Radon-Nikodym Theorem, Many Ways

Radon-Nikodym Theorem. *Let (X, Σ) be a measurable space, and let μ, ν be σ-finite measures on (X, Σ). If $\nu \ll \mu$, then there exists a Σ-measurable function $g : X \to [0, \infty)$ such that*

$$\nu(A) = \int_A g \, d\mu \quad \forall A \in \Sigma.$$

Moreover, g is unique μ-almost everywhere (a.e.) and it is called the Radon-Nikodym derivative of ν with respect to μ, denoted by $g = \frac{d\nu}{d\mu}$.

Contents

1 Preliminaries

We follow standard conventions, for a background on Measure Theory, see Folland [3] and Tao [5].

1.1 Measure-theoretic notions

Definition 1.1 (Null sets and a.e. properties)**.** Let μ be a measure on (X, Σ). A set $N \in \Sigma$ is μ-null if $\mu(N) = 0$. A property $P(x)$ holds μ-a.e if $\mu(\{x \in X : P(x) \text{ fails}\}) = 0$. If $f, g : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ are measurable, we write $f = g$ μ-a.e. if $\mu(\{f \neq g\}) = 0$.

Definition 1.2 (Absolute continuity)**.** For measures μ, ν on (X, Σ), ν is absolutely continuous with respect to μ, denoted by $\nu \ll \mu$, if $\mu(A) = 0$ implies $\nu(A) = 0$ for all $A \in \Sigma$.

Definition 1.3 (σ-finiteness)**.** A measure μ on (X, Σ) is σ-finite if there exist sets $X_1, X_2, \dots \in \Sigma$ such that $X = \bigcup_{n \geq 1} X_n$ and $\mu(X_n) < \infty$ for all n.

Definition 1.4 (Restriction)**.** Let μ be a measure on (X, Σ) and $E \in \Sigma$. The restriction $\mu|_E$ is the measure on $(E, \Sigma|_E)$ defined by

$$(\mu|_E)(A \cap E) := \mu(A \cap E), \quad A \in \Sigma,$$

where $\Sigma|_E := \{A \cap E : \ A \in \Sigma\}$.

Lemma 1.5 (Restriction preserves absolute continuity)**.** *If $\nu \ll \mu$, then for every $E \in \Sigma$ we have $\nu|_E \ll \mu|_E$.*

Proof. Let $B \in \Sigma|_E$ with $(\mu|_E)(B) = 0$. Then $B = A \cap E$ for some $A \in \Sigma$ and $\mu(A \cap E) = 0$, hence $\nu(A \cap E) = 0$ since $\nu \ll \mu$. Therefore $(\nu|_E)(B) = 0$. $\square$

1.2 Results used throughout

Lemma 1.6 (Patching from finite to σ-finite)**.** *Assume the Radon-Nikodym theorem holds whenever both measures are finite. Let μ, ν be σ-finite measures on (X, Σ) with $\nu \ll \mu$. Then there exists a Σ-measurable $g : X \to [0, \infty)$ such that*

$$\nu(A) = \int_A g \, d\mu \quad \forall A \in \Sigma,$$

and g is unique μ-a.e.

Proof. Choose $(A_i)_{i \geq 1}$ with $X = \bigcup_i A_i$ and $\mu(A_i) < \infty$, and $(B_j)_{j \geq 1}$ with $X = \bigcup_j B_j$ and $\nu(B_j) < \infty$. Enumerate $E_n := A_i \cap B_j$ so that $X = \bigcup_n E_n$ and $\mu(E_n), \nu(E_n) < \infty$. Let $F_1 := E_1$ and $F_n := E_n \setminus \bigcup_{k<n} E_k$ for $n \geq 2$, so that $X = \biguplus_n F_n$ and still $\mu(F_n), \nu(F_n) < \infty$. By restriction, $\nu|_{F_n} \ll \mu|_{F_n}$ for each n. Apply the finite Radon-Nikodym theorem on $(F_n, \Sigma|_{F_n})$ to obtain $g_n : F_n \to [0, \infty)$ with

$$\nu(A \cap F_n) = \int_{A \cap F_n} g_n \, d\mu \quad \forall A \in \Sigma.$$

Extend each g_n by 0 outside F_n and set $g := \sum_{n \geq 1} g_n$. Then for any $A \in \Sigma$,

$$\int_A g \, d\mu = \sum_{n \geq 1} \int_{A \cap F_n} g_n \, d\mu = \sum_{n \geq 1} \nu(A \cap F_n) = \nu(A),$$

since $A = \biguplus_n (A \cap F_n)$. Uniqueness follows by restricting to each F_n and using uniqueness in the finite case (if g, h both work, then $g = h$ μ-a.e. on every F_n, hence on X). $\square$

Lemma 1.7 (Uniqueness of the Radon-Nikodym derivative)**.** *Let $g, h : X \to [0, \infty]$ be measurable. If*

$$\int_A g \, d\mu = \int_A h \, d\mu \quad \forall A \in \Sigma,$$

then $g = h$ μ-a.e.

Proof. For $n \geq 1$, set $B_n := \{g > h + \frac{1}{n}\} \in \Sigma$. If $\mu(B_n) > 0$, then

$$\int_{B_n} g \, d\mu \ \geq \ \int_{B_n} \left(h + \tfrac{1}{n}\right) d\mu = \int_{B_n} h \, d\mu + \tfrac{1}{n} \, \mu(B_n) > \int_{B_n} h \, d\mu,$$

contradicting the hypothesis with $A = B_n$. Hence $\mu(B_n) = 0$ for all n. Since $\{g > h\} = \bigcup_{n \geq 1} B_n$, we get $\mu(\{g > h\}) = 0$. By symmetry, $\mu(\{h > g\}) = 0$, so $\mu(\{g \neq h\}) = 0$. $\square$

Corollary 1.8 (Change of measure)**.** *Let μ, ν be measures on (X, Σ) with $\nu \ll \mu$, and let $g = \frac{d\nu}{d\mu}$. Then for every measurable $f \geq 0$,*

$$\int_X f\, d\nu \;=\; \int_X f\, g\, d\mu.$$

In particular, if $f \in L^1(\nu)$, then $fg \in L^1(\mu)$ and the same identity holds.

Proof. First assume f is a simple function, $f = \sum_{k=1}^m a_k \mathbf{1}_{A_k}$ with $a_k \geq 0$. Then by linearity and the defining property of g,

$$\int f\, d\nu = \sum_{k=1}^m a_k \nu(A_k) = \sum_{k=1}^m a_k \int_{A_k} g\, d\mu = \int f g\, d\mu.$$

Now let $f \geq 0$ be measurable. Choose simple $f_n \uparrow f$ pointwise. By the previous step, $\int f_n\, d\nu = \int f_n g\, d\mu$ for all n. Apply monotone convergence to both sides to get

$$\int f\, d\nu = \lim_{n\to\infty} \int f_n\, d\nu = \lim_{n\to\infty} \int f_n g\, d\mu = \int f g\, d\mu.$$

Finally, if $f \in L^1(\nu)$, write $f = f^+ - f^-$ and apply the $f \geq 0$ case to f^+ and f^-. This yields $\int |f|\, d\nu = \int |f| g\, d\mu < \infty$, hence $fg \in L^1(\mu)$, and $\int f\, d\nu = \int f g\, d\mu$. □

In what follows, each proof route establishes Radon-Nikodym Theorem in the finite measures case. The general σ-finite measures case follows directly from Lemma 1.6. Two consequences are uniqueness (Lemma 1.7) and the change-of-measure identity (Corollary 1.8).

2 Proof I: Measure-Theoretic

This section gives the classical measure-theoretic proof of the Radon-Nikodym Theorem, following Folland [3]. The argument has two complementary components: (i) a supremum construction that produces a maximal density $g \geq 0$ satisfying

$$\int_A g\, d\mu \leq \nu(A) \quad (\forall A \in \Sigma),$$

and (ii) a signed-measure step (via Hahn decomposition) that upgrades this maximal inequality into the exact identity

$$\nu(A) = \int_A g\, d\mu \quad (\forall A \in \Sigma).$$

We first prove the theorem in the finite-measure case by constructing g from a maximizing sequence and then showing that the residual measure $\lambda := \nu - g\mu$ must vanish. The key point is that if $\lambda \neq 0$, Hahn decomposition identifies a measurable region on which one can increase g while preserving admissibility, contradicting maximality. The extension to the σ-finite case and the uniqueness μ-a.e. property then follow from Lemma 1.7 and Lemma 1.6.

2.1 Signed measures and Hahn decomposition

Definition 2.1 (Signed measure)**.** A signed measure on (X, Σ) is a set function $\rho : \Sigma \to \overline{\mathbb{R}}$ that can be written as $\rho = \rho^+ - \rho^-$ for two finite measures $\rho^\pm$ and is countably additive on disjoint unions (with the usual convention excluding $\infty - \infty$).

Theorem 2.2 (Hahn decomposition)**.** *Let ρ be a signed measure on (X, Σ). Then there exists a measurable partition $X = P \sqcup N$ such that $\rho(E) \geq 0$ for all $E \subset P$ and $\rho(E) \leq 0$ for all $E \subset N$.*

Proof. Write $\rho = \rho^+ - \rho^-$ where $\rho^\pm$ are finite measures. Define

$$\mathcal{A} := \{A \in \Sigma : \ \rho(E) \geq 0 \text{ for all } E \in \Sigma \text{ with } E \subset A\}.$$

Let $\alpha := \sup_{A\in\mathcal{A}} \rho(A)$. By definition of α, for each n choose $A'_n \in \mathcal{A}$ with $\rho(A'_n) > \alpha - \frac{1}{n}$, and set $A_n := \bigcup_{k=1}^n A'_k$. Then (A_n) is increasing, $(A_n)_{n\geq1} \subset \mathcal{A}$ and $\rho(A_n) \uparrow \alpha$. Set

$$P := \bigcup_{n\geq1} A_n, \quad N := X \setminus P.$$

Since $A_n \uparrow P$ and $\rho^{\pm}$ are continuous from below as they are measures,

$$\rho(P) = \rho^+(P) - \rho^-(P) = \lim_{n\to\infty} \rho^+(A_n) - \lim_{n\to\infty} \rho^-(A_n) = \lim_{n\to\infty} \rho(A_n) = \alpha.$$

$\rho(E) \geq 0$ for all measurable $E \subset P$: Let $E \subset P$ be measurable and define $E_n := E \cap A_n \subset A_n$. Then $\rho(E_n) \geq 0$ for each n since $A_n \in \mathcal{A}$. Also $E_n \uparrow E$ as $n \to \infty$. Since $\rho^{\pm}$ are continuous from below,

$$\rho(E) = \rho^+(E) - \rho^-(E) = \lim_{n\to\infty} \rho^+(E_n) - \lim_{n\to\infty} \rho^-(E_n) = \lim_{n\to\infty} \big(\rho^+(E_n) - \rho^-(E_n)\big) = \lim_{n\to\infty} \rho(E_n) \geq 0.$$

Thus $P \in \mathcal{A}$, i.e. $\rho(E) \geq 0$ for all measurable $E \subset P$.

$\rho(E) \leq 0$ for all measurable $E \subset N$: Suppose not. Then there exists a measurable $E \subset N$ with $\rho(E) > 0$. Define $\beta := \sup\{\rho(F) : F \in \Sigma,\ F \subset E\}$. Then $\beta \geq \rho(E) > 0$. For each $n \geq 1$ choose $F_n' \subset E$ measurable such that $\rho(F_n') > \beta - \frac{1}{n}$. Set $F_n := \bigcup_{k=1}^n F_k'$. Then $F_n \subset E$, (F_n) is increasing, and $\rho(F_n) \uparrow \beta$. Define

$$B := \bigcup_{n\geq 1} F_n \subset E.$$

We claim $B \in \mathcal{A}$. First, since $F_n \uparrow B$, by continuity from below for $\rho^{\pm}$, $\rho(B) = \lim_{n\to\infty} \rho(F_n) = \beta$. Next, let $G \subset B$ be measurable. Put $G_n := G \cap F_n$, so $G_n \uparrow G$. If $\rho(G) < 0$, then $\rho(G_n) \to \rho(G) < 0$, so for n large enough $\rho(G_n) < 0$. But then $F_n \setminus G_n \subset E$ and $\rho(F_n \setminus G_n) = \rho(F_n) - \rho(G_n) > \rho(F_n)$, contradicting the definition of β as the supremum of $\rho(\cdot)$ over subsets of E. Hence $\rho(G) \geq 0$ for all $G \subset B$, i.e. $B \in \mathcal{A}$.

Now $B \subset E \subset N$ implies $B \cap P = \varnothing$. For any measurable $H \subset P \cup B$, write $H = (H \cap P) \cup (H \cap B)$ as disjoint union. Since $P, B \in \mathcal{A}$, we have $\rho(H \cap P) \geq 0$ and $\rho(H \cap B) \geq 0$, so $\rho(H) \geq 0$. Thus $P \cup B \in \mathcal{A}$ and $\rho(P \cup B) = \rho(P) + \rho(B) > \rho(P) = \alpha$, contradicting the definition of α. This contradiction shows $\rho(E) \leq 0$ for all measurable $E \subset N$.

All in all, $X = P \sqcup N$ is a Hahn decomposition. □

2.2 Proof of Radon-Nikodym Theorem

Lemma 2.3 (Lattice property)**.** *Let μ, ν be measures on (X, Σ) and define*

$$\mathcal{C} := \Big\{ f : X \to [0, \infty] \text{ measurable} : \int_A f \, d\mu \leq \nu(A) \quad \forall A \in \Sigma \Big\}.$$

If $f, h \in \mathcal{C}$, then $f \vee h \in \mathcal{C}$, where $f \vee h := \max\{f, h\}$.

Proof. With $f, h \in \mathcal{C}$, for any $A \in \Sigma$,

$$\int_A (f \vee h) \, d\mu = \int_{A \cap \{f \geq h\}} f \, d\mu + \int_{A \cap \{h > f\}} h \, d\mu \leq \nu(A \cap \{f \geq h\}) + \nu(A \cap \{h > f\}) = \nu(A).$$

Thus $f \vee h \in \mathcal{C}$. □

Theorem 2.4 (Radon-Nikodym, finite case)**.** *Let μ, ν be finite measures on (X, Σ) with $\nu \ll \mu$. Then there exists a measurable function $g : X \to [0, \infty)$ such that*

$$\nu(A) = \int_A g \, d\mu \quad \forall A \in \Sigma.$$

Proof. By Lemma 2.3, let $\alpha := \sup_{f \in \mathcal{C}} \int_X f \, d\mu$, choose $f_n \in \mathcal{C}$ such that $\int_X f_n \, d\mu \uparrow \alpha$, and set $g_n := f_1 \vee \cdots \vee f_n \in \mathcal{C}$. Then (g_n) is increasing and $\int_X g_n \, d\mu \uparrow \alpha$. Define $g := \sup_{n \geq 1} g_n = \lim_{n\to\infty} g_n$. By monotone convergence, g is measurable and for every $A \in \Sigma$,

$$\int_A g \, d\mu = \lim_{n\to\infty} \int_A g_n \, d\mu \leq \nu(A),$$

so $g \in \mathcal{C}$ and $\int_X g \, d\mu = \alpha$. Define the set function

$$\lambda(A) := \nu(A) - \int_A g \, d\mu \quad (A \in \Sigma).$$

Since $g \in \mathcal{C}$, we have $\lambda(A) \geq 0$ for all A, and since both $A \mapsto \nu(A)$ and $A \mapsto \int_A g\,d\mu$ are measures, λ is a finite measure. Moreover, $\lambda \ll \mu$ because $\nu \ll \mu$ and $\int_A g\,d\mu = 0$ whenever $\mu(A) = 0$.

Assume for contradiction that $\lambda(X) > 0$. Choose $\varepsilon > 0$ such that $\rho(X) := \lambda(X) - \varepsilon\mu(X) > 0$, and define the signed measure $\rho := \lambda - \varepsilon\mu$. Let $X = P \sqcup N$ be a Hahn decomposition of ρ (Theorem 2.2). Then $\rho(P) \geq 0$ and $\rho(N) \leq 0$, hence $\rho(P) = \rho(X) - \rho(N) > \rho(X) > 0$. In particular, $\mu(P) > 0$: otherwise $\mu(P) = 0$ would imply $\lambda(P) = 0$ (since $\lambda \ll \mu$), contradicting $\rho(P) = \lambda(P) - \varepsilon\mu(P) = \lambda(P) > 0$.

Now define $g' := g + \frac{\varepsilon}{2}\mathbf{1}_P$, we claim $g' \in \mathcal{C}$. Indeed, fix any $A \in \Sigma$. Since $A \cap P \subset P$ and P is ρ-positive, we have

$$\rho(A \cap P) = \lambda(A \cap P) - \varepsilon\mu(A \cap P) \geq 0 \implies \lambda(A \cap P) \geq \varepsilon\mu(A \cap P).$$

Using $\lambda(A) \geq \lambda(A \cap P)$ (monotonicity of the measure λ),

$$\int_A g'\,d\mu = \int_A g\,d\mu + \frac{\varepsilon}{2}\mu(A \cap P) = \nu(A) - \lambda(A) + \frac{\varepsilon}{2}\mu(A \cap P) \leq \nu(A) - \lambda(A \cap P) + \frac{\varepsilon}{2}\mu(A \cap P) \leq \nu(A).$$

Thus $g' \in \mathcal{C}$. But then

$$\int_X g'\,d\mu = \int_X g\,d\mu + \frac{\varepsilon}{2}\mu(P) > \int_X g\,d\mu,$$

contradicting the maximality $\int_X g\,d\mu = \alpha$. Hence $\lambda(X) = 0$. Since $\lambda \geq 0$, it follows that $\lambda \equiv 0$, i.e.

$$\nu(A) = \int_A g\,d\mu \quad \forall A \in \Sigma.$$

□

By Theorem 2.4, Radon-Nikodym holds when both measures are finite. Lemma 1.6 then yields the σ-finite case. Uniqueness μ-a.e. follows from Lemma 1.7.

3 Proof II: Partitions and Martingales

This section gives a probabilistic proof of the Radon–Nikodym theorem following Durrett [2]. The guiding idea is that the density $f = d\nu/d\mu$ is the limit of increasingly fine local averages. Given a finite partition $\mathcal{P}$, the best $\mathcal{P}$-measurable approximation to f is the piecewise-constant function that on each atom $A \in \mathcal{P}$ equals the average ratio $\nu(A)/\mu(A)$. Refining the partitions produces a filtration $(\mathcal{F}_n)$ and a sequence (g_n) of such approximations; coherence under refinement is exactly the martingale property. The Martingale Convergence Theorem then yields an almost sure limit g, and we verify that $\nu(B) = \int_B g\,d\mu$ for all measurable B.

3.1 Conditional expectation and martingale convergence

Throughout this section, (X, Σ, μ) is a finite measure space (in particular, a probability space is allowed), and $(\mathcal{F}_n)_{n \geq 1}$ denotes an increasing sequence of sub-σ-algebras of Σ.

Definition 3.1 (Conditional expectation)**.** Let $\mathcal{F} \subseteq \Sigma$ be a sub-σ-algebra and let $Y \in L^1(\mu)$. A conditional expectation of Y given $\mathcal{F}$ is any $\mathcal{F}$-measurable function $Z \in L^1(\mu)$ such that

$$\int_B Z\,d\mu = \int_B Y\,d\mu \quad \forall B \in \mathcal{F}.$$

Any two such functions agree μ-a.e. and we denote (a chosen version of) Z by $\mathbb{E}_\mu[Y \mid \mathcal{F}]$.

Lemma 3.2 (Conditional expectation for a finite partition)**.** *Let $\mathcal{P}$ be a finite measurable partition of X and let $\mathcal{F} = \sigma(\mathcal{P})$ be a σ-algebra generated by $\mathcal{P}$. For $Y \in L^1(\mu)$, define*

$$\mathbb{E}_\mu[Y \mid \sigma(\mathcal{P})] := \sum_{A \in \mathcal{P}} \left(\frac{1}{\mu(A)}\int_A Y\,d\mu\right)\mathbf{1}_A, \text{ with the convention } \frac{1}{\mu(A)}\int_A Y\,d\mu := 0 \text{ if } \mu(A) = 0.$$

Then this function is $\mathcal{F}$-measurable, belongs to $L^1(\mu)$, and satisfies

$$\int_B \mathbb{E}_\mu[Y \mid \mathcal{F}]\,d\mu = \int_B Y\,d\mu \quad \forall B \in \mathcal{F}.$$

Hence it is a version of the conditional expectation from Definition 3.1.

Proof. Set

$$Z := \sum_{A\in\mathcal{P}} \left(\frac{1}{\mu(A)} \int_A Y\,d\mu\right) \mathbf{1}_A,$$

with the stated convention on $\{\mu(A)=0\}$. Then Z is $\sigma(\mathcal{P})$-measurable and

$$\int_X |Z|\,d\mu = \sum_{A\in\mathcal{P}} \int_A |Z|\,d\mu = \sum_{A\in\mathcal{P}} \left|\frac{1}{\mu(A)} \int_A Y\,d\mu\right| \mu(A) \le \sum_{A\in\mathcal{P}} \int_A |Y|\,d\mu = \int_X |Y|\,d\mu < \infty,$$

so $Z \in L^1(\mu)$.

Now let $B \in \sigma(\mathcal{P})$. Since $\mathcal{P}$ is a partition, B is a disjoint union of atoms $A \in \mathcal{P}$ (possibly up to null sets), so

$$\int_B Z\,d\mu = \sum_{\substack{A\in\mathcal{P}\\ A\subseteq B}} \int_A Z\,d\mu = \sum_{\substack{A\in\mathcal{P}\\ A\subseteq B}} \frac{1}{\mu(A)} \left(\int_A Y\,d\mu\right) \mu(A) = \sum_{\substack{A\in\mathcal{P}\\ A\subseteq B}} \int_A Y\,d\mu = \int_B Y\,d\mu.$$

Thus Z satisfies the defining property in Definition 3.1. ☐

Definition 3.3 (Martingale)**.** Let (X,Σ,μ) be a finite measure space and let $(\mathcal{F}_n)_{n\ge1}$ be an increasing sequence of sub-σ-algebras of Σ (a *filtration*). A sequence of integrable functions $(M_n)_{n\ge1} \subset L^1(\mu)$ is called an $L^1(\mu)$-*martingale* with respect to $(\mathcal{F}_n)$ if for every n:

1. M_n is $\mathcal{F}_n$-measurable;
2. $\mathbb{E}_\mu[\,|M_n|\,] < \infty$ (equivalently $M_n \in L^1(\mu)$);
3. $\mathbb{E}_\mu[M_{n+1} \mid \mathcal{F}_n] = M_n$ μ-a.e.

If moreover $M_n \ge 0$ μ-a.e. for all n, we call it a *nonnegative martingale.*

Definition 3.4 (Submartingale and supermartingale)**.** Let (X,Σ,μ) be a finite measure space and let $(\mathcal{F}_n)_{n\ge1}$ be a filtration. A sequence $(M_n)_{n\ge1} \subset L^1(\mu)$ is called

1. a *submartingale* (w.r.t. $(\mathcal{F}_n)$) if M_n is $\mathcal{F}_n$-measurable for all n and
$$\mathbb{E}_\mu[M_{n+1} \mid \mathcal{F}_n] \ge M_n \ \mu\text{-a.e. } \forall n;$$
2. a *supermartingale* if M_n is $\mathcal{F}_n$-measurable for all n and
$$\mathbb{E}_\mu[M_{n+1} \mid \mathcal{F}_n] \le M_n \ \mu\text{-a.e. } \forall n.$$

Lemma 3.5 (Martingales are both sub- and supermartingales)**.** *Every martingale is both a submartingale and a supermartingale.*

Proof. If $\mathbb{E}_\mu[M_{n+1} \mid \mathcal{F}_n] = M_n$ a.e., then both inequalities in Definition 3.4 hold. ☐

Lemma 3.6 (Basic identities for martingales)**.** *If (M_n) is an $L^1(\mu)$-martingale with respect to $(\mathcal{F}_n)$, then:*

1. *for all n, $\int_X M_{n+1}\,d\mu = \int_X M_n\,d\mu$;*
2. *more generally, for any $B \in \mathcal{F}_n$, $\int_B M_{n+1}\,d\mu = \int_B M_n\,d\mu$.*

In particular, $\|M_n\|_{L^1(\mu)} = \int_X |M_n|\,d\mu$ is not necessarily constant, but $\int_X M_n\,d\mu$ is constant if $M_n \ge 0$.

Proof. By Definition 3.3, $\mathbb{E}_\mu[M_{n+1} \mid \mathcal{F}_n] = M_n$ μ-a.e. Integrating both sides over any $B \in \mathcal{F}_n$ and using the defining property of conditional expectation (Definition 3.1) yields

$$\int_B M_{n+1}\,d\mu = \int_B M_n\,d\mu.$$

Taking $B = X \in \mathcal{F}_n$ gives the first claim. ☐

Theorem 3.7 (Martingale convergence for nonnegative martingales)**.** *Let (X, Σ, μ) be a finite measure space and let $(\mathcal{F}_n)$ be a filtration. If $(M_n)_{n\geq 1}$ is a nonnegative $L^1(\mu)$-martingale, then there exists $M \in L^1(\mu)$ such that $M_n \to M$ μ-a.e. Moreover,*

$$\int_X M\,d\mu = \lim_{n\to\infty} \int_X M_n\,d\mu,$$

and if in addition (M_n) is uniformly integrable (for example, if it is bounded in $L^p(\mu)$ for some $p > 1$), then $M_n \to M$ in $L^1(\mu)$ as well.

Proof. Since $M_n \geq 0$ and (M_n) is a martingale, Lemma 3.6 implies that $\int_X M_n\,d\mu$ is constant and finite, hence the family $\{M_n : n \geq 1\}$ is bounded in $L^1(\mu)$. Define the maximal function $M^* := \sup_{n\geq 1} M_n$. A standard maximal inequality (Doob's inequality) for nonnegative submartingales yields, for all $\lambda > 0$,

$$\mu(\,M^* > \lambda\,) \leq \frac{1}{\lambda} \sup_{n\geq 1} \int_X M_n\,d\mu.$$

In particular, $M^* < \infty$ μ-a.e. This implies $(M_n(x))_{n\geq 1}$ is bounded for μ-a.e. x. Next, apply the upcrossing inequality (again standard and proved using only the martingale property) to conclude that, for any rationals $a < b$, the number of upcrossings of the interval $[a, b]$ by the sequence $(M_n(x))$ is finite for μ-a.e. x. Since $\mathbb{Q}^2$ is countable, for μ-a.e. x there are finitely many upcrossings for every rational interval, which forces $M_n(x)$ to converge. Define $M(x) := \lim_{n\to\infty} M_n(x)$ on this full-measure set, and set $M = 0$ elsewhere. Finally, Fatou's lemma gives

$$\int_X M\,d\mu \leq \liminf_{n\to\infty} \int_X M_n\,d\mu < \infty,$$

so $M \in L^1(\mu)$, and the identity for $\int_X M\,d\mu$ follows since $\int_X M_n\,d\mu$ is constant and $M_n \to M$ a.e. with $M_n \geq 0$. The L^1 convergence under uniform integrability is a standard implication of Vitali's theorem. □

3.2 Proof of Radon-Nikodym Theorem

We first prove the theorem for finite measures and then extend to the σ-finite case using Lemma 1.6. Assume $\mu(X) < \infty$, $\nu(X) < \infty$, and $\nu \ll \mu$. Assume Σ is countably generated, i.e. $\Sigma = \sigma(E_1, E_2, \dots)$ for some measurable sets $(E_k)_{k\geq 1}$.[1] For each n, let $\mathcal{P}_n$ be the finite partition of X into the atoms of the algebra generated by $E_1, \dots, E_n$:

$$\mathcal{P}_n := \Big\{ \bigcap_{k=1}^{n} F_k : F_k \in \{E_k, E_k^c\} \Big\} \setminus \{\emptyset\}.$$

Let $\mathcal{F}_n := \sigma(\mathcal{P}_n) = \sigma(E_1, \dots, E_n)$; then $(\mathcal{F}_n)_{n\geq 1}$ is a filtration and $\sigma\Big(\bigcup_{n\geq 1} \mathcal{F}_n\Big) = \Sigma$.

Definition 3.8 (Partition approximations)**.** For each $n \geq 1$, define the $\mathcal{F}_n$-measurable function

$$g_n := \sum_{A\in\mathcal{P}_n} \frac{\nu(A)}{\mu(A)} \mathbf{1}_A, \quad (0/0 := 0).$$

This is well-defined since $\nu \ll \mu$ implies $\mu(A) = 0 \Rightarrow \nu(A) = 0$.

Lemma 3.9 (Representation on $\mathcal{F}_n$)**.** *For every $n \geq 1$ and every $B \in \mathcal{F}_n$,*

$$\nu(B) = \int_B g_n\,d\mu.$$

In particular, $\int_X g_n\,d\mu = \nu(X)$ and $g_n \in L^1(\mu)$.

Proof. Every $B \in \mathcal{F}_n = \sigma(\mathcal{P}_n)$ is a disjoint union of atoms $A \in \mathcal{P}_n$, hence

$$\int_B g_n\,d\mu = \sum_{\substack{A\in\mathcal{P}_n\\ A\subseteq B}} \frac{\nu(A)}{\mu(A)}\mu(A) = \sum_{\substack{A\in\mathcal{P}_n\\ A\subseteq B}} \nu(A) = \nu(B).$$

Taking $B = X$ gives $\int_X g_n\,d\mu = \nu(X)$. □

[1] This holds in essentially all standard probability/analysis settings (e.g. standard Borel spaces). The fully general case can be handled by replacing the sequence below with a directed family of finite partitions.

Lemma 3.10 (Martingale property)**.** *The sequence $(g_n)_{n\geq 1}$ is a nonnegative $L^1(\mu)$-martingale with respect to the filtration $(\mathcal{F}_n)_{n\geq 1}$.*

Proof. By construction, g_n is $\mathcal{F}_n$-measurable and $g_n \geq 0$. Also $g_n \in L^1(\mu)$ by Lemma 3.9. It remains to show $\mathbb{E}_\mu[g_{n+1} \mid \mathcal{F}_n] = g_n$ μ-a.e. Since g_n is $\mathcal{F}_n$-measurable, it suffices (Definition 3.1) to check that

$$\int_B g_{n+1}\, d\mu = \int_B g_n\, d\mu \quad \forall\, B \in \mathcal{F}_n.$$

But $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$, so Lemma 3.9 applied at levels n and $n+1$ yields

$$\int_B g_{n+1}\, d\mu = \nu(B) = \int_B g_n\, d\mu.$$

□

Lemma 3.11 (Uniform integrability from absolute continuity)**.** *If $\nu \ll \mu$, then the martingale (g_n) is uniformly integrable. Equivalently,*

$$\lim_{K\to\infty} \sup_{n\geq 1} \int_{\{g_n > K\}} g_n\, d\mu = 0.$$

Proof. Fix $\varepsilon > 0$. Since $\nu \ll \mu$ and ν is finite, there exists $\delta > 0$ such that $\mu(E) < \delta \Rightarrow \nu(E) < \varepsilon$. By Markov's inequality and Lemma 3.9,

$$\mu(\{g_n > K\}) \leq \frac{1}{K} \int_X g_n\, d\mu = \frac{\nu(X)}{K} \quad \forall n.$$

Choose K large enough that $\nu(X)/K < \delta$. Then $\mu(\{g_n > K\}) < \delta$, hence $\nu(\{g_n > K\}) < \varepsilon$ for every n. Finally, $\{g_n > K\} \in \mathcal{F}_n$, so Lemma 3.9 gives

$$\int_{\{g_n > K\}} g_n\, d\mu = \nu(\{g_n > K\}) < \varepsilon \quad \forall n.$$

Taking the supremum over n proves uniform integrability. □

Proof of Theorem 2.4. By Lemma 3.10, (g_n) is a nonnegative $L^1(\mu)$-martingale. By Theorem 3.7, there exists $g \in L^1(\mu)$ such that $g_n \to g$ μ-a.e. Moreover, by Lemma 3.11, (g_n) is uniformly integrable, hence Theorem 3.7 yields $g_n \to g$ in $L^1(\mu)$ as well.

Let $\mathcal{A} := \bigcup_{n\geq 1} \mathcal{F}_n$. This is an algebra and $\sigma(\mathcal{A}) = \Sigma$. Fix $B \in \mathcal{A}$; then $B \in \mathcal{F}_m$ for some m. For all $n \geq m$, Lemma 3.9 gives $\nu(B) = \int_B g_n\, d\mu$. Passing to the limit using L^1-convergence,

$$\nu(B) = \lim_{n\to\infty} \int_B g_n\, d\mu = \int_B g\, d\mu.$$

Define

$$\mathcal{C} := \Big\{ B \in \Sigma :\ \nu(B) = \int_B g\, d\mu \Big\}.$$

Then $\mathcal{C}$ is a Dynkin system (closed under complements and countable disjoint unions), and it contains the algebra $\mathcal{A}$. By the π-λ / monotone class principle, $\mathcal{C}$ contains $\sigma(\mathcal{A}) = \Sigma$. Therefore,

$$\nu(B) = \int_B g\, d\mu \quad \forall B \in \Sigma.$$

□

Now assume μ and ν are σ-finite and $\nu \ll \mu$. Choose measurable sets $X_1 \subseteq X_2 \subseteq \cdots$ with $X = \bigcup_{k\geq 1} X_k$ and $0 < \mu(X_k) < \infty$, $\nu(X_k) < \infty$ for all k. Apply the finite-measure case above to the restricted measures $\mu|_{X_k}$ and $\nu|_{X_k}$ to obtain measurable $g_k : X_k \to [0, \infty)$ such that

$$\nu(B) = \int_B g_k\, d\mu \quad \forall\, B \in \Sigma \text{ with } B \subseteq X_k.$$

By Lemma 1.6, these local densities patch together to a measurable $g : X \to [0, \infty)$ satisfying

$$\nu(B) = \int_B g \, d\mu \quad \forall \, B \in \Sigma.$$

Uniqueness μ-a.e. follows from the usual argument in Lemma 1.7: if g' also represents ν, then $\int_B (g - g') \, d\mu = 0$ for all B, hence $g = g'$ μ-a.e. This completes the proof of Radon-Nikodym Theorem.

4 Proof III: Strong von Neumann

This section presents von Neumann's Hilbert-space proof of the finite-measure Radon-Nikodym theorem, following Shapiro [4]. The key idea is to regard

$$f \;\longmapsto\; \int_X f \, d\nu$$

as a bounded linear functional on a suitable Hilbert space. By the Riesz representation theorem, this functional is realized as an inner product with some $h \in L^2(\rho)$, where $\rho := \mu + \nu$. One then solves the identity $d\nu = h \, d(\mu + \nu)$ for $d\nu$ in terms of $d\mu$ to obtain $d\nu = g \, d\mu$.

4.1 Hilbert space setup and Riesz representation

Throughout this section assume $\mu(X) < \infty$, $\nu(X) < \infty$, and $\nu \ll \mu$. Set

$$\rho := \mu + \nu.$$

Then ρ is a finite measure and $\nu \ll \rho$. Let $H := L^2(\rho)$ with inner product $\langle f, k \rangle := \int_X f \, k \, d\rho$. Define a linear functional $L : H \to \mathbb{R}$ by

$$L(f) := \int_X f \, d\nu.$$

This is well-defined: if $f = 0$ ρ-a.e., then $f = 0$ ν-a.e. since $\nu \ll \rho$.

Lemma 4.1 (Boundedness of L)**.** *The functional L is bounded on $L^2(\rho)$ and satisfies*

$$|L(f)| \le \sqrt{\nu(X)} \, \|f\|_{L^2(\rho)} \quad \forall f \in L^2(\rho).$$

Proof. For $f \in L^2(\rho)$, Cauchy-Schwarz gives

$$|L(f)| \le \int_X |f| \, d\nu \le \left(\int_X |f|^2 \, d\nu \right)^{1/2} \nu(X)^{1/2} \le \left(\int_X |f|^2 \, d\rho \right)^{1/2} \nu(X)^{1/2} = \sqrt{\nu(X)} \, \|f\|_{L^2(\rho)}.$$

□

Theorem 4.2 (Riesz representation in Hilbert spaces)**.** *Let H be a Hilbert space (over $\mathbb{R}$). For every bounded linear functional $L : H \to \mathbb{R}$, there exists a unique $h \in H$ such that*

$$L(f) = \langle f, h \rangle \quad \forall f \in H.$$

Applying Theorem 4.2 to $H = L^2(\rho)$ and the functional L above, we obtain $h \in L^2(\rho)$ such that

$$\int_X f \, d\nu \;=\; \int_X f \, h \, d\rho \quad \forall f \in L^2(\rho). \tag{1}$$

In particular, taking $f = \mathbf{1}_A$ for $A \in \Sigma$ yields

$$\nu(A) = \int_A h \, d\rho \quad (A \in \Sigma),$$

so h is a Radon-Nikodym derivative of ν with respect to ρ.

Lemma 4.3 (Pointwise bounds on h)**.** *The function h satisfies $0 \le h \le 1$ ρ-a.e.*

Proof. First, for any measurable A we have $\nu(A) = \int_A h\,d\rho \geq 0$. If $N := \{h < 0\}$ had positive ρ-measure, then $\nu(N) = \int_N h\,d\rho < 0$, a contradiction. Hence $h \geq 0$ ρ-a.e. Next, for each $n \geq 1$ set $A_n := \{h > 1 + \frac{1}{n}\}$. Then

$$\nu(A_n) = \int_{A_n} h\,d\rho > \left(1 + \frac{1}{n}\right)\rho(A_n) \geq \rho(A_n).$$

But $\nu(A_n) \leq \rho(A_n)$ since $\rho = \mu + \nu$. Hence $\rho(A_n) = 0$ for all n, and therefore $\rho(\{h > 1\}) = 0$. □

4.2 Back-substitution to recover $d\nu/d\mu$

Let $N := \{h = 1\}$. Since $\nu(N) = \int_N h\,d\rho = \rho(N)$, we have

$$\rho(N) = \mu(N) + \nu(N) = \nu(N) \Rightarrow \mu(N) = 0,$$

and hence also $\nu(N) = 0$ because $\nu \ll \mu$. Define the measurable function

$$g := \frac{h}{1-h}\mathbf{1}_{\{h<1\}} \quad \text{(with any value on } N\text{; it is } \mu\text{-null).}$$

Lemma 4.4 (A useful identity)**.** *For every bounded measurable φ,*

$$\int_X \varphi(1-h)\,d\nu = \int_X \varphi h\,d\mu.$$

Proof. Bounded φ belongs to $L^2(\rho)$ since $\rho(X) < \infty$, so (1) applies:

$$\int_X \varphi\,d\nu = \int_X \varphi h\,d\rho = \int_X \varphi h\,d\mu + \int_X \varphi h\,d\nu.$$

Rearranging yields the claim. □

Proof of Theorem 2.4. For $n \geq 1$ let $E_n := \{h \leq 1 - \frac{1}{n}\}$. Then $E_n \uparrow \{h < 1\}$ and on E_n we have $\frac{1}{1-h} \leq n$. Fix $A \in \Sigma$ and apply Lemma 4.4 with

$$\varphi := \frac{\mathbf{1}_{A\cap E_n}}{1-h},$$

which is bounded and measurable. Then

$$\nu(A \cap E_n) = \int_X \frac{\mathbf{1}_{A\cap E_n}}{1-h}(1-h)\,d\nu = \int_X \frac{\mathbf{1}_{A\cap E_n}}{1-h}h\,d\mu = \int_{A\cap E_n} \frac{h}{1-h}\,d\mu = \int_{A\cap E_n} g\,d\mu.$$

Letting $n \to \infty$ and using monotone convergence on both sides (note that $A \cap E_n \uparrow A \cap \{h < 1\}$ and $g\mathbf{1}_{E_n} \uparrow g$), we obtain

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

Finally, $\nu(A \cap N) = 0$ and $\mu(N) = 0$, hence

$$\nu(A) = \nu(A \cap \{h < 1\}) = \int_{A\cap\{h<1\}} g\,d\mu = \int_A g\,d\mu.$$

□

As in the previous sections, the extension from finite measures to σ-finite measures follows directly from Lemma 1.6. Uniqueness μ-a.e. is handled by Lemma 1.7.

5 Failure Modes

Radon-Nikodym requires absolute continuity $\nu \ll \mu$. When this fails, the obstruction is not mysterious: ν can contain a *singular part* that lives on a μ-null set, and no density with respect to μ can describe that portion. The correct global statement is the *Lebesgue decomposition* (Folland [3] and Tao [5]).

Definition 5.1 (Mutual singularity)**.** Let μ, ν be measures on (X, Σ). We say ν is *singular with respect to* μ, written $\nu \perp \mu$, if there exists $N \in \Sigma$ such that

$$\mu(N) = 0 \text{ and } \nu(N^c) = 0.$$

Equivalently, μ and ν are supported on disjoint measurable sets. From this definition, a basic uniqueness principle follows: if η is a measure such that $\eta \ll \mu$ and $\eta \perp \mu$, then $\eta = 0$. Indeed, choose N with $\mu(N) = 0$ and $\eta(N^c) = 0$. Absolute continuity gives $\eta(N) = 0$, hence $\eta(X) = \eta(N) + \eta(N^c) = 0$.

Theorem 5.2 (Lebesgue decomposition)**.** *Let μ, ν be σ-finite measures on (X, Σ). Then there exist measures ν_{ac} and ν_{s} such that*

$$\nu = \nu_{\mathrm{ac}} + \nu_{\mathrm{s}}, \quad \nu_{\mathrm{ac}} \ll \mu, \quad \nu_{\mathrm{s}} \perp \mu.$$

Moreover, the pair $(\nu_{\mathrm{ac}}, \nu_{\mathrm{s}})$ is unique.

Proof. Set $\lambda := \mu + \nu$. Then λ is σ-finite and $\mu \ll \lambda$, $\nu \ll \lambda$. By the Radon-Nikodym Theorem, there exist Σ-measurable functions $g, f \geq 0$ such that

$$\mu(A) = \int_A g \, d\lambda, \quad \nu(A) = \int_A f \, d\lambda \quad (\forall A \in \Sigma).$$

Let $E := \{g > 0\}$ and $N := E^c = \{g = 0\}$. Note that

$$\mu(N) = \int_N g \, d\lambda = 0.$$

Define, for $A \in \Sigma$,

$$\nu_{\mathrm{s}}(A) := \nu(A \cap N) = \int_{A \cap N} f \, d\lambda, \quad \nu_{\mathrm{ac}}(A) := \nu(A \cap E) = \int_{A \cap E} f \, d\lambda.$$

Then $\nu = \nu_{\mathrm{ac}} + \nu_{\mathrm{s}}$ is immediate from the partition $A = (A \cap E) \sqcup (A \cap N)$. Also, $\nu_{\mathrm{s}} \perp \mu$ because $\mu(N) = 0$ and $\nu_{\mathrm{s}}(N^c) = \nu(N^c \cap N) = 0$. It remains to see $\nu_{\mathrm{ac}} \ll \mu$. On E we have $g > 0$, so we may set

$$h := \frac{f}{g} \mathbf{1}_E \quad \text{(with any value on } N),$$

which is Σ-measurable. For $A \in \Sigma$,

$$\int_A h \, d\mu = \int_A \frac{f}{g} \mathbf{1}_E \, d\mu = \int_{A \cap E} \frac{f}{g} \, d\mu = \int_{A \cap E} \frac{f}{g} \, g \, d\lambda = \int_{A \cap E} f \, d\lambda = \nu_{\mathrm{ac}}(A).$$

Thus $\nu_{\mathrm{ac}} = h\mu$, hence $\nu_{\mathrm{ac}} \ll \mu$. For uniqueness, suppose also $\nu = \eta + \rho$ with $\eta \ll \mu$ and $\rho \perp \mu$. Then $\nu_{\mathrm{ac}} - \eta = \rho - \nu_{\mathrm{s}}$ is a signed measure which is absolutely continuous w.r.t. μ (difference of $\ll \mu$ measures) and singular w.r.t. μ (difference of $\perp \mu$ measures). By the basic uniqueness principle above, it must be 0. Hence $\nu_{\mathrm{ac}} = \eta$ and $\nu_{\mathrm{s}} = \rho$. $\square$

Corollary 5.3 (Lebesgue-Radon-Nikodym form)**.** *Under the hypotheses of Theorem 5.2, there exist a measurable $h : X \to [0, \infty)$ and a measure $\nu_{\mathrm{s}} \perp \mu$ such that*

$$\nu(A) = \int_A h \, d\mu + \nu_{\mathrm{s}}(A) \quad (\forall A \in \Sigma).$$

Moreover, h is unique μ-a.e. and equals $d\nu_{\mathrm{ac}}/d\mu$.

Proof. This is Lebesgue decomposition plus Radon-Nikodym applied to the absolutely continuous part $\nu_{\mathrm{ac}} \ll \mu$. $\square$

The obstruction to writing $\nu(A) = \int_A h\, d\mu$ for all A is precisely the singular component ν_{s}. In particular,

$$\nu \ll \mu \quad \Longleftrightarrow \quad \nu_{\mathrm{s}} = 0.$$

So Radon-Nikodym works *exactly* on the ν_{ac} part and has nothing to say about ν_{s}.

Example 5.4 (Density plus atoms). Let $X = \mathbb{R}$ with Borel σ-algebra, $\mu = dx$ (Lebesgue measure), and

$$\nu := f(x)\, dx + \sum_{k \geq 1} a_k\, \delta_{x_k}, \quad f \geq 0,\ a_k \geq 0,\ x_k \in \mathbb{R}.$$

Then $\nu_{\mathrm{ac}} = f\, dx$ and $\nu_{\mathrm{s}} = \sum_{k \geq 1} a_k\, \delta_{x_k}$. Indeed, the atomic part is supported on the countable set $\{x_k\}$, which has μ-measure 0, hence is singular w.r.t. dx.

Example 5.5 (Cantor measure vs Lebesgue). Let $\mu = dx$ on $[0, 1]$ and let ν be the standard Cantor-Lebesgue (Cantor) measure. Then $\nu \perp \mu$, so $\nu_{\mathrm{ac}} = 0$ and $\nu_{\mathrm{s}} = \nu$. Equivalently, any putative density h must satisfy $\int h\, d\mu = 0$ while $\nu([0, 1]) = 1$, so the mass of ν cannot be captured by a μ-density.

Example 5.6 (Purely atomic vs atomless). Let μ be an atomless measure (e.g. Lebesgue measure on $[0, 1]$), and let

$$\nu = \sum_{k \geq 1} a_k\, \delta_{x_k} \text{ with } a_k \geq 0.$$

Then $\nu \perp \mu$: take $N := \{x_k : \ k \geq 1\}$, which is countable, hence $\mu(N) = 0$, and $\nu(N^c) = 0$ by construction. Thus the decomposition is $\nu_{\mathrm{ac}} = 0$, $\nu_{\mathrm{s}} = \nu$.

6 Calculus of Radon-Nikodym Derivatives

This section records the standard "calculus rules" for Radon-Nikodym derivatives that one uses constantly in practice; see Folland [3] and Durrett [2]. Throughout, (X, Σ) is a measurable space and all measures are assumed σ-finite.

6.1 Algebraic rules

Lemma 6.1 (Chain rule). *Let λ, μ, ν be measures on (X, Σ) with $\nu \ll \mu$ and $\mu \ll \lambda$. Then $\nu \ll \lambda$ and*

$$\frac{d\nu}{d\lambda} = \frac{d\nu}{d\mu} \frac{d\mu}{d\lambda} \quad \lambda\text{-a.e.}$$

Proof. Let $f := \frac{d\nu}{d\mu}$ and $h := \frac{d\mu}{d\lambda}$. For any $A \in \Sigma$,

$$\int_A fh\, d\lambda = \int_A f\, d\mu = \nu(A),$$

so fh is a Radon-Nikodym derivative of ν with respect to λ, and uniqueness holds λ-a.e. □

Corollary 6.2 (Linearity). *Let μ be a measure and let $\nu_1, \nu_2 \ll \mu$. Then $(\nu_1 + \nu_2) \ll \mu$ and*

$$\frac{d(\nu_1 + \nu_2)}{d\mu} = \frac{d\nu_1}{d\mu} + \frac{d\nu_2}{d\mu} \quad \mu\text{-a.e.}$$

More generally, for $a, b \geq 0$,

$$\frac{d(a\nu_1 + b\nu_2)}{d\mu} = a\frac{d\nu_1}{d\mu} + b\frac{d\nu_2}{d\mu} \quad \mu\text{-a.e.}$$

Proof. Let $f_i := \frac{d\nu_i}{d\mu}$. For any $A \in \Sigma$,

$$\int_A (f_1 + f_2)\, d\mu = \nu_1(A) + \nu_2(A) = (\nu_1 + \nu_2)(A),$$

and similarly for $a\nu_1 + b\nu_2$. Uniqueness gives the identities μ-a.e. □

Corollary 6.3 (Reciprocal rule)**.** *Let μ, ν be measures on (X, Σ) with $\mu \ll \nu$ and $\nu \ll \mu$ (i.e. $\mu \sim \nu$). Write $g := \frac{d\nu}{d\mu}$ and $h := \frac{d\mu}{d\nu}$. Then*

$$g\,h = 1 \quad \nu\text{-a.e.} \qquad \text{and} \qquad h\,g = 1 \quad \mu\text{-a.e.}$$

In particular,

$$\frac{d\mu}{d\nu} = \frac{1}{d\nu/d\mu} \quad \nu\text{-a.e.} \qquad \text{and} \qquad \frac{d\nu}{d\mu} = \frac{1}{d\mu/d\nu} \quad \mu\text{-a.e.}$$

Proof. Apply Lemma 6.1 twice: first with $\nu \ll \mu \ll \nu$ to get

$$\frac{d\nu}{d\nu} = \frac{d\nu}{d\mu}\frac{d\mu}{d\nu},$$

which is $1 = gh$ ν-a.e.; and second with $\mu \ll \nu \ll \mu$ to get $1 = hg$ μ-a.e. □

6.2 Change of measure in probability language

Corollary 6.4 (Expectation under a change of measure)**.** *Let $(\Omega, \mathcal{F})$ be a measurable space and let $\mathbb{P}, \mathbb{Q}$ be probability measures with $\mathbb{Q} \ll \mathbb{P}$. Define the likelihood ratio (density)*

$$L := \frac{d\mathbb{Q}}{d\mathbb{P}}.$$

Then for every measurable $Y \geq 0$,

$$\mathbb{E}_{\mathbb{Q}}[Y] = \mathbb{E}_{\mathbb{P}}[Y L].$$

In particular, $\mathbb{E}_{\mathbb{P}}[L] = 1$.

Proof. This is Corollary 1.8 applied with $(\mu, \nu) = (\mathbb{P}, \mathbb{Q})$; taking $Y \equiv 1$ gives $\mathbb{E}_{\mathbb{P}}[L] = \mathbb{Q}(\Omega) = 1$. □

Corollary 6.5 (Bayes formula for conditional expectations)**.** *In the setting of Corollary 6.4, let $\mathcal{G} \subseteq \mathcal{F}$ be a sub-σ-algebra and let $Y \in L^1(\mathbb{Q})$. Define*

$$Z := \frac{\mathbb{E}_{\mathbb{P}}[Y L \mid \mathcal{G}]}{\mathbb{E}_{\mathbb{P}}[L \mid \mathcal{G}]},$$

with the convention $Z = 0$ on the set $\{\mathbb{E}_{\mathbb{P}}[L \mid \mathcal{G}] = 0\}$. Then

$$\mathbb{E}_{\mathbb{Q}}[Y \mid \mathcal{G}] = Z \quad \mathbb{Q}\text{-a.e.}$$

In particular, for $A \in \mathcal{F}$,

$$\mathbb{Q}(A \mid \mathcal{G}) = \frac{\mathbb{E}_{\mathbb{P}}[\mathbf{1}_A L \mid \mathcal{G}]}{\mathbb{E}_{\mathbb{P}}[L \mid \mathcal{G}]} \quad \mathbb{Q}\text{-a.e.}$$

Proof. Let $B \in \mathcal{G}$. Using Corollary 6.4 and the defining property of conditional expectation (Definition 3.1),

$$\int_B Z\,d\mathbb{Q} = \int_B Z L\,d\mathbb{P} = \int_B \frac{\mathbb{E}_{\mathbb{P}}[Y L \mid \mathcal{G}]}{\mathbb{E}_{\mathbb{P}}[L \mid \mathcal{G}]}\,L\,d\mathbb{P}.$$

Now note that, by the tower property and $\mathcal{G}$-measurability,

$$\int_B \frac{\mathbb{E}_{\mathbb{P}}[Y L \mid \mathcal{G}]}{\mathbb{E}_{\mathbb{P}}[L \mid \mathcal{G}]}\,L\,d\mathbb{P} = \int_B \frac{\mathbb{E}_{\mathbb{P}}[Y L \mid \mathcal{G}]}{\mathbb{E}_{\mathbb{P}}[L \mid \mathcal{G}]}\,\mathbb{E}_{\mathbb{P}}[L \mid \mathcal{G}]\,d\mathbb{P} = \int_B \mathbb{E}_{\mathbb{P}}[Y L \mid \mathcal{G}]\,d\mathbb{P} = \int_B Y L\,d\mathbb{P} = \int_B Y\,d\mathbb{Q}.$$

Thus Z satisfies the defining property of $\mathbb{E}_{\mathbb{Q}}[Y \mid \mathcal{G}]$, and uniqueness holds $\mathbb{Q}$-a.e. □

7 Duality Perspective

This section records a useful "two-sides-of-the-same-coin" viewpoint (see Folland [3]): the Radon-Nikodym theorem and the identification of $(L^1(\mu))^*$ with $L^\infty(\mu)$ are tightly linked.

Remark 7.1 (Dependencies). In many texts, L^1-L^∞ duality is proved *using* Radon-Nikodym (so presenting it as a separate "proof route" can create a logical trap). For that reason, we state the relationship explicitly as an *equivalence*: we give (A) a standard derivation of duality from Radon-Nikodym, and (B) a derivation of Radon-Nikodym from duality *under the standing assumption* that the duality theorem has been proved independently.

7.1 Statement: $(L^1(\mu))^* \cong L^\infty(\mu)$

Let (X, Σ, μ) be a measure space. For $g \in L^\infty(\mu)$ define a linear functional

$$\Phi(g) : L^1(\mu) \to \mathbb{R}, \quad \Phi(g)(f) := \int_X f\, g\, d\mu.$$

By Hölder's inequality, $|\Phi(g)(f)| \le \|g\|_{L^\infty(\mu)} \|f\|_{L^1(\mu)}$, hence $\Phi(g) \in (L^1(\mu))^*$.

Theorem 7.2 (L^1-L^∞ duality; equivalence form)**.** *Assume μ is σ-finite. The following are equivalent.*

(RN) *(Radon-Nikodym) Whenever $\nu \ll \mu$ are σ-finite measures, there exists g with $\nu(A) = \int_A g\, d\mu$ for all $A \in \Sigma$.*

(D) *(Duality) The map $\Phi : L^\infty(\mu) \to (L^1(\mu))^*$ is an isometric isomorphism; i.e. every $T \in (L^1(\mu))^*$ has the form*

$$T(f) = \int_X f\, g\, d\mu \quad (f \in L^1(\mu))$$

for a unique $g \in L^\infty(\mu)$ (unique μ-a.e.), and $\|T\| = \|g\|_{L^\infty(\mu)}$.

7.2 Route A: Radon-Nikodym $\Rightarrow$ duality

We prove (D) from (RN) (this is the standard derivation in Folland [3]).

Proof of (RN)$\Rightarrow$(D). We already noted $\Phi(g) \in (L^1(\mu))^*$ and $\|\Phi(g)\| \le \|g\|_\infty$. For the reverse inequality, let $c := \|g\|_{L^\infty(\mu)}$. By definition of essential supremum, for each n the set $A_n := \{|g| > c - \frac{1}{n}\}$ has positive μ-measure. Define

$$f_n := \frac{\mathbf{1}_{A_n} \operatorname{sgn}(g)}{\mu(A_n)} \in L^1(\mu), \quad \|f_n\|_{L^1(\mu)} = 1.$$

Then

$$\|\Phi(g)\| \ \ge\ |\Phi(g)(f_n)| = \left| \frac{1}{\mu(A_n)} \int_{A_n} |g|\, d\mu \right| \ge c - \frac{1}{n}.$$

Letting $n \to \infty$ yields $\|\Phi(g)\| \ge c$, hence $\|\Phi(g)\| = \|g\|_\infty$ and Φ is an isometry.

Let $T \in (L^1(\mu))^*$ be arbitrary. Define a set function $\nu : \Sigma \to \mathbb{R}$ by

$$\nu(A) := T(\mathbf{1}_A), \quad A \in \Sigma.$$

If μ is finite, then for any disjoint $A_1, A_2, \ldots$ we have $\mathbf{1}_{\cup_{k=1}^n A_k} \to \mathbf{1}_{\cup_{k\ge1} A_k}$ in $L^1(\mu)$, so continuity of T implies countable additivity; hence ν is a finite signed measure. Moreover, for all A,

$$|\nu(A)| = |T(\mathbf{1}_A)| \le \|T\|\, \|\mathbf{1}_A\|_{L^1(\mu)} = \|T\|\, \mu(A),$$

so in particular $\nu \ll \mu$. By applying (RN) to the positive and negative parts of ν, there exists $g \in L^1(\mu)$ with $\nu(A) = \int_A g\, d\mu$ for all A; the domination $|\nu(A)| \le \|T\|\mu(A)$ forces $g \in L^\infty(\mu)$ with $\|g\|_\infty \le \|T\|$.

Now check that $T(f) = \int fg\, d\mu$ for all $f \in L^1(\mu)$. This holds first for simple functions by linearity and the identity for indicators; then extend to all of L^1 by density of simple functions and continuity of both sides. Finally, isometry gives $\|T\| = \|g\|_\infty$.

If μ is merely σ-finite, reduce to the finite case on a partition $X = \biguplus_n E_n$ with $\mu(E_n) < \infty$ (as in Lemma 1.6) and patch the resulting representing functions. □

7.3 Route B: duality $\Rightarrow$ Radon-Nikodym

Here we go in the opposite direction, *assuming* the duality statement (D) has been proved independently.

Proof of (D)$\Rightarrow$(RN). Let μ, ν be σ-finite measures on (X, Σ) with $\nu \ll \mu$. Set $\rho := \mu + \nu$ (still σ-finite). Define a linear functional on $L^1(\rho)$ by

$$L(f) := \int_X f\, d\nu, \qquad f \in L^1(\rho).$$

This is well-defined and bounded because $|L(f)| \le \int |f|\, d\nu \le \int |f|\, d\rho = \|f\|_{L^1(\rho)}$, so $\|L\| \le 1$.

By (D) applied to the measure ρ, there exists $h \in L^\infty(\rho)$ such that

$$\int_X f\, d\nu \;=\; \int_X f\, h\, d\rho \qquad \forall f \in L^1(\rho).$$

Taking $f = \mathbf{1}_A$ shows $\nu(A) = \int_A h\, d\rho$ for all $A \in \Sigma$, i.e. $d\nu = h\, d\rho$. Since L is positive, we may take $h \ge 0$ ρ-a.e.; and since $\|L\| \le 1$, the isometry in (D) gives $\|h\|_{L^\infty(\rho)} \le 1$, hence $0 \le h \le 1$ ρ-a.e.

Now expand $\rho = \mu + \nu$:

$$\nu(A) = \int_A h\, d\rho = \int_A h\, d\mu + \int_A h\, d\nu,$$

so for all A,

$$\int_A (1-h)\, d\nu = \int_A h\, d\mu.$$

In other words, as measures we have

$$(1-h)\, d\nu = h\, d\mu.$$

Because $\nu \ll \mu$, we have $\nu(\{h=1\}) = 0$ (since $\mu(\{h=1\}) = 0$ from the identity above). Define

$$g := \frac{h}{1-h} \quad \text{on } \{h<1\}, \qquad g := 0 \quad \text{on } \{h=1\}.$$

Then g is measurable and for every $A \in \Sigma$,

$$\int_A g\, d\mu = \int_{A\cap\{h<1\}} \frac{h}{1-h}\, d\mu = \int_{A\cap\{h<1\}} \frac{1}{1-h}\, h\, d\mu = \int_{A\cap\{h<1\}} \frac{1}{1-h}\, (1-h)\, d\nu = \nu(A),$$

where the last step uses $\nu(\{h=1\}) = 0$. This is exactly the Radon-Nikodym conclusion. $\square$

8 Geometric Differentiation Viewpoint

This viewpoint treats the Radon-Nikodym derivative as a *local ratio of masses*. It requires additional geometric structure (a topology/metric and a differentiation basis), so it is not available on an arbitrary measurable space. Following the discussion in Tao [5], we can make this precise by examining the Euclidean setting.

Let $X = \mathbb{R}^d$ with its Borel σ-algebra, and let μ, ν be locally finite Borel measures (e.g. Radon measures). For $x \in \mathbb{R}^d$ and $r > 0$, write $B(x,r)$ for the open Euclidean ball. If $\nu \ll \mu$ and $g = \frac{d\nu}{d\mu}$, then

$$\frac{\nu(B(x,r))}{\mu(B(x,r))} = \frac{1}{\mu(B(x,r))} \int_{B(x,r)} g\, d\mu$$

whenever $\mu(B(x,r)) > 0$. Thus, identifying $g(x)$ amounts to a differentiation theorem for the averages of g over shrinking neighbourhoods. To resolve this, we rely on the fact that balls in $\mathbb{R}^d$ form a differentiation basis. This allows us to invoke the Lebesgue differentiation theorem:

Theorem 8.1 (Lebesgue differentiation for measures on $\mathbb{R}^d$)**.** *Let μ be a locally finite Borel measure on $\mathbb{R}^d$ and let $f \in L^1_{\mathrm{loc}}(\mu)$. Then for μ-a.e. x with $\mu(B(x,r)) > 0$ for all sufficiently small r,*

$$\lim_{r\downarrow 0} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} f\, d\mu \;=\; f(x).$$

A standard proof route uses a maximal inequality together with a covering lemma (the geometry of balls in $\mathbb{R}^d$ is what makes this work); see Tao [5]. Applying Theorem 8.1 directly to the density function $f = g = \frac{d\nu}{d\mu}$ yields the promised geometric identification.

Corollary 8.2 (Geometric differentiation formula)**.** *Let μ, ν be locally finite Borel measures on $\mathbb{R}^d$ with $\nu \ll \mu$, and let $g = \frac{d\nu}{d\mu} \in L^1_{\mathrm{loc}}(\mu)$. Then*

$$g(x) = \lim_{r\downarrow 0} \frac{\nu(B(x,r))}{\mu(B(x,r))} \quad \textit{for } \mu\textit{-a.e. } x.$$

Proof. For any x, r with $\mu(B(x,r)) > 0$,

$$\frac{\nu(B(x,r))}{\mu(B(x,r))} = \frac{1}{\mu(B(x,r))} \int_{B(x,r)} g \, d\mu.$$

The conclusion follows by applying Theorem 8.1 to $f = g$. $\square$

This "zooming-in" principle has a discrete parallel in the martingale viewpoint, often described as information refinement. Instead of Euclidean balls, one can use a nested partition, such as dyadic cubes. Let $\mathcal{D}_n$ be the dyadic cubes of sidelength 2^{-n} and let $\mathcal{F}_n$ be the σ-algebra they generate. For $x \in \mathbb{R}^d$ let $Q_n(x) \in \mathcal{D}_n$ be the unique cube containing x, and define

$$g_n(x) := \frac{\nu(Q_n(x))}{\mu(Q_n(x))} \quad (\text{with } g_n(x) = 0 \text{ if } \mu(Q_n(x)) = 0).$$

If $\nu \ll \mu$ with density $g = \frac{d\nu}{d\mu}$, then one checks that $g_n = \mathbb{E}_\mu[g \mid \mathcal{F}_n]$ a.e., so $(g_n)_{n\geq 1}$ is a martingale and the martingale convergence theorem gives $g_n \to g$ μ-a.e.

Ultimately, these formulas highlight why "extra structure" is required. On a general measurable space (X, Σ), the Radon-Nikodym theorem guarantees the existence of a density g as an equivalence class, but there is no canonical notion of a "shrinking neighbourhood." The geometric differentiation formula only becomes available when X carries enough structure (like $\mathbb{R}^d$ or specific metric measure spaces) to support a valid differentiation basis.

9 Radon–Nikodym Property

This section records the Banach-space frontier behind the slogan "Radon–Nikodym = density." For *vector-valued* measures, absolute continuity does *not* by itself guarantee the existence of a density; instead, this becomes a geometric property of the target Banach space. We follow the standard presentation in Diestel and Uhl [1].

9.1 Vector-valued measures and variation

Let (X, Σ) be a measurable space and let $(B, \|\cdot\|)$ be a Banach space.

Definition 9.1 (Vector measure)**.** A *B-valued (countably additive) measure* is a map $m : \Sigma \to B$ such that for every pairwise disjoint family $(A_n)_{n\geq 1} \subset \Sigma$ we have

$$m\Big(\biguplus_{n\geq 1} A_n\Big) = \sum_{n\geq 1} m(A_n) \quad (\text{convergence in } \|\cdot\|).$$

Definition 9.2 (Total variation)**.** If $m : \Sigma \to B$ is a vector measure, define its *variation* on $A \in \Sigma$ by

$$|m|(A) := \sup \left\{ \sum_{k=1}^{n} \|m(A_k)\| \;:\; A = \biguplus_{k=1}^{n} A_k,\ A_k \in \Sigma \right\} \in [0, \infty].$$

We say that m has *bounded variation* if $|m|(X) < \infty$.

It is standard that $|m|$ is a (finite) measure whenever m has bounded variation; it plays the role of "$|dm|$" in the scalar theory.

Definition 9.3 (Absolute continuity for vector measures)**.** Let μ be a (scalar) measure on (X, Σ). We write $m \ll \mu$ if $\mu(A) = 0$ implies $m(A) = 0$ for all $A \in \Sigma$. Equivalently (when m has bounded variation), $m \ll \mu$ iff $|m| \ll \mu$.

9.2 Bochner densities

To even state a Radon-Nikodym theorem, we must specify what we mean by a "density".

Definition 9.4 (Bochner integrability)**.** A function $f : X \to B$ is *Bochner integrable* with respect to μ if it is strongly (Bochner) measurable and

$$\int_X \|f\| \, d\mu < \infty.$$

In this case one can define $\int_A f\,d\mu \in B$ for all $A \in \Sigma$ by approximation with simple B-valued functions, and the map

$$A \longmapsto \int_A f\,d\mu$$

is a B-valued measure of bounded variation.

9.3 Vector Radon-Nikodym theorem and the RNP

The scalar Radon-Nikodym theorem says: if $\nu \ll \mu$, then ν has a density $g \in L^1(\mu)$. For vector measures, the analogous statement may fail even when $m \ll \mu$.

Theorem 9.5 (Vector Radon-Nikodym property [1]). *Let (X, Σ, μ) be a* finite *measure space and let $m : \Sigma \to B$ be a B-valued measure of bounded variation with $m \ll \mu$. If B has the Radon-Nikodym property, then there exists a Bochner integrable function $f \in L^1(\mu; B)$ such that*

$$m(A) = \int_A f\,d\mu \qquad (\forall A \in \Sigma).$$

In this case f is unique μ-a.e. and is denoted by $f = \frac{dm}{d\mu}$.

This motivates the central definition.

Definition 9.6 (Radon-Nikodym property (RNP)). A Banach space B has the *Radon-Nikodym property* if for every finite measure space (X, Σ, μ) and every B-valued measure m of bounded variation with $m \ll \mu$, there exists a Bochner integrable density $f \in L^1(\mu; B)$ such that $m(A) = \int_A f\,d\mu$ for all $A \in \Sigma$.

Remark 9.7 (Failure mode). If B fails the RNP, then there exist a finite measure space (X, Σ, μ) and a B-valued measure m of bounded variation with $m \ll \mu$ for which *no* Bochner density exists. Thus, unlike the scalar case, "$m \ll \mu$" is not the right hypothesis by itself; one must also assume a structural property of the target space.

9.4 Operator viewpoint

A convenient reformulation packages vector measures as operators on L^1.

Definition 9.8 (Representable operators). Let (X, Σ, μ) be finite. A bounded linear operator $T : L^1(\mu) \to B$ is *representable* if there exists an essentially bounded, strongly measurable $g : X \to B$ such that

$$T(\varphi) = \int_X \varphi(x)\, g(x)\, d\mu(x) \qquad (\forall \varphi \in L^1(\mu)).$$

Given $T : L^1(\mu) \to B$, the set function $m_T(A) := T(\mathbf{1}_A)$ is a B-valued measure of bounded variation and satisfies $m_T \ll \mu$ (indeed $\|m_T(A)\| \le \|T\|\mu(A)$). Conversely, any such m induces a bounded operator by $T_m(\varphi) = \int \varphi\, dm$. One of the standard equivalences in Diestel and Uhl [1] is that B has the RNP iff every bounded $T : L^1(\mu) \to B$ is representable (it suffices to test this on $(X, \Sigma, \mu) = ([0, 1], \mathcal{B}, \lambda)$).

9.5 Examples and non-examples

We list a few standard facts; proofs and many more characterizations (dentability, martingales, differentiability of Lipschitz maps, ...) are in Diestel and Uhl [1].

- Reflexive spaces have the RNP. In particular, Hilbert spaces and $L^p(\mu)$ for $1 < p < \infty$ have the RNP.
- Separable duals have the RNP. If $B = Y^*$ for some separable Banach space Y, then B has the RNP; e.g. $\ell^1 = c_0^*$ has the RNP.
- A canonical non-example is $L^1(0, 1)$ (Lebesgue measure), which does *not* have the RNP; consequently there exist $L^1(0, 1)$-valued measures $m \ll \lambda$ of bounded variation with no Bochner density. Other familiar spaces without the RNP include $L^\infty(0, 1)$ and $C[0, 1]$.
- Closed subspaces of an RNP space have the RNP, and equivalent renormings preserve the RNP.

Remark 9.9 (Big picture). At the scalar level, the Radon-Nikodym theorem says: absolute continuity $\Rightarrow$ density. At the vector level, the correct statement is:

$$m \ll \mu \text{ and } B \text{ has RNP} \quad \Longrightarrow \quad m \text{ has a Bochner density.}$$

So the RN phenomenon persists precisely in those Banach spaces whose geometry rules out the "pathological" vector measures.

References

[1] Joseph Diestel and J. Jerry Uhl. *Vector Measures*. Vol. 15. Mathematical Surveys and Monographs. Providence, RI: American Mathematical Society, 1977. ISBN: 978-0821815151.

[2] Rick Durrett. *Probability: Theory and Examples*. 5th ed. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge: Cambridge University Press, 2019. ISBN: 978-1108473682.

[3] Gerald B. Folland. *Real Analysis: Modern Techniques and Their Applications*. 2nd ed. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. New York: John Wiley & Sons, 1999. ISBN: 978-0471317166.

[4] Joel H. Shapiro. *The Radon–Nikodym Theorem "Made Easy"*. Lecture Notes. 2018.

[5] Terence Tao. *An Introduction to Measure Theory*. Vol. 126. Graduate Studies in Mathematics. Providence, RI: American Mathematical Society, 2011. ISBN: 978-0821869192.