediate (it is of course a particular case of that for integrals done in exercise 8.5.3): we have

$$\sum_{k=0}^{n} a_k\, b_k = b_n \sum_{k=0}^{n} a_k - \sum_{k=1}^{n} \left(\sum_{j=0}^{k-1} a_j \right) (b_k - b_{k-1}); \qquad (*)$$

since $b_n \to 0$ and the sum $\sum_{k=0}^{n} a_k$ remains bounded, say by M, the first term in the right hand side of (*)tends to 0; and since

$$\sum_{k=1}^{\infty} \left| \sum_{j=0}^{k-1} a_j \right| |b_k - b_{k-1}| \le M \sum_{k=1}^{\infty} |b_k - b_{k-1}| < \infty,$$

the second term is the $n^{th}-$partial sum of an absolutely converging series. □

Exercise 8.5.12. (*Folland, exercises 39, 40*) Let F_n be a sequence of positive increasing functions such that $F(x) = \sum_{n=0}^{\infty} F_n(x)$ is finite for every $x \in [a,b]$. Then $F'(x) = \sum_{n=0}^{\infty} F_n'(x)$ for a.e $x \in [a,b]$ (consider the series of measures $\mu_n = dF_n$ in the space $M([a,b])$ of finite measures ...).

Let now $f : \mathbb{R} \to \mathbb{R}$ be the Cantor function. Let $n \mapsto [a_n, b_n]$ be a bijection of $\mathbb{N}$ onto the set of subintervals of $[0,1]$ with rational endpoints, and let $f_n(x) = f((x - a_n)/(b_n - a_n))$. Then

$$F(x) = \sum_{n=0}^{\infty} \frac{f_n(x)}{2^{n+1}}$$

is continuous, it is strictly increasing on $[0,1]$, but $F'(x) = 0$ for a.e. $x \in \mathbb{R}$.

Solution. We can consider finite measures in $M(\mathbb{R})$, by extending the functions as $F_n(x) = F_n(a)$ for $x < a$ and $F_n(x) = F_n(b)$ for $x > b$: this is equivalent to extending the measures $\mu_n = dF_n$ by declaring $]-\infty, a[$ and $]b, +\infty[$ null sets for them. Then $\mu_n = F_n'(x)\,dx + \nu_n$, with $\nu_n \perp m$. The series $\sum_{n=0}^{\infty} \|\mu_n\|(\mathbb{R}) = \sum_{n=0}^{\infty}(F_n(b) - F_n(a))$ converges; then (see exercise 7.2.7) the series $\sum_{n=0}^{\infty} F_n'$ converges normally in $L^1_m(\mathbb{R})$ and a.e. to $F'(x)$, where $F'(x)\,dx$ is the absolutely continuous part of $dF = \mu_F$. The first part is proved.

For the second part, setting $F_n(x) = f_n(x)/2^{n+1}$ we clearly have $F_n'(x) = 0$ for a.e. $x \in \mathbb{R}$ and for every n, so that by the first part $F'(x) = 0$ a.e.; only the fact that F is strictly increasing is to be checked; and this is trivial: if $x_1 < x_2$ then $x_1 < a_p < b_p < x_2$ for some $p \in \mathbb{N}$ (in fact for infinitely many $p \in \mathbb{N}$), so that

$$F(x_2) - F(x_1) = \sum_{n=0}^{\infty} F_n(x_2) - \sum_{n=0}^{\infty} F_n(x_1) = \sum_{n=0}^{\infty} (F_n(x_2) - F_n(x_1)) \ge F_p(x_2) - F_p(x_1) \ge \frac{1}{2^{p+1}}.$$

□

Before the next exercise read this: if I is an interval of $\mathbb{R}$, and $f : I \to \mathbb{R}$ is increasing, the set $S(f) = \{y \in \mathbb{R} : \varkappa(f^{\leftarrow}(y)) \ge 2\}$, where $\varkappa$ is the counting measure, is an at most countable set. In fact, $f^{\leftarrow}(y)$ is an interval, and as soon as it contains more than one element it has a non empty interior; and any disjoint family of open intervals of $\mathbb{R}$ is at most countable (pick a rational number in each interval, and obtain a one–to–one function of the family into $\mathbb{Q}$). Then, for any pair A, B of disjoint subsets of I we have that $f(A) \cap f(B)$ is at most countable, being contained in $S(f)$. And also, for any $A \subseteq I$ we have that $f(I \setminus A) \setminus (f(I) \setminus f(A))$ is at most countable (it coincides with $f(I \setminus A) \cap f(A)$). Taking account of these observations:

Exercise 8.5.13. Let I be an interval of $\mathbb{R}$ and let $F : I \to \mathbb{R}$ be increasing and continuous.

(i) Prove that for every $B \in \mathcal{B}_I$ the set $F(B)$ is a Borel set. Hint: consider the set

$$\mathcal{M} = \{B \in \mathcal{B}_I : F(B) \quad \text{is Borel}\};$$

prove that this set is a $\sigma-$algebra of parts of I containing the subintervals of I.

(ii) Prove that the formula (m is Lebesgue measure)

$$\mu(B) = m(F(B))$$

defines a measure on $\mathcal{B}_I$, and that this measure coincides with $\mu_F = dF$. Prove that m is the image measure of $\mu = \mu_F$ on $F(I)$.

(iii) Deduce from (ii) that for every $g \in L^+_m(F(I))$ we have

$$\int_{F(I)} g(y)\, dm(y) = \int_I g \circ F(x)\, dF(x),$$

in particular if $F \in AC(I)$ we get

$$\int_{F(I)} g(y)\, dm(y) = \int_I g \circ F(x)\, F'(x)\, dx.$$

Solution. (i) Clearly $\mathcal{M}$ contains all intervals, since any continuous image of an interval is an interval, hence a Borel set. Since the image preserves countable unions, $\mathcal{M}$ is closed under countable unions. Finally, if $E \in \mathcal{M}$, then $F(I \setminus E)$ differs from $F(I) \setminus F(E)$ by a countable set, hence is also a Borel set.

(ii) If $(A_n)_{n\in\mathbb{N}}$ is a disjoint sequence of Borel subsets of I, then $(F(A_n))_{n\in\mathbb{N}}$ is an almost disjoint sequence of Borel sets ($F(A_k) \cap F(A_j)$ is at most countable, hence of zero Lebesgue measure), hence

$$m\left(\bigcup_{k=0}^{\infty} F(A_k)\right) = \sum_{k=0}^{\infty} m(F(A_k)),$$

proving countable additivity of μ. For an interval $[a, b] \subseteq I$ we clearly have

$$\mu([a, b]) = m(F([a, b]) = m([F(a), F(b)] = F(b) - F(a) = \mu_F([a, b]);$$

since compact intervals are a generating system closed under intersection, by uniqueness we get $\mu = \mu_F$.

Let us now prove that for every compact interval $[c, d] \subseteq F(I)$ the inverse image $F^{\leftarrow}([c, d])$ is an interval $J \subseteq I$ with $\mu_F(J) = d - c$. In fact by monotonicity of F this inverse image is an interval, closed in the relative topology of I because F is continuous; if $F(a) = c$ and $F(b) = d$ then $a < b$ and $\mu_F([a, b]) = d - c = m([c, d])$; but $\mu_F([a, b]) = \mu_F(J)$; in fact, if $x \in J$ and $x \le a$ we have $F(x) = F(a)$ and if $x \in J$ and $x \ge b$ then $F(x) = F(b)$, so that $F(\sup J) - F(\inf J) = F(b) - F(a) = d - c$.

(iii) Simply recall 4.1.12.

□

EXERCISE 8.5.14. (⊙⊙ *quite difficult*) Let $[a, b]$ be a compact interval, $f : [a, b] \to \mathbb{R}$ a continuous function. We define $N = N_f : \mathbb{R} \to [0, \infty]$ as $N(y) = \varkappa(f^{\leftarrow}(y))$, where $\varkappa$ is the counting measure; that is $N(y)$ is the number of points in the fiber of f over y if this fiber is finite, otherwise $N(y) = \infty$. We consider a sequence of partitions of $[a, b]$: the n^{th} partition is

$$I(n1) = [a, a + (b - a)/2^n];\ I(nk) =]a + ((k - 1)/2^n)(b - a), a + (k/2^n)(b - a)],\ \ k = 2, \dots, 2^n;$$

we also set $J(nk) = f(I(nk))$, $k = 1, \dots, 2^n$, and $g_n = \sum_{k=1}^{2^n} \chi_{J(nk)}$.

(i) Prove that

$$g_n(y) = \varkappa\left(\{k \in \{1, \dots, 2^n\} : f^{\leftarrow}(y) \cap I(nk) \ne \emptyset\}\right),$$

and that $g_n \uparrow N$.

(ii) Let $\lambda(nk) = \inf J(nk)$ and $\Lambda(nk) = \sup J(nk)$. Express $\int_{\mathbb{R}} g_n(y)\, dy$ with these constants, and prove that

$$\int_{\mathbb{R}} N(y)\, dy = Vf([a, b]) \quad \text{finite or } \infty.$$

Assume now that f is of bounded variation on $[a, b]$.

(iii) Prove that there is unique positive measure μ on the Borel subsets of $[a, b]$ such that $\mu([c, d]) = Vf([c, d])$ for every subinterval $[c, d]$ of $[a, b]$.

(iv) For every Borel subset B of $[a, b]$ we define the function $N_B : \mathbb{R} \to [0, \infty]$ as

$$N_B(y) = \varkappa(f^{\leftarrow}(y) \cap B).$$

Prove that for every such set B:

$$\int_{\mathbb{R}} N_B(y)\, dy = \mu(B).$$

(v) Assume now that f is absolutely continuous. We say that $y \in \mathbb{R}$ is a *regular value* for f if for every $x \in f^{\leftarrow}(y)$ the derivative $f'(x)$ exists and is non zero (in particular, any $y \notin f([a,b])$ is a regular value). Prove that $m-$almost every $y \in \mathbb{R}$ is a regular value.

Solution. (i) is trivial: for every $y \in \mathbb{R}$ we have that $g_n(y)$ is the cardinality of the set $\{k \in \{1,\dots,2^n\} : y \in J(nk)\}$; and clearly $y \in J(nk) = f(I(nk))$ if and only if $f^{\leftarrow}(y) \cap I(nk) \neq \emptyset$. Clearly $g_n(y) \le g_{n+1}(y)$; every $J(nk)$ is the union of two $J((n+1)k)$, and the cardinality of the set $\{k \in \{1,\dots,2^n\} : y \in J(nk)\}$ can then only increase with n. And given a finite subset $A \subseteq f^{\leftarrow}(y)$, as soon as $(b-a)/2^n$ is smaller than the smallest distance of a pair of distinct points of A we have $g_n(y) \ge |A|$, so that $g_n(y) \uparrow N(y)$.

(ii) Clearly we have

$$\int_{\mathbb{R}} g_n(y)\,dy = \sum_{k=1}^{2^n} m(J(nk)) = \sum_{k=1}^{2^n} (\Lambda(nk) - \lambda(nk));$$

of course $\Lambda(nk) = \max f(\bar{I}(nk))$ and $\lambda(nk) = \min f(\bar{I}(nk))$; assume that $\xi(nk) \le \eta(nk)$ are points of the closed interval $\bar{I}(nk)$ where these values are assumed, that is either $\lambda(nk) = f(\xi(nk))$ and $\Lambda(nk) = f(\eta(nk))$, or viceversa $\lambda(nk) = f(\eta(nk))$ and $\Lambda(nk) = f(\xi(nk))$. Then

$$\int_{\mathbb{R}} g_n(y)\,dy = \sum_{k=1}^{2^n} (\Lambda(nk) - \lambda(nk)) = \sum_{k=1}^{2^n} |f(\xi(nk)) - f(\eta(nk))| \le$$
$$\le \sum_{k=1}^{2^n} |f(\xi(nk)) - f(\eta(nk))| + \sum_{k=2}^{2^n} |f(\xi(nk)) - f(\eta(n(k-1)))| \le Vf([a,b]).$$

As $n \uparrow \infty$, by monotone convergence we have $\int_{\mathbb{R}} g_n \uparrow \int_{\mathbb{R}} N$, so that

$$\int_{\mathbb{R}} N(y)\,dy \le Vf([a,b]),$$

and we need the reverse inequality. Given α, with $0 < \alpha < Vf([a,b])$ get a subdivision $a_0 = a < a_1 < \dots < a_p = b$ such that $\sum_{j=1}^{p} |f(x_j) - f(x_{j-1})| \ge \alpha$. Take $n \in \mathbb{N}$ so large that $(b-a)/2^n < \max\{a_j - a_{j-1} : j = 1,\dots,p\}$. We have, using the notations previously introduced, and setting $x(nk) = \sup I(nk)$, $x(n(k-1)) = \inf I(nk)$:

$$\int_{\mathbb{R}} g_n(y)\,dy = \sum_{k=1}^{2^n} |f(\xi(nk)) - f(\eta(nk))| \ge \sum_{k=1}^{2^n} |f(x(nk)) - f(x(n(k-1)))|;$$

let $S(j) = \{k \in \{1,\dots,2^n\} : a_{j-1} \le x(n(k-1)) < x(nk) \le a_j\}$, and let $T = \{k \in \{1,\dots,2^n\} : x(n(k-1)) < a_{j(k)} < x(nk),\ \text{for some } j(k) \in \{1,\dots,p-1\}\}$. Then

$$\int_{\mathbb{R}} g_n(y)\,dy \ge \sum_{k=1}^{2^n} |f(x(nk)) - f(x(n(k-1)))| = \sum_{j=1}^{p} \left(\sum_{k \in S(j)} |f(x(nk)) - f(x(n(k-1)))| \right) +$$
$$+ \sum_{k \in T} |f(x(nk)) - f(x(n(k-1)))|;$$

If to this we add and subtract $\sum_{k \in T} (|f(a_{j(k)}) - f(x(n(k-1)))| + |f(x(nk)) - f(a_{j(k)})|)$ we get the variation of f on the subdivision obtained by joining the points $\{a_0,\dots,a_p\}$ with the points $\{x(nk) : k = 0,\dots,2^n\}$ so that

$$\int_{\mathbb{R}} g_n(y)\,dy \ge \alpha - \sum_{k \in T} (|f(a_{j(k)}) - f(x(n(k-1)))| + |f(x(nk)) - f(a_{j(k)})| - |f(x(nk)) - f(x(n(k-1)))|);$$

since f is continuous at all points a_j, as n tends to infinity this sum tends to 0. We then get

$$\int_R N(y)\,dy \ge \alpha,$$

and since α is an arbitrary number smaller than $Vf([a,b])$ we conclude.

(iii) Since f is continuous we have $Vf([c,d]) = Vf([c,d[) = Vf(]c,d]) = Vf(]c,d[) = T_a(d) - T_a(c)$, so that the measure is the Radon–Stieltjes measure associated to the monotone function $T_a(x) = Vf([a,x]$.

(iv) We know, from (ii), that the formula is true when B is a compact subinterval of $[a,b]$. We simply have to prove that $\int_{\mathbb{R}} N_B(y)\,dy$ is defined for every Borel subset B of $[a,b]$ (in other words $y \mapsto N_B(y)$ is Lebesgue measurable) and that $B \mapsto \int_{\mathbb{R}} N_B(y)\,dy$ is a measure. Simply consider

$$\mathcal{D} = \{B \in \mathcal{B}_{[a,b]} : y \mapsto N_B(y) \quad \text{is Lebesgue measurable, and } \int_{\mathbb{R}} N_B(y)\,dy = \mu(B)\};$$

it is easy to see that $\mathcal{D}$ is Dynkin class of parts of $[a,b]$: for the complement observe that the set $\{y \in \mathbb{R};\ N_{[a,b]}(y) = \infty\}$ has Lebesgue measure 0, since $\int_{\mathbb{R}} N_{[a,b]}(y)\,dy = Vf([a,b])$ is finite. Since $\mathcal{D}$ contains the intervals we have $\mathcal{D} = \mathcal{B}_{[a,b]}$.

(v) The set N of points of $[a,b]$ at which f' does not exist is contained in a Borel set M of measure 0; if $A = M \cup Z(f')$ we clearly have $\int_A |f'(x)|\,dx = 0$; but the measure $B \mapsto \int_B |f'(x)|\,dx$ coincides with μ (because on intervals it coincides with μ); then

$$\int_{\mathbb{R}} N_A(y)\,dy = 0,$$

and since $N_A \ge 0$ this implies $N_A(y) = 0$ for a.e. $y \in \mathbb{R}$; now $N_A(y) = 0$ means that there is no $x \in A$ such that $y = f(x)$; that is, if $y = f(x)$ then $x \notin A$, so that $f'(x)$ exists and is non–zero, i.e., y is a regular value for f. □

8.6. The metric density of a Radon measure. If $\mu : \mathcal{B}_n^* \to \mathbb{K}$ is a locally finite Radon measure on the bounded Borel subsets of $\mathbb{R}^n$, given an open euclidean ball $B(x,r[$ we define the *average density* of μ on $B(x,r[$ as

$$Q_r\mu(x) = \fint_{B(x,r[} d\mu := \frac{\mu(B(x,r[)}{m(B(x,r[)} = \frac{\mu(B(x,r[)}{v_n\, r^n} \qquad v_n = m(B) = \frac{2\,\pi^{n/2}}{n\,\Gamma(n/2)}.$$

Notice that if $\mu \ll m$ and $d\mu = f\,dm$, with $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, then $Q_r\mu(x) = A_r f(x)$, as in 8.1.1. The *(metric) density* at $x \in \mathbb{R}^n$ is

$$D\mu(x) = \lim_{r \to 0^+} Q_r\mu(x), \quad \text{provided that the limit exists.}$$

EXERCISE 8.6.1. Observe that $r \mapsto Q_r\mu(x)$ is left–continuous, and that

$$\lim_{t \to r^+} Q_r(x) = \frac{\mu(B(x,r])}{m(B(x,r])} \left(= \frac{\mu(B(x,r])}{v_n\, r^n} \right),$$

so that $Q_r\mu(x)$ has finite left and right limits at every $r > 0$, and $D\mu(x)$ can be computed also using closed balls in place of open balls.

EXERCISE 8.6.2. Let $\mu : \mathcal{B}_n \to [0,\infty]$ be a positive Radon measure. Keeping $r > 0$ fixed, prove that $x \mapsto Q_r\mu(x)$ is a lower semicontinuous function from $\mathbb{R}^n$ to $[0,\infty[$. If $\mu \ll m$ then this function is continuous.

Solution. Assume that x_k is a sequence in $\mathbb{R}^n$ converging to $x \in \mathbb{R}^n$. If $\chi_k = \chi_{B(x_k,r[}$ and $\chi = \chi_{B(x,r[}$ then we have $\lim_{k\to\infty} \chi_k(y) = \chi(y)$ if $y \in B(x,r[)$ or if $|y-x| > r$ (since $\lim_{k\to\infty}(|y-x_k|-r) = |y-x|-r$, eventually $|y-x_k|-r$ has the same sign as $|y-x|-r$, if this is non–zero). In particular we have:

$$\chi(x) \le \liminf_{k\to\infty} \chi_k(x) \Longrightarrow \int_{\mathbb{R}^n} \chi\,d\mu \le \int_{\mathbb{R}^n} (\liminf_{k\to\infty} \chi_k)\,d\mu \le \liminf_{k\to\infty} \int_{\mathbb{R}^n} \chi_k\,d\mu,$$

by Fatou's lemma. In other words:

$$\mu(B(x,r[) \le \liminf_{k\to\infty} \mu(B(x_k,r[), \text{ (dividing both sides by } m(B(x,r[) = v_n\, r^n)\ Q_r\mu(x) \le \liminf_{k\to\infty} Q_r\mu(x_k).$$

If $\mu \ll m$ then χ_k converges μ–a.e. to χ, since $m(\{|y-x| = r\}) = 0$ implies $\mu(\{|y-x| = r\}) = 0$, and we can apply Lebesgue's dominated convergence theorem to the sequence $\chi_k\, \chi_{B(x,R]}$, where $R = \sup\{r + |x-x_k| : k \in \mathbb{N}\}$.

□

Our aim is to prove the differentiation theorem stated in 8.1. Because of the local character of metric derivatives, we need to prove the theorem only for finite measures: if μ is locally finite, given a big bounded open set, say $B(0,R[$ with R large, we can consider the finite Radon measure $\mu_R : \mathcal{B}_n \to \mathbb{K}$

given by $\mu_R(E) = \mu(E \cap B(0,R[)$; then, given $x \in B(0,R[$ we have that $D\mu(x)$ exists iff $D\mu_R(x)$ exists, and they coincide. So we shall prove that if $f \in L^1_m(\mathbb{R}^n)$ then

$$\lim_{r\to 0^+} A_r f(x) = f(x) \quad m\text{−a.e. in } \mathbb{R}^n,$$

and that if ν is a finite measure with $\nu \perp m$ then

$$D\nu(x) = D|\nu|(x) = 0 \quad m\text{−a.e. in } \mathbb{R}^n;$$

from these facts the differentiation theorem follows immediately. A convenient tool for the proof, of independent interest, is the notion of maximal function, with the maximal theorem which is a generalization of Čebičeff's inequality. First we need a technical lemma.

8.6.1.

LEMMA. *Let $(B_j)_{1\le j\le N}$ be a finite family of open balls in $\mathbb{R}^n$. Then there is a subset $J \subseteq \{1,\dots,N\}$ such that the balls $(B_j)_{j\in J}$ are pairwise disjoint and*

$$m\left(\bigcup_{j=1}^N B_j\right) \le 3^n \sum_{j\in J} m(B_j).$$

Proof. Let $B_j = c_j + r_j\, B$, with B the open unit ball; assume that the indexing is with decreasing radii, i.e. such that $r_1 \ge r_2 \ge \cdots \ge r_N$. We define the subset J by induction in the following way: $j(1) = 1 \in J$; consider now the set J_2 of all $j \in \{1,\dots,N\}$ such that $B_j \cap B_1 = \emptyset$; if this set is empty, then $J = \{1\}$, otherwise $j(2) = \min J_2$. Assuming that $\{j(1),\dots,j(p)\}$ have been defined, set $J_{p+1} = \{j \in \{1,\dots,N\} : B_j \cap B_{j(k)} = \emptyset,$ for all $k \in \{1,\dots,p\}\}$; if J_{p+1} is empty, then $J = \{j(1),\dots,j(p)\}$, otherwise $j(p+1) = \min J_{p+1}$. Since $\{1,\dots,N\}$ is finite the process must end at some p; we define $J = \{j(1),\dots,j(p)\}$. We now prove that

$$\bigcup_{j=1}^N B_j \subseteq \bigcup_{j\in J}(c_j + 3r_j\, B);$$

in fact, assume that $k \in \{1,\dots,N\}$ is not in J; then $B_k = c_k + r_k\, B$ has non–empty intersection with some ball $c_j + r_j\, B$, with $r_j \ge r_k$, and $j \in J$; then $c_k + r_k\, B \subseteq c_j + 3r_j\, B$. By monotonicity and subadditivity we then get

$$m\left(\bigcup_{j=1}^N B_j\right) \le m\left(\bigcup_{j\in J}(c_j + 3r_j\, B)\right) \le \sum_{j\in J} m(c_j + 3r_j\, B) = 3^n \sum_{j\in J} m(B_j).$$

□

8.6.2. *The maximal theorem.* If $\mu : \mathcal{B}_n \to \mathbb{K}$ is a finite $\mathbb{K}-$valued measure on the Borel sets of $\mathbb{R}^n$, its *maximal function* is the function $M\mu : \mathbb{R}^n \to [0,\infty]$ defined by

$$M\mu(x) = \sup\{Q_r|\mu|(x) : r > 0\} = \sup\left\{\frac{|\mu|(B(x,r[)}{m(B(x,r[)} : r > 0\right\}.$$

Notice that $M\mu = M|\mu|$: the notion is really a notion for finite positive measures. For any given $x \in \mathbb{R}^n$ we of course have $\lim_{r\to\infty} |\mu|(B(x,r[)/m(B(x,r[) = 0$, since $|\mu|(\mathbb{R}^n) < \infty$; but it is not at all apparent that $M\mu(x)$ is finite $m-$a.e., or even only for some $x \in \mathbb{R}^n$. We observe that:

. $M\mu : \mathbb{R}^n \to [0,\infty]$ *is lower semicontinuous: that is, for every $\alpha > 0$ the set $\{M\mu > \alpha\}$ is open in $\mathbb{R}^n$.*

Proof. Let $a \in \{M\mu > \alpha\}$; then $|\mu|(B(a,t[)/m(B(a,t[) > \alpha$ for some $t > 0$; given $\delta > 0$, if $|x - a| < \delta$ then $B(x, t+\delta[\supseteq B(a,t[$, so that

$$|\mu|(B(x,t+\delta[) \ge |\mu|(B(a,t[) \Longrightarrow \frac{|\mu|(B(x,t+\delta[)}{m(B(x,t+\delta[)} \ge \frac{|\mu|(B(a,t)}{m(B(a,t+\delta[)};$$

as $\delta \to 0^+$ the last quotient tends to $|\mu|(B(a,t[)/m(B(a,t[) > \alpha$, and hence it is larger than α for δ small. For this δ we clearly have $M\mu(x) > \alpha$, for all $x \in B(a,\delta[$. Alternatively, we can use exercise 8.6.2: for fixed $r > 0$ the function $x \mapsto Q_r|\mu|(x)$ is lsc, so that $M\mu$ is lsc, as a supremum of lsc functions (immediate). □

. The maximal theorem *Let $\mu : \mathcal{B}_n \to \mathbb{K}$ be a finite $\mathbb{K}-$valued measure on the Borel sets of $\mathbb{R}^n$. Then, for any $\alpha > 0$ we have*

$$m\left(\{M\mu > \alpha\}\right) \le \frac{3^n}{\alpha}\,\|\mu\| \qquad (\|\mu\| = |\mu|(\mathbb{R}^n)).$$

Proof. For each $x \in E(\alpha) = \{M\mu > \alpha\}$ there is an open ball $B(x, r(x)[$ such that

$$|\mu|(B(x, r(x)[) > \alpha\, m(B(x, r(x)[).$$

Given a compact $K \subseteq E(\alpha)$, there is a finite set $\{B(x_j, r(x_j)[:\ 1 \le j \le N\}$ of these balls which covers K; by lemma 8.6.1 there is a disjoint subset of this set, say $B_1, \dots, B_p$, such that $m\left(\bigcup_{j=1}^N B(x_j, r(x_j)[\right) \le 3^n \sum_{k=1}^p m(B_k)$. Then

$$m(K) \le m\left(\bigcup_{j=1}^N B(x_j, r(x_j)[\right) \le 3^n \sum_{k=1}^p m(B_k) \le 3^n \sum_{k=1}^p \frac{|\mu|(B_k)}{\alpha} = \frac{3^n}{\alpha}\,|\mu|\left(\bigcup_{k=1}^p B_k\right) \le \frac{3^n}{\alpha}\,|\mu|(\mathbb{R}^n).$$

We have proved that for every compact subset K of $E(\alpha)$ we have $m(K) \le (3^n/\alpha)\,|\mu|(\mathbb{R}^n)$; since $m(E(\alpha)) = \sup\{m(K) :\ K \subseteq E(\alpha),\ K \text{ compact}\}$, we conclude. □

8.6.3. *Differentiation of a singular measure.*

Proposition. *Let $\nu : \mathcal{B}_n \to \mathbb{K}$ be a finite $\mathbb{K}-$valued measure on the Borel sets of $\mathbb{R}^n$. Assume that ν is singular with respect to Lebesgue measure. Then, for every substantial family $E_r(x)$ and $m-$almost every $x \in \mathbb{R}^n$ we have*

$$\lim_{r\to 0^+} \frac{\nu(E_r(x))}{m(E_r(x))} = 0.$$

Proof. We have

$$\frac{|\nu(E_r(x))|}{m(E_r(x))} \le \frac{|\nu|(E_r(x))}{m(E_r(x))} \le \frac{|\nu|(B(x, r])}{m(E_r(x))} \le \frac{1}{\alpha(x)}\frac{|\nu|(B(x, r])}{m(B(x, r])},$$

so that it is enough to prove that the metric density of $|\nu|$ is a.e. zero. So we assume that $\nu = |\nu|$, i.e., that ν is positive. There is a Borel set $A \subseteq \mathbb{R}^n$ such that $\nu(\mathbb{R}^n \setminus A) = 0$ and $m(A) = 0$. Given $\varepsilon > 0$ we find a compact subset K of A such that $\nu(K) > \nu(A) - \varepsilon = \nu(\mathbb{R}^n) \setminus \varepsilon = \|\nu\| - \varepsilon$. Let us split ν as the sum of two measures, $\nu_1(E) = \nu(E \cap K)$ and $\nu_2(E) = \nu(E \setminus K)$. Then $\|\nu_2\| = \nu(\mathbb{R}^n \setminus K) < \varepsilon$; clearly $D\nu_1(x) = 0$ for every $x \in \mathbb{R}^n \setminus K$; given $\alpha > 0$, the set $\{x \in \mathbb{R}^n :\ \limsup_{r\to 0^+} \nu_2(B(x, r[)/m(B(x, r[) > \alpha\}$ is clearly contained in $\{M\nu_2 > \alpha\}$; the maximal theorem implies

$$m\left(\{M\nu_2 > \alpha\}\right) \le \frac{3^n}{\alpha}\,\|\nu_2\| \le \frac{3^n}{\alpha}\,\varepsilon.$$

Then

$$\{x \in \mathbb{R}^n :\ \limsup_{r\to 0^+} Q_r\nu(x) > \alpha\} \subseteq K \cup \{M\nu_2 > \alpha\},$$

a set of Lebesgue measure smaller than $3^n\varepsilon/\alpha$. By the arbitrariness of ε, we get

$$m\left(\{x \in \mathbb{R}^n :\ \limsup_{r\to 0^+} Q_r\nu(x) > \alpha\}\right) = 0.$$

Since

$$\{x \in \mathbb{R}^n :\ \limsup_{r\to 0^+} Q_r\nu(x) > 0\} = \bigcup_{k\ge 1}\{x \in \mathbb{R}^n :\ \limsup_{r\to 0^+} Q_r\nu(x) > 1/k\},$$

we get

$$m\left(\{x \in \mathbb{R}^n :\ \limsup_{r\to 0^+} Q_r\nu(x) > 0\}\right) = 0,$$

as required □

8.6.4. *Density for absolutely continuous measures.* Assume that $f \in L^1(\mathbb{R}^n)$ is given. If $g \in L^1_{\rm loc}(\mathbb{R}^n)$ it is easy to prove that at every point x of continuity for g we have:

$$\lim_{r\to 0^+} |A_r g(x) - g(x)| = \lim_{r\to 0^+} A_r|g - g(x)|(x) = 0.$$

Integrable step–functions are continuous a.e., so that the preceding formula is a.e. true in $\mathbb{R}^n$ if g is a step–function in $L^1(\mathbb{R}^n)$. Now fix $\alpha > 0$; we prove that the set

$$F(\alpha) = \{x \in \mathbb{R}^n : \limsup_{r\to 0^+} |A_r f(x) - f(x)| > \alpha\}$$

has Lebesgue measure 0; since the set $\{x \in \mathbb{R}^n : D\nu_f(x) \neq f(x)\}$ is contained in $\bigcup_{k=1}^{\infty} F(1/k)$, we conclude (here ν_f is the indefinite integral of f, $d\nu_f = f\,dm$). Given $\varepsilon > 0$ pick a step–function $g \in L^1(m)$ such that $\|f - g\|_1 \le \varepsilon$, and write

$$\begin{aligned}|A_r f(x) - f(x)| =& |A_r(f-g)(x) + A_r g(x) - g(x) + g(x) - f(x)| \le \\ &\le A_r|f-g|(x) + A_r|g - g(x)|(x) + |g(x) - f(x)| \le \\ &\le M|f-g|(x) + A_r|g-g(x)|(x) + |g(x) - f(x)|,\end{aligned}$$

where $M|f-g|(x) = \sup\{A_r|f-g| : r > 0\}$ is the maximal function of the measure $|f-g|\,dm$; then

$$\limsup_{r\to 0^+} |A_r f(x) - f(x)| \le M|f-g|(x) + |g(x) - f(x)| \qquad \text{for } m\text{−a.e. } x \in \mathbb{R}^n.$$

Then

$$\{x \in \mathbb{R}^n : \limsup_{r\to 0^+} |A_r f(x) - f(x)| > \alpha\} \subseteq \{M|f-g| > \alpha/2\} \cup \{|f-g| > \alpha/2\} \cup N;$$

with $m(N) = 0$; the maximal theorem says that

$$m\left(\{M|f-g| > \alpha/2\}\right) \le 3^n\,\frac{2}{\alpha}\,\|f-g\|_1 \le 3^n\,\frac{2}{\alpha}\,\varepsilon,$$

and by Čebičeff's inequality:

$$m(\{|f-g| > \alpha/2\}) \le \frac{2}{\alpha}\,\|f-g\|_1 \le \frac{2}{\alpha}\,\varepsilon.$$

We have proved that

$$m(F(\alpha)) \le 3^n\,\frac{2}{\alpha}\,\varepsilon + \frac{2}{\alpha}\,\varepsilon,$$

and by the arbitrariness of ε we conclude that $m(F(\alpha)) = 0$, as required.

The proof of theorem 8.1.1 is concluded.

Remark. Given $f \in L^1_{\rm loc}(\mathbb{R}^n)$ we have that $\lim_{r\to 0^+} A_r f(x) = f(x)$ a.e. in $\mathbb{R}^n$ and also that almost every $x \in \mathbb{R}^n$ is a Lebesgue point for f, that is

$$\lim_{r\to 0^+} \fint_{B(x,r]} |f(y) - f(x)|\,dy = 0.$$

These notions depend on the particular representative f, and are not invariant if we substitute f with another function a.e. equal to it. However $\lim_{r\to 0^+} A_r f(x)$, when it exists finite, and a value $[f](x)$ such that $\lim_{r\to 0^+} \fint_{B(x,r]} |f(y) - [f](x)|\,dy = 0$, when such a value exists, depend only on x and on the class $[f]$ of functions a.e. equal to f, and not on the particular representative f: the proposition in 8.1.1 asserts that $[f](x)$ exists for a.e. $x \in \mathbb{R}^n$, and that given any representative $g \in [f]$ we have $g(x) = [f](x)$ for a.e. $x \in \mathbb{R}^n$. We can of course get a representative a.e. defined by the formula $x \mapsto [f](x)$: this is the one with the largest Lebesgue set.

Exercise 8.6.3. Assume that $f \in L^1(\mathbb{R}^n)$ is non zero. Prove that there is $k > 0$ such that, for $|x|$ large enough

$$Mf(x) \ge \frac{k}{|x|^n};$$

then $Mf \in L^1(\mathbb{R}^n)$ if and only if $f = 0$.

Solution. Since f is nonzero, we have $\lim_{r\to\infty}\int_{rB}|f|\,dm = \|f\|_1 > 0$, so that there is $r > 0$ such that $\int_{rB}|f|\,dm = a > 0$. If $|x| > r$, then $rB \subseteq B(x, 2|x|]$ so that

$$⨍_{B(x,2|x|]}|f(y)|\,dy = \frac{1}{2^n\,v_n|x|^n}\int_{B(x,2|x|]}|f(y)|\,dy \ge \frac{1}{2^n\,v_n|x|^n}\int_{rB}|f(y)|\,dy \ge \frac{a}{2^n\,v_n|x|^n} = \frac{a/(2^n\,v_n)}{|x|^n},$$

so that, if $k = a/(2^n\,v_n)$ we have

$$Mf(x) \ge \frac{k}{|x|^n} \qquad (|x| > r).$$

For the last assertion, simply recall that $x \mapsto 1/|x|^n$ is not in $L^1(\mathbb{R}^n \setminus B)$.

□

EXERCISE 8.6.4. (*Folland, exercise 23 pag. 100*) For $f \in L^1(\mathbb{R}^n)$ define another maximal function $M^*f : \mathbb{R}^n \to [0,\infty]$ as

$$M^*f(x) = \sup\left\{⨍_B |f(y)|\,dy : \ x \in B,\ B \text{ an euclidean open ball}\right\}.$$

Prove that $Mf(x) \le M^*f(x) \le 2^n\,Mf(x)$.

Solution. The first inequality is trivial ($Mf(x)$ is a supremum taken over a smaller set of averages, those over balls centered at x). If B is an open ball containing x, of radius $\rho > 0$, then $B(x, 2\rho[\supseteq B$, so that

$$⨍_B |f| = \frac{1}{v_n\,\rho^n}\int_B |f| \le \frac{1}{v_n\,\rho^n}\int_{B(x,2\rho[}|f| = \frac{2^n}{v_n\,(2\rho)^n}\int_{B(x,2\rho[}|f| = 2^n ⨍_{B(x,2\rho[}|f| \le 2^n\,Mf(x).$$

□

EXERCISE 8.6.5. Prove that if $f \in L^1(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$ is a Lebesgue point for f, then

$$|f(x)| \le Mf(x).$$

Solution. If x is a Lebesgue point for f, then x is also a Lebesgue point for $|f|$, as is clear from the inequality $||f(y)| - |f(x)|| \le |f(y) - f(x)|$. Then

$$|f(x)| = \lim_{r\to 0^+} ⨍_{B(x,r[}|f(y)|\,dm(y) \quad \text{which implies} \quad |f(x)| \le \sup_{r>0} ⨍_{B(x,r[}|f(y)|\,dm(y) = Mf(x).$$

□

8.7. Some complements.

8.7.1. *The Severini–Egoroff's theorem on almost uniform convergence.* On finite measure spaces pointwise convergence implies uniform convergence outside sets of arbitrarily small measure:

. SEVERINI– EGOROFF'S THEOREM *Let $(X, \mathcal{M}, \mu)$ be a finite measure space. If $f_n \in L(X, \mathbb{K})$ is a sequence of measurable functions which converges pointwise a.e. to $f : X \to \mathbb{K}$, then f_n converges to f* almost uniformly *on X: in other words, given $\delta > 0$ there is a set $E \in \mathcal{M}$ with $\mu(E) \le \delta$, such that on $X \setminus E$ the sequence f_n converges uniformly to f.*

Proof. For a given $x \in X$ it is false that $f_n(x)$ converges to $f(x)$ if and only if there is $\varepsilon > 0$ such that $|f(x) - f_n(x)| > \varepsilon$ for infinitely many $n \in \mathbb{N}$; this last set is exactly $\limsup_{n\to\infty}\{|f - f_n| > \varepsilon\}$; in other words the set $M = \{x \in X : \ f_n(x) \text{ does not converge to } f(x)\}$ is

$$M = \bigcup_{\varepsilon>0}\limsup_{n\to\infty}\{|f - f_n| > \varepsilon\} \quad \text{or also} \quad M = \bigcup_{k=1}^{\infty}\limsup_{n\to\infty}\{|f - f_n| > 1/k\}.$$

Since by hypothesis $\mu(M) = 0$, each set $A_k = \limsup_{n\to\infty}\{|f - f_n| > 1/k\}$ has zero measure. This set is

$$A_k = \bigcap_{p=0}^{\infty} B_p(k), \quad \text{where} \quad B_p(k) = \bigcup_{n=p}^{\infty}\{|f - f_n| > 1/k\};$$

now $B_0(k) \supseteq B_1(k) \supseteq \dots$ is a decreasing sequence of sets of finite measure with intersection of measure zero, so we can find $p(k) \in \mathbb{N}$ such that $E(k) = B_{p(k)}$ has measure smaller than $\delta/2^{k+1}$. If $E = \bigcup_{k=0}^{\infty} E(k)$, then $\mu(E) \le \delta$; and if $x \in X \setminus E$, then $x \in X \setminus E(k)$, so that if $n \ge p(k)$ we have $|f(x) - f_n(x)| \le 1/k$; this says that f_n converges uniformly to f on $X \setminus E$. □

8.7.2. *Severini–Egoroff's theorem, version with dominated convergence.* To make the preceding proof work one needs only to know that for every $k \ge 1$ there is p large enough to make the measure of $B_p(k) = \bigcup_{n=p}^{\infty}\{|f - f_n| > 1/k\}$ finite; this is trivially true if the entire space has finite measure; another possibility is the following:

. *Let $(X, \mathcal{M}, \mu)$ be a measure space. If $f_n \in L(X, \mathbb{K})$ is a sequence of measurable functions which converges pointwise a.e. to $f : X \to \mathbb{K}$, and $|f_n| \le g$ for some $g \in L^1(\mu)$, then f_n converges to f* almost uniformly *on X.*

Proof. Simply note that $\{|f - f_n| > 1/k\} \subseteq \{g > 1/(2k)\}$, since $|f - f_n| \le 2g$, hence $\bigcup_{n=0}^{\infty}\{|f - f_n| > 1/k\} \subseteq \{g > 1/(2k)\}$; by Čebičeff's inequality $\mu(\{g > 1/(2k)\}) \le 2k \int_X g$.

□

8.7.3. *A continuity property for L^1 functions.* We have seen that functions in $L^1(\mathbb{R})$ can be nowhere continuous, and even have essential supremum infinite on every non–empty open set. The following result may therefore come as a surprise:

. *Let $f \in L^1(\mathbb{R}^n)$. Then, for every $\delta > 0$ there exists a closed set $F \subseteq \mathbb{R}^n$ with $m(\mathbb{R}^n \setminus F) \le \delta$, such that $f|F)$ is continuous on F.*

Proof. Choose a sequence g_k of functions in $C_c(\mathbb{R}^n)$ which converges in $L^1(\mathbb{R}^n)$ to f and is of bounded variation in $L^1(\mathbb{R}^n)$, ie. such that $\sum_{k=1}^{\infty} \|g_{k+1} - g_k\|_1 < \infty$. We know that in this hypothesis the sequence g_k converges also a.e. to f; moreover, if $g = |g_0| + \sum_{k=1}^{\infty} |g_{k+1} - g_k|$ we have $g \in L^1(\mathbb{R}^n)$, and clearly $|g_k| \le g$ for every $k \in \mathbb{N}$. Then, given $\delta > 0$ there is a subset of measure less than δ, and by outer regularity we may suppose it to be an open set U, such that on the complement $F = \mathbb{R}^n \setminus U$ the sequence g_k converges uniformly to f. □

Some people express the above situation by saying that every L^1 function is *almost continuous*. Of course this *does not mean that f has points of continuity*; it is the *restriction* of f to F that is continuous *on F*.

8.7.4. *Lusin's theorem.* Strictly related to the preceding result is

. Lusin's theorem. *Let $f : \mathbb{R}^n \to \mathbb{K}$ be measurable. Then*

(i) *Assume that f is zero outside a set C of finite measure. Then for every $\delta > 0$ there is a compact subset K of C such that $f|K$ is continuous, and $m(C \setminus K) \le \delta$.*

(ii) *For every $\delta > 0$ there is a closed subset F of $\mathbb{R}^n$, with $m(\mathbb{R}^n \setminus F) \le \delta$, such that $f|F$ is continuous.*

Proof. (i) The sets $E_k = \{|f| > k\}$, $k \ge 1$, are a decreasing sequence of sets of finite measure ($E_k \subseteq C$) with empty intersection, so that $m(E_k) \le \delta/3$ for k large; we may reset f to be 0 on E_k, so that, calling f_1 the new function, f_1 is now bounded and zero outside a set of finite measure, and hence $f_1 \in L^1(\mathbb{R}^n)$. The preceding result says that there is a closed subset $F \subseteq \mathbb{R}^n$, with $m(\mathbb{R}^n \setminus F) \le \delta/3$, such that $f_1|F$ is continuous; next, simply pick a compact $K \subseteq F \cap \mathrm{Coz}(f_1)$, with $m(F \cap \mathrm{Coz}(f_1) \setminus K) \le \delta/3$.

(ii) Write $\mathbb{R}^n = \bigcup_{j=0}^{\infty} j\,B$. For $j = 1, \dots$ pick a compact subset $L_j \subseteq j\,B \setminus (j-1)\,\bar{B}$ such that $f|L_j$ is continuous, and $m((j\,B \setminus (j-1)\,\bar{B}) \setminus L_j) \le \delta/2^{j+1}$. Then $F = \bigcup_{j=1}^{\infty} L_j$ is a closed set: if $x \notin F$, and j is the smallest integer strictly larger than $|x|$, we have $x \in j\,B \setminus \bigcup_{l=1}^{j} L_l$, and open set disjoint from F. And $f|F$ is continuous, since each L_j is open–and–closed in the relative topology of F (we have $L_j = F \cap (j\,B \setminus (j-1)\,\bar{B})$; clearly $m(\mathbb{R}^n \setminus F) \le \delta$ □

There is a theorem of general topology which says that any continuous function from a closed subset F of $\mathbb{R}^n$ to $\mathbb{K}$ may be extended to a $\mathbb{K}-$valued continuous function on all of $\mathbb{R}^n$ (the Tietze–Urysohn extension theorem). Then the preceding theorem implies that if $f : \mathbb{R}^n \to \mathbb{K}$ is measurable, and $\delta > 0$, there is $g \in C(\mathbb{R}^n, \mathbb{K})$ and a closed subset $F \subseteq \mathbb{R}^n$, with $m(\mathbb{R}^n \setminus F) \le \delta$, such that $f|F = g|F$.

8.7.5. *More exercises.*

Exercise 8.7.1. Prove that the function $F : \mathbb{R} \to \mathbb{R}$ defined by $F(x) = 2x + \int_0^x \sin(1/t)\,dt$ is strictly increasing and has a finite derivative everywhere, but F' not continuous.

Solution. Clearly we have $F'(x) = 2 + \sin(1/x) > 0$ for $x \ne 0$, and F is continuous, so that it is strictly increasing; we prove that $F'(0) = 2$. In fact, assuming $x \ne 0$:

$$\left|\int_0^x \sin(1/t)\,dt\right| = \left|\int_0^x (t^2)(-\sin(1/t))(-1/t^2)\,dt\right| = \left|\big[t^2\cos(1/t)\big]_{t=0}^{t=x} - \int_0^x 2t\cos(1/t)\,dt\right| \le$$

$$\leq \left|x^2\,\cos(1/x)\right| + \left|\int_0^x 2t\,\cos(1/t)\,dt\right| \leq x^2\,|\cos(1/x)| + \left|\int_0^x 2|t|\,dt\right| \leq$$
$$\leq x^2 + x^2 = 2x^2,$$

so that we have, for $x \neq 0$:

$$\left|\frac{1}{x}\int_0^x \sin(1/t)\,dt\right| \leq 2|x| \quad \text{hence} \quad \lim_{x\to 0}\frac{1}{x}\int_0^x \sin(1/t)\,dt = 0.$$

□

Exercise 8.7.2. Prove that for every subset E of $\mathbb{R}$ of zero Lebesgue measure there is an absolutely continuous increasing function $F : \mathbb{R} \to \mathbb{R}$ such that $F'(x) = \infty$ for every $x \in E$ (Hint: there is a decreasing sequence $U_0 \supseteq U_1 \supseteq U_2 \supseteq \dots$ of open sets such that $E \subseteq \bigcap_{n=0}^\infty U_n$ and $m(U_n) \leq 1/2^{n+1}\ \dots$).

Solution. Let f_n be the characteristic function of U_n, let $F_n(x) = \int_{-\infty}^x f_n(t)\,dt$, and

$$F(x) = \sum_{n=0}^\infty F_n(x) = \sum_{n=0}^\infty \int_{-\infty}^x f_n(t)\,dt = \int_{-\infty}^x \left(\sum_{n=0}^\infty f_n(t)\right)\,dt;$$

Then clearly F is increasing, finite–valued ($F(+\infty) = \sum_{n=0}^\infty m(U_n) \leq 1$) and absolutely continuous. We prove that $F'(c) = \infty$ for every $c \in \bigcap_{n=0}^\infty U_n$. For $x \neq c$ let $\nu(x)$ be the cardinality of the set

$$N(x) = \{n \in \mathbb{N} : x \text{ is contained in an interval contained in } U_n\}.$$

If $n \in N(x)$ we clearly have $(F_n(x) - F_n(c))/(x-c) = 1$, so that

$$\frac{F(x)-F(c)}{x-c} = \sum_{k=0}^\infty \frac{F_k(x)-F_k(c)}{x-c} \geq \sum_{k\in N(x)} \frac{F_k(x)-F_k(c)}{x-c} = \nu(x).$$

It is plain that $\lim_{x\to c}\nu(x) = \infty$ (given an integer N, pick an interval $]c-\delta, c+\delta[$ contained in U_N; then $\nu(x) \geq N$ for every x in this interval); then

$$F'(c) = \lim_{x\to c}\frac{F(x)-F(c)}{x-c} \geq \lim_{x\to c}\nu(x) = \infty.$$

□

Exercise 8.7.3. Given a function $f : \mathbb{R}^n \to \mathbb{K}$ and $a \in \mathbb{R}^n$, we say that a is a *period* for f if $f(x+a) = f(x)$, for every $x \in \mathbb{R}^n$; trivially the zero vector is a period for every function; we say that a function is periodic when it has a nonzero period. We may also weaken the notion of period by asking $f(x+a) = f(x)$ only for $m-$a.e. $x \in \mathbb{R}^n$, with m Lebesgue measure.

(i) Prove that the set of all periods of a function is an additive subgroup of $\mathbb{R}^n$; this subgroup is denoted $\mathrm{Per}(f)$, the group of periods of f, so that f is periodic iff $\mathrm{Per}(f)$ is not the trivial subgroup. Notice that if G is a non–trivial additive subgroup of $\mathbb{R}^n$ then its characteristic function χ_G is periodic and $\mathrm{Per}(\chi_G) \supseteq G$.

(ii) Prove that a continuous periodic function $f : \mathbb{R}^n \to \mathbb{K}$ with a group of periods dense in $\mathbb{R}^n$ is a constant.

(iii) Let $f : \mathbb{R}^n \to \mathbb{K}$ be periodic and in $L^1_{\mathrm{loc}}(\mathbb{R}^n)$; prove that for every $r > 0$ the function $A_r f : \mathbb{R}^n \to \mathbb{K}$ defined by $A_r f(x) = ⨍_{B(x,r[} f$ is also periodic, and $\mathrm{Per}(A_r f) \supseteq \mathrm{Per}(f)$.

(iv) Prove that a locally summable periodic function f with a dense group of periods is a.e. constant.

(v) Prove that a measurable periodic function f with a dense group of periods is a.e. constant (hint: to use (iv), consider $\arctan f\ \dots$).

(vi) Prove that a proper additive subgroup of $\mathbb{R}^n$ is measurable if and only if it has zero $m-$measure (hint: use the difference theorem 2.7.3, and observe that an additive subgroup that contains a nbhd of 0 necessarily coincides with $\mathbb{R}^n$).

Solution. (i) Easy; to prove that if $a \in \mathrm{Per}(f)$ then also $-a \in \mathrm{Per}(f)$ simply observe that $f(x) = f(x-a+a) = f((x-a)+a) = f(x-a)$ for every $x \in \mathbb{R}$. Clearly $\mathrm{Per}(\chi_G) \supseteq G$: if $a \in G$ then $x + a \in G$ if and only if $x \in G$ (we also have $\mathrm{Per}(\chi_G) = G$; but if periodicity is in the weaker sense of a.e. equality then $\mathrm{Per}(\chi_G)$ may be strictly larger, e.g. the Dirichlet function $\chi_\mathbb{Q}$ is a.e. equal to the constant 0, hence its period group is $\mathbb{R}$, in the weaker sense).

(ii) If $\mathrm{Per}(f) = G$ with G dense, then $f(a) = f(0)$ for every $a \in G$, so that f is constant on G; and clearly a continuous function constant on a dense subset is constant ($f^{\leftarrow}(\{f(0)\})$ is a closed subset of $\mathbb{R}^n$ containing the dense subset G, hence $f^{\leftarrow}(\{f(0)\}) = \mathbb{R}^n$). Alternatively we might observe that for a continuous function $\mathrm{Per}(f)$ is closed in $\mathbb{R}^n$ (proof immediate), and that a function with period group $\mathbb{R}^n$ is necessarily constant.

(iii) We have

$$A_r f(x) = \frac{1}{v_n\, r^n} \int_{B(x,r]} f(y)\, dy = \frac{1}{v_n\, r^n} \int_{rB} f(x+t)\, dt,$$

and when written in this form it is clear that $A_r f(x+a) = A_r f(x)$ for every $x \in \mathbb{R}^n$ and every $a \in \mathrm{Per}(f)$.

(iv) Considering $A_r f$, by (iii) and (ii) we have that $A_r f$ is a constant, for every $r > 0$. The sequence of constants $A_r f$ converges a.e. in $\mathbb{R}^n$ to f by the differentiation theorem; of course a pointwise limit of a sequence of constant functions is necessarily a constant function, so that f is constant a.e.

(v) Clearly $\arctan f$ is in $L^\infty(\mathbb{R}^n)$, and hence is locally summable. It is also periodic with the same period group of f; by (iv) $\arctan f(x) = k$ for a.e. $x \in \mathbb{R}^n$; then $f(x) = \tan k$ for a.e. $x \in \mathbb{R}^n$.

(vi) Let G be an additive subgroup of $\mathbb{R}^n$. If G is measurable and $m(G) > 0$ the difference theorem says that $G - G$ is a nbhd of 0; but $G - G = G$ since G is a subgroup. Now if $\delta B \subseteq G$ for some $\delta > 0$ then $G = \mathbb{R}^n$: given $x \in \mathbb{R}^n$ there is a positive integer k such that $x/k \in \delta B$ (simply take $k > |x|/\delta$) so that $x = k(x/k) \in G$.

Remark. Nonmeasurable subgroups do exist. We know that $\mathbb{R}^n$ is a $\mathbb{Q}-$vector space of dimension $\mathfrak{c}$. If we consider a proper $\mathbb{Q}-$vector subspace of $\mathbb{R}^n$ of finite or countable codimension over $\mathbb{Q}$, its additive group G has countable index, i.e. its set $\mathbb{R}^n/G$ of cosets is countable (it is a $\mathbb{Q}-$vector space of finite dimension); if G were measurable, it could not have zero measure, since $\mathbb{R}^n$ is a countable union of translates of G; but then G cannot be a proper subgroup of $\mathbb{R}^n$; the contradiction can be avoided only if G is non measurable.

□

8.8. Change of variables in multiple integrals.

Proposition. *Let U, V be open subsets of $\mathbb{R}^n$, and let $\phi : U \to V$ be a C^1 diffeomorphism. Then the image measure of the measure $d\mu = |\det \phi'(x)|\, dx$ on U is Lebesgue measure on V.*

This proposition will be proved shortly by induction on the dimension n. First observe that the Proposition is equivalent to any one of the following statements (the proof of this fact is left as an exercise):

- *For every compact interval Q of $\mathbb{R}^n$ contained in U we have*

$$m(\phi(Q)) = \int_Q |\det \phi'(x)|\, dx.$$

- *For every positive Borel measurable function $f : V \to [0, \infty]$ we have*

$$\int_V f(y)\, dy = \int_U f(\phi(x))\, |\det \phi'(x)|\, dx.$$

- Variable change in multiple integrals *The function $f : V \to \mathbb{K}$ is in $L^1_m(V)$ if and only if we have $f \circ \phi\, |\det \phi'| \in L^1_m(U)$, and:*

$$\int_V f(y)\, dy = \int_U f(\phi(x))\, |\det \phi'(x)|\, dx.$$

Proof. For $n = 1$ see 4.1.13. Assume that $\phi : U \to V$ and $\psi : V \to W$ are C^1 diffeomorphisms, both verifying the proposition. Then $\psi \circ \phi : U \to W$ also verifies the proposition. In fact, if $f : W \to [0, \infty]$ is positive measurable we have

$$\int_W f(z)\, dz = \int_V f(\psi(y))\, |\det \psi'(y)|\, dy$$

this is because ψ verifies the proposition; and since also ϕ verifies the proposition we have

$$\int_V f(\psi(y))\, |\det \psi'(y)|\, dy = \int_U f(\psi(\phi(x)))\, |\det \psi'(\phi(x))|\, |\det \phi'(x)|\, dx,$$

so that

$$\int_W f(z)\, dz = \int_U f \circ \psi \circ \phi(x)\, |\det \psi'(\phi(x))\, \det \phi'(x)|\, dx,$$

and since $\det \psi'(\phi(x)) \det \phi'(x) = \det(\psi \circ \phi)'(x)$, we conclude that $\psi \circ \phi$ verifies the proposition.

We say that a diffeomorphism $\phi : U \to V$ verifies locally the Proposition if for every $p \in U$ there is an open nbhd U_p of p in U such that, denoting by $\phi_p : U_p \to V_p = \phi(U_p)$ the induced diffeomorphism, ϕ_p verifies the proposition. If ϕ verifies the proposition locally, then ϕ verifies the proposition. In fact, given a compact interval Q contained in U, pick a nbhd U_p for every $p \in U$ where the proposition is verified; for every p pick then a compact cube Q_p centered at p and contained in U_p. By compactness of Q there is a finite subset $p(1), \dots, p(m) \in Q$ such that $Q \subseteq \bigcup_{k=1}^m Q_{p(k)}$. Considering the subalgebra of parts of Q generated by the intervals $Q_{p(k)} \cap Q$, we can write Q as a finite disjoint union of intervals $I(1), \dots, I(r)$, each contained in some $U_{p(k)}$, where the proposition holds. Then

$$m(\phi(Q)) = m\left(\bigcup_{j=1}^r \phi(I(j))\right) = \sum_{j=1}^r m(\phi(I(j))) = \sum_{j=1}^r \int_{I(j)} |\det \phi'(x)|\, dx = \int_Q |\det \phi'(x)|\, dx.$$

Next, assuming the inductive hypothesis that the proposition holds for every C^1 diffeomorphism between open subsets of $\mathbb{R}^n$ we prove that the proposition holds if U, V are open sets in $\mathbb{R}^{n+1}$ and $\phi : U \to V$ is of the form (we denote by (t, x) the independent variable of $\mathbb{R}^{n+1}$, i.e. $t \in \mathbb{R}$ and $x \in \mathbb{R}^n$)

$$\phi(t, x) = (\phi_0(t, x), \dots, \phi_n(t, x)) \qquad \text{with } \phi_0(t, x) = t,$$

that is the first component of ϕ is the identity of $\mathbb{R}$, $\phi_0(t, x) = t$, in other words we have

$$\phi(t, x) = (t, \phi_1(t, x), \dots, \phi_n(t, x)).$$

We may write ϕ as $\phi(t, x) = (t, \psi_t(x))$, where for each $t \in \mathrm{pr}_0(U) = J (= \mathrm{pr}_0(V))$ the function $\psi_t(x) = (\phi_1(t, x), \dots, \phi_n(t, x))$ is a diffeomorphism of the $t-$section $U(t) = \{x \in \mathbb{R}^n : (t, x) \in U\}$ of U onto the $t-$section $V(t) = \{y \in \mathbb{R}^n : (t, y) \in V\}$ of V. Given a positive Borel measurable $f : V \to [0, \infty]$ we have, by Tonelli's theorem:

$$\int_V f(t, y)\, dm(t, y) = \int_{t \in J} \left(\int_{V(t)} f(t, y)\, dy \right) dt;$$

Now $\psi_t : U(t) \to V(t)$ is a diffeomorphism of open subsets of $\mathbb{R}^n$, so that by the inductive hypothesis we get, for every $t \in J$:

$$\int_{V(t)} f(t, y)\, dy = \int_{U(t)} f(t, \psi_t(x))\, |\det \partial_x \psi_t(x)|\, dx,$$

where $\partial_x \psi_t(x)$ is the jacobian matrix of ψ_t with respect to the $x-$variables. Then $\det \partial_x \psi_t(x) = \det \phi'(t, x)$ for every $(t, x) \in U$. We have obtained:

$$\int_V f(t, y)\, dm(t, y) = \int_{t \in J} \left(\int_{U(t)} f(\phi(t, x))\, |\det \phi'(t, x)|\, dx \right) dt;$$

and again by Tonelli's theorem we get

$$\int_{t \in J} \left(\int_{U(t)} f(\phi(t, x))\, |\det \phi'(t, x)|\, dx \right) dt = \int_U f(\phi(t, x))\, |\det \phi'(t, x)|\, dm(t, x).$$

We now prove that locally every C^1 diffeomorphism $\phi : U \to V$ between open subsets U, V of $\mathbb{R}^{n+1}$ may be factored as the composition of two diffeomorphisms each of which is the identity on some variable. Given $p \in U$, by a suitable permutation of coordinates on domain and range we may assume that $\partial_1 \varphi(p) \neq 0$. The map $\alpha : U \to \mathbb{R}^{n+1}$ defined by $\alpha(x) = (\varphi_1(x), x_2, \dots, x_{n+1})$ is locally invertible at p, having jacobian determinant $\partial_1 \varphi(p) \neq 0$, and hence induces a C^1 diffeomorphism of an open nbhd U_p of p onto an open subset W of $\mathbb{R}^{n+1}$. Let $\beta : W \to U_p$ be the inverse diffeomorphism, and set $\psi = \varphi \circ \beta$. Then ψ is a C^1 diffeomorphism from W onto $V_p = \varphi(U_p)$, and $\varphi = \psi \circ \alpha$. Now α is the identity on the last n variables, and ψ is the identity on the first variable, so they both verify the Proposition. The proof is completed.

$\square$

Contents