19 The Radon-Nikodým Theorem and other applications of martingales Solutions to Problems 19.1–19.18

Problem 19.1 This problem is intimately linked with problem 19.7.

Without loss of generality we assume that μ and ν are finite measures, the case for σ-finite μ and arbitrary ν is exactly as in the proof of Theorem 19.2.

Let $(A_j)_j$ be as described in the problem and define the finite σ-algebras $\mathcal{A}_n := \sigma(A_1, \dots, A_n)$. Using the hint we can achieve that

$$\mathcal{A}_n = \sigma\big(C_1^n, \dots, C_{\ell(n)}^n\big)$$

with mutually disjoint C_j^k's and $\ell(n) \leqslant 2^n + 1$ and $\dot\bigcup_j C_j^n = X$. Then the construction of Example 19.5 yields a countably-indexed martingale since the σ-algebras $\mathcal{A}_j$ are increasing.

This means, that the countable version of the martingale convergence theorem is indeed enough for the proof.

Problem 19.2 Using simply the Radon-Nikodým theorem, Theorem 19.2, gives

$$\forall\, t \quad \exists\, p_t(x) \;\text{ such that }\; \nu_t(dx) = p_t(x) \cdot \mu_t(dx)$$

with a measurable function $x \mapsto p_t(x)$; it is, however, far from being clear that $(t, x) \mapsto p_t(x)$ is jointly measurable.

A slight variation of the proof of Theorem 19.2 allows us to incorporate parameters provided the families of measures are measurable w.r.t. these parameters. Following the hint we set (notation as in the proof of 19.2)

$$p_\alpha(t, x) := \sum_{A \in \alpha} \frac{\nu_t(A)}{\mu_t(A)}\, I_A(x)$$

with the agreement that $\frac{0}{0} := 0$ (note that $\frac{a}{0}$ with $a \neq 0$ will not turn up because of the absolute continuity of the measures!). Since $t \mapsto \mu_t(A)$ and $t \mapsto \mu_t(A)$ are measurable, the above sum is measurable so that

$$(t, x) \mapsto p(t, x)$$

is a jointly measurable function. If we can show that

$$\lim_{\alpha} p_\alpha(t,x) = p(t,x)$$

exists (say, in L^1, t being fixed) then the limiting function is again jointly measurable.

Using exactly the arguments of the proof of Theorem 19.2 with t fixed we can confirm that this limit exists and defines a jointly measurable function with the property that

$$\nu_t(dx) = p(t,x) \cdot \nu_t(dx).$$

Because of the a.e. uniqueness of the Radon-Nikodým density the functions $p(t,x)$ and $p_t(x)$ coincide, for every t a.e. as functions of x; without additional assumptions on the nature of the dependence on the parameter, the exceptional set may, though, depend on t!

Problem 19.3 We write $u^\pm$ for the positive resp. negative parts of $u \in \mathcal{L}^1(\mathcal{A})$, i.e. $u = u^+ - u^-$ and $u^\pm \geqslant 0$. Fix such a function u and define

$$\nu^\pm(F) := \int_F u^\pm(x)\,\mu(dx), \quad \forall\, F \in \mathcal{F}.$$

Clearly, $\nu^\pm$ are measures on the σ-algebra $\mathcal{F}$. Moreover

$$\forall\, N \in \mathcal{F},\ \mu(N) = 0 \implies \nu^\pm(N) = \int_N u^\pm\,d\mu = 0$$

which means that $\nu^\pm \ll \mu$. By the Radon-Nikodým theorem, Theorem 19.2 and its Corollary 19.6, we find (up to null-sets unique) positive functions $f^\pm \in \mathcal{L}^1(\mathcal{F})$ such that

$$\nu^\pm(F) = \int_F f^\pm\,d\mu \quad \forall\, F \in \mathcal{F}.$$

Thus, $u^{\mathcal{F}} := f^+ - f^- \in \mathcal{L}^1(\mathcal{F})$ clearly satisfies

$$\int_F u^{\mathcal{F}}\,d\mu = \int_F u\,d\mu \quad \forall\, F \in \mathcal{F}.$$

To see uniqueness, we assume that $w \in \mathcal{L}^1(\mathcal{F})$ also satisfies

$$\int_F w\,d\mu = \int_F u\,d\mu \quad \forall\, F \in \mathcal{F}.$$

Since then

$$\int_F u^{\mathcal{F}}\,d\mu = \int_F w\,d\mu \quad \forall\, F \in \mathcal{F}.$$

we can choose $f := \{w > u^{\mathcal{F}}\}$ and find

$$0 = \int_{\{w>u^{\mathcal{F}}\}} (w - u^{\mathcal{F}})\,d\mu$$

which is only possible if $\mu(\{w > u^{\mathcal{F}}\}) = 0$. Similarly we conclude that $\mu(\{w < u^{\mathcal{F}}\}) = 0$ from which we get $w = u^{\mathcal{F}}$ almost everywhere.

Reformulation of the submartingale property.

Recall that $(u_j, \mathcal{A}_j)_j$ is a submartingale if, for every j, $u_j \in \mathcal{L}^1(\mathcal{A}_j)$ and if

$$\int_A u_j\,d\mu \leqslant \int_A u_{j+1}\,d\mu \quad \forall A \in \mathcal{A}_j,\ \forall\, j.$$

We claim that this is equivalent to saying

$$u_j \leqslant u_{j+1}^{\mathcal{A}_j} \quad \text{almost everywhere, } \forall\, j.$$

The direction '⇒' is clear. To see '⇐' we fix j and observe that, since

$$\int_A u_j\,d\mu \leqslant \int_A u_{j+1}\,d\mu = \int_A u_{j+1}^{\mathcal{A}_j}\,d\mu \quad \forall A \in \mathcal{A}_j,$$

we get, in particular, for $A := \{u_{j+1}^{\mathcal{A}_j} < u_j\} \in \mathcal{A}_j$,

$$0 \leqslant \int_{\{u_{j+1}^{\mathcal{A}_j} < u_j\}} (u_{j+1}^{\mathcal{A}_j} - u_j)\,d\mu$$

which is only possible if $\mu(\{u_{j+1}^{\mathcal{A}_j} < u_j\}) = 0$.

Problem 19.4 The assumption $\nu \leqslant \mu$ immediately implies $\nu \ll \mu$. Indeed,

$$\mu(N) = 0 \implies 0 \leqslant \nu(N) \leqslant \mu(N) = 0 \implies \nu(N) = 0.$$

Using the Radon-Nikodým theorem, Theorem 19.2 we conclude that there exists a measurable function $f \in \mathcal{M}^+(\mathcal{A})$ such that $\nu = f \cdot \mu$. Assume that $f > 1$ on a set of positive μ-measure. Without loss of generality we may assume that the set has finite measure, otherwise we would consider the intersection $A_k \cap \{f > 1\}$ with some exhausting sequence $A_k \uparrow X$ and $\mu(A_k) < \infty$.

Then, for sufficiently small $\epsilon > 0$ we know that $\mu(\{f \geqslant 1+\epsilon\}) > 0$ and so

$$\begin{aligned}\nu(\{f \geqslant 1+\epsilon\}) &= \int_{\{f\geqslant 1+\epsilon\}} f\,d\mu \\ &\geqslant (1+\epsilon)\int_{\{f\geqslant 1+\epsilon\}} d\mu \\ &\geqslant (1+\epsilon)\mu(\{f \geqslant 1+\epsilon\}) \\ &\geqslant \mu(\{f \geqslant 1+\epsilon\})\end{aligned}$$

which is impossible.

Problem 19.5 Because of our assumption both $\mu \ll \nu$ and $\nu \ll \mu$ which means that we know

$$\nu = f\mu \quad \text{and} \quad \mu = g\nu$$

for positive measurable functions f, g which are a.e. unique. Moreover,

$$\nu = f\mu = f \cdot g\nu$$

so that $f \cdot g$ is almost everywhere equal to 1 and the claim follows.

Because of Corollary 19.6 it is clear that $f, g < \infty$ a.e. and, by the same argument, $f, g > 0$ a.e.

Note that we do not have to specify *w.r.t. which measure* we understand the 'a.e.' since their null sets coincide anyway.

Problem 19.6 Take Lebesgue measure $\lambda := \lambda^1$ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and the function $f(x) := x + \infty \cdot \mathbf{1}_{[0,1]^c}(x)$. Then $f \cdot \lambda$ is certainly not σ-finite.

Problem 19.7 Since both μ and ν are σ-finite, we can restrict ourselves, using the technique of the Proof of Theorem 19.2 to the case where μ and ν are finite. All we have to do is to pick an exhaustion $(K_\ell)_\ell$, $K_\ell \uparrow X$ such that $\mu(K_\ell), \mu(K_\ell) < \infty$ and to consider the measures $\mathbf{1}_{K_\ell}\mu$ and $\mathbf{1}_{K_\ell}\nu$ which clearly inherit the absolute continuity from μ and ν.

Using the Radon-Nikodým theorem (Theorem 19.2) we get that

$$\mu_j \ll \nu_j \implies \mu_j = u_j \cdot \nu_j$$

with an $\mathcal{A}_j$-measurable positive density u_j. Moreover, since μ is a finite measure,

$$\int_X u_j\,d\nu = \int_X u_j\,d\nu_j = \int_X d\mu_j = \mu_j(X) < \infty$$

so that all the $(u_j)_j$ are ν-integrable. Using exactly the same argument as at the beginning of the proof of Theorem 19.2 (ii)⇒(i), we get that $(u_j)_j$ is even uniformly ν-integrable. Finally, $(u_j)_j$ is a martingale (given the measure ν), since for $j, j+1$ and $A \in \mathcal{A}_j$ we have

$$\begin{aligned}\int_A u_{j+1}\,d\nu &= \int_A u_{j+1}\,d\nu_{j+1} \\ &= \int_A d\mu_{j+1} && (u_{j+1}\cdot\nu_{j+1} = \mu_{j+1}) \\ &= \int_A d\mu_j && (A \in \mathcal{A}_j) \\ &= \int_A u_j\,d\nu_j && (\mu_j = u_j \cdot \nu_j) \\ &= \int_A u_j\,d\nu\end{aligned}$$

and we conclude that $u_j \to u_\infty$ a.e. and in $L^1(\nu)$ for some limiting function u_∞ which is still $L^1(\nu)$ and also $\mathcal{A}_\infty := \sigma(\bigcup_{j\in\mathbb{N}} \mathcal{A}_j)$-measurable. Since, by assumption, $\mathcal{A}_\infty = \mathcal{A}$, this argument shows also that

$$\mu = u_\infty \cdot \nu$$

and it reveals that

$$u_\infty = \frac{d\mu}{d\nu} = \lim_j \frac{d\mu_j}{d\nu_j}\,.$$

Problem 19.8 This problem is somewhat ill-posed. We should first embed it into a suitable context, say, on the measurable space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Denote by $\lambda = \lambda^1$ one-dimensional Lebesgue measure. Then

$$\mu = \mathbf{1}_{[0,2]}\lambda \quad\text{and}\quad \nu = \mathbf{1}_{[1,3]}\lambda$$

and from this it is clear that

$$\nu = \mathbf{1}_{[1,2]}\nu + \mathbf{1}_{(2,3]}\nu = \mathbf{1}_{[1,2]}\lambda + \mathbf{1}_{(2,3]}\lambda$$

and from this we read off that

$$\mathbf{1}_{[1,2]}\nu \ll \mu$$

while

$$\mathbf{1}_{(2,3]}\nu \perp \mu.$$

It is interesting to note how 'big' the null-set of ambiguity for the Lebesgue decomposition is—it is actually $\mathbb{R} \setminus [0,3]$ a, from a Lebesgue (i.e. λ) point of view, huge and infinite set, but from a μ-ν-perspective a negligible, name null, set.

Problem 19.9 Since we deal with a bounded measure we can use $F(x) := \mu(-\infty, x)$ rather than the more cumbersome definition for F employed in Problem 7.9 (which is good for locally finite measures!).

With respect to one-dimensional Lebesgue measure λ we can decompose μ according to Theorem 19.9 into

$$\mu = \mu^\circ + \mu^\perp \quad \text{where} \quad \mu^\circ \ll \lambda, \ \ \mu^\perp \perp \lambda.$$

Now *define* $\mu_2 := \mu^\circ$ and $F_2 := \mu^\circ(-\infty, x)$. We have to prove property (2). For this we observe that μ° is a finite measure (since $\mu^\circ \leqslant \mu$ and that, therefore, $\mu^\circ = f \cdot \lambda$ with a function $f \in L^1(\lambda)$. Thus, for every $R > 0$

$$\begin{aligned} F(y_j) - F(x_j) &= \mu^\circ(x_j, y_j) \\ &= \int_{(x_j,y_j)} f(t)\,\lambda(dt) \\ &\leqslant R \int_{(x_j,y_j)} \lambda(dt) + \lambda\big(\{f \geqslant R\} \cap (x_j, y_j)\big) \\ &\leqslant R \int_{(x_j,y_j)} \lambda(dt) + \frac{1}{R}\int_{(x_j,y_j)} f\,d\lambda \end{aligned}$$

where we used the Markov inequality, cf. Proposition 10.12, in the last step. Summing over $j = 1, 2, \ldots, N$ gives

$$\sum_{j=1}^{N} |F_2(y_j) - F_2(x_j)| \leqslant R \cdot \delta + \frac{1}{R}\int f\,d\lambda$$

since $\dot\bigcup_j (x_j, y_j) \subset \mathbb{R}$. Now we choose for given $\epsilon > 0$

$$\text{first} \quad R := \frac{\int f\,d\lambda}{\epsilon} \qquad \text{and then} \quad \delta := \frac{\epsilon}{R} = \frac{\epsilon^2}{\int f\,d\lambda}$$

to confirm that

$$\sum_{j=1}^{N} |F_2(y_j) - F_2(x_j)| \leqslant 2\epsilon$$

this settles (2).

Now consider the measure $\mu^\perp$. Its distribution function $F^\perp(x) := \mu^\perp(-\infty, x)$ is increasing, left-continuous but not necessarily continuous. Such a function has, by Lemma 13.12 at most countably many discontinuities (jumps), which we denote by J. Thus, we can write

$$\mu^\perp = \mu_1 + \mu_3$$

with the jump (or saltus) $\Delta F(y) := F(y+) - F(y-)$ if $y \in J$.

$$\mu_1 := \sum_{y \in J} \Delta F(y) \cdot \delta_y, \quad \text{and} \quad \mu_3 := \mu^{\perp} - \mu_1;$$

μ_1 is clearly a measure (the sum being countable) with $\mu_1 \leqslant \mu^{\perp}$ and so is, therefore, μ_2 (since the defining difference is always positive). The corresponding distribution functions are

$$F_1(x) := \sum_{y \in J, y < x} \Delta F(y)$$

(called the jump or saltus function) and

$$F_2(x) := F^{\perp}(x) - F_1(x).$$

It is clear that F_2 is increasing and, more importantly, continuous so that the problem is solved.

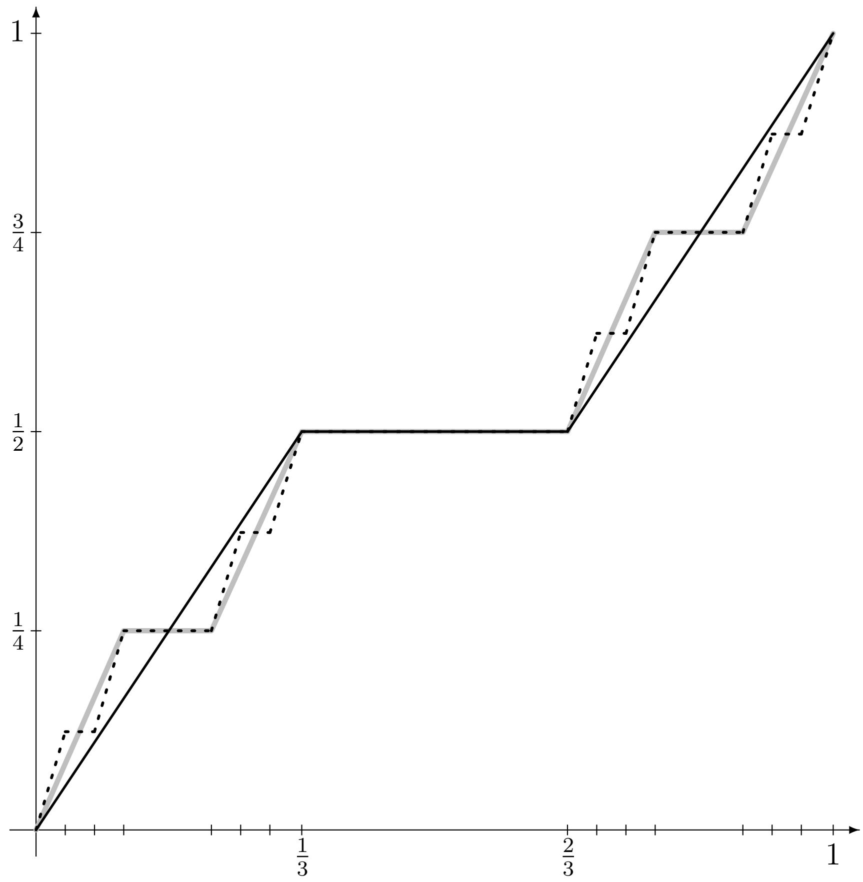
It is interesting to note that our problem shows that we can decompose every left- or right-continuous monotone function into an absolutely continuous and singular part and the singular part again into a continuous and discontinuous part:

$$g = g_{\text{ac}} + g_{\text{sc}} + g_{\text{sd}}$$

where

- g —is a monotone left- or right-continuous function;
- g_{ac} —is a monotone absolutely continuous (and in particular continuous) function;
- g_{sc} —is a monotone continuous but singular function;
- g_{sd} —is a monotone discontinuous (even: pure jump), but nevertheless left- or right-continuous, and singular function.

Problem 19.10 (i) In the following picture F_1 is represented by a black line, F_2 by a grey line and F_3 is a dotted black line.



(ii),(iii) The construction of the F_k's also shows that

$$|F_k(x) - F_{k+1}(x)| \leqslant \frac{1}{2^{k+1}}$$

since we modify F_k only on a set J_{k+1}^{ℓ} by replacing a diagonal line by a combination of diagonal-flat-diagonal and all this happens only within a range of 2^{-k} units. Since the flat bit is in the middle, we get that the maximal deviation between F_k and F_{k+1} is at most $\frac{1}{2} \cdot 2^{-k}$. Just look at the pictures!

Thus the convergence of $F_k \to F$ is uniform, i.e. it preserves continuity and F is continuous as all the F_k's are. That F is increasing is already inherited from the pointwise limit of the F_k's:

$$\begin{aligned} x < y \implies \forall\, k \,:\, F_k(x) \leqslant F_k(y) \\ \implies F(x) = \lim_k F_k(x) \leqslant \lim_k F_k(y) = F(y). \end{aligned}$$

(iv) Let C denote the Cantor set. Then for $x \in [0,1] \setminus C$ we find k and ℓ such that $x \in J_k^\ell$ (which is an open set!) and, since on those pieces F_k and F do not differ any more

$$F_k(x) = F(x) \implies F'(x) = F_k'(x) = 0$$

where we used that $F_k | J_k^\ell$ is constant. Since $\lambda(C) = 0$ (see Problem 7.10) we have $\lambda([0,1] \setminus C) = 1$ so that F' exists a.e. and satisfies $F' = 0$ a.e.

(v) We have $J_k^\ell = (a_\ell, b_\ell)$ (we suppress the dependence of a_ℓ, b_ℓ on k with, because of our ordering of the middle-thirds sets (see the problem):

$$a_1 < b_1 < a_2 < \cdots < a_{2^k-1} < b_{2^k-1}$$

and

$$\sum_{\ell=1}^{2^k-1} \big[F(b_\ell) - F(a_\ell)\big] = F(b_{2^k-1}) - F(a_1) \xrightarrow{k\to\infty} F(1) - F(0) = 1$$

while (with the convention that $a_0 := 0$)

$$\sum_{\ell=1}^{2^k-1} (a_\ell - b_{\ell-1}) \xrightarrow{k\to\infty} 0.$$

This leads to a contradiction since, because of the first equality, the sum

$$\sum_{\ell=1}^{2^k-1} \big[F(a_\ell) - F(b_{\ell-1})\big]$$

will never become small.

Problem 19.11 We can assume that $VX_j < \infty$, otherwise the inequality would be trivial.

Note that the random variables $X_j - EX_j$, $j = 1, 2, \ldots, n$ are still independent and, of course, centered (= mean-zero). Thus, by Example 17.3(x) we get that

$$M_k := \sum_{j=1}^{k} (X_j - EX_j) \quad \text{is a martingale}$$

and, because of Example 17.3(v), $(|M_k|)_k$ is a submartingale. Applying (19.12) in this situation proves the claimed inequality since

$$VM_n = E(M_n^2) \qquad \text{(since } EM_n = 0\text{)}$$

$$= \sum_{j=1}^{n} E(X_j^2)$$

where we used, for the last equality, what probabilists call *Theorem of Bienaymé* for the independent random variables X_j:

$$\begin{aligned}
E(M_n^2) &= \sum_{j,k=1}^{n} E\big[(X_j - EX_j)(X_k - EX_k)\big] \\
&= \sum_{j=k=1}^{n} E\big[(X_j - EX_j)^2\big] + \sum_{j\neq k} E\big[(X_j - EX_j)\big]E\big[(X_k - EX_k)\big] \\
&\qquad\qquad \text{(by independence)} \\
&= \sum_{j=k=1}^{n} E\big[(X_j - EX_j)^2\big] \\
&= \sum_{j=1}^{n} E\big[M_j^2\big] \\
&= \sum_{j=1}^{n} VM_j.
\end{aligned}$$

Problem 19.12 (i) As in the proof of Theorem 19.12 we find

$$\begin{aligned}
\int u^p\,d\mu &\overset{(13.8)}{=} p\int_0^\infty s^{p-1}\,\mu\left(\{u \geqslant s\}\right)\,ds \\
&\leqslant p\int_0^\infty s^{p-2}\bigg(\int \mathbf{1}_{\{u\geqslant s\}}(x)w(x)\mu(dx)\bigg)\,ds \\
&= p\int \bigg(\int_0^\infty \mathbf{1}_{[0,u(x)]}(s)s^{p-2}\,ds\bigg)w(x)\,\mu(dx) \\
&= p\int \frac{u(x)^{p-1}}{p-1}\,w(x)\,\mu(dx) \\
&= \frac{p}{p-1}\int u^{p-1}w\,d\mu
\end{aligned}$$

Note that this inequality is meant in $[0,+\infty]$, i.e. we allow the cases $a \leqslant +\infty$ and $+\infty \leqslant +\infty$.

(ii) Pick conjugate numbers $p,q \in (1,\infty)$, i.e. $q = \frac{p}{p-1}$. Then we can rewrite the result of (i) and then apply Hölder's inequality to get

$$\|u\|_p^p \leqslant \frac{p}{p-1}\int u^{p-1}w\,d\mu$$

$$\begin{aligned}
&\leqslant \frac{p}{p-1}\Big(\int u^{(p-1)q}\,d\mu\Big)^{1/q}\Big(\int w^p\,d\mu\Big)^{1/p}\\
&= \frac{p}{p-1}\Big(\int u^p\,d\mu\Big)^{1-1/p}\|w\|_p\\
&= \frac{p}{p-1}\|u\|_p^{p-1}\cdot\|w\|_p
\end{aligned}$$

and the claim follows upon dividing both sides by $\|u\|_p^{p-1}$. (Here we use the finiteness of this expression, i.e. the assumption $u \in \mathcal{L}^p$).

Problem 19.13 Only the first inequality needs proof. Note that

$$\max_{1\leqslant j\leqslant N}\int |u_j|^p\,d\mu \leqslant \int \max_{1\leqslant j\leqslant N}|u_j|^p\,d\mu = \int u_N^*\,d\mu$$

from which the claim easily follows.

Problem 19.14 Let $(A_k)_k \subset \mathcal{A}_0$ be an exhausting sequence, i.e. $A_k \uparrow X$ and $\mu(A_k) < \infty$. Since $(u_j)_j$ is L^1-bounded, we know that

$$\sup_j \|u_j\|_p \leqslant c < \infty$$

and we find, using Hölder's inequality with $\frac{1}{p}+\frac{1}{q}=1$

$$\int |\mathbf{1}_{A_k}u_j|\,d\mu \leqslant \big(\mu(A_k)\big)^{1/q}\cdot\|u_j\|_p \leqslant c\,\big(\mu(A_k)\big)^{1/q}$$

uniformly for all $j \in \mathbb{N}$. This means that the martingale $(\mathbf{1}_{A_k}u_j)_j$ (see the solution to Problem 18.8) is L^1-bounded and we get, as in Problem 18.8 that for some unique function u

$$\lim_j \mathbf{1}_{A_k}u_j = \mathbf{1}_{A_k}u \quad \forall\, k$$

a.e., hence $u_j \xrightarrow{j\to\infty} u$ a.e. Using Fatou's Lemma we get

$$\begin{aligned}
\int |u|^p\,d\mu &= \int \liminf_j |u_j|^p\,d\mu\\
&\leqslant \liminf \int |u_j|^p\,d\mu\\
&\leqslant \sup_j \int |u_j|^p\,d\mu \;<\; \infty
\end{aligned}$$

which means that $u \in L^p$.

For each $k \in \mathbb{N}$ the martingale $(\mathbf{1}_{A_k} u_j)_j$ is also uniformly integrable: using Hölder's and Markov's inequalities we arrive at

$$\begin{aligned}\int_{\{\mathbf{1}_{A_k}|u_j|>\mathbf{1}_{A_k}R\}} \mathbf{1}_{A_k}|u_j|\,d\mu &\leqslant \int_{\{|u_j|>R\}} \mathbf{1}_{A_k}|u_j|\,d\mu \\ &\leqslant \big(\mu\{|u_j|>R\}\big)^{1/q}\|u_j\|_p \\ &\leqslant \left(\frac{1}{R^p}\,\|u_j\|_p^p\right)^{1/q}\|u_j\|_p \\ &\leqslant \frac{C^{p/q+1}}{R^{p/q}}\end{aligned}$$

and the latter tends, uniformly for all j, to zero as $R \to \infty$. Since $\mathbf{1}_{A_k} \cdot R$ is integrable, the claim follows.

Thus, Theorem 18.6 applies and shows that for $u_\infty := u$ and every k the family $(u_j \mathbf{1}_{A_k})_{j\in\mathbb{N}\cup\{\infty\}}$ is a martingale. Because of Example 17.3(vi) $(|u_j|^p \mathbf{1}_{A_k})_{j\in\mathbb{N}\cup\{\infty\}}$ is a submartingale and, therefore, for all $k \in \mathbb{N}$

$$\int |\mathbf{1}_{A_k} u_j|^p\,d\mu \leqslant \int |\mathbf{1}_{A_k} u_{j+1}|^p\,d\mu \leqslant \int |\mathbf{1}_{A_k} u_\infty|^p\,d\mu = \int |\mathbf{1}_{A_k} u|^p\,d\mu,$$

Since, by Fatou's lemma

$$\int |\mathbf{1}_{A_k} u|^p\,d\mu = \int \liminf_j |\mathbf{1}_{A_k} u_j|^p\,d\mu \leqslant \liminf_j \int |\mathbf{1}_{A_k} u_j|^p\,d\mu$$

we see that

$$\int |\mathbf{1}_{A_k} u|^p\,d\mu = \lim_j \int |\mathbf{1}_{A_k} u_j|^p\,d\mu = \sup_j \int |\mathbf{1}_{A_k} u_j|^p\,d\mu.$$

Since suprema interchange, we get

$$\begin{aligned}\int |u|^p\,d\mu &= \sup_k \int |\mathbf{1}_{A_k} u|^p\,d\mu \\ &= \sup_k \sup_j \int |\mathbf{1}_{A_k} u_j|^p\,d\mu \\ &= \sup_j \sup_k \int |\mathbf{1}_{A_k} u_j|^p\,d\mu \\ &= \sup_j \int |u_j|^p\,d\mu\end{aligned}$$

and Riesz's convergence theorem, Theorem 12.10, finally proves that $u_j \to u$ in L^p.

Problem 19.15 Since f_k is a martingale and since

$$\begin{aligned}\int |f_k|\,d\mu &\leqslant \sum_{z\in 2^{-k}\mathbb{Z}Z^n} \frac{1}{\lambda^n(Q_k(z))} \int_{Q_k(z)} |f|\,d\lambda^n \int \mathbf{1}_{Q_k(z)}\,d\lambda^n \\ &= \sum_{z\in 2^{-k}\mathbb{Z}Z^n} \int_{Q_k(z)} |f|\,d\lambda^n \\ &= \int |f|\,d\lambda^n < \infty\end{aligned}$$

we get from the martingale convergence theorem 18.2 that

$$f_\infty := \lim_k f_k$$

exists almost everywhere and that $f_\infty \in \mathcal{L}^1(\mathcal{B})$. The above calculation shows, on top of that, that for any set $Q \in \mathcal{A}_k^{[0]}$

$$\int_Q f_k\,d\lambda^n = \int_Q f\,d\lambda^n$$

and

$$\int_Q |f_k|\,d\lambda^n \leqslant \int_Q |f|\,d\lambda^n$$

which means that, using Fatou's Lemma,

$$\int_Q |f_\infty|\,d\lambda^n \leqslant \liminf_k \int_Q |f_k|\,d\lambda_n \leqslant \int_Q |f|\,d\lambda_n$$

for all $Q \in \mathcal{A}_k^{[0]}$ and any k. Since $\mathcal{S} = \bigcup_k \mathcal{A}_k^{[0]}$ is a semi-ring and since on both sides of the above inequality we have measures, this inequality extends to $\mathcal{B} = \sigma(\mathcal{S})$ (cf. Lemma 15.6) and we get

$$\int_B |f_\infty|\,d\lambda^n \leqslant \int_B |f|\,d\lambda^n.$$

Since f_∞ and f are $\mathcal{B}$-measurable, we can take $B = \{|f_\infty| > |f|\}$ and we get that $f = f_\infty$ almost everywhere. This shows that $(f_k)_{k\in\mathbb{N}\cup\{\infty\}}$ is a martingale.

Thus all conditions of Theorem 18.6 are satisfied and we conclude that $(f_k)_k$ is uniformly integrable.

Problem 19.16 As one would expect, the derivative at x turns out to be $u(x)$. This is seen as follows (without loss of generality we can assume that $y > x$):

$$\begin{aligned}
&\left|\frac{1}{x-y}\left(\int_{[a,x]} u(t)\,dt - \int_{[a,x]} u(t)\,dt\right) - u(x)\right| \\
&= \left|\frac{1}{x-y}\int_{[x,y]} \big(u(t)-u(x)\big)\,dt\right| \\
&\leqslant \frac{1}{|x-y|}\int_{[x,y]} \big|u(t)-u(x)\big|\,dt \\
&\leqslant \frac{1}{|x-y|}\,|x-y|\,\sup_{t\in[x,y]}\big|u(t)-u(x)\big| \\
&= \sup_{t\in[x,y]}\big|u(t)-u(x)\big|
\end{aligned}$$

and the last expression tends to 0 as $|x-y| \to 0$ since u is uniformly continuous on compact sets.

If u is not continuous but merely of class L^1, we have to refer to Lebesgue's differentiation theorem, Theorem 19.20, in particular formula (19.21) which reads in our case

$$u(x) = \lim_{r\to 0}\frac{1}{2r}\int_{(x-r,x+r)} u(t)\,dt$$

for Lebesgue almost every $x \in (a,b)$.

Problem 19.17 We follow the hint: first we remark that by Lemma 13.12 we know that f has at most countably many discontinuities. Since it is monotone, we also know that $F(t) := f(t+) = \lim_{s>t,s\to t} f(s)$ exists and is finite for every t and that $\{f \neq F\}$ is at most countable (since it is contained in the set of discontinuities of f), hence a Lebesgue null set.

If f is right-continuous, $\mu(a,b] := f(b) - f(a)$ extends uniquely to a measure on the Borel-sets and this measure is locally finite and σ-finite. If we apply Theorem 19.9 to μ and $\lambda = \lambda^1$ we can write $\mu = \mu^\circ + \mu^\perp$ with $\mu^\circ \ll \lambda$ and $\mu^\perp \perp \lambda$. By Corollary 19.22 $D\mu^\perp = 0$ a.e. and $D\mu^\circ$ exists a.e. and we get a.e.

$$D\mu(x) = \lim_{r\to 0}\frac{\mu(x-r,x+r)}{2r} = \lim_{r\to 0}\frac{\mu^\circ(x-r,x+r)}{2r} + 0$$

and we can set $f'(x) = D\mu(x)$ which is a.e. defined. Where it is not defined, we put it equal to 0.

Now we get

$$\begin{aligned} f(b) - f(a) &= \mu(a, b] \\ &\geqslant \mu(a, b) \\ &= \int_{(a,b)} d\mu \\ &\geqslant \int_{(a,b)} d\mu^\circ \\ &= \int_{(a,b)} D\mu(x)\,\lambda(dx) \\ &= \int_{(a,b)} f'(x)\,\lambda(dx). \end{aligned}$$

The above estimates show that we get equality if f is continuous and also absolutely continuous w.r.t. Lebesgue measure.

Problem 19.18 Without loss of generality we may assume that $f_j(a) = 0$, otherwise we would consider the (still increasing) functions $x \mapsto f_j(x) - f_j(a)$ resp. their sum $x \mapsto s(x) - s(a)$. The derivatives are not influenced by this operation. As indicated in the hint call $s_n(x) := f_1(x) + \cdots + f_n(x)$ the nth partial sum. Clearly, s, s_n are increasing

$$\frac{s_n(x+h) - s_n(x)}{h} \leqslant \frac{s_{n+1}(x+h) - s_{n+1}(x)}{h} \leqslant \frac{s(x+h) - s(x)}{h}.$$

and possess, because of Problem 19.17, almost everywhere positive derivatives:

$$s_n'(x) \leqslant s_{n+1}'(x) \leqslant \cdots s'(x), \qquad \forall\, x \notin E$$

Note that the exceptional null-sets depend originally on the function s_n etc. but we can consider their (countable!!) union and get thus a universal exceptional null set E. This shows that the formally differentiated series

$$\sum_{j=1}^{\infty} f_j'(x) \quad \text{converges for all } x \notin E.$$

Since the sequence of partial sums is increasing, it will be enough to check that

$$s'(x) - s_{n_k}'(x) \xrightarrow{k\to\infty} 0 \quad \forall x \notin E.$$

Since, by assumption the sequence $s_k(x) \to s(x)$ we can choose a subsequence n_k in such a way that

$$s(b) - s_{n_k}(b) < 2^{-k} \quad \forall\, k \in \mathbb{N}.$$

Since

$$0 \leqslant s(x) - s_{n_k}(x) \leqslant s(b) - s_{n_k}(b)$$

the series

$$\sum_{k=1}^{\infty} (s(x) - s_{n_k}(x)) \leqslant \sum_{k=1}^{\infty} 2^{-k} < \infty \quad \forall\, x \in [a,b].$$

By the first part of the present proof, we can differentiate this series term-by-term and get that

$$\sum_{k=1}^{\infty} (s'(x) - s'_{n_k}(x)) \quad \text{converges} \quad \forall\, x \in (a,b) \setminus E$$

and, in particular, $s'(x) - s'_{n_k}(x) \xrightarrow{k\to\infty} 0$ for all $x \in (a,b) \setminus E$ which was to be proved.

20 Inner Product Spaces
Solutions to Problems 20.1–20.6

Problem 20.1 If we set $\mu = \delta_1 + \cdots + \delta_n$, $X = \{1, 2, \ldots, n\}$, $\mathcal{A} = \mathcal{P}(X)$ or $\mu = \sum_{j\in\mathbb{N}} \delta_j$, $X = \mathbb{N}$, $\mathcal{A} = \mathcal{P}(X)$, respectively, we can deduce 20.5(i) and (ii) from 20.5(iii).

Let us, therefore, only verify (iii). Without loss of generality (see Scholium 20.1 and also the complexification of a real inner product space in Problem 20.3) we can consider the real case where $L^2 = L^2_{\mathbb{R}}$.

- L^2 is a vector space — this was done in Remark 12.5.
- $\langle u, v\rangle$ is finite on $L^2 \times L^2$ — this is the Cauchy-Schwarz inequality 12.3.
- $\langle u, v\rangle$ is bilinear — this is due to the linearity of the integral.
- $\langle u, v\rangle$ is symmetric — this is obvious.
- $\langle v, v\rangle$ is definite, and $\|u\|_2$ is a Norm — cf. Remark 12.5.

Problem 20.2 (i) We prove it for the complex case—the real case is simpler. Observe that

$$\begin{aligned}\langle u \pm w, u \pm w\rangle &= \langle u, u\rangle \pm \langle u, w\rangle \pm \langle w, u\rangle + \langle w, w\rangle \\ &= \langle u, u\rangle \pm \langle u, w\rangle \pm \overline{\langle u, w\rangle} + \langle w, w\rangle \\ &= \langle u, u\rangle \pm 2\mathrm{Re}\,\langle u, w\rangle + \langle w, w\rangle.\end{aligned}$$

Thus,

$$\langle u + w, u + w\rangle + \langle u - w, u - w\rangle = 2\langle u, u\rangle + 2\langle w, w\rangle.$$

Since $\|v\|^2 = \langle v, v\rangle$ we are done.

(ii) (SP_1): Obviously,

$$0 < (u, u) = \tfrac{1}{4}\,\|2v\|^2 = \|v\|^2 \implies v \neq 0.$$

(SP_1): is clear.

(iii) Using at the point (*) below the parallelogram identity, we have

$4(u + v, w) = 2(u + v, 2w)$

$$
\begin{aligned}
&= \tfrac{1}{2}\left(\|u+v+2w\|^2 - \|u+v-2w\|^2\right)\\
&= \tfrac{1}{2}\left(\|(u+w)+(v+w)\|^2 - \|(u-w)+(v-w)\|^2\right)\\
&\overset{*}{=} \tfrac{1}{2}\Big[2\big(\|u+w\|^2 + \|v+w\|^2 - \|u-w\|^2 - \|v-w\|^2\big)\Big]\\
&= 4(u,w) + 4(v,w)
\end{aligned}
$$

and the claim follows.

(iv) We show $(qv,w) = q(v,w)$ for all $q \in \mathbb{Q}$. If $q = n \in \mathbb{N}_0$, we iterate (iii) n times and have

$$(nv,w) = n(v,w) \qquad \forall\, n \in \mathbb{N}_0 \tag{*}$$

(the case $n = 0$ is obvious). By the same argument, we get for $m \in \mathbb{N}$

$$(v,w) = \big(m\,\tfrac{1}{m}\,v, w\big) = m\big(\tfrac{1}{m}\,v, w\big)$$

which means that

$$\big(\tfrac{1}{m}\,v, w\big) = \tfrac{1}{m}\,(v,w) \qquad \forall\, m \in \mathbb{N}. \tag{**}$$

Combining (*) and (**) then yields $(\frac{n}{m}\,v, w) = \frac{n}{m}\,(v,w)$. Thus,

$$(pu + qv, w) = p(u,w) + q(v,w) \qquad \forall\, p,q \in \mathbb{Q}.$$

(v) By the lower triangle inequality for norms we get for any $s,t \in \mathbb{R}$

$$
\begin{aligned}
\big|\|tv \pm w\| - \|sv \pm w\|\big| &\leqslant \|(tv \pm w) - (sv \pm w)\|\\
&= \|(t-s)v\|\\
&= |t-s| \cdot \|v\|.
\end{aligned}
$$

This means that the maps $t \mapsto tv \pm w$ are continuous and so is $t \mapsto (tv,w)$ as the sum of two continuous maps. If $t \in \mathbb{R}$ is arbitrary, we pick a sequence $(q_j)_{j\in\mathbb{N}} \subset \mathbb{Q}$ such that $\lim_j q_j = t$. Then

$$(tv,w) = \lim_j (q_j v, w) = \lim_j q_j (qv,w) = t(v,w)$$

so that

$$(su + tv, w) = (su,w) + (tv,w) = s(u,w) + t(v,w).$$

Problem 20.3 This is actually a problem on complexification of inner product spaces... .

Since v and iw are vectors in $V \oplus iV$ and since $\|v\| = \|\pm iv\|$, we get

$$\begin{aligned}(v, iw)_{\mathbb{R}} &= \tfrac{1}{4}\left(\|v + iw\|^2 - \|v - iw\|^2\right)\\ &= \tfrac{1}{4}\left(\|i(w - iv)\|^2 - \|(-i)(w + iv)\|^2\right)\\ &= \tfrac{1}{4}\left(\|w - iv\|^2 - \|w + iv\|^2\right) \qquad (*)\\ &= (w, -iv)_{\mathbb{R}}\\ &= -(w, iv)_{\mathbb{R}}.\end{aligned}$$

In particular,

$$(v, iv) = -(v, iv) \implies (v, iv) = 0 \quad \forall v,$$

and we get

$$(v, v)_{\mathbb{C}} = (v, v)_{\mathbb{R}} > 0 \implies v = 0.$$

Moreover, using (*) we see that

$$\begin{aligned}(v, w)_{\mathbb{C}} &= (v, w)_{\mathbb{R}} + i(v, iw)_{\mathbb{R}}\\ &\overset{*}{=} (w, v)_{\mathbb{R}} - i(w, iv)_{\mathbb{R}}\\ &= (w, v)_{\mathbb{R}} + \bar{i} \cdot (w, iv)_{\mathbb{R}}\\ &= \overline{(w, v)_{\mathbb{R}} + i(w, iv)_{\mathbb{R}}}\\ &= \overline{(w, v)_{\mathbb{C}}}.\end{aligned}$$

Finally, for real $\alpha, \beta \in \mathbb{R}$ the linearity property of the real scalar product shows that

$$\begin{aligned}(\alpha u + \beta v, w)_{\mathbb{C}} &= \alpha(u, w)_{\mathbb{R}} + \beta(v, w)_{\mathbb{R}} + i\alpha(u, iw)_{\mathbb{R}} + i\beta(v, iw)_{\mathbb{R}}\\ &= \alpha(u, w)_{\mathbb{C}} + \beta(v, w)_{\mathbb{C}}.\end{aligned}$$

Therefore to get the general case where $\alpha, \beta \in \mathbb{C}$ we only have to consider the purely imaginary case:

$$\begin{aligned}(iv, w)_{\mathbb{C}} = (iv, w)_{\mathbb{R}} + i(iv, iw)_{\mathbb{R}} &\overset{*}{=} -(v, iw)_{\mathbb{R}} - i(v, -w)_{\mathbb{R}}\\ &= -(v, iw)_{\mathbb{R}} + i(v, w)_{\mathbb{R}}\\ &= i\big(i(v, iw)_{\mathbb{R}} + (v, w)_{\mathbb{R}}\big)\\ &= i(v, w)_{\mathbb{C}},\end{aligned}$$

where we used twice the identity (*). This shows complex linearity in the first coordinate, while skew-linearity follows from the conjugation rule $(v, w)_{\mathbb{C}} = \overline{(w, v)_{\mathbb{C}}}$.

Problem 20.4 The parallelogram law (stated for L^1) *would* say:

$$\Big(\int_0^1 |u+w|\,dx\Big)^2 + \Big(\int_0^1 |u-w|\,dx\Big)^2 = 2\Big(\int_0^1 |u|\,dx\Big)^2 + 2\Big(\int_0^1 |w|\,dx\Big)^2.$$

If $u \pm w, u, w$ have always only ONE sign (i.e. +ve or −ve), we could leave the modulus signs $|\bullet|$ away, and the equality would be correct! To show that there is no equality, we should therefore choose functions where we have some sign change. We try:

$$u(x) = 1/2, \qquad w(x) = x$$

(note: $u - w$ does change its sign!) and get

$$\begin{aligned}
\int_0^1 |u+w|\,dx &= \int_0^1 (\tfrac12 + x)\,dx = [\tfrac12(x+x^2)]_0^1 = 1\\
\int_0^1 |u-w|\,dx &= \int_0^{1/2} (\tfrac12 - x)\,dx + \int_{1/2}^1 (x-\tfrac12)\,dx\\
&= [\tfrac12(x-x^2)]_0^{1/2} + [\tfrac12(x^2-x)]_{1/2}^1\\
&= \tfrac14 - \tfrac18 - \tfrac18 + \tfrac14 = \tfrac14\\
\int_0^1 |u|\,dx &= \int_0^1 \tfrac12\,dx = \tfrac12\\
\int_0^1 |w|\,dx &= \int_0^1 x\,dx = [\tfrac12 x^2]_0^1 = \tfrac12
\end{aligned}$$

This shows that

$$1^2 + (\tfrac14)^2 = \tfrac{17}{16} \neq 1 = 2(\tfrac12)^2 + 2(\tfrac12)^2.$$

We conclude, in particular, that L^1 cannot be a Hilbert space (since in any Hilbert space the Parallelogram law is true....).

Problem 20.5 (i) If $k = 0$ we have $\theta = 1$ and everything is obvious. If $k \neq 0$, we use the summation formula for the geometric progression to get

$$S := \frac1n \sum_{j=1}^n \theta^{jk} = \frac1n \sum_{j=1}^n \left(\theta^k\right)^j = \frac{\theta}{n}\,\frac{1-(\theta^k)^n}{1-\theta^k}$$

but $(\theta^k)^n = \exp(2\pi \frac{i}{n} \cdot k \cdot n) = \exp(2\pi i k) = 1$. Thus $S = 0$ and the claim follows.

(ii) Note that $\overline{\theta^j} = \theta^{-j}$ so that

$$\begin{aligned}\|v + \theta^j w\|^2 &= \langle v + \theta^j w, v + \theta^j w\rangle \\ &= \langle v, v\rangle + \langle v, \theta^j w\rangle + \langle \theta^j w, v\rangle + \langle \theta^j w, \theta^j w\rangle \\ &= \langle v, v\rangle + \theta^{-j}\langle v, w\rangle + \theta^j\langle w, v\rangle + \theta^j\theta^{-j}\langle w, w\rangle \\ &= \langle v, v\rangle + \theta^{-j}\langle v, w\rangle + \theta^j\langle w, v\rangle + \langle w, w\rangle.\end{aligned}$$

Therefore,

$$\begin{aligned}&\frac{1}{n}\sum_{j=1}^{n}\theta^j\|v + \theta^j w\|^2 \\ &= \frac{1}{n}\sum_{j=1}^{n}\theta^j\langle v, v\rangle + \frac{1}{n}\sum_{j=1}^{n}\langle v, w\rangle + \frac{1}{n}\sum_{j=1}^{n}\theta^{2j}\langle w, v\rangle + \frac{1}{n}\sum_{j=1}^{n}\theta^j\langle w, w\rangle \\ &= 0 + \langle v, w\rangle + 0 + 0\end{aligned}$$

where we used the result from part (i) of the exercise.

(iii) Since the function $\phi \mapsto e^{i\phi}\|v + e^{i\phi}w\|^2$ is bounded and continuous, the integral exists as a (proper) Riemann integral, and we can use *any* Riemann sum to approximate the integral, see 11.6–11.10 in Chapter 11 or Corollary E.6 and Theorem E.8 of Appendix E. Before we do that, we change variables according to $\psi = (\phi+\pi)/2\pi$ so that $d\psi = d\phi/2\pi$ and

$$\frac{1}{2\pi}\int_{(-\pi,\pi]} e^{i\phi}\left\|v + e^{i\phi}w\right\|^2 d\phi = -\int_{(0,1]} e^{2\pi i\psi}\left\|v - e^{2\pi i\psi}w\right\|^2 d\psi.$$

Now using equidistant Riemann sums with step $1/n$ and nodes $\theta_n^j = e^{2\pi i\cdot\frac{1}{n}\cdot j}$, $j = 1, 2, \dots, n$ yields, because of part (ii) of the problem,

$$\begin{aligned}-\int_{(0,1]} e^{2\pi i\psi}\left\|v - e^{2\pi i\psi}w\right\|^2 d\psi &= -\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}\theta_n^j\|v - \theta_n^j w\|^2 \\ &= -\lim_{n\to\infty}\langle v, -w\rangle \\ &= \langle v, w\rangle.\end{aligned}$$

Problem 20.6 We assume that V is a $\mathbb{C}$-inner product space. Then,

$$\|v + w\|^2 = \langle v + w, v + w\rangle$$

$$
\begin{aligned}
&= \langle v, v\rangle + \langle v, w\rangle + \langle w, v\rangle + \langle w, w\rangle \\
&= \|v\|^2 + \langle v, w\rangle + \overline{\langle v, w\rangle} + \|w\|^2 \\
&= \|v\|^2 + 2\mathrm{Re}\,\langle v, w\rangle + \|w\|^2.
\end{aligned}
$$

Thus

$$
\|v+w\|^2 = \|v\|^2 + \|w\|^2 \iff \mathrm{Re}\,\langle v, w\rangle = 0 \iff v\perp w.
$$